The Assessment of Basic Learning Abilities Test for Predicting Learning of Persons With Intellectual Disabilities

A Review

Garry L. Martin
Jennifer R. Thorsteinsson
University of Manitoba
C. T. Yu
Toby L. Martin
St. Amant
Tricia Vause
Brock University, St. Catharines, Ontario, Canada

The Assessment of Basic Learning Abilities (ABLA) Test uses standard prompting and reinforcement procedures to assess the ease or difficulty with which a testee is able to learn a simple imitation and five two-choice discriminations. The authors review studies that have examined performance of participants with developmental disabilities (DD) on the ABLA test to predict (a) performance on a variety of simple imitations and two-choice discriminations, (b) performance on three-choice and four-choice discriminations, (c) the relative efficacy of three presentation modes (objects vs. photographs vs. verbal descriptions) for assessing preferences, (d) compliance of adults with DD and children with and without DD, and (e) participants’ ability to learn to respond to the spoken names of pictures of common objects. Across all five types of studies, the predictive validity of the ABLA test has been very high.

Keywords: behavioral assessment; discrimination learning; intellectual disabilities assessment; basic learning abilities

Authors’ Note: Preparation of this article was supported by Grant MOP 6353 from the Canadian Institutes of Health Research. Please address correspondence to Garry L. Martin, University of Manitoba, 129 St. Paul’s College, 70 Dysart Road, Winnipeg, Manitoba, R3T 2M6; e-mail: gmartin@cc.umanitoba.ca.
Direct-care staff working with persons with moderate, severe, or profound mental retardation (MR, as defined in the *Diagnostic and Statistical Manual of Mental Disorders* [American Psychiatric Association, 2000]) often do not know why such individuals learn some tasks with relative ease while they learn other superficially similar tasks very slowly or not at all. For example, an adolescent with severe MR may be able to learn to restock a salad bar in a cafeteria but when asked for lettuce versus tomatoes is not able to retrieve the appropriate item. After observing these types of problems, Kerr, Meyerson, and Flora (1977) developed an assessment tool called the Assessment of Basic Learning Abilities (ABLA) test to assess an individual’s deficits in learning the prerequisite motor, visual, and auditory discriminations thought to underlie such tasks. The ABLA test is a dynamic assessment during which a teacher, using standardized prompting and reinforcement (SPR) procedures, attempts to teach a student to learn a simple imitation, a two-choice position discrimination, a two-choice visual discrimination, a two-choice visual-visual quasi-identity match-to-sample discrimination, a two-choice auditory discrimination, and a two-choice auditory-visual discrimination. One or more of these tasks appear to be required for a person to readily learn a large number of self-care, academic, prevocational, and vocational tasks in programs at facilities for individuals with MR (DeWiele & Martin, 1996). It is not surprising then, that, since the ABLA test was first published, a number of studies have examined the extent to which the performance of participants on the ABLA test is predictive of the ease or difficulty that participants experience in learning a variety of self-care, educational, and prevocational tasks. The purpose of the present article is to review that research.

The ABLA Test

Table 1 describes the six tasks, called levels, that make up the ABLA test. The test is conducted using a yellow can, a small yellow wooden cylinder, a red box, a small red wooden cube, and a white irregularly shaped piece of foam. Kerr et al. (1977) designed a set of testing rules so that the six minilearning tasks could be administered in 30 min or less with a testee. The testing rules involve three SPR procedures. First, a task is introduced with a demonstration, a guided practice trial, and an opportunity for a testee to respond independently. Then, scoring begins on that task. Second, every correct response is reinforced with praise and an edible (or some other preferred item). Third, errors are followed by an error correction procedure consisting of a demonstration, a guided practice trial, and an
Table 1
A Description of the ABLA Levels, Discriminations Required, and Sample Everyday Behaviors Requiring the Discriminations

<table>
<thead>
<tr>
<th>ABLA Levels</th>
<th>Discriminations</th>
<th>Sample Behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1: Imitation. A tester puts an object into a container and asks the testee to do likewise.</td>
<td>A simple imitation.</td>
<td>Children playing Follow-the-Leader.</td>
</tr>
<tr>
<td>Level 2: Position Discrimination. When a red box and a yellow can are presented in a fixed position, a testee is required to consistently place a piece of white foam in the container on the left when the tester says &quot;put it in.&quot;</td>
<td>A simultaneous visual discrimination with position, color, shape, and size as relevant visual cues.</td>
<td>Turning on the cold (vs. the hot) water tap.</td>
</tr>
<tr>
<td>Level 3: Visual Discrimination. When a red box and a yellow can are randomly presented in left-right positions, a testee is required to consistently place a piece of white foam in the yellow can when the tester says &quot;put it in.&quot;</td>
<td>A simultaneous visual discrimination with color, shape, and size as relevant visual cues.</td>
<td>Locating one’s coat from among other coats hung in a closet, with the coats in no fixed position.</td>
</tr>
<tr>
<td>Level 4: Match-to-Sample. When a yellow can and a red box are presented in random left-right positions and a testee is presented with a yellow cylinder or a red cube, he/she consistently places the cylinder in the yellow can and the cube in the red box.</td>
<td>A conditional visual-visual quasi-identity discrimination with color, shape, and size as relevant visual cues.</td>
<td>Sorting socks into pairs.</td>
</tr>
<tr>
<td>Level 5: Auditory Discrimination. When presented with a yellow can and a red box (in fixed positions), a testee is required to consistently place a piece of white foam in the appropriate container when the tester randomly says &quot;red box&quot; (in a high-pitched rapid fashion) or &quot;yellow can&quot; (in a low-pitched drawn-out fashion).</td>
<td>A conditional auditory-visual nonidentity discrimination, with pitch, pronunciation, and duration as relevant auditory cues and with position, color, shape, and size as relevant visual cues.</td>
<td>Responding to instructions to go left or right, to go to different rooms, or to open different drawers.</td>
</tr>
<tr>
<td>Level 6: Auditory-Visual Combined Discrimination. The same as Level 5 except that the left-right position of the containers is randomly alternated.</td>
<td>A conditional auditory-visual nonidentity discrimination, with the same auditory cues as Level 5 and with only color, shape, and size as relevant visual cues.</td>
<td>Responding appropriately to requests such as &quot;pass the salt&quot; versus “pass the pepper” when the salt and pepper shakers are in different places on the table from meal to meal.</td>
</tr>
</tbody>
</table>

Source: Reprinted with permission from Martin and Yu (2000).
Note: ABLA = Assessment of Basic Learning Abilities.
opportunity for an independent response. Testing continues on a particular level until the testee emits eight consecutive correct responses, not counting correct responses on error correction trials, or until a testee shows eight cumulative errors on that task. The level is considered “passed” if eight consecutive correct responses occur first and “failed” if eight cumulative errors occur first.

A Summary of Research on the ABLA Test

Research on the ABLA test was reviewed by Martin, Yu, and Vause (2004). That research indicates that (a) the ABLA tasks are ordered as listed in Table 1 such that most individuals who pass a certain level of discrimination also pass at lower levels of the hierarchy and those who fail a certain level of discrimination also fail at higher levels, and this has been demonstrated for persons with MR (Kerr et al., 1977; Martin, Yu, Quinn, & Patterson, 1983), children with Pervasive Developmental Disorder (PDD); Ward & Yu, 2000), and typically developing children (Casey & Kerr, 1977); (b) the ABLA test has very high test-retest and intertester reliability (Martin et al., 1983); (c) if a testee fails an ABLA level, then tasks requiring that ABLA level are very difficult to teach using the SPR procedures of the ABLA test (Meyerson, 1977; Stubbings & Martin, 1995; Wacker, Kerr, & Carroll, 1983; Wacker, Steil, & Greenebaum, 1983; Witt & Wacker, 1981; Yu & Martin, 1986); (d) mismatching of ABLA test level of clients to ABLA difficulty of training tasks causes aberrant behaviors (Vause, Martin, Cornick et al., 2000; Vause, Martin, & Yu, 1999); (e) direct-care staff with no knowledge of the ABLA test often mismatch the ABLA test level of clients and the ABLA difficulty level of training tasks (DeWiele & Martin, 1996; Vause, Martin, Cornick et al. 2000); and (f) the ABLA has concurrent validity with language, reading, and adaptive behavior assessments (Barker-Collo, Jamieson, & Boo, 1995; Casey & Kerr, 1977; Meyerson, 1977; Richards, Williams, & Follette, 2002; Vause, Martin, & Yu, 2000).

The degree of intellectual disability as measure by IQ is roughly associated with ABLA test performance, with 73% of persons with moderate MR, 35% of persons with severe MR, and no persons with profound MR able to learn ABLA Levels 5 and 6 (the auditory discriminations; Kerr et al., 1977). However, IQ tests provide a much more global measure of performance than the ABLA test, and IQ does not have the predictive validity of the ABLA test for the types of motor, visual, and auditory tasks examined to date (Kerr et al., 1977; Richards et al., 2002; Stubbings & Martin, 1995, 1998).
Overview of Five Types of Predictive Validity Studies on the ABLA Test

One type of predictive validity study evaluates the extent to which ABLA test results predict performance on a variety of simple imitations and two-choice discriminations. There are seven predictive validity studies of this type. The second type of study examined the extent to which ABLA test results predict learning performance on three-choice and four-choice discriminations. There are two studies of this type. The third type of study examined the extent to which ABLA test results predict the relative efficacy of three presentation modes (objects vs. photographs vs. verbal descriptions) for assessing preferences of persons with MR. There are three studies of this type. The fourth type of study examined the extent to which ABLA test results predict compliance to instructions alone (IA) versus instructions plus modeling plus guidance (IMG). There are two studies of this type. The fifth type of study examined the extent to which ABLA test results predict ability to learn spoken names of pictures of common objects. There is one study of this type. The research to be reviewed will be grouped according to these five categories of predictive validity studies.

The ABLA Test Predicts Learning Performance on Imitative and Two-Choice Discriminations

Studies in this category typically identified participants as being “at” each of the ABLA levels. For example, a participant who passed ABLA Levels 1, 2, and 3 and failed Levels 4, 5, and 6 would be considered to be “at” Level 3. Then, a variety of imitative and two-choice discrimination tasks, referred to as criterion tasks, were identified for each ABLA level. For example, learning to match a cup to another cup (rather than to a plate) and matching a plate to another plate (rather than to a cup) would be considered an ABLA Level 4 criterion task involving a visual-visual two-choice identity-matching task. After a variety of criterion tasks were identified, the researchers attempted, using the SPR procedures of the ABLA test, to teach each of the criterion tasks to each participant until one of the ABLA pass or fail standards of eight consecutive correct responses or eight cumulative errors had been met for each task. This enabled the authors of such studies to address two questions: (a) if a participant was at a particular level on the ABLA test (e.g., Level 3), then, following SPR training on criterion tasks at that or lower levels (i.e., at Levels 1, 2, or 3),
would the participant achieve the pass standard before meeting the failure standard on those criterion tasks, and (b) if a participant failed a particular level (e.g., Level 4), then, using SPR procedures to teach criterion tasks thought to be at the failed level or a higher level, would the client meet the failure standard before achieving the pass standard on those criterion tasks? To date, seven such studies have been published.

**Characteristics of Imitative and Two-Choice Criterion Tasks**

As stated previously, the ABLA test materials include a large yellow can, a small yellow cylinder, a large red box (but not as tall as the yellow can), a small red cube, and a white irregularly shaped piece of foam. The materials for the imitative and two-choice criterion tasks that have been studied can be divided into two main categories. The first category (referred to in this article as A-tasks) included containers and materials similar to the ABLA test. For example, one criterion A-task included a white box, a small white cube, a black can, a small black cylinder, and a beige piece of foam. Another criterion A-task included materials that differed only in terms of shape, such as a white box, a white cube, a white can, a white cylinder, and a white piece of foam.

The second category of criterion tasks (referred to as B-tasks) used materials commonly available in prevocational, educational, and everyday training tasks with persons with MR. For example, requiring a testee to imitate the tester in turning on a light by flipping a light switch would be a Level 1 B-task. Requiring a participant to place a coin into a stationary piggy bank (versus in an alternate stationary container) would be considered a Level 2 B-task. Requiring a testee to place a card into a card file versus a bowl when the bowl and file are in randomly alternating positions would be considered a Level 3 B-task. Requiring a participant to put a piece of white paper into a white envelope (rather than a brown envelope) and to put a piece of brown paper into a brown envelope (rather than a white envelope) would be considered a Level 4 B-task. A Level 5 B-task involved presenting a hat and sock to a testee in stable positions on a table and the tester randomly saying “touch hat” or “touch sock.” In the previous example, if the hat and sock were randomly placed in alternating left/right positions, then the task would be considered a Level 6 B-task. Table 2 summarizes the imitative and two-choice criterion A-tasks and B-tasks used in each of the seven predictive validity studies in this category.
Results With Imitative and Two-Choice Criterion Tasks

In each predictive validity study, the authors made two predictions: (a) that participants at a particular ABLA level would pass all criterion tasks at that and lower levels, and (b) that participants at a particular ABLA level would fail all criterion tasks above that level.

Participant characteristics and results. The results of the seven studies that examined imitative and two-choice tasks are presented in Table 3. Across six of the studies (excluding Condillac, 2002, in which the participants had a dual diagnosis of MR plus autism or PDD–not otherwise specified [NOS]), 77 participants had a diagnosis of MR with a range of profound to mild MR and with ages ranging from 4 to 55. Across those six studies, there were 23 participants who experienced criterion A-tasks, and
<table>
<thead>
<tr>
<th>Source</th>
<th>Total number</th>
<th>Age range (years)</th>
<th>Level of MR</th>
<th>ABLA Levels</th>
<th>Mean % of Predictions That Were Accurate</th>
</tr>
</thead>
</table>
| Tharinger, Schallert, & Kerr (1977)         | 11           | 4 to 14           | Mild to profound | 2, 3, 4, 6 | A-tasks: 84%  
B-tasks: 81% (tested on 6 of the 11 Ps) |
| Wacker, Kerr, & Carroll (1983)              | 12           | 19 to 55          | Moderate to profound | 2, 4, 6 | A-tasks: 96%  
B-tasks: 100% (tested on 9 of the 12 Ps) |
| Wacker, Steil, & Greenebaum (1983)          | 7            | 3 to 7            | Mild to severe | 2, 4, 6 | B-tasks: 90% |
| Stubbings & Martin (1995)                   | 9            | 7 to 35           | Mild to profound | 2, 4, 6 | B-tasks: 93% |
| Stubbings & Martin (1998)                   | 18           | 9 to 39           | Moderate to profound | 1, 2, 3, 4, 5, 6 | B-tasks: 90% |
| Condillac (2002)                            | 46           | 6 to 31           | Mild to profound, 38 with autistic disorder, 8 with PDD-NOS | 0, 1, 2, 3, 4, 6 | A-tasks: 95%  
B-tasks: 78% |
| Thorsteinsson et al. (2007)                 | 20           | 24 to 49          | Moderate to profound | 2, 3, 4, 6 | B-tasks: 94% |

Note: ABLA = Assessment of Basic Learning Abilities; PDD-NOS = Pervasive Developmental Disorder–not otherwise specified.
the mean accuracy of predictions about performance on the those tasks was 90%. This high level of predictive validity for A-tasks is not surprising in that such tasks involved materials very similar to the ABLA test but with containers of different colors and sizes. Prediction accuracy was also high, however, for B-tasks involving a variety of prevocational and educational materials. Across the six studies (excluding Condillac), 69 participants experienced criterion B-tasks, and the mean prediction accuracy about performance on these was 92%.

The study by Condillac (2002) included 46 participants with a dual diagnosis of MR and autistic disorder (or PDD-NOS for 8 participants). As indicated in Table 2, Condillac assessed the predictive validity on A-tasks involving similar but different colored containers and materials than the ABLA test, and on B-tasks using everyday materials. Mean prediction accuracy across both types of tasks was 86%. Mean accuracy on A-tasks was 95%, and mean accuracy on B-tasks was 78%. On the six everyday tasks, predictions were confirmed for 72% of the participants for the B-task at Level 1, 57% of the participants for the B-task at Level 2, 80% of the participants for the B-task at Level 3, 85% of the participants for the B-task at Level 4, and 87% of the participants for the B-task at Level 6. Although these results suggest that the predictive validity of the ABLA test for children with autism is somewhat lower than for persons with MR (who do not have a diagnosis of PDD), several limitations of the Condillac study should be noted. First, the study was limited by having only one everyday criterion task for each ABLA level. Second, the Level 1 criterion B-task, making a stirring motion in an empty cup with a spoon, may have been inconsistent with everyday experiences of the participants (i.e., stirring with a spoon only when there is something to stir), especially given that lack of imaginary play is one of the difficulties in children with PDD. This may explain why the mean correct predictions for the Level 1 B-task was only 72%. Third, the Level 2 criterion B-task required participants to put a cup in the upper right quadrant of a placemat above/beside a plate, which appears to be a four-choice discrimination rather than a two-choice discrimination. This may explain why the mean accuracy of predictions for the Level 2 B-task was only 57%.

Prediction accuracy by ABLA level of participants. Table 4 shows the performance on criterion tasks by the ABLA level of participants across all seven studies using imitative and two-choice criterion tasks. As can be seen in Table 4, the ABLA test was very accurate for predicting that participants would pass criterion tasks at or below their ABLA level as well as for
predicting that participants would fail criterion tasks that were above their ABLA level. The ABLA test was least accurate for predicting performance of participants at ABLA Level 3, the two-choice visual discrimination. This was due to the tendency of some participants at ABLA Level 3 to pass criterion tasks at ABLA Level 4, the visual-visual quasi-identity match-to-sample discrimination.

**Predictive accuracy of the ABLA test versus experienced staff.** Two studies compared the ABLA test’s predictive accuracy to the predictive accuracy of experienced staff. In one study, Stubbings and Martin (1998) compared three sets of predictions concerning the ability of 18 persons with severe or moderate MR to learn 12 training tasks, two tasks at each of the six ABLA levels. Predictions were made by (a) experienced instructors who had worked with the clients for a minimum of 8 months, (b) experienced instructors who did not know the clients personally but who interacted with them for 30 min, and (c) consideration of client performance on the ABLA test, each of whom were tested in less than 30 min. Before making predictions, caregivers were given a written description of each criterion task that included the materials required to perform each task, the position of each of the objects for each task, the instructions to be given to the participant to
perform the task, and the response required of the participant to perform each task correctly. Caregivers were also given a written description of the SPR procedures of the ABLA test to be used to teach each task to a participant. A caregiver was asked to predict whether a participant would learn the task, before failing it, using the SPR testing procedure and pass/fail standards of the ABLA test. After the predictions were made, the SPR procedures of the ABLA test were applied with each individual participant until each participant met the pass or fail standard on each of the 12 tasks about which predictions were made. The ABLA test results were significantly more accurate for predicting client performance than either group of experienced staff.

In another study, Thorsteinsson et al. (2007) compared two sets of predictions concerning the ability of 20 adults with profound, severe, or moderate MR to learn 15 everyday B-tasks. Predictions were made by (a) caregivers who had worked with the participants for a minimum of 24 months, and (b) consideration of participant performance on the ABLA test. The procedures for asking caregivers to make their predictions replicated those of Stubbings and Martin (1998). The ABLA test was significantly more accurate for predicting client performance than the caregivers.

Summary

Seven studies have examined performance on the ABLA test for predicting performance on imitative and two-choice criterion tasks. The studies included 123 participants with mild to profound MR, 46 of whom also had a diagnosis of autistic disorder or PDD-NOS. Across all studies, ABLA test results predicted learning performance on the criterion tasks with 89% accuracy. Moreover, two of the studies demonstrated that the ABLA test results of participants were significantly more accurate for predicting learning performance of clients than experienced staff who knew the clients well.

The ABLA Test Predicts Performance on Three-Choice and Four-Choice Tasks

As indicated previously, the ABLA test assesses the ease or difficulty with which a person is able to learn a simple imitation and five two-choice discriminations. But many everyday tasks involve three-choice, four-choice, or many-choice discriminations. Thus far, two studies have examined the predictive validity of the ABLA test for such tasks. Doan, Martin, Yu, and
Martin (in press) examined the ABLA’s ability to predict three-choice discrimination performance with 12 adults with moderate to profound MR and ranging in age from 23 to 49. The sample included three individuals who were at ABLA Level 2, three who were at ABLA Level 3, three who were at ABLA Level 4, and three who were at ABLA Level 6. Using the standard SPR procedures and pass/fail standards of the ABLA test, an attempt was made to teach each participant three three-choice tasks at their ABLA level and three three-choice tasks immediately above their ABLA level (except for participants at Level 6). The three-choice tasks used various materials and containers analogous to the ABLA test. For example, one of the three-choice Level 2 tasks required a participant to place a clothespin in the left-hand container when a red box, a gold can, and a rectangular green box were placed in a row in front of the participant. As another example, one of the Level 4 three-choice tasks required a participant to place a piece of paper (matched in color to an envelope) in either a large white envelope, a medium brown envelope, or a small blue envelope when all three were placed in front of the participant. The authors predicted that participants would pass three-choice tasks at their ABLA level and fail three-choice tasks above their ABLA level. The nine participants at ABLA Levels 2, 3, and 4 passed a mean of 1.78 three-choice tasks at their ABLA level and a mean of 0.11 three-choice tasks above their ABLA level, which was statistically significant. Overall, 71% of the predictions were confirmed.

Wacker, Kerr et al. (1983) examined the predictive validity of the ABLA test on four-choice discriminations with 12 adults with profound to moderate MR. The sample included four participants who were at ABLA Level 2, four who were at ABLA Level 4, and four who were at ABLA Level 6. Each of the four-choice discriminations were similar to the ABLA test discriminations except that the containers included a white box, a green box, a yellow box, and a red box, and when testing Level 4, the containers included a white cube, a green cube, a yellow cube, and a red cube. These materials were used to assess the clients on a Level 2 criterion task, a Level 4 criterion task, and a Level 6 criterion task. Participants were tested on the four-choice task at their ABLA level and the four-choice task immediately above their ABLA level (except for participants at Level 6). All participants passed the criterion task at their ABLA level, and the participants at ABLA Levels 2 and 4 failed the criterion tasks above their ABLA level. In other words, 100% of the predictions were confirmed.
The ABLA Test Predicts the Relative Efficacy of Three Presentation Modes for Assessing Preferences

It is widely accepted that the opportunity to make choices is important for improving the quality of life of persons with MR. Choices may be presented to persons with MR using objects, pictures of objects, or spoken descriptions of objects. There are obvious practical advantages for using pictures over objects and for using spoken words over pictures. However, the presentation mode (objects vs. pictures vs. spoken words) during preference assessments must be appropriate for the discrimination ability of the participant. Thus, a participant with few or no language skills might be able to select their preferred food items consistently when the actual items are presented but not when the items are described verbally. Consideration of the levels of the ABLA test suggests that a participant who is at ABLA Level 3 (the two-choice visual discrimination) would be able to make consistent choices when objects are presented but not when photographs (which seem to involve ABLA Level 4, a visual-visual match-to-sample discrimination) and not when spoken words (which would require ABLA Level 6, an auditory discrimination) are used. Participants at ABLA Level 4 would be expected to make consistent choices with objects or photographs but not with spoken words. Participants at ABLA Level 6 would be expected to make consistent choices when any of the three presentation modes are used. Three studies have investigated the ABLA test for predicting the relative efficacy of the three presentation modes (objects vs. pictures vs. spoken words) for assessing the consistency of preferences of persons with MR.

Conyers et al. (2002) assessed the consistency of choices presented to nine individuals with profound to moderate MR using a two-choice paired-stimulus preference assessment format. A preferred and a nonpreferred item were initially identified for each participant using actual food items and then using small manipulable nonfood items. Then, over a number of trials, each pair of items was presented to each participant in three conditions (actual items, pictures of the items, and spoken name presentation) using a reversal design. When choices were presented in a modality that was comparable to their ABLA level (or lower), the participants’ ability to make consistent choices with food and nonfood items were predicted with 94% accuracy by their performance on the ABLA test. When choices were presented in a modality above their ABLA level, their choices were at approximately chance level. This study was replicated by DeVries et al. (2005) but using leisure activities instead of easily manipulable food and
nonfood items. Leisure activities included such things as hitting a golf ball with a putter, listening to music, painting, and watering plants. Participants were nine adults with moderate to profound MR. During the presentation of choices, a participant was presented with either an object associated with the leisure activity (such as a golf club) or a photograph of someone performing the activity or a verbal description. The procedures were similar to that described for Conyers et al. but using an alternating-treatments design (tangible, pictorial, and verbal stimuli). For eight of the nine participants, they were able to make consistent choices in the modalities comparable to their ABLA level (or lower), with a mean accuracy of 94%. When choices were presented in a modality above their ABLA level, their choices were at approximately chance level. The results were less consistent for the ninth participant.

Reyer and Sturmey (2006) replicated the above two studies, using a preference assessment for work tasks. Participants were nine adults with profound to mild MR, including three at ABLA Level 2, three at Level 4, and three at Level 6. For five of the participants (two at Level 2, one at Level 4, and two at Level 6), their consistent choice of their preferred work task was predicted by their performance on the ABLA test. For four of the participants, the ABLA test did not predict consistency of choices. One difference in participants between this and the previous two studies is that this study included three participants at ABLA Level 2 (not ABLA Level 3) whereas the previous two studies included three participants at ABLA Level 3. All three studies included three participants at ABLA Level 4 and three participants at ABLA Level 6.

Across the three studies, the ABLA test predicted, very accurately, the relative efficacy of three presentation modes (objects, pictures, and spoken words) for assessing preferences for 22 of 27 participants.

The ABLA Test Predicts Compliance

Noncompliance is a common problem exhibited by children with MR (Walker, 1993). In a sample of children with MR interacting with their caregivers, Berry et al. (2003) found that noncompliance to instructions averaged 51% across participants even though caregiver reinforcement of compliance was high. Considering that responding to instructions requires an auditory-visual discrimination (ABLA Levels 5 and 6) and responding to instructions plus modeling and/or gestures requires a visual discrimination (ABLA Levels 3 and 4), two studies examined whether performance on the ABLA test would predict compliance to IA versus instructions with...
modeling and/or gestures. In a sample of 35 adults with MR, 18 of whom were at ABLA Levels 2 through 4 and 14 of whom were at ABLA Levels 5 and 6, participants were instructed to complete three tasks under conditions of IA or instructions plus gestures (IG). The group that passed the visual ABLA levels responded significantly more frequently to IA than to IG, whereas the participants who passed the auditory ABLA levels responded equally well to IA and IG.

In a systematic replication of LaForce and Feldman (2000), Hiebert, Martin, Yu, Thorsteinsson, and Martin (2007) assessed compliance of a group of children to IA and instructions plus modeling plus guidance (IMG) administered by their caregivers. The results indicated that (a) children at ABLA Level 6 performed significantly better on IA than those at ABLA Level 3 and/or 4, (b) children at ABLA Level 3 and 4 performed significantly better when given IMG than IA, and (c) children at ABLA Level 6 performed approximately equally well to both IA and IMG. These results were consistent across all of the children combined and for the two subgroups of children without disabilities and children with MR. These results provide important information to caregivers on how best to instruct their children in an effort to increase their compliance.

The ABLA Test Predicts Object Name Recognition

As indicated previously, research has shown that performance on the ABLA test correlates with language assessments. Verbeke, Martin, Yu, and Martin (2007) examined whether performance on ABLA Level 6 might predict the ability of persons with severe MR to recognize spoken names of pictures of common objects. Five participants who were at ABLA Level 6 and five who were at ABLA Level 4 were each tested on five two-choice discriminations that required them to point to pictures of common objects after hearing their names. The testing format was essentially the same as the SPR procedures for testing ABLA Level 6. For a pair of pictures, such as a picture of a cup and a plate, the pictures were randomly placed in left-right positions across trials. On each trial, the experimenter randomly said the name of one of the two pictures (e.g., either “cup” or “plate”), and the correct response was to point to the appropriate picture. Trials continued on a pair of pictures until either the ABLA pass or fail criterion was met. Four of the five participants who had failed ABLA Level 6 failed all of the receptive name recognition tasks. All five participants who had passed ABLA Level 6 passed all of the name recognition tasks. For the 50 predictions (5 Pairs of Pictures × 10 Participants), 90% of them were accurate.
Areas for Future Research

The ABLA appears to be a very useful tool for matching the learning ability of individuals with MR with the difficulty of various training tasks. Regarding the predictive validity of the ABLA test, however, additional research is needed in several areas. First, as indicated previously, only one study (Condillac, 2002) has thus far examined the predictive validity of the ABLA test for children with PDD, and that study was limited in that it included only one everyday criterion task for each ABLA level. Replication of that study with a greater variety of everyday criterion tasks at each of the ABLA levels is needed.

Only two studies have examined the predictive validity of the ABLA test for three- and four-choice discriminations as criterion tasks, and those two studies each had a small number of participants. Additional research is needed in this area.

Across all of the studies reviewed in this article, 241 participants were included. Among those, only 2 participants were assessed at ABLA Level 5. Stated differently, as indicated in a previous review (Martin & Yu, 2000), the great majority of participants who pass ABLA Level 5 also pass ABLA Level 6. DeWiele and Martin (1998), therefore, in the preparation of a self-instructional manual for direct-care staff, recommended that Level 5 of the ABLA test be deleted. Concerning a possible replacement task for ABLA Level 5, consideration should be given to a visual-visual nonidentity matching (VVNM) discrimination. Many functional everyday tasks involve relating two or more physically dissimilar stimuli, such as matching a pencil to a piece of paper or a sock to a shoe. Sakko, Martin, Vause, Martin, and Yu (2004) assessed a prototype VVNM task for possible addition to the ABLA test. To assess their prototype VVNM task, like ABLA Level 4, a red box and a yellow can are randomly placed in left-right positions across trials. A testee is randomly presented with either a silver-colored piece of wood shaped into the capital-lettered word “BOX” (to be placed in the red box) or a purple-colored piece of wood shaped into the upper and lowercase word “Can” (to be placed in the yellow can). All test procedures and pass/fail standards are those of the ABLA test described previously. In their research involving eight participants at ABLA Level 3, nine participants at ABLA Level 4, and six participants at ABLA Level 6, the results indicated that the VVNM prototype task fell between ABLA Levels 4 and 6 in difficulty. The prototype task also demonstrated high test-retest reliability. Sakko et al. also assessed the predictive validity of their VVNM prototype task by assessing all participants on five sets of
VVNM criterion tasks. The 12 individuals who failed the VVNM prototype task failed 87% of the criterion tasks. The 11 participants who passed the VVNM prototype task passed 98% of the criterion tasks, yielding an overall 92% of the predictions confirmed. Although these results suggest that the VVNM prototype task is a potentially useful replacement for ABLA Level 5 on the ABLA test, replication of the findings of Sakko et al. is needed with additional participants.

Are there additional tasks (i.e., levels) that are beyond ABLA Level 6 that might make the ABLA test useful for a greater number of individuals with intellectual disabilities? Considering that the ability to imitate sounds is one of the first expressive language skills that children acquire and considering that the ability to recognize that two sounds are the same is part of accurate vocal imitation, two prototype auditory matching tasks appear to be good candidates. An auditory-auditory identity matching (AAIM) prototype task requires a tester to say a word (e.g., “pen”) on some trials and a different word (e.g., “block”) on other trials, while one assistant says “pen” and a second assistant says “block.” The assistants randomly alternate speaking order and the word spoken. The testee must point to the assistant who spoke the word that matched that of the tester. An auditory-auditory nonidentity matching (AANM) prototype task requires the tester to say “ball” on some trials and “ice” on other trials. Two assistants say either “field” or “rink.” The testee must learn to point to the assistant who says “rink” when the tester says “ice” and to point to the assistant who says “field” when the tester says “ball.” During testing, the standard SPR and pass/fail criteria of the ABLA test are used on both the AAIM and AANM prototype tasks. Thus far, research indicates that (a) AAIM is more difficult than ABLA Level 6 (Harapiak, Martin, & Yu, 1999; Vause, Martin, Cornick et al., 2000), (b) AANM is more difficult than AAIM (Harapiak et al., 1999; Vause, Martin, Cornick et al., 2000), and (c) performance on the AAIM and AANM prototype tasks has good predictive validity for performance on similar tasks (Vause, Harapiak, Martin, & Yu, 2003). Moreover, the addition of the auditory matching tasks to the ABLA test increases the correlation of the ABLA test with communication ability (Vause, Martin, Cornick et al., 2000). Another study examined the relationship between performance on the ABLA test, AAIM, and AANM and the ability to perform specific verbal operants, namely, echoics, tacts, and mands (Marion et al., 2003). It was found that individuals who passed ABLA Level 6 performed better on the test of the three verbal operants than those unable to perform this discrimination, and individuals who passed ABLA Level 6 plus the two auditory matching tasks performed better on the test of the three verbal
operants than those unable to pass the auditory matching tasks. Additional research is needed to confirm these findings. Additional research might also examine whether AAIM and AANM facilitates subsequent teaching of echoics, tacts, and mands to children with autism and persons with MR.

References


Garry L. Martin is a distinguished professor of psychology at the University of Manitoba. He is internationally known for his books and journal articles on various areas of behavior modification and in particular in the areas of developmental disabilities and sport psychology.

Jennifer R. Thorsteinsson is a graduate of the Department of Psychology at the University of Manitoba and is a research assistant at the St. Amant Research Centre in Manitoba. Her research interests include behavioral assessment and training with persons with developmental disabilities.

C. T. (Dickie) Yu is the research director of the St. Amant Research Centre and an associate professor of psychology at the University of Manitoba. He is internationally known for his research and publications on behavioral assessment and treatment of persons with developmental disabilities, including children with autism.

Toby L. Martin is the research program manager at the St. Amant Research Centre and a PhD candidate in the Department of Psychology at the University of Manitoba. His research interests include behavioral assessment and training with persons with developmental disabilities.

Tricia Vause is an assistant professor of clinical psychology at Brock University. Her research interests include behavioral assessment and training with persons with developmental disabilities, including children with autism.