Response to Intervention Series (RTI Series)

Level 6 Career Assessment (L-6)

Technical Manual

DEVELOPED BY:
Larry K. Irvin, Ph.D. and Russell M. Gersten, Ph.D.

with

Valerie E. Taylor, M.A.
Thomas J. Heiry, Ph.D.
Daniel W. Close, Ph.D.
G. Thomas Bellamy, Ph.D.

Further research and development by Oakwood Solutions, LLC.

Where We Are and Where We Want To Go

This instrument has its origins back in the early 1980s. The conclusion at that time was that most assessment instruments designed for severely impaired individuals were focused in the wrong direction. Most assessment instruments for this population attempt to assess a learner's prior knowledge rather than response to intervention or training. The question is not can this individual succeed in learning career/vocational tasks, but rather what is the type and amount of resources needed to help this person succeed. That type of thinking is more appropriate today than it was in the 1980s. This instrument is the beginning of a whole new set of assessment and skill intervention instruments that we are calling our **Response to Intervention Series (RTI Series)**. This instrument is the first in this exciting new series.

We have recently expanded the original assessment to include a new scaled – Video Modeling. This new scale takes into account the use of smart phone technology such as iPods and iPads.

In the months ahead we will be continuing the research on this instrument and when we release the results of our research you will receive this new information at no additional cost. We would appreciate any of your thoughts on improving this instrument. Please send your comments to us and we will incorporate the best ideas into the next version of this instrument.

Terry Schmitz, Owner
The Conover Company, a Division of Oakwood Solutions, LLC
tschmitz@conovercompany.com

PHASE 1 STANDARDIZATION

The test was initially standardized and normed exclusively on a sample of over 300 significantly impaired adults at a state institution in the Northwest (Bellamy & Snyder, 1976). Incorporation of a modified process assessment approach was an innovative feature of the test. Each of the original items represented a skill required by a variety of benchwork tasks, and included a training procedure (verbal instructions, modeling or physical guidance) within the test item in order to measure both examinees' prior knowledge and ability to benefit from instruction. Psychometric analyses demonstrated remarkably sufficient reliability and validity for demonstrating required training resources (see Bellamy and Snyder, 1976). These initial standardization data are also detailed in the RTI SERIES TECHNICAL MANUAL (Irvin, et al., 1982).

PHASE 2 STANDARDIZATION

1. PURPOSE: In 1979 the assessment was revised. The major purposes of the revision were: (a) to increase content validity by employing an item-sampling methodology that, in a systematic fashion, better represented the relevant types of tasks and instructions commonly encountered in skill training of severely impaired individuals; (b) to increase the range of appropriate examinees to include individuals classified in the low range of moderate impairment; and (c) to examine thoroughly the reliability and validity of the revised scale.

2. SAMPLE: Thus, 149 adults classified as severely or low moderately impaired were identified as a final standardization group. Selection for participation was based on willingness of participants and facility personnel to participate. Administrators of all ten facilities that were initially contacted agreed to be involved; eight of the facilities were community-based work facilities and two were institutions—one state and one private. Fifty-four percent of the individual participants were enrolled in community work-activity centers and 46% in residential institutions. All were over the age of 16 and had no major physical or visual impairments that would preclude testing. A sample of 26 community-based participants was retested two weeks after the original test for purposes of establishing the test-retest reliability of the instrument.

3. CONTENT AND FORMAT: A test blueprint for test items was developed based on the skill training methodology of the Specialized Training Program—a nationally recognized model training program (Bellamy et. al. 1979). The three dimensions considered for each test item were: type of instruction, nature of the task, and number of objects involved in the task. Item development/revision was designed to assure adequate representation of the components of each of these dimensions in the revised instrument.

   Three types of instruction—verbal directions, physical model/match-to-sample, and physical guidance or prompts—were varied systematically in revised test items. The task attributes that were represented across all revised test items were: (a) difficult discriminations, (b) coordinated two-hand movement requirements, and (c) sequencing of steps. According to Bellamy et al., (1979) these attributes are common elements of skill tasks that severely impaired persons are likely to encounter in training settings. Finally, the number of objects within test items was systematically varied. This served to guarantee a variety in types of test items; e.g., physical manipulation without objects as well as more difficult tasks with one or two objects.

Via use of the blueprint, a total of 25 items were finalized. Nine of these items involved only verbal instructions or corrections, nine involved verbal as well as modeled instructions and corrections, and seven provided physically prompted corrections in addition to verbal and modeled instructions.

Vocabulary was selected for inclusion in verbal instructions on the basis of its typical use with a variety of benchwork tasks. The vocabulary sample included nouns/pronouns (e.g., table, side, hand, it, this) verbs (e.g., are, stand, screw, stop, turn, put), and several commonly-used adjectives, adverbs, and prepositions (e.g., smallest, red, when, over, in, together). These words were combined into three- and four-word sentences for use in test items, e.g., "Turn it over." and "Stand it on its side."

The 25 individual items in the final version of the RTI Series–Level 6 career assessment involve a variety of skills and materials: (a) sorting various objects (wires, wood blocks, plastic containers) by color or size, (b) assembling various objects (e.g., nuts and bolts, bottles and caps, cards and clips, zip-back bags, metal angles, boxes and lids, axle posts and washers), (c) following verbal instructions with and without objects, and (d) manipulating a variety of single objects in various ways (e.g., bending wires, inserting pegs).

4. PSYCHOMETRIC SUMMARY FOR PHASE 2 STANDARDIZATION: A variety of test and item statistics, including difficulty and variability, as well as reliability and validity indices, are presented and discussed in detail in the RTI Series Technical Manual (Irvin et al., 1982). In summary, the internal consistency reliability (coefficient alpha) for the 25 items was $r\alpha = .95$, and test-retest reliability was $r_{tt} = .93$. The median concurrent validity correlation between RTI Series scores and ratings by supervisors at work activity centers was r .78. The predictive validity correlations between test scores and training trials- and time-to-criterion on two different training tasks (sorting and assembly) ranged from r = -.57 to 4 = -.68. These were appropriately negative in that higher test scores were related to lower trials- and times-to-criterion on the training tasks.

PHASE 3 STANDARDIZATION

A new phase of standardization is currently underway. This new standardization has three clear purposes. The first is to bring the original instrument up to today's standards to reflect the changes that may have taken place since Phase 2 Standardization was completed. The second is to develop a set of standards that can be used with the new teaching/training techniques of video modeling and smart phone technology. The third purpose is to take into account recently identified populations that could benefit from this type of assessment. The Phase 1 and Phase 2 Standardizations were built around a targeted population of individuals who were indentified as cognitively impaired based upon standardized IQ or intelligence tests of that period. Since then a whole new population has been identified that could possibly benefit from RTI Series testing – those individuals identified with Autism Spectrum Disorder (ASD) including autism, Asperger's syndrome and other developmental disorders. The drive for a solid decision making tool that is research based and can assist these individuals in modeling and learning appropriate behaviors drives our Phase 3 Standardization. The use of unscientific, non-standardized and untested approaches to teaching and training this population have lead to the use of ineffectual and even harmful intervention techniques (Heflin & Simpson, 1998; Simpson, 2005). Reactions to these ineffective practices have led to our court system, where judges and due process hearing officers decide the appropriateness of teaching/training strategies (Yell & Drasgow, 2000). As results of this standardization become available, they will be made available at no additional cost.

TABLE OF CONTENTS

SECTON I
INTRODUCTION

A. PHASE 1 RESEARCH AND DEVELOPMENT

Despite the progress in providing successful career vocational training for significantly impaired adolescents and adults (e.g., Bellamy, Horner, & Inman, 1979; Gold, 1973), little has been achieved in the way of career/vocational assessment instrumentation for this population. The general tendency with regard to assessment of significantly impaired persons is toward idiographic, task-specific behavioral assessment (Bellamy et al., 1979; Gold, 1973; Haywood, Filler, Shifman, & Chatelanat, 1975). While this methodology certainly is superior to traditional norm-referenced testing for ongoing analysis of individual worker performance, it does not ameliorate the problems associated with appropriate placement of significantly impaired youth and adults in residential and career/vocational training programs. Our major focus in the research reported here has been to address this concern via development and refinement of the Response To Intervention Series – Level 6, a career/vocational skills assessment instrument for use with significantly impaired adolescents and adults.

A survey of career/vocational assessment instruments that are potentially useful for this population yielded rather disheartening results. The type of instrument typically advocated for career/vocational assessment is the work sample (Kulman, Stephens, Armstrong, May, Rosenberg, Sax, Spoonster, & Thurgood, 1975), a “well defined work activity involving tasks, materials and tools which are identical or similar to those in an actual job or cluster of jobs” (Kulman et al., p. 55). Despite its surface appeal, there are numerous problems in applying work sample methodology with significantly impaired persons. For one thing, as Halpern (1978) observed, the validity of work sample assessment has not been established even for use with mildly impaired individuals.

Even if there were empirically validated work samples available for this population, they would be of dubious value for significantly impaired adults for several reasons. First, work samples measure the products of prior learning. Individuals with impoverished educational and training backgrounds are heavily penalized on such product measures (Irvin & Halpern, 1979).

Another problem with work samples is that they adopt a macroscopic approach toward evaluating performance. Current practice demonstrates that successful career/vocational training for severely and profoundly impaired workers depends on the microscopic breakdown of large tasks into small components so that chains of behavior can be taught (Bellamy et al., 1979). Similarly focused assessment would have direct implications for subsequent training.

A final concern regarding work samples is that the criterion behavior they purport to predict, ultimate level of productivity, may not be the most appropriate criterion variable, especially in light of accumulating evidence of the effects of task and situation variables on both learning and productivity of significantly impaired persons (Bellamy & Snyder, 1976). It seems more sensible to focus on the amount of time or effort required for training an individual given a detailed training procedure and a specific work task. This approach focuses attention on the fact that level of vocational achievement is not static but can be altered through appropriate training.

A process approach has been proposed as an alternative assessment methodology with significantly impaired persons (Budoff & Hamilton, 1976; Feuerstein, 1979; Haywood et al., 1975; Irvin & Halpern, 1979). In process assessment, an individual’s ability to benefit or learn from instruction (or training) is measured. Initial validity results with significantly impaired students indicate that process measures predict subsequent acquisition of new skills more precisely than do conventional intelligence tests (Budoff & Hamilton, 1976).

The term process approach to assessment has been recently replaced with the term response to intervention. For all practical purposes process approach and response to intervention have the same meaning and outcomes – the type and amount of resources required for individual skill training.

B. PHASE 2 RESEARCH AND DEVELOPMENT

The original instrument (Bellamy & Snyder, 1976) was a variation on work sample career/vocational assessment instrumentation for use with significantly impaired persons that also incorporated a process measurement approach. The development of the test was guided by three major considerations: (a) use of brief tasks as test items, rather than longer work samples; (b) use of process assessment rather than a product approach, i.e., use of measures of an individual's ability to benefit from current instruction rather than measures of prior learning; and (c) use of subsequent acquisition of new skills rather than "ultimate" student/client productivity as criterion validation.

The premise upon which the original instrument was based was that if a test were constructed of items that representatively sampled the domain of tasks and instructional methods commonly used in career/vocational training for significantly impaired individuals, it should have some predictive validity for illustrating training needs in subsequent acquisition of new skills.

The test was standardized and normed exclusively on a sample of 300 significantly impaired youth and adults at a state institution in the Northwest. Incorporation of a modified learning-potential model was an innovative feature of the test. Each of the 30 items represented a skill required by a variety of benchwork tasks and included a training procedure (modeling or physical prompting) within the test item. To some extent, the original instrument measured ability to benefit from instruction as well as prior knowledge.

The original instrument also incorporated a correction procedure within the item format. If individuals performed the item correctly on the first trial, they were simply asked to "do it again". If they were unsuccessful the first time, the examiner presented a second trial, repeating the model or prompt, and offering examinees an opportunity to benefit from a correction procedure. Item scores were scaled 0, 1, and 2 to reflect amount of assistance required, with 0 reflecting maximal assistance.

Psychometric analyses demonstrated a remarkably high internal consistency reliability (coefficient alpha = .961) and high test-retest reliability (r_{tt} = .91). Criterion-related validity results were promising. There were reasonably strong, appropriately negative correlations between the original scores and the training time, trials-, and errors-to-criterion necessary to teach several career/vocational assembly tasks via a variety of types of training procedures (median r = -.72).

In the original instrument (Bellamy & Snyder, 1976) the use of instruction by models was emphasized, with only a few items involving physical prompts and none involving verbal directions only. The recent articulation of a comprehensive career/vocational training methodology for significantly impaired individuals (Bellamy et al., 1979) enabled a more systematic approach to the generation of test items. Accordingly, in 1979 we revised the original instrument. The major purposes of our revision and the focus of the research reported here was:

(a) to increase content validity by employing an item-sampling methodology that, in a systematic fashion, represented the relevant arrays of types of tasks and types of instructions commonly encountered in career/vocational training of significantly impaired individuals; (b) to increase the range of appropriate examinees to include individuals classified in the low range of moderate cognitive disabilities; and (c) to examine thoroughly the reliability and validity of the revised scale.

C. PHASE 3 RESEARCH AND DEVELOPMENT

The original development of a career assessment instrument for significantly impaired individuals that would assess an individual's response to training was certainly a novel idea. When one takes into consideration that this original research took place almost 30 years ago, one cannot help to think that this work was <u>way ahead of it's time</u>. With all the current emphasis on proper assessment techniques for identifying an individual's Response to Intervention (RTI) it is easy to conclude that this instrument is more appropriate today than it was when it was first introduced in the early 1980s. You can, if so desired, participate in this research to help us further our knowledge on the use of this instrument to assist significantly impaired learners in successful transition from education and training to the workplace.

Step 1: We will first look at the utility of the instrument with our current knowledge of effective instructional and training strategies used with significantly impaired youth and adults.

Step 2: The second phase of this research will be to implement changes identified from the first phase.

Step 3: Thirdly, with the completion of Step 1 and Step 2 new normative data will be collected.

Step 4: Finally phase 4 will be the development of technology to assist in the administration and scoring of the updated instrument.

Please contact us if you are interested in participating in this much needed research and development.

SECTION II
CONTENT VALIDITY

A. METHOD

For the original version of this instrument, a test blueprint was developed. The three dimensions of the blueprint were type of instruction, task attributes, and number of objects involved in the task. Item development/revision was designed to assure adequate representation of the components of each of these dimensions in the revised instrument.

Three types of instruction--verbal directions, model/match-to-sample, and physical prompts--were varied systematically in revised instrument items. The task attributes that were represented across all revised test items were: (a) difficult discriminations, (b) coordinated two-hand movement requirements, (c) sequencing of steps, and (d) conjunctive stimulus control. Bellamy et al. (1979) stated that these attributes serve as useful difficulty descriptors for career/vocational tasks. Finally, the number of objects within each task was systematically varied. This served to guarantee a variety in types of test items, e.g., physical manipulation without objects as well as more difficult tasks with one or two objects. Via use of the blueprint for item generation or revision, we prepared a total of 25 revised or newly developed items for piloting. Seven of these items included only verbal instructions or corrections, nine included only modeled instructions and corrections, and seven included only physically prompted corrections.

Vocabulary was selected for inclusion in verbal instructions or corrections on the basis of its representativeness of simple vocabulary that could be used across a variety of benchwork tasks. The vocabulary sample included 10 nouns/pronouns (e.g., table, stand, screw, top, this), 10 verbs (e.g., are, stand, screw, stop, turn, put), and several adjectives, adverbs, and prepositions (e.g., smallest, red, when, over, in, together). These words were combined into 3- and 4-word sentences for use in original test items, e.g., “Turn it over”, and “Put this together”.

Thus, the individual items in the revised instrument involve a variety of skills and materials: (a) sorting various objects (wires, wood blocks, plastic containers) by color or size; (b) assembling various objects (e.g., nuts and bolts, bottles and caps, cards and clips, zip-back bags, metal angles, boxes and lids, axle posts and washers); (c) following verbal instructions with and without objects; and (d) manipulating a variety of single objects in various ways (e.g., bending wires, inserting pegs).

B. SAMPLE ITEMS

Two examples of test item scripts are:

(1) Materials: Four 6-inch pieces of copper wire

Trial 1

SETTING: Examinee sits facing or beside Tester. Tester has placed the four wires on the table out of examinee’s reach.

CUE: Say “DO THIS” while picking up one wire, placing your thumbs and index fingers around each end of the wire, and bending it into an inverted U-shape, and say “GOOD”. Say “YOU DO IT” while giving a second wire to examinee.

SCORING: If examinee bends wire at least 30° from the horizontal within 10 seconds, score + in the O provided for Trial 1 of Item 1. If resistance or NO RESPONSE from examinee within 10 seconds, score 0 in the O provided for Trial 1 of Item 1, and place a checkmark in the space next to “Resist Prompts”, if appropriate, at the bottom center of the Scoring Form. If examinee bends wire incorrectly, score - in the O provided for Trial 1 of Item 1.

Trial 2

IF TRIAL 1 CORRECT: Say “GOOD, DO IT AGAIN” while giving another wire to examinee.

IF ERROR ON Trial 1: (prompt) Say “O.K. DO THIS” while taking a third wire and positioning examinee’s thumbs and index fingers around each end of the wire and moving examinee’s hands so that the wire bends into an inverted U-shape. Then say “YOU DO IT” while giving the fourth wire to examinee.

(22) Materials: One wood base with two metal posts in it and four rubber washers

Trial 1

SETTING: Examinee sits facing or beside Tester. Tester has placed the base and two rubber washers on the table in front of the examinee.

CUE: Say “DO THIS” while placing one washer on the post with the SMALL END UP. Point to the model and say “GOOD”. Place a second washer so that it is on top of the first washer, SMALL END DOWN. Point to washers and say “GOOD”. Say “YOU DO IT” while placing the remaining 2 washers on the table in front of examinee. Leave first set of washers as a model.

SCORING: If examinee places both other washers on the other post, BOTTOM WASHER SMALL END UP and TOP WASHER SMALL END DOWN within 10 seconds, score + in the □ provided for Trial 1 of Item 22. If resistance or NO RESPONSE within 10 seconds, score 0 in the □ provided for Trial 1 of Item 22, and place a checkmark in the space next to “Resist Prompts”, if appropriate, at the bottom center of the Scoring Form. If examinee places the washers incorrectly, score - in the □ provided for Trial 1 of Item 22.

Trial 2

IF TRIAL 1 CORRECT: Say “GOOD, DO IT AGAIN” while giving the second pair of washers to examinee again.

IF ERROR ON TRIAL 1 (mode 1) Say “NO, TRY AGAIN” and remove incorrect washers, leaving correct ones as a model. Say “DO THIS” while placing one washer with the small end up on the post. Point to the model and say “GOOD”. Place second washer so that it is on top of the first washer, small end down. Point to washers and say “GOOD”. Disassemble washers to examinee’s left and place them on table in front of examinee. Point to empty post and say “YOU DO IT” leaving pair of correctly placed washers as a model to the right.

Item number 1, as detailed in the preceding description, involves a simple manipulation with an initially modeled instruction and with a physically prompted correction if necessary. Item number

22 involves a small assembly requiring sequential steps, with an initially modeled instruction and a repeated model as a correction if required. Other test items employ the same general format, the content described previously, and systematically varied verbal, modeled, and prompted initial instructions and subsequent corrections.

SECTION III
COMPARSON GROUPS

A. SAMPLING

Delineating operational definitions for the samples proved to be problematic. The most appropriate population for standardization and validation of the revised test was adolescents and adults classified as significantly impaired or those in the low moderately cognitively impaired. Many of the individuals in institutions and work activity centers, however, have not taken intelligence tests in years. Furthermore, the appropriateness of a Wechsler Adult Intelligence Scale is questionable because the high verbal loadings of these tests make results relatively uninterpretable for this population (Anastasi, 1976). Thus, a decision was made to define the population in the following manner: (a) In general, Mercer's (1975) "community-based definition" of severe and low moderate retardation was used. Individuals so classified by staff members of schools, institutions, and work activity centers would be considered eligible candidates for the revised test (b) Raw scores on the Peabody Picture Vocabulary Test (PPVT, Dunn, 1973) were recorded for all individuals. Throne (1972) argued that raw scores on standardized tests can offer valuable information for evaluation of instruction. Similarly, raw scores can offer a reasonably reliable distribution of verbal intellectual abilities of a test sample for purposes of validity analyses.

Thus, in 1978-79,171 adults and 56 adolescents classified as severely or low moderately cognitively disabled were identified as participants. Selection for participation was based on willingness of school and facility personnel to participate. Administrators of all 13 facilities serving adults that were initially contacted agreed to be involved; 11 of the facilities were community-based work facilities and two were institutions--one state and one private. Fifty-four percent of the individual adult participants were enrolled in community work-activity centers and 46 percent in residential institutions. All were over the age of 16 and had no major physical or visual impairments. A sample of 26 community-based adult participants was retested two weeks after the original test. For the study of test utility with significantly impaired adolescents, a large secondary-level educational services district in Pennsylvania (services to 19 school districts) was selected on the basis of widely available and willing participation of the target population. Students' classification as severely or profoundly handicapped by the school district was on the basis of IQ/SQ, adaptive behavior, and institution referral records. All participating students' IQ were below 40. Ages of Pennsylvania participants ranged from 12 to 22, with mean, median, and mode between 17.0 and 17.5. The standard deviation was 2.4.

All participants were individually tested on the revised test instrument by project staff members during 1978-79. Testing required one 30-minute session per participant and occurred at the participant's residence/work/school site.

B. DESCRIPTIVE STATISTICS

Mean difficulty and variability data for test scales for all sample groups are presented in Tables 1, 2, 3, and 4 below:

Table 1
Mean Difficulty and Variability Indices
of Test Scales for Total Northwest Sample*

	# of Points Possible	Mean Correct Score	Mean % Correct	Raw Score Standard Deviation
RTI Series Total	50	29.4/50	58.8%	12.8
Verbal	14	7.5/14	53.3%	4.2
Model	18	10.2/18	56.6%	5.2
Prompt	14	9.2/14	65.7%	3.5

* Institution and Community-Based Combined (N = 171)

Table 2
Mean Difficulty and Variability Indices
For Northwest Institution Sample*

	# of Points Possible	Mean Correct Score	Mean % Correct	Raw Score Standard Deviation
RTI Series Total	50	22.1/50	44.1%	11.5
Verbal	14	5.6/14	40.0%	3.7
Model	18	7.1/18	40.0%	4.6
Prompt	14	7.6/14	54.3%	3.4

* (N = 68)

Table 3
Mean Difficulty and Variability Indices
for Northwest Community-Based Sample*

	# of Points Possible	Mean Correct Score	Mean % Correct	Raw Score Standard Deviation
RTI Series Total	50	34.3/50	68.6%	11.3
Verbal	14	8.9/14	63.5%	4.0
Model	18	12.3/18	68.3%	4.6
Prompt	14	10.3/14	73.5%	3.2

* (N = 103)

Table 4
Mean Difficulty and Variability Indices
for Pennsylvania School-Based Sample*

	# of Points Possible	Mean Correct Score	Mean % Correct	Raw Score Standard Deviation
RTI Series Total	50	20.5/50	41.0%	12.1
Verbal	14	3.6/14	25.7%	2.8
Model	18	7.1/18	40.0%	5.4
Prompt	14	7.7/14	55.0%	4.0

* (N = 56)

Comparison of the data in Tables 1, 2, 3, and 4 reveals that, as expected, community-based adults (Table 3) outperform all other comparison groups on all test scales. Institution-based adults (Table 2) do slightly better than institutional school-based adolescents (Table 4) on the total test, but demonstrate better verbal skills.

Another result evidenced in Tables 1-4 is that, again as expected, the seven test items in which only verbal instructions were provided were the most difficult for all groups. The nine test items involving modeled instructions/corrections were next-most-difficult; and the seven items involving physically prompted corrections after modeled instructions were clearly the easiest for all groups.

Variability indices demonstrate that the test, and each of its scales, is sensitive to all levels of examinee functioning. In addition, all variability indices remain reasonably stable across all groups.

SECTION IV
RELIABILITY

A. TOTAL TEST AND SUBSCALES

To determine reliability indices for the revised instrument, we employed several procedures. After the test was administered, a coefficient alpha internal consistency reliability estimate (Nunnally, 1967) was calculated for all test scales for all samples. Temporal stability was estimated by calculation of a conventional test-retest Pearson product-moment correlation (Anastasi, 1976) on the 26 adults for whom the test was administered twice. The test-retest reliability was r_{tt} = .93 for the 26 individuals randomly selected for retesting in the total Northwest sample. Means and standard deviations at initial and later (2-3 weeks) testings were equivalent (Means = 41.4 [S.D. = 12.7] and 41.2 [S.D. = 12.6], respectively).

The internal consistency coefficients for all groups for all test scales are presented in Table 5.

Table 5
Coefficient Alpha Internal Consistency Reliabilities

	Number of Items	Total N.W. Sample (N = 171)	N.W. Institution Sample (N = 68)	N.W. Community Sample (N = 103)	Pennsylvania School Sample (N = 56)
RTI Series Total	25	.93	.92	.91	.91
Verbal	7	.83	.77	.83	.64
Model	9	.86	.82	.82	.81
Prompt	7	.75	.71	.72	.77

As is apparent in Table 5, the internal consistency reliability of the test is, on the whole, quite high for all groups. The three coefficients below .75 appear to be due, at least in part, to reduced variability in those samples on those subscales (see Tables 2, 3, & 4). For example, the internal consistency of the verbal scale for the school-age sample was .64, and the standard deviation was 2.8, compared to r?'s = .77 to .83 for other groups where standard deviations ranged from 3.7 to 4.2. To some extent this is to be expected because the school-age sample comprised the lowest scoring group on all test scales, <u>and</u> the verbal subscale was the most difficult of all test scales.

B. ITEM/TEST CORRELATIONS

Correlations were calculated between items and the total test score, and between items and their respective test subscales in order to determine item discriminability indices. The ranges of item/test correlations for all test scales for all groups are presented in Table 6.

Table 6
Test Item/Total Correlations

	# of Items	Northwest Institution (N = 68)	Northwest Community (N = 103)	Pennsylvania Schools (N = 56)
RTI Series Total	25	.35-.66	.25-.73	.23-.67
Verbal	7	.37-.60	.39-.69	.29-.68
Model	9	.41-.62	.41-.61	.39-.65
Prompt	7	.15-.64	.19-.65	.34-.62

As is evident in Table 6, item discriminability indices are adequate for RTI Series and all test scales.

SECTION V
VALIDITY

A. CONCURRENT VALIDITY

Palmer's (1974) factor analysis of his vocational-rating scale for mildly cognitively disabled persons was used as a basis for ensuring broad, representative sampling of behavior in a newly created Supervisor Checklist, which included types of career/vocational skill behavior appropriate for significantly impaired individuals. Palmer found four major categories of work behavior that correlated with career/vocational success: stability of worker's performance, ability to follow instructions, social/cooperation and peer-relation skills, and physical dexterity skills. We used Palmer's categories as a guide for generating new career/vocational skill checklist items. These items were then organized into a concurrent validation scale in a manner ensuring that Palmer's (1974) categories were adequately sampled. Each item was a behavioral statement such as "responds to simple verbal directions without requiring further assistance" or "perform tasks which require the use of a hand tool, such as a screwdriver, with minimal assistance". Each statement was followed by a 3-point rating scale, most of which involved frequency ratings ("usually or always", "sometimes", "seldom or never"). To determine concurrent validity we correlated examinees' test scores with scores on the newly created Supervisor Checklist. The most appropriate work supervisor completed the Supervisor Checklist for each examinee. Directions included an instruction to rate individual's performance within the context of the performance of other individuals at that work site, Completion required about 5 to 10 minutes per examinee.

An additional index for understanding concurrent validity, albeit a tentative one, is correlation of scores and PPVT raw scores, This provides a rough estimate of how revised Sample scores would correlate with scores taken from a standardized intelligence test. The PPVT testing was conducted by project staff members in a 15-minute session with individual examinees just prior to testing with the RTI Series.

Prior to validity analyses, the psychometric characteristics of the Supervisor Checklist were evaluated. The average item rating was 2.21 out of a possible 3, slightly above the midpoint value. Since supervisors were asked to rate each worker compared to other workers at that shop, the expected mean per shop would be 2.0. The deviation between actual mean and expected mean was .2, indicating a slight tendency toward leniency on the part of the staff members who provided ratings. Median item SD was .75 (possible range = 1 to 3). The coefficient alpha internal consistency reliability estimate was r? = .95 for the 23 items.

The results of the correlational analyses of RTI Series test scores, Supervisor Checklist scores, and IQs (PPVT) were encouraging. Two sets of revised RTI Series/Supervisor Checklist validity coefficients (Pearson product-moment correlations) were calculated--those for institutionalized participants ($n = 68$) and those for community-based participants ($n = 81$). The relationship between all institutionalized residents' revised RTI Series and Supervisor Checklist scores was $r = .40$ ($p < .001$), while that for all community-based participants was $r = .64$ ($p < .001$).

Validity coefficients between revised RTI Series and Supervisor Checklist scores were also computed for each of the 10 sites, since the vocational training staff members at each site had been asked to rate each worker in terms of other workers at that site. The range of these 10 correlations between Supervisor Checklist and revised Sample scores was from $r = .97$ to $r = .34$, with a median value of $r = .78$ ($p < .01$); two of these correlations were nonsignificant (r below .50). The eight statistically significant correlations between the revised RTI Series and Supervisor Checklist scores were lower at both institutions than at any of the community facilities.

Correlations between PPVT/Supervisor Checklist scores were also calculated for every site; all PPVT/Supervisor Checklist correlations (median r = .40) were appropriately lower than the corresponding revised RTI Series/Supervisor Checklist rating correlations (median r = .61).

Revised RTI Series scores correlate moderately well with both raw scores on a standardized intelligence test and supervisor judgment of worker performance. The two sites with nonsignificant correlations between revised Sample and supervisor's scores were the two of the eight community-based sites with the highest functioning clients; the nonsignificant correlations between those RTI Series scores and supervisor's ratings were probably due to the restricted range of test scores at these sites. The fact that all PPVT/Supervisor Checklist correlations were lower than corresponding (same site) RTI Series/Supervisor Checklist correlations suggests that though the test was moderately correlated with a general index of IQ, it offers information about important criterion vocational behaviors (those on the Supervisor Checklist) that is different than that provided by intelligence tests.

In addition, the correlation of test scores with staff members' judgments of current vocational performance was greater at the smaller workshops than at large institutions, indicating some potential utility for the instrument in deinstitutionalization decisions; i.e., at the smaller centers where there is more often individualized training and supervision, the revised RTI Series supplies less new information than at the larger institutions where most placement decisions will have to be made.

B. CRITERION-RELATED VALIDITY

To determine criterion-related validity of the revised RTI Series, several steps were taken in an effort to demonstrate that: (1) differential applications of instructional variables based on test results produce varying achievement results with different learners; and/or (2) that level of training effort (time, trials, cost, etc.) varies predictably according to RTI Series test performance. An approach that focused on predicting significantly impaired persons' responses to instruction on well-defined tasks was chosen to ameliorate at least two traditional validity-related problems associated with wider-band approaches. First, by measuring a set of clearly operationalized "learning" skills, we could better spread out for purposes of selection, placement, and training, a group of people who are not distinguishable from one another on broad IQ-type measures, because they all score low on those broader measures. And second, by measuring examinee response to small task-analyzed units of content, we accomplished better criterion-sampling of likely subsequent training events (step-by-step repeated practice on small units of content).

Clearly defined, replicable career/vocational tasks with specified training procedures were the starting point for development of a criterion validation measure. Initial pilot validity studies conducted by Bellamy and Snyder (1976) and Irvin and Bellamy (1977) indicated reasonably high negative correlations between scores on an initial version of the RTI Series and actual training time, i.e., the higher the scores the less training time subsequently necessary for the client. In the first of these studies, subjects learned to assemble an 11-piece cam switch subassembly similar to that described by Bellamy, Peterson, and Close (1975) using either a large step forward chaining procedure similar to that described by Gold (1972) or a small step forward chaining procedure like that used by Crosson (1967). The two procedures appeared to have similar efficiency. Training continued in each case until a criterion of two consecutive errorless assemblies was achieved. Training time (minutes) required before this criterion was the dependent variable of interest. In the second study, subjects were trained using either forward or reverse chaining procedures to assemble a three-piece heat sink assembly. Training time in minutes before a criterion of two consecutive perfect assemblies was again the criterion variable. In the third study, conducted by Irvin and Bellamy (1977) subjects were taught a difficult visual discrimination via three different stimulus enhancement and fading techniques. In each condition subjects learned to place a hex nut on a bicycle axle with the slightly raised face of the nut facing upwards. Training continued until a criterion of 20 consecutive correct discriminations, and the number of trials before reaching this criterion was the dependent variable.

Correlations between test scores and training time ranged between - .21 and -.81, with a median value of -.72. In over half the samples, at least 50 percent of the variance was explained by test scores. The results were promising considering the range of tasks and types of training strategies utilized. However, the sample sizes were too small (median sample size of 14) to allow for generalizable results.

Thus, additional validity studies were designed to involve larger numbers of examinees/trainers so that correlations could be considered stable reflectors of such relationships across the population. In addition, two different criterion tasks were selected to enable examination of task specific effects. The overall intent was to investigate criterion-related validity of the test for demonstrating differential subsequent training needs of significantly impaired students/clients when a clearly specified narrowband methodology was employed in the development of criterion tasks. Approximately seventy-five severely/profoundly/low-moderately handicapped adults in community vocational training settings in Oregon and Washington participated in the criterion-validation study.

As in the standardization study, selection for participation was based on facility and client willingness to participate. Approximately fifty-five clients from the eight facilities that participated in the standardization study participated in the additional criterion-related validation effort. Also, approximately twenty clients were included from three additional community-based vocational training sites in the Northwest.

Clients were tested on the instrument during March-April, 1979, and trained during May-August, 1979, on two different prototypes of vocationally relevant tasks; i.e., tasks which systematically involve a variety of task attributes which are commonly encountered in contemporary career/vocational skill training settings for significantly impaired adolescents and adults. Testing was conducted by trained project staff, and criterion-validation training was implemented by both trained workshop staff and trained project staff.

The first task required the sorting of four small objects (solderless wire terminals) into separate bins on the basis of shape. Color and size were held constant. Two of the objects were quite similar in shape--a flat U-shaped terminal and a flanged U-shaped terminal. The other two objects were quite different in shape from the first two and only somewhat similar in shape to each other--a flat spade-shaped terminal and a curled-edge terminal. See Figure 1 for an illustration of all four objects.

Validity Study Sorting Task Materials

Twenty objects were placed in a small container on a table in front of the trainee who was seated at the table, and a 4-bin tray was placed behind the container on the table. One of each of the four sorting objects was attached to the back wall of each bin as a match-to-sample model intended to show the trainee where to place each type of object.

Standardized training and sorting procedures derived from contemporary training practice as presented in Bellamy, Homer, and Inman (1979) were employed. Initial training consisted of a demonstration of correct performance by the trainer, with three different objects, to individual trainees who were seated next to the trainer. Subsequent verbal instructions were given to trainees to “do the rest”, followed by full physical guidance corrections for all errors, with verbal and gestural references to the match-to-sample models.

Ten consecutive trials correct (each object picked up represented one trial) was criterion performance. If a trainee made any errors in the first twenty trials, a massed practice procedure was implemented after those twenty trials. In this procedure, twenty trials of “massed” training were focused on the two objects that produced the most errors. Training and correction procedures were as described above. After twenty trials of this massed practice procedure, another twenty trials of training was provided with all four objects. This process of alternating between massed practice and standard training occurred every twenty trials until criterion was achieved with standard training procedures. Scoring involved counting total trials-to-criterion.

The second criterion-validation task required the assembly of a five- piece cable connector. See Figure 2 for an illustration of an assembled and an unassembled demonstration of this second task.

Validity Study Assembly Task Materials

Standardized training procedures involving trainer demonstration and full physical guidance for all errors were employed. For this task, massed practice was not used for errors. Ten of each of the five assembly parts were placed left-to-right in a five compartment bin, in front of the trainee, in the order in which they were to be picked up and assembled (See Figure 2). Trainers used a fifteen-step task analysis to guide error correction and data collection. Whole task, forward chaining training continued until any two of three consecutive complete assemblies were achieved correctly.

To determine criterion-related validity of the instrument for illustrating subsequent training time and resources required by examinees, trials-, errors-, and time-to-criterion for both validity tasks were correlated with the total test score, with verbal subscale score, with model subscale scores, and with prompt subscale score. In addition, total and subscale scores were intercorrelated to enable clearer interpretation of the predictive validity data.

The predictive validity coefficients are presented in Tables 7 and 8. Table 7 shows the intercorrelations among test scores, and trials-, errors-, and time- (in minutes) to-criterion performance on the criterion sorting task.

Table 7
Validity Coefficients for Predicting Performance
on the Criterion Sorting Task

	Criterion Sorting Task Measures		
	Trials-to-Criterion	Errors-to-Criterion	Time-to-Criterion
RTI Series Scale			
Total	-.59	-.59	-.66
Verbal	-.24*	-.26*	-.25*
Model	-.56	-.55	-.63
Prompt	-.56	-.47	-.69

All r's > .40 are statistically significant at p < .001
*n.s. (p's = .03)

Table 8 presents the intercorrelations among scores and trials-, errors-, and time-to-criterion performance on the criterion assembly task.

Table 8
Validity Coefficients for Predicting Performance
on the Criterion Assembly Task

	Criterion Assembly Task Measures		
	Trials-to-Criterion	Errors-to-Criterion	Time-to-Criterion
RTI Series Scale			
Total	-.58	-.66	-.57
Verbal	-.19*	-.24*	-.20*
Model	-.57	-.64	-.55
Prompt	-.59	-.61	-.57

All r's above .50 are statistically significant at p < .001

*n.s. (p's = .03 to .09)

Examination of Tables 7 and 8 reveals that total scores are valid predictors of subsequent career/vocational training performance of severely handicapped examinees. Validity coefficients are moderately high (r's range from -.57 to -.66) between total scores and subsequent performance on two different training tasks (sorting and assembly). The negative correlations are appropriate in that as scores are higher, subsequent training requires fewer trials-, errors-, and time-to-criterion.

Essentially, these results suggest that test scores can be used as indicators of likely training resources required by individual examinees with less one-to-one trainer time, and the more likely it is that verbal instructions, as well as modeling and prompting, will work effectively as training tools. Conversely, low scorers will clearly require maximal one-to-one trainer attention, repeated training trials and modeled and physically prompted instructions and corrections.

The utility of this type of career/vocational assessment information is high for placement and instructional design decisions regarding significantly impaired individuals. High scorers are good candidates for commonly available sheltered workshop and activity center settings. Low scorers, on the other hand, cannot be expected to achieve habilitative career/vocational skills outside of a training environment where maximal assistance is always available. Acquiring and using such information prior to placement would enable a habilitation specialist to enhance the likelihood of APPROPRIATE career/vocational training for all significantly handicapped individuals.

Of interest also, in Tables 7 and 8, is the apparent disparity between the predictive ability of “verbal” items and “model/prompt” items in the test. Examinee performances on the “model” and “prompt” items are clearly more indicative of subsequent training performance than is their performance on “verbal” items.

Two observations appear warranted with regard to this disparity. First, receptive verbal competencies in examinees, as measured by the “verbal” subscale, are not strong predictors of subsequent responsiveness to training; whereas responsiveness to examiner modeling and prompting is a moderately strong predictor of “trainability”. That is, examinees who evidence high verbal receptivity do not necessarily respond to career/vocational training any better than do those who evidence low verbal receptivity. But the individuals who show responsiveness to examiner models and prompts during testing are more predictably those who demonstrate subsequent responsiveness to training.

Second, a reasonable explanation for the disparity in predictive ability between “verbal” and “model” and “prompt” subscales is that verbal directions and corrections do not function as instruction for examinees in the manner that examiner models and prompts do. Rather, it appears that verbal directions may operate only to signal examinees to begin performing, whether they know how to do the task or not. Then those examinees who are familiar with the task materials perform the task and those who are not do not, relatively regardless of verbal receptivity skills. Modeled and prompted directions and corrections, on the other hand, supply examinees with additional information than they themselves bring to the testing situation, and thus provide direct instruction. The stronger predictive ability of “model” and “prompt” items emerges because those items are actually more of a relevant sample of the criterion events that occur during career/vocational training.

These observations are even further supported by the results of correlating the subscales. Table 9 shows those correlations.

Table 9
Intercorrelations Among
Subscales and Total Scores

	Verbal	Model	Prompt	Total
Verbal	1.0	.27[a]	.24[b]	.65
Model		1.0	.75[c]	.88
Prompt			1.0	.82
TOTAL				1.0

a $p = .01$
b $p = .02$
c $p = .001$

As is apparent in Table 9, “verbal” items are substantially less correlated with “model” and “prompt” items than are those model” and “prompt” items with each other. The implication is that “verbal” items supply somewhat different information than do “model” or “prompt” items. “Verbal” items appear to reflect already acquired examinee familiarity with either verbal stimuli or test tasks, whereas “model” and “prompt” items appear to reflect responsiveness to training regardless of prior levels of acquired skills.

In summary, the current results, combined with the earlier predictive validity results, confirm that the RTI Series can be used to identify examinee response to subsequent career/vocational training. The total “model” and “prompt” scores can be used as effective indicators of levels and types of training required by individual examinees. In addition, performance on the verbal subscale can be used to gauge to what extent verbal instructions will be helpful in subsequent training settings.

C. CONSTRUCT VALIDITY

One method of assessing construct validity was by contrasted groups (Martuza, 1977). One such index is the disparity between scores of institutionalized and community-based examinees. Experiences related to revised content are clearly more readily available in community workshop/activity center settings than in typical institutional prevocational training programs. The revised instrument appears to be quite sensitive to these experience-related effects; the mean total score for eighty-one community-based residents was 71 percent correct, while that for sixty-eight institutionalized residents was 45 percent (SDs were similar--13.2 and 14.0 respectively). A test for differences between independent means affirmed that these means were significantly different ($t = 9.29$, 147 df, $p < .0001$). Additionally, as can be seen in Table 10, distributions of scores across the possible range of 0 to 60 were distinctly different for the two groups in the predicted direction, suggesting that extreme scores of a few do not account for the mean difference.

Table 10
Percentages of Community/Institution-Based
Trainee Performance Sample Examinees in
Different Score Intervals

Score Intervals	Community-Based[a]	Institution-Based[b]
0 to 15	4.9	25.0
16 to 30	13.6	29.4
31 to 45	29.6	35.3
46 to 60	51.9	10.3

a N = 81.
b N = 68.

Two other construct validity studies are underway--a factor analysis of the instrument with different subject groups, and an examination of the utility of the verbal subscale for identifying examinees for whom verbal instructions are/are not facilitative during training. Preliminary results suggest that: (1) the factor structure is stable across groups and generally supports the verbal-model-prompt design of the test; and (2) the verbal scale is a useful tool for identifying examinees for whom verbal instruction will be helpful during subsequent training. Full descriptions of these construct validity studies, and their implications for test users, will appear in the initial revision of this Technical Report.

SECTION VI
BIBLIOGRAPHY

Anastasi, A. Psychological Testing (4th ed.). New York: Macmillan, 1976.

Baharav, E., & Darling, R. (2008). Case report: Using an auditory trainer with caregiver video modeling to enhance communication and socialization behaviors in autism. *Journal of Autism and Development Disorders*, 38, 771-775.

Banda, D., Matuszny, R.M., Turkan, S. (2007). Video modeling strategies to enhance appropriate behaviors in children with Autism Spectrum Disorders. *Teaching Exceptional Children, 39*, 47-52.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran. (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). New York: Academic Press. (Reprinted in *Encyclopedia of mental health*, by H. Friedman, Ed., 1998, San Diego: Academic Press.)

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Bellamy, G. T., & Snyder, S. The Trainee Performance Sample: Toward the Prediction of Habilitation Costs for Severely Handicapped Adults. AAESPH Review, 1976, 1, 17-36.

Bellamy, G. T., Homer, R. H., & Inman, D. P. Vocational Habilitation of Severely Retarded Adults: A Direct Service Technology. Baltimore: University Park Press, 1979.

Bellamy, T., Peterson, L., & Close, D. Habilitation of the Severely and Profoundly Retarded: Illustrations of Competence. Education and Training of the Mentally Retarded, 1975, 10, 174-186.

Bellini, A., Akullin, J. (2007). A meta-analysis of video modeling and video self-modeling intervention for children and adolescents with autism spectrum disorders. *Exceptional Children, 75(3)*, 264-387.

Bellini, S., Akullian, J., & Hopf, A. (2007). Increasing social engagement in young children with autism spectrum disorders using video self-modeling. *School Psychology Review*, 36, 80-90.

Bellini, S., Akullian, J., & Hopf, A. (2007). Increasing Social Engagement in Young Children with Autism Spectrum Disorders Using Video Self-Modeling. *School Psychology Review:* Volume 36, Issue No. 1.

Bidwell, M. A. , & Rehfeldt, R. A. (2004). Using Video Modeling to Teach a Domestic Skills With an Embedded Social Skill to Adults with Sever Mental Retardation, *Behavior Interventions, 19,* 263-274.

Budoff, M., & Hamilton, J. Optimizing Test Performance of Moderately and Severely Mentally Retarded Adolescents and Adults. American Journal of Mental Deficiency, 1976, 81, 49-57.

Buggey, T. (2005). Video self-modeling applications with students with autism spectrum disorder in a small private school setting. *Focus on Autism and Other Developmental Disabilities, 20*, 52-63.

Charlop, M. H., & Milstein, J. P. (1989). Teaching autistic children conversational speech using video modeling. *Journal of Applied Behavior Analysis, 22*, 275-285.

Charlop-Christy, M. H., & Daneshvar, S. (2003). Using video modeling to teach perspective taking to children with autism. *Journal of Positive Behavior Interventions, 5*, 12-21.

Charlop-Christy, M. H., Le, L., & Freeman, K. A. (2000). A comparison of video modeling with in vivo modeling for teaching children with autism. *Journal of Autism and Developmental Disabilities, 30*, 537-552.

Charlop-Christy, M.H., & Daneshvar, S. (2003). Using Video Modeling to Teach Perspective Taking to Children with Autism. *Journal of Positive Behavior Interventions:* Volume 5, Issue No. 1, pp. 12-21.

Charlop-Christy, M.H., Le, L., & Freeman, K.A. (2000). A Comparison of Video Modeling with In Vivo Modeling for Teaching Children with Autism. *Journal of Autism and Developmental Disorders:* Volume 30, Issue No. 6, pp. 537-552.

Clarke, A., & Hermelin, F. Adult Imbeciles: Their Abilities and Trainability. The Lancer, 1955, 2, 337-339.

Corbett, B.A. & Abdullah, M. (2005) Video Modeling: Why Does It Work for Children with Autism? *Journal of Early and Intensive Behavior Intervention:* Volume 2, Issue No. 1, pp. 2-8.

Corbett, B.A. (2003). Video Modeling: A Window into the World of Autism. *The Behavior Analyst Today:* Volume 4, Issue No. 3.

Crosson, J. The Experimental Analysis of Vocational Behavior in Severely Retarded Males. Final Report, Grant No. 0EG32-47-0230-6024. Washington, D.C.: Department of Health, Education and Welfare, 1967.

D’Ateno, P., Mangiapanello, K., & Taylor, B.A. (2003). Using video modeling to teach complex play sequences to a preschooler with autism. *Journal of Positive Behavior Interventions, 5*, 5-11.

D'Ateno, P., Mangiapanello, K., & Taylor, B. A. (2003). Using Video Modeling to Teach Complex Play Sequences to a Preschooler with Autism. *Journal of Positive Behavior Interventions:* Volume 5, Issue No. 1, pp. 5-11.

Feuerstein, R. The Dynamic Assessment of Retarded Performers: The Learning Potential Assessment Device, Theory, Instruments and Techniques. Baltimore: University Park Press, 1979.

Gold, M. Research on the Vocational Habilitation of the Retarded: The Present, the Future. In N. R. Ellis (Ed.), International Review of Research in Mental Retardation (Vol. 6). New York: Academic Press, 1973.

Gold, M. Stimulus Factors in Skill Training of the Retarded on a Complex Assembly Task: Acquisition Transfer and Retention. American Journal of Mental Deficiency, 1972, 76, 517-526.

Gold, M., & Barclay, C. The Learning of Difficult Visual Discrimination by the Moderately and Severely Retarded. Mental Retardation, 1973, 11, 9-11.

Goldsmith, T.R. & LeBlanc, L.A. (2004) Use of Technology in Interventions for Children with Autism. *Journal of Early and Intensive Behavioral Intervention:* Volume 1, Issue No. 2, pp. 166-178.

Greenspan, S., & Love, P. F. (1997). Social intelligence and developmental disorder: Mental retardation, learning disabilities, and autism. In W. E. MacLean Jr. (Ed.), *Ellis' handbook of mental deficiency, psychological theory and research* (3rd ed., pp. 311-342). New York: Erlbaum.

Haymes, L. K. (1995). Using video modeling to improve conversation skills with autistic children: A comparison of obsessions and non-obsessions as topics. *Dissertation Abstract International: Section B: The Sciences and Engineering, 55*, 5556.

Haywood, H. C., Filler, J., Shifman, M., & Chatelanat, G. Behavioral Assessment in Mental Retardation. In P. McReynolds (ed.), Advances in Psychological Assessment (Vol. 3). Washington, D.C.: Jossey-Bass, 1975.

Heflin, J., & Simpson, R. (1998). Interventions for children and youth with autism: Prudent choices in a world of exaggerated claims and empty promises. Part II: Legal policy analysis and recommendations for selecting interventions and treatments. *Focus on Autism and Other Developmental Disabilities, 13*, 212-220.

Hine, J.F. & Wolery, M. (2006). Using Point-of-View Video Modeling to Teach Play to Preschoolers with Autism. *Topics in Early Childhood Special Education:* Volume 26, Issue No. 2, pp. 83–93.

Irvin, L. K., & Gersten, R. Trainee Performance Sample Technical Manual. New Berlin, WI: Ideal Developmental Labs, 1982.

Irvin, L. K., & Halpern, A. The Process Model of Diagnostic Assessment. In G. T. Ballamy, G. O'Connor, & O. Karan (Eds.), Vocational Rehabilitation of Severely Handicapped Persons: Contemporary Service Strategies. Baltimore: University Park Press, 1979.

Irvin, L. K., Gersten, R., et al. Assessing Vocational Skills of Severely Retarded Adults. American Journal of Mental Deficiency, 1981, 85(6), 631-638.

Lasater, M. W., & Brady, M. P. (1995). Effects of video self-modeling and feedback on task fluency: A home-based intervention. *Education and Treatment of Children, 18*, 389-407.

LeGrice, B., & Blampied, N. M. (1994). Training pupils with intellectual disability to operate educational technology using video prompting. *Education and Training in Mental Retardation and Developmental Disabilities, 29*, 321-330.

Lovett, D. L., & Harris, M. B. (1987). Important skills for adults with mental retardation: The client's point of view. *Mental Retardation, 25*, 351-356.

Neff, W. Vocational Assessment—Theories and Models. Journal of Rehabilitation, 1970, 36, 27-29.

Nikopoulos, C. K., & Keenan, M. (2003). Promoting social initiation children with autism using video modeling. *Behavioral Interventions, 18*, 87-108.

Nikopoulos, C.K., & Keenan, M. (2007). Using video modeling to teach complex social sequences to children with autism. *Journal of Autism and Development Disorders, 37*, 678-693.

Norman, J. M., Collins, B.C., & Schuster, J. W. (2001). Using an instructional package including video technology to teach self-help skills to elementary students with mental disabilities. *Journal of Special Education, 16*, 5-18.

Pierce, K., & Schreibman, L. (1994). Teaching daily living skills to children with autism in unsupervised settings through pictorial management. *Journal of Applied Behavior Analysis, 27*, 471-481.

Quill, K. A. (1997). Instructional considerations for young children with autism: The rationale for visually cued instruction. *Journal of Autism and Developmental Disabilities, 27*, 697-714.

Rehfeldt, R. A., Dahman, D., Young, A., Cherry, H., & Davis, P. (2003). Teaching a simple, meal preparation skill to adults with severe and mental retardation using video modeling. *Behavioral Intervention, 18*, 209-218.

Shipley-Benamou, R., Lutzker, J. R., & Taubman, M. (2002). Teaching daily living skills to children with autism through instructional video modeling. *Journal of Positive Behavior Interventions, 4*, 165-175.

Simpson, R. L. (2005). Evidence-based practices and students with autism spectrum disorders. *Focus on Autism and Other Developmental Disorders, 20(3)*, 140-149.

Smith, C., Williamson, R. & Siegel-Robertson, J. (2005). Implementing Technology to Teach Social Skills to Students with Multiple High-Incidence Disabilities. Unpublished University of Memphis research study, 11 pp.

Tate, B., & Barhoff, C. Training the Mentally Retarded in the Production of a Complex Product: A Demonstration of Work Potential. Exceptional Children. 1967, 33, 405-408.

Taylor, B. A., Levin, L., & Jasper, S. (1999). Increasing play-related statements in children with autism toward their siblings: Effects of video modeling. *Journal of Developmental and Physical Disabilities, 11*, 253-264.

Wert, B. Y., & Neisworth, J. T. (2003). Effects of video self-modeling on spontaneous requesting in children with autism. *Journal of Positive Behavior Interventions, 5*, 30-34.

Wert, B. Y., & Neisworth, J. T. (2003). Effects of Video Self-Modeling on Spontaneous Requesting in Children with Autism. *Journal of Positive Behavior Interventions:* Volume 5, Issue No. 1, pp. 30-34.

Williams, C., Wright, B., Callaghan, G., & Coughlan, B. (2002). Do Children with Autism Learn to Read More Readily by Computer Assisted Instruction or Traditional Book Methods? *Journal of Autism and Developmental Disabilities*, Volume 6, pp. 71-91.

Williams, P. Industrial Training and Remunerative Employment of the Profoundly Retarded. Journal of Mental Subnormality, 1967, 13, 14-23.

Wolfensberger, W. Vocational Preparation and Occupation. In A. Baumeister (Ed.), Mental Retardation. Chicago: Aldine, 1967.

Yell, M. L., & Drasgow, E. (2000). Litigating a free appropriate public education: The Lovas hearings and cases. *The Journal of Special Education, 33*, 205-214.