Journal on Developmental Disabilities
Le journal sur les handicaps du développement
JoDD
Volume 17, Number 1, 2011

Does the Assessment of Basic Learning Abilities Predict Four-Choice Discrimination Learning for Persons with Developmental Disabilities?

This research was supported by Grant MOP6353 from the Canadian Institutes of Health Research.

Authors

Lee F. MacPherson,[1]
Christine G. P. Sousa,[1]
Kendra M. Thomson,[1]
Toby L. Martin,[2]
Sandra Salem,[1]
Garry L. Martin,[3]
C. T. Yu[2]

[1] Department of Psychology, University of Manitoba, Winnipeg, MB

[2] St. Amant Research Centre, 440 River Road, Winnipeg, MB

[3] St. Paul's College, University of Manitoba, Winnipeg, MB

Correspondence

gmartin@cc.umanitoba.ca
or Garry L. Martin,
Rm. 129 St Paul's College,
University of Manitoba,
Winnipeg, MB R3T 2M6.

Keywords

Assessment of Basic Learning Abilities (ABLA), developmental disabilities, predicting learning performance, learning four-choice discriminations

Abstract

The Assessment of Basic Learning Abilities (ABLA) measures the ease or difficulty with which persons with developmental disabilities are able to learn a simple imitation and 5 two-choice discrimination tasks (Kerr, Meyerson, & Flora, 1977). The pass/fail performance of clients on the ABLA predicts their ability to learn a variety of two-choice training tasks (Martin, Thorsteinsson, Yu, Martin, & Vause, 2008). The current study assessed whether pass/fail performance of clients on the ABLA predicted their learning of four-choice tasks, and whether those predictions were as accurate as direct-care staff who had worked with the participants for at least three months. Participants passed significantly more four-choice tasks at their ABLA level than four-choice tasks immediately above their highest passed ABLA level, supporting the predictive validity of the ABLA. Staff predictions were slightly more accurate than ABLA predictions, although the difference was not statistically significant. Implications of these findings are discussed.

Why is it that some people with developmental disabilities can learn to perform some tasks with ease, but struggle with other tasks of seemingly similar difficulty? This question motivated Kerr, Meyerson, and Flora (1977) to develop the Assessment of Basic Learning Abilities (ABLA) test. The ABLA uses standard prompting and reinforcement procedures to evaluate whether a testee is able to learn each of six tasks, called *levels*, including a simple imitation (Level 1) and 5 two-choice discriminations. Training trials on a task are conducted until a testee reaches a pass standard of eight consecutive correct responses or a fail standard of eight cumulative errors–whichever comes first. When a participant passes one level but fails the next, they are considered to be "at" the highest level that they passed. ABLA performance accurately predicts success at a variety of two-choice training tasks (Martin, Thorsteinsson, Yu, Martin, & Vause, 2008). In this study, we examined the predictive validity of the ABLA test for participants to learn a sample of four-choice tasks.

The ABLA Test

Level 1, a simple imitation is the easiest of the six levels; a testee must drop a piece of foam into a can on request, immediately after the experimenter models the same behaviour. Level 2, labelled a position discrimination, requires a testee

to place a piece of foam into the container on the left, when a yellow can and a red box are placed in fixed left-right positions. This ability should generalize to persons who pass Level 2 should also be able to perform tasks such as being able to turn on the cold (vs. hot) water tap. Level 3, labelled a visual discrimination, requires a testee to place a piece of foam into the yellow can when the can and the red box are randomly alternated in left-right positions. A testee who passes Level 3 should be able to locate their coat among other coats in a closet, no matter where it is placed. Level 4, labelled a visual match-to-sample discrimination, requires a testee to place a red cube or a yellow cylinder into the corresponding red box or yellow can, depending on whether the cube or cylinder are presented to the individual. The containers in this task are also randomly alternated in left-right positions. This ability should generalize to the individual being able to perform a task such as sorting socks into matching pairs. Level 5, labelled an auditory discrimination requires that the red box and yellow can are placed in fixed left-right positions, and the testee must place a piece of foam into the box or can depending on whether the experimenter says "yellow can" very slowly and in a low pitch or "red box" more quickly in a higher pitch. This ability should generalize to an individual being able to respond to instructions to go right or left. Level 6, labelled an auditory-visual combined discrimination, involves the same procedure as in Level 5, except that the can and box are randomly alternated in right-left positions. A more detailed description of the types of discriminations required at each ABLA level is presented in Table 1. It has been found that people who can pass Level 5 most often also pass Level 6 (Martin & Yu, 2000). Therefore, Level 5 is no longer tested in some clinical settings.

Two-, Three-, and Four-Choice Discriminations

ABLA performance accurately predicts success at training tasks that involve two-choice discriminations such as those described above (Martin, et al., 2008), suggesting that the ABLA is an important tool for informing those who work with persons with developmental disabilities in a clinical or educational context of the learning ability of the clients. However, many everyday tasks presented to persons with developmental disabilities involve three-choice, four-choice, and many-choice discriminations. Doan, Martin, Yu and Martin (2007) examined the effectiveness of the ABLA for predicting participants' ability to perform three-choice discrimination tasks. Participants included nine adults with moderate-to-profound intellectual disabilities; three participants at each of ABLA Levels 2, 3, and 4. Using the standard test-

Table 1. A Description of theTypes of Discriminations Required for the ABLA Levels (From Martin & Yu, 2000).

ABLA levels	*Types of discriminations*
Level 1, imitation	A simple imitation
Level 2, position discrimination	A simultaneous visual discrimination with position, color, shape, and size as relevant cues
Level 3, visual Discrimination	A simultaneous visual discrimination with color, shape, and size as relevant cues
Level 4, match-to-sample discrimination	A conditional visual-visual identity discrimination with color, shape, and size as relevant cues
Level 5, auditory discrimination	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
Level 6, auditory-visual discrimination	A conditional auditory-visual discrimination, with the same auditory cues as Level 5, and with only color, shape, and size as relevant visual cues

ing procedures and pass/fail standards of the ABLA test, they attempted to teach each participant 3 three-choice criterion tasks at their ABLA level, and 3 three-choice criterion tasks immediately above their level. Doan, Martin, Yu and Martin (2007) predicted that participants would pass the three-choice tasks at their level, and fail the three-choice tasks above their level. Overall, 78% of their predictions were confirmed.

The purpose of the current study was to extend the research done by Doan et al. (2007) by evaluating whether participants' ABLA pass/fail performance would predict their performance on a sample of four-choice tasks, and whether those predictions were as accurate as staff members' predictions based on their experience with the participants.

Method

Participants and Settings

This research was approved by the Psychology/Sociology Research Ethics Board of the University of Manitoba. Twelve adults with developmental disabilities were recruited from St. Amant, a residential and community resource facility for individuals with developmental disabilities in Winnipeg, Manitoba, Canada. Participants who had not been assessed on the standard ABLA test during the three months before the study began were re-assessed to confirm their ABLA level. There were three participants who scored at ABLA Level 2, and three at Level 3, and six at Level 4. The mean age of the participants was 43 years, and there were six males and six females in total (see Table 2).

Twelve direct-care staff members, one for each participant and who had at least 3 months experience working with the respective participants, were also recruited from St. Amant to provide predictions of each participant's performance on the four-choice criterion tasks.

Research sessions took place in a quiet testing room at St. Amant. At least one of the experimenters (first and second authors) was present at every session and a trained observer was present at most of the sessions for each participant to conduct reliability checks (described below).

Table 2. Results for Four-Choice Criterion Tasks

ABLA level	*Participant number*	*Functioning level*	*Age*	*Sex*	*Tasks at participant's level*			*Tasks above participant's level*		
					1	*2*	*3*	*1*	*2*	*3*
2	1	Severe	48	F	F	F	P	P	F	F
	2	Severe	39	M	P	P	P	F	F	F
	3	Severe	41	F	F	F	F	F	F	F
3	4	Profound	32	M	P	F	F	F	F	F
	5	Profound	45	M	P	F	F	F	P	F
	6	Severe	45	F	P	F	F	F	F	N/A
4	7	Severe	40	F	P	P	P	F	F	F
	8	Profound	43	F	F	P	P	F	F	F
	9	Severe	40	F	P	P	P	F	F	F
	10	Severe	53	M	P	P	P	F	F	F
	11	Severe	47	M	F	F	F	F	F	F
	12	Severe	39	M	F	P	P	F	F	P

Note. P = pass, F = fail.

Materials

Before conducting any research sessions with a participant, a pre-assessment questionnaire was given to a direct-care staff member who had worked with the respective participant for at least three months. This questionnaire gathered pertinent information about the participant, such as dietary restrictions and interests. The questionnaire also contained a description of the teaching procedure and materials for each four-choice criterion task followed by two phrases, "PASS" and "NOT PASS." A staff member was instructed to read the summary descriptions, and then circle one of the phrases to indicate what outcome he/she predicted regarding the client's performance.

Standard ABLA (Levels 2, 3, and 4). Standard ABLA test materials (i.e., to assess ABLA Levels 1 through 6) consisted of two containers and three manipulanda. The containers included a red box measuring 14 cm × 14 cm × 10 cm, and a yellow can measuring 15 cm in diameter and 17 cm in height. The manipulandum for Levels 2 and 3 was an irregularly-shaped piece of white foam measuring approximately 4.5 cm × 4.5 cm × 4.5 cm. For Level 4, a yellow cylinder measuring 9 cm long and 4 cm in diameter, and a red cube measuring 5 cm × 5 cm × 5 cm were used as the manipulanda.

Analogue four-choice ABLA tasks. The materials for the four-choice tasks are presented in Table 3. One of the four-choice tasks for each level in the current study was analogous (in procedure and materials) to the standard ABLA Levels 2, 3, 4, and 6 tasks, except with additional stimuli (i.e., four containers). The other 2 four-choice tasks at each level included practical stimuli (e.g., cutlery, clothing, garbage receptacles, etc.) relevant to the participants' everyday functioning. Each four-choice task was designed to be administered like one of the ABLA levels, differing only due to physical variations in stimuli (e.g., placing a flash card in one of four envelopes of varying size and color, versus inserting a crumpled piece of paper into one of four garbage receptacles of varying colors).

Procedure

Standard and analogue ABLA assessments. During testing of both the standard two-choice ABLA levels and analogue four-choice tasks, the experimenter sat across the table from the participant. First, the experimenter presented various reinforcers (usually edibles, depending on the preferences indicated from the pre-assessment questionnaires). Next, the experimenter asked the participant to "pick one." The item that the participant chose was the item with which he/she was reinforced following correct independent responses. Testing of a participant on an ABLA level or a criterion task followed the standardized ABLA test procedures as described by Martin and Yu (2000). Testing of an ABLA level or criterion task began with the experimenter giving the testee a demonstration trial (i.e., modelling the correct behaviour), a guided trial, and then an opportunity for an independent response. For the ABLA task and the criterion tasks at Levels 4 and 6, these three trials were repeated for each correct option. Recording of responses commenced after the client's first correct independent response for tasks at Levels 2 and 3, and after all correct independent responses for tasks at Levels 4 and 6. Correct independent responses were always followed by an edible (e.g., a small piece of a cookie) and verbal praise (e.g., "good job"), except in cases when the independent response followed the initial guided trial or an error correction trial, in which case the testee was praised, but no edible was given. Incorrect responses were always followed by a 3-component error correction procedure (i.e., a demonstration trial, a guided trial, and an opportunity for an independent response). If the participant responded incorrectly on an opportunity for an independent response on the error correction procedure, the error correction procedure was repeated, and the error was counted toward the fail criterion (eight cumulative errors for both standard ABLA and four-choice tasks). A correct independent response made during the error correction procedure did not count toward the pass criterion (eight consecutive correct responses for a standard ABLA task but only four consecutive correct responses for the four-choice analogue tasks). Trials for each task were presented until the person met the pass criterion or the fail criterion, whichever came first. Assuming responses are independent, the

Table 3. Four-ChoiceTask Materials

ABLA Level	*Task*	*Materials*
2	1	Yellow can, red box, triangular prism-shaped blue container, clover-shaped green container and a piece of white foam.
	2	Four envelopes of varying size, varying in colour (yellow, pink, blue, white), and a small cardboard flashcard.
	3	Four garbage receptacles of varying colours (green, red, blue, white) but the same size and shape, and the white bin has a plastic bag inside it; and a crumpled up piece of paper.
3	1	Yellow can, red box, triangular prism-shaped blue container, clover-shaped green container and a piece of white foam.
	2	Four books of varying colours and size, and a small cardboard flashcard.
	3	Four colourful DVD cases varying in colour (yellow, pink, blue, green) and a blank disk.
4	1	Yellow can, red box, triangular prism-shaped blue container, clover-shaped green container and smaller matching manipulanda for each container.
	2	Knife, fork, spoon, and napkin, each varying in colour (gold, silver, blue, white), and four short identical containers.
	3	Identical/matching socks, slippers, gloves, and toque, each varying in colour (grey, pink, black, green).
6	1	Round yellow can, square red box, triangular prism-shaped blue container, clover-shaped green container, and a piece of white foam. Auditory stimuli: "yellow can" (L), "red box" (H), "blue triangle" (N), "green clover" (C).
	2	Plastic ketchup bottle, plastic mustard bottle, plastic relish bottle, and plastic mayonnaise jar. Auditory stimuli: "ketchup" (H), "mustard" (L), "relish" (N), and "mayonnaise" (C).
	3	Plastic dog figurine, plastic cat figurine, plastic horse figurine, and plastic bird figurine. Auditory stimuli: "dog" (L), "cat" (H), "horse" (N), and "bird" (C).

Note. H= high pitch and fast, L= low pitch and slow, N= normal pitch/speed, C= computerized voice played via iPod.

probability of passing an ABLA level purely by chance is approximately 0.03 (Doan et al., 2007). This probability applies to the passing criteria of the standard ABLA and the four-choice tasks in the present study. By doubling the amount of stimuli on each four-choice task as compared to a standard ABLA task, only half as many correct responses were required in order to pass; however, the failure criterion remained at eight incorrect responses for both each four-choice task and each standard ABLA task.

Each participant was tested on a total of 6 four-choice tasks: three tasks *at* their ABLA level, and three tasks immediately *above* their ABLA level. The 3 four-choice tasks at each level included one task which was analogous to the ABLA (differing only in extra stimuli), and two tasks that used everyday materials. The position of the stimuli remained stationary for Level 2 tasks, whereas the position of the stimuli alternated from trial to trial for every other level. The positions of stimuli for tasks above Level 2 were randomly selected before testing began, based on two conditions: At least one stimulus had to

change after each trial, and individual stimuli could not remain in the same position for more than two successive trials. For each standard ABLA task and each four-choice task, the experimenter gave the manipulandum to the participant at the start of each trial and asked, "Where does it go?" The correct response for Level 2 was placing the manipulandum in or on (depending on the materials of the task) the left-most stimulus. The correct response for tasks at Levels 3, 4, and 6 was placing the manipulandum in or on the pre-designated correct stimulus.

Staff prediction assessments. Staff questionnaires were given to the unit coordinators at St Amant, who then administered them to experienced staff members as they saw fit, according to their schedules and perceived likelihood of responsible completion. Based on their experience with the participants, the staff members were asked to indicate whether they believed the specified client would pass or fail each task. The completion of the questionnaires was not monitored by the experimenter, and included an address to which the staff members were able to submit the completed questionnaires while remaining anonymous. The results of the questionnaires remained in sealed envelopes until all testing of participants was complete.

Reliability assessments. Interobserver agreement (IOA) assessments were conducted across all phases of the study and across all participants. During an IOA assessment, the tester and a trained observer independently recorded the participant's responses for each trial. A trial was scored as an agreement if the tester and the observer recorded the same response for a trial, and a disagreement if they did not record the same response for a trial. An IOA score for a session was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100% (Martin & Pear, 2011). IOA assessments were conducted on 59% of all sessions and averaged 99.85%.

Procedural integrity (PI) assessments on the experimenter's behaviour were conducted across all phases of the study and across all participants. An observer used a checklist to record whether the experimenter did the following on each trial: (a) placed the containers and stimuli in the correct positions; (b) gave the correct instruction to the participant; (c) delivered the error correction procedure immediately and accurately after an error; and (d) presented reinforcement immediately following a correct response. The observer recorded how many of these steps the tester performed correctly on a trial. The procedural integrity score was calculated for a session by dividing the number of steps performed correctly by the total number of steps and multiplying by 100%. Procedural integrity checks were performed on 59% of all sessions and the mean procedural integrity score was 99.85%.

Results

The 12 participants passed 55.6% (*SD* = 38.49) of the 3 four-choice criterion tasks at their ABLA levels, and only 8.3% (*SD* = 15.07) of the 3 four-choice criterion tasks above their ABLA levels. A paired sample t-test was used to compare participants' performance on the four-choice tasks at and above their ABLA levels. The difference was statistically significant, $t(11) = 3.74$, $p = .002$, one-tailed. The three Level 2 and three Level 3 participants passed 38.89% of the tasks at their respective levels, compared to the six Level 4 participants who passed 72.22% of the tasks at their level. See Figure 1 for the percentage of tasks passed at and above the participants' levels.

The mean ABLA prediction accuracy score was 73.06% (*SD* = 22.13) and the mean staff prediction accuracy score was 76.39% (*SD* = 14.89). The difference was not statistically significant, $t(11) = -.325$, $p = 0.376$, one-tailed. See Figure 2 for a comparison of correct ABLA predictions vs. correct predictions made by staff.

Discussion

The results supported the ABLA's predictive validity to the extent that participants passed significantly more four-choice tasks *at* their ABLA level than four-choice tasks immediately *above* their highest passed ABLA level. These results are consistent with the findings of Doan et al. (2007), in which three-choice criterion tasks were used. Since participants in this study passed only slightly more than half of all four-choice tasks at their ABLA level, one should not predict with confidence that a student will be able to perform a four-choice

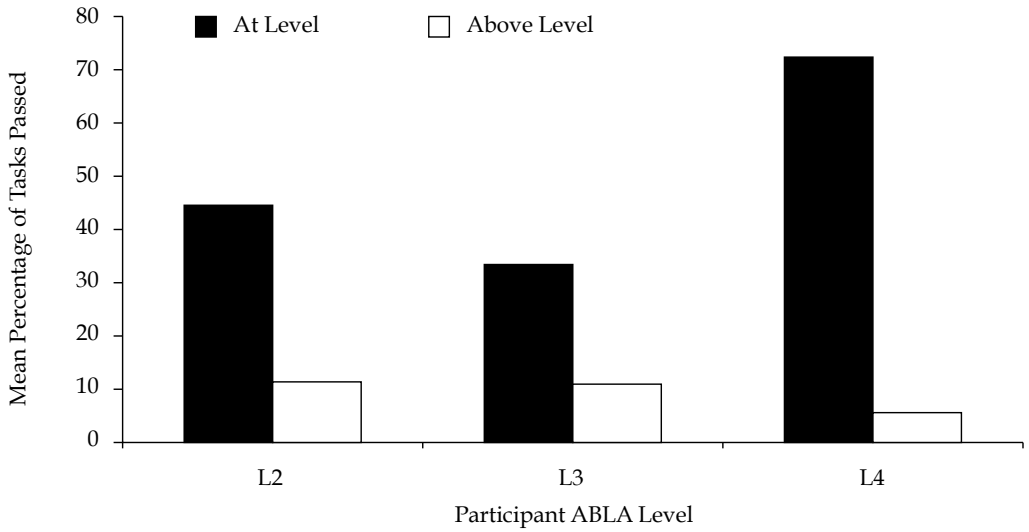

Figure 1. Mean percentage of four-choice tasks passed at participants' abla levels versus four-choice tasks immediately above participants' abla levels, for participants at abla levels 2, 3, and 4.

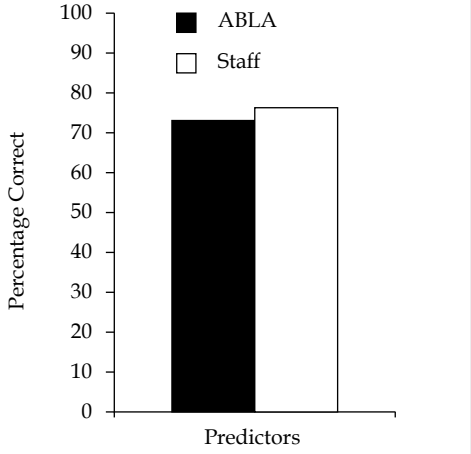

Figure 2. Percentage of correct predictions that were made based on abla performance vs. The direct-care staff members

discrimination that matches his or her ABLA level. Predictions that a student will not be able to perform a four-choice discrimination above his or her ABLA level may be made with much greater confidence.

Our findings differed from those of Doan et al. in that the three Level 2 participants in Doan et al. passed 100% of the tasks at their level, whereas our Level 2 participants passed only 44% of the tasks at their level. A potential reason for Level 2 participants' lower performance in the present study could be due to the participants' being at a lower level of intellectual functioning relative to the Level 2 individuals studied by Doan et al. This speculation is supported in part by the experimenters' subjective experience with the participants in testing situations, and in part from the results of the staff questionnaires. That is, the staff members predicted that the Level 2 participants would uniformly fail all of the tasks presented. It is conceivable that the ABLA may not take into account the variance in learning ability within a particular ABLA level of functioning, whereas a direct-care staff member who has worked with a participant for at least three months may have additional experiences and information to make such predictions.

Since prior research has found that the ABLA is able to predict standard two-choice performance more accurately than experienced staff (Thorsteinsson, Martin, Spevack, Martin, & Lee, 2007), we hypothesized that the ABLA would have greater predictive validity than the direct-care staff in the present study. Results indicated otherwise, in that the direct-care staff members were slightly (not statistically significantly) better predictors than the ABLA. In the context of the larger body of research that has found the ABLA to have significantly higher predictive

validity than experienced staff (Thorsteinsson et al., 2007; Stubbings & Martin, 1998), the findings from the staff predictions in the current study should be interpreted with caution, and future research is warranted in this area.

For certain participants, slight modifications had to be made to the everyday four-choice tasks based on the participant's functioning. For example, due to the gross-motor skills restrictions of Participant 7, the procedure often had to be modified slightly (depending on the size of the stimuli) so that the correct response was pointing at the correct stimulus for 5 seconds, rather than placing a manipulandum with the correct stimulus. Participant 6 had similar restrictions, which is why Task 3 (above her level) was terminated and excluded from the analyses. That is, the participant appeared to be reinforced by the feeling of soft clothing, and since the task involved clothing, holding onto a manipulandum was presumably more reinforcing than placing it with the corresponding stimulus (i.e., releasing it). Future research should take into account the variability of participant functioning when selecting everyday tasks to assess.

Future replications of this study might use a research design in which each task is assessed on a best-of-three basis (i.e., two passes and one fail means a pass, and vice versa). This type of assessment might yield a more reliable measure of the participants' ability to perform certain tasks at various levels. This was particularly exemplified by Participant 1, who failed an analogue task at her level, but passed it above her level. It is conceivable that Participant 1 can pass that analogue task at her level under most circumstances, but on the particular day when her performance was recorded, she did not, perhaps for extraneous reasons such as fatigue, etc.

Overall, the present findings provide evidence that an individual's performance on ABLA Levels 2, 3, and 4 can predict four-choice discrimination learning for persons with developmental disabilities, which has practical implications for designing training programs in clinical settings.

Key Messages from This Article

People with disabilities: Application of the present findings can enhance the learning experiences of persons with severe and moderate intellectual disabilities.

Professionals: The present findings provide further evidence that teachers of persons with intellectual disabilities can use the Assessment of Basic Learning Abilities to match the difficulty of training tasks to the learning abilities of students.

Policy Makers: The present findings provide further evidence that instructors of persons with intellectual disabilities should be knowledgeable concerning the use of the Assessment of Basic Learning Abilities

References

Doan, L. A., Martin, T. L., Yu, C. T., & Martin, L. G. (2007). Do ABLA test results predict performance on three-choice discriminations for persons with developmental disabilities? *Journal on Developmental Disabilities, 13,* 1–11.

Kerr, N., Meyerson, L., & Flora, J. A. (1977). The measurement of motor, visual, and auditory discrimination skills. *Rehabilitation Psychology, 24,* 95–112.

Martin, G., & Pear, J. (2011). *Behavior modification: What it is and how to do it* (9th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Martin, G. L., Thorsteinsson, J. R., Yu, C. T., Martin, T. L., & Vause, T. (2008). The assessment of basic learning abilities test for predicting learning of persons with intellectual disabilities: A review. *Behavior Modification 32,* 228–247.

Martin, G. L., & Yu, C. T., (2000). Overview of research on the Assessment of Basic Learning Abilities test. *Journal of Developmental Disabilities, 7,* 10–36.

Stubbings, V., & Martin, G. L. (1998). Matching training tasks to abilities of people with mental retardation: A learning test versus experienced staff. *American Journal on Mental Retardation, 102,* 473–484.

Thorsteinsson, J. R., Martin, G. L., Spevack, S., Martin, T. L., & Lee, M. S. (2007). Predicting learning ability of people with intellectual disabilities: Assessment of Basic Learning Abilities test versus caregivers' predictions. *American Journal of Mental Retardation, 112,* 130–139.