Education and Training in Developmental Disabilities

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

Volume 42  Number 3  September 2007
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          David L. Gast          John McDonnell          Laurence R. Sargent
Reuben Altman        Herbert Goldstein       Gale M. Morrison       Gary M. Sasso
Phillip J. Belfiore  Robert Henderson       Gabriel A. Nardi       Tom E. C. Smith
Sharon Borthwick-Duffy Carolyn Hughes       John Nietupski         Scott Sparks
Michael P. Brady     Larry K. Irvin         James R. Patton        Fred Spooner
Fredda Brown         James V. Kahn          Edward A. Polloway      Robert Stodden
Mary Lynne Calhoun   H. Earle Knowlton       Thomas G. Roberts      Keith Storey
Sharon F. Cramer     Barry W. Lavay        Robert S. Rueda        David L. Westling
Caroline Dunn        Rena Lewis            Diane L. Ryndak        John J. Wheeler
Lise Fox             Kathleen J. Marshall     Edward J. Sabornie     Mark Wolery

Education and Training in Developmental Disabilities is sent to all members of the Division on Developmental Disabilities of The Council for Exceptional Children. All Division members must first be members of The Council for Exceptional Children. Division membership dues are $25.00 for regular members and $13.00 for full time students. Membership is on a yearly basis. All inquiries concerning membership, subscription, advertising, etc. should be sent to the Division on Developmental Disabilities, 1110 North Glebe Road, Arlington, VA 22201.

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 Phil Parette
President Polly Parish
President-Elect J. David Smith
Vice President Emily Bouch
Secretary Toni Merfeld
Treasurer Amanda Boutot

Members

Leslie Broun
Linda Lazar
Nikki Murdick
Angie Stone-MacDonald (Student Governor)
Dianne Zager
Deborah Wichmanowski

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. $175.00 per year; Canada, PUAS, and all other countries $179.50; 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.
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.
No Child Left Behind: Issues of Assessing Students with the Most Significant Cognitive Disabilities
DARLENE E. PERNER

Assistive Technology as a Self-Management Tool for Prompting Students with Intellectual Disabilities to Initiate and Complete Daily Tasks: A Literature Review
LINDA C. MECHLING

Self-Determination Interventions’ Effects on the Academic Performance of Students with Developmental Disabilities
CATHERINE H. FOWLER, MOIRA KONRAD, ALLISON R. WALKER, DAVID W. TEST, and WENDY M. WOOD

Assessing Work Task Preferences among Persons with Intellectual Disabilities: An Integrative Review of Literature
VIRGINIE COBIGO, DIANE MORIN, and YVES LACHAPELLE

Review of Recent Treatment Acceptability Research
STACY L. CARTER

Examining the Career Paths and Transition Services of Students with Disabilities Exiting High School
ROBERT M. BAER, ROBERT W. FLEXER, and LAWRENCE DENNIS

“Am I supposed to understand this stuff?” Youth with Special Health Care Needs Readiness for Transition
TANIS BRYAN, NORA STILES, KAREN BURSTEIN, CEVRIYE ERGUL, and PEN-CHIANG CHAO

Training Paraeducators to Promote the Use of Augmentative and Alternative Communication by Students with Significant Disabilities
MARY ANNA BINGHAM, FRED SPOONER, and DIANE BROWDER

Attitudes of Preservice Teachers Enrolled in an Infusion Preparation Program regarding Planning and Accommodations for Included Students with Mental Retardation
DAVID L. CAMERON and BRYAN G. COOK

Use of Adapted Bicycles on the Learning of Conventional Cycling by Children with Mental Retardation
TAMMY L. BURT, DAVID L. PORRETTA, and RICHARD E. KLEIN

Manuscripts Accepted for Future Publication in Education and Training in Developmental Disabilities
MANUSCRIPTS ACCEPTED FOR FUTURE PUBLICATION IN EDUCATION AND TRAINING IN DEVELOPMENTAL DISABILITIES
Manuscripts Accepted for Future Publication in Education and Training in Developmental Disabilities

December 2007

Lessons learned through implementing a positive behavioral support intervention at home: A case study on self-management with a student with autism and his mother. **Suk-Hyang Lee**, Denise Poston, and AJ Poston, Beach Center on Disability, University of Kansas, Haworth Hall, 1200 Sunnyside Ave., Room 3136, Lawrence, KS 66045-7534.

Addressing communication needs of young adults with autism in a college-based inclusion program. **Carol S. Alpern** and Dianne Zager, Communication Sciences and Disorders Program, Pace University, 1 Pace Plaza, New York, NY 10038.

Using comic strip conversations to increase social satisfaction and decrease loneliness in students with autism spectrum disorder. **Melinda R. Pierson** and Barbara C. Glaeser, Department of Special Education, California State University, Fullerton, PO Box 6868, Fullerton, CA 92834-6868.

Effect of tutoring on reading achievement for students with cognitive disabilities, specific learning disabilities, and students receiving Title I services. Jan Osborn, Amy Freeman, Margaret Burley, **Rich Wilson**, Eric Jones, and Stacey Rychener, Bowling Green State University, 451 Education Building, Bowling Green, OH 43403.


A large-scale study of the characteristics of Asperger syndrome. **Brenda Smith Myles**, Hyo Jung Lee, Sheila M. Smith, Kai-Chien Tien, Yu-Chi Chou, and Jill Hudson, Ohio Center for Autism and Low Incidence, 5220 N. High Street, Building C1, Columbus, OH 43221.


Facilitating Student Achievement with Assistive Technology. **Howard P. Parette** and George R. Peterson-Karlan, Department of Special Education, Illinois State University, Campus Box 5910, Normal, IL 61790-5910.

Technologies for Self-Determination for Youth with Developmental Disabilities. **James R. Skouge**, Mary L. Kelly, Kelly D. Roberts, David W. Leake, and Robert A. Stodden, College of Education, University of Hawai‘i at Manoa, 1776 University Avenue, Honolulu, HI 96822.

Address is supplied for author in boldface type.
No Child Left Behind: Issues of Assessing Students with the Most Significant Cognitive Disabilities

Darlene E. Perner
Bloomsburg University of Pennsylvania

Abstract: Council for Exceptional Children, Division on Developmental Disabilities (DDD), Position Paper approved by the DDD Board on January 31, 2007

Neither school reform nor the focus on developing accountability systems that measure and evaluate the success of schooling are new or innovative educational practices. Even before the No Child Left Behind (NCLB) Act of 2001, many states had changed or were changing their accountability and assessment systems. Similarly, individual school districts or local education authorities, often in the name of school reform, were using their own standards or benchmarks to assess progress for school improvement purposes and for fulfilling long-term district goals. However, with the NCLB mandate, state departments of education have had to specify what all students must achieve through state standards; and for students with the most significant cognitive disabilities, through state departments alternative achievement standards (U.S. State Department of Education, 2003). Many of the content standards, developed to date, continue to be expanded upon and/or refined. Likewise, states have developed or refined statewide assessments based on their own standards to show and monitor student and school achievement.

While the Division on Developmental Disabilities (DDD) supports accountability systems to evaluate school and student achievement and progress, as well as the use of statewide assessments based on achievement and content standards, we are concerned with how states determine accountability, and how they develop and implement assessments that affect students with disabilities. In particular, this DDD position paper will address a number of issues related to statewide assessments used to evaluate students with the most significant cognitive disabilities. The main issues that concern DDD are as follows: the variation of alternate assessments from state to state; the need to prove that the assessment instruments and procedures used are technically sound and appropriately identify proficient (and above) or non-proficient performance; and, the lack of awareness of or attention to the need for pre-service and in-service teacher training in assessing students with alternate assessments and using the results of the alternate assessments to guide ensuing instruction. These and other issues presented in this paper lead to the evaluation of how states are determined accountable for the achievement of students with the most significant cognitive disabilities. Additionally, many of the issues raised relate to all students who may not be able to achieve grade-level standards because of their special instructional needs.

Since the Individuals with Disabilities Education Act (IDEA) Amendments of 1997, states have been required to include students with disabilities in state and system-wide assessments. This includes students with significant cognitive disabilities. In addition, NCLB has identified the specific provisions for the alternate assessments required for students with the most significant disabilities (Flowers, Ahlgrim-Delzell, Browder, & Spooner, 2005). Alternate assessment has been defined by the U.S. Department of Education (2003) as:

Correspondence concerning this article should be addressed to Darlene E. Perner, Dept. of Exceptionality Programs, Bloomsburg University of Pennsylvania, 400 East Second Street, Bloomsburg, PA 17815.
An alternate assessment is an assessment designed for the small number of students with disabilities who are unable to participate in the regular State assessment, even with appropriate accommodations. An alternate assessment may include materials collected under several circumstances, including (1) teacher observation of the student, (2) samples of student work produced during regular classroom instruction that demonstrate mastery of specific instructional strategies in place of performance on a computer-scored multiple-choice test covering the same content and skills, or (3) standardized performance tasks produced in an “on demand” setting, such as completion of an assigned task on test day. To serve the purposes of assessment under title I, an alternate assessment must be aligned with the State’s content standards, must yield results separately in both reading/language arts and mathematics, and must be designed and implemented in a manner that supports use of the results as an indicator of AYP. (p. 68699)

For many states and local education authorities (LEAs), system-wide assessments based on content that reflect standards have become a way of school life. Although the requirements of NCLB have been outlined and are being implemented by state Departments of Education and LEAs, the U.S. Department of Education has allowed for flexibility in the design and implementation of these requirements. As a result, states and LEAs have been challenged with the responsibility of trying to meet all the requirements of NCLB and the many issues and decisions related to accountability, as well as student achievement and system-wide assessments. On January 14, 2004, the U.S. Department of Education identified 40 issues that states and LEAs have taken responsibility for in their systems design and implementation to meet average yearly progress (AYP) and the required 2014 timeline for all students to meet academic proficiency including students with the most significant cognitive disabilities. As a result, the design of accountability systems (including assessments) and the implementation of various provisions of NCLB have led to great variation from one state to another (see Table 1). In some states, this variation can also be from one district to another (e.g., Wyoming). This has made it difficult to identify and discuss some of the issues in a general manner, and to some extent, share “good practices” or “good systems.” DDD believes that reviewing what individual states do or do not do in terms of their responsibilities and decisions causes concern and heightens the awareness of disparity. We feel this holds equally true for those in special education. The issues, responsibilities, and decisions encountered have their own challenges and vary from state to state. These concerns and disparities also exist for students requiring special instruction.

Clearly and simply, the intent of high stakes testing for all students is to raise expectations for student achievement (i.e., to have all students achieve educational standards set by states and districts) and to monitor student and school progress in order to help raise student achievement on arbitrary standards. In May 2005, Secretary of Education Margaret Spellings announced “A New Commonsense Approach to Raising Achievement for Students with Disabilities” directed toward 2% of school population (in addition to the 1% of students with the most significant cognitive disabilities) who are working toward but may not be able to “reach grade-level achievement standards” (U.S. Department of Education, May 10, 2005). This gives to the states who have applied the flexibility to modify achievement standards and assessments for 2% of their student population. Different responsibilities and decisions that each individual state must make come with this added flexibility.

The federal regulations governing the inclusion of students with disabilities in the NCLB mandated assessments have prompted reconsideration of the role of assessment in special education. Some of these considerations related to assessment have included developing alternate standards and/or modified standards, benchmarks and/or performance indicators, extending special educators’ skills and practices to accommodate the teaching of standards-based content, ensuring that students have access to general curriculum, providing appropriate test accommodations for statewide and school district assessments, and designing and implementing alternate assessments directed toward students with the most
significant cognitive disabilities and “modified” assessments toward students with disabilities who may not be able to achieve grade-level standards.

Issues and Concerns about Statewide Alternate Assessments

It is important to recognize that having high expectations of achievement for students with the most significant cognitive disabilities and involving students in statewide assessments that are meaningful (i.e., alternate assessments) is highly supported by educators, parents and associations like the Council for Exceptional Children (CEC) and our Division (DDD) (CEC, 2005; CEC, 2004; Ysseldyke et al., 2004). In devising these state assessments, the U. S. Department of Education (2003) has stated in its regulations that alternate assessments: “should have a clearly defined structure, guidelines for which students may participate, clearly defined scoring criteria and procedures, and a report format that clearly communicates student performance in terms of the academic achievement standards defined by the State. The requirements for high technical quality set forth in §§ 200.2(b) and 200.3(a)(1), including validity, reliability, accessibility, objectivity, and consistency with nationally recognized professional and technical standards, apply to alternate assessments as well as to regular State assessments.” (p. 68699) However, there are various concerns about state alternate assessments and alternate achievement standards that have been developed, or are currently being developed or redefined (e.g., Browder, Fallin, Davis, & Karvonen, 2003; Browder, Wakeman, & Flowers, 2006; Crawford & Tindal, 2006; Kohl, McLaughlin, & Nagle, 2006; Yovanoff & Tindal, 2007). Some of the issues related to state alternate assessments include how they are developed, administered, scored, and reported, and whether they are useful to teachers for improving instruction. Some areas that alternate achievement standards are being questioned on are the validity of their alignment with content standards, and their application to life skills curriculum.

As DDD reviewed the literature, listened to teachers who are responsible for assessing students using their state’s alternate assessment, and discussed critical issues regarding alternate assessments, a number of themes emerge. A major one is the factors that can affect student scores. Browder, Fallin, et al. (2003) identified variables that they felt may influence alternate assessment outcome scores. One of these variables included the technical quality of the alternate assessment format. Yovanoff and Tindal (2007) have agreed stating that “very few research studies have been published on the technical adequacy of alternate assessments…” (p. 186). According to Browder et al. about half of the states have chosen a portfolio format for their alternate assessment. Other states use checklists, IEP analysis and other types of performance-based assessments. More recently, Browder et al. (as cited in Flowers et al., 2005) have identified the most frequently used alternate assessment approaches to be portfolio, performance-based and checklist. A portfolio alternate assessment is a collection of student work gathered to demonstrate the student’s performance on specific skills and knowledge, generally linked to state content standards. Performance-based assessments are items or tasks administered by a teacher or test administrator to a student. These items/tasks are the same for all students. Checklist assessments are a list of skills that an individual familiar with the student’s performance rates according to a specific proficiency scale (Flowers et al.). Yovanoff and Tindal have identified portfolios, observations (teachers are asked to select a behavior that represents student need and then observe it in a functional environment) and performance assessments as the responsive formats used for alternate assessments (see Table 1 for descriptive examples of state alternate assessments). Each of these types of formats has a subjective component and thus we raise a number of questions about assessment considerations. Issues such as rater reliability, the content being assessed and its relation [alignment] to state standards, construct validity, measurement error, assessors’ training, length of assessment, and procedures for administering, scoring and interpreting results must all be addressed and analyzed. It appears that all types of commonly used alternate assessment approaches have problems with technical aspects. Tindal et al. (2003) have been critical of portfolio assess-
<table>
<thead>
<tr>
<th>State</th>
<th>Portfolio/Collection of Work</th>
<th>Performance-based</th>
<th>Alternate Assessment Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arizona</td>
<td><strong>Portfolio (optional):</strong> IEP teams are encouraged to maintain student portfolios or specific examples of pertinent student work. Evidence of student achievement on selected skills may be demonstrated and collected using a variety of formats. Some examples include lists of survival words or logos the student knows, sample of personal budget, and graphs of performance. While the creating of a portfolio or the maintenance of student work products is important for instructional decisions and communication with parents, it is optional for Arizona’s alternate assessment.</td>
<td>Data forms are provided for the assessment of student performance. They include: Teacher Evaluation of Student Performance; Activity-Based Performance Assessment (Level I contains reading, writing, listening and speaking and mathematics skills that are infused into functional activities; listening and speaking skills are not assessed for Level II); and Parent Interview.</td>
<td>AIMS-A: Level I—students attending 2nd–12th grades working on Level I Functional and Kindergarten level assessments; Level II—students attending 6th–12th grades [whose performance on the Level I assessments meets or exceeds the performance standards] working on articulated standards at the 1st–3rd grades [at a beginning level]. Level I includes a teacher evaluation of student performance on selected standards, a parent interview, and, an evaluation of student performance on four activities in four domains—school or vocational (daily schedule), recreation and leisure (interactive game), community (making purchases at a fast food restaurant or school snack bar), and domestic living (entertaining friends). Level II includes a teacher appraisal of the student’s status on selected standards and a performance evaluation of a student’s ability to perform activities in four domains—school or vocational (applying for a job or volunteer activity), recreation and leisure (choosing a free time activity), community (determining transportation options), and domestic living (entertaining friends). Science standards and assessments are currently under development.</td>
</tr>
<tr>
<td>Illinois</td>
<td><strong>Portfolio of Student Work</strong> The Illinois Alternate Assessment (IAA) requires a portfolio of student work collected over the course of the school year. Evidence is being collected for two purposes: 1) to document the student’s participation in standards-based activities focusing on alternate performance indicators (APIs) and 2) to document student’s performance on those APIs.</td>
<td>Alternate performance indicators (APIs) are listed for each state goal. A team is identified to select the specific APIs that will be focused on for assessment and instruction.</td>
<td>Illinois Alternate Assessment consists of a portfolio of student work and other materials with data collection a minimum of three times during the year. Data must be collected at three points (for baseline, midpoint and ending point) in the school year (with defined dates starting in September and ending early February). The materials can include samples of student work, photos of the student doing work in school or at home and teachers’ summaries of what students have learned. Each portfolio is evaluated on several dimensions.</td>
</tr>
</tbody>
</table>
TABLE 1—(Continued)

<table>
<thead>
<tr>
<th>State</th>
<th>Portfolio/Collection of Work</th>
<th>Performance-based</th>
<th>Alternate Assessment Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kansas</td>
<td>Data Folio</td>
<td>Indicators are selected from each of the state’s extended standards, two in reading, four in math, and three in writing. The other indicator(s) may be selected from any standard in that curricular area. The IEP team chooses five extended indicators for each of the content areas: reading, mathematics and writing. The Kansas Alternate Assessment consists of a collection of data (i.e., data folio) of an individual student’s performance of skills and content outlined in the Kansas Extended Standards. Documents are compiled, assembled, scored, and submitted to the state during the spring assessment window. Instruction needs to occur before the assessment window opens and data collection begins. Data for the data folio are collected during the testing window designated by the state department. A scoring rubric (skill level) is used and scored locally by three scorers. Data folio consists of three pieces of evidence for each selected indicator. Formats submitted may include work samples, worksheet, videotape/DVD, photographs, audiotape and data sheets. These products document the student’s knowledge of concepts and content and the performance of skills outlined in the Kansas General and Extended Curricular Standards.</td>
<td></td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>Forty (40) on demand performance tasks per year. Approximately 20 items are related to literacy standards and approximately 20 items are related to numeracy standards. Each test item represents an authentic, relevant, and age-appropriate activity related to reading or mathematics. The science alternate assessment is currently being developed.</td>
<td>The Pennsylvania Alternate System of Assessment (PASA) measures the attainment of knowledge and skills of students with significant cognitive disabilities through a series of approximately 40 discrete items. The PASA is designed to take a snapshot of students’ typical performance on a small sample of academic skills in the areas of reading and mathematics (and soon science) derived from the PA standards. In order for all students with significant disabilities to participate, these tasks are designed with three levels of difficulty to assess skills that are useful in their daily lives. Teachers either scribe or videotape their student’s participation in the on demand performance assessment administered by a test administrator (usually the teacher), one-to-one.</td>
<td></td>
</tr>
</tbody>
</table>
ments in particular because this type of assessment is “difficult to implement with any level of technical adequacy” (p. 483). Similarly, Flowers et al. identified checklists as relying on subjective judgments about student performance instead of direct observation. Also, after working with teachers who conduct performance-based assessments on pre-determined tasks, they have stated that the teachers learn which task levels to implement so that their students perform at the proficient level. The teachers have learned how to use prompts and therefore, feel that the alternate assessment is assessing the teachers’ ability to conduct the assessment and not the student’s level of achievement. Zatta and Pullin (2004) reached similar conclusions after reviewing research related to teachers designing and scoring portfolio assessments. Browder et al. (2006) indicated that one of the needs in improving alternate assessments is “how to establish valid and reliable alternate assessment scores” (p. 256).

Another issue of concern has been identi-
fied as data collection. One of the major ques-
tions heard initially about alternate assess-
ment was, “How will the assessment and the
data collected help inform teachers for in-
structional purposes?” This has not only been
stated by teachers and other educators who
are involved in implementing these assess-
ments but also by researchers (Browder, 2001;
Browder, Spooner, et al., 2003; Flowers et al.,
2005; Kleinert & Kearns, 2001; Kleinert &
Thurlow, 2001). DDD, and others in the field,
feel that a major outcome of the alternate
assessment should be to improve the quality of
our programs for students with the most sig-
nificant cognitive disabilities. Although in
some cases this may be happening, in many
cases it is not. For states that used, or continue
to use, assessments that are given only once
and are not linked directly to the student’s
IEP and those standards particularly relevant
for that student (e.g., pre-determined perfor-
mance tasks given to all students with the most
significant cognitive disabilities in that state)
the results have been discouraging. Some
states that use a portfolio to collect data for a
number of months (e.g., New Hampshire,
New York) or over the course of the school
year (e.g., Illinois, North Carolina) are likely
to see benefits to this on-going assessment
process. However, the data collected must be
appropriately linked to both the individual
student’s IEP goals and the state standards or
alternate content standards and the assess-
ment must be technically sound.

Another major issue of alternate assess-
ments is training which is in turn strongly
related to other variables. This training on
alternate assessments is multi-faceted and
takes in a range of individuals from the stu-
dents themselves (e.g., self-determination as-
sessments as part of the alternate assessment;
involveinent in the portfolio assessment pro-
cess) to teams (e.g., the IEP team’s determi-
nation of whether a student is required to take
the alternate assessment, and informed deci-
sion making regarding the IEP goals).

Although alternate assessment training af-
fected many individuals, the teacher most often
has the major responsibility for using alter-
ate standards to guide instruction and en-
sure that students are accessing the general
curriculum. Training includes helping teach-
ers with methodologies and other techniques
to incorporate alternate assessments and stan-
dards in their day-to-day instruction with stu-
dents with the most significant cognitive dis-
abilities. Teachers are also responsible for
conducting and scoring the alternate assess-
ments, interpreting and reporting the results,
and consequently, making decisions based on
these results and ensuring that students are
involved in the assessment process (e.g., Wash-
burn-Moses, 2003). Teachers and students
must be motivated in this process and there-
fore need to receive the appropriate training
and support to bring about successful assess-
ment. For example, based on the research of
Kampfer, Hovrath, Kleinert, and Kearns
(2001), Kleinert, Green, Hurte, Clayton and
Oetinger (2002) noted that the time teachers
spent on assessing alternate assessment port-
folios was a slight predictor of student scores.
“A far more powerful predictor of student
scores was the extent to which the alternate
assessment was integrated into daily instruc-
tion, as well as the extent to which students
were actively involved in the construction of
their own assessment portfolios” (Kleinert et
found that only 30% of the teachers they sur-
vveyed indicated that state department “test
results were frequently or always useful in guid-
ing instruction” (p. 215). In 2003, Browder,
Fallin, et al. stated that unless teachers are
trained to make appropriate connections be-
tween the alternate assessments and daily in-
struction, the assessments may contribute to
teacher stress due to time constraints.

In two recent studies to assess teachers’ per-
ceptions about alternate assessments, Flowers
et al. (2005) and Toweles-Reeves and Kleinert
(2006) found that the alternate assessments
infringe on teaching time and teachers’ per-
sonal time. As well, teachers using the three
main assessment approaches, portfolio, per-
formance-based and checklist, felt that there
was an increase demand on paperwork. Flow-
ers and her colleagues recommend that teach-
ers be taught ways to streamline their time and
how to better organize and manage the paper-
work. They also suggest that states “focus on
the most parsimonious ways to document
progress on state academic content standards”
(p. 90). The same should be noted for school
districts as some may be requiring additional

NCLB: Issues of Assessing Students / 249
Implications and Recommendations

In addressing only some of the many issues related to alternate assessments, DDD’s position is apparent. States need more time, feedback and resources to fully develop alternate assessments that are relevant to students as well as their teachers, parents and IEP teams. Other needs and implications identified by DDD from the research and the discussions of teachers are: ensuring that functional skills and curriculum are part of the alternate standards and assessments; including parents in the assessment process; providing teachers with more resources and skills; providing training to assist both teachers and students to make connections between the assessment and daily instruction; assisting teachers in helping students be more involved in the assessment process (e.g., self-determination, progress monitoring); ensuring that statewide and district assessments are technically sound (e.g., valid and reliable) and actually assess student’s performance (i.e., not the teacher’s or assessment administrator’s skill in conducting the assessment). Tindal et al. (2003) conclude that more development is needed for alternate assessments before the system can effectively serve teachers and students. The foci should be on teacher training for both implementation and decision making, and on assessments that use “brief measures on a range of relevant behaviors” (p. 493). Browder, Fallin, et al. (2003) state that these measures should be “across time to capture student performance” (p. 263) and IEPs should be based on students’ individual needs and the state standards or alternate state standards. More and more research is being conducted on issues and concerns related to alternate assessments. With this research come additional recommendations as states continue to face the challenge of raising expectations and meeting the goal of proficiency for all of their students.

References


Kleinert, H., Green, P., Hurte, M., Clayton, J., & Oetinger, C. (2002). Creating and using mean-
Assistive Technology as a Self-Management Tool for Prompting Students with Intellectual Disabilities to Initiate and Complete Daily Tasks: A Literature Review

Linda C. Mechling
University of North Carolina Wilmington

Abstract: This paper summarizes the results of a review of the empirical literature (1990-2005) focusing on use of assistive technology as a self-management tool for persons with intellectual disabilities. Forty investigations were identified which provided information on assistive technology to assist persons with disabilities to initiate and complete daily tasks. Four areas of research were defined and analyzed through the identified studies: (a) pictorial prompts; (b) tactile prompts; (c) auditory prompts; and (d) computer-aided systems. Research supports assistive technology as an effective tool for providing antecedent prompts that can be self-operated by persons with intellectual disabilities. Implications of the research and suggestions for future research are discussed.

Increased independence and the ability to manage one’s own behavior and task performance continue to be areas of concern when working with persons with disabilities. Included in these behaviors are abilities to self initiate, self-instruct, self-maintain, and self-monitor one’s behavior and task performance (Post & Storey, 2002). Student self-management strategies provide a means to increase the independence of persons with disabilities, decrease the need for continuous supervision and prompting by others, and free the teacher to attend to other instruction. While self management procedures may be in the form of antecedents or consequences to target behaviors (Browder & Shapiro, 1985), the focus of this review is the application of self management tools in the form of antecedents to influence behavior. Antecedent self-management strategies may cue or guide a person’s behavior by using stimuli such as pictures or audio cues that precede occurrence of a behavior (Harchik, Sherman, & Sheldon, 1992). Antecedent prompts such as photographs may increase a person’s ability to respond independently without: (a) waiting for another person to tell them what to do (Riffel et al., 2005); and (b) the need to memorize step sequences (Lancioni, Van den Hof, Furniss, O’Reilly, & Cunha, 1999). Such prompts may further improve performance, increase fluency of responding (productivity, duration of behavior), and promote maintenance of skills already acquired.

Since Congress outlined guidelines for providing assistive technology in the Technology-Related Assistance for Individuals with Disabilities Act of 1988, the use of assistive technology has been applied across a wide range of settings and activities to meet the needs of students with varying disabilities (Gregor & Pachuski, 1996). Increased advancements, awareness, and availability of materials, equipment, and adaptations make the application of assistive technology an ever evolving field, which includes new inventions, and innovate use of existing tools.

The term assistive technology device means “any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified, or customized, that is used to increase, maintain, or improve functional capabilities of a child with a disability.” 20 U.S.C. 1401 [25], Sec. 504.5 (Individuals with Disabilities Education Act, 1990). Lewis (1998) describes assistive technology as any
technology that can enhance the performance of persons with disabilities by augmenting an individual’s strengths or providing an alternative mode of performing a task to compensate for the effects of a disability. Assistive technology devices to aid persons with disabilities can be technological (vibrating pager, Palmtop PC) or nontechnological (picture prompts, activity schedules) (Bryant, Bryant, & Raskind, 1998). They can be devices designed specifically for persons with disabilities (alternative and augmentative communication devices) or generic devices developed for the general public to use (cassette tape players, Palmtop PC) (Lewis).

Due to the potential of assistive technology for improving the lives of persons with disabilities (Wehmeyer, 1999) and facilitating participation in society (Parette, 1991), it is no wonder that practitioners are evaluating its use to increase self-management skills. The purpose of this review was to examine the published, empirical literature evaluating assistive technology to assist persons with disabilities to initiate and complete daily tasks. Assistive technology, as a self-management tool, was evaluated as a strategy for controlling antecedents used by persons with disabilities to influence their own behavior (Browder & Shapiro, 1985). The review included studies that used both light tech (picture prompts, picture schedules, recording devices) and high tech devices (Palmtop PC, computer-based systems) to increase self-management skills by transferring behavior control to a prompting device operated by the student.

**Method**

The purpose of this paper was to conduct a comprehensive review of literature published from over the past 15 years to summarize use of assistive technology as a self-management tool for persons with intellectual disabilities. Electronic and ancestral searches of journals focusing on the education of persons with disabilities were conducted from 1990-2005. The electronic search of ERIC used the keywords: assistive technology, self-management, picture prompts, auditory prompts, tactile prompts, activity schedules, antecedent prompts, prompts, computer-assisted, computer-based, and Palmtop PC, for locating articles. A manual search was then conducted of the table of contents of the journals listed in Table 1 followed by an ancestral search of the reference lists of the identified articles.

In order to be included in the review, studies met the following criteria:

1. Use of experimental design
2. Publication in peer-reviewed journal
3. Evaluation of assistive technology for self-management
4. Participants were diagnosed with an intellectual disability

Although assistive technology has been shown as an effective tool for teaching new skills (Hagiwara & Myles, 1999; Mechling & Gast, 1997; Norman, Collins, & Schuster, 2001), studies which used trainer-directed cues or instructional prompts were not included in this review which focused on use of assistive technology as an independent prompting device, used by the person alone, rather than a tool for presenting instruction. Nor were studies included which focused on the effects of assistive technology on aberrant behaviors (Alberto, Taber, & Fredrick 1999), social skills (Shabani et al., 2002) and communication skills (Taylor & Levin, 1998).

**TABLE 1**

<table>
<thead>
<tr>
<th>Journals Reviewed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior Modification</td>
</tr>
<tr>
<td>Education and Training in Developmental Disabilities</td>
</tr>
<tr>
<td>Education and Treatment of Children</td>
</tr>
<tr>
<td>Exceptional Children</td>
</tr>
<tr>
<td>Exceptionality</td>
</tr>
<tr>
<td>Focus on Autism and Other Developmental Disabilities</td>
</tr>
<tr>
<td>Focus on Exceptional Children</td>
</tr>
<tr>
<td>Journal of Applied Behavior Analysis</td>
</tr>
<tr>
<td>Journal of the Association for Persons with Severe Handicaps</td>
</tr>
<tr>
<td>Journal of Autism and Developmental Disabilities</td>
</tr>
<tr>
<td>Journal of Developmental and Physical Disabilities</td>
</tr>
<tr>
<td>Journal of Early Intervention</td>
</tr>
<tr>
<td>Journal of Special Education</td>
</tr>
<tr>
<td>Journal of Special Education Technology</td>
</tr>
<tr>
<td>Journal of the Association for Persons with Severe Handicaps</td>
</tr>
<tr>
<td>Mental Retardation</td>
</tr>
<tr>
<td>Remedial and Special Education</td>
</tr>
<tr>
<td>Topics in Early Childhood Special Education</td>
</tr>
</tbody>
</table>

Assistive Technology as a Self-Management Tool / 253
Results

Forty studies (see Table 2) were identified in addition to one literature review by Post and Storey (2002). The literature review focused on the use of auditory prompting. The technology was primarily visual, auditory, and tactile in nature and devices included: (a) picture prompts (17 studies); (b) audio cassette players (11 studies); (c) hand held computer-based systems (palmtop personal computers) (11 studies); and (d) vibration (1 study).

Identified studies measured the effects of assistive technology on: (a) independent performance of tasks (task completion) (25 studies); (b) task engagement (5 studies); (c) on-task behavior (4 studies); (d) accuracy of task performance (3 studies); (e) initiation of tasks (2 studies); (f) transitioning between tasks (2 studies); and (g) fluency of work performance (1 study).

The review found assistive technology to be effective in teaching task initiations and completion to persons with disabilities. Students were taught to self-operate devices to guide their own behaviors independent of instructor prompts. The first part of the paper reviews each of the identified studies and results of the strategies used while the discussion section addresses questions raised from the current research and suggestions for future research.

Picture Prompts

Pictures as visual supports were the most frequently evaluated form of antecedent prompts (17 studies) and reported to be the most commonly used and economic material (Lancioni, O’Reilly, & Oliva, 2001). Visual supports may provide aids to maintain attention, understand spoken language, sequence events, organize environments, or increase independent task performance by persons with disabilities (Hodgdon, 1995). As defined by Quill (1997), visual cues as environmental prompts, are available to persons with disabilities after external physical, gestural, or spoken adult prompts are faded and may become a permanent reference for completing tasks (Browder & Shapiro, 1985)

Use of picture cues to prompt initiation and completion of tasks can be divided into two subgroups according to Lancioni and O’Reilly (2001): those which target multi-step tasks and those which sequence activities through a schedule.

Picture cued multi-step tasks. Eight studies were identified using picture cues to prompt students to complete multi-step tasks. Each step of a task analysis is depicted in a picture or photograph and presents the sequential steps comprising a single task (Allen, White, & Test, 1992). Students then look at each photograph prior to or during completing of a step. Picture prompts were demonstrated as effective in increasing vocational or job tasks (6 studies), daily living (1 study), and meal preparation (1 study).

Picture referencing in a study by Martin, Mithaug, and Frazier (1992) provided five students with moderate intellectual disabilities the necessary prompting to independently assemble a 13-step, 31 piece chair. Four of the five students further completed a more complex assembly of a love seat and settee using picture referencing. Students were able to look at a picture prior to attempting a step and self-correct using line drawings.

Use of a photograph task analysis permitted a 40-year-old man with profound mental retardation and atypical psychosis to perform vocational chores (dusting tables, setting a table, and vacuuming) independent of the instructor and to maintain this behavior for six months (Steed & Lutzker, 1997). Johnson and Miltenberger (1996) likewise found that pictures were effective in prompting three workers with mild to moderate mental retardation to complete packaging tasks which were changed daily while Copeland and Hughes (2000) found photographs to be effective prompts for two high school students with severe disabilities to initiate and complete job tasks. Agran, Fodor-Davis, Moore, and Martella (1992) found one of three participants in their study unable to respond to peer delivered instruction until picture prompts (paired with instructor prompts) were added to the task of taking customer orders and preparing a sack lunch. In a comparison study, photographs and line drawings for completing steps of a vocational assembly task were found to be more effective than physical demonstration for students with moderate intellectual disabilities. No difference was found for students
with severe and profound intellectual disabilities (Martin, Mithaug, & Burger, 1990).

Three students with autism were successfully able to use pictures to self-manage their behaviors to complete daily living skills (i.e. setting a table, making a bed) in the absence of an instructor, generalize their behavior across settings and tasks, and maintain behaviors after two months. Color photographs of steps of the task analysis were inserted in a photo book, one picture per page, with a “smiley face” sticker on the last page to prompt self-reinforcement (Pierce & Schreibman, 1994).

In the only study evaluating picture prompts for food preparation, Singh, Oswald, Ellis, and Singh (1995) used a picture based cookbook to prompt three adults with mental retardation to independently prepare a dessert within community-based settings. Follow-up checks spread over 6 months indicated maintenance of task performance using the picture prompts and generalization to a novel setting.

**Picture cued activity schedules.** Picture activity schedules sequence different daily activities or tasks familiar to the student in order to increase engagement, participation, and independent transitioning from one activity to the next (Lancioni & O'Reilly, 2001). A sequential series of pictures or photographs may further illustrate multiple tasks for completing a more complex activity (Copeland & Hughes, 2000). Persons may follow a preset activity schedule or participate in determining what they want to do and the order in which they will do it (self-scheduling). Nine studies were identified which used picture schedules to increase on-task, task engagement (7), transitioning between activities (1), and task initiation (1).

During classroom center activities, on-task and on-schedule behaviors of four children with high functioning autism improved when a picture activity book was present and decreased when the book was removed during a study by Bryan and Gast (2000) using an A-B-A-B design.

Within both home and school settings, Irvine, Erickson, Singer, and Stahlberg (1992) used picture schedules as a self-management tool to prompt four high school students with moderate to severe mental retardation to initiate behavioral tasks. Visual daily schedules were used to facilitate transitions for two children with autism at home and in the community using both photographs and line drawings. Portable activity schedules in the form of a photo album and strip of foam board attached to an automobile dashboard resulted in decreased latency between presentation of instruction and time to begin a task and decreased teacher-delivered verbal and physical transitioning prompts (Dettmer, Simpson, Myles, & Ganz, 2000). Likewise, photographic activity schedules were used by MacDuff, Krantz, and McClannahan (1993) to teach four boys with autism to self-initiate leisure activities within a group home setting. Photographs were placed in a three ring binder and participants repeated the steps of turning a page, initiating and completing an activity until three tasks were completed.

An auditory feature was added to a set of picture cards by Lancioni, Brouwer, Bouter, and Coninx (1993). A watch, worn by two adults with intellectual disabilities, emitted an auditory cue at prescheduled times to prompt participants to take a card from a box and perform the scheduled activity. Results indicated that the combination of auditory cues for initiating activities and picture cards describing the activities were effective in increasing activity engagement for both participants.

Parents and classroom assistants have also been taught to use activity schedules to increase the independence of persons with disabilities. Parents successfully implemented photographic schedules within home settings to increase sustained and independent participation of their children with autism in a study by Krantz, MacDuff, and McClannahan (1993) while teacher assistants decreased their use of prompts to encourage three students with varying disabilities to follow photograph activity schedules (Hall, McClannahan, & Krantz, 1995).

Rather than learning to follow an existing schedule, Brown (1991) addresses the importance of designing daily activity schedules with regard to a person’s preferences and individual life style including the choices of what to do, what not to do, and when to do it. Two studies evaluated the effects of self-scheduling when using picture activity schedules. In a group home setting, Bambara and Ager
<table>
<thead>
<tr>
<th>Reference</th>
<th>Participants</th>
<th>Self-Management Skill (Dependent Variable)</th>
<th>Design</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Picture Cued Multi-Step Tasks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agran et al. (1992)</td>
<td>n = 3 CA = 14–16 yrs Moderate, Severe ID</td>
<td>Complete job task at school: make sack lunch</td>
<td>Multiple baseline across participants</td>
<td>2 of 3 students made sack lunches in correct sequence. Generalization across novel customers.</td>
</tr>
<tr>
<td>Copeland &amp; Hughes (2000)</td>
<td>n = 2 CA = high school Severe ID</td>
<td>Initiate and complete job task in high school faculty dining room: clean &amp; set tables, sweep</td>
<td>Multiple baseline across participants</td>
<td>Independent task initiations increased. Task completion increased for 2 of 3 students.</td>
</tr>
<tr>
<td>Martin et al. (1990)</td>
<td>n = 20 CA = 13–17 yrs Mild, Moderate, Severe, Profound ID</td>
<td>Complete job task in classroom: vocational assembly</td>
<td>Latin Square</td>
<td>Mild, Moderate: photographs more effective than demonstration. Severe, Profound: no difference between photographs and demonstration.</td>
</tr>
<tr>
<td>Martin et al. (1992)</td>
<td>n = 5 CA = 15–20 yrs Moderate ID</td>
<td>Complete job task in school room: furniture assembly</td>
<td>Multiple probe across tasks, replicated across students</td>
<td>Picture &amp; video referencing more effective than assembly photographs. Picture referencing effective in 13 step assembly task. 4 of 5 students completed more complex tasks with picture referencing.</td>
</tr>
<tr>
<td>Singh et al. (1995)</td>
<td>n = 3 CA = adults Profound ID</td>
<td>Meal preparation in apartment setting: e.g. pineapple mouse</td>
<td>Multiple baseline across subjects</td>
<td>Acquisition and maintenance of multi-step meal preparation.</td>
</tr>
<tr>
<td>Steed &amp; Lutzker (1997)</td>
<td>n = 1 CA = 40 yrs Profound ID</td>
<td>Complete vocational chores in community &amp; day center: dusting, setting table, vacuuming</td>
<td>Multiple probe across tasks</td>
<td>Picture prompts effective across vocational tasks.</td>
</tr>
<tr>
<td>Reference</td>
<td>Participants</td>
<td>Self-Management Skill (Dependent Variable)</td>
<td>Design</td>
<td>Results</td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------------------------------</td>
<td>--------------------------------------------</td>
<td>---------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Anderson et al.</td>
<td>$n = 3$</td>
<td>Engagement in group home daily activities</td>
<td>Alternating</td>
<td>Greater participation for 2 of 3 adults when self-scheduling available.</td>
</tr>
<tr>
<td>(1997)</td>
<td>CA = adults Moderate, Severe ID</td>
<td></td>
<td>treatment</td>
<td></td>
</tr>
<tr>
<td>Bambara &amp; Ager</td>
<td>$n = 3$</td>
<td>Engagement in home &amp; community leisure activities</td>
<td>Multiple probe across participants</td>
<td>Increased participation in self-directed leisure activities and diversity of activities using self-scheduling.</td>
</tr>
<tr>
<td>(1992)</td>
<td>CA = 31–57 yrs Moderate ID</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bryan &amp; Gast</td>
<td>$n = 4$</td>
<td>On-task &amp; on-schedule for classroom center activities</td>
<td>A-B-A-B withdrawal</td>
<td>Student performance rose with graduated guidance and visual schedule package.</td>
</tr>
<tr>
<td>(2000)</td>
<td>CA = 7–8 yrs Autism</td>
<td></td>
<td></td>
<td>Performance generalized with schedule to novel activities.</td>
</tr>
<tr>
<td>Dettmer et al.</td>
<td>$n = 2$</td>
<td>Transitions between activities in home or community</td>
<td>A-B-A-B withdrawal</td>
<td>Decrease latency to begin activities with visual schedule.</td>
</tr>
<tr>
<td>(2000)</td>
<td>CA = 5, 7 yrs Autism</td>
<td></td>
<td></td>
<td>Decrease in teacher verbal &amp; physical prompts for transitioning with visual schedule for 1 student.</td>
</tr>
<tr>
<td>Hall et al.</td>
<td>$n = 2$</td>
<td>Classroom aides prompting of students &amp; student engagement in inclusive classroom</td>
<td>Nonconcurrent multiple-baseline</td>
<td>Decrease in multiple prompts used by aides after instruction. Independent engagement improved for both students.</td>
</tr>
<tr>
<td>(1995)</td>
<td>CA = 7, 8 yrs Fragile X Autism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irvine et al.</td>
<td>$n = 4$</td>
<td>Initiation of school tasks &amp; completion of home chores</td>
<td>Multiple baseline across subjects, replicated across settings</td>
<td>Independent use of schedules at home &amp; school. Maintained use of schedules across both settings.</td>
</tr>
<tr>
<td>(1992)</td>
<td>CA = 15–18 yrs Moderate, Severe ID</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krantz et al.</td>
<td>$n = 3$</td>
<td>Engagement in home living tasks in child’s home</td>
<td>Multiple baseline across subjects</td>
<td>Parents were taught use of photographic schedules. Children’s engagement increased with work &amp; play materials</td>
</tr>
<tr>
<td>(1993)</td>
<td>CA = 6–8 yrs Autism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Participants</td>
<td>Self-Management Skill (Dependent Variable)</td>
<td>Design</td>
<td>Results</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------------</td>
<td>----------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Lancioni et al. (1993)</td>
<td>$n = 2$</td>
<td>Activity engagement in institutional setting</td>
<td>Multiple baseline across subjects</td>
<td>Auditory cues from watch plus box of picture cards increased frequency of independent participation for 1 student. Visual cues introduced for 2nd student who did not respond to auditory watch.</td>
</tr>
<tr>
<td>MacDuff et al. (1993)</td>
<td>$n = 4$</td>
<td>On-task &amp; on-schedule after school activities</td>
<td>Multiple baseline across participants</td>
<td>Sustained engagement, independent activity change. Generalization to new photographs &amp; group home settings.</td>
</tr>
<tr>
<td><strong>Tactile Prompts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Auditory Prompts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Briggs et al. (1990)</td>
<td>$n = 4$</td>
<td>Multiple step task completion in home living suite &amp; school gymnasium: e.g. laundry, cleaning</td>
<td>Multiple baseline across settings &amp; tasks</td>
<td>Independent use of self-operated audio prompting system. Generalized use to new setting. Maintenance of task performance.</td>
</tr>
<tr>
<td>Davis et al. (1992)</td>
<td>$n = 3$</td>
<td>Fluency of vocational task performance in community food preparation site</td>
<td>Multiple baseline across participants</td>
<td>Musical tapes with interspersed performance cues increased vocational task fluency.</td>
</tr>
<tr>
<td>Grossi (1998)</td>
<td>$n = 2$</td>
<td>Accuracy of work performance, time spent working at community setting: e.g. cleaning</td>
<td>Reversal</td>
<td>Musical tapes with interspersed auditory prompts increased work performance for each employee.</td>
</tr>
<tr>
<td>Lancioni et al. (1995)</td>
<td>$n = 2$</td>
<td>Independent task performance (cooking) in school kitchen &amp; work room</td>
<td>Alternating treatments</td>
<td>Auditory and pictorial prompting systems equally effective.</td>
</tr>
<tr>
<td>Reference</td>
<td>Participants</td>
<td>Self-Management Skill (Dependent Variable)</td>
<td>Design</td>
<td>Results</td>
</tr>
<tr>
<td>----------------------------</td>
<td>-------------------------</td>
<td>---------------------------------------------</td>
<td>--------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Mechling &amp; Gast, (1997)</td>
<td>$n = 4$</td>
<td>Completion of daily living tasks in school setting</td>
<td>ABAB across two behaviors &amp; 4 students</td>
<td>Combination audio &amp; picture prompting system led to increased number of steps completed independently. Promoting device used to complete tasks &amp; generalization to novel (faculty) bathroom. Mixed results for maintenance of skills when auditory prompting faded. “Dramatic” increases in independent task completion. Generalization to untrained settings &amp; maintenance of performance.</td>
</tr>
<tr>
<td>Mitchell et al. (2000)</td>
<td>$n = 3$</td>
<td>Accuracy of vocational task performance completed in bathroom of self-contained classroom: cleaning mirror, sink, toilet</td>
<td>Multiple probe across behaviors, replicated across 3 students</td>
<td></td>
</tr>
<tr>
<td>Steed &amp; Lutzker (1999)</td>
<td>$n = 2$</td>
<td>Task completion in day program setting: making coffee, washing dishes, watering garden</td>
<td>Multiple baseline across behaviors, replicated across 2 adults</td>
<td></td>
</tr>
<tr>
<td>Taber et al. (1998)</td>
<td>$n = 5$</td>
<td>Transition through job tasks at community retail pet store and church</td>
<td>Alternating treatments, multiple probe across settings, with embedded withdrawal</td>
<td>Comparison of single &amp; multiple word prompts produced no significant differences. Both strategies effective in increasing independent task initiations. Significant decrease in # of teacher-delivered prompts. Generalization across settings.</td>
</tr>
<tr>
<td>Taber et al. (1999)</td>
<td>$n = 1$</td>
<td>On task behavior, teacher prompts during classroom writing activity &amp; school cafeteria vocational task</td>
<td>Multiple probe across settings, embedded withdrawal</td>
<td></td>
</tr>
<tr>
<td>Trask-Tyler et al. (1994)</td>
<td>$n = 3$</td>
<td>Completion of cooking skills within 2 on campus kitchens</td>
<td>Multiple baseline across behaviors</td>
<td>Students completed steps using tape recorded recipes. Generalization to untaught recipes.</td>
</tr>
<tr>
<td>Palmtop Personal Computers</td>
<td>$n = 40$</td>
<td>Completion of order-fulfillment tasks at software developers’ office</td>
<td>Beta testing, two-group, within subjects design</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Participants</td>
<td>Self-Management Skill (Dependent Variable)</td>
<td>Design</td>
<td>Results</td>
</tr>
<tr>
<td>---------------------------</td>
<td>--------------</td>
<td>--------------------------------------------</td>
<td>-------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Davies et al. (2002a)</td>
<td>n = 12</td>
<td>Completion of 8-item schedule within software developers' company</td>
<td>Two-group, within subjects design</td>
<td>Palmtop PC with Schedule Assistant software more effective than written schedule for increasing independent performance of vocational &amp; daily living tasks.</td>
</tr>
<tr>
<td></td>
<td>CA = 19–46 yrs Mild, Moderate ID</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Davies et al. (2002b)</td>
<td>n = 10</td>
<td>Completion of vocational tasks: pizza box assembly, software packaging</td>
<td>Beta testing, two group, within subjects design</td>
<td>Palmtop PC with Visual Assistant software using audio &amp; picture cues effective tool for improving task accuracy &amp; decreasing external adult prompting.</td>
</tr>
<tr>
<td></td>
<td>CA = 18–70 yrs Mild, Moderate, Severe ID</td>
<td></td>
<td></td>
<td>Personal Digital Assistant with alarm &amp; calendar effective in prompting student to complete tasks.</td>
</tr>
<tr>
<td>Ferguson et al. (2005)</td>
<td>n = 1</td>
<td>Time Management &amp; completion of routine daily tasks at school &amp; home</td>
<td>Multiple baseline across settings</td>
<td>VICAID (simplified Palmtop PC) with audio reminder and pictorial line drawings supported higher levels of work accuracy &amp; pace than picture prompt systems.</td>
</tr>
<tr>
<td></td>
<td>CA = 14 yrs Asperger Syndrome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Furniss et al. (1999)</td>
<td>n = 6</td>
<td>Completion of packaging &amp; assembly jobs in community work settings</td>
<td>6 single case experiments</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CA = 31–47 yrs Severe ID</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hersh &amp; Treadgold (1994)</td>
<td>n = 8</td>
<td>Completion of tasks at home (take out trash, yard maintenance), attending meetings, taking apartment key to work, taking medication</td>
<td>AB</td>
<td>NeuroPage with vibration text messages effectively provided reminders for participants to complete routine tasks.</td>
</tr>
<tr>
<td></td>
<td>CA = 21–49 yrs Traumatic Head Injury</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lancioni et al. (2000)</td>
<td>n = 6</td>
<td>Completion of tasks at day program: cleaning, table setting, food preparation</td>
<td>Alternating treatment</td>
<td>Palmtop PC with vibration, auditory prompts, &amp; picture cues was more effective than picture cards as measured by percentage of steps performed correctly.</td>
</tr>
<tr>
<td></td>
<td>CA = 23–47 yrs Severe Developmental Disability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n = 3</td>
<td>Completion of tasks from Study 1</td>
<td>Alternating treatment</td>
<td>Clustering of picture instructions on Palmtop PC effective for maintaining task performance.</td>
</tr>
<tr>
<td></td>
<td>Selected based on 90% maintenance performance from Study 1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(1992) used picture schedules of recreational activities each week with three adults with moderate intellectual disabilities. Adults choose picture cards of preferred activities and placed them in a sequenced activity book for each day. All participants learned to self-schedule their own leisure activities, participated in a wider variety of activities, and maintained their performance for up to 6 months after training. Anderson, Sherman, Sheldon, and McAdam (1997) took color photographs of three adults with mental retardation participating in housekeeping, personal care, and recreational activities within a group home setting. Adults made active choices about what to put on their schedules. Engagement time increased for participants when schedules were available.

In summary, picture cues as a form of light tech assistive technology, can increase skills for independently participating in activities whether it involves: (a) performance of a multi-task behavior; (b) initiation, participation, or sustained participation in an activity;
(c) participation in a wider variety of activities without reminders or prompts from others; or (d) decreased dependency on adults. Use of picture activity schedules have been shown to: (a) increase understanding of routines and expectations; (b) increase engagement time; (c) promote transition from one activity to the next; and (d) provide opportunities to make choices, and increase control over their daily lives.

What does not appear to be as clear from the available literature is whether the type of picture and/or form of presentation influence these behaviors. Although the majority of studies used color photographs of actual steps or activities (Johnson & Miltenberger, 1996; Pierce & Schreibman, 1994; Steed & Lutzker, 1997) line drawings were also reported as effective tools (Irvine et al., 1992; Martin et al., 1992; Singh et al., 1995). Digital photography and the ability to download images directly to print programs or edit them via personal computers make this form of imagery competitive with the ease of drawing pictures by hand. If there is a possibility that some students may not understand line drawings or may receive more concrete information from photographs, the use of color photographs may begin to dominate the field.

Assuming photographs are the visual image of choice for presenting pictures to students with disabilities a second question is whether photographs should depict finished products or individual steps of a task analysis. Martin et al. (1992) was the only identified study that addressed this issue. Results indicated that video and picture referencing to both be effective tools compared to photographs of completed items. A further question is who should be in the photographs and whether students respond better to seeing themselves performing a task or engaged in an activity (Anderson et al., 1997).

Picture cues have been organized onto poster board (Johnson & Miltenberger, 1996), binders (Steed & Lutzker, 1997), booklets and photograph albums (Copeland & Hughes, 2000), folders, and single sheets of papers with little attention to whether one system may be more effective than another. Questions remain as to how many pictures should be on one page and whether a more active response (touching a picture or turning a page) may have differing effects than looking at stationary pictures. While Irvine et al. (1992) required students to mark off pictures of completed steps, Johnson and Miltenberger found it necessary for one of three participants to touch photographs in addition to looking at them. Copeland and Hughes found that requiring students to touch the photograph increased initiation of a task step, but not completion of the task. Turning over the picture, however, resulted in greater task completion for both students in the study. Agran et al. (1992) likewise found turning over photographs to be an effective strategy while others (MacDuff et al., 1993) found placing photographs in binders and turning pages to be useful for keeping track of completed steps.

**Tactile Prompting**

The use of vibration as a self-management tool to prompt persons with disabilities to independently complete tasks has received minimal attention in the literature. Only two such studies were identified. A vibrating pager was used to prompt a student to read a message on the pager and complete daily tasks in a study by Epstein, Willis, Conners, and Johnson (2001) however, the study was conducted with a student with a diagnosis of AD/HD rather than an intellectual disability which is the focus of this review. Lancioni et al. (1991) used a small vibrator connected to a hearing aid ear mold to decrease the number and length of breaks in performance made by three persons with severe intellectual disabilities. The vibrator was wireless and activated by a computer system that registered inactivity. The only other identified studies using tactile prompting with students with intellectual disabilities (autism) focused on verbal initiations (Taylor & Levin, 1998) and social initiations (Shabani et al., 2002). The potential of this form of assistive technology as an antecedent self-management strategy may warrant more attention as either stand alone systems or to prompt students to use a visual prompting system.

**Auditory Prompting**

Post and Storey (2002) define auditory prompting as any prerecorded antecedent cue
(single or multiple word phrases), which increases the probability that a desired behavior will occur. In their review of the literature they identified nine studies between the years 1990-1999 using auditory prompting by persons with moderate to severe disabilities. In addition to these, three additional articles were identified in the current review (Lancioni et al., 2001; Lancioni, Klaase, & Goossens, 1995; Mitchell, Schuster, Collins, & Gasaway, 2000). Of the eleven identified studies, ten used auditory prompts recorded on portable cassette players with an attached headset worn by the participant. One used a combination of recorded speech and picture overlays on an augmentative communication device to provide visual and auditory prompts (Mechling & Gast, 1997). The majority of the studies (seven) evaluated self-instruction using a task analysis and recorded step-by-step prompts for completing a task. For example, Briggs et al. (1990) recorded a verbal script of 22 steps for operating a washing machine and used a bell tone to signal when to stop the cassette player and perform a step.

Three studies used recorded music interspersed with verbal prompts such as “Keep working,” and “Are you working?” to encourage participants to remain on task and complete vocational tasks (Davis, Brady, Williams, & Burta, 1992; Grossi, 1998; Taber, Seltzer, Heflin, & Alberto, 1999). Although only one student participated in the Taber et al. study, results have implications for the use of auditory prompts by persons with autism who are traditionally described as being stronger visual learners.

Taber, Alberto, and Fredrick (1998) compared single- and multiple-word prompts through an alternating treatment design. No significant difference between independent task performance was found however, when workers used multiple-word prompts they had a slightly shorter duration time for transitioning between tasks and more independent task changes than when following single-word prompts. A study by Lancioni et al. (2001) also supports the use of multiple word prompts on auditory systems. They found that self-operated audio systems could successfully present single word instructions corresponding to each step of a task analysis as well as clusters of instructions (i.e. two task steps presented in succession when “play” was selected on the audio cassette player).

In an interesting comparison study, Lancioni et al., (1995) found no significant difference between the use of picture prompts and auditory prompts for presenting steps of cleaning and food preparation to two adolescents with multiple disabilities. One difference noted in the results was that one participant found manipulation of the picture system to be more difficult due to physical disabilities. Results highlight the importance of an auditory system for persons who may have motor delays. These devices have received only limited evaluation with persons who have visual impairments (Lancioni et al., 2001) or may be non-readers and thus find such a system to be an effective tool for presenting stimulus prompts without an emphasis on visual systems.

Limited concern has been presented about the amount of time to develop the auditory prompting system in comparison to creating line drawing picture systems (Lancioni et al., 1995). Others report ease in preparing auditory systems compared to photograph systems (Mitchell et al., 2000).

A point about the design of auditory systems was raised by Taber et al. (1999) concerning whose voice should be recorded on such systems and would authoritative figures, persons liked, or well known by the student, have differing effects on response. Trask-Tyler, Grossi, and Heward (1994) suggest that the student’s voice (i.e. “Have I checked my___?”) may be an effective prompt for some.

Another feature to consider is the ability of subjects to replay auditory prompts when needed. Some systems allow the cassette recorder to stop automatically after each prompt and repeat prompts (Lancioni et al., 1995) while less expensive cassette players may not have these features and may be more difficult for persons with severe intellectual disabilities to operate. Such systems may only present a “beep” to cue the student to stop the cassette player. Digital cassette players allow students to start and stop at specific points during the recording.

Mechling and Gast (1997) addressed this issue in their selection of an augmentative communication device to present both visual (photographs on paper overlays) and auditory
(voice recording) prompts on one device to teach multiple steps of home living skills to four students with intellectual disabilities. Students could concretely see which step to complete and hear additional information about how to perform the step. Students could also repeat presentation of information by repressing a photograph for any step of the task analysis and listening again to the information. This form of assistive technology closely resembles the benefits described in the next section, which reports the effects of handheld computers as an antecedent prompting system.

**Palmtop Personal Computers**

Palmtop personal computers are portable handheld systems that employ features of a touch screen for input and multimedia capabilities (text, sound, digital photographs, and video clips) (Swan, Swan, Van Hover, & Bell, 2002) which have been used to promote self-management by persons with disabilities. Basic use of these devices has included only the features of text, sound, and light. A flashing light, and/or auditory “beep” have prompted students to look at the device displaying a text message. Messages reminded students to: complete tasks within a morning, school, or evening routine (Ferguson, Smith Myles, & Hagiwara, 2005); and attend a meeting or complete daily living skills such as taking out the trash (Hersh & Treadgold, 1994). Each increased independence while decreasing dependence on adult reminders. In a comparison study, Davies, Stock, and Wehmeyer (2002a) found a portable computer system with auditory prompts and text reminders to be more effective in prompting 12 students with mental retardation to perform tasks according to a schedule compared to a traditional written schedule.

Digital images have also been imported onto Palmtop personal computers to provide visual antecedent prompts. Furniss et al. (1999) used line drawings representing each step of a vocational task, along with auditory and vibration features, to prompt students. In addition to showing pictures, the Palmtop aid vibrated or emitted a recorded voice (“Press the key now”) to prompt students to use the device or advance to a new step if they were idle for a pre-determined amount of time. Lancioni, O’Reilly, et al. (1999) evaluated single pictures for each task step and multiple steps clustered into a picture. Step-by-step instructions presented in pictures were effective in prompting completion of tasks such as table setting, food preparation, and cleaning. Results of the second part of the study showed that students were able to maintain high levels of correct responding for these same tasks when instruction was clustered together under picture prompts.

Three separate comparison studies found pictures presented on a Palmtop personal computers to be more effective than pictures presented manually on cards as antecedent prompts for completing such tasks as table setting, food preparation, and cleaning (Lancioni, O’Reilly, Seedhouse, Furniss, & Cunha, 2000; Lancioni, Van den Hof, Boelens, Rocha, & Seedhouse, 1998; Lancioni, Van den Hof, et al., 1999). Students had a higher level of correct performance on task completion when using the portable handheld computer systems. Results were attributed to students mishandling cards and therefore skipping steps of the task analysis and failing to initiate steps when using the manual card system. The portable computer systems used in the studies also included auditory and vibrating features to prompt students to initiate the next step in the task sequence.

Additional information in the form of digital auditory and picture prompts have been delivered to study participants through Palmtop personal computers. Using touch screen features, participants accessed pictures corresponding to steps of a task analysis while hearing a description of how to complete the steps. Using these features, Riffel et al. (2005) decreased the need for external instructor prompts while increasing the number of steps students completed independently. Students further decreased the time for completing vocational and independent living tasks. Likewise, Davies, Stock, and Wehmeyer (2002b; 2003) incorporated digital images of each step of a task analysis and digital recordings of instructions corresponding to each step to prompt completion of vocational tasks. Each study incorporated features that allowed students to repeat auditory step directions when needed by pressing a “Play” or “Start” button.
and moving to the next step by pressing a “Done” or “Next” button on the touch screen. As a newly emerging technology, Palmtop personal computers presently represent the “highest tech” form of technologies used as antecedent prompts for self-management among persons with intellectual disabilities. These systems are commercially available and their use is becoming more prevalent in society, making them socially acceptable and less stigmatizing than systems developed specifically for a person with disabilities (Davies et al., 2002a). Unlike some technologies, the capabilities of Palmtop personal computers will continue to increase (Swan et al., 2002) including software that can be individualized to meet the needs of students.

Similar issues arise when using Palmtop personal computers as when using auditory or picture systems alone. Whose voice should be on the system? What images should appear on pictures? Lancioni et al. (2000) were able to reduce the number of instruction “occasions” by clustering multiple steps into one picture once students reached independent performance with individual steps presented on a Palmtop system. The ability to readily make adjustments in the complexity of antecedent prompts as the learner’s performance improves or tasks change has further been described as an advantage to computer-based systems (Lancioni et al.).

The use of computer-based systems compared to paper-based systems such as picture cards, photograph albums, and lists, may hold some distinct advantages. As demonstrated by Lancioni, O’Reilly et al. (1999) and Lancioni et al. (1998) students may lose their place or mishandle manual systems. Computer-based systems may allow better coordination and completion of steps of a task analysis and may decrease the chances that a student becomes lost in the sequence of task (Furniss et al., 1999). Further, auditory cuing can be added to minimize loss of attention to the task (Lancioni & O’Reilly, 2001; Montgomery et al., 1996) by prompting students to initiate the next step or activity. Compared to traditional cassette players, handheld computer-based systems readily provide: (a) repetition of steps; (b) control over auditory and visual cuing (Davies et al., 2002b); and (c) task specific auditory instructions with addition of visual supports through digital pictures or video clips specific to the task. These features may be important for persons who are not strong auditory learners.

For persons with physical disabilities, computer-based systems may employ a touch screen or be accessed through single switch activation or other assistive devices. Alerting and reminding capabilities of computer-based systems can provide self-cuing, problem solving, and information storage for persons with memory or attentional difficulties associated with traumatic brain injury or intellectual disabilities (Hersh & Treadgold, 1994).

Argued disadvantages of hand-held computer-based systems are the cost and maintenance of devices, but with rapid change in technology, prices will continue to decrease (Davies et al., 2002a; Swan et al. 2002). These systems also require some basic computer skills to set up the devices and select the best system. Although more time is needed initially to set up the systems and personalize them (Davies et al., 2002a), the process of taking pictures and recording audio files is described as very “straightforward” by Davies et al. (2003).

Finally, although Kimball, Kinney, Taylor, and Stromer (2003; 2004) outline instructions on how to create computer-based activity schedules with photographs and video models using Microsoft Power Point, to date no research base exists to support this new and creative use of high tech systems in providing students with visual, auditory, and animated cues for following and transitioning between activities or use in other forms of self-management. The promising effects of video prompting for teaching new skills (watching a video segment, making an immediate response, and advancing or repeating a video clip on a computer-based program) (Mechling, 2005) sets the stage for using such a system to create self-management systems within a computer-based program.

Discussion

The purpose of this paper was to review the application of assistive technology to increase the independent, self-management skills of persons with intellectual disabilities through self-prompting devices. Each of the 40 identi-
fied studies reported positive results when using assistive technology as an antecedent prompt for eliciting a target response while reducing the need for instructor prompts. Permanent prompts, whether auditory, visual, or automated, provided the assistance and support needed for persons with intellectual disabilities to participate in everyday activities that they could not perform with total independence. Twelve studies reported generalized responding to new activities or tasks and fifteen reported prompting strategies that guided and maintained performance over time.

Research has shown that assistive technology can facilitate learning, increase access, and serve as a tool to compensate for specific challenges associated with a disability (McGregor & Pachuski, 1996). The question appears to no longer be whether assistive technology can benefit persons with disabilities, but rather how best to use it and how it can be improved. The challenge for many educators is to remain abreast of technology. Many interventionists remain unaware of the types of devices available (Wehmeyer, 1999) and the varying uses outside those identified by the manufacturer. Barriers to funding and lack of information prevent proper identification and use of devices (Wehmeyer) and teachers continue to express concern about equipment complexity and lack of support and training (McGregor & Pachuski). It does little good when assistive technology devices are kept on a shelf or in a closet because practitioners do not know how to use them. Since its early inception, technology has become smaller, cheaper, more powerful, and easier to use (Lewis, 1998), yet in spite of these benefits and steadily decreasing costs, persons with intellectual disabilities continue to underutilize assistive technology (Wehmeyer).

A continued challenge is to examine applications of various technologies to determine which properties are the most beneficial and suitable to different individuals (Lancioni & O’Reilly, 2001; Lewis, 2000). One system does not fit all and future research and application may need to examine students’ characteristics that relate to the effectiveness of one procedure over others. For example, students with autism are characterized as being stronger visual learners (Quill, 1997) which supports the use picture systems or automated systems incorporating digital images. Language or intellectual disabilities may hinder the use of strictly auditory systems while persons with attentional difficulties may require auditory cuing, and those with physical disabilities may have difficulty manipulating picture systems. Martin et al. (1990) found that while photographs and line drawings were effective self-management tools for persons with moderate intellectual disabilities, there were no differential effects between these systems and physical demonstration for persons with severe and profound intellectual disabilities. Multiple use and forms of assistive technology make it a promising avenue for meeting the individual and diverse needs of learners. Its potential for providing additional, animated information, incorporating video and multi-media features may further address the unique and multiple learning styles of persons with disabilities.

Just as one system does not fit all, even one system may not permanently fit each person. Even as a person’s independence increases, it may not be appropriate to remove the system, yet it may not be appropriate to continue the identical system. A learner may need a more advanced system which can: (a) adjust with learning of a task; (b) provide fewer prompts; (c) cluster steps under one visual prompt (Lancioni, O’Reilly, et al., 1999); or (d) combine auditory prompts into multiple messages. Use of assistive technology as a form of antecedent prompting can address the need for systems to adjust as student needs change.

The nature of the task may also influence the effectiveness of one form of prompting over another. Nailos, Thomas, Whitman, and Maxwell (1994) found that students performed better with verbal-visual instructions when tasks were familiar to them, but visual cues alone were more effective when task stimuli were unfamiliar to the student. Attention needs to be paid to matching the most effective strategy to the task and the learner in order to promote independence across environments.

In conclusion, this review supports the current use of assistive technology as a self-management tool and recognizes the need for further expansion of its use as the field of technology continues to evolve. There ap-
pears to be a need for careful consideration of the capabilities of assistive technology as stimulus prompts that can be self-operated by individuals with disabilities. Creative applications should continue as persons with disabilities seek means to exert control over their lives, decrease dependence on others, and promote greater inclusion in community settings and daily activities.

References


Hall, L. J., McLannahan, L. E., & Krantz, P. J.
Hersh, N. A., & Treadgold, L. G. (1994). Neu-
Harchik, A. E., Sherman, J. A., & Sheldon, J. B.
Irvine, A. B., Erickson, A. M., Singer, G. H., &
Lancioni, G. E., Coninx, R., Manders, N., Driessen,
Kimball, J. W., Kinney, E. M., Taylor, B. A., & Stro-
Kimball, J. W., Kinney, E. M., Taylor, B. A., & Stro-
30, 208–217.
30, 211–227.
NeuroPage: The rehabilitation of memory dysfunc-
In K. A. Quill (Ed.), Teaching children with autism: Strategies to enhance communi-
185–197.
problems through the use of visually supported communication.
A coordinated program to transfer self-management skills from school to home. Education and Training in Mental Retarda-
Samplin, R. G. (1996). The direct and generalized effects of self-instructions and picture prompts on vocational task perform-
Behavioral Interventions, 11, 19–34.
Kimbell, J. W., Kinney, E. M., Taylor, B. A., & Stro-
5–12.
15, 235–238.
16, 235–238.
Behavioral Interventions, 10, 237–244.
187–197.
dents through automatic prompting or peer supervision. Journal of Developmental and Physical Dis-
bilities, 3, 115–128.


Received: 20 February 2006
Initial Acceptance: 20 April 2006
Final Acceptance: 1 July 2006
Abstract: Federal laws mandate that students with cognitive disabilities receive instruction in academic skills and also support the importance of teaching self-determination. The purpose of this literature review was to synthesize intervention research examining effects of self-determination interventions on academic skills for students with cognitive disabilities. Findings indicated that the majority of self-determination interventions affected skills that directly support academic performance (e.g., organization of academic assignments); however, traditional academic skills such as math productivity and spelling accuracy were also positively affected. The review included analysis of the strength of results, research design quality, and discussion of implications for researchers and practitioners.

Outcomes for students with mental retardation represent some of the poorest postschool outcomes of any disability group, including being the least likely to (a) leave high school with a diploma, (b) be involved in organized community groups, and (c) be engaged in work, postsecondary education, or work preparation (National Longitudinal Transition Study 2 [NLTS2], 2005). However, this population of students is the most likely disability group to be living on their own and parenting two years out of high school (NLTS2).

One factor that has been associated with positive outcomes for school leavers with mental retardation is self-determination. Wehmeyer and Schwartz (1998) noted a correlation between level of self-determination for adults with cognitive disabilities, as measured on a normed scale, and quality of life, measured on a questionnaire. Wehmeyer and Schwartz (1997) also demonstrated a correlation between high levels of self-determination and financial autonomy as demonstrated by maintenance of a checking or savings account, employment (full or part-time), and wages earned. Wehmeyer and Palmer (2003) extended these results for the same group of participants in a follow-up study and found a positive relationship between self-determination and work-related benefits earned, such as vacation and sick leave.

There have been several literature reviews measuring the effects of self-determination interventions for students with disabilities, including students with mental retardation or developmental disabilities. First, Harchik, Sherman, and Sheldon (1992) reviewed the effects of various self-management interventions including self-evaluation, self-regulation, and self-instruction on various behaviors (e.g., on-task, disruptive, social skills, in seat, work rate) of individuals with developmental disabilities. Authors concluded that most often these interventions were combined with other procedures such as token motivational systems or prompts. However, they found that when “self” managed interventions were compared to “other” managed interventions, the “self” interventions were more effective (Harchik et al.). Second, Algozzine, Browder, Karvonen, Test, & Wood (2001) conducted a meta-analysis of self-determination interventions across
all disability groups and ages and determined that students with mental retardation were most frequently taught choice-making skills. Goal-setting in conjunction with other instructional strategies was determined to be effective for improving task performance in Copeland and Hughes’ (2002) review of goal-setting interventions with individuals with mental retardation. Next, Shogren, Fagella-Luby, Bae, and Wehmeyer (2004) conducted a meta-analysis of choice interventions within academic, daily living, and vocational activities for individuals with developmental disabilities and found questionable results across the studies. Finally, Test, Fowler, Brewer, and Wood (2005) examined the effects of self-advocacy interventions across disability categories and determined that students with mental retardation were most frequently taught self-awareness and assertive communication skills in an effort to improve their participation in their IEP meetings.

The majority of studies examined in these literature reviews investigated self-determination intervention effects on behaviors related to classroom success such as “on-task behavior” (e.g., Copeland & Hughes, 2002; Harchik et al., 1992), aggressive behavior (e.g., Shogren et al., 2004), assertive communication skills (e.g., Test et al., 2005), or specific self-determination behaviors such as participation in IEP meetings (e.g., Test et al.). Although some of these literature reviews included studies that measured academic dependent variables, none looked exclusively at effects of self-determination interventions on academic outcomes.

In addition to having self-determination skills, another predictor of postschool success for students with disabilities is academic performance in such areas as reading, writing, and math (Benz, Yovanoff, & Doren, 1997). Benz et al. and Raskind, Goldberg, Higgins, and Herman (1999) correlated school factors and postschool success for students with disabilities. Raskind et al. concluded that higher levels of education attained and reading and math achievement were among predictors of post-school success for students with disabilities. In a post-school survey of students with and without disabilities, Benz concluded that the level of reading, writing, math, and problem-solving skills students had at the time they exited school were each significantly correlated with being competitively employed. These findings underscore the importance of academic achievement for all students.

While students with mental retardation and other developmental disabilities have most frequently been excluded from academic instruction and access to the curriculum experienced by their peers without disabilities (NLTS2, 2003), the emphasis on academic skill instruction for all students, including those with mental retardation and developmental disabilities (MRDD), is increasing in light of legal mandates for access to the general curriculum (IDEA, 2004) and participation by all students in the state and district assessment process (NCLB, 2001). In addition, strong academic skills appear to contribute to post-school success (Benz et al., 1997; Raskind et al., 1999).

Wehmeyer, Field, Doren, Jones, and Mason (2004) analyzed this push for students to engage in the general curriculum and concluded that self-determination skills can be taught within the current educational framework. The results of the Agran, Blanchard, Wehmeyer, and Hughes (2002) study designed to investigate the effects of self-regulated problem-solving instruction on specific classroom behaviors demonstrated that self-determination can be used as a mechanism for increasing access to the general curriculum by students with disabilities. The authors concluded that students with MRDD can be taught to problem-solve and self-manage their performance within academic content areas. Thus infusing self-determination skills into the general curriculum can be one method to overcome the finding of Wehmeyer, Agran, and Hughes’ (2000) national survey of special education teachers that teachers did not have enough time to teach self-determination skills, due to the need to focus instruction on other skills.

As a result, teachers may need tools that help them accomplish teaching both academic and self-determination skills simultaneously. Specifically, there is a need for examples of research-based strategies designed to combine self-determination and academic instruction for students with MRDD. Therefore, the purpose of this review was to identify, describe, and synthesize studies that have ex-
examined the effects of self-determination interventions on the academic skills of students with mental retardation and developmental disabilities.

Method

Selection Procedures

Studies were identified for this literature review through a four-step process. First, authors conducted a computer search of the ERIC database using full and truncated versions of the following terms: student, child, self determination, self monitoring, choice making, problem solving, goal setting, decision making, self regulation, self advocacy, self understanding, self awareness, self efficacy, self directed, self instruction, self evaluation, student involvement, reading, mathematics, writing, written expression, science, social studies, behavior, general education, general curriculum, social skills, special education, and disabilities. Second, authors completed a hand search for the years 2000–2005 of 10 major special education journals: Behavioral Disorders, Career Development for Exceptional Individuals, Education and Training in Developmental Disabilities, Exceptional Children, Journal of Applied Behavior Analysis, Journal of Positive Behavior Interventions, Journal of Special Education, Learning Disability Quarterly, Learning Disabilities Research and Practice, and Remedial and Special Education. Third, an additional electronic search was conducted to add the terms self-advocacy, self awareness, self efficacy, personal assistance, help, request, and recruit, after noting the term self-recruitment in association with self-advocacy during the hand search. Finally, we examined reference lists from literature reviews on self-determination (i.e., Algozzine et al., 2001), self-recruitment (i.e., Alber & Heward, 2000), goal-setting (i.e., Copeland & Hughes, 2002), self-management (i.e., Shogren et al., 2004) and self-advocacy (i.e., Test et al., 2005).

To be included in our review, studies needed to meet the following criteria. The criteria included articles (a) published through May 2005 in peer-reviewed journals, (b) that included at least one student participant (pre-K through post-secondary education) with an identified disability, as defined by IDEA (2004) or a diagnosis of ADHD, (c) in which the intervention and data collection took place in a school setting, including college or university settings, (d) in which the independent variable taught a self-determination skill, (e) that measured at least one dependent variable that was an academic skill, and (f) that were experimental (e.g., single-subject or group experimental), pre-experimental (e.g., pretest-posttest), or qualitative analyses of an intervention.

With regard to the independent variables of included studies, a self-determination intervention included one or more of the following skills: choice making, decision making, problem solving, goal setting and attainment, self-advocacy, self-awareness, self-management, or self-efficacy (Algozzine et al., 2001; Wehmeyer, 1999). Self-determination had to be the primary component or a portion of an intervention package for which a component analysis allowed for observation of the incremental effects of the self-determination intervention.

For dependent variables, academic skills were defined as tasks within any course or subject area that involved academic skills such as reading, writing, math, or spelling. For example, a student’s ability to read the labels of parts in an automotive maintenance course was considered an academic task; however, a student’s ability to assemble pens in a vocational preparation class that did not require reading, counting, or writing was not. Academic skills were noted as measures of: (a) quality, defined as accuracy and/or fluency on an academic task or as indicated by a holistic score or rating on a quality checklist; (b) productivity, defined as amount of task completed; (c) standardized academic assessment performance; and/or (d) other (e.g., asking for help on an academic task, following instructions to complete an academic task, organizational skills directly associated with completing an academic task). See Table 1 for detailed explanations of each study’s academic dependent variable(s).

While behaviors such as “on-task” may affect classroom performance, in this literature review we chose only to include studies that measured behaviors directly related to academic performance. We chose to narrow our focus for two reasons. First, much of the self-determination literature to date for students with mental retardation has focused on social behavior (e.g., Harchik et al., 2000).
or non-academic skill performance (e.g., Copeland & Hughes, 2002). Second, teachers are expected to teach academic skills as well as self-determination skills to students with disabilities (Wehmeyer et al., 2004). Narrowing the focus was consistent with the purpose of our review; an attempt to locate examples of evidence-based strategies that teach both.

Scope of Review

Fifty-eight studies were identified that met the inclusion criteria. Due to the large number of studies identified, findings were divided into disability categories. Fourteen studies included participants with MRDD; however, for three studies the results were not reported in a manner to determine which participants were labeled MRDD (i.e., Anderson-Inman, Paine, Deutchman, 1984; Palmer & Wehmeyer, 2003; Taylor-Ritzler, Balcazar, Keys, Hayes, Garate-Serafini, & Espino, 2001). Therefore, these three studies were excluded from this review. This review summarizes the findings from the 11 studies that clearly delineated intervention effects for the target population.

Procedures for Descriptive Review

For each study a researcher-developed coding form was used to gather information on age, disability, and ethnicity of participants; setting; research design; dependent variables; independent variables; and results. Inter-rater reliability for this information was calculated in two stages. The first two authors read titles and abstracts of the 191 articles identified in the initial computer search and decided independently whether or not the articles should be included. They agreed on the inclusion of 91.6% of articles before proceeding to the individual review stage. All disagreements were discussed and agreement was reached. As a result, 58 studies met the inclusion criteria and 11 were identified that analyzed results specifically for students with MRDD.

Next, inter-rater agreement was calculated for the descriptive information gathered from 4 (36.4%) of the 11 MRDD articles by two of the authors. Inter-rater agreement was calculated by dividing the total number of agreements by the total number of items reviewed and converted to a percentage. Consensus was reached on disagreements before information was recorded as results. Inter-rater reliability for this portion of the review ranged from 92.4% to 95.7% with a mean of 94.4%.

Procedures for Evaluation of Study Quality

Quality indicators for group designs described by Gersten et al. (2005) and for single subject designs described by Horner et al. (2005) were used to analyze the research quality of each study. Gersten et al. (2005) delineated 10 essential quality indicators for group and quasi-experimental research and Horner et al. (2005) identified 21 for single-subject research. Two authors independently reviewed four (36.4%) of the studies using forms based on Horner et al. (2005), since the vast majority of studies (n = 10) used single-subject designs. The same methods for calculating reliability were used as described previously. Reliability ranged from 76.0 to 100% with a mean of 86.0%.

Calculation of Intervention Effects

Gersten et al. (2005) noted the importance of reporting the power of an intervention, because it indicates the probability of an accurately analyzed statistically significant result. In single-subject research, systematically analyzing the effects of an intervention is critical to understanding its external validity (Scruggs, Mastropieri, & Casto, 1987). Several methods have been used to assess the strength of the results of single-subject studies (Campbell, 2003; Scruggs et al.; Scruggs & Mastropieri, 2001) and effect sizes are a frequently used power analysis in group experiments (Gersten et al.). This review reports the strength of the effects on the academic variables for the studies that used single subject designs and reported data in a manner that allowed for such calculations; however, the one group design study did not provide information to allow calculations of an effect size.

Based on the single-subject meta-analysis works of Scruggs and Mastropieri (2001), Scruggs et al. (1987), and Algozzine et al. (2001), we decided to use percentage of non-overlapping data points (PND) to analyze the
### TABLE 1
Study Methods

<table>
<thead>
<tr>
<th>Reference</th>
<th>Independent Variable</th>
<th>Academic Dependent Variable(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brooks, Todd, Tofflemoyer, &amp; Horner (2003)</td>
<td>Self-management: Intervention determined through results of functional behavior assessment and resulting behavior support plan; taught (1) discrimination of on-task behavior through examples and nonexamples of on-task seat work and on-task group work; (2) operation of a cassette player using task analysis; (3) self-monitoring of on-task behavior instruction: (a) use of picture cue of tape player on student’s schedule, (b) record a + or – on card for on-task when tape prompts “now”; and (4) self-recruitment of praise with prompt of hand icon at every 6th checkpoint; received praise for accurate recording; earned activity with peer for on-task behavior; 12 30-minute instructional sessions.</td>
<td>1. Work completion noted as yes or no regarding completion of seat work in 4th grade class [Productivity on various academic assignments]</td>
</tr>
<tr>
<td>Hughes, Copeland, Agran, Wehmeyer, Rodi, &amp; Presley (2002)</td>
<td>Self-management: Students provided with (1) rationale for self-monitoring; (2) modeling, direct instruction, guided practice, corrective feedback on self-monitoring procedures; students were taught to place a check on self-monitoring form for (a) “find picture,” (b) “find picture caption,” (c) write name of tool” after performing each task on a modified worksheet in auto-mechanics class; 2–3 training sessions.</td>
<td>1. Percentage of correct written responses per worksheet, labeling auto repair tools [Quality of language arts assignments]</td>
</tr>
<tr>
<td>O’Reilly, Lacioni, Gardiner, Tiernan, &amp; Lacy (2002)</td>
<td>Decision making, problem-solving, self-management: Training scripts were modeled and role-played with verbalization of steps (self-instruction) to teach student to (1) decide the situation to determine the problem, (2) decide what to do to resolve the problem, (3) perform behavior, and (4) self-evaluate the effectiveness of the performed behavior; 40-minute sessions using 4–5 social situations; 1 session/day.</td>
<td>1. Percentage appropriate classroom behavior, indicated by (1) compliance with teacher instruction within 5 s and (2) when completing a written English or Gaelic assignment, raising hand to ask for assistance [Other]</td>
</tr>
<tr>
<td>Agran, Blanchard, Wehmeyer, &amp; Hughes (2001)</td>
<td>Problem-solving, goal-setting, self-management: Students learned to set a goal, monitor performance, evaluate the outcome, and provide self-reinforcement for desired responding; students received problem-solving instruction based on SDLMI that involved three instructional phases: setting a goal, taking action, adjusting goal or plan; included verbal instruction, modeling, guided practice, independent performance, and review.</td>
<td>1. Percent correct performance of organizational skills [Other] 2. Goal Attainment Scaling (GAS) [Other]</td>
</tr>
<tr>
<td>Agran, Blanchard, &amp; Wehmeyer (2000)</td>
<td>Goal-setting, self-management: Students (a) selected transition-related goals to work on, (b) learned to use problem-solving strategies to answer questions that would help them create and implement action plan, (c) implemented action plan, and (d) self-evaluated progress in achieving their goals.</td>
<td>1. Percentage correct on alphabetizing word cards [Quality of language arts assignment]</td>
</tr>
</tbody>
</table>

strength of effects of single-subject interventions. Data were reported in a manner that allowed for calculation of PND for individuals identified as MRDD in seven (64%) of the
studies reviewed. For these studies, PNDs were calculated as follows. First, the most extreme (highest for targeting high performance and lowest for targeting low performance) data point in the baseline phase was noted. Second, all points in the intervention phase that exceeded that point were counted and divided by the total number of points in the intervention phase (Scruggs et al.). This number was converted to a percentage and reported as the PND.

In reversal designs, phases with self-determination interventions were compared to all phases without self-determination (e.g., Moes, 1998). For studies using a multiple baseline design, we only calculated percentage of non-overlapping data points for students, skills, or settings in which an academic dependent variable was measured. For example, in a study that used a multiple baseline across students, the dependent variable may not have been an academic skill for all participants (e.g., Hughes et al., 2002). In those cases, we only counted data points in the academic tiers. Similarly, percentage of non-overlapping data points could only be calculated when the academic variable was graphed, rather than described in the results as a secondary or tertiary

<table>
<thead>
<tr>
<th>Reference</th>
<th>Independent Variable</th>
<th>Academic Dependent Variable(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Craft, Alber, &amp; Heward (1998)</td>
<td><strong>Self-advocacy:</strong> One-on-one self-recruitment training; included three parts: (a) instruction and role play, (b) morning prompts, and (c) end-of-day check and reward; instruction included discussion of rationale, modeling, “think aloud” procedures, and role-playing with feedback; 20-minute sessions; 2 sessions.</td>
<td>1. Percentage of items completed on spelling assignment [Productivity on language arts assignments] 2. Percentage of items correct on spelling assignment [Quality of language arts assignments]</td>
</tr>
<tr>
<td>Moes (1998)</td>
<td><strong>Choice-making:</strong> Two conditions—(1) no choice: tutor chose order of homework activities, sequence of problems, and stimulus items (e.g., pens, scissors); (2) choice: child chose order of homework activities, sequence of problems, and stimulus items.</td>
<td>1. Percentage of correct homework responses [Quality of homework assignments] 2. Number of homework trials (i.e., instances of child output on homework sheets that were legible and consistent with instructions) completed per minute [Productivity on homework assignments]</td>
</tr>
<tr>
<td>McCarl, Svobodny, &amp; Beare (1991)</td>
<td><strong>Self-management:</strong> Students taught during 20–30 minute sessions for 4 days: (1) rationale for target behavior discussed, (2) demonstration of examples and non examples of on task behavior, (3) participants discriminate between on and off task behaviors observed, (4) role play of on and off task behaviors, (5) evaluation of participant understanding of target behavior; participants self-assessed behavior during independent work and self-recorded “Y” or “N” when prompted through audio tape every 10–120 s.</td>
<td>1. Rate/minute digits (movements) written in answering arithmetic problems [Productivity on math assignments] 2. Percent correct of math problems completed during independent seatwork [Quality of math assignments]</td>
</tr>
<tr>
<td>Warner &amp; dejung (1971)</td>
<td><strong>Goal-setting:</strong> Under goal setting condition, participants were instructed to identify the number of words they thought they could get correct from a list of 10 words: (a) each participant recorded responses, (b) participant given 15 seconds after the experimenter’s second verbal reading of the word to respond, (c) verbal praise given for correct response, no verbal expression given for incorrect response.</td>
<td>Number of items correct on spelling task (when student received correct score on 10 out of 10 items on 5 consecutive trials testing ended) [Quality of language arts assignments]</td>
</tr>
</tbody>
</table>
variable. PND was only calculated for studies using an alternating treatment design, if a baseline measure was included.

For multiple baselines across students that measured the same skill, non overlapping data points were added across phases and were divided by the total number of intervention points across phases to determine PND for that variable. However, if more than one academic dependent variable was measured (e.g., math and reading), PNDs were calculated and reported for each dependent variable.

Median PNDs were calculated across the seven single-subject studies for which data were calculated for all academic dependent variables. In addition, median PNDs were calculated for specific self-determination components and specific academic dependent variables. Scruggs and Mastropieri (2001) noted that a PND from 70 – 90% can be interpreted as an effective intervention, 50-70% as questionable, and below 50% as ineffective.

Results

Student Demographic Information

One hundred fifty-six participants with MRDD were included in the 11 studies, with a range of 1 to 80 participants. Three studies (27.2%) included students from other disability populations or students without disabilities (i.e., Agran, Blanchard, Wehmeyer, & Hughes, 2001; Agran, Blanchard, & Wehmeyer, 2000; Rock, 2005). Participants with MRDD represented 21 (61.8%) of the 34 participants in these studies with disability and non-disability populations. All studies reported data on gender. Male participants numbered 104 (66.7%) and 52 (33.3%) were female. Only two (18.2%) studies reported data on the ethnicity of participants (i.e., Hughes et al., 2002; Rock; Warner & deJung, 1971). For these two studies 50% of participants were identified as White and 50% as African-American.

Setting

Seven (63.6%) took place in general education classrooms at the elementary, middle, and high school levels (i.e., Agran et al., 2000, 2001, 2002; Brooks, Todd, Tofflemoyer, & Horner, 2003; Craft, Alber, & Heward, 1998; Hughes et al., 2002; Rock, 2005), five (45.5%) were in resource or self-contained special education classrooms at various grade levels (i.e., Brooks et al.; Craft et al.; Hughes et al.; McCarl, Svobodny, & Beare, 1991; O’Reilly, Lancer, Gardiner, Tiernan, & Lacy, 2002), and three (27.3%) occurred in non-specific locations in public schools (i.e., Hughes et al.; Rock; Warner & deJung, 1971). Finally, two (18.2%) studies were conducted in clinic and non-integrated settings for individuals with disabilities (i.e., Moes, 1998; Warner & deJung, 1971), and one (9.1%) took place in a community work site (i.e., Agran et al., 2000).

Academic Variables Measured

Eighteen academic variables were measured across the 11 studies. This represents just less than half (45%) of all the dependent variables (academic and non-academic) measured in the studies in this review. Quality (accuracy or fluency) of various language arts assignments was measured in four (36.4%) studies (e.g., Hughes et al., 2002; Warner & deJung, 1971). While quality of math assignments, productivity in math assignments, and productivity of general assignments were each measured in two (18.2%) studies. Productivity on language arts assignments (i.e., Craft et al., 1998) and general assignment quality (i.e., Moes, 1998) were each measured in one (9.1%) study. Finally, other measures of academic performance that were not direct measures of accuracy or productivity in a reading, writing, or math task (e.g., Agran et al., 2002; O’Reilly et al., 2002) were measured in three (27.3%) studies, and represented six (33.3%) of the academic variables measured in the review.

Language arts dependent variables. Hughes et al. (2002) measured the accuracy of writing the labels for parts in an auto-mechanics
course, which was categorized as a quality measure on a language arts assignment. Accuracy of an alphabetizing assignment (Agran et al., 2000) and Warner and deJung’s (1971) accuracy on spelling assignments, were also placed in this category of dependent variables. One study (Craft et al., 1998) measured the percentage of items completed on a spelling assignment, after providing self-recruitment training to students.

Math dependent variables. Rock (2005) measured the accuracy and productivity of student independent math work with a computer software tool that generated unlimited practice assignments, individualized for each student. Two of the nine students in the study were diagnosed with developmental disabilities. McCarl et al. (1991) measured accuracy of math seatwork and calculated the number of digits written in completing math seatwork for three elementary age students diagnosed with mild and moderate mental retardation. Neither study specified the specific math operations completed by students in the study.

General academic measures. General measures of productivity included both homework (Moes, 1998) and seatwork (Brooks et al., 2003) assignment completion rates in math and language arts subject areas. General academic accuracy measures also included both accuracy on homework assignments (Moes) and accuracy on seatwork assignments (Brooks et al.).

Other dependent variables. Measures categorized as “other” for this review included the Goal Attainment Scaling applied to measuring academic goals met (e.g., Agran et al., 2002). Three studies (27.3%) examined students’ use of organizational and communication skills directly linked to academic performance including following directions to complete assignments (e.g., O’Reilly et al., 2002), organizational skills for assignment completion (Agran et al., 2001), and verbal contributions to class (Agran et al., 2002).

Self-Determination Interventions and Results

Self-management interventions were used in four (36.4%) studies. Choice-making, goal setting, and self-advocacy were the primary self-determination components of the intervention for one (9.1%) study each. Four (36.4%) studies implemented self-determination interventions using multiple component skills.

Self-management interventions. Four studies used self-management as their primary independent variable, First, McCarl et al. (1991) used tape recorded auditory prompts to teach students to self-monitor their on-task behaviors to improve academic accuracy and productivity. Hughes et al. (2002) taught students to use a checklist to self-monitor accuracy of assignment completion. Brooks et al. (2003) used auditory prompts to teach self-monitoring of on-task behaviors, accuracy, and productivity. Rock (2005) taught students a 6-step strategy for self-recording attention to task and performance using auditory prompts to increase student accuracy and productivity on math problems completed. Self-management interventions, which included self-recording attention to task or on-task behavior (e.g. McCarl et al., 1991), self-recording productivity on assignments (e.g., Rock), and use of a self-monitoring checklist (e.g. Hughes et al.), were used to increase (a) productivity on math assignments, (b) accuracy on language arts assignments, and (c) productivity on general class assignments. Self-management interventions also (a) decreased off-task behavior, (b) increased on-task behaviors, (c) increased accuracy on non-academic behaviors, and (d) improved self-monitoring skills.

Choice-making intervention. One study used choice-making as their independent variable. Moes (1998) provided students with autism with choices regarding the order in which assignments were to be completed. The intervention increased student productivity and accuracy on homework assignments, decreased disruptive behaviors, and improved the affect of participants in the intervention.

Goal setting intervention. Students were instructed to identify a goal regarding the number of spelling words they might get correct in the one study that used goal-setting alone as intervention to promote academic performance (Warner & deJung, 1971). The intervention increased student productivity and accuracy on homework assignments, decreased disruptive behaviors, and improved the affect of participants in the intervention.

Self-advocacy intervention. Role play, direct instruction, and beginning and end-of-day monitoring by instructors were used to teach students to self-recruit teacher assistance to improve spelling assignment accuracy and
productivity (Craft et al., 1998). The intervention increased student productivity and accuracy on class assignments and increased the frequency of self-recruiting for assistance and of teacher praise for the student.

**Multi-component interventions.** Of the four studies that used a multi-component approach to promoting students’ academic skills, three (75%) used the *Self-Determined Learning Model of Instruction* to guide students through a three step process of (a) setting a goal by identifying their strengths and needs (self-awareness) and engaging in choice-making; (b) creating a plan to meet the goal which may include decision-making, problem-solving, and self-advocacy; and (c) adjusting the plan based on self-evaluation of progress, problem-solving barriers, and self-monitoring steps toward the goal. In addition, O’Reilly et al. (2002) used role play and student self-instruction/self-verbalization of steps to move one student through a decision-making and problem-solving process to increase the student’s ability to follow directions to complete academic assignments. Multi-component interventions yielded positive results on (a) organizational skills to complete assignments (e.g., Agran et al., 2001), (b) asking for teacher assistance on academic assignments (e.g., Agran et al., 2002; O’Reilly), (c) following instructions on assignments (e.g., Agran et al., 2002; O’Reilly), and (d) accuracy on a language arts assignment (i.e., Agran et al., 2000). Multi-component interventions also demonstrated positive effects on non-academic goals such as social, daily living, or employment goal attainment in the three studies using the *Self-Determined Learning Model of Instruction* (i.e., Agran et al., 2000, 2001, 2002).

**Additional results.** Maintenance data were collected for five (45.5%) studies (Agran et al., 2001, 2002, 2000; Craft et al., 1998; Hughes et al., 2002). Studies demonstrated generally positive results on the academic variable for 2 to 24 days after intervention. Four (36.4%) studies also collected generalization data. Generalization data were positive for three (75%) of these studies (i.e., Agran et al., 2002; Craft et al.; Hughes et al.). Finally, social validity measures were gathered in five (45.5%) studies (e.g., Agran et al., 2002; Moes, 1998). All social validity data were positive, noting the value of the intervention, the impact on student skills, or the practicality of the intervention.

**Percent Non-Overlapping Data Points**

The PNDs of the academic dependent variable were able to be calculated for seven (63.6%) of the 11 studies and results and strengths of effects are summarized in Table 2. PNDs could not be calculated for four (36.4%) studies because (a) data were not graphed for the academic variable (Rock, 2005), (b) graphed results could not be clearly associated with the participants with MRDD, but results were described narratively for the population (i.e., Agran et al., 2000; Agran et al., 2001), or (c) the study used a group design and effect sizes could not be calculated with the data provided (Warner & deJung, 1971). PNDs for four (57.2%) of the seven studies were above 90%, indicating very strong results. Three of these studies demonstrated 100% non-overlapping data points between the phase of the study without the self-determination intervention and the use of the self-determination on the academic variable measured (Agran et al., 2002; Hughes et al., 2002; O’Reilly et al., 2002). In addition, Moes (1998) and Mo and Rolf (1998) demonstrated PNDs above 70%, whereas Craft et al. (1998) demonstrated poor results below 25%. The median PND across the seven studies was 85%, indicating strong effects (Scruggs & Mastropieri, 2001).

**PNDs for dependent variables.** We calculated the median PND across studies within categories of dependent variables to get a sense of the strengths of interventions in these academic areas. The median PND for measures of other academic skills, including assignment organization skills, was 100%. For general measures of academic productivity, median PND was 86%, while general measures of academic accuracy had a median PND of 85%. Math assignment productivity measures yielded a PND of 74%, while language arts assignment productivity yielded a PND of 21%. Finally, language arts quality yielded a median PND of 50% and math quality a PND of 11%.

**PNDs for independent variables.** Median PNDs for studies that used particular components of self-determination in the indepen-
dent variable were also calculated. Studies that used self-management interventions yielded a median PND for the academic variable of 84%. The choice-making intervention yielded a median PND of 81% homework assignment accuracy and productivity, while the self-advocacy intervention yielded a median PND of 11% on language arts productivity and quality. The goal-setting intervention that was used in the one group design study did not report data in a manner that allowed for calculation of effect sizes. Median PNDs for studies that used multiple components of self-determination in the intervention were 100%.

Quality Indicators

The quality indicators from Horner et al.'s (2005) and Gersten et al.'s (2005) descriptions of quality research in special education addressed by each study are described below.

Single subject designs. A condensed list of single-subject quality indicators is provided in Table 3. For example if a study’s “dependent variables are described with operational precision . . . (are) measured with a procedure that generates a quantifiable index . . . measurement . . . (are) valid and described with replicable precision . . . dependent variables were measured repeatedly over time” (Horner et al., 2005, p.174), the study is reported to have a “dependent variable that is operationally defined”. Similarly, if a study met indicators for social validity, including the “dependent variable is socially important . . . magnitude of change in the dependent variable . . . is socially important . . . implementation . . . is practical and cost effective” (Horner et al., p.174), the study was reported to have a “socially valid intervention.” A full description of quality indicators can be obtained from the first author.

Ten (100%) of the single-subject studies described the participants and setting for the intervention thoroughly. While all studies included gender information, only two (20.0%) studies included data on participants’ ethnicity. Five (50.0%) also described how participants were selected in a manner that would allow for replication.

The dependent variable was defined so that it could be quantified and measured and interobserver agreement was measured and above 80% in all single-subject studies. All studies also described the independent variable clearly, including a clear description of baseline and treatment conditions. However, procedural fidelity was measured in only one (10.0%) of the 10 studies (O’Reilly et al., 2002). The baseline phase demonstrated a stable pattern of behavior in eight (80.0%) instances. Seven (70.0%) studies also provided three replications of effect of the independent variable on the dependent variable through multiple baseline designs across three or more participants (e.g. Hughes et al., 2002) or through a reversal design with three demonstrations of effect (Moes, 1998). In one of the three studies that did not demonstrate three replications of effect, the study was a two-tiered multiple baseline design across settings (O’Reilly et al., 2002). All studies had interventions that were socially valid based on the social importance of the dependent variable, as well as the practicality of implementation and cost of the intervention.

Group designs. The one group intervention study provided a compelling rationale for its research and clearly articulated the purpose of the study. It also clearly described the participants in the study, but did not provide information on ethnicity of participants. Warner and deJung (1971) used an experimental design, and researchers took steps to ensure the equivalency of the groups. The intervention was clearly described for replication; however, procedural fidelity data were not collected. Methods of data analysis corresponded to the research questions and data gathered. Effect sizes were not reported in the analysis, nor were data provided amenable to calculation of effect sizes for the current review.

Discussion

The purpose of this review of literature was to describe studies that have examined the effects of self-determination interventions on the academic skills of students with mental retardation and other developmental disabilities. While the results summarized may not provide definitive instructions for practitioners, it is a solid beginning and provides the field with a direction for future research and suggestions for practice.
Eleven studies were analyzed for this review, and effects of self-determination interventions on academic skills for students with MRDD were positive. There were a total 156 participants, spanning 34 years of studies. Effects were generally stronger for organization skills in academic assignments than for more traditional academic measures, such as spelling as-

<table>
<thead>
<tr>
<th>Reference</th>
<th>Results for Academic Variable</th>
<th>Percent Non-Overlapping Data Points</th>
</tr>
</thead>
</table>
| Rock (2005) | 1. Some effect on number of math problems completed each day (i.e., mixed results)  
2. No effect on accuracy of math work | Unable to calculate strength of effects because results for academic dependent variable not graphed |
| Brooks et al. (2003) | 1. Work completion appears to have increased during intervention; however, minimal data collection make results inconclusive  
2. Generalization of skills occurred in one of two generalization settings | 94% PND for completed seatwork assignments w/self-management intervention |
| Agran et al. (2002) | 1. All 4 students demonstrated 100% accuracy on target behaviors, including following directions to complete assignments and contributions to class  
2. All students surpassed expectations on GAS | 100% PND for following directions & 100% PND for frequency of verbal contributions to class w/ SD intervention |
| Hughes et al. (2002) | 1. Student demonstrated correct written responses at 100% accuracy during training | 100% PND for writing correct answer w/ self-management intervention for students with mental retardation |
| O’Reilly et al. (2002) | 1. Increases in appropriate behaviors in both settings | 100% PND for frequency of behaviors w/ SD intervention |
| Agran et al. (2001) | 1. Marked increase in target performance for all students  
2. After changing reinforcer, each group’s data stabilized during the training condition  
3. Positive maintenance results for all three groups | Unable to calculate strength of effects for participants with MRDD due to manner in which data on the academic variable was reported |
| Agran et al. (2000) | 1. Mean performance increased for students with academic target behaviors  
2. Maintenance scores increased for majority of students | Unable to calculate strength of effects for participants with MRDD due to manner in which data on the academic variable was reported |
| Craft et al. (1998) | 1. Increases in completion of academic work  
2. Improved accuracy of academic work | 21% PND for completed spelling tasks w/ self-advocacy intervention 0% PND for accurate spelling items w/ self-advocacy intervention |
| Moes (1998) | 1. Increases in percentage of correct responses in choice condition for all participants  
2. Increases in number of homework trials completed per minute in choice condition for all participants | 77% PND for completed homework assignments w/ choice-making intervention 85% PND for correct homework w/ choice-making intervention |
assignment accuracy or math assignment productivity. Similarly effects were stronger for productivity of both math and language arts assignments than for accuracy measures.

Multi-component self-determination and self-management interventions were the most frequently studied independent variables. Choice making, goal-setting, and self-advocacy were the only other self-determination interventions used. While this review focused on academic dependent variables, such variables represented just under half of the total number of dependent variables measured by the studies, suggesting the lack of focus on academic skills in the self-determination literature for students with MRDD. However, while the wealth of the studies reviewed were from this century, the inclusion of studies from 1971 and 1984 indicate that using self-determination as an intervention to promote academic success for this population is not a completely new idea.

In terms of self-determination components taught to individuals with MRDD, our findings differ from Algozzine et al. (2001), who found that students with mental retardation were

---

**TABLE 2—(Continued)**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Results for Academic Variable</th>
<th>Percent Non-Overlapping Data Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>McCarl et al. (1991)</td>
<td>1. Two participants demonstrated a variable but clear increase from baseline to intervention in rate/minute digits written; one participant demonstrated an average increase of 38%</td>
<td>74% PND for math assignment productivity w/ self-management intervention</td>
</tr>
<tr>
<td></td>
<td>2. Two participants continued high rates of accuracy in intervention (no change from baseline); third participant demonstrated a clear increase in accuracy during intervention, however, was beginning increase in accuracy in baseline phase</td>
<td>11% PND for math assignment quality w/ self-management intervention</td>
</tr>
<tr>
<td>Warner &amp; deJung, (1971)</td>
<td>1. Overall mean score of participants making goal statements was greater than participants not making goal statements</td>
<td>Unable to calculate effect sizes with data provided</td>
</tr>
</tbody>
</table>

---

**TABLE 3**

**Study Design Quality**

<table>
<thead>
<tr>
<th>Quality Indicator</th>
<th>Participants described</th>
<th>Setting described</th>
<th>Selection described</th>
<th>Operationally defined DV</th>
<th>Inter-rater reliability</th>
<th>Replicable IV</th>
<th>Procedural fidelity</th>
<th>Experimental control</th>
<th>Demonstration of effects</th>
<th>Socially valid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock (2005)</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Brooks et al. (2003)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Agran et al. (2002)</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Hughes et al. (2002)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>O’Reilly et al. (2002)</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Agran et al. (2001)</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Agran et al. (2000)</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Craft et al. (1998)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Moes (1998)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>McCarl et al. (1991)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>
most frequently taught the self-determination component of choice-making. In this review of self-determination as a method of promoting academic skills, combined strategies of self-management, goal-setting, and problem-solving were found to be most effective and used most frequently used with this population. Self-management strategies alone were also used widely with this population to promote academic skills. It is important to note that in the one study in which choice-making was used as the intervention, the results on academic productivity and accuracy were powerful. However, while traditionally associated with self-determination interventions for students with MRDD, choice-making was used alone in only one of the academic studies reviewed.

The majority of interventions used single-subject designs, primarily multiple baseline designs. Generally, the studies demonstrated attention to quality research implementation; however, particular weaknesses, such as an absence of measures of procedural fidelity and limited information on participants were noted for many studies. These same weaknesses in methodological rigor were noted in the Test et al. (2005) review of self-advocacy intervention studies and the Wood, Fowler, Uphold, & Test (2006) review of self-determination interventions with students with severe disabilities. One difference in this review was the presence of more studies that demonstrated stronger control for internal validity through use of multi-tiered multiple baseline designs.

Limitations

The first limitation of this review is the paucity of research of self-determination interventions on academic performance of this population, which limits the ability to generalize the results. Only 11 studies were located in which results for individuals with MRDD could be clearly demonstrated.

A second limitation is the lack of specific terminology regarding the components of self-determination (e.g., self-monitoring, self-evaluation, self-recruitment, self-advocacy). Because authors did not always define their interventions using the components of self-determination that we applied to this research, we cannot be certain that we always made the correct decisions in categorizing the interventions. However, inter-rater agreement indicated that we were consistent.

Third, this review focused on self-determination intervention effects on academic skills, even though this may not have been the primary purpose of the original studies. Therefore, we were applying standards such as review of study design quality and strength of effects to measurement of variables that were not always the primary variable of concern in the original study. For example, the strength of effects could not be examined for all studies because the data on the academic variables were not graphed allowing for a calculation of PND (e.g., Rock, 2005).

Finally, the results of calculations of median PNDs across studies for categories of dependent variables and components of interventions should be interpreted cautiously. Calculations of PND for specific independent and dependent variables only represented one to three studies, due to the small number of studies reviewed. Because these studies were conducted with different participants and used various measures, the PNDs reported can provide the field with a sense of the strengths in these areas, but the PND calculations for individual studies are more definitive. However, the best way to build a case for external validity using single-subject research is through systematic replication (Horner et al., 2005).

Implications for Research

There is still a need for further research on how self-determination interventions affect academic skills for students with MRDD. Particularly given that students with MRDD are expected to access the same curriculum as students without disabilities and be assessed using the same measures as their peers without disabilities in many cases. However, since research also indicates that self-determination is an important skill for enhancing an individual’s quality of life (e.g., Benz et al., 1997; Wehmeyer & Palmer, 2003), it is important to focus on interventions that help student learn both academic and self-determination skills. While the current body of literature provides a good start, further research is needed that
investigates effects on both self-determination and academic skills.

In such research there is a need to examine academic variables such as accuracy and fluency in academic assignments. In the current review, organizational skills to complete assignments and asking for assistance to complete assignments were measured more frequently than productivity or accuracy on academic content assignments. This call for investigating more traditional academic variables is not intended to undermine the importance of organizational skills; it is simply an observation that such skills were less frequently the focus of research. Further, academic skills should be a primary, not simply a collateral measure in future research on self-determination interventions. The lack of graphing of the academic variable in some studies did not allow for a visual inspection of the results, critical to the interpretation of single-subject research.

Interventions with multiple self-determination components were most effective. However, these were also typically used to measure the dependent variables categorized as “other”. It is unclear if student performance on those “other” skills were stronger than accuracy and productivity, or if the measures of less traditional academic skills were simply more sensitive. Future research should investigate the effects presented in this review with replications of studies as well as investigations of similar interventions on different academic skills.

Further, some interventions included contingencies for performance of skills (e.g., Brooks et al., 2003). Researchers need to investigate the effects of self-determination interventions with extrinsic rewards, such as praise and tangible rewards, compared to self-determination interventions that do not reward students for completing the intervention, to better understand the strength of self-determination skill development alone. Further investigation is warranted by the generally positive results of this review to support infusion of self-determination and academic instruction as a worthy and effective method of instruction for all students.

Finally, future research must attend to rigor in design. Only one study in this review measured procedural fidelity. Replication of research is strengthened by these measures. Participant descriptions should include information on ethnicity and geography as another means of enhancing replication efforts, as well as responding to the need to examine self-determination interventions for diverse populations as suggested by Test et al. (2005) and Wood et al. (2006). Single-subject studies must use rigorous designs that insure internal validity so that conclusions may be drawn from the research. Horner et al. (2005) recommend demonstration of three replications of effect through use of three-tiered multiple baseline or reversal designs with multiple replications demonstrated. Similarly, group designs need to use and describe control group conditions to strengthen their designs. Within group studies should provide pre and post test measures and account for changes that may occur over time (Gersten et al., 2005). Finally, only about one-third of the studies in this review collected maintenance, generalization, or social validity data. Such measures are critical to demonstrating the strength and value of interventions and should be included in future research of self-determination interventions.

Implications for Practice

Educators can use this emerging body of research to support teaching self-determination skills as a means to also increase performance of academic skills for students with MRDD. Therefore, self-determination does not need to be considered an “add-on,” but can be used to support attainment of academic skills. In every study, students demonstrated an increase in the performance of the measured academic skill. For educators this means two things. First, teaching self-determination does not detract from academic skill attainment, and second, when teaching self-determination it is important to measure academic outcomes. Educators may use this review as a guide for the types of interventions that have been successful.

A word of caution for educators is that self-determination interventions alone can only increase behaviors that are already in students’ repertoires. Although they may support students in learning an academic skill, they have not yet been designed to teach academic
skills. In the studies included in this review, explicit instruction (direct instruction, modeling, explicit feedback on performance, opportunities to practice) of both the self-determination component skill and the academic skill demonstrated positive results (e.g., Agran et al., 2002; Hughes et al., 2002).

In conclusion, since self-determination is a complex construct, it will be important for researchers and practitioners to determine which components of self-determination are more closely linked to improved academic performance. The results of the current study provide initial evidence linking the self-determination components of self-management, goal setting, problem solving, decision making, choice making, self-advocacy to increased student performance on academic skills.

References
(References marked with an * were included in the review.)


Received: 27 January 2006
Initial Acceptance: 20 March 2006
Final Acceptance: 1 July 2006
Assessing Work Task Preferences among Persons with Intellectual Disabilities: An Integrative Review of Literature

Virginie Cobigo and Diane Morin
Université du Québec à Montréal

Yves Lachapelle
Université du Québec à Trois-Rivières

Abstract: Quality of life and self-determination are important values in the field of intellectual disabilities however they may be difficult to facilitate with persons who have limited communication skills. In fact, many studies provide evidence that these persons have less opportunity to make choices and express their preferences. To help practitioners understand what their clients with non verbal mode of communication prefer, direct observation methods were developed over the last few decades. However, these methods are used more often to determine best reinforcers for behavioral interventions, and not domains of the persons’ quality of life. Vocational activities are an important aspect of life planning, but assessing preferences for these complex stimuli requires specific methods. This integrative review of literature aims at gathering information about work task preference assessment, and consequently offering recommendations to practitioners on how to do it.

According to some authors, an adequate support should be person-centered because respecting and promoting self-determination of persons with intellectual disabilities is a critical factor of their quality of life (Lachapelle et al., 2005; Wehmeyer & Schwartz, 1998). Consequently, professionals should determine their clients’ preferences to plan important aspects of their lives such as vocational activities. Nevertheless, it might be difficult for persons with intellectual disabilities to express their preferences because of limited communication skills (Cameron & Murphy, 2002). Limitations in their cognitive skills and adaptive behaviors also impact the reliability of interviews with them. For example, they can have limited abstraction or insight skills (Luckasson et al., 2002). They have a tendency to acquiesce when presented with Yes or No questions, and tend to choose the last option in a multiple choice question (Fortin & Carrier, 1994; Guillemette & Boisvert, 2003). Furthermore, their answers are likely to be influenced by social desirability (Fortin & Carrier). On the other hand, a lack of concordance is highlighted between preferences identified by parents or caregivers and actual preferences expressed by persons with intellectual disabilities (Foxx et al., 1993; Reid et al. 1999; Stanchilfe, 1995). In fact, interviews with clients or with their relatives seem to be insufficient to determine their preferences. Thus, preferences assessment should rely on direct observation of a person selecting, consuming or using the presented items. Published reviews of literature described different assessment methods developed in the last decades (Cannella et al., 2005; Hagopian et al., 2004; Hughes et al., 1998; Kearney & Mc Knight, 1997; Lohrmann-O’Rourke & Browder, 1998). These reviews gathered information about the context and the format of the assessment, the behaviors considered as expression of a preference, and the stimuli presented. In summary, assessment methods are described as repeated sessions in which items are proposed to the person whether alone, paired or in groups (Hagopian et al.). In the single-stimulus presentation, each item is placed in front of the person and his or her behaviors are recorded. It is also possible to present the stimuli two by two or all together and observe which item is selected. In the group presentation, researchers recommend not to replace a selected item for the subsequent trials. This procedure is called multiple-stimuli without replacement presentation.
(MSWO). It is the recommended method because it takes less than half the time to administer compared to the paired presentation in which the evaluator must present all the possible pairs in each session. Paired and MSWO presentations are also convenient because they allow to determine a hierarchy of preferences and discriminate high from non-preferred items (Hagopian et al.; Lohrmann-O’Rourke & Browder). Previous research also described several behaviors that can be considered as an expression of a preference. Hughes et al. (1998) identified six kinds of behaviors: approach or physical selection of an item; task performance that allows obtaining the item for which the level of preference is evaluated; time engaged with the stimulus; signs, gestures, positive verbalizations or vocalizations; and activation of a microswitch. Some researchers also consider repair behaviors. It means that they do not record the selection of an item replaced on the table just after its selection. It is also possible to measure refusals of an item which consist in pushing the stimulus away or throwing it, exhibiting negative vocalizations, negative facial expressions or aberrant behaviors (Lohrmann-O’Rourke & Browder).

Thus, previous reviews of literature provide evidence that methods are available to assess preferences of people who cannot communicate verbally. However, most of the published studies assessed preferences to determine best reinforcers but not aspects of the persons’ life plans, including social or vocational activities (Hughes et al., 1998; Lohrmann-O’Rourke & Browder, 1998). Consequently, stimuli are more often tangible items that can be consumed or used as soon as presented. Assessment method for items with higher level of iconicity (such as activities) should address special issues because of the stimuli’s characteristics (Lohrmann-O’Rourke & Browder). For example, a symbol has to be used to represent a complex stimulus like an activity. It might be a representative object, a picture or a word. Furthermore, we cannot assume that the same behaviors than those observed with tangible items could be considered as expression of a preference. Nevertheless, none of the published reviews or studies tried to determine the specific characteristics of an assessment method depending on the type of stimuli presented, especially for more complex stimuli. The purpose of this review is to analyze assessment methods developed and used to determine one’s vocational preferences. The method used for this review is inspired by the guidelines for integrative review as proposed by Jackson (1989) who recommends trying to answer some questions when conducting such analysis. Thus, this review aims at answering the following questions: 1) how are stimuli presented (tangible, verbal, pictorial presentation) in the efficient methods to assess work tasks preferences of individuals with intellectual disabilities and limited communicative skills, 2) in these efficient methods, which behaviors could be considered as expression of preference or refusal, 3) in these efficient methods, which stimuli presentation are used: single-, paired- or multiple-stimuli presentation?

Method

Selection Criteria

Articles were selected according to the following six criteria: (a) articles reported empirical studies in which an intervention was applied to assess work tasks preferences, (b) studies’ purposes were to assess work tasks preferences or to demonstrate and evaluate a method to assess vocational preferences, (c) methods used to assess work tasks preferences had to be based on direct observation of the participants while they selected a task and were working on it (were excluded studies in which work tasks preferences were assessed only by questionnaires or interviews completed with the person or his or her caregivers or parents), (d) participants were adults with developmental disabilities, (e) participants were described as having limitations in their language or communicative skills, (f) selected studies had to be published in a journal with peer revision process.

Search Procedure

The search procedure included: (a) a search in ERIC and PsycInfo databases from 1970 to 2005 using the following 10 descriptors preferences, assessment, evaluation, work, job, vocational tasks, prevocational tasks, mental retardation, in-
intellectual disabilities and developmental disabilities, and (b) a review of secondary references (those cited in primary articles identified).

Nine studies met the six selection criteria (Lancioni et al., 1995; 1998a; 1998a; Mithaug & Hanawalt, 1978; Mithaug & Mar, 1980; Parsons et al., 1998; Reid et al., 1998; Stock et al., 2003; Worsdell et al., 2002).

Analysis Procedure

Articles were first analyzed to determine the participants’ characteristics. Then, we focused on purpose of the study, research design, and results presented. Subsequently, particular attention was provided to the preference assessment procedure. Relevant information was identified about the presented stimuli (what), the context of the assessment (who, when, where), the kind of behaviors considered as a means to express preference or refusal, and the format of the preference assessment (how).

Results

Participants Characteristics

All studies had between one and four participants except one that involved 14 persons. Their characteristics are reported in Table 1. In summary, all were adults with ages between 18 and 73, though two studies only described the participants as being adults. Most participants had severe to profound intellectual disabilities, but in three studies, no levels of intellectual disabilities were available. Most of them had secondary disabilities including sensory or physical limitations. Four studies gave information about their level of communication skills. Their participants did not speak at all or used only few-word sentences.

Studies’ Purposes

Four of the nine studies aimed at assessing participants’ preferences. Only one study assessed work tasks preferences (Mithaug &

<table>
<thead>
<tr>
<th>Reference</th>
<th>n</th>
<th>Age</th>
<th>Level of ID</th>
<th>Other disabilities</th>
<th>Communication skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mithaug &amp; Hanawalt, 1978</td>
<td>3</td>
<td>19 to 21</td>
<td>Severe</td>
<td>James and Mary: Down’s syndrome.</td>
<td>James: 1/2-word sentences Gary: 3/4-word sentences Mary: no language</td>
</tr>
<tr>
<td>Mithaug &amp; Mar, 1980</td>
<td>2</td>
<td>19 to 20</td>
<td>Severe</td>
<td>Mary: Down’s syndrome</td>
<td>Gary: 3/4-word sentences Mary: no language</td>
</tr>
<tr>
<td>Reid et al., 1998</td>
<td>3</td>
<td>30, 49, 73</td>
<td>Severe</td>
<td>Nonambulatory</td>
<td>Mr. West: 1/2-word sentences Mr. Freeman: short utterances, articulation problems Ms. Edwards: speech very difficult to understand</td>
</tr>
<tr>
<td>Parsons et al., 1998</td>
<td>1</td>
<td>38</td>
<td>Profound</td>
<td>Totally blind</td>
<td>No formal method of expressive communication other than approach and avoidance responses</td>
</tr>
<tr>
<td>Stock et al., 2003</td>
<td>14</td>
<td>18+</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Worsdell et al., 2002</td>
<td>4</td>
<td>18+</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Lancioni et al., 1995</td>
<td>4</td>
<td>20 to 34</td>
<td>NA</td>
<td>Deafness or blindness</td>
<td>NA</td>
</tr>
<tr>
<td>Lancioni et al., 1998a</td>
<td>4</td>
<td>24 to 31</td>
<td>Profound</td>
<td>Deafness</td>
<td>NA</td>
</tr>
<tr>
<td>Lancioni et al., 1998a</td>
<td>3</td>
<td>27, 38, 29</td>
<td>Profound</td>
<td>Deafness and severe visual impairment or blindness.</td>
<td>NA</td>
</tr>
</tbody>
</table>
Hanawalt, 1978), but three of them assessed preferences between work conditions: working alone versus working with a peer (Lancioni et al., 1995), task variation versus task repetition (Lancioni et al., 1998a), mobility versus sedentariness (Lancioni et al., 1998a). The other five studies evaluated the effectiveness of an assessment method or a part of a method (e.g. number of stimuli presented, type of behavior observed).

Research Design

Most of these studies used a single-case design (see Table 2). The authors compared the preferences assessed under different conditions using ABAB design or its variations or alternating treatment design. One article reported a case study (Parsons et al., 1998). The last study made a group comparison to check the method reliability and also used two questionnaires to collect data about the method effectiveness (Stock et al., 2003).

Assessment Methodology

The purpose of this review is to provide information about how to implement a work task preference assessment by analyzing the procedures used in previous studies. Table 3 presents characteristics of these assessment methods.

Stimuli presented. As noted earlier, one difficulty encountered during preference assessment is due to stimuli’s level of iconicity since, by definition, stimuli with high level of iconicity are not tangible items (Lohrmann-O’Rourke & Browder, 1998). In this review, studies included only stimuli with high level of iconicity: vocational activities or tasks arrangements. Consequently, this review provides evidence that practitioners could assess client’s preferences for various daily tasks and activities, and thus enhance his or her quality of life. The description of the procedures also provides guidelines for this kind of preference assessment.

To assess preferences for work tasks, most experimenters presented representative objects of tasks to participants following a method developed by Mithaug and Hanawalt (1978). Reid et al. (1998) recommended using, as representative objects, the same materials as those used on the job to facilitate the generalization of the assessed preferences to the job situation. Given the level of iconicity of the stimuli, use of object-signals requires a preliminary phase in which the evaluator makes sure the person understands the link between the object and the task it represents, and between choosing the object and subsequently working on the task it represents (Lancioni et al., 1995; 1998a; 1998a). However, if the person understands these links, the selection of an object-signal seems to be a valid indicator of his or her preferences. Its validity is demonstrated because participants tended to avoid a punishing item or frequently selected a reinforcing box when their representative objects were presented (Mithaug & Mar, 1980).

In the Stock et al. (2003) study, stimuli presented were different. They used still pictures depicting activities presented in video clips that would play when the corresponding picture was clicked. Results showed that the reliability of this method varies considerably among job categories depicted. Authors attributed these variations to the efficacy of the video clips to communicate tasks, but same issues than with object-signals could be addressed. We cannot assume that the person understands the link between the picture and the video clip and between selecting a picture and subsequently working on the task depicted on the video clip. Indeed, in contradiction with other studies, the selection of an item allowed the participant to view a video clip depicting the corresponding work task instead of allowing him or her to subsequently work on this task. However, a video clip seems to be an effective method in giving information regarding the work tasks, and thus allows the person to make informed choices when choosing one of the presented tasks. This idea is supported by a study in which the purpose was to evaluate the effects of vocational experience on vocational preference (Rudrud et al., 1986). First, the experimenters showed video clips to 13 participants and evaluated their preferences (using a vocational preference inventory). Then they trained the participants to complete these tasks and reassessed their preferences. The results demonstrated no change in preferences assessed before and after vocational experience. It suggests that
<table>
<thead>
<tr>
<th>Purpose and Design</th>
<th>Results and Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mithaug &amp; Hanawalt, 1978</strong></td>
<td>Mixed</td>
</tr>
<tr>
<td>Assess the prevocational task preferences.</td>
<td>9 of 12 preferences validated, 2 partially validated, and 1 reversed.</td>
</tr>
<tr>
<td>Baseline condition: Identified each subject’s most preferred, least preferred and moderately preferred tasks.</td>
<td>Validation of the preferences assessed in phase 1 for 2 participants (n = 3).</td>
</tr>
<tr>
<td>Reversal design over four conditions: 1) the preferred task with moderately-preferred task A and the non-preferred task with the moderately-preferred task B; 2) the preferred task with moderately-preferred task B and the non-preferred task with the moderately-preferred task A; 3 and 4 are replications of 1st and 2nd conditions.</td>
<td>Might reveal real changes in task preferences. We cannot assume that the subjects understood that picking up a choice object and placing it on the table meant that they would subsequently work the corresponding task.</td>
</tr>
<tr>
<td><strong>Mithaug &amp; Mar, 1980</strong></td>
<td>Positive</td>
</tr>
<tr>
<td>Investigate the functional relation between choice behaviors and the responses required to complete the task.</td>
<td>The reinforcing and punishing effects of working different tasks were demonstrated. The relation between selecting objects and working tasks suggests that the selection response is a valid indicator of the person’s preferences for different types of work.</td>
</tr>
<tr>
<td>ABABA design</td>
<td></td>
</tr>
<tr>
<td>Baseline condition: The participants worked on the tasks indicated by their object choice.</td>
<td></td>
</tr>
<tr>
<td>Task substitution conditions (alternated with baselines in a separate condition sequences): 1) the selection of a task object resulted in work on a relatively more preferred task; 2) or the selection resulted in work on a relatively less preferred task.</td>
<td></td>
</tr>
<tr>
<td><strong>Reid et al., 1998</strong></td>
<td>Positive</td>
</tr>
<tr>
<td>Evaluate a method of assessing work-task preferences by comparing the prework preference assessment to the preferences expressed on the job.</td>
<td>When provided with a choice of work tasks, each participant chose the task that the prework assessment had indicated as most preferred (75% of the choice opportunities). The procedure developed by Mithaug and Mar is a valid means to assess preferences.</td>
</tr>
<tr>
<td>Alternating treatment design</td>
<td></td>
</tr>
<tr>
<td>Phase 1: prework preference assessment.</td>
<td></td>
</tr>
<tr>
<td>Phase 2: 3 conditions: 1) assigning a participant to work on his or her least preferred task based on the prework preference assessment; 2) assigning a participant to work on his or her most preferred task; 3) allowing a participant to choose to work on either his or her least or most preferred task.</td>
<td></td>
</tr>
<tr>
<td><strong>Parsons et al., 1998</strong></td>
<td>Positive</td>
</tr>
<tr>
<td>Demonstrate a means of assessing work task preferences.</td>
<td>When Mr. Dunn made a choice 47% of all the presentations. He was engaged in work 100% of the observations.</td>
</tr>
<tr>
<td>Case study</td>
<td>Phase 2: He made a choice 75% of the choice opportunities. He chose the most preferred task on 75% of the choice opportunities. Prework preference assessment was a valid means of predicting preferences for work tasks. Only 1 participant.</td>
</tr>
<tr>
<td>Phase 1: Prework preference assessment</td>
<td></td>
</tr>
<tr>
<td>Phase 2: Choice between his most and least preferred tasks.</td>
<td></td>
</tr>
</tbody>
</table>
the critical factor to give to the persons the opportunity to make informed vocational choices is to provide enough information not vocational experience. It also supports that using video clips can be a good way to provide information. However, because of the com-

**TABLE 2—(Continued)**

<table>
<thead>
<tr>
<th>Purpose and Design</th>
<th>Results and Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stock et al., 2003</strong></td>
<td><strong>Mixed</strong></td>
</tr>
<tr>
<td>Evaluate the effectiveness of utilizing multimedia software (WorkSight, WS) presented in a self-directed format to enable persons with intellectual disabilities to independently express their job interests.</td>
<td>Significant correlations ($p&lt;0.05$) between the WS results and evaluator predictions (6/12 categories). WS more effective than 3 existing job assessment tools.</td>
</tr>
<tr>
<td><strong>Two questionnaires</strong> filled out by teachers and agency support staff: 1) The <em>pre-test form</em> examined the efficacy of the WS to predict good jobs match. 2) The <em>post-test form</em> compared current assessment tools and WS. <strong>Test-retest</strong> checked for reliability ($n=11$)</td>
<td>7 job categories showed correlations ($p&lt;0.10$) between the first test session and the retest. WS able to predict job preferences.</td>
</tr>
<tr>
<td><strong>Worsdell et al., 2002</strong></td>
<td><strong>Mixed</strong></td>
</tr>
<tr>
<td>Determine whether results obtained from single-stimulus assessment (SS) or multiple-stimulus assessment (MS) would be predictive of task engagement in a 60-min. session.</td>
<td>MS assessment: all 4 participants engaged in one task exclusively. SS assessment: High levels of engagement. These results were similar to those obtained during a 60-min criterion test.</td>
</tr>
<tr>
<td><strong>Alternating treatment design</strong></td>
<td>Exclusive engagement in one activity during the MS assessment reflected participants' preferences for a particular task. Results obtained from the SS assessment showed greater correspondence to participants' engagement in sustained activity under typical work conditions. An alternative: Multiple stimuli without replacement procedure.</td>
</tr>
<tr>
<td><strong>Condition 1</strong>: SS assessment</td>
<td><strong>Mixed</strong></td>
</tr>
<tr>
<td><strong>Condition 2</strong>: MS assessment</td>
<td><strong>Lancioni et al., 1995</strong></td>
</tr>
<tr>
<td><strong>Condition 3</strong>: Extended criterion test. 60-min test sessions.</td>
<td><strong>Mixed</strong></td>
</tr>
<tr>
<td><strong>2 subjects were purposeful in their selection and their preferences were stable across the phases and sets of object-signals.</strong> Subject 3: Increasing selection of the punishing item, but selection of the reinforcer box almost consistently. Differences in the 2 sets (varied of 20 or more percentage points). Subject 4: Decline in her peer selection, some variability across object-signals but continued to select the reinforcer box and to avoid the punishing item.</td>
<td>2 subjects were purposeful in their selection and their preferences were stable across the phases and sets of object-signals. Subjects’ preferences about work conditions could be assessed and those preferences were fairly consistent across sets of object-signals. The selections were quite purposeful.</td>
</tr>
<tr>
<td><strong>Phase 1</strong>: Establishing the first set of object-signals. <strong>Phase 2</strong>: Assessment with the first set of object-signals. <strong>Phase 3</strong>: Establishing the second set of object-signals. <strong>Phase 4</strong>: Assessment with both sets of object-signals.</td>
<td><strong>ABC design</strong></td>
</tr>
</tbody>
</table>

---

Work Task Preference Assessment  /  291
plex procedure developed by Stock et al. (2003) to assess preferences, it seems to be difficult to implement with persons with severe to profound intellectual disabilities. This article did not give enough information about the participants to know with whom it could be used, but the Rudrud et al. study (1986) involved participants with an IQ ranging from 51 to 71. For persons with more significant disabilities, practitioners should make sure the persons have experience with the tasks presented as experimenters did in the other studies of this review.

Moreover, studies in the field of preference assessment evaluated the relative efficacy of tangible, pictorial or verbal presentation methods and demonstrated that the choice of a presentation method should be determined by the person’s discriminative skills (Conyers et al., 2002; De Vries et al., 2005). If the method used is too complex for the person, the preference assessment is likely to be compromised. However, these studies never involved preference assessment for work tasks, but for food items and leisure activities. Nevertheless, De Vries et al. used representative objects of the leisure activities as it was done in most of the work task preference assessment. In the studies analyzed for this article, none of the experimenters assessed discriminative skills before choosing stimuli presentation method. Consequently, it could explain some differences observed between participants. However, in three studies, phases were implemented to establish the association between tasks and a set of representative objects (Lancioni et al., 1995; 1998a; 1998a). No information was provided about the level of discriminative skills before and after these trainings and no research exists concerning the influ-

<table>
<thead>
<tr>
<th>Purpose and Design</th>
<th>Results and Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lancioni et al., 1998a</td>
<td>Asses preferences between task variation and task repetition</td>
</tr>
<tr>
<td><strong>ABAB</strong></td>
<td><strong>Clear preferences expressed.</strong></td>
</tr>
<tr>
<td>Phase 1: Establishing object-signals.</td>
<td>Additional evidence of the possibility to enable people with severe and profound intellectual disabilities to express preferences.</td>
</tr>
<tr>
<td>Phase 2: Experimenter’s assignment.</td>
<td><strong>Participants did not differ in performance</strong> in the two task conditions during phases 2 and 4.</td>
</tr>
<tr>
<td>Phase 3: Participants’ choices.</td>
<td>Participant’s familiarity with the tasks and the presence of regular reinforcement ensured a relatively stable responding and work interest in both task conditions.</td>
</tr>
<tr>
<td>Phases 4 and 5: Replications of phases 2 and 3.</td>
<td><strong>Mixed</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Clear preferences expressed.</strong></td>
</tr>
</tbody>
</table>

Lancioni et al., 1998a
Assess preferences between a task arrangement involving mobility and a task arrangement involving sedentariness.

**ABAB**
Phase 1: Establishing object-signals.
Phase 2: Experimenter’s assignment.
Phase 3: Introduction to choice.
Phase 4: Participants’ choices.
Phases 5 and 6: Replications of phases 3 and 4.

**Mixed**
**Clear preferences expressed.**
Participants’ preferences about task arrangements could be successfully assessed.
Mean percentages of on-task behavior varied slightly for 2 participants, but were higher in the arrangement with mobility for participant 3 (about 75% vs. 55%).

**Mood expressions**
Participant 1: few positive scores; participant 2: about 26% in mobility arrangement and about 11% in sedentariness arrangement; participant 3: about 10% and 5% in the two task arrangements.

It may be useful to supplement the recording of the persons’ choice with checks on other aspects of the persons’ behavior such as on-task behavior, mood expression, and response time and performance accuracy.
ence of training on discriminative skills (De Vries et al.). Therefore, we cannot assume this method has an influence on the participant’s discriminative skills and should be recommended.

**Assessment context.** Only one identified study included comparison between information obtained from caregivers’ interview with those from direct assessment (Stock et al., 2003). Results of this study showed significant correlations ($p < 0.05$) between the two types of assessment for only six of the 12 job categories. This lack of concordance gives further evidence that practitioners should not rely only on information obtained by caregivers interview when assessing clients’ preferences, but should prioritize direct assessment.

Another consideration regarding the assessment context is who the evaluator is. In most of the identified studies, the experimenter also collected data, and one used future job coaches as evaluators (Reid et al., 1998). As noticed in a previous review (Lohrmann-O’Rourke & Browder, 1998), little information is provided about the relationship between the evaluator and the participant. However, it could influence the interpretation of the response and the participant’s motivation to respond because non symbolic communication is more complicated to interpret than symbolic communication as it implies the use of a limited behavioral repertoire including idiosyncratic behaviors (Campbell & Fletcher, 1993).

In the identified studies, the assessment was often embedded in the participant’s typical routine. This is contradictory with the results of previous reviews (Cannella et al., 2005; Lohrmann-O’Rourke & Browder, 1998). Experimental sessions were limited and their results were often compared with on-the-job preferences. It provides evidence that work tasks preference can be assessed daily in the person’s typical work environment.

**Responses considered as expression of a preference.** Physical selection of the representative object seems to be a valid indicator of participants’ preferences as demonstrated by the Mithaug and Hanawalt method replications. Physical selection was defined as picking up one object from those presented and setting it on the table (Mithaug & Hanawalt, 1978; Mithaug & Mar, 1980), pointing to or touching one of the object (Reid et al., 1998), or holding or manipulating the object (Parsons et al., 1998). Parsons et al. specified a limited duration of manipulation of at least 10 seconds. Moreover, very high interobserver agreements approaching 100% were recorded for behaviors described as physical selection. Mithaug and Hanawalt’s assessment procedure and its replications also included record of repair behaviors defined as replacing an object previously selected. In Stock et al. (2003), the choice was expressed by clicking on a picture depicting a task. It might be categorized as physical selection or activation of a microswitch.

After the participant selected an object, he or she was allowed to work on the corresponding task for a few minutes. While he or she was completing the task, some experimenters observed task engagement or on-task behaviors defined as manipulating materials in a manner to complete the work task (Lancioni et al., 1998a; 1998a; Parsons et al., 1998; Reid et al., 1998; Worsdell et al., 2002). Results of Worsdell et al. study gave evidence that on-task behaviors are good indicators of one’s preferences, but only in a choice condition. When the tasks were presented singly, all participants showed high level of task engagement in most tasks, even during a 60-minutes session. Same results were described in a study in which the purpose was to compare levels of on-task behaviors and disruptive behaviors between choice and assignment conditions (Parsons, Reid, Reynolds, & Bumgarner, 1990). It is also possible to record the length of the session corresponding to the time engaged in one task instead of the number of on-task behaviors (Lancioni et al., 1998a).

Lancioni et al. (1998a) also recorded the number of task-related responses carried out correctly (performance to the task) between two work task arrangements. It seems that the performance to a task did not provide information about one’s preferences because no difference was observed between the two work conditions although clear preferences were expressed.

In an other study, the same authors observed mood expressions defined as smiling, laughing, or emitting excited vocal sounds (Lancioni et al., 1998a). For two of the three participants, more mood expressions were ob-
<table>
<thead>
<tr>
<th>Assessment Methods</th>
<th>Context, stimuli, format</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mithaug &amp; Hanawalt, 1978</strong></td>
<td><strong>Paired presentation</strong> of two representative <strong>objects</strong> taken from each of the tasks. Objects randomly arranged on the left and the right side. The subject picked up one object. The experimenter removed the tray with remaining object, recorded the choice and supplied the materials to work the selected task for a 7-min. period (at the end a buzzer sounded). A 2 min-break before the next trial. All choice pairs were selected randomly from the available pairs to be presented. The task selected to work on one trial was paired at random with a different task on the next trial. This procedure was repeated for all possible pairs as long as the subject continued to select the same task or were presented. <strong>Evaluator</strong>: experimenter. <strong>Length of the assessment</strong>: 6 tasks, 15 pairs, 1 session per day, 2 hours per session, 34 sessions.</td>
<td>Choice responses: The subject picked up one object from the tray and set it on the table <strong>(physical selection)</strong>. A choice was not recorded when the subject picked up an object and replaced it on the tray <strong>(repair)</strong>. <strong>Interobserver agreement</strong>: 100%</td>
</tr>
<tr>
<td><strong>Mithaug &amp; Mar, 1980</strong></td>
<td>Replication of the procedure used by Mithaug and Hanawalt (1978). <strong>Evaluator</strong>: experimenter. <strong>Choice responses</strong>: See Mithaug and Hanawalt (1978). <strong>Interobserver agreement</strong>: 91%</td>
<td></td>
</tr>
<tr>
<td><strong>Reid et al., 1998</strong></td>
<td>Replication of the procedure used by Mithaug and Hanawalt (1978). <strong>4 modifications</strong>: 1) work materials used during the assessment were identical to the materials to be used on the job, 2) Assessment conducted by staff members who would later function as job coaches, 3) Each participant spent time working on each task during the preference assessment in the same manner as would be expected on the job, 4) 3 min.-period to work on the selected task. <strong>Evaluator</strong>: future job coaches. <strong>Length of assessment</strong>: 1) <strong>prework assessment</strong>: 5 tasks, 10 pairs, 30 min. per session, 3 to 4 sessions; 2) <strong>on the job</strong>: 2 tasks, 9 to 16 sessions, 8 to 11 weeks.</td>
<td>Choice responses: Pointing to or touching one of the task materials <strong>(physical selection)</strong>. <strong>Collateral responses</strong>: On-task behaviors defined as manipulating materials in a manner to complete the work task. <strong>Interobserver agreement</strong>: 100%</td>
</tr>
<tr>
<td><strong>Parsons et al., 1998</strong></td>
<td>Replication of the procedure developed by Mithaug and Hanawalt (1978) and modified by Reid et al. (1998). <strong>Evaluator</strong>: NA. <strong>Length of assessment</strong>: 1) <strong>prework assessment</strong>: 3 tasks, 3 pairs, 1 session per day, 9 minutes per session, 10 days; 2) <strong>on the job</strong>: 20 minutes per session.</td>
<td>Choice responses: Holding or manipulating a material at least 10 sec. <strong>(physical selection)</strong>. <strong>Collateral responses</strong>: On-task behaviors defined as manipulating the materials in a manner to complete the work task. <strong>Interobserver agreement</strong>: 100%</td>
</tr>
<tr>
<td>Context, stimuli, format</td>
<td>Response</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>Stock et al., 2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two equal sized windows displayed side by side (paired presentation). Each window contained a still picture depicting the video clip that would play when the picture was clicked. The user “selected the picture for job video he or she would like to watch first” (audio prompt). Once one of the pictures was selected, the other picture was hidden. After the first video clip finished playing, it was hidden from the screen and the remaining picture reappeared, the user “now selects the other picture to watch a different job video” (audio prompt). After the second video clip finished playing, both pictures reappeared, and two buttons with a “thumb up” icon on them appeared under each picture. The user then clicked the button directly under the picture depicting the job they liked the best between the two choices available (audio prompt). Then two new pictures depicting the next two video clips appeared and the procedure was repeated. Evaluator: computer. Length of assessment: 30 trials, 12 tasks, 5 distinct video clips for each task (10–15 sec. in length).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choice responses: Clicking on the corresponding image (activation of a microswitch).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collateral responses: Number of request for help.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Wordsett et al., 2002 |          |
| Signal stimulus assessment: Each task was presented for 5 min. During each 5-min trial, the participant was instructed to do whatever he or she wanted, no consequences were provided for task engagement. |
| Multiple-stimulus assessment: The seven tasks were presented in a concurrent arrangement during a 5-min assessment. Sets of task materials were available in an arc in front of participant. During each 5-min trial, the participant was instructed to do whatever he or she wanted, no consequences were provided for task engagement. Evaluator: NA |
| Length of assessment: 7 tasks, 1 to 3 sessions daily, 4 to 5 days per week. SS = 35 min per session, MS = 5 min per session. |
| Choice responses: On-task behaviors defined as manipulating materials in a manner required to complete the task, physical selection (only in MS). Interobserver agreement >85%. |

| Lancioni et al., 1995 |          |
| Object-signals were established as cues for the choice alternatives offered to the subject: working alone, a reinforcer box, a low chair (kneeling on the chair during the activity), and two of the peers. Two sets of object-signals were used. |
| Paired presentation of two object-signals. The choices including the reinforcer box and working alone (subject expected to select the reinforcer box) or the low chair and working alone (subject were expected to select working alone) were for checking whether subjects were purposeful in their preference responses. The execution of the tasks required about 2 min. Evaluator: Experimenter |
| Length of assessment: 6 or 8 choice trials per session, 39 to 62 sessions per phase. |
| Choice responses: selection of an object-signal (physical selection) Interobserver agreement: 83% to 100%, mean > 96% |
served while they were completing the preferred task. The third participant showed less than 2.5% of positive mood expressions in the two task conditions although very strong preferences were expressed by choosing one of the conditions almost 90% of its presentations. The authors provided no explanation for this discrepancy. Nevertheless, one plausible explanation is that the third participant might have verbal communication skills. Results of an analysis of nonverbal behaviors as expressions of preferences showed a significantly higher number of this kind of behaviors exhibited by participants who had limited verbal communication skills. Results of an analysis of nonverbal behaviors as expressions of preferences showed a significantly higher number of this kind of behaviors exhibited by participants who had limited verbal communication skills (Campbell & Fletcher, 1993). Nonverbal persons are likely to express preferences with mood expressions, as they have no other behavior in their repertoire to communicate better.

Unhappiness behaviors are not frequent enough to be considered in the assessment (Spevack, Martin, Hiebert, Yu, & Martin, 2004). No clear evidence was found about the influence of working on a non-preferred task versus a preferred task on the amount of disruptive behaviors exhibited (Parsons et al., 1990; Spevack et al.).

Although trends seem to appear in the selection of behaviors as expression of one’s preferences, results are not equivalent for all persons. Unfortunately, too little information was provided about participants to propose explanations. We can argue that participants who did not exhibit different behaviors while completing a preferred or a non-preferred task might be those who had not experienced enough opportunities to make choices (Martin, Martin, Spevack, Verbeke, & Yu, 2002). This idea supports the importance of a preliminary phase to introduce choice condition (Lancioni et al., 1998a), but research is needed to evaluate how much training is necessary for a choice effect to be established (Spevack et al., 2004).

**Assessment format.** Eight studies involved a paired-presentation procedure in which experimenters presented two representative objects or two pictures simultaneously where participants were expected to select one. This method has good psychometric properties but requires a relatively long period to administer notably because all the possible pairs have to be presented in each assessment session (Hagopian et al., 2004). Worsdell et al. (2002)
used two methods to assess participants’ preferences: the single-stimulus presentation and the multiple-stimuli with replacement presentation. The single-stimulus presentation seems to be inappropriate to assess work task preferences. As task engagement is high for all tasks even in a 60-minutes work period, it is difficult to determine the participant’s preferences. The multiple-stimuli presentation appears to be more accurate. Nevertheless, this method informs only on high- and non-preferred tasks but not on the level of preference for each of the presented tasks. Like the actual studies in the field of preference assessment, Worsdell et al. recommended the use of a multiple-stimuli without replacement procedure. This method is helpful to determine a hierarchy of preferences that provides several options of vocational activities respecting the client’s likes. Moreover, switching among several preferred tasks is more likely to maintain high level of work engagement over time (Worsdell et al.) and to enhance the person’s quality of life.

A last but important aspect to consider in the assessment procedure is the duration of access to the stimuli given to the participant. In the analyzed studies, the duration was determined by the experimenters and varied between three to 20 minutes. As the task engagement seems to be equivalent in a 5-minute or in a 60-minute session (Worsdell et al., 2002), an access period of more than five minutes does not seem to be necessary. A short period of assessment is also recommended because the observation is easier (Lohrmann-O’Rourke & Browder, 1998). Moreover, the assessment length is related to the access duration. So the shorter the access is, the shorter the assessment is, and consequently the easier it is to implement. Considering that a preference assessment has to be based on several sessions, the length of sessions is strongly related to the willingness of the practitioners to use direct measures of their clients’ preferences.

Discussion

**Guidelines to Assess Work Task Preferences**

This review provides evidence that methods exist to assess work task preferences for persons with intellectual disabilities and limited communication skills. These methods might also be useful to assess work conditions preferences. The method developed by Mithaug and Hanawalt (1978) and modified in its replications seems to be a valid assessment procedure. Consequently, to assess vocational preferences, practitioners should present pairs of representative objects and record which of them the person picks up. They should consider repair behaviors, so a task is not recorded as selected when the person replaces on the table the object he has picked up 10 seconds or less before. The selection of a task will be reinforced by the subsequent opportunity to work on the chosen task. Then, the task selected on one trial is presented again with a different remaining task in each trial. This procedure continues as long as the same task is chosen again or until all possible pairs including that task are presented. A session ends when all possible pairs were presented. After each trial, the person is allowed to work on the chosen task during a period previously determined by the evaluator. Research seems to demonstrate that a 5-minute period is enough. While the person is completing a task, evaluator should record on-task behaviors defined as how many behaviors are exhibited to complete the task or the amount of time engaged in completing that task. Mood expressions could be considered too but with caution. It seems that this kind of behaviors provides relevant information about one’s preferences only with persons using non-verbal communication. Moreover, positive mood expressions seem to be exhibited only when a very high-preferred item is presented (Green, Gardner, & Reid, 1997; Green & Reid, 1996). Performance while completing the task, unhappiness and disruptive behaviors exhibited seems not to be valid indicators of one’s preferences.

Before starting the assessment, some precautions must be taken. First, it is recommended to choose an evaluator who knows the person and frequently interacts with him or her because of the difficulty to interpret non-symbolic communication and to recognize idiosyncratic behaviors. Second, an assessment of the person’s discriminative skills is useful to make sure he or she will be able to associate an object to the task it represents. Some authors recommended training the per-
son to understand the link between an object and a task and between choosing an object and subsequently working on the corresponding task. We cannot pretend this training has an impact on the person's discriminative skills, but it is likely to have an influence on the behaviors expressed during the preference assessment. It is particularly noteworthy with people with severe to profound intellectual disabilities who have a history of activity assignment. Consequently, a preliminary phase could be necessary to show the person that he or she is allowed to make choices. Furthermore, the preliminary phase and the relationship between the evaluator and the person with intellectual disabilities might address another issue since studies showed evidence that parents' and professionals' are likely to overestimate receptive communication skills of the persons with intellectual disabilities (Baker, Freeman, & High, 2000; Banat, Summers, & Pring, 2002; Bradshaw, 2001; McConkey, Morris, & Purcell, 1999).

MSWO seems to be interesting to assess work task preferences because it is a quicker procedure (Hagopian et al., 2004) but none of the published studies evaluated the implementation of this procedure in a vocational setting. This is particularly noteworthy because the duration of the assessment is related to the willingness of the practitioners to use it. Without quicker means to assess work task preferences, professionals are likely to rely only on the parents' or caregivers' attributions that are not always a reliable assessment method.

**Guidelines for Future Research**

Future research must address special issues in the validation of preferences assessment. First, parents' or caregivers' attributions often are not a reliable way to determine preferences. Therefore, further research should not evaluate method effectiveness by comparing its results to those obtained with questionnaires or interviews with relatives. Furthermore, research should determine the influence of a history of tasks assignment on the preference expression. If an impact exists, alternating treatment designs are likely to be inadequate to evaluate method effectiveness as the assignment phase could influence the preference expression on the following choice phase.

Future research is needed to further determine characteristics of work task preference assessment. First, nine studies are not enough to allow for any generalizations and clearly recommend one procedure. Then, practitioners need more information about how to implement such methods. For example, studies should try to determine how many sessions are necessary to accurately assess work task preferences and how long these sessions should last. Information is also lacking about the characteristics of the evaluator and the person who is evaluated. Their relationship can have an impact on the assessment because of the non-verbal mode of communication. Furthermore, researchers should determine what should be the characteristics of the representative objects or pictures to accurately depict vocational tasks. This element is particularly important to give to the person with intellectual disabilities the opportunity to make informed choices (Rudrud et al., 1986).

Future research is also needed to determine the easiest and quickest means to assess work task preferences. The easier the assessment is, the more frequently it will be used. Consequently, practitioners could offer to the persons with intellectual disabilities and limited communication skills more opportunities to express their preferences and, thus influence their environment. This is also important because preferences change over time, so their assessment should be done regularly (Mason, McGee, Farmer-Dougan, & Risely, 1989). One way to answer these questions is by interviewing caregivers after they have used preference assessment methods to collect their perceptions about them. Results should allow researchers to make recommendations to better implement these methods. It might be useful to know how they adapt assessment methods to deal with natural work conditions, like the presence of other persons who need supervision. The implementation of these assessment methods is an important purpose of research because results will offer practitioners new tools to enhance participation and quality of life of persons with intellectual disabilities who have limited communication skills.
References


ton, DC: American Association on Mental Retardation.


Received: 27 March 2006
Initial Acceptance: 12 May 2006
Final Acceptance: 20 July 2006
Review of Recent Treatment Acceptability Research

Stacy L. Carter
Texas Tech University

Abstract: With recent increases in the use of positive approaches to treatment for individuals with developmental disabilities, it seems appropriate to review the variables that have been found to influence the acceptability of various treatments. Programmatic treatments for problematic behaviors that incorporate primarily positive (reinforcement) components rather than negative (punishment) components may still be susceptible to variables found to influence the acceptability of treatments. Although more positive reinforcement based approaches are certainly preferred, the need to consider the right to effective treatment is also an important component of any intervention for problematic behavior. To continually assure the right to effective treatment, the examination of variables affecting the acceptance of treatments continues to be an important area of research. This paper reviews the instruments that have been used to evaluate the acceptability of treatments as well as the variables that have shown demonstrated influence on the acceptability of treatments for problematic behavior.

Research in the area of treatment acceptability is relatively new to the area of education and psychology with a noticeable increase in treatment acceptability studies occurring during the 1980’s (Elliott, 1988). Recognition of the need for research in the area of social validity became more evident with an elaboration of the topic in an article by Baer, Wolf, and Risley (1968). Prior to this publication, the primary emphasis with regard to educational and behavioral treatments had been treatment efficacy (Witt, 1986). Wolf (1978) stated that social validity refers to decisions based on the social importance of interventions on three related levels. First, objectives of the treatment must be socially significant in which the desired outcome is both important and pertinent. Second, treatment procedures must be deemed as socially appropriate in that they are generally considered acceptable by society. Third, effects of the treatment should have clinical significance. Kazdin (1980) noted that social validity research has mainly focused on the appropriateness of treatment procedures, an area of research also known as treatment acceptability. The term treatment acceptability, a component of social validity, was defined by Kazdin as judgments of treatments by actual or potential consumers of the treatments, such as nonprofessionals, clients, laypersons, and others. For the purposes of this review, treatment acceptability will be defined according to the definition proposed by Kazdin and the terms treatment and intervention will be used interchangeably.

Treatment Acceptability Measurement Instruments

Two of the most frequently used instruments for measuring treatment acceptability are the Treatment Evaluation Inventory (TEI; Kazdin, 1980) and the Intervention Rating Profile (IRP; Witt & Elliott, 1985). Both of these instruments have been modified and revised in multiple ways by researchers in attempts to address target raters more effectively or to reduce the administration time of the instruments. Because modifications to these instruments have been implemented in various acceptability studies, a review of these two instruments and their subsequent modifications is provided (refer to Table 1).

Treatment Evaluation Inventory. The TEI was originally designed to assess treatments recommended for children with behavior disorders (Kazdin, 1980). It consists of 15 items that ask direct questions regarding specific treatment procedures. Ratings are made on a seven point Likert-type scale, with descriptive anchor points that vary depending on the
<table>
<thead>
<tr>
<th>Instrument Name</th>
<th>Type of Instrument</th>
<th>Validity</th>
<th>Reliability</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Evaluation Inventory (TEI; Kazdin, 1980)</td>
<td>Consists of 15 items, uses a seven point Likert-type scale, descriptive anchor points vary depending on question</td>
<td>Item loadings from .61 to .95 on a unitary factor accounting for 51.4% of variance</td>
<td>Internal consistency ranged from .35 to .96</td>
<td>Considered the first and one of the most frequently used instruments</td>
</tr>
<tr>
<td>Treatment Acceptability Rating Form (TARF; Reimers &amp; Wacker, 1988)</td>
<td>Consists of 15 items, uses a seven point Likert-type scale, varying descriptive anchor points</td>
<td>Internal consistency ranged from .80 to .91</td>
<td>Developed from the TEI for use with parents, incorporated factors of effectiveness, and cost of treatment</td>
<td></td>
</tr>
<tr>
<td>Treatment Evaluation Inventory Short Form (TEI-SF; Kelly et al., 1989)</td>
<td>Consists of nine items rated using a five point Likert-type scale with fixed anchor points</td>
<td>Internal consistency of .85</td>
<td>Revised version of the TEI, requires less time to complete than the TEI</td>
<td></td>
</tr>
<tr>
<td>Treatment Acceptability Rating Form-Revised (TARF-R; Reimers et al., 1991)</td>
<td>Consists of 20 items rated using a seven point Likert-type scale, varying descriptive anchor points for each item</td>
<td>Internal consistency of .92</td>
<td>Revised version of the TARF, includes items addressing problem severity and understanding of treatment</td>
<td></td>
</tr>
<tr>
<td>Intervention Rating Profile (IRP; Witt &amp; Elliott, 1985)</td>
<td>Consists of 20 items evaluated by a six point Likert-type scale</td>
<td>One primary factor (41% of the variance) and four secondary factors</td>
<td>Internal consistency of .89</td>
<td>Designed to evaluate educational treatments</td>
</tr>
<tr>
<td>Intervention Rating Profile-15 (IRP-15; Martens et al., 1985)</td>
<td>Consists of 15 items, uses a six point Likert-type scale</td>
<td>Principal component analysis resulted in a unitary factor</td>
<td>Internal consistency of .98</td>
<td>Modified the IRP to increase internal consistency</td>
</tr>
<tr>
<td>Children’s Intervention Rating Profile (CIRP; Witt &amp; Elliott, 1985)</td>
<td>Consists of seven questions evaluated using a seven point Likert-type scale</td>
<td>Principal component analysis resulted in a unitary factor</td>
<td>Internal consistency ranged from .75 to .89</td>
<td>Modified the IRP for children, items written at fifth grade reading level</td>
</tr>
<tr>
<td>Behavior Intervention Rating Scale (VonBrock &amp; Elliott, 1987)</td>
<td>Consists of 24 items, uses a six point Likert-type scale</td>
<td>Internal consistency of .97</td>
<td>Modified version of the IRP-15</td>
<td></td>
</tr>
<tr>
<td>Abbreviated Acceptability Rating Profile (AARP; Tarnowski &amp; Simonian, 1992)</td>
<td>Consists of eight items rated using a six point Likert-type scale</td>
<td>Unitary factor accounted for 84.9% of the variance with item loadings ranging from .89 to .96</td>
<td>Internal consistency of .98</td>
<td>Modified the IRP-15 to reduce completion time and improve readability</td>
</tr>
</tbody>
</table>
question. Total scores are obtained by summing all items with higher summed total scores indicating greater levels of treatment acceptability. A principal component analysis of the TEI rotated to varimax criterion resulted in item loadings from .61 to .95 on a unitary factor accounting for 51.4% of the variance (Tarnowski & Simonian, 1992). The TEI is the most frequently used measure of treatment acceptability, but it has been used very little in actual clinical intervention studies to measure acceptability of behavioral interventions with children and families. Modifications to the TEI have been developed in order to decrease administration time and to make it more comprehensive.

Treatment Evaluation Inventory-Short Form. The TEI was abridged and published as the Treatment Evaluation Inventory-Short Form (TEI-SF; Kelley, Heffer, Gresham, & Elliott, 1989). This revised form used to measure acceptability of treatments for children with behavior disorders was reduced to nine items to make it more expedient to complete. Items are rated using a five point Likert-type scale with fixed, anchored points. Descriptions on the Likert-type scale range from 1 (strongly disagree) to 5 (strongly agree). Total scores are obtained by summing all items with higher summed scores indicating greater levels of acceptability. The internal consistency of this instrument was reported to be .85 (Kelley et al.).

Treatment Acceptability Rating Form. The Treatment Acceptability Rating Form (TARF; Reimers & Wacker, 1988) was devised from Kazdin’s (1980) original measure of treatment acceptability to measure parent’s acceptability of treatments devised within a clinical setting. It consists of 15 items on a seven point Likert-type scale. The Likert-type scale ranges from 1 (strongly disagree) to 6 (strongly agree). Total scores are obtained by summing all items with higher summed scores indicating greater levels of acceptability. Principal component analysis with varimax rotation revealed a primary factor accounting for 41% of the variance and four secondary factors (Tarnowski & Simonian, 1992). The internal consistency of this instrument was reported to be .92.

Intervention Rating Profile. The IRP was developed to extend research in treatment acceptability to educational treatments and make practitioners more aware of interventions considered acceptable by teachers. It consists of 20 items that are rated on a six-point Likert type scale. The Likert-type rating scale ranges from 1 (strongly disagree) to 6 (strongly agree). Total scores are obtained by summing all items with higher summed scores indicating greater levels of acceptability. Principal component analysis revealed a primary factor accounting for 41% of the variance and four secondary factors (Tarnowski & Simonian, 1992). The internal consistency of this instrument was reported to be .98.

Intervention Rating Profile – 15. Modifications were made to the IRP to shorten the instrument and to increase item loading on a single factor. This modified version of the IRP was published as the Intervention Rating Profile – 15 and used to measure educational interventions (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985). Modifications to improve the internal consistency of the instrument included generating eight new items and removing other items resulting in a total of 15 items. Items are rated using a six point Likert-type rating scale with ranges from 1 (strongly disagree) to 6 (strongly agree). Total scores are obtained by summing all items with higher summed scores indicating greater levels of acceptability. The internal consistency of this instrument was reported to be .98.

Children’s Intervention Rating Profile. The Children’s Intervention Rating Profile (CIRP; Witt & Elliott, 1985) was a modification of the IRP designed to assess treatment acceptability of educational interventions by children. The instrument consists of seven questions related to the perceived fairness and expected effective-
ness of a treatment. Questions on this instrument were written at a fifth-grade reading level. The items are rated using a seven point Likert-type rating scale. Total scores are obtained by summing all items with higher summed scores representing greater levels of acceptability. The internal consistency of this instrument ranged from .75 to .89.

Behavior Intervention Rating Scale. The Behavior Intervention Rating Scale (BIRS; Brock & Elliott, 1987) was developed to create a new instrument for measuring treatment acceptability in school settings. The BIRS is a modification of the IRP-15, which added nine items for a total of 24 questions included on the instrument. Items are rated using a six point Likert-type rating scale that ranges from 1 (strongly disagree) to 6 (strongly agree). Total scores are obtained by summing all items with higher summed scores indicating greater levels of acceptability. The internal consistency of this instrument was reported to be .97.

Abbreviated Acceptability Rating Profile. An additional abbreviated form of the IRP-15 was developed to measure educational interventions and published as the Abbreviated Acceptability Rating Profile (AARP; Tarnowski & Simonian, 1992). They developed the AARP through modification of the IRP-15 by eliminating seven items and maintaining eight items that were reworded to improve readability. Items are rated using a six point Likert-type rating scale that ranges from 1 (strongly disagree) to 6 (strongly agree). Total scores are obtained by summing all items with higher summed scores indicating greater levels of acceptability. The internal consistency of this instrument was reported to be .98.

Psychometric properties of the AARP were initially evaluated using a sample of 60 parents seen for routine pediatric outpatient visits at a large urban hospital. The initial sample when subjected to a principal component analysis resulted in all items loading on a unitary factor (Acceptability) that accounted for 84.9% of the variance with item loadings ranging from .89 to .96. The initial sample was cross-validated in a second independent sample of 80 mothers by subjecting the data to identical analysis and resulting in item loadings ranging from .89 to .98 on a single factor that accounted for 90.3% of the variance. The AARP resulted in reduced completion time to approximately one-half the ten minutes required to complete the IRP-15. Readability analyses conducted using the Harris-Jacobson Wide Range Readability Formula (Harris & Jacobson, 1982) resulted in indices of 5.0 for the AARP compared to 7.9 for the IRP-15.

Comparison of Treatment Acceptability Measures

Finn and Sladeczek (2001) critiqued nine treatment acceptability measures which were the: (a) TEI, (b) TEI-SF, (c) TARF, (d) TARF-R, (e) IRP, (f) IRP for teachers, (g) AARP, (h) CIRP, and (i) BIRS. These nine measures were selected after an extensive database search of studies investigating behavioral intervention acceptability. The evaluation covered eight areas of treatment acceptability as follows: (a) definition of treatment acceptability; (b) content and purpose; (c) test reliability; (d) test validity; (e) statistical analysis; (f) sample characteristics; (g) scoring procedures; and (h) uses of the measure in research and practice. The evaluation indicated no single measure of treatment acceptability to be more comprehensive than any other. Conversely, in comparison to the IRP-15, Tarnowski and Simonian (1992) indicated that the AARP was more simplistic, easier to read, and required less administration time. In addition, Kelley et al. (1989) considered the AARP to be an improvement over the IRP-15 as well as other measures of acceptability.

Method

Although treatment acceptability research methodology has primarily consisted of analogue research, several variations of this methodology have been incorporated to extend the ecological validity of the findings. Variations have included presentation of case descriptions of problem behavior and treatments by use of audio tape (Kazdin, 1980), written summaries (Cavell, Frentz, & Kelley, 1986), video presentations (Martens et al., 1985), and combinations of presentation formats including actual implementation of treatments (Reimers, Wacker, Cooper, & De Raad, 1992). Participants may rate the treatment acceptability of one treatment applied to several different case descriptions or rate treatment ac-
ceptability of several treatments applied to the same or several different case descriptions. While analogue research methodology has provided insight into the acceptability of several treatments when hypothetically applied to numerous different cases covering a wide range of variables, the findings may be limited in ecological validity when compared to a clinical research methodology (Miltenberger, 1990). Clinical research methodology in treatment acceptability has typically involved having clients and/or consultees rate treatment acceptability prior to, during, and/or after implementation of a treatment. While clinical research methodology to investigate treatment acceptability appears to provide more ecologically valid findings and naturalistic evaluation of treatment acceptability in relationship to other variables, the analogue method of investigating treatment acceptability provides a much larger and more representative data base more quickly especially with treatments for low incidence populations.

Previous Literature Reviews of Treatment Acceptability

Reimers, Wacker, and Koepple (1987) conducted a review of the literature on behavioral interventions and found five primary factors that were considered to affect treatment acceptability. These factors were problem severity, treatment approach, time needed to implement treatment, side effects of treatment, and cost. Treatments for more severe problem behaviors were generally considered more acceptable than treatments for less severe problem behaviors, with some evidence for increased acceptability when more restrictive treatments were matched with more severe problems and less restrictive treatments were matched to less severe problems. Reinforcement-based procedures were generally rated more acceptable than punishment-based procedures. Treatments that required less implementation time were generally rated as more acceptable in comparison to treatments requiring more implementation time. One study reported that treatments with adverse side effects were rated less acceptable than those reporting no side effects. Cost was a factor suggested by Reimers et al. to influence treatment acceptability, although no studies were reported to have examined this factor as a variable.

Miltenberger (1990) conducted a review of the literature on treatment acceptability research conducted during the 1980’s and suggested that the most acceptable treatments would be those which were least restrictive, required little time, have the fewest side effects, are least disruptive to other students, are consistent with the rater’s training or orientation, are presented with the most appropriate rationales, are considered to be necessary for behavioral improvement, and are considered to be most effective.

Elliott (1988) reviewed 20 empirical studies on treatment acceptability of behavioral interventions for school children as rated by teachers, children, and psychologists. Elliott concluded that treatment acceptability was a complex construct influenced by several variables. Additionally, Elliott indicated that educational consumers rated positive treatments as more acceptable than reductive treatments and that there was a moderate-to-strong relationship between pretreatment acceptability ratings and perceived treatment effectiveness.

The present literature review does not attempt to provide a comprehensive review of literature on treatment acceptability conducted prior to 1990 (See Elliott, 1988; Miltenberger, 1990; and Reimers et al., 1987 for detailed reviews of treatment acceptability research published prior to 1990). Representative studies conducted prior to 1990 are reviewed in brief as examples of previously established themes found within the treatment acceptability literature. These representative studies are provided as a comparison to more recent research on treatment acceptability.

Review of Recent Treatment Acceptability Research

This review focuses on recent treatment acceptability research, which was defined as studies published during the 1990’s through 2005, while comparing these recent studies with some representative studies conducted prior to 1990. This segment of the literature was chosen in order to update the most recent review of treatment acceptability, which was considered to be Miltenberger’s (1990) review. Studies were identified by conducting a search in December 2005 for the terms “treat-
ment acceptability” using the search engine EBSCO Host. Additional articles were identified through the reference section of articles reviewed. An emphasis was placed on discussing treatment acceptability research that directly involved studies focusing on treatments for problem behaviors. Studies primarily focusing on academic skill enhancement, counseling techniques, and consultation methods were excluded from this review.

The recent review involved examining the research on treatment acceptability with regard to three types of variable manipulation. Although these three types of research often overlap with multiple variable manipulations, for the purposes of this review, the specific variables primarily focused on in the study were examined along these three separate lines of research (i.e., manipulations involving treatment variables, client/case description variables, and rater variables/individuals from which treatment acceptability ratings are obtained). Further clarification of these lines of research is provided below.

**Treatment Variable Manipulation**

Research involving acceptability of interventions with the manipulation of treatment variables appears to have received the greatest attention within the literature on treatment acceptability (Kazdin, 1980; Spreat & Walsh, 1994). Prior to 1990, treatment variable manipulation primarily consisted of manipulations of different types of treatment, varying effectiveness of treatment, and different mediators of treatment. Manipulations of different types of treatment have included factors such as comparing the acceptability of treatments that are highly intrusive versus less intrusive (Kazdin), based on reinforcement techniques versus more punishment based techniques (Kazdin; Witt, Elliott, & Martens, 1984), apparent appropriateness (Cavell et al., 1986), and time required to implement treatment (Witt et al.). Research involving manipulations of the effectiveness of treatments has evaluated the acceptability of treatments by: (a) including or excluding information on treatment efficacy (Brock & Elliott, 1987); (b) providing information on the efficacy of the treatment during previous treatment (Kalfus & Burk, 1989); and (c) providing information on the efficacy of the treatment in comparison to other treatments (Clark & Elliott, 1988). Manipulations involving the mediator of treatment have compared acceptability of treatments proposed to be implemented by parents and teachers versus school psychologists (Kalfus & Burk). Manipulations involving the professional recommending a treatment have involved comparisons of treatments recommended by a teacher, school psychologist, and a pediatrician (Carter, 2005). The following sections will discuss studies that represent the different types of treatment variable manipulation conducted prior to 1990 and provide a review of more recent literature (1990–2005), which extends or complements the research conducted prior to 1990.

**Intrusiveness of treatments.** Kazdin (1980) conducted a treatment variable manipulation experiment to assess acceptability of alternative treatments for inappropriate behaviors of children. The experiment involved 68 female and 20 male undergraduates who rated acceptability of four different treatments (reinforcement, time-out from reinforcement, drug treatment, and electric shock) each of which represented increasing levels of intrusiveness, respectively. Case vignettes were presented via a cassette tape player operated by the experimenter in a separate observation room from the participant. Treatment acceptability data was obtained by using the TEI. Results indicated statistically significant differences among all treatments with the following hierarchy of acceptability respectively, reinforcement, time-out from reinforcement, drug treatment, and electric shock. This study demonstrated that highly intrusive treatments were considered less acceptable than less intrusive treatments. Spreat and Walsh (1994) extended the examination of the intrusiveness of treatment on treatment acceptability. They surveyed members of American Association of Mental Retardation (AAMR). Treatment intrusiveness variables described within the case vignette included previously used procedures, likely side effects, and different treatments. Intrusiveness factors found to influence acceptability were restrictiveness of the proposed treatments and whether other procedures had been previously tried.

Jones, Eyberg, Adams, and Boggs (1998) continued the examination of intrusiveness of
treatments. They assessed the treatment acceptability of six child management techniques (positive reinforcement, response cost, differential attention, time-out, overcorrection, and spanking) using the TEI-SF with 20 mothers of children referred for treatment of disruptive behavior. Mothers rated the positive reinforcement technique as more acceptable than other techniques while spanking was rated as less acceptable than the other techniques.

**Reinforcement versus punishment.** The previously described study by Kazdin (1980) also revealed a relationship between treatment acceptability with reinforcement based techniques and punishment based techniques. The reinforcement procedure as described in the study represented a reinforcement based treatment and all other treatments represented punishment based techniques. The reinforcement based treatment was found to be the most acceptable of all the treatments used in the study. Burgio et al. (1995) continued the research on reinforcement techniques compared to punishment techniques. They evaluated the acceptability of two behavioral treatments (differential reinforcement and time-out) and a drug therapy regime among elderly individuals using a modified version of the TEI. Ratings demonstrated significant differences between all treatments for behavioral disturbances of geriatric individuals with differential reinforcement receiving the highest ratings followed by the time-out, and drug therapy. In addition, significantly higher acceptability ratings were given to time-out procedures described as taking place in a nursing home versus a community setting. Several other studies have examined procedures consisting of reinforcement compared to punishment techniques although this comparison was not the primary focus of the study. These studies are described under the category that appeared to be the primary focus of the research.

**Apparent appropriateness of treatments.** Cavell et al. (1986) used the TEI with 120 middle and high school teachers to assess the acceptability of four paradoxical treatments with different rationales and an ineffective contingency contract for a hypothetical case of school disruption and truancy. They found that continuing the ineffective contingency contract was rated as more acceptable than all paradoxical treatments regardless of the rationale provided. Betts and Remer (1993) evaluated the acceptability of paradoxical versus nonparadoxical interventions. Paradoxical interventions were defined as techniques designed to eliminate undesirable behavior by encouraging the undesirable behavior. Participants in this study were 97 undergraduates taking part in a semester long simulation of a family arguing with a rebellious adolescent daughter. Participants were grouped into “families” and participated in four role-play exercises and a family therapy session prior to evaluating either a paradoxical or nonparadoxical intervention using the TEI-SF. Findings indicated that paradoxical interventions were less acceptable than nonparadoxical interventions, although paradoxical interventions were considered to be an acceptable intervention overall.

**Time required to implement treatment.** Witt et al. (1984) examined the influence that varying amounts of time associated with implementing a treatment had on the acceptability of the treatment. They found that the less time required to implement a treatment, the more acceptable the treatment would be rated. 

**Previous treatment effectiveness.** Kalfus and Burk (1989) examined the influence of previous treatment effectiveness on treatment acceptability. They administered the TEI to 105 graduate level education students and 53 undergraduate level psychology students to evaluate the acceptability of five treatment procedures for a case vignette of a child who engaged in pica. Five treatment procedures were rated in the following order of most acceptable to least acceptable: positive reinforcement, overcorrection, contingent removal of the pica item, differential attention, and time-out. Treatment history information was manipulated by providing either a long case description or a shorter case description. The absence of treatment history information increased acceptability ratings by undergraduate psychology students and decreased acceptability ratings by graduate education students. Findings from this research supported previous research by identifying positive reinforcement as the most acceptable treatment, time-out as the least acceptable treatment,
and other treatments receiving ratings falling between positive reinforcement and time-out.

Comparison of alternative effectiveness of treatments. Clark and Elliott (1988) examined the influence of manipulating the level of treatment effectiveness on treatment acceptability. They distributed a case vignette, a treatment acceptability measure, and a general knowledge of techniques questionnaire to 133 elementary school teachers from Nebraska and Louisiana. The case vignettes manipulated two potential treatments: modeling-coaching (a form of) versus overcorrection method (a form of) and two levels of outcome effectiveness (weak vs. strong therapeutic effects). Treatment acceptability was rated using the BIRS. Findings indicated a statistically significant preference for the modeling-coaching treatment when compared to the overcorrection treatment and statistically significant positive influences on acceptability by strong therapeutic outcome effects.

Mediator of treatment. Kalfus and Burk (1989), in the study previously described, evaluated the effects of different treatment mediators on treatment acceptability. In their study, treatment mediator identity was manipulated by presenting the treatment mediator as either a psychologist or as a combination of parents and teachers. The identity of the treatment mediator (parents and teachers, or psychologist) did not influence acceptability ratings.

Professional recommending treatment. Carter (2005) found that the title of the individual recommending a treatment for problem behaviors influenced the acceptability of treatments as rated by college students. The study compared the acceptability of medication, token economy with response cost, and a time-out procedure for problem behavior when recommended by a special education teacher, a school psychologist, and a pediatrician. Findings revealed significantly lower acceptability ratings for the token economy with response cost when recommended by a physician.

The treatment acceptability literature that has focused on treatment variable manipulation has determined some specific treatment variables that influence acceptability ratings. Another line of research within the treatment acceptability literature has involved manipulating variables associated with the client. Research involving client variable manipulations will now be described.

Client Variable Manipulation

Another line of research has involved the manipulation of variables associated with the client described within the case description and their impact on treatment acceptability. This line of research has consisted primarily of manipulations of: (a) the severity of client problem (Kazdin, 1980), (b) the age/gender of the client (Elliott & Fuqua, 2002; Sprent & Walsh, 1994), and (c) the diagnostic label of the client (Fairbanks & Stinnett, 1997; Stinnett, Crawford, Gillespie, Crue, & Langford, 2001). Research prior to 1990 involving client variable manipulation consisted of manipulations of the severity of client problem (Kazdin) and indirectly evaluating treatments for several different problem behavior such as noncompliance, self-injury, aggression, etc. (see Miltenberger, 1990 for comprehensive list).

Severity of client problem behavior. Kazdin (1980) evaluated the impact of the severity of client problem behavior on treatment acceptability. The experiment included 68 female and 26 male undergraduate participants who rated acceptability of four different treatments (reinforcement, time-out from reinforcement, drug treatment, electric shock) using the TEI. Additionally, manipulations of the clients’ problem behavior severity (moderate vs. severe) were conducted within the case vignettes. The case vignettes were presented via cassette tape operated by the experimenter in a separate observation room from the participant. Results indicated the reinforcement treatment as statistically significantly more acceptable than other treatments while electric shock was statistically significantly less acceptable than other treatments. The drug therapy and the time-out from reinforcement treatments did not differ on acceptability. The severity of the problem behavior produced a statistically significant increase in overall acceptability ratings.

Tingstrom (1990) used the IRP-15 to assess 103 teacher ratings of a time-out procedure when implemented by a teacher or a school psychologist for either mild or severe problem behaviors.
behavior. Findings revealed no significant differences based on the individual described as implementing the treatment. Significantly higher ratings were provided for time-out when applied to severe problem behavior vs. mild problem behavior.

Reimers et al. (1992) conducted both analog and clinical examinations of 40 parent ratings of treatment acceptability using the TARF-R. Parents rated the acceptability of one of three treatments (positive reinforcement, time-out, or medication) for a case description of a fictional child displaying either mild or severe problem behaviors. In addition, parents were given written descriptions and provided demonstrations of positive treatment packages (e.g., verbal praise, differential reinforcement, token systems, etc.) which were recommended for implementation with their child. Parents rated the acceptability of the recommended treatment package prior to implementing the package and at 1, 3, and 6 month follow-ups. Findings of the analog evaluations revealed higher acceptability ratings for positive reinforcement and time-out when recommended for the mild problem behaviors and higher ratings for medication when recommended for the severe problem behaviors. The influence of child problem severity on acceptability for the clinical ratings was conducted by dividing the parents into two groups based on ratings obtained from a problem behavior checklist. Findings from the clinical group indicated that parents of children with less severe problems rated the positive treatment packages more acceptable than parents of children with more severe problem behavior on all ratings obtained except for the 3-month follow-up which revealed no differences in ratings. Findings demonstrated similar ratings of treatment acceptability by parents across analog and clinical contexts. Additional analyses revealed that treatment acceptability ratings may influence treatment compliance and treatment efficacy.

Age/gender of client. The impact of client age on treatment acceptability was evaluated by Elliott and Fuqua (2002) using the AARP. They evaluated the acceptability of four interventions (habit reversal, hypnosis, medication, and punishment) for treating trichotillomania. The study presented case vignettes to 239 college students in which the age of the client (8 years, 16 years, 26 years) and the severity of hair pulling (mild vs. severe) were manipulated. Results showed significant differences among the four treatments with hypnosis and habit reversal being rated as the most acceptable interventions. The age of client and severity of trichotillomania did not significantly alter the ratings.

Spreat and Walsh (1994) assessed factors associated with decisions regarding acceptability of behavior modification programs by members from differing divisions of the AAMR. Surveys were mailed to 400 members of Region IX of the AAMR. Of these surveys, 198 were returned representing approximately a 50% return rate. The survey consisted of a case vignette with manipulations occurring among nine variables with randomly assigned values. Variables manipulated for the client described in the case vignette included sex, age, level of mental retardation, restrictiveness of residence, behavior descriptors, severity of self-injury, and frequency of behavior. A modified TEI was used to rate the acceptability of each case vignette. None of the client variables were found to be statistically significant in influencing treatment acceptability.

Diagnostic label of the client. Stinnett et al. (2001) manipulated the variable of attaching a label to the client described in the case description and the impact this manipulation had on treatment acceptability. Their study also evaluated the acceptability by teachers-in-training of two treatments, Ritalin versus special education placement, for students with a label of ADHD versus those without a diagnosis. Participants were recruited from undergraduate teacher education courses at a medium-sized university in the southwest and consisted of 27 males and 117 females. Participants were predominantly Caucasian (87.5%) with 50% having graduated from a rural high school and 49% having graduated from an urban high school. Participants were presented with a case vignette with manipulations involving label (ADHD vs. no label) and treatment (Ritalin vs. special education placement). Participants self-reported high school location at graduation and rated acceptability of interventions using the IRP-15 (Martens et al., 1985). These data were analyzed with other data collected using three 3-way ANO-
VAs. The impact of labeling on the child in the case vignette with ADHD versus no label was found to have no statistical significance and no statistically significant difference was found between the two treatments (Ritalin vs. special education placement).

Miller, Manne, and Palevsky (1998) used the TEI-SF to examine the acceptability of five behavioral treatments (positive reinforcement, chair time-out, response cost, overcorrection, and reprimands) for general or medically related noncompliance of a hypothetical child described as either healthy or with cancer. Ratings were obtained from parents of children receiving medical treatment for cancer, pediatric nurses, and parents of medically healthy children. Results revealed no significant differences in acceptability ratings based on the type of noncompliance (general vs. medically related) or based on the child being described as healthy or with cancer. Parents of children receiving medical treatment for cancer rated the response cost and the time-out procedure significantly lower than the participants. Parents of healthy children rated the positive reinforcement procedure significantly less acceptable than the other groups.

Fairbanks and Stinnett (1997) evaluated treatment acceptability of different behavior interventions associated with different diagnostic labels as rated by members of different professional groups. A vignette was presented to 31 teachers, 33 school psychologists, and 33 school social workers from three school districts in southwestern Illinois. A copy of the IRP-15 was completed by the participants to rate the acceptability of the proposed treatment described within the vignette. Variables manipulated included the diagnostic label of the child described within the case vignette. Diagnostic labels manipulated were learning disabled (LD), behavior disordered (BD), and Attention Deficit Disorder (ADD). The case vignette described a third grade boy who displayed behaviors such as excessive talking, out-of-seat, and overly active. A three-way analysis of variance was used to evaluate the data. Findings indicated no significant differences based on the diagnostic label of the child described in the case vignette.

In summary, research has shown that client variable manipulations may influence results concerning treatment acceptability. At the same time, there has been research that has illustrated that consumer variable manipulation can affect perceptions of treatment acceptability.

Rater Variable Manipulation

A third line of research identified within the literature on treatment acceptability involved examining the impact of different characteristics of raters. Prior to 1990, examination of different characteristics of raters involved: (a) gender of raters (Kazdin, 1980), (b) raters’ knowledge of treatment (Singh & Katz, 1985), and (c) affiliation of raters (Heffer & Kelley, 1987). Recent variable examination of different raters has involved: (a) geographic location of raters’ high schools (Stinnett et al., 2001), (b) raters’ knowledge of treatment (Gage & Wilson, 2000; Rasnake, Martin, Tarnowski, & Mulick, 1993; Singh & Katz), and (c) professional affiliation of raters (Fairbanks & Stinnett, 1997; Spreat & Walsh, 1994).

Gender of raters. Supplementary analyses within Kazdin’s (1980) study focused on the gender of the raters and revealed a statistically significant difference in male ratings versus female ratings, with males rating electric shock as more acceptable than females and reinforcement as less acceptable than females. The hierarchical ordering of treatments did not differ by gender. Miller and Kelley (1992) evaluated three rater variables (gender, marital adjustment, and child behavior) on treatment acceptability. The TEI was administered to 69 married couples with young children to assess the acceptability of six interventions (positive reinforcement, response cost, medication, room timeout, chair timeout, and spanking) for a hypothetical child displaying noncompliance and aggressiveness. The couples also completed a marital adjustment scale and a child behavior problem inventory. Findings revealed that gender influenced treatment acceptability ratings for all interventions except chair time-out, with mothers providing higher ratings than fathers for all interventions except spanking and medication. When compared to non-distressed couples, those couples in marital distress provided significantly higher acceptability ratings for room time-out and significantly lower ratings for positive reinforcement. Couples of children
displaying problem behaviors rated medication as more acceptable and spanking as less acceptable than couples of children without behavior problems.

*Raters’ knowledge of treatment.* Singh and Katz (1985) presented case descriptions to 96 undergraduate psychology students and asked them to rate the acceptability of four different treatments (differential reinforcement of incompatible behavior, positive practice overcorrection, time-out, and humanistic parenting) using the TEI. Results indicated the following ranking for acceptability with the greatest acceptability first: differential reinforcement of incompatible behavior, humanistic parenting, and positive practice. Time-out received ratings as least acceptable. Following this study, the same participants were provided formal educational training on three behavioral treatments (differential reinforcement of incompatible behavior, positive practice overcorrection, and time-out) and then the participants re-evaluated all four treatments. The training provided information such as specific details of each treatment, empirical data outlining the effectiveness of each treatment, and potential side effects of the treatments. At post-training, the participants’ treatment rating revealed the rankings to be differential reinforcement of incompatible behavior, positive practice overcorrection, and time-out. The humanistic parenting was rated lower at post education than at pre-education.

Rasnake et al. (1993) examined the influence of knowledge of behavioral principles on treatment acceptability. They administered the IRP to 57 institutional staff members to evaluate the acceptability of six treatments (differential reinforcement of other behavior, differential reinforcement of incompatible behavior, stimulus control, overcorrection, physical restraint, and contingent shock) for self-injurious behavior. In addition, the staff members completed a measure of their understanding of general behavioral principles. The stimulus control procedure was rated as the most acceptable of the treatments while the other treatments were not rated as acceptable treatments. Findings revealed that older residential staff rated the stimulus control procedure as significantly more acceptable than younger residential staff. Neither variables of knowledge of general behavioral principles or work experience were found to influence acceptability ratings.

**Affiliation of raters.** Heffer and Kelley (1987) found that specific affiliations of raters influenced ratings of treatment acceptability. They used the TEI to assess mothers’ ratings of five child management techniques. They found that raters from different socioeconomic classes and from different races rated acceptability of treatments differently. In a previously described study, Spreat and Walsh (1994) found differences in treatment acceptability among members of different divisions of the American Association on Mental Retardation (AAMR) assessed factors associated with decisions regarding acceptability of behavior modification programs by members from differing divisions of the AAMR. Findings indicated that the strongest indicator of treatment acceptability was the respondents’ personal estimates of probable treatment success. Members of the Psychology Division of the AAMR rated treatments as slightly more acceptable than members of other divisions of AAMR.

In a previously described study by Fairbanks and Stinnett (1997), participants evaluated treatment acceptability of different behavior interventions as rated by members of three different professional groups. Findings indicated that teachers rated the negative intervention (time-out from reinforcement with praise) more acceptable than school psychologists or school social workers. Also, school psychologists rated the negative intervention more acceptable than school social workers. Waas and Anderson (1991) used the CIRP to reveal differences in treatment acceptability ratings of 2nd graders, 5th graders, and college students. Participants rated the acceptability of a behavior contingency intervention, group counseling, and special education class placement. Findings revealed that college students rated the behavior contingency intervention and special education class placement significantly lower than the 2nd and 5th graders. No significant differences in acceptability ratings were found for the group counseling treatment.

**Geographic location of raters’ high school.** In a previously described study in which diagnostic label of the client was also examined, Stinnett
et al. (2001) evaluated the impact of geographic location of raters’ high schools on treatment acceptability. Teachers-in-training were asked information regarding the high school from which they graduated and grouped by graduation from a rural versus an urban high school. Findings indicated that location of the high school of raters influenced acceptability of treatments, with participants who graduated from rural high schools rating treatments as more acceptable than participants who graduated from urban high schools.

<table>
<thead>
<tr>
<th>Studies</th>
<th>Treatment Variable Manipulation</th>
<th>Client Variable Manipulation</th>
<th>Rater Variable Manipulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tingstrom, 1990</td>
<td>Mild vs. severe problem behavior revealed significantly higher acceptability of a time-out procedure for more severe problem behavior</td>
<td>Teacher vs. school psychologist ratings of time-out revealed no significant differences on acceptability</td>
<td>2nd graders, 5th graders, and college student revealed similarities and differences</td>
</tr>
<tr>
<td>Waas &amp; Anderson, 1991</td>
<td>2nd graders, 5th graders, and college student revealed similarities and differences</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miller &amp; Kelley, 1992</td>
<td>Behavior interventions and medication revealed similarities and differences</td>
<td>Mothers vs. fathers, materially distressed couples vs. non-materially distressed couples, and parents of children with and without behavior problems revealed similarities and differences</td>
<td></td>
</tr>
<tr>
<td>Reimers et al., 1992</td>
<td>Mildly intrusive treatments rated more acceptable for mild problem behaviors and highly intrusive treatments rated more acceptable for severe problem behaviors</td>
<td>Analog case descriptions and clinical cases revealed similar ratings</td>
<td></td>
</tr>
<tr>
<td>Betts &amp; Remer, 1993</td>
<td>Paradoxical treatment found less acceptable than nonparadoxical treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rasnake et al., 1993</td>
<td>Differences and similarities revealed among six treatments for self-injurious behavior</td>
<td>Age of institutional staff influenced acceptability; knowledge of behavioral principles and years of work experience did not</td>
<td></td>
</tr>
</tbody>
</table>
Gage and Wilson (2000) randomly assigned 30 parents of children with ADHD (as determined by the Child Behavior Check List and a diagnosis by a professional) and 30 parents of children without an ADHD diagnosis to three different conditions. The three experimental conditions consisted of: (a) medication, (b) behavioral treatment, and (c) a combination

<table>
<thead>
<tr>
<th>Studies</th>
<th>Treatment Variable Manipulation</th>
<th>Client Variable Manipulation</th>
<th>Rater Variable Manipulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreat &amp; Walsh, 1994</td>
<td>Highly intrusive treatments less acceptable than less intrusive treatments</td>
<td>Sex, age, level of mental retardation, restrictiveness of residence, behavior descriptors, severity of self-injury, and frequency of behavior revealed no influence on acceptability</td>
<td>Raters’ personal estimates of treatment success and membership in the Psychology Division of AAMR were found to influence acceptability</td>
</tr>
<tr>
<td>Burgio et al., 1995</td>
<td>Treatment setting influenced acceptability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fairbanks &amp; Stinnett, 1997</td>
<td></td>
<td>Diagnostic label of Learning Disabled, Behavior Disordered, &amp; Attention Deficit Disorder did not influence acceptability</td>
<td>Teachers rated negative intervention more acceptable than school psychologists and school social workers</td>
</tr>
<tr>
<td>Jones et al., 1998</td>
<td>Reinforcement technique rated more acceptable and spanking rated less acceptable than other techniques</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miller et al., 1998</td>
<td>Treatment described as medically related did not influence acceptability</td>
<td>Child described as healthy vs. child described as having cancer, did not influence acceptability</td>
<td>Pediatric nurses, parents of healthy children, and parents of children with cancer revealed some similarities and differences in acceptability</td>
</tr>
<tr>
<td>Stinnett et al., 2001</td>
<td></td>
<td>Label of ADHD vs. no label did not influence acceptability</td>
<td>Graduating from an urban vs. a rural high school influenced acceptability</td>
</tr>
<tr>
<td>Elliot &amp; Fuqua, 2002</td>
<td>Differences found among four treatments for hair pulling</td>
<td>Age and severity of problem did not influence acceptability</td>
<td></td>
</tr>
<tr>
<td>Carter, 2005</td>
<td>Differences found based on professional recommending a treatment</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
of both. Following a case vignette, the TEI was used to measure acceptability of treatments. Data were analyzed using three ANOVAs for each treatment scenario across TEI ratings. Results indicated that parents of children with ADHD rated medications and combination of both treatments statistically significantly higher than parents of children without ADHD. Parents of children without ADHD rated the behavioral treatment statistically significantly higher than parents of children with ADHD. Parents of children with ADHD rated the combination of treatments statistically significantly higher than the other treatments. No other significant effects were found within the parents of children with ADHD group. Among parents of children without ADHD, results revealed statistically significant differences among all treatments with the behavioral treatment receiving the highest acceptability treatment, followed by the combination treatment, and then the medication treatment.

Summary of Research on Treatment Acceptability

A summary of the research on treatment acceptability reveals that different types of treatment may produce different levels of treatment acceptability (Betts & Remer, 1993; Kazdin, 1980). The examination of the individual implementing a treatment (Kalfus & Burk, 1989) appears to have been replicated (Tingstrom, 1990) with consistent findings. The setting in which a treatment is proposed to be implemented has also been found to influence acceptability ratings (Burgio et al., 1995). Several other factors that have been found to impact the acceptability of some treatments include level of restrictiveness of treatments (Kazdin; Spreat & Walsh, 1994), information on positive outcome of treatments (Clark & Elliott, 1988), treatment history/severity (Kalfus & Burk; Kazdin; Spreat & Walsh), location of rater’s high school (rural vs. urban; Stinnett et al., 2001), raters’ professional affiliation (Fairbanks & Stinnett, 1997; Spreat & Walsh), professional recommending treatment (Carter, 2005), and raters’ experiences as a parent (having or not having a child diagnosed with ADHD; Gage & Wilson, 2000). Factors that have been found to have no effect on treatment acceptability include treatment mediators (Kalfus & Burk), client age/gender (Elliott & Fuqua, 2002), client labeling (Fairbanks & Stinnett; Stinnett et al.), and increasing raters’ knowledge of treatment (Rasnake et al., 1993). Refer to Table 2 for recent studies reviewed.

With the availability of numerous treatments for problem behaviors treatment variables appear to be the most frequently examined variables within the literature on treatment acceptability although directly examining acceptability of different treatments in isolation from other variable manipulations rarely appears as the primary focus of recent research. Alberto and Troutman (1999) have stated that a hierarchy of intrusiveness among behavior interventions is commonly accepted within the literature. Treatments that are considered more intrusive are generally considered less acceptable, but can be influenced by descriptions of client problem severity (Miltenberger, 1990).

Client/case description variables that appear to consistently influence acceptability ratings include the severity of the problem behavior being described. While in general, higher acceptability ratings are typical for more severe problem behavior, the research has shown a tendency for higher acceptability ratings when more intrusive treatments are matched to more severe behaviors and less intrusive treatments to less severe behaviors. Additionally, recent literature has evaluated several client/case description variables that do not appear to consistently influence acceptability ratings. These variables include the sex of the client, age of the client, diagnostic label of the client, and descriptions of client health.

Rater variables have been extensively examined in the recent literature and numerous interactions have been shown between rater variables and the acceptability of specific treatments. Some of the rater variables found to influence acceptability ratings include gender, age, current educational level, membership in a specific division of a professional organization, personal estimates of treatment success, marital distress, parenting a child with behavior problems, professional title (teacher vs. school psychologist), parenting a child with or without a health problem, graduating from a rural or urban high school. Knowledge of
behavioral principles and years of work experience were not found to influence acceptability ratings.

It appears that researchers have considered numerous factors that could influence the acceptability of treatments. The manipulation of variables related to treatments, clients, and raters has shown that all types of variable manipulations can influence the acceptability of ratings of treatments. With the advances that have been made with treatment acceptability instruments, the evaluation of specific treatments, specific populations of clients, and raters from specific areas of the country with varying backgrounds should become more prevalent. Miltenberger (1990) recommended that professionals could collect representative acceptability ratings from the area or school in which they work for a number of different treatments and problem behaviors. The professional could then use this information toward making treatment decisions. With the improved instruments for assessing treatment acceptability and the advances in treatment acceptability research, professionals have a well-developed basis from which to make treatment decisions and conduct further research.

References


Received: 6 March 2006
Initial Acceptance: 20 April 2006
Final Acceptance: 15 July 2006
Examining the Career Paths and Transition Services of Students with Disabilities Exiting High School

Robert M. Baer and Robert W. Flexer
Kent State University

Lawrence Dennis
Ohio Office for Exceptional Children

Abstract: Career paths and transition services were investigated for students with disabilities who were exiting special education for comparison with two models of transition developed by Siegel (1998) and Greene (2003). Teachers and parent-mentors from 52 local education agencies (LEAs) conducted a record review and an exit interview of 741 students with disabilities in their final year of high school. An SPSS two-step cluster analysis was used to group these students according to twenty-four variables relating to their postschool goals, secondary education services, and level of performance on proficiency tests. A Tamhane’s post-hoc analysis was used to identify the variables that were significantly different for each of seven identified clusters. These clusters resembled the career paths identified by Greene and had some elements in common with the transition intensity model developed by Siegel. The authors suggest that both transition intensity and career paths are key aspects of transition systems.

Seven years after Madeleine Will (1983) developed her conceptual model of how to support students with disabilities in their transition to adulthood, Halpern (1990) noted that this model represented “old wine in new bottles” starting with the work study programs of the 1960s and the career education programs of the late 70s. Will’s transition initiative has undergone a similar “re-bottling” over the past two decades. When transition was first conceptualized, it was defined in terms of the need for overlapping supports and linkages between school and the world of work. Will identified three levels of transition services—no special supports, time-limited supports, and ongoing supports. This model was expanded by Halpern (1985) to include linkages to interpersonal adjustment and residential outcomes.

Transition policy shifted considerably in the 1990s due, in part, to policy research that identified a disconnect between transition services and the rest of the curriculum (Stodden & Leake, 1994). This led prominent researchers in the field of transition to call for a “transition perspective of education” (Kohler, 1998). This shift toward a curricular focus was mirrored in the IDEA of 1997 which required both a statement of needed transition services focused on postschool outcomes and a statement of transition service needs focused on a student’s secondary course of study. The IDEA of 2004 moved transition more significantly toward a curricular focus by defining transition services as a coordinated set of activities “focused on improving the academic and functional achievement of the child with a disability to facilitate the child’s movement from school to postschool activities” [H.R. 1350 § 602 (34)].

This shift in the definition of transition services has been portrayed in both a positive and negative light. Some policy-makers have maintained that focusing transition services on improving the academic and functional achievement of students with disabilities would promote higher expectations, improve educational opportunities, encourage better teaching, and increase accountability; and that the consequent improvement of student achievement would lead to better postschool outcomes (deFur, 2002). Others have argued that this change in focus was a return to remedial academics and that progress in the general curriculum was not enough to move students with disabilities into meaningful postschool

Since the IDEA of 1997, transition advocates have tried to unify curriculum focused models of transition with earlier linkage and support models. Siegel (1998) developed a model that considered both the intensity of transition supports and the curricular needs of secondary students with disabilities. This model focused on the general curriculum as defined by the School-to-Work Opportunities Act (STWOA) of 1994, which included work-based learning, school-based learning, and connections to meaningful adult activities. By choosing the STWOA as his model for the general curriculum, Siegel questioned the value of focusing transition on a strict academic curriculum "that at best produces a 20 percent college graduation rate" (p. 149). His model identified five levels of transition intensity tailored to the needs of five distinct groups of students: (a) average to high achieving students from high income families, (b) average to high achieving students from low income families, (c) low to average academic students, (d) students at moderate risk of unemployment and underemployment, and (e) students at high risk of unemployment and underemployment due to severe disabilities and/or challenging behaviors.

Greene (2003) developed a model of transition services that emphasized the career choices of students with disabilities and the types of transition services as defined by the IDEA of 1997 (i.e., instruction, community experiences, development of employment and other postsecondary living objectives, and functional vocational education) needed to support those choices. This transition model was designed to develop transition services around pathways that were "most typically available to youth in schools today" (p. 200). Identifying four career pathways that should be available to students with disabilities, Greene posited that students could follow pathways that were: (a) fully integrated college preparatory, (b) full or semi-integrated community college preparatory, (c) semi-integrated employment and independent living preparatory, and (d) semi-integrated supported living and supported employment preparatory.

These conceptual models of transition have yet to be researched, and the purpose of this study was to examine how well these models corresponded to actual transition practices occurring in schools in Ohio. For this study, 741 exiting students were interviewed to determine postsecondary goals and transition services received. A descriptive typology was statistically derived in order to determine patterns in the exit data for comparison to Siegel’s (1998) and Greene’s (2003) transition models. Specifically, the research questions that guided this study were:

1. What were the transition pathways of students with disabilities who were exiting special education?
2. How did career pathways and service patterns of students in this sample correspond to the transition services models described by Greene and Siegel?

Method

Participants

The sample was 741 students randomly selected from the population of Ohio special education students who were graduating in the Year 2004 from the local education agencies (LEAs) in five of sixteen Special Education Regional Resource Center (SERRC) regions in Ohio. In each SERRC region, LEAs were recruited to represent a cross-section of urban, suburban, and rural areas. Urban school settings accounted for 31% of the sample, suburban 52%, and rural 17%. High schools accounted for 76% of the surveys, joint vocational schools (or career and technical centers) for 19%, and other schools (e.g., separate schools for students with severe/profound disabilities, the School for the Blind, the School for the Deaf, etc.) for 5%. The average free school lunch rate for districts that participated was 23.5% compared to 20.0% for students in Ohio schools in general (Ohio Department of Education Similar Districts Raw Data, 2004).

Within each participating LEA, exiting students were randomly selected. Table 1 shows how the exiting students with disabilities in this sample compared to the population of students with disabilities exiting Ohio in the Year 2000 according to the Twenty-Fourth Annual Report to Congress (U.S. Department of
The sample generally matched the profile of students who had exited in the Year 2000 with minor overrepresentation of minorities and some underrepresentation of students with multiple disabilities.

**Procedure**

Students with disabilities were surveyed in the spring of their final year in school. Surveys were conducted by the students’ special education teachers in person, generally as part of the students’ exit IEP. In most cases, students provided all of the information on their own, but for approximately 6% this information was provided by parents or significant others. Prior to the interview, teachers conducted a record review to obtain information about student disability designations and the secondary education programs in which they were involved. Teachers were trained in two half-day sessions on how to administer the surveys and code student responses. Teachers were allowed to paraphrase questions to make them more understandable to the students as needed.

**Instrument**

The survey used in this study consisted of two parts—a student record review and a student/family interview. The survey used in this study was derived from a follow-up survey developed by Ohio’s systems change project for transition (Baer et al., 2003), and from follow-up surveys developed for the National Longitudinal Study for Transition (Wagner & Blackorby, 1996). The survey’s face validity was tested over four years at ten LEAs (Baer et al.). During piloting, the teachers administering the surveys evaluated the survey process and each question in the survey for clarity, content, and reliability. Survey questions were dis-

**TABLE 1**

Comparison of Participants in Study to Students Graduating with Diplomas

<table>
<thead>
<tr>
<th>Demographic Characteristics</th>
<th>Study Sample</th>
<th>Ohio Exiters 2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>n(741) %</td>
<td>N (9,709) %</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>432 58.5</td>
<td>*</td>
</tr>
<tr>
<td>Female</td>
<td>307 41.4</td>
<td>*</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>592 81.1</td>
<td>8,257 85.0</td>
</tr>
<tr>
<td>African-American</td>
<td>101 13.6</td>
<td>1,283 13.2</td>
</tr>
<tr>
<td>Other Ethnic</td>
<td>37 5.3</td>
<td>169 1.8</td>
</tr>
<tr>
<td>Primary Disability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning disabilities</td>
<td>416 57.9</td>
<td>5,351 55.1</td>
</tr>
<tr>
<td>Mental Retardation</td>
<td>179 24.9</td>
<td>2,605 26.8</td>
</tr>
<tr>
<td>Other Health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impaired</td>
<td>42 5.8</td>
<td>476 4.9</td>
</tr>
<tr>
<td>Emotional Disability</td>
<td>28 3.9</td>
<td>481 4.9</td>
</tr>
<tr>
<td>Hearing Impairments</td>
<td>15 2.1</td>
<td>112 1.1</td>
</tr>
<tr>
<td>Multiple Disabilities</td>
<td>11 1.5</td>
<td>362 3.7</td>
</tr>
<tr>
<td>Visual Impairment</td>
<td>9 1.2</td>
<td>53 0.5</td>
</tr>
<tr>
<td>Autism</td>
<td>5 0.7</td>
<td>21 0.2</td>
</tr>
<tr>
<td>Orthopedic Impairment</td>
<td>5 0.7</td>
<td>116 0.5</td>
</tr>
<tr>
<td>Traumatic Brain Injury</td>
<td>5 0.7</td>
<td>53 0.5</td>
</tr>
<tr>
<td>Speech and Language</td>
<td>3 0.4</td>
<td>79 0.8</td>
</tr>
<tr>
<td>Deaf-Blindness</td>
<td>1 0.1</td>
<td>0 0.0</td>
</tr>
</tbody>
</table>

* These data were not available from the 24th Annual Report to Congress
carded or revised if the surveyors found they were hard for students to understand, lacked consistent interpretation, or elicited unreliable information based on what the surveyors knew about their students. Additionally, some interview questions were cross-checked with student records to identify whether student responses were congruent with information known about the student. Items with less than 95% agreement were discarded. After refining and editing the survey questions, the questionnaire was formatted in a bubble format for easier coding.

Analysis

Cluster analysis was chosen to address the first question of this study: What were the transition pathways of students with disabilities who were exiting special education? Results from this analysis were also a way of developing descriptive typologies and classification for comparison with the Siegel (1998) and Greene (2003) models (the second question). For this set of data, the authors chose the SPSS two-step cluster analysis because it works well with both continuous and categorical data (Al-denderfer & Blashfield, 1984; SPSS, 2003). Because there was overlap in two of the groups identified by Greene and Siegel, seven clusters were generated to provide sufficient detail for comparison with the five transition intensity levels of Siegel and the four career paths of Greene. Survey questions that were used in developing the cluster analysis were selected according to whether the same variables were discussed in the Greene and Siegel models. Table 2 provides a list of the survey questions that were used in the development of this cluster analysis. Generally, items were chosen to develop a picture of student postschool goals, career pathways, and types and intensity of transition services. The clustering technique used to group these variables was the log-likelihood criterion where the distance between two clusters depended on the decrease in log-likelihood when they were combined in a single cluster (Norusis, 2003).

After clustering, an analysis of variance (ANOVA) and a Tamhane’s T2 post hoc analysis were used to identify variables in each cluster that were significantly different from the other clusters. The Tamhane’s T2 conducts conservative pair wise comparisons based on a t test and is appropriate when the variances are unequal (SPSS, 2003). While analysis of variance should not be used to validate clusters (since the process itself is designed to yield non-overlapping variables), post hoc tests can be useful in identifying variables that were unique to each cluster (Al-denderfer & Blashfield, 1984).

Results

Cluster Analysis

Table 3 shows how participants in each cluster responded to the defining items. The clusters ranged in size from 68 to 135 graduating students. Fifteen students did not fit into any cluster and were classified as “outliers.”

Many of the defining variables were shared among the clusters, while the variables that made clusters unique generally fell into one of three categories: (a) postschool goals/career pathways, (b) disability, and (c) the academic proficiencies of students.

In regard to the first question, “What were the transition pathways of students with disabilities who were exiting special education?” Cluster 1 had a few students (7.1%) with postsecondary education goals but mainly focused on students planning to enter employment after graduation. It had a high number of students with cognitive disabilities and very low numbers of students passing their ninth grade proficiency tests. This cluster was also distinguished by a high rate of participation in work study and supported employment programs, many students planning to use SSI disability benefits as a source of income, a high level of alternate assessment, and a low rate of participation in regular academics.

Cluster 2 had a higher proportion of students planning to enter two-year colleges with the remainder planning to enter four-year colleges. It was among the three clusters that had a significantly lower rate of students passing their proficiency tests, and it also was among two clusters that had a high number of students with cognitive disabilities. Services included semi-integrated academics and some work study. More IEP focus was on residential goals. This cluster also had a significantly lower proportion of male and white students.
and a relatively high number of urban students. Cluster 3 had a high proportion of high academic achieving male students generally planning to enter employment after graduation. Services included high levels of fully in-

TABLE 2
Survey variables used in the cluster analysis

1. Did the student plan on any postsecondary education? This variable was selected to identify students in Siegel (1998) and Greene’s (2003) postsecondary education career paths.
2. Did the student plan on four-year college as an outcome? This variable was selected to separate students for whom regular and advanced academics would be critical as per Siegel’s level 1 transition intensity and Greene’s first career path.
3. Did the student plan on two-year college as a postschool outcome? This variable was selected to separate students who may need a combination of academics and career/technical classes per Siegel’s level 2 and Greene’s second career path.
4. Did the student plan on a technical school as a postschool outcome? This variable was selected to separate students who would generally need high levels of technical training and proficiency.
5. Did the student plan to work full-time after graduation? This variable was used to separate students who would typically benefit from work study, supported employment, and career/technical education.
6. Was the student from an urban setting? This variable was used because urban schools typically have high poverty levels and unique student concerns.
7. Was the student white? This variable was used to separate students who were from majority or minority cultures.
8. What gender was the student? This variable was designed to identify gender issues?
9. Was the student categorized as having a learning disability? This variable was used to separate students from the predominate disability group from other students in this study?
10. Did the student attend regular academic classes? This variable was designed to separate students who did or did not participate or have access to inclusive academic classes.
11. Did the student attend regular career/technical classes? This variable was designed to separate students who did or did not have access to inclusive career/technical education.
12. How many semesters did the student take career/technical education? This variable was used to separate out students who were career/technical concentrators.
13. Did the student participate in work study? This variable was used to separate out students who received school-supervised work experiences.
14. Did the student participate in supported employment (Option 4 in Ohio)? This variable was used to separate out students who needed intensive supports to work.
15. Did the student plan to have a driver’s license? This variable was used to separate out students who would need or not need alternative forms of transportation.
16. Did the student’s transition plan (ITP) also address residential needs? This variable was used to separate out students who had residential needs.
17. Did the student’s transition plan (ITP) also address community and social needs? This variable was used to separate out students who had community and social/interpersonal needs.
18. Did the student plan to apply for disability benefits such as SSI? This variable was used to separate out students who were eligible and in need of financial supports.
19. Did the student report passing mathematics proficiency tests? This variable was used to separate out students who had good mathematics proficiency.
20. Did the student report passing reading proficiency tests? This variable was used to separate out students who had good reading proficiency.
21. Did the student report passing writing proficiency tests? This variable was used to separate out students who had good writing proficiency.
22. Did the student report passing science proficiency tests? This variable was used to separate out students who had good science proficiency.
23. Did the student report passing citizenship proficiency tests? This variable was used to separate out students who had good social studies proficiency.
24. Did the student report taking an alternate assessment? This variable was used to separate out students who had difficulty taking state proficiency tests.
<table>
<thead>
<tr>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
<th>Cluster 4</th>
<th>Cluster 5</th>
<th>Cluster 6</th>
<th>Cluster 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n = 127)</td>
<td>(n = 68)</td>
<td>(n = 93)</td>
<td>(n = 128)</td>
<td>(n = 104)</td>
<td>(n = 135)</td>
<td>(n = 71)</td>
</tr>
<tr>
<td>Full-time work goal (63%)</td>
<td>Four-year college goal (24%)</td>
<td>Full-time work goal (82%)</td>
<td>Urban student (45%)</td>
<td>Two-year college goal (100%)</td>
<td>Four-year college goal (99%)</td>
<td>Technical school goal (100%)</td>
</tr>
<tr>
<td>Cognitive disabilities (76%)</td>
<td>Two-year college goal (72%)</td>
<td>Urban student (15%)</td>
<td>Learning disabilities (89%)</td>
<td>Urban student (15%)</td>
<td>Full-time work goal (18%)</td>
<td>Male student (83%)</td>
</tr>
<tr>
<td>Work study program (72%)</td>
<td>Cognitive disabilities (81%)</td>
<td>Male student (73%)</td>
<td>Regular academics (84%)</td>
<td>Regular academics (93%)</td>
<td>Regular academics (98%)</td>
<td>Regular academics (92%)</td>
</tr>
<tr>
<td>Supported work (23%)</td>
<td>Regular academics (57%)</td>
<td>Learning disabilities (83%)</td>
<td>Regular career/tech (66%)</td>
<td>Regular career/tech (49%)</td>
<td>Regular career/tech (32%)</td>
<td>Regular career/tech (76%)</td>
</tr>
<tr>
<td>ITP community goals (80%)</td>
<td>Work study program (46%)</td>
<td>Regular academics (93%)</td>
<td>Mean Proficiency (18%)</td>
<td>Work study program (40%)</td>
<td>Work study program (22%)</td>
<td></td>
</tr>
<tr>
<td>Plan to use SSI etc. (40%)</td>
<td>ITP residential goals (81%)</td>
<td>Regular career/tech (67%)</td>
<td>Mean Proficiency (93%)</td>
<td>Mean Proficiency (97%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Proficiency (8%)</td>
<td>Mean Proficiency (9%)</td>
<td>Mean Proficiency (92%)</td>
<td>Mean Proficiency (93%)</td>
<td>Mean Proficiency (97%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alternate Assessment (31%)</td>
<td>Male Student (37%)</td>
<td>Urban Student (41%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Items listed under a cluster differed significantly from at least one other cluster.
tegrated academic classes, and higher levels of career/technical education. This cluster had fewer urban students.

Cluster 4 did not distinguish itself from all the other clusters on any given variable. It appeared to be the product of two variables—learning disabilities and low rates of passage of the ninth grade proficiency exams. Services generally included semi to fully-integrated coursework and higher levels of career/technical education. This cluster had a higher proportion of urban students.

“Two-year college goal” clearly defined cluster 5 with all students in this cluster indicating this postschool goal. This cluster was also among the four clusters whose students generally passed their ninth grade proficiency tests. Services generally included regular academics, some career and technical education, and often work study. On a descriptive level, this cluster was also distinguished by a relatively low number of urban students.

“Four-year college goal” clearly defined cluster 6 which had significantly higher numbers of students listing this goal than all of the other clusters. This cluster was also among the four clusters whose students generally passed their ninth grade proficiency tests and had the highest passing rate. On a descriptive level, this cluster was also characterized by few students who planned to work full-time after leaving high school, a high participation in regular academics, and low rate of participation in work study programs and in career/technical education.

“Technical school goal,” was significantly higher for cluster 7. This cluster was also among the four clusters that scored well on the ninth grade proficiency test. On a descriptive level, this cluster was also characterized by a high level of participation in career/technical education and a high proportion of male students.

The second research question was “How did career pathways and service patterns of students in this sample correspond to the transition services models described by Greene (2003) and Siegel (1998)?” Table 4 shows how the clusters compared with their most similar counterparts in the Siegel and Greene models. Each cluster is labeled with the characteristics that best describe the unique features of students in each group. Cluster 1 included low academic achieving students planning to enter work and likely to use SSI. This cluster showed much similarity to Siegel’s level 5 (most intense) transition intensity and Greene’s career pathway 4 which included a semi-integrated instructional program focused on daily living skills, community instruction, career exploration, and paid work experiences.

Cluster 2 included low academic achieving students primarily with cognitive disabilities planning to enter 2 and 4-year colleges and showed marginal similarity to Siegel’s (1998) level 5 transition intensity in that these students probably had career paths with lower likelihood of success due to lack of academic proficiency. Cluster 2 did not appear to match Greene’s (2003) pathway 2 that focused on community colleges and professional school outcomes because of their failure to achieve academic proficiency. Additionally, many of these students planned to use SSI benefits after graduation.

Cluster 3 students consisted of high academic achieving students generally planning to enter employment after graduation and showed some similarity to Siegel’s (1998) level 2 transition intensity in that they were generally high academic achievers and had higher levels of career and technical education. Cluster 3 was not paired with Greene’s (2003) career pathway 3 because these students appeared to have a high likelihood of independent rather than semi-independent living outcomes after graduation.

Cluster 4 students consisted of low achieving students mainly with learning disabilities planning to enter college and/or employment, and these students showed similarity to Siegel’s (1998) level 3 transition intensity grouping in that they were low to average academic achievers, some of whom were college bound. Many of these students had community participation goals in their IEP suggesting fewer community resources. These students showed a great deal of similarity to Greene’s (2003) career pathway 3 that focused on students in semi-integrated curriculum who would need additional supports to live independently.

Cluster 5 students consisted of fairly high achieving students planning to enter 2-year colleges. These students fit well with Siegel’s (1998) level 2 transition intensity in that they had community college goals and some career and technical education. Cluster 5 students also


<table>
<thead>
<tr>
<th>Cluster and Features</th>
<th>Siegel 5 Level Model</th>
<th>Greene 4 Pathways Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1 - low academic achieving students planning to enter work and likely to use SSI. Services included semi-integrated academics more work study and supported employment. More IEP focus on residential and community goals.</td>
<td>Level 5 - Low achieving or youth with significant disabilities at high risk of unemployment and marginalization. Often have illegitimate career paths. Services include model community programs, immersion programs, and intensive follow-along services.</td>
<td>Pathway 4 - semi-integrated instructional program focused on daily living skills, community-based instruction, career exploration, and paid work experiences. Need daily living skills for independent residential and community living.</td>
</tr>
<tr>
<td>Cluster 2 - low academic achieving students primarily with cognitive disabilities planning to enter 2 and 4-year colleges and somewhat more likely to use SSI. Services included semi-integrated academics and some work study. More IEP focus on community goals.</td>
<td>No equivalent transition intensity level</td>
<td>No equivalent career path</td>
</tr>
<tr>
<td>Cluster 3 - high academic achieving students generally planning to enter employment after graduation. Services included high levels of fully integrated academic classes, higher levels of career/technical education.</td>
<td>Level 2 - Average to high achievers, some college-bound. Services include tech-prep, job and job training activities, unsupervised work experience and targeted college scholarships and options.</td>
<td>No equivalent career path</td>
</tr>
<tr>
<td>Cluster 4 - low achieving students mainly with learning disabilities planning to enter college and/or employment. Services generally included semi to fully-integrated coursework and higher levels of career/technical education.</td>
<td>Level 3 - Low to average achievers; some college bound, more from a low economic status. Services include vocational/special education, cooperatives, work experience, help with college education.</td>
<td>Pathway 3 - Semi-integrated high school curriculum leading to passage with differential standards, if necessary, of district proficiency exams and graduation. Services include career exploration, paid work experiences, and vocational evaluation.</td>
</tr>
<tr>
<td>Cluster 5 - fairly high achieving students planning to enter 2-year colleges. Services generally included regular academics, some career and technical education, and often work study.</td>
<td>Level 2 - Average to high achievers, some college-bound. Services include tech-prep, job and job training activities, unsupervised work experience and targeted college scholarships and options.</td>
<td>Pathway 2 - Semi-integrated high school curriculum, completion of graduation requirements for entrance into a community college or professional school. Services include paid work experiences.</td>
</tr>
</tbody>
</table>
showed a good fit with Greene’s (2003) career pathway 2 which included semi-integrated high school curriculum and completion of high school requirements for entrance into a community college.

Cluster 6 students consisted of very high achieving students who generally plan to enter 4-year colleges. These students fit well with Siegel’s (1998) level 1 transition intensity which included students who focused on regular academics to enter four year colleges. They also fit well with Greene’s (2003) career pathway 1 that included students who were fully integrated into college preparatory curriculum for a four year university.

Cluster 7 students consisted of high achieving students who plan to enter technical school. This cluster fit well with Siegel’s (1998) level 1 transition intensity that included tech-prep students. These students fit career pathway 1, but they engaged in career and technical education as well as college preparatory academics.

**Discussion**

The primary variables that predicted student career pathways and the intensity of transition services centered around three general types of variables tapped by the survey used in this study. Students’ academic proficiency and postsecondary education goals were the primary differentiating variables for three of the clusters—proficient students planning to enter four year colleges (cluster 6), proficient students planning to enter two year colleges (cluster 5), and proficient students planning to enter technical schools (cluster 7). Students in these three clusters generally showed a good alignment between secondary education and postsecondary goals and as a group represented 42% of the sample. These students appeared to have benefited from standards-based reforms and a transition focus of academic and functional achievement. They achieved academic proficiency and were able to participate in career and technical education as needed.

The discrepancy between postschool goals and academic proficiency defined two clusters—proficient students planning to enter employment (cluster 3) and non-proficient students planning to enter college (cluster 2). Proficient students planning to enter employment represented a fairly high number of career and technical education concentrators and a high proportion of this cluster was male students. Non-proficient students planning to enter college had a low proportion of male students and a high proportion of minority students. Cluster 3 students showed a better match between their secondary education and
postschool goals because they had a higher level of career and technical education which could lead to employment after graduation. Cluster 2 students showed a poor match between their postschool goals and secondary program. These students were not enrolled in programs that were likely to lead to college with many participating in non-integrated academics and focusing on work study. When cluster 2 was discussed with the teachers who conducted the surveys, they suggested that many of these students were young women planning to have children and work part-time after graduation.

One cluster was defined by the discrepancy between the disability and academic proficiency—non-proficient students with learning disabilities (cluster 4). The secondary program for this cluster included fairly high average levels of career and technical education. When this cluster was discussed with the teachers who conducted the surveys, they suggested that many of these students had behavioral issues that interfered with their learning.

One cluster (cluster 1) was defined by the disability (cognitive) and by high levels of school and postschool support needs. This cluster had significantly higher averages of students participating in alternate assessment, supported work in high school, work study, community goals as part of their IEP, and planning to use SSI after graduation. These students also had lower than average numbers of students participating in any regular academics.

The second research question asked, “How did the intensity of transition services and career paths of students in this sample correspond to the transition services models described by Greene (2003) and Siegel (1998)?” To some degree, this analysis tried to answer the question of whether school transition systems were operating in an exemplary fashion.

For low academic achieving students planning to enter work and likely to use SSI (cluster 1), the LEAs in this sample generally appeared to be meeting their needs as defined by Greene (2003) and Siegel (1998) in regard to transition intensity and student career pathways. These students were much more likely to receive supported employment, community-based training, and a greater emphasis on residential and community goals in their IEP. This suggested that the transition system was meeting the needs of a high proportion of these students.

For low academic achieving students planning to enter college (cluster 2), very little from either the Greene (2003) or Siegel (1998) models appeared to be happening. Though marginally similar to Siegel’s level 5 transition intensity student (in that these students probably had undeveloped career paths), there was no evidence of the degree of transition intensity that these students would need to succeed under the Siegel model. These students also appeared to be outside any of the exemplary practices suggested in the Greene career pathway leading to 2-year colleges. The Siegel and Greene models suggest two alternative ways of approaching students in this cluster. Under the Siegel model, the intensity of transition services would need to considerably increase both in school and postschool to provide any reasonable chance of success. Under the Greene model, a focus on improving self-determination would help these students choose career paths better aligned with their academic achievement.

High academic achieving students planning to enter employment (cluster 3) showed a fairly high level of agreement with Siegel’s (1998) transition intensity level 2 model. Students in this cluster generally took more career and technical education and, therefore, had a reasonable chance of success under the Siegel model. None of the career pathways described by Greene (2003) matched this cluster very well. Under Greene’s model, these students would generally be encouraged to consider postsecondary education through college visits, functional vocational evaluation, and self-determination training.

Low achieving students with learning disabilities (cluster 4) generally fell in line with Siegel’s (1998) transition intensity 3 and Greene’s (2003) career pathway 3. Siegel suggested that these students needed vocational special education, which in light of the low proficiency levels of these students was probably provided in their career and technical education settings. Greene would emphasize career exploration, paid work experiences, and vocational evaluation to help these students find settings where their lack of academic proficiency would be less critical.
High achieving students planning to enter two year colleges (cluster 5) showed many similarities to Siegel’s (1998) transition level 2 and Greene’s (2003) pathway 3. Siegel’s model emphasized the need for career and technical education for these students, while the Greene model focused more on progress in the general curriculum. Both emphasized the need for transition supports related to finding appropriate colleges and financial support.

High achieving students planning to enter four year colleges (cluster 6) showed a high level of agreement with Siegel’s (1998) transition intensity level 1 and with Greene’s (2003) career pathway 1. Both models agreed that these students should be highly focused on progress in the general academic curriculum. Greene’s model had greater emphasis on career exploration and work experiences than was evident for this cluster. Many of the teachers conducting the surveys suggested that these students needed more self-determination training to request the needed accommodations while in college, since they noted that special education had only very limited contact with these students.

High achieving students planning to enter technical school after graduation (cluster 7) were lumped into Siegel’s (1998) transition level 2 and Greene’s (2003) career pathway 2. Participants in this sample appeared to be higher functioning academically than suggested in the Greene model, and most participated in regular academic classes. The high level of proficiency among these students suggested that tech-prep and programs leading to professional schools may be as demanding academically as the regular college preparatory tracks.

**Limitations**

It should be emphasized that cluster analysis is heavily influenced by its interpretation. Al-enderfer and Blashfield (1984) have presented three general precautions that must be considered in interpreting cluster analyses:

1. Cluster analyses are inevitably biased by the questions being asked of the data.
2. Different clustering methods can generate different solutions to the same data set.
3. The strategy of cluster analysis is structure seeking and will always place objects into groups.

In regard to the first limitation of cluster analysis, the Siegel (1998) and Greene (2003) models were used to reduce the bias in questions being asked of the data. Additionally, the results and conclusions of this study were discussed extensively with teachers conducting the surveys through two follow-up meetings and their interpretations were used to triangulate the conclusions drawn by the researchers. The choice of seven clusters undoubtedly affected the interpretation of the data. More clusters would have generated more alternative structures and would have probably provided less congruence with the five and four group models of Greene and Siegel.

Care should also be taken in generalizing these findings to sub-groups, especially students with severe disabilities. Most of the students in this study had learning disabilities and coupled with students with other health impairments (such as attention deficit-disorder), students with mild disabilities were a large majority of the sample. Students with more severe disabilities tended to be clustered with students with cognitive disabilities, so conclusions about sub-groups (such as multiple disabilities, autism) were not possible with this form of analysis. Also, it should be noted that this sample only included students with disabilities who aged out or graduated from secondary education. Consequently, these findings can not be generalized to students who dropped out.

**Future Research**

Additional research will be required to determine how students in each of the career pathways and transition intensity levels fared after graduation. Also, additional research is needed to track how decisions about career paths and the intensity of transition services were made. As Greene (2003) emphasized in his discussion, career paths can easily become career tracks if choices are not based on the decisions of students and family who are self-determined and empowered.

Additional research is also needed on the relationship between the intensity of transition services and the career paths of students.
with disabilities. Should the primary point of intervention be to ensure students choose career paths that have a high probability of success, or should secondary educators enhance the intensity of transition services for students who choose unconventional career paths?

**Conclusion and Recommendations**

1. The concept of transition will need to develop a new vision to rapidly adapt to the skyrocketing postsecondary education goals of students with disabilities.

The skyrocketing expectation for postsecondary education identified in this study raised a number of issues. First, transition coordinators will need to broaden their focus and develop expertise in helping students with disabilities transition to college. Teachers who conducted the surveys reported that they were surprised by the numbers of students in postsecondary paths. They noted that many of these students were not receiving transition services through special education and/or that special educators were not aware of the postsecondary education goals of their students. A related question was whether postsecondary educational systems will be able to accommodate these students. Currently, national data suggests that only about half of students with disabilities anticipating postsecondary education ever enroll, and even fewer completed postsecondary education (Cameto, 2003). It is essential that transition supports be infused into both the secondary and postsecondary environments in order to support these students, including continued commitment to support access to standards-based education and appropriate academics in secondary school.

2. New models of transition focused on academic and functional achievement will require greater involvement of regular and career/technical educators, school guidance counselors, and other school and adult services personnel.

The number of students with disabilities planning to enter postsecondary education suggested there is a continuing need to build capacity within the general curriculum by raising the expectations for students with disabilities and by developing a shared responsibility throughout the educational system for improving results for all learners. These new opportunities to collaborate will enhance the transition planning process, providing applied learning experiences outside the classroom and integrating work experiences into the students’ educational experiences. Continued expansion of this collaboration will be required to ensure that students are provided the breadth of experiences that lead to positive post school outcomes for students.

3. This vision will require continued professional development for the individuals coordinating transition services for students with disabilities along with other educators and agency personnel that support the student’s transition plan.

To promote collaboration in providing transition services in the general curriculum, it will be necessary to train regular educators, guidance counselors, and other ancillary staff regarding the transition needs of student with disabilities. Guidance counselors will need to become more familiar with the major adult services and disabilities supports available in employment and postsecondary settings for students with disabilities. There is also a critical need to develop assessment, curriculum, and instructional strategies that are relevant for all students. This will require training to promote high expectations for students with disabilities, systemic use of instructional accommodation and strategies, and assuring that students with disabilities have access to the full range of secondary education curricula to reduce the gaps between students with disabilities and their non-disabled peers. This move toward shared responsibility will require a focus on outcomes and the continued development of data collection that will allow state, regional, and local educational agencies to promote policies, procedures, and practices associated with improved postsecondary outcomes for all students.

**References**


(NLTS2): Collecting and utilizing postsecondary outcome data to improve transition programs and services at the national, state, and local levels.” Capacity Building Institute, Division of Career Development and Transition, Council for Exceptional Children, Roanoke, Virginia.

dcFur, S. H. (2002). Education reform, high-stakes assessment, and students with disabilities: One state’s approach. Remedial and Special Education


Received: 27 March 2006
Initial Acceptance: 10 May 2006
Final Acceptance: 1 September 2006
In this study, 55 youth with special health care needs (YSHCN) were surveyed to understand their transition experiences. The survey asked about educational and vocational goals, current health care (including access to adult care providers, health insurance, and medications), life experiences (exercise, doing chores, cooking, types of assistance needed in everyday routines), and social life (activities for spare time). The survey results showed that many YSHCN have plans for continuing education and clear goals for future employment. However, most did not have transition plans, or were not aware of having made transition plans. Many who identified goals were unable to explain how they were going to reach their goals. Although most described themselves as happy, about 16% evidenced signs associated with depression. The study concludes that more work is needed in both research and practice to help youth with special needs make the best transition possible in education, vocation, and everyday living arrangements.

“Special health care needs” is an umbrella term that refers to about 200 chronic conditions affecting children and youth, including physical problems (e.g., asthma, diabetes, cerebral palsy, spina bifida), cognitive deficits (e.g., down syndrome, fragile X syndrome), developmental delays, and behavioral or emotional conditions (Newacheck & Stoddard, 1994). Eighteen percent of children in the United States under 18 years of age, or 12.6 million children nationally, were reported to have chronic medical conditions that required health and related services of a type or amount beyond that required by children generally (Newacheck et al., 1998). It was estimated that more than 500,000 youth with special health care needs (YSHCN) will turn 18 each year (Center for Disease Control and Prevention, 2004).

Correspondence concerning this article should be addressed to Cevriye Ergul, 5111 North Scottsdale Road, Suite 158, Scottsdale, AZ 85250.
Only 13% of students with learning disabilities (LD) have attended a 4-year post-secondary school program within two years of leaving high school compared to 53% of students in the general population (National Center for Learning Disabilities, 2004). Given the strong correlation between years of completed education and employment status, it is not surprising that youth with disabilities have a particularly difficult time finding a satisfactory job or even being employed. From 40% to 70% of youth with disabilities were found to be unemployed (Blackorby & Wagner, 1996; Wehman, 1992) with the rate much higher for those with moderate and severe mental retardation, ranging from 85% to 90% (Wehman, Kregel, & Seyfarth, 1985). Similarly, approximately 73% of students with multiple disabilities and orthopedic impairments were not employed (Wagner, 1993). Being unemployed in turn reduces the opportunity to lead an independent life. About 65% to 75% of students with disabilities have been unable to live independently (Blackorby & Wagner; Wehman); only 13% of students with multiple disabilities and 38% of students with orthopedic impairments were able to achieve independent-living status 3 to 5 years after exiting high school (Wagner).

Attention is now being directed to implementing policy and developing programs that ensure youth have the opportunity to be “healthy and ready to work;” that is, that they have the knowledge and skills necessary to make successful transitions to young adulthood. The 1997 amendments to the IDEA mandated that students with disabilities must have a statement of transition service needs in their individualized education programs (IEPs) by age 14. The 2005 re-authorization stipulated age 16. The term “transition services” refers to a coordinated set of activities for a student with a disability that: (a) is designed within an outcome-oriented process that promotes movement from school to post-school activities, including post-secondary education, vocational training, integrated employment (including supported employment), continuing and adult education, adult services, independent living, or community participation; (b) is based on the individual student’s needs, taking into account the student’s preferences and interests; and (c) includes instruction, related services, community experiences, the development of employment and other post-school adult living objectives, and when appropriate, acquisition of daily living skills and a functional vocational evaluation. IDEA requires that students be involved in transition planning and that students’ preferences and interests be taken into account when transition services are planned.

Although effective transition practices have been identified (Blackorby & Wagner, 1996; Kohler & Field, 2003), the extent to which they have been implemented varies widely. Wehmeyer and Schwartz (1997) found that none of the transition goals of students with mental retardation or developmental disabilities targeted instruction related to self-determination skills. Trainor’s (2005) study of high school postsecondary transition planning (ITP) meetings for students with LD found that exit goals did not match students’ interview responses, and students were largely uninvolved in their transition planning process. Although most students signed their ITPs, none could later recall the specifics of the meeting, the contents of the document, or the definition of the term transition plan. McMahon and Baer (2001) and Everson, Zhang, and Guillory (2001) found that school districts were technically complying with IDEA but best transition practice was not in place.

With respect to youth with special health care needs, little is known about the transition services they receive, their transition plans, or their knowledge of how to go about achieving the goals they set for themselves. What is known is that youth with special health care needs achieve limited outcomes as they leave high school and attempt to access adult health, independent living, post-secondary education, employment, and adult community services (Gloss, Reiss, & Hackett, 2000; Thompson, McGrew, Johnson, & Bruininks, 2000). Specific problems that contribute to poor outcomes for youth with special health care needs have been identified by HRSA’s survey: Special Programs of Regional and National Significance (SPRANS). The problems include: (1) no experience managing their own health care, making medical appointments or even discussing the specifics of their medical conditions, (2) barriers in seeking...
education and employment opportunities, (3) adults around them have extremely low expectations of their abilities and future prospects, (4) present barriers to attaining a degree of independence that would be considered matter of fact for youth without disabilities, (5) families may be unaware of the existence of programs and resources and how to access them (6) pediatric and health care professionals often do not communicate, much less collaborate, to achieve successful transition of care as children with special health needs mature, and (7) health care systems do not interact with education, rehabilitation or insurance systems in planning or facilitating transition.

Because the category special health care needs includes youth who are extremely heterogeneous in both types and severity of physical, health and cognitive limitations, it is difficult to know if the SPRANS data can be generalized to all youth with special health care needs. Moreover, little is known about the transition experiences of youth with special health care needs. For example, youth with special health care needs who do not receive special education services may not be provided any transition services, and professionals and families may not consider transition planning a priority for youth who have severe physical and cognitive impairments.

The purpose of this study was to examine the transition to the various areas of young adult life of youth with special health care needs who vary in their capacity to be “healthy and ready to work.” The intent was to determine whether they had received transition services and their success in making transitions to adult health care, postsecondary education, and their assumption of responsibility for daily life routines. Thus the focus was on educational and vocational goals, current health care (e.g., access to adult care providers, health insurance, medications), life experiences (exercise, doing chores, cooking, types of assistance needed in everyday routines), and social life (e.g., activities for spare time).

Method

Participants

Youth. A pool of potential participants included 323 youth who were aging out of Children’s Rehabilitative Services (CRS) in Phoenix, Arizona. Youth were recruited from a list of youth invited to attend a yearly CRS Transition Fair that provides information and resources to CRS patients who had reached the age of 18 years and their families. Thus, the pool included youth preparing to transition out of high school and the CRS.

The 323 youth were contacted by phone to complete the phone survey: 129 had outdated phone numbers, 86 were unwilling to participate (didn’t speak English, deaf, deceased, unable to speak), and interviewers were unable to connect with 53. Fifty-five youth completed the survey. Table 1 presents the primary diagnostic categories, age, gender, ethnicity, and primary language of participants.

Survey. The Survey included 65 items that assessed: Transition Plans, Insurance and SSI, Health (problems, exercise, diet, smoking, and alcohol usage), Education, Independent Living, Employment, and Social Activities. It was designed using the Healthy and Ready to Work (Gloss et al., 2000) Center’s proposed outcomes measure of services and supports that youth with special health care needs require to make successful transitions (see Table 2). Items also assessed attributions for successful and unsuccessful social outcomes, general mood and indications of depression (e.g., recent changes in activity level and appetite).

Procedure

Youth with special health care needs who had participated in a Youth Leadership Program conducted the survey two nights per week over a four-week period. Interviewers created a call log and made initial calls to each of the 323 participants. Interviewers recorded the results on the call log as answered busy, call back later, disconnected number, gave up after 10 rings, fax machine, answering machine, and answered but refused immediately. Based on these results, interviewers sorted the call logs indicating whether additional calls were needed and made a second round of calls. Again, results were recorded and sorted based on whether additional calls were needed. Interviewers then conducted a third and fourth round of calls to remaining participants. In-
terviewers stopped calling after four unsuccessful attempts to reach participants.

Results

Education. The results are summarized in Table 3. About 50% of the participants were attending high school and 20% were attending college. Although the youth were all above the age at which they should have received transition services, only 30% indicated that they had a written transition plan. The majority indicated they had future educational goals. However, when asked what they needed to achieve their goals, only 25% were specific in identifying financial aid. The responses of the remainder were nonspecific (e.g., “try hard, don’t know, education”).

Health. Although many concerns have been voiced about youth with special health care needs not having health care and insurance, the participants reported that they were largely covered. Most had health insurance through their parents’ employment, and a few, through their own employment. About 50% of the youth received Medicaid. There should be some concern for whether youth covered by their parents’ insurance will lose their benefits when they reach majority. The majority of youth indicated they had a primary care physician and about 66% saw specialists. However, about 20% reported that they did not have a dentist.

Most rated their health care as good or excellent. Nonetheless, 27% had been hospitalized at least one during the previous year, and 27% had at least one visit to the emergency room. About 50% take medications regularly, and 24% admit to sometimes forgetting. More than 50% indicated they exercise regularly, and are not on special diets. They do not smoke and few reported drinking. About 16% related having recent changes in energy, sleep, concentrating, and headaches.

Employment. About 20% indicated they had jobs. The jobs were minimum wage (e.g., bussing tables, Burger King) or volunteering. About 66% had identifiable career and employment goals. Some were nonspecific (e.g., work with children, get a job and be self supporting, rich), while others were very specific.

TABLE 1
Participant Demographic Characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Male (n = 22)</th>
<th>Female (n = 33)</th>
<th>Total N = 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnostic Category</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurology</td>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Cognitive</td>
<td>7</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Genetic</td>
<td>1</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Unknown</td>
<td>9</td>
<td>20</td>
<td>29</td>
</tr>
<tr>
<td>Age</td>
<td>17–29</td>
<td>17–28</td>
<td>17–29</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>19.36 (3.26)</td>
<td>18.82 (2.98)</td>
<td>19.04 (3.08)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>14</td>
<td>20</td>
<td>34</td>
</tr>
<tr>
<td>Hispanic</td>
<td>6</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>African American</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Primary Language</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>19</td>
<td>28</td>
<td>47</td>
</tr>
<tr>
<td>Spanish</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>English and Spanish</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

1 (e.g., Cerebral Palsy, Spina Bifida, Hydrocephalus, etc).
2 (e.g., Autism, Down Syndrome, Mental Retardation, etc).
3 (e.g., Fragile X Syndrome, Lesch-Nyan Syndrome, etc).

Readiness for Transition / 333
(lawyer, CPA, teacher, doctor, customer service).

**Life Experiences.** With respect to independent living, about 60% described themselves as independent, and 11% need a lot of assistance. While 80% reported doing chores around the house, less than 50% shop for groceries, do laundry, or cook meals. In about 75% of the cases, transportation is provided by parents or public transportation with 14% driving to school. Youth described themselves as very social, making lots of phone calls, and spending time on the internet. Fewer than half reported having difficulty with relationships. In general, youth described themselves as happy most of the time.

**SSI.** 40% reported they received SSI and an additional 4% reported receiving SSDI. 50% indicated that they received either Medicaid or Arizona Health Care Cost Containment System (AHCCCS). AHCCCS is Arizona’s Medicaid Program that provides health care coverage to those who can not afford to pay for health care.

**Discussion**

In general, the results of the survey were positive. Many of the youth are attending secondary and post-secondary schools with plans for continuing education and have clear goals for future employment. They have health insurance and regularly visit their physicians. They describe themselves as being in good health, being independent, having positive moods, having friends, and doing chores around the house.

There were several problematic findings as well. Most of the respondents did not have transition plans, or were not aware of having...
| TABLE 3  
Youth Survey Results (N = 55) |
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Education</td>
</tr>
<tr>
<td>● 71% attending school</td>
</tr>
<tr>
<td>● 46% receive general education, 24% special education, 4% both general and special ed</td>
</tr>
<tr>
<td>● 33% have a written transition plan</td>
</tr>
<tr>
<td>● 71% have future educational plans</td>
</tr>
<tr>
<td>2. Health</td>
</tr>
<tr>
<td>● 93% see a primary care physician</td>
</tr>
<tr>
<td>● 98% have medical insurance</td>
</tr>
<tr>
<td>● 67% see a specialist</td>
</tr>
<tr>
<td>● 80% see a dentist</td>
</tr>
<tr>
<td>● 9% had problems finding a doctor</td>
</tr>
<tr>
<td>● 2% rated their health as poor, 15% fair, 67% good, 15% excellent</td>
</tr>
<tr>
<td>● 67% had not been hospitalized since Jan/03, 15% 1 time, 6% 2 times, 6% 3 times</td>
</tr>
<tr>
<td>● 66% had no emergency visits since Jan/03, 15% 1 time, 6% 2 times, 6% 3 times</td>
</tr>
<tr>
<td>● 49% take vitamins/supplements/herbal medications, 49% do not</td>
</tr>
<tr>
<td>● 51% take medications regularly, 24% sometimes forget</td>
</tr>
<tr>
<td>● 11% had problems getting their prescriptions filled, 73% had no problems</td>
</tr>
<tr>
<td>● 44% have allergies</td>
</tr>
<tr>
<td>Healthy Habits</td>
</tr>
<tr>
<td>● 66% exercise regularly</td>
</tr>
<tr>
<td>● 16% are on a special diet</td>
</tr>
<tr>
<td>● 6% smoke</td>
</tr>
<tr>
<td>● 7% drink regularly</td>
</tr>
<tr>
<td>● 16% have experienced changes in appetite/body weight, 36% have not</td>
</tr>
<tr>
<td>Mood Changes</td>
</tr>
<tr>
<td>● 13% recently have experienced energy loss</td>
</tr>
<tr>
<td>● 16% recently have experienced difficulty sleeping</td>
</tr>
<tr>
<td>● 18% have experienced difficulty concentrating</td>
</tr>
<tr>
<td>● 18% recently have experienced headaches, muscle aches, or stomach aches</td>
</tr>
<tr>
<td>3. Employment</td>
</tr>
<tr>
<td>● 20% have a job</td>
</tr>
<tr>
<td>4. Life Experiences</td>
</tr>
<tr>
<td>● Self-care: 60% independent, 26% need minimal assistance, 11% need a lot of assistance</td>
</tr>
<tr>
<td>● Household chores: 80% do chores</td>
</tr>
<tr>
<td>● Grocery Shopping: 9% grocery shop, 66% parents, 22% youth with parents</td>
</tr>
<tr>
<td>● Laundry: 38% youth, 35% parents, 4% attendants, 20% youth with parents</td>
</tr>
<tr>
<td>● Cook Meals: 15% youth, 53% parents, 2% attendants, 27% youth with parents</td>
</tr>
<tr>
<td>● Transportation: 14% drive to school, 38% use public transportation, 37% parents drive, 2% take school bus, 4% walk</td>
</tr>
<tr>
<td>Social</td>
</tr>
<tr>
<td>● 2% rate themselves as lonely, 16% not very social, 42% social, 36% very social</td>
</tr>
<tr>
<td>● 40% make 0–5 phone calls Weekly, 22% 6–10 calls, 11% 11–15 calls, 4% 16–20 calls, 20% more than 20 calls</td>
</tr>
<tr>
<td>● Internet Weekly: 67% spend 0–5 hours; 17%, 6–10 hours; 6%, 11–15 hours; 7%, 20+ hours</td>
</tr>
<tr>
<td>● Relationships: 7% recently have experienced difficulty with relationships, 47% have not</td>
</tr>
<tr>
<td>● Mood in general, 76% feel happy most of the time, 15% some of the time, 2% once in a while</td>
</tr>
<tr>
<td>● Mood in general, 6% feel sad most of the time, 22% some of the time, 55% once in a while, 13% never</td>
</tr>
<tr>
<td>5. SSI</td>
</tr>
<tr>
<td>● 40% get SSI</td>
</tr>
<tr>
<td>● 4% get SSDI</td>
</tr>
<tr>
<td>● 49% get Medicaid/AHCCCS</td>
</tr>
</tbody>
</table>
transition plans. It is possible that they did not receive transition plans because they were not eligible; that is, they did not have IEPs and were not receiving special education services. But assuming that at least some were eligible for transition services, it would be difficult to forget receiving the types of services specified by IDEA (e.g., vocational training or integrated employment), and IDEA requires that students be involved in transition planning and that students’ preferences and interests be taken into account when transition services are planned. Perhaps they received these services but it was not called transition planning. It also is plausible that transition services are fairly new to school districts, and they are scrambling to do their best to provide these services. Moreover, the types of support included in transition services have not been traditional components of high school curricula, and school personnel may not be fully aware of how to tap into community resources that are supposed to provide complementary services (e.g., vocational rehabilitation services). It also is likely that there is considerable heterogeneity in school districts approaches to transition. The guidelines are broad and there is no particular curriculum to guide transition services.

Although many youth could specify future educational and vocational goals, many did not. Many who identified goals were unable to explicate how they were going to reach their goals. Aside from recognizing the need for financial support to continue their education, there were few specific responses. This type of planning typically starts when children are young, and these results raise the concern that families and schools have low aspirations or expectations for these youth and neglect starting this type of planning process. At the same time, many youth (without disabilities) also complete school without knowing what they plan to do. But because so much more effort is required for youth with special needs to maximize their opportunities, it is crucial that planning start much earlier, that families and teachers raise their expectations for what youth can achieve and systematically build a foundation for transition.

Youth rated themselves as healthy, but they reported incidence rates of hospitalizations and visits to the emergency room that are much higher than the average rate in Arizona. The report, Status of Adolescent Health in Arizona (Arizona Adolescent Health Coalition, 2004), indicated that 5.87% of youth, ages 15-19, were hospitalized in Arizona in 2001. This is much lower than the 27% reported by youth in this survey for hospitalizations and visits to the emergency room. To the extent that increased hospitalizations and emergency room visits affect school attendance, capacity to participate in, or attention to, academic activities, continuing health problems (and medications) may well be a significant influence in transition outcomes. More research is needed on the interactive role of health and learning and how to ensure that health is included in transition planning.

Youth rated themselves as happy, about 16% evidenced signs associated with depression. The structure of the interview did not permit follow up questions to the 16% of youth who described themselves with symptoms of depression. This is unfortunate as 4 out of 5 cases of depression among youth with special needs go undetected and only half of those detected get referred to mental health specialists (Costello, 1986). Youth with special needs experience a higher incidence rate of mental health problems than their peers; indeed, depression is the most frequently cited secondary problem for several special health care problems (Lollar, 1994). There are many plausible explanations for the increased rate of depression among youth with special needs, from the secondary effects of medication to the growing realization that the disability is a lifelong condition. This is clearly a significant issue that can affect youth participation in the transition process. But, like health, mental health is not an issue that many teachers are prepared to deal with, and Arizona has limited resources for mental health referrals.

The study has a number of limitations. First, the size of the sample is relatively small and limited information is available about the characteristics of the youth. So the generalizability to other populations is not evident. Furthermore, survey data always have to be interpreted taking into account the reliability and validity of respondents. The respondents did not have advance notice of the calls, so it is possible that their responses to questions...
about transition plans, education and career goals would have been more detailed had they had prior notice of the questions. Nonetheless, these data are consistent with others who have found that the implementation of best practices in transition has not yet been realized.

The interviews were conducted by youth, which may have biased responses, at least in the areas of mood and social encounters. The youth who responded to questions about their moods and social lives with extremely positive answers (e.g., “happy all the time,” “make more than 20 calls a week”) may have been trying to impress the interviewers. By this age, socially aware individuals understand that moods fluctuate and no one is always happy. Glowing reports about a high frequency of social encounters are similarly suspect. This raises the possibility that the incidence of depression was understated.

Much progress has been made to establish transition planning as an integral component of secondary education for youth with disabilities. These findings suggest, however, greater efforts need to be made to ensure that YSHCN have transition plans, that they have been included in developing their plans, and that the plans provide specific strategies for realizing proposed outcomes. Youth’s murky ideas about how to reach their goals suggest that having a transition plan in place is the first step, that the process should start at a young age, and that youth need continuing guidance to translate their dreams into reality.

References


Readiness for Transition / 337
analysis of postschool follow-up data. *Exceptionality*, 8, 73–90.


Received: 27 March 2006
Initial Acceptance: 23 May 2006
Final Acceptance: 15 August 2006
Training Paraeducators to Promote the Use of Augmentative and Alternative Communication by Students with Significant Disabilities

Mary Anna Bingham, Fred Spooner, and Diane Browder
University of North Carolina at Charlotte

Abstract: The purpose of this study was to determine the impact of training paraeducators on (a) paraeducator prompting use of augmentative communication (AAC) systems, (b) paraeducator responding to student requests, (c) student use of AAC, and (d) student problem behavior via a series of multiple probe designs. Participants were three paraeducators and students. Paraeducators were trained on (a) importance of communication, (b) relationship between behavior and communication, (c) use of AAC, (d) how to prompt students to use AAC and respond to communications, and (e) how to self evaluate their behavior. All paraeducators increased the number of times they prompted student use of AAC and responded to student requests. All students increased use of AAC and exhibited fewer problem behaviors.

All people communicate in order to express their wants and needs. In fact, the ability to communicate in meaningful and acceptable ways is fundamental to participation in our society (Drasgo & Halle, 1996; Kaiser & Grim, 2006). Students with significant disabilities almost always have limitations in their receptive and expressive language skills (McLean, Brady, & McLean, 1996; Siegel & Wetherby, 2006; Sigafos & Pennell, 1995). They may have difficulty understanding spoken language as well as expressing their wants and needs. An inability to communicate with others presents serious problems for students with significant disabilities including the use of problem behaviors as communicators. Problem behaviors may take the form of tantrums, self-injury, aggression toward others, and property destruction (Day, Horner, & O’Neill, 1994; Dunlap, Ferro, & dePerczel, 1994; Durand, 1990; Ferro, Foster-Johnson, & Dunlap, 1996; Horner, Albin, Todd, & Sprague, 2006; Repp, Felce, & Barton, 1988).

Students with challenging behaviors are likely to be excluded from general education programs and social activities. If children with significant disabilities are to grow up and hold meaningful jobs in the community, take part in leisure events and activities with the general public, and live as independently as possible, they need to communicate effectively with other people in all these environments (Sack & McLean, 1997). The inability to communicate and challenging behaviors also interfere with students’ ability to be included in regular education programs, to participate in social activities, and to lead full, productive lives. Teachers and paraeducators who support these students need strategies for promoting appropriate communication and decreasing occurrences of problem behaviors.

Bird, Dores, Moniz, and Robinson (1989), Carr and Durand (1985), Durand and Carr (1991), Horner and Budd (1985), and Wacker et al. (1990) suggest that functional communication training and the use of augmentative and alternative communication benefit students with communication deficits who exhibit challenging behaviors. On the other hand, many teachers and paraeducators supporting students with significant disabilities face challenges when implementing these practices (Johnston, Reichle, & Evans, 2004).
For example, some students (a) have AAC systems but are not using them, (b) have AAC systems but their communicative partners are not using them, or (c) use alternative, but socially inappropriate strategies for communication. Additionally, some students have learned that behaviors such as tantrums gain attention more quickly than pointing to a symbol. Students may not have the opportunity to use AAC systems due to limited interactions with others and lack of understanding on how to use the systems. Children with severe disabilities who use AAC systems experience difficulties in their daily interactions (Light, 1988) and often have few opportunities to communicate with others (Calculator & Luchko, 1983). Moreover, some adults who interact with these students do not understand the importance of communication in controlling behavior. If significant others cannot understand efforts to communicate, they may not be able to respond appropriately, which results in increases in problem behaviors (Durand, 1990; Keen, Sigafoos, & Woodyatt, 2001; Reichle, 1997).

Education and support of students with significant disabilities is often the responsibility of paraeducators. Many paraeducators feel unprepared to meet the needs of these students (Giangreco, 2003). As a result, the need for investigations to determine effective interventions for paraeducators is critical. In order for paraeducators to support student use of AAC, they must fully understand the relationship between behavior and communication (Durand, 1999).

The roles of paraeducators providing supports to children with disabilities have changed significantly from being teacher aides who provide clerical assistance to being support personnel in the areas of instruction, tutoring, and management of classroom behavior (Harrington & Mitchelson, 1986; Jones & Bender, 1993; Pickett, 1997). Paraeducators are now providing direct services to students including behavior management, implementing and evaluating treatment and intervention programs, and instruction (French & Cabell, 1993). A critical consideration is the level of staff training, given that paraeducators are typically responsible for ensuring that students with disabilities have access to and are able to effectively utilize AAC devices (Butterfield & Arthur, 1995). Unless properly trained, paraeducators may disregard communicative actions of students who use AAC devices (Calculator & Jorgensen, 1991). When students’ communicative attempts are ignored, increases in problem behaviors occur (Durand, 1999; Lalli, Casey, & Kates, 1995).

Paraeducators provide much needed supports and are critical to student success. The literature suggests that many paraeducators continue to be undertrained or untrained (Giangreco, 2003) and are often assigned to work with students who have the most complex learning or behavioral challenges (Giangreco, Edelman, & Broer, 2001). Paraeducators continue to engage in a broad range of roles, many of which they are untrained or insufficiently trained to perform (Blalock, 1991; French & Pickett, 1997). Training paraeducators in communication use may increase student use of AAC and reduce aberrant behavior. Yet paraeducators who support these students in special education and general education classrooms often do not receive appropriate training on how to implement AAC.

As more students with significant disabilities are included in general education classrooms, the need for additional training for paraeducators is evident. Training staff using a treatment package that includes self-evaluation may increase the likelihood of success for students with significant disabilities. The purpose of this study was to determine if training paraeducators using a treatment package including self-evaluation results in increased prompting of student use of augmentative communication devices and increased responding to student attempts to communicate.

Method

Participants and Settings

Staff (paraeducators). Three paraeducators who support students who have (a) severe/profound or multiple disabilities, (b) use augmentative communication systems to communicate, and (c) exhibit problem behaviors participated in this study. These paraeducators worked in self-contained classrooms operated by a community agency under a contrac-
tual arrangement with a large urban school district.

Letters describing the study and inviting participation were sent to paraeducators. Paraeducators volunteered to participate in the study by returning a letter of informed consent. Three paraeducators were randomly selected from the group of 18 paraeducators who returned letters (see Table 1).

Students. Three students who attended the classrooms of the paraeducators who were randomly selected participated in the study. Students selected had (a) a diagnosis of severe/profound or multiple disability, (b) a history of challenging behavior (e.g., aggression, self-injury, property destruction) as determined by teacher report, individualized education plan (IEP), and behavioral intervention plan, (c) a functional behavioral assessment in place, and (d) an augmentative or alternative communication (AAC) system for communicating (e.g., picture boards, pictures or symbols, objects, or voice output devices). Student characteristics information is presented in Table 2.

The researcher met with the administrator of the community program to identify students who met the inclusion criteria. Prior to the meeting, the administrator and researcher gathered information based on teacher report, student records including IEP’s, functional behavioral assessments, behavioral intervention plans, and classroom observations.

Setting

Paraeducators received training in an agency staff development classroom. The classroom measured 30 x 30 feet and was equipped with 10 tables and 40 chairs; a TV and VCR, a video camera, a dry erase board, and other training materials (e.g., videos, flip charts, training modules).

The study took place in each student’s classroom and school environment. Classrooms measured approximately 60 x 60 feet and were self-contained within elementary, middle, and high schools. Classes had six to eight students, one teacher, and one or two paraeducators. School environments included areas such as the lunchroom, media center, gym, and playground.

Experimental Design

A multiple probe across paraeducator design was used to evaluate the effectiveness of the intervention (training using a treatment package including self-evaluation). Paraeducators

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics of Staff</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Paraeducators</td>
</tr>
<tr>
<td>Paraeducator 1</td>
</tr>
<tr>
<td>Paraeducator 2</td>
</tr>
<tr>
<td>Paraeducator 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics of Students</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Student</td>
</tr>
<tr>
<td>Student 1</td>
</tr>
<tr>
<td>Student 2</td>
</tr>
<tr>
<td>Student 3</td>
</tr>
</tbody>
</table>
were initially assessed to determine baseline levels of behavior. Baseline continued until student responding was stable for at least three days. The first paraeducator was then trained using the treatment package including self-evaluation. Once a change in responding was noted for the first paraeducator, the second was trained; and similarly, the third began training once the second showed a change.

Procedure

Baseline. Student performance data were collected for three hours between the hours of 9:00 A.M. and 12:00 P.M. (noon) on each paraeducator for a minimum of three days to establish baseline performance. Observers utilized data collection instruments to record time, problem behavior, attempts to communicate using AAC, prompts, and responses. During baseline, nothing in the classroom was altered except that the observer was present in the classroom. Paraeducators knew that the observer was interested in communication, but were naïve about the intensities of the study and what behaviors were being observed. The fact that the paraeducators were not knowledgeable about the premise of the investigation helped control for the potentially confounding variable of reactivity to observer presence (e.g., Hay, Nelson, & Hay, 1980; White, 1977).

Intervention: Training on AAC and self-monitoring. Paraeducators were trained (a) on the importance of communication, (b) the relationship between behavior and communication, (c) the use of AAC, (d) how to prompt students to use AAC, (e) how and when to respond to students’ communicative attempts, and (f) how to self-evaluate their own behavior. One two hour session was held with each paraeducator in which the components of behavior, communication, prompting, and self-evaluation were taught.

During the first half of the training, paraeducators were provided with a summary of intervention research on communication and its importance for students who have significant disabilities (e.g., the impact that communication has on students’ lives including outcomes of holding meaningful jobs in the community, taking part in leisure events and activities with the general public, and living as independently as possible (Sack & McLean, 1997). The relationship between behavior and communication was discussed (Carr & Durand, 1985).

During the second half of the training, paraeducators were introduced to AAC. Training focused on the definition of AAC and the types of AAC that students use to communicate (picture boards, pictures, symbols, voice output devices, or other assistive communication devices).

Through role-playing, paraeducators learned how to prompt students to use AAC systems (e.g., verbal, gestures, and physical prompts). Paraeducators were instructed to prompt student use of AAC immediately when the student attempted to communicate. The researcher modeled appropriate prompting strategies including (a) verbal prompts, (b) gestures, and (c) physical prompts. Prompts were defined as assisting the student to communicate with AAC using least to most prompting (verbal, gestures, physical). Paraeducators were instructed to prompt students to use AAC each time that student attempted to communicate using AAC or to communicate through problem behavior (e.g., hitting, biting, and yelling). Paraeducators role played prompting the use of AAC and were given feedback until they were able to achieve 100% accuracy. The training lasted for three hours.

Paraeducators were trained through modeling and role-playing to respond to student attempts to communicate. Examples of AAC systems (e.g., big mac switches, voice output devices, communication boards) were used to demonstrate how students communicate using the systems. The researcher modeled how to respond immediately to requests. Responses were defined as reacting within three seconds to student requests in a manner consistent with the communicative intent of the request (i.e., presenting student with requested object or activity). Paraeducators role-played responses to requests until 100% accuracy was achieved. The role-playing lasted for three hours.

Paraeducators were then trained on how to self-evaluate their own behavior. They were instructed to count the number of times they prompted or responded to communication and to record this information on a data
sheet. The researcher checked for staff mastery of the procedure through use of a video exercise. The paraeducators viewed videos of themselves taped during baseline. They were instructed to record the number of times they observed themselves prompting student use of AAC or responding to student requests. Through the use of videotape, this data was compared to the data collected by the observer during baseline. A criterion of 100% accuracy was required for mastery. Paraeducators were retrained and repeated the video exercise until criterion was met. The sessions lasted for one hour and all paraeducators achieved 100% accuracy after repeating the exercise one time.

Paraeducators then implemented the training in their classrooms. They prompted students to use their AAC system each time they attempted to communicate. Prompts were defined as assisting the student to communicate with AAC using least to most prompting (verbal, gestures, physical). The paraeducators recorded each time they prompted students to use AAC on a data sheet.

Paraeducators then responded immediately to student requests. Responses were defined as reacting within three seconds to student requests in a manner consistent with the communicative intent of the request (presenting student with requested object or activity). The paraeducators recorded each time they responded to student requests on a data sheet.

Interobserver Agreement

Interobserver agreement was recorded based on the total number of occurrences of the behaviors. The mean interobserver agreement for paraeducator prompting was 92% (88% - 100%). The mean interobserver agreement for paraeducator responding was 96% (88% - 100%). The mean interobserver agreement for student use of AAC was 100% (100% - 100%). The mean interobserver agreement for student problem behavior was 91% (86% - 100%). Reliability on the researcher’s data sheets and the observer’s data sheets was calculated by recording the number of agreements and disagreements, then multiplying by 100% (Cooper, Heron & Heward, 1987).

Procedural fidelity was determined by comparing the paraeducators’ self-evaluation data sheets to the observers’ data sheets. The mean score for paraeducator 1 for prompting was 92% (88% - 100%). The mean score for paraeducator 1 for responding was 97% (92% - 100%). The mean score for paraeducator 2 for prompting was 93% (90% - 100%). The mean score for paraeducator 2 for responding was 91% (88% - 100%). The mean score for paraeducator 3 for prompting was 94% (90% - 100%). The mean score for paraeducator 3 for responding was 94% (90% - 100%).

Results

Data are presented for both paraeducator behavior (prompting student use of AAC and response to student requests) and student behavior (AAC use and problem behavior) for three conditions: baseline, intervention, and maintenance. As can be see in Figures 1 and 2, paraeducators behavior is recorded and analyzed via a multiple probe across participants design. Similarly, student behavior is presented in Figures 3 and 4.

Paraeducator Behavior

Prompt student use of AAC. The data collected in the baseline condition showed that Paraeducator 1 did not prompt the student to use AAC. After receiving the training, Paraeducator 1 immediately began prompting the student to use AAC. During baseline the mean number of prompts for student use of AAC was 0. After the intervention was implemented on day 4, the mean number of prompts for student use of AAC was 11 with a range of 7-16 prompts. During maintenance, Paraeducator 1 prompted student use of AAC 15 times on day 25 and 16 times on day 29.

Data collected in the baseline condition showed that Paraeducator 2 prompted the student to use AAC two or less times per session. After receiving the training, Paraeducator 2 immediately increased the number of prompts per session. During baseline the mean score for prompting student use of AAC was one time per session with a range of 0-2 prompts. After training the mean score for prompting student use of AAC was 13 with a range of 10-17 prompts. During maintenance, Paraeducator 2 prompted student use of AAC 16 times on day 29.
Data collected in the baseline condition showed that Paraeducator 3 did not prompt the student to use AAC. After receiving the training, Paraeducator 3 immediately began prompting the student to use AAC. During baseline the mean score for prompting student use of AAC was 0. After training the mean score for prompting student use of AAC was 7 with a range of 4-14 prompts. During maintenance, Paraeducator 3 prompted student use of AAC 14 times on day 29 and 15 times on day 31 (see Figure 1).

**Respond to student requests.** Data collected in the baseline condition showed that Paraeducator 1 did not respond to student requests. After the introduction of the intervention, Paraeducator 1 immediately began responding to student requests. During baseline the mean score for responding to student requests was 0. After the intervention was introduced on day 4, the mean score for responding to student requests was 10 with a range of 3-16 responses. During maintenance, Paraeducator 1 responded to student requests 15 times on day 25 and 15 times on day 29.

Data collected in the baseline condition showed that Paraeducator 2 did not respond to student requests. After introduction of the intervention on day 9, Paraeducator 2 immediately began responding to student requests. During baseline the mean score for responding to student requests was 0. After training the mean score for responding to student requests was 10 with a range of 5-15 responses. During maintenance, Paraeducator 2 responded to student requests 17 times on day 29.

Data collected in the baseline condition showed that Paraeducator 3 responded to student requests less than 1 time per session. After receiving the training, Paraeducator 3 immediately began responding to student requests. During baseline the mean score for responding to student requests was .2 with a range of 0-1 responses. After training the mean score for responding to student requests was 9 with a range of 5-16 responses. During maintenance, Paraeducator 3 responded to student requests 15 times on day 29 and 16 times on day 31 (see Figure 2).

**Student Behavior**

**Use of AAC.** Data collected in the baseline condition showed that Student 1 did not attempt to use AAC independently. After the paraeducator received training, Student 1 did attempt to use AAC during several sessions. During baseline the mean score for use of AAC was 0. After training the mean score for use of AAC was .5 with a range of 0-2 attempts. During maintenance, Student 1 attempted to use AAC 2 times on day 25 and 2 times on day 29 (see Figure 3).

Data collected in the baseline condition showed that Student 2 attempted to use AAC less than 1 time per session. After the paraeducator received training, Student 2 increased attempts to use AAC. During baseline the mean score for use of AAC was 0. After training the mean score for use of AAC was 4.5 with a range of 2-7 attempts. During maintenance, Student 2 attempted to use AAC 8 times on day 29.

Data collected in the baseline condition showed that Student 3 did not attempt to use AAC independently. After the paraeducator received training, Student 3 began to use AAC. During baseline the mean score for use of AAC was 0 with a range of 0-0 attempts. After training the mean score for use of AAC was 5 with a range of 4-9 attempts. During maintenance, Student 3 attempted to use AAC 6 times on day 29 and 8 times on day 31 (see Figure 3).

**Problem behavior.** Data collected in the baseline condition showed that Student 1 exhibited a high level of problem behaviors. After the paraeducator received training occurrences of problem behavior decreased. During baseline the mean score for problem behavior was 10 with a range of 8-11 behaviors. After training the mean score for problem behavior was 2 with a range of 0-6 behaviors. During maintenance, Student 1 exhibited 3 problem behaviors on day 25 and 1 problem behavior on day 29. The data collected in the baseline condition showed that Student 2 exhibited a high level of problem behaviors. After the paraeducator received training occurrences of problem behavior decreased. During baseline the mean score for problem behavior was 11 with a range of 8-16 behaviors. After training the mean score for problem
behavior was 5 with a range of 3-9 behaviors. During maintenance, Student 2 exhibited four problem behaviors on day 29 (see Figure 4).

Data collected in the baseline condition showed that Student 3 exhibited a moderate level of problem behaviors. After the paraeducator received training occurrences of problem behavior decreased. During baseline the mean score for problem behavior was 4 with a range of 3-4 behaviors. After training the mean score for problem behavior was 1 with a
Figure 2. Number of times paraeducators responded to students requests across paraeducators 1, 2, and 3.
range of 0-4 behaviors. During maintenance, Student 3 exhibited 2 problem behaviors on day 29 and 1 problem behavior on day 31 (see Figure 4).

**Discussion**

The overarching outcomes of this investigation indicated that all three paraeducators in-
increased the number of times that they prompted the use of augmentative communication systems and the number of times they responded to student requests. The primary finding was that the training of paraeducators using a treatment package including self-evaluation increased the number of times they prompted students to use AAC and the number of times they responded to student requests. All three paraeducators demonstrated immediate and dramatic change in behavior after receiving training. The paraeducators went from prompting the use of augmentative and alternative communication (AAC) two or
less times during baseline to prompting the use of AAC an average of nine times per session. Responses to student requests increased from less than one time to an average of nine times per session.

Augmentative and alternative communication has become a method of choice to facilitate language and to address challenging behavior (Mirenda & Erickson, 2000). Variables related to the efficiency of AAC include: (a) time to introduce and get different parties up and going with AAC system, (b) time that lapses before an individual is able to use AAC to accomplish predetermined needs and purposes, (c) extent to which use of an AAC system complements other methods used by the individual to communicate, and (d) time necessary to solicit contingent responses from partners with and without AAC (Calculator, 1999).

Johnston et al. (2004) identified the need for methods to train educators to promote the use of AAC in the classroom. They suggested that students have AAC systems but are not using them and that their communicative partners are not using them. They also suggested that students use alternative, but socially inappropriate, strategies for communication. The present study strengthened this finding in the following ways. Although all three students who participated in this study had AAC systems in place, the paraeducators did not consistently prompt students to use the systems prior to intervention. Without this prompting, students seldom used AAC devices and relied instead on problem behaviors in order to communicate. Further, paraeducators responded more often to student requests when the students communicated through problem behaviors than when students attempted to use AAC to make requests.

Paraeducator Training

Several prior studies demonstrated the effectiveness of paraeducator training to support students with severe communication disabilities (e.g., Light et al., 1992; McNaughton, & Light, 1989; Payne & Ogletree, 1995). While other investigations have reported positive changes in staff behavior following training (Light et al.; McNaughton & Light; Payne & Ogletree), this study shows that training using self-evaluation resulted in immediate behavior change. After receiving training that included verbal instruction, modeling, role playing, and video examples, all three paraeducators demonstrated dramatic changes in prompting and responding behavior in the first probe after intervention.

Self-Evaluation

The present study used self-evaluation as a strategy to improve paraeducator performance. Self-evaluation has proven to be an effective strategy (Allinder, 2000; Belfiore & Browder, 1992; Browder, Liberty, Heller, & D’Huyvetters, 1986). Although the treatment package contained several components that could have contributed to changes in paraeducator behavior (e.g., instruction on communication, students with significant disabilities, problem behaviors, augmentative and alternative communication (AAC); demonstration of AAC devices; modeling prompting and responding behaviors), the use of self-evaluation strengthened the overall outcomes of the study. When paraeducators were instructed to monitor their behavior (number of times they prompted student use of AAC and responded to student requests) these behaviors increased dramatically.

This current study makes several contributions to the literature in the areas of self-evaluation and paraeducator training. First, the present study contributes to the literature by extending research in the area of staff self-evaluation. Although the treatment package contained several components that could have contributed to changes in paraeducator behavior, the use of self-evaluation strengthened the findings. When paraeducators were instructed to monitor their behavior (number of times they prompted student use of AAC and responded to student requests) these behaviors increased dramatically.

Second, this study contributes to the literature by showing that training paraeducators using a treatment package including self-evaluation resulted in increases in their prompting students to use AAC and responding to student requests. Paraeducator behavior resulted in increases in student use of AAC and decreases in problem behaviors.

Third, the study contributes to the litera-
ture base on the training of paraeducators by demonstrating that when paraeducators are provided explicit training on key instructional behaviors aligned with actual responsibilities, major gains in paraeducator performance and related gains in student performance are attainable.

Student Behavior Change

The outcomes suggest that two students increased the number of times they attempted to use AAC to communicate. The third student attempted to use AAC one time during intervention. Student problem behaviors decreased in all three cases. The effectiveness of teaching functionally equivalent communicative behaviors to replace problem behaviors has been documented throughout the literature (Bird et al., 1989; Carr & Durand, 1985; Durand, 1999; Durand & Carr, 1991, 1992; Gerra & Dorfman, 1995; Horner & Budd, 1985; Lalli, Browder, & Mace, 1993; Wacker et al., 1990). The present study is similar to prior research in that the use of AAC resulted in fewer occurrences of problem behavior.

In conclusion, this research demonstrated that training paraeducators using a treatment package including self-evaluation resulted in increased prompting of student use of augmentative communication devices and increased responding to student attempts to communicate. This is important because effective training programs for paraeducators who support students with severe disabilities who use AAC are limited. If immediate and dramatic change in behavior can be achieved through training paraeducators in a one to two hour training session as demonstrated in this investigation, additional training may result in significant improvements in practice. Functional communication training using assistive technology is an effective strategy for reducing serious problem behaviors with students who have significant disabilities. Future research is needed to assist in understanding and overcoming the barriers to effectively training paraeducators in supporting students with severe disabilities.

References


Received: 10 April 2006
Initial Acceptance: 30 May 2006
Final Acceptance: 25 August 2006
Attitudes of Preservice Teachers Enrolled in an Infusion Preparation Program regarding Planning and Accommodations for Included Students with Mental Retardation

David L. Cameron
Agder University College

Bryan G. Cook
University of Hawaii

Abstract: The beliefs, skills, and intended practices of general (n = 34) and special education (n = 23) preservice teachers regarding planning and making accommodations for included students with mental retardation were investigated. Participants were drawn from teacher preparation programs that infused content related to inclusion into pre-existing courses. Results of a repeated-measures ANOVA indicated a significant main effect of teacher type—that preservice special educators rated their beliefs, skills, and intended practices significantly higher than their general education counterparts. A significant main effect of attitudinal category showed that participants also rated their beliefs and intended practices significantly higher than their skills. A significant interaction effect was evidenced due to the varying discrepancies between general and special educators’ ratings of their beliefs, skills, and intended practices. Specifically, general educators’ ratings were further below those of special educators in skills and closest in the area of beliefs. Findings suggest that this infusion teacher preparation program was more effective at generating positive beliefs and intentions than skills, especially among general educators. Implications for teacher preparation and practice are discussed.

The diversity of student characteristics in general education classrooms in the United States has consistently increased in recent years. In particular, the prevalence of inclusive placements for students with mental retardation has been on the rise. For example, the 10 year period from 1990 to 2000 represents an increase from approximately 41,000 to 86,000 children with this disability having spent greater than 79% of the school day in general education settings (U.S. Department of Education, 2002). During this same period the number of children with mental retardation who spent between 40% and 79% of the day in inclusive settings increased by nearly 43,000 students. In contrast, the number of children with mental retardation who were placed primarily in separate classrooms and facilities (greater than 60% of the day) decreased by approximately 40,000 from 1990 to 2000 (U.S. Department of Education). These data clearly demonstrate the importance of examining teachers’ preparedness for addressing the needs of children with disabilities in general education settings.

A great deal of the extant research on inclusion has focused on general educators’ attitudes toward the general concept of inclusion (see Scruggs & Mastropieri, 1996 for a review of this literature). However, teachers’ attitudes toward inclusion appear to be a multifaceted construct. For example, teachers may believe in a practice associated with effective inclusive instruction, but not actually implement it. Schumm, Vaughn, Gordon, and Rothlein (1994) conducted a study related to this issue that investigated general educators’ beliefs, skills, and practices regarding plan-
ning and adapting their instruction for included students with learning disabilities. Their findings indicated that inclusive teachers reported themselves to practice adaptations and planning techniques significantly less than the degree to which they believed and were skilled in the approaches. As such, it is important to investigate these aspects of teachers’ inclusive attitudes (i.e., beliefs, skills, and practices) as distinct constructs.

That teachers believe in, are skilled in, and engage in appropriate planning and adaptations for students with mental retardation appears critical to the educational progress of the increasing numbers of students with this disability who receive their education, partially or wholly, in general education classrooms. It is widely accepted that students with mental retardation require their inclusive teachers to make adaptations and accommodations to typical planning and instruction to meet their unique learning needs (e.g., Westling & Fox, 2000). Practicing teachers have, however, reported a lack of competence and the need for more training in planning and making adaptations for included students with disabilities (Scruggs & Mastropieri, 1996; Wolery, Werts, Caldwell, Snyder, & Lisowski, 1995). One possible explanation for reports of teachers’ lack of skill in this regard is that their teacher preparation did not place sufficient emphasis on training or skills (Griffin, Jones, & Kilgore, in press).

Research has indicated that teachers’ attitudes can be influenced by their training experiences. For example, Leyser (1988) successfully increased positive teacher attitudes toward inclusion among 15 general education teachers through an in-service program emphasizing inclusive practices and concepts. Similarly, Hutchinson and Martin (1999) found that a nine-month field based course incorporating discussions of case studies increased understanding of equity issues and developed the expectations among participants that they would adapt their teaching for included students. A significant task for researchers, then, is to determine which approaches to training preservice teachers produce desired attitudes, skills, and practices by continuing to investigate the degree to which different types of preparation programs are associated with desired inclusive attitudes and aptitudes.

One approach for preparing general educators that has been instituted by many teacher preparation programs has been dubbed a curricular infusion model (see Akasmit, 1990; Cook, 2002; Strawderman & Lindsey, 1995; Voltz, 2003). The infusion approach involves presenting curricular content related to inclusion throughout an existing teacher education program rather than devoting an entire course to the topic. A primary advantage of the model appears to be its ease of implementation; it requires few significant changes to an existing program when new content is introduced (Strawderman & Lindsey). Another reported benefit of the infusion model is that special education issues are presented in unison with other important aspects of teaching, reinforcing the idea that children with disabilities should not be viewed as a separate concern to be dealt with in isolation (Voltz). However, some have expressed concern that curricular content is likely to be presented with inadequate consistency and depth in infused programs (Akasmit; Gay, 1997).

Although the bulk of the research base on inclusive planning and adaptations has focused on general educators (e.g., Schumm et al., 1994), special educators also play a prominent and critical role in planning and providing inclusive education (Fennick & Liddy, 2001). Indeed, a central focus of traditional teacher preparation in special education has been on planning instruction and making instructional adaptations for students with disabilities. Yet, it is only recently that the focus of these procedures has been the inclusive classroom. In order to prepare preservice special educators for inclusive teaching, the special education teacher education program in which we conducted the present study had also adopted an infusion model regarding inclusion. We speculate that many other special education teacher preparation programs have implemented a curricular infusion model for their inclusive content, for many of the same reasons that general education programs have done so. In particular, the interspersing of inclusion issues throughout an existing set of
courses introduces little disruption to the pre-existing program. Thus, it is important to conduct research on the impact of inclusive curricular infusion on preservice teachers in special, as well as general, education teacher preparation programs.

As the trend towards including students with mental retardation shows little sign of abating, the value of ensuring that new teachers are prepared for these contexts is clear. Despite the growing popularity of infusing content related to inclusion into existing teacher preparation coursework, little research has investigated the effectiveness of this approach. As both general and special education teachers are actively involved in planning and making accommodations for students with mental retardation in inclusive classrooms, we chose to study the attitudes (i.e., beliefs, skills, and intended practices) of preservice general and special education teachers who had been trained in infusion teacher preparation programs toward planning and accommodations for included students with mental retardation. As we were unable to identify any research examining this combination of variables, this research study will, we believe, be the first to explore this important set of issues.

**Method**

**Participants**

The study sample comprised 57 preservice general and special educators enrolled in separate 4-year teacher preparation programs at a large mid-western university. Participants were in their final semester of undergraduate teacher training. Two comparison groups were formed according to whether students would be entering the teaching field as general (n = 34) or special educators (n = 23). Although the specific licensures sought by participants varied within comparison groups (e.g., some special education majors sought licensure in mild/moderate disabilities, others in moderate/intensive disabilities; general education majors sought licensure in early childhood, physical, elementary, and secondary education), participants were easily identified and categorized as either general education or special education on the basis of their degree program and licensure track. Individuals pursuing dual-certification in general and special education concentrations, and/or previously certified teachers, were not included in the study. Demographic information regarding participants’ gender, ethnicity, age, and intended teaching area are presented in Table 1.

**Program Setting**

The special and general education teacher preparation programs existed as largely separate entities in the university setting in which data were collected. Both programs led to a bachelor’s degree and state teaching licensure after four years of training. Approximately one-third of program requirements for both groups consisted of coursework in the liberal arts and sciences. In addition, undergraduate education majors devoted at least one-fifth of their training to professional preparation through field experiences and student teaching. With respect to field experiences (including student teaching), placement in inclusive classrooms was encouraged but not mandated or guaranteed. The remaining time was allotted for mastery of the content or specialization area that they would be teaching.

Shared classroom experiences for general and special educators were typically limited to universal liberal arts and science requirements and introductory or foundations coursework (e.g., educational theory/psychology, language and literacy, child development). These courses were most often completed in the early stages of training. In contrast, the majority of programming pertaining to the student’s concentration area(s) occurred in the final two years of training, with increasing exposure each year. For example, by their senior year, preservice special educators were typically enrolled only in courses pertaining to special needs populations.

With regard to special education related content, the general education training program was consistent with previous descriptions of infusion models for teacher preparation (see Aksamit, 1990; Cook, 2002; Voltz, 2003). Students were not required to take separate courses in which special education or inclusion was a primary focus. Rather, inclusion
and special education issues were infused into a small number of introductory teacher preparation courses. Participating preservice general educators were enrolled in the final (fourth) year of a requisite series of cumulative seminar courses. It was within these seminar courses, in particular, that content related to special education and inclusive schooling was provided to general educators. Among other course topics infused throughout these seminars were learning theories, diversity awareness, technology use, and historical and current issues in schooling.

Preservice special education students were not required to take a specific course that was uniquely intended to address the issue of inclusion. However, the majority of courses within the special education concentration covered, to varying degrees, issues related to inclusion. The course that focused most extensively on strategies for adapting general education curriculum and teaching methods for inclusive classrooms was the required methods course (which was taught separately for those enrolled in mild/moderate and moderate/intensive licensure tracks). In addition, the special education participants had taken the first three courses in the required educational seminar series and were enrolled in the fourth seminar alongside their general education counterparts.

**Instrumentation**

Participants completed a 3-page survey containing demographic questions and a modified version of the Teachers’ Beliefs and Attitudes toward Planning for Mainstreamed Students instrument (TBAP) (Schumm et al., 1994). The modified instrument (P-TBAP) contains 30 statements intended to measure preservice teacher attitudes in three areas (attitudinal categories) with respect to 10 effective inclusive classroom practices. Definitions for the attitudinal categories are as follows:

**Beliefs** – participants’ ratings of their belief in the value of planning and making adaptations for included students with mental retardation.

**Skills** – participants’ ratings of their skill in planning and making adaptations for included students with mental retardation.

### Table 1

**Characteristics of Preservice General and Special Education Teachers**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>General Education</th>
<th>Special Education</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>P&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>25</td>
<td>73.53</td>
</tr>
<tr>
<td>Male</td>
<td>9</td>
<td>26.47</td>
</tr>
<tr>
<td>Ethnicity&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>33</td>
<td>97.06</td>
</tr>
<tr>
<td>Unspecified</td>
<td>1</td>
<td>2.94</td>
</tr>
<tr>
<td>Intended Teaching Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary</td>
<td>22</td>
<td>64.71</td>
</tr>
<tr>
<td>Intermediate</td>
<td>1</td>
<td>2.94</td>
</tr>
<tr>
<td>High School</td>
<td>4</td>
<td>11.76</td>
</tr>
<tr>
<td>K-12</td>
<td>7</td>
<td>20.59</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>23.76</td>
<td>5.91</td>
<td>23.48</td>
<td>3.98</td>
</tr>
<tr>
<td>Special Education/Inclusion Courses&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.54</td>
<td>1.71</td>
<td>10.96</td>
<td>4.45</td>
</tr>
</tbody>
</table>

*Note. *<sup>a</sup>P = percent of each comparison group. <sup>b</sup>Only ethnic categories reported by participants are listed in this table. <sup>c</sup>Classes taken in which special education or inclusion related issues were a major component.*
Intended Practices – participants’ ratings of their will (intention) to plan and make adaptations for included students with mental retardation when they become teachers.

Participants rated their level of agreement with each statement using a 4-point Likert-type scale (1 = low agreement, 4 = high agreement). Statements from the P-TBAP instrument are presented in Table 2. In their investigation of general education teachers’ attitudes, Schumm et al. (1994) reported that the TBAP demonstrated adequate reliability and validity. The P-TBAP differs from the original scale in two ways. First, the scale was adapted for preservice teachers to reflect intended rather than current practices. Second, each item was changed to refer to “included students with mental retardation,” rather than the broader category of “mainstreamed students.” Both the original and modified scales were shown to exhibit adequate internal reliability. Analyses of the P-TBAP revealed Cronbach coefficient alphas of 0.82, 0.90, 0.97, and 0.93 for Beliefs, Intended Practices, Skills, and the total scale, respectively.

In regard to validity, Schumm et al. (1994) generated items for the TBAP through a multi-stage procedure consisting of (a) a review of the literature on teacher planning, (b) a survey of teachers’ planning practices (Schumm & Vaughn, 1992), and (c) a focus group interview with teachers. In addition, the content validity of the scale was assessed through an analysis of 25 classroom teachers’ responses to open-ended questions pertaining to items on the instrument. Schumm et al. then coded, analyzed, and compared the teachers’ responses to potential instrument items. Although the P-TBAP differs slightly from the original instrument with respect to the population of students investigated, the planning strategies and adaptations depicted on the scale are consistent with the literature on effective inclusive practices for students with mental retardation (e.g., Westling & Fox, 2000). Thus, the modified instrument appears to be valid for this population as well.

Procedure

The entire survey, comprising the P-TBAP instrument and demographic questions, takes approximately 15 minutes to complete. All surveys were administered by one of the authors over six sessions during the initial 20 minutes of university courses in which participants were enrolled. Directions for completion of the survey were scripted to ensure uniformity of administration. During each phase of data collection students were read a general description of the study and instructed that participation was voluntary and anonymous. A total of 72 students completed the survey (response rate = 96%), of which 57 (79% of those surveyed) met the criteria for participation. The majority of individuals who completed the survey but were not included in the study were graduate students in special education classes that were attended by both undergraduate and master’s level students. Three sessions of data collection were conducted for each comparison group during courses for which the majority of students were in their final year of training. Specific classes were randomly selected for participation from identified clusters of courses in which the target groups were enrolled. The researchers were not involved in instruction for any of the courses in which data were collected.

Research Design and Analysis

The investigation employed a causal comparative research design to evaluate teachers’ ratings of their beliefs, intended practices, and skills regarding planning and making accommodations for included students with mental retardation. Mean ratings of the 10 effective practices for each attitudinal category (i.e., beliefs, intended practices, and skills) served as the unit of analysis in a repeated-measures ANOVA in which attitudinal category was the within-subjects factor and teacher type (general and special education preservice teachers) was the between-subjects factor. If the main effect of attitudinal category or the teacher-type by attitudinal category interaction effect were statistically significant, within-subject contrasts were conducted to examine between which levels of the variables the dif-
TABLE 2

P-TBAP Statements Rated by Participants

1. I BELIEVE that information sources such as the student’s IEP and psychological reports are useful in planning for included students with mental retardation.
2. I BELIEVE that long-range instructional planning (e.g., yearly, semester, monthly) needs to be adapted for included students with mental retardation.
3. I BELIEVE that short-range planning (e.g., daily, weekly) needs to be adapted for included students with mental retardation.
4. I BELIEVE that varying group composition (e.g., small group, large group, whole class) is important for included students with mental retardation.
5. I BELIEVE that course content needs to be adapted for included students with mental retardation.
6. I BELIEVE that the time and the pace of a lesson must be adjusted to meet the needs of included students with mental retardation.
7. I BELIEVE that the tests I design for my general education students need to be adapted for included students with mental retardation.
8. I BELIEVE that frequent checks with individual students are an effective way to monitor the progress of included students with mental retardation.
9. I BELIEVE that providing individualized instruction according to the students’ needs helps included students with mental retardation.
10. I BELIEVE that using individualized (or different) criteria for grading assignments and tests is important when evaluating included students with mental retardation.
11. As a teacher, I WILL use information sources such as the student’s IEP to plan for included students with mental retardation in my classroom.
12. As a teacher, I WILL adapt my long-range instructional planning for included students with mental retardation.
13. As a teacher, I WILL adapt daily planning for included students with mental retardation.
14. As a teacher, I WILL vary group composition for included students with mental retardation.
15. As a teacher, I WILL adapt course content for included students with mental retardation.
16. As a teacher, I WILL adjust the time and pace of a lesson for included students with mental retardation.
17. As a teacher, I WILL adapt tests for my included students with mental retardation.
18. As a teacher, I WILL use frequent checks with the individual students to monitor the progress of included students with mental retardation.
19. As a teacher, I WILL provide individualized instruction for included students with mental retardation.
20. As a teacher, I WILL use individualized/different criteria when evaluating the assignments and tests of included students with mental retardation.
21. I am SKILLED at using a variety of information sources (e.g., IEPs, parents, student feedback) to assist me in planning for my included students with mental retardation.
22. I am SKILLED at designing long-range plans that meet the needs of my included students with mental retardation.
23. I am SKILLED at appropriately pacing and timing the presentation of content material for included students with mental retardation.
24. I am SKILLED at grouping for instruction so that the needs of included students with mental retardation may be met.
25. I am SKILLED at designing tests that effectively monitor progress of included students with mental retardation.
26. I am SKILLED at using individualized/different criteria when evaluating the assignments and tests of students with mental retardation.
27. I am SKILLED at designing short-range plans that meet the needs of included students with mental retardation.
28. I am SKILLED at adapting course content to meet the needs of included students with mental retardation.
29. I am SKILLED at using frequent checks to monitor the progress of included students with mental retardation.
30. I am SKILLED at providing individualized instruction for included students with mental retardation.
ferences or interactions existed. A conventional alpha level of .05 was used to determine statistical significance for all analyses.

Results

Means and standard deviations of attitudinal category ratings for comparison groups are presented in Table 3. A repeated measures ANOVA was performed with attitudinal category as the within subjects factor and teacher type as the between subjects factor (special and general education preservice teachers in their final year of training). A significant main effect of teacher type was found, $F(1, 55) = 33.34, p < .001, \eta^2 = .38$, where future special educators rated their beliefs, skills, and intended practices higher than those of preservice general educators. A significant main effect of attitudinal category was also found, $F(2, 110) = 96.79, p < .001, \eta^2 = .64$, indicating that significant differences existed in participants’ self-ratings regarding their beliefs, intended practices, and skills related to effective planning and accommodations for included students with mental retardation. Within-subjects contrasts indicated that skills were rated significantly lower than both beliefs, $F(1, 55) = 118.00, p < .001$, and intended practices, $F(1, 55) = 93.79, p < .001$. No significant differences were found in preservice educators’ beliefs and intended practices ratings, $F(1, 55) = .004, p = .95$. A significant attitudinal category-by-teacher type interaction was also found, $F(2, 110) = 6.96, p = .001, \eta^2 = .11$. Within subjects contrasts indicated that significant teacher type-by-attitudinal category interactions existed between beliefs and intended practices ($F(1, 55) = 4.56, p = .04$), beliefs and skills ($F(1, 55) = 10.59, p = .002$), and intended practices and skills ($F(1, 55) = 4.40, p = .04$). These significant interactions can be attributed to the variability in differences between general and special educators across the attitudinal categories. That is, general educators’ mean ratings related to beliefs were 0.39 lower than special educators’, 0.58 lower for ratings of intended practice, and 1.09 lower than special educators in regard to skill ratings.

Discussion

Across all attitudinal categories (i.e., beliefs, skills, and intended practices), future special educators’ ratings regarding planning and accommodations for included students with mental retardation were higher than those of preservice general educators. Ratings for all participants were significantly higher for beliefs and intended practices than skills. Furthermore, the disparity between special and general education preservice teachers varied significantly across attitudinal categories. It should be noted that participants’ ratings were generally positive—above the theoretical neutral rating of 2.5 in all but one instance. The one exception was that general educators reported themselves to be, on the whole, less than moderately skilled related to survey items. Interpretations and implications of these findings as well as limitations and suggestions for future research are discussed below.

Explanation and Interpretation of Findings

Findings indicate that participants, future general and special educators nearing completion of teacher preparation programs that infused content related to inclusion, believe planning and instruction should be adapted

---

**Table 3**

<table>
<thead>
<tr>
<th>Attitudes</th>
<th>General Educators</th>
<th>Special Educators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beliefs</td>
<td>Mean 3.25, SD 0.54</td>
<td>Mean 3.64, SD 0.34</td>
</tr>
<tr>
<td>Intended Practices</td>
<td>Mean 3.15, SD 0.63</td>
<td>Mean 3.73, SD 0.36</td>
</tr>
<tr>
<td>Skills</td>
<td>Mean 1.72, SD 0.85</td>
<td>Mean 2.81, SD 0.66</td>
</tr>
</tbody>
</table>

---
for students with mental retardation, intend on practicing these adaptations, yet do not feel commensurately skilled in using these strategies. The positive beliefs of participants regarding the importance of effectively planning and adapting their instruction for included students with mental retardation may be an outcome of a larger trend among educators in recognizing the benefits of inclusion as inclusive placements become more commonplace (Scruggs & Mastropieri, 1996). An alternative explanation for participants’ positive beliefs may be their lack of professional experience. Given that these undergraduate students have not yet experienced the demands of working professionally in heterogeneous classrooms first hand, it may be easier for them to maintain positive beliefs about the value of inclusive practices. In fact, some evidence suggests a negative correlation between years of teaching and support for inclusion (Soodak, Podell, & Lehman, 1998). Another possible reason for preservice teachers’ positive beliefs and intentions is that the infused content related to inclusion that they received in both teacher preparation programs was sufficient to make participants feel that engaging in effective inclusive practices was a good idea and something that they intended to do.

However, it appears that the infused training did not provide these future teachers with skills commensurate with their beliefs and intentions. It may be that infusing content related to inclusion does not allow sufficient time for the intensive training needed to develop inclusive teaching skills. It is interesting to note that special education preservice teachers—who were enrolled in a program that infused content related to inclusion more pervasively than the program in which general education participants were involved (general educators reported taking 1.5 courses on average in which inclusion or special education content was a major focus, as opposed to approximately 11 courses for special educators)—reported relatively higher skills (although still not as high as their reported beliefs or intentions). Although participants’ relatively low level of skills may be due to their infused teacher preparation, it is also possible that teacher preparation in general does a better job of developing beliefs and intentions than producing actual skills; or that participants felt more skilled in planning and adapting instruction for students with disabilities other than mental retardation (e.g., learning disabilities).

Findings of this study may appear somewhat inconsistent with those of Schumm et al. (1994) that practicing teachers believed in the value of planning and making adaptations, perceived themselves as skilled in these areas, but confessed that they did not practice these adaptations. Although preservice teachers in the present study held similarly positive beliefs, they reported possessing lower levels of skills, but, nonetheless, intended to implement these practices when they became teachers. These discrepant findings may be due to the lack of relevant experience of preservice teachers. Teachers who have worked in inclusive classrooms may simply be more aware of the barriers (such as lack of resources, time, and support) encountered in adapting instruction and planning for included students. Intended practices instead appear to be a function of beliefs for preservice teachers, unmediated by their perceived skills. In other words, it may be difficult for these future teachers to separate the idea that something is the right thing to do from the reality that they do not have the requisite skills to implement the techniques.

Prospective general educators’ ratings across categories were pervasively lower than those of special educators. One logical interpretation of these between group differences is that they are a function of the teacher preparation programs in which participants were enrolled. Although both programs infused content related to inclusion, the special education program did so in a prominent manner in many courses. Inclusion was a central topic in most, if not all, special education courses. For example, the introduction to exceptionalties course, the course on professional collaboration, and the special education instructional methods courses all had multiple class sessions specifically devoted to inclusion. In contrast, although instructors were encouraged to relate course content to inclusion whenever possible in the general education preparation program, inclusion was a prominent focus in only one seminar course. It is not surprising, then, that Cook (2002) reported that many preservice general educa-
tion teachers previously enrolled in this same program perceived that the amount of content and time dedicated to inclusion and students with disabilities was insufficient. It is also possible that differences other than the degree to which inclusion content was infused in the teacher preparation programs contributed to the attitudinal discrepancies between preservice general and special educators. For example, special educators undoubtedly were exposed to more coursework about and involved in more field work with students with mental retardation, which could have impacted their relevant beliefs, intentions, and skills toward these students. Preservice special educators may also have had experiences and developed dispositions before entering their teacher preparation program that predisposed them to obtain high beliefs and skills related to the inclusion of students with mental retardation.

In regard to the significant interaction effect, in which differences in ratings between general and special education preservice teachers varied across the three attitudinal categories, it is noteworthy that the gap between the two groups widens as the prompts become more practical or concrete. That is, the ratings provided by the two groups of participants were most similar on beliefs—a construct that reflects participants’ convictions and is not bound by what can actually be accomplished. The gap in ratings expanded when participants rated their intended practices, which asks them to speculate about what will happen in the real world; but in a future that participants may have idealized or not known enough about to realistically appraise. The most concrete category of questions, regarding their actual skills, is the area in which the two groups are the most discrepant. We speculate that the positive relation between rating disparities and the concreteness of the attitudinal category may have been due to (a) the differences in the degree to which inclusion was infused in the two programs (i.e., greater infused content results in greater concrete skills, whereas even a minimal amount of infused content may be sufficient for producing positive beliefs), (b) the general and special education nature of the two programs (i.e., perhaps the focus on individualizing instruction in special education teacher preparation, regardless of the degree to which inclusion content is infused in the curriculum, is responsible for this interaction effect), or (c) both.

Implications and Recommendations

That participating preservice teachers appear to hold positive beliefs and intentions regarding appropriate planning and instructional adaptations for included students with mental retardation is an undeniably positive finding and suggests that infused teacher preparation programs can accomplish these important goals for both general and special educators. Believing that these practices are important and intending to implement them likely predisposes teachers to actually practicing the techniques and overcoming the barriers often encountered in applying effective practices. However, positive beliefs and intentions are not sufficient for the effective application of effective instructional practices. Teachers certainly have to be skilled in a technique to use it correctly. Given their relatively low skills ratings, it seems highly unlikely that these teachers will begin their careers using effective planning and instructional adaptations for included students with mental retardation, regardless of their beliefs and intentions. Teachers not engaging in appropriate planning and instructional adaptations have obvious negative implications for the educational opportunities and outcomes of the increasing number of students with mental retardation included in general education classes. The lack of skills appears especially problematic for general educators, who had rated their skills much lower than special education participants and often find themselves having primary or even sole responsibility for instructing included students with disabilities.

Although this study does not address what changes in teacher education programs might result in higher skill levels, it is interesting to note that special education teachers reported skill levels that were above a theoretical neutral rating of 2.5 and were markedly higher than those of their general education counterparts. So, it seems that infused teacher preparation programs can be somewhat successful in engendering adequate skills when a great deal of content related to inclusion is
presented throughout the curriculum and when the infused content is integrally tied to the focus of the program (as inclusion is to special education). As such, if it is to produce teachers who are at least moderately skilled in practicing what they believe in and intend to do regarding the inclusion of students with mental retardation, we speculate that the general education teacher preparation program may need to (a) increase the frequency and intensity that content related to inclusion is presented and (b) communicate that being skilled in inclusion is a fundamental aspect of being a general education teacher. Alternatively, neither program produced future teachers with skills equal to their beliefs and intentions. It may be that, as opposed to infused content, specific coursework devoted to inclusion and linked to high quality field experiences is necessary to generate high levels of inclusive teaching skills such as appropriate planning and instructional adaptations.

Limitations and Suggestions for Future Research

The self-reported nature of participants’ ratings poses an important limitation to these findings. It is possible that participants did not accurately perceive their own skills, for example, or that they may have reported what they deemed to be socially desirable responses. The limited nature of the sample poses another noteworthy limitation. Because participants were drawn from only two teacher preparation programs at a single university, results cannot be generalized to other teacher preparation programs that infuse content related to inclusion in pre-existing courses. Results indicated that preservice special educators rated their beliefs, skills, and intended practices significantly higher than general educators; across groups, participants rated their beliefs and intended practices significantly higher than their skills; and discrepancies varied significantly between general and special educators’ ratings of their beliefs, skills, and intended practices—with skills being the category on which the two groups differed most. Important limitations notwithstanding, we conjecture that infusing content related to inclusion into pre-existing coursework, rather than devoting one or more entire courses to the topic, may preclude most prospective teachers from developing high levels of skills in this area. Conversely, infused instruction, even when occurring infrequently (i.e., in one seminar course), appears to be sufficient to engender positive beliefs and intentions.

Future research could compare the degree to which these findings hold across other infused teacher preparation programs, and if the degree to which inclusion content is infused systematically varies with future teachers’ beliefs, intentions, and skills. Future research should also seek to broaden the sample to include participants from more diverse backgrounds. In addition, future researchers could compare educators’ beliefs, practices, and skills regarding planning and adapting instruction for students with different disabilities. Ultimately, experimental research is necessary to determine what approach to teacher preparation produces teachers with the highest beliefs, intended practices, and skills.

Conclusion

Fifty-seven future general and special education teachers rated their beliefs, intentions, and skills related to planning and making accommodations for included students with mental retardation. Participants were nearing completion of separate teacher preparation programs that infused content related to inclusion into pre-existing courses. Results indicated that preservice special educators rated their beliefs, skills, and intended practices significantly higher than general educators; across groups, participants rated their beliefs and intended practices significantly higher than their skills; and discrepancies varied significantly between general and special educators’ ratings of their beliefs, skills, and intended practices—with skills being the category on which the two groups differed most. Important limitations notwithstanding, we conjecture that infusing content related to inclusion into pre-existing coursework, rather than devoting one or more entire courses to the topic, may preclude most prospective teachers from developing high levels of skills in this area. Conversely, infused instruction, even when occurring infrequently (i.e., in one seminar course), appears to be sufficient to engender positive beliefs and intentions.

References


Preservice Teacher Attitudes / 363


Received: 10 April 2006
Initial Acceptance: 22 May 2006
Final Acceptance: 1 September 2006
Use of Adapted Bicycles on the Learning of Conventional Cycling by Children with Mental Retardation

Tammy L. Burt
Southern Illinois University Edwardsville

David L. Porretta
The Ohio State University

Richard E. Klein
University of Illinois

Abstract: This study investigated the use of adapted bicycles on the acquisition, maintenance, and generalization of conventional cycling by seven children with mild mental retardation. Feedback was used in addition to the adapted bicycles and consisted of pedal rate, head position, and steering participation. A multiple probe design was used. Participants were required to ride as far as possible for each trial. Results indicated that 100% of participants demonstrated acquisition of conventional cycling. Maintenance was demonstrated by 71.4% of participants, and generalization was demonstrated by 42.9% of participants. Results are interpreted from an ecological perspective.

Considerable attention has been given over at least the past two decades on the learning of motor skills by individuals with mental retardation (MR) (Bouffard, 1990; Hoover & Wade, 1985). It is generally known that individuals with MR can typically acquire simple motor skills (e.g., Gillespie, 2002, 2003; Yang & Porretta, 1999). However, these same individuals have difficulty learning complex skills (Porretta, 1990; Reid, 1993). In motor performance tests of eye/hand coordination, balance, and body coordination, Eichstaedt, Wang, Polacek, & Dohrmann, (1991) found that children with mild or moderate MR scored significantly lower than children without MR. Due to deficits in balance and coordination, children with MR experience difficulty in learning and performing motor tasks, especially when task complexity increases (Wade, Newell, & Wallace, 1978) and the task requires a temporal response (Kail, 1992; Newell, Wade, & Kelly, 1979).

Historically motor learning research on persons with MR lacked the use of ecologically appropriate tasks (Reid, 1993). However, more recently, research on the learning of motor skills has focused on tasks that have direct relevance to daily life (Davis & Burton, 1991; Sherrill, 2004). Bicycling is such a task. It allows individuals the opportunity for independent travel or travel with family or friends. However, many individuals, especially those with disabilities, have never learned to ride a conventional bicycle (Joules, 1996). This is especially true for persons with mental retardation (MR). It is difficult to pinpoint why there is such a low number of cycling participants compared to other sports opportunities. It could be attributed to the lack of equipment, finances, or programming. Special Olympics International offers bicycling as a competitive event, however few athletes participate. Statistics from Ohio Special Olympics indicate that only 313 individuals with MR compete in cycling events compared to 7,469 competitors in track & field events and 8,448 competitors in bowling events (M.S. Allen, personal communication, December, 2005). These figures can be considered representative of participants in other states. However, perhaps a greater number of individuals with...
MR might actively participate in cycling programs if learning to ride were not such a complicated motor task. Riding a conventional bicycle requires appropriate levels of eye-hand coordination, balance, and response time.

A variety of training aids and devices have been developed and patented over the years in an attempt to simplify the learning of cycling (e.g., Egley, 1994; Harrison, 1995; Henrichs, 1996; Kalmus, 1994; Rieber, 1992). However, the use of training wheels (Egley, 1994; Rieber, 1992) is probably the most popular method for introducing cycling skills to young children. Training wheels keep the novice rider in an upright position, however, they are actually counterproductive in learning how to balance and ride a conventional bicycle. The use of training wheels elicits inappropriate feedback responses. When using training wheels, the tendency is to lean to one side or the other as the bike moves forward on three wheels (typically two cycle wheels and one training wheel). This in turn, results in a preoccupation with the upper torso as a balancing mechanism. This tendency is noticeable as upper-body articulation (shifting of shoulders to one side or the other) is pronounced when observing a child attempting to balance a two-wheel bicycle after many years on training wheels.

In terms of achieving balance on a conventional bicycle without training wheels, the appropriate technique requires one to steer in the direction of the fall or lean. Roberts (1995) explains that the path of a bicycle is not a straight line, but one of constant lateral oscillations. The reason for these oscillations is that any tendency to fall to one side or the other is promptly arrested by a corresponding horizontal movement in the same direction. With this in mind, balancing a bicycle is much easier and more productive when the arms are loose and participate in the steering process. Again, the use of training wheels may tend to inhibit this necessary action. The arms are often observed as being rigid and inflexible as a result of a preoccupation with the upper-body as an alternate balancing mechanism.

To our knowledge, no empirical research has been completed that has addressed the learning of conventional bicycling skills for individuals with MR. Therefore, the purpose of this paper is for children with MR to use a series of adapted bicycles in order to acquire, maintain and generalize the skill of riding a conventional bicycle. During the acquisition phase, we had participants use a series of adapted bicycles and provided participants with feedback relative to head position, pedal rate and steering participation. According to Bouffard (1990) appropriate feedback is an essential element in learning motor skills for individuals with MR.

This study adopts an ecological perspective (Davis & Burton, 1991) relative to the learning of conventional bicycling. The interaction of all elements in the environment is taken into consideration upon learning a new skill. More specifically, the perception-action approach introduced by Gibson (1976) uses the term affordance to describe the function that an environmental object might provide a particular individual. For example, a conventional two-wheel bicycle might be the appropriate learning tool for a child without disability. However, for a child with body coordination and balancing difficulties, a modified bicycle may be more appropriate during the initial stages of learning. Achievement of movement goals are related to the characteristics of the individual as well as to environmental objects, which afford certain movements to that individual (Haywood & Getchell, 2001). The use of a series of adapted bicycles would represent a progressive change in the task at hand, in that forward speed and tipping rate are tempered to allow children the opportunity to experience the dynamics of cycling in more forgiving and appropriate conditions.

**Method**

**Participants**

Participants for this study were purposefully selected based upon age, disability and the inability to ride a conventional bicycle. Seven elementary school students (6 males and 1 female) with mild MR (ages 7-11, M age = 9.3) participated in this study. These participants are those who would qualify for intermittent support services (Luckasson et al., 2002). All participants were receiving public school special education services in a Midwest-
ern state and were identified by the school district as mildly mentally retarded. In accordance with the definition of MR, all participants exhibited deficits in adaptive behaviors and significantly subaverage intellectual functioning (68 to 84 IQ scores). Participants had no known physical or sensory impairments that would inhibit performance of cycling skills. None of the participants were able to independently ride a conventional bicycle at the beginning of the study. Written informed consent was obtained from parents/guardians of each participant, prior to data collection. In addition, participants provided verbal assent to participate. Approval of research protocol was obtained from the author’s Institutional Review Board.

**Bicycles**

In this study, adapted bicycles as designed by the third author, an emeritus professor of engineering at the University of Illinois are used (Klein, McHugh, Harrington, Davis, & Lieberman 2005). Additional information about the adapted bicycles can be obtained on Dr. Klein’s website (www.losethetrainingwheels.org). Each adapted bicycle is described below:

**Bicycle A.** The “double roller” bicycle features two rollers in lieu of conventional pneumatic tires. The rollers have been mounted onto a 16 in or 20 in (40.64 cm. or 50.8 cm. tire diameter), single-speed bicycle frame. One roller is located in the front position and the other is located in the rear position of the bicycle. Both rollers to be used on bike A have crowns of 36.25 arc-inches (92.08 arc-cm.) (roller #5). The drive train is modified to convert the design to that of fixed gear as opposed to free wheel or “coaster brake” system.

**Bicycle B.** The rear roller bicycle is a variation on the two-roller bicycle above (see Figure 1). It features only one roller in the rear position; this roller has a crown of 36.25 arc-inches (92.08 arc-cm.) (roller #5). The front fork is configured with a conventional front fork and front tire. The drive train is a fixed gear system.

**Bicycle C.** The second rear roller bicycle is a variation on Bicycle B, in that the sprocket in the front is larger and the sprocket in the rear
position is smaller, allowing the bicycle to move forward at a quicker pace relative to pedal cadence. In addition, the rear roller used on Bicycle C has a crown of 26.35 arc-inches (66.93 arc-cm.) (roller #7).

**Bicycle D.** The “fat tire” bicycle is a standard 16 in or 20 in (40.64 cm or 50.8 cm) size frame, which features a wide, 16 x 8 x 5 in (40.64 x 20.32 x 12.7 cm) inflatable garden tractor style tire located in the front position. The rear tire is a conventional bicycle tire, with a conventional drive train.

The conventional bicycle used in this study is a single speed, 16 in or 20 in (40.64 cm or 50.8 cm) bicycle with pneumatic tires and coaster brakes. This is a child’s size bicycle and will appropriately accommodate children in the 7 to 11 age range.

**Design**

A multiple probe design (Cooper, Heron, & Heward, 1987) was used for this study. The multiple probe design is appropriate for evaluating the relationship between the independent variable and the acquisition of successive approximations of a skill. Rather than collecting prolonged baseline data on skills that were not yet developed, intermittent measurements provided the information necessary to determine whether or not a change had occurred prior to intervention. All participants were introduced to conventional cycling through a series of four progressively challenging adapted bicycles. Social validity was obtained by a questionnaire designed to acquire parents'/guardians' opinions relative to their perceptions of learning conventional cycling via a progressive approach to cycling. The social validity questionnaire consisted of 10 questions with a four-item rating scale: strongly agree, agree, disagree, and strongly disagree.

Participants received individualized instruction three times a week, 45 minutes per session. Initial probes, baseline, acquisition, generalization, and maintenance took place inside a gymnasium. Participants were instructed not to practice cycling skills outside of study sessions. Parents/guardians were seated outside the gymnasium during all practice sessions. The researcher, two video-camera operators, and one participant, were present in the gymnasium at the time of data collection. This arrangement permitted participants to learn in an environment relatively free from distractions.

Each trial began at the “ready” line near the wall at one end of the gymnasium. The participant was instructed to mount the bicycle, position hands on handlebars, and feet on pedals. The instructor pushed the bicycle forward 1.5 m to the starting line. As the front tire of the bicycle touched the starting line, the lead researcher released her hand from the bicycle’s seat. This is the point where measurement of each trial began. The trial ended when: (a) the investigator touched the participant or the bicycle, (to intervene prior to a fall) (b) the participant touched the floor, or (c) the front tire of the bicycle touched a side boundary line. Upon completion of each trial, the child was instructed to dismount the bicycle and walk back to the starting position. Participants did not advance to the next bicycle in the series until a criterion level of performance was attained on the current bicycle. Criterion was set at 12 m of independent riding on 3 out of 5 consecutive trials.

**Initial Probes.** At the first training session, the participant’s helmet was appropriately adjusted for his/her head size. Once the helmet was in place, all five bicycles were adjusted to fit the participant so that both feet could easily touch the ground flat-footed while seated. Once helmet and bicycle fitting were complete, the participant attempted one trial on each of five bicycles in the series. All five probes were completed without feedback. The initial probes served to establish a starting point relative to the relationship between the independent and dependent variables of interest.

**Baseline.** Following initial probes, a baseline was established for the first bicycle in the series (Bike A). The baseline trials were performed without feedback. Each participant was told to ride the bicycle as far as he/she could. Once a steady state of responding was established for a particular bicycle under baseline conditions, the instructor began providing feedback following each trial. The purpose of requiring a stable pattern of response was to eliminate extraneous influences on performance, thereby providing an accurate picture of the full efforts of that condition prior to the introduction of another variable. The
number of trials under baseline conditions differed from participant to participant and from bicycle to bicycle.

**Acquisition.** Once a baseline was established for a particular bicycle, the acquisition phase began. At the completion of each trial, the instructor provided positive, positive/corrective, or positive/specific feedback. Feedback addressed pedal rate, steering participation, or head position. During the acquisition phase, participants rode the bicycles in sequential order (1) Bicycle A, (2) Bicycle B, (3) Bicycle C, (4) Bicycle D, and (5) Bicycle E (conventional bicycle). Participants did not advance to the next bicycle in the series until a criterion level of performance was attained on the current bicycle. Criterion was set at 12 m of independent riding on 3 out of 5 consecutive trials. Once criterion level performance was attained on a particular bicycle, one generalization probe was taken on that bicycle. Response generalization required the participant to perform a variation of the task. In this case, the participant was required to navigate through a series of cones. Following this, a single trial (probe) of each bicycle remaining in the series was taken to determine whether performance changes had occurred in any other steps. Once all probes were completed, the next bicycle in the series would be introduced in a similar manner.

**Maintenance.** Maintenance is defined as the extent to which a learner continues to perform a target behavior upon completion of the intervention (Cooper et al., 1987). In this study, a maintenance session was conducted 2-3 days following acquisition of cycling skills on the conventional bicycle (Bike E). During the maintenance session, the participant was asked to complete a maximum of 15 trials. Maintenance trials were not accompanied by feedback. Following maintenance trials, during the same session, generalization trials were completed.

**Generalization.** The generalization session took place 2 to 3 days following acquisition of cycling skills on the conventional bicycle. During generalization trials, response generalization was targeted to elicit the participant’s ability to perform a variation of the cycling task. The cones were set up in a weaving pattern and the participant was requested to complete a maximum of 15 generalization trials navigating through the cone arrangement.

**Procedure**

Interobserver agreement data were collected by two independent observers through videotaped sessions. Treatment integrity was ascertained through onsite observations to ensure that the independent variable was presented in an accurate and consistent manner. Social validity data were obtained from parents who provided opinions relative to the efficacy of utilizing adapted bicycles as a learning tool and the importance of cycling skills for recreation and leisure.

Visual inspection of graphs was used to interpret data. Individual graphics were utilized for the distance traveled independently across five bicycles (Bicycles A-E). Data analysis consisted of both within-phase and between-phase analysis. Visual analysis involved a determination of the number and variability of data points, in addition to the level of performance and the direction and degree of trends. Social validity data were analyzed using percentage of responses for each question.

**Results**

**Interobserver Agreement**

Interobserver agreement data for bicycling acquisition were collected across all seven participants. Session-by-session interobserver agreement data of distance by trial were taken from seven sessions across all five bicycles (Bicycles A-E). Across all sessions for all participants, procedural reliability was calculated at 99% (range 97 to 100%).

**Procedural Integrity**

Procedural integrity was established by the use of a procedural reliability checklist consisting of seven questions. The video-camera operator completed the procedural reliability checklist at the end of each session. Across all sessions for all participants, procedural reliability was calculated at 99% (range 97 to 100%).
Performance of Participants across Five Bicycles

Participant 1 (male) participated in one session for acquisition. Total trials were 29 for acquisition of conventional cycling skills (see Figure 2). The acquisition phase encompassed initial probes, baseline, and training trials. Initial probes on Bicycles A-E were 12, 4, 1, 12, and 3 m respectively. Participant 1 immediately reached criterion levels on Bicycles A-E during baseline. Participant 1 demonstrated maintenance consistently, cycling a distance of 12 m, during 100% of the maintenance trials. Participant 1 demonstrated generalization intermittently, cycling through ob-
stacles a distance of 12 m, during 50% of the generalization trials. Participant 2 (male) participated in two sessions for acquisition. Total trials were 50 for acquisition of conventional cycling skills (see Figure 3). Initial probes on Bicycles A-E were 12, 0, 9, 1, and 3 m respectively. Participant 2 immediately reached criterion levels on Bicycles A-C during baseline conditions. Baseline for Bike D consisted of 6 trials with an average distance of 2.5 m. The data path for training trials on Bike D was extremely variable, ranging from 0-12 m. After 11 trials with feedback, participant 2 successfully reached criterion.

Figure 3. Participant #2 performance across bikes A-E.
level on Bike D. Baseline for Bike E consisted of 5 trials with an average distance of 11 m. After 5 trials with feedback, participant 2 successfully reached criterion levels on Bike E. Participant 2 did not demonstrate maintenance, although he did cycle a distance of 12 m on 3 separate trials (20%). Participant 2 demonstrated generalization consistently at criterion level (80%).

Participant 3 (male) participated in four sessions for acquisition. Total trials were 156 for acquisition of conventional cycling skills (see Figure 4). Initial probes for Bicycles A-E were 12, 0, 0, 0, and 0 m respectively. Partici-
Participant 3 immediately reached criterion level on Bike A during baseline conditions. Participant 3 met criterion levels during baseline conditions on Bikes B and C within 6 trials. Baseline for Bike D consisted of 3 trials at 1, 1, and 0 m respectively. Participant 3 demonstrated a variable data path (range of 0-7 m) throughout the first 74 training trials (two-thirds of all training trials) on Bike D. For the remaining 27 training trials, the data path became extremely variable, (range of 2-12 m). During this period, Participant 3 met criterion levels on Bike D. Baseline for Bike E consisted of 3 trials with an average distance of 4.6 m. After 12 trials with feedback, Participant 3 successfully reached criterion levels on Bike E. Participant 3 demonstrated maintenance consistently, cycling a distance of 12 m during consecutive trials (90%). Participant 3 demonstrated generalization intermittently, cycling through obstacles a distance of 12 m, during 40% of the generalization trials.

Participant 4 (male) participated in six sessions for acquisition. Total trials were 227 for acquisition of conventional cycling skills (see Figure 5). Initial probes for Bicycles A-E were 12, 1, 1, 0, 0 m respectively. Participant 4 immediately reached criterion level on Bike A during baseline conditions. Baseline for Bike B consisted of 3 trials at 0 m. Participant 4 demonstrated a steady data path (range of 0-1 m) for the next 19 trials on Bike B. It was evident, that Participant 4 was not making adequate progress on Bike B. At this point, a roller change was made. A less contoured roller replaced the original roller for Bike B. Participant 4 reached criterion level with the less contoured roller in place after 20 trials. The data path for Bike B was extremely variable (range of 1-12 m). At this point, the original roller was reinserted on Bike B. The data path for Bike B was extremely variable (range of 1-12 m). Participant 5 met criterion level on Bike B with the original roller reinserted after 63 trials. Baseline for Bike C consisted of 3 trials at 4, 6, and 3 m respectively. Participant 5 met criterion level on Bike C after 8 trials. The data path during training on Bike C was extremely variable (range of 2-12 m). Baseline for Bike D consisted of 3 trials at 0, 1, and 0 m respectively. During training trials, the data path was initially extremely variable at low levels (range of 0-8 m), followed by extreme variability at higher levels (3-12 m). Participant 5 met criterion level on Bike D after 68 trials. Participant 5 met criterion level on Bike E during baseline conditions after 8 trials of extreme variability (range of 3-12 m). Participant 5 demonstrated maintenance, cycling a distance of 12 m during 45% of the maintenance trials. Participant 5 did not demonstrate generalization, although he did navigate a distance of 8 m on one attempt.

Participant 6 (female) participated in seven sessions for acquisition. Total trials were 389 for acquisition of conventional cycling skills (see Figure 7). Initial probes for Bicycles A-E were 12, 0, 3, 0, and 0 respectively. Participant 6 met criterion level on Bikes A-C immediately during baseline conditions. Baseline on Bike
D consisted of 3 trials; each at 0 m. During the first 165 trials of intervention, the data path was variable (range of 0-4 m). During the remaining 108 training trials, the data path became increasingly variable (range 0-12). Participant 6 met criterion level on Bike D after a total of 273 trials. Baseline for Bike E consisted of 5 trials of 4, 6, 12, 4, and 5 m respectively. Participant 6 met criterion level after 32 trials of intervention, demonstrating an extremely variable data path (range of 3-12 m). Participant 6 did not demonstrate maintenance, although she did cycle a distance of 12 m during on one attempt. Participant 6 did
not demonstrate generalization. Her best attempt at navigating through obstacles was 4 m.

Participant 7 (male) participated in seven sessions for acquisition. Total trials were 232 for acquisition of conventional cycling skills (see Figure 8). Initial probes for Bicycles A-E were 12, 1, 3, 0, and 1 respectively. Participant 7 immediately met criterion level on Bike A during baseline conditions. Baseline measures on Bike B consisted of 8 trials with an average distance of approximately 2 m. Participant 7 met criterion level on Bike B after 40 trials of intervention. During acquisition, the data path was extremely variable and gradually increasing on Bike B. Baseline measures on Bike C consisted of 5 trials with an average distance...
of approximately 5 m. Participant 7 met criterion level on Bike C after 28 trials of intervention. During acquisition, the data path was extremely variable on Bike C. Baseline measures on Bike D consisted of 3 trials at 3, 2, 3 m respectively. Participant 7 met criterion level on Bike D after 63 trials of intervention. During acquisition on Bike D, the data path was initially variable (range of 2-5 m), this was followed by an extremely variable data path (range 2-12 m). Baseline data on Bike E consisted of 7 trials with an average distance of approximately 8 m. Participant 7 met criterion level on Bike E after 38 trials of intervention.

Figure 7. Participant #6 performance across bikes A-E.
During acquisition on Bike E, the data path was extremely variable (range of 2-12 m). Participant 7 demonstrated maintenance, cycling a distance of 12 m during 47% of the maintenance trials. Participant 7 did not demonstrate generalization, although he did navigate a distance of 4 m on several attempts.

Summary of Performance across Five Bicycles for all Seven Participants

The average number of trials to criterion level for each of the five bicycles was: Bike A, $M = 4.4$ trials (range 4 to 6); Bike B, $M = 40$ trials (range 5 to 106); Bike C, $M = 12.4$ trials.
(range 6 to 36); Bike D, $M = 92.9$ trials (range 7 to 280); Bike E, $M = 35.3$ trials (range 7 to 94). The average number of total trials to acquisition of conventional cycling skill was 185 trials (range 29 to 389). In a comparison by gender, the average number of total trials to acquisition of conventional cycling skill was $M = 151$ trials (male) and $M = 389$ trials (female). On average it required 4.9 sessions to acquire conventional cycling skills. Gender comparison of total number of sessions reveals $M = 4.5$ sessions (males), and $M = 7$ sessions (female). Five out of seven participants (71.4%) were able to demonstrate maintenance of conventional cycling skill 2-3 days following acquisition. Three out of seven participants (42.9%) were able to demonstrate generalization of conventional cycling skill 2-3 days following acquisition. The only female in this study did not demonstrate maintenance or generalization of conventional cycling skill.

**Discussion**

The purpose of this study was to investigate the effects of the use of adapted bicycles on the acquisition, maintenance, and generalization of conventional cycling by children with mild MR. Acquisition of conventional cycling skills was achieved by all seven participants (100%) supporting the notion that individuals with MR can acquire motor skills (Gillespie, 2002; Porretta & Surburg, 1995; Yang & Porretta, 1999). Moreover, performance by athletes in Special Olympics events such as gymnastics, cycling, and swimming provides evidence that individuals with MR can become highly proficient at a variety of sports skills (Sherrill, 2004). Results also support the notion proposed by several investigators (Belmont & Mitchell, 1987; Bouffard, 1990; Wall, 1990) that skill progressions should be used. For example, the presentation of a series of adapted bicycles, representing a gradual continuum of change allowed the participants in this study to successfully progress toward the goal of riding a two-wheeled, conventional bicycle. In contrast, an instructional program that does not make appropriate equipment modifications to accommodate for the needs of a child with MR may not be as successful in terms of success rate and length of training. One example of such a program has been underway in Michigan (see www.bikeprogram.org). The director, John Waterman indicated (personal communication, June, 2004) that the Michigan D.O.T. sponsored program requires approximately sixteen weeks of training sessions and yet achieves lower success rates, typically fifty percent, as compared with the adapted bicycle based program used in this study. Follow-up empirical research is needed to verify this observation. Certainly, the interaction of the learner, the task, and the environment should be taken into consideration upon learning a new skill. An ecological perspective as used in this study (Davis & Burton, 1991) is important when teaching children with disabilities, especially those with MR.

Maintenance of conventional cycling skills was demonstrated by five out of seven participants (71.4%) following acquisition. Although this level of maintenance was fairly high, considering the difficulty that many individuals with MR have in maintaining motor skills, there is certainly room for improvement. In this study, the acquisition phase was terminated as soon as a participant met the criterion of three out of five consecutive trials at a distance of 12 m. This cut-off did not provide sufficient practice time riding the conventional bicycle. It has been recommended that to improve maintenance of motor skills, ample practice time should be provided (Bouffard, 1990; Porretta & O’Brien, 1991; Ulrich, Ulrich, & Angulo-Kinzler, 1998).

The generalization of conventional cycling skills was demonstrated by three out of seven participants (42.9%) following acquisition. Response generalization was utilized, in that participants were required to navigate through a series of cones while riding a conventional bicycle. The setting remained the same (inside a gymnasium) but the task was altered. There may be several reasons for the results. Perhaps the placement of the cones was too challenging after just learning to ride a bicycle from one end of the gym to the other. In this case, the task did not appropriately match the current ability of the learner. This observation supports Singer’s (1986) notion that positive transfer (generalization) effects will occur when the related task involves minimal change. Another reason may be related to practice time. Participants received minimal practice riding the conventional bi-
cycle. Therefore they did not have the opportunity to develop navigational skills.

There was a noticeable difference between male and female participants relative to acquisition, maintenance, and generalization. Overall, males acquired conventional cycling skills much faster in comparison to the one female participant in the study. On average, males acquired the skill in less than half the total trials, $M = 151$ trials (male) and $M = 389$ trials (female). This trend was also evident when examining total number of sessions, $M = 7$ sessions (male), and $M = 15$ sessions (female). Moreover, the female participant was the only participant in this study not to demonstrate either maintenance or generalization of conventional cycling skills. One possible explanation for the results relative to gender comparison would be that past experiences and personal attributes of the female participant may have differed considerably from her male counterparts. Potential characteristics that may have had a detrimental effect would be lack of opportunity, decreased muscular strength, lack of motivation, or negative experiences leading to increased levels of anxiety and fear. The female participant in this study was an only child. Perhaps she would have been more motivated or had more opportunities if she were exposed to siblings riding bicycles. However, anecdotal evidence appears to contradict the present findings relative to gender differences. The third author, in supervising approximately 1,000 children with disabilities (many with varying degrees of MR) in his adapted cycling programs for the period of 2001-2005, has not observed any significant gender differences as related to the ability to master bicycling skills.

In conclusion, the results of this experimental study demonstrate that the use of a series of adapted bicycles can lead to acquisition of conventional cycling skills for children with mild MR. It is suggested that future research in this area emphasize additional practice trials to increase both maintenance and generalization of conventional cycling skills. This would necessitate raising criterion levels. In addition, generalization to more applied settings such as outdoors or perhaps on a bicycle trail would be beneficial. Also, it is imperative to match the ability of the child to the task at hand. Not only does this practice lead to successful experiences, but it can also have a positive effect on movement confidence (Griffin & Keogh, 1982) and motivation (Harter, 1978). Finally, the use of specially designed bicycles as described and as used in a progression affords children with MR increased opportunities to succeed in the challenging task of riding a two-wheeled bicycle without training wheels.

References


Received: 15 January 2006
Initial Acceptance: 16 March 2006
Final Acceptance: 1 June 2006
Search the entire archives of
Education and Training in Developmental Disabilities
at
http://www.dddcec.org/search.htm

Mark your calendar!
The 11th International DDD Conference
Sheraton San Diego Hotel & Marina
San Diego, California
October 8–10, 2008

For further information, please contact
Cindy Perras, Conference Coordinator
cindy.perras@cogeco.ca

Visit the official Website of the
Division on Developmental Disabilities:
http://www.dddcec.org