Do ABLA Test Results Predict Performance on Three-Choice Discriminations for Persons With Developmental Disabilities?

Lori A. Doan, Toby L. Martin, C.T. Yu, and Garry L. Martin

Abstract

The Assessment of Basic Learning Abilities (ABLA) test assesses the ease or difficulty with which persons with developmental disabilities are able to learn 2-choice visual and auditory discriminations. We examined the ABLA's ability to predict 3-choice discrimination performance with 12 adults with developmental disabilities. Participants performed significantly better on 3-choice tasks that require discriminations that they passed on the 2-choice tasks than on 3-choice tasks that require discriminations that they failed on the 2-choice tasks, strengthening the ABLA's predictive validity. Theoretical and applied implications of these results are discussed.

The Assessment of Basic Learning Abilities (ABLA) test, developed by Kerr, Meyerson, and Flora (1977), evaluates the ability of individuals with moderate to profound developmental disabilities to learn a simple imitation and five two-choice discriminations that are required in many educational and vocational tasks. The six levels of the ABLA are hierarchically ordered in difficulty, with Level 1 being the easiest task to learn and Level 6 being the most difficult task to learn. It has high test-retest, intertester, and interobserver reliability, and accurately predicts the ability of individuals with developmental disabilities to learn two-choice visual and auditory discriminations (Martin, Yu, & Vause, 2004). Teachers that know a student's ABLA level can therefore choose academic, work, and leisure activities that match his or her discrimination skills, increasing the likelihood of successful performance and often reducing problem behaviours (Martin et al., 2004).

Other things being equal, discrimination assessments with three-choice options on each trial may have greater validity than assessments that include only two-choice options. On a two-choice task a testee can earn reinforcement by choosing the correct stimulus (selection control), or by rejecting the incorrect stimulus (rejection control; Boelens, 2002; Carrigan & Sidman, 1992). On a three-choice task, however, learning by selection

is more efficient because it requires attending to only one stimulus, whereas learning by rejection requires attending to two stimuli. Reducing learning by rejection should reduce experimenter error in data analysis and allow experimenters to make more accurate predictions from assessment results (Boelens, 2002; Carrigan & Sidman, 1992). The selection-rejection distinction may also have practical consequences, as seen in the example of a person who is taught to put garbage in a blue bin, and not the juxtaposed red bin (which is for dirty laundry). The person could respond correctly by attending to the stimulus characteristics of the red bin, which comprise an S-delta. If the red bin were replaced, however, the person's frequency of correct responses (to the blue bin) might drop sharply, leading to inappropriate garbage disposal.

Does the ABLA's two-choice format limit the test's validity? Many discriminations required in everyday life involve more than two choices (Carrigan & Sidman, 1992). Including only two-choice tasks may therefore reduce the test's predictive validity in that performance on the test's two-choice discrimination tasks may not correlate well with performance on three-choice tasks. The purpose of this study was to evaluate this question experimentally. Given the ABLA's high predictive validity for two-choice discriminations, we hypothesized that the predictive validity of the ABLA test would extend to three-choice discriminations. If this were true, it would broaden the applications of the ABLA for assessment and training that have been demonstrated in previous research (Martin et al., 2004).

Method

Participants and Setting

Twelve adults with moderate to profound developmental disabilities and ranging in age from 23 to 49 years (M=32 years) participated in this study. Participants were recruited from the St. Amant Centre, a residential and community resource facility for individuals with developmental disabilities in Winnipeg, Manitoba, Canada. Participants were selected based on ABLA test results, such that three participants passed Level 2 and failed all higher levels, three passed Level 3 and failed all higher levels, three passed Level 4 and failed all higher levels, and three passed all six ABLA levels (ABLA levels are described in more detail below). Testing took place in a quiet assessment room at the St. Amant Centre or in a quiet room at the participant's residence, and participants sat at a table across from the experimenter.

Procedure

Standard ABLA test (levels 2, 3, 4 and 6). Standard ABLA test materials consisted of two containers and three manipulanda. The containers included a red box measuring 14 cm x 14 cm x 10 cm, and a yellow can measuring 15 cm in diameter and 17 cm in height. The manipulanda included an irregularly shaped piece of white foam measuring approximately 4.5 cm x 4.5 cm x 4.5 cm, a yellow cylinder measuring 9 cm long and 4 cm in diameter, and a red cube measuring 5 cm x 5 cm x 5 cm (Martin & Yu, 2000).

Level 2 is a position discrimination. The two containers (yellow can and red box) are placed in a fixed position, and the client is required to place the foam into the container on the left. Level 3 is a visual discrimination. The left-right positions of the containers change randomly from trial to trial, and the person is required to put a piece of foam into one of the containers (i.e., the yellow can), regardless of its position. Level 4 is a visual match-to-sample discrimination. The two containers again randomly change positions across trials. The person is given a yellow cylinder or a red cube on each trial and the correct response is to place the yellow cylinder in the yellow can, and the red cube in the red box. Level 5 is an auditory discrimination. The positions of the two containers remain fixed across trials. The tester says "red box" or "yellow can", and the client is required to put a piece of foam into the appropriate container. Level 6 is an auditory-visual combined discrimination. The procedure is identical to Level 5 except that the containers' positions are altered randomly across trials.

During testing, the tester sat across the table from the individual being tested, and introduced all levels with a demonstration trial, a guided trial, and a practice trial. Testing and response recording began following the client's first independent correct response on the practice trial. An edible and verbal praise (e.g., "good work") were presented after each correct response. Errors led to another set of demonstration, guided, and independent trials. Each task was presented until the person met the pass criterion (eight consecutive correct responses) or fail criterion (eight cumulative errors). Correct responses made during the error correction procedure were not counted toward the pass criterion. The probability of reaching this pass criterion in a 2-choice task by chance, with independent responses across trials (similar to flipping a coin), is 0.03. The test took approximately 30 minutes to conduct. Over the course of six studies, 96% of 197 participants who passed Level 5 also passed Level 6 (Martin & Yu, 2000). As a result, DeWiele and Martin (1998) recommended omitting Level 5 from ABLA testing, and participants in this study were not tested on Level 5.

Analogue three-choice tasks. Participants whose highest passed level on the standard 2-choice ABLA was 2, 3, 4, or 6 proceeded to three-choice task assessments. Participants at levels 2, 3, and 4 received six 3-choice task assessments: three tasks that required the highest passed ABLA discrimination (referred to as tasks at the participants' ABLA levels), and three tasks that required the first failed ABLA discrimination (referred to as tasks above the participants' ABLA levels). For example, Level 2 participants received three tasks at Level 2, and three tasks at Level 3. Level 6 participants received three 3-choice discrimination task assessments at their ABLA Levels and no tasks above their levels because Level 6 is the highest level in the ABLA test. Table 1 describes all three-choice task materials.

Table 1. Three-choice task materials and stimuli

ABLA Level	*Task*	*Materials*
2	1	Round yellow can, square red box, triangular prism-shaped blue container, and white foam.
	2	Square black box, round gold can, rectangular green box, and red poker chip.
	3	Square black box, round gold can, rectangular green box, and clothespin.
3	1	Round yellow can, square red box, triangular prism-shaped blue container, and white foam.
	2	Neon pink index card (largest size), white index card with picture of a dog (medium size), neon green index card with text (small size), and heart-shaped block.
	3	Square black box, round gold can, rectangular green box, and a bolt.
4	1	Round yellow can, square red box, triangular prism-shaped blue container, and smaller matching manipulanda for each container.
	2	Large white envelope, medium brown envelope, small blue envelope, and smaller matching pieces of paper.
	3	Large blue box with a blue spoon taped to the back, medium gold box with a gold knife taped to the back, small silver box with a silver fork taped to the back, and matching cutlery.
6	1	Round yellow can, square red box, triangular prism-shaped blue container, and white foam. Auditory stimuli: "yellow can" (L), "blue triangle" (N), and "red box" (H).

continued

Table 1. (cont'd.)

ABLA Level	*Task*	*Materials*
6	2	A one-dollar coin "loonie" (H), "quarter" (N), and "penny" (L), each glued to separate index cards.
	3	Three index cards with the letter "A" (H), "B" (L), and "C" (N) written in different colours.

Note. ABLA = Assessment of Basic Learning Abilities. (L) indicates words spoken at low pitch and speed. (N) indicates words spoken at normal pitch and speed. (H) indicates words spoken at high pitch and speed.

Testing procedures for the three-choice discrimination tasks were similar to those for the standard ABLA test. Task presentation varied across participants, with at and above tasks presented in an alternating order. For the three-choice Level 2 tasks, the containers remained in the same position and the participant was required to put the manipulandum in the container to the right when the tester asked "where does it go?"

The position of the containers or stimuli changed from trial to trial for tasks at levels 3, 4, and 6. These position changes were randomly selected before testing commenced, with two restrictions: a minimum of one position change was required for each trial, and individual containers or stimuli could not remain in the same position for more than two successive trials. For the three-choice Level 3 tasks, when the tester asked "Where does it go?" the participant was required to put the manipulandum in the same container or on the same stimulus for each trial, but the positions of the containers or stimuli changed from trial to trial.

For the Level 4 match-to-sample task, on each trial, the experimenter gave one of the three manipulanda to the participant, who was then required to place that object on top of the matching stimulus or in the matching container when the tester asked "Where does it go?" Unlike the standard ABLA test, in which the manipulanda are smaller versions of the larger containers, the manipulanda presented for Level 4 analogue testing were objects that matched, but were not necessarily identical to, the test objects or containers (e.g., blue spoon in blue container). There were two reasons for this. First, it was difficult to find smaller versions of some of the objects. Second, the purpose of the analogue test was to assess whether or not the two-choice discrimination was predictive of three-choice educational and prevocational tasks, some of which involved matching on only one dimension (e.g., colour).

For the Level 6 three-choice tasks, the experimenter gave a verbal prompt to indicate the target container or object for each trial, and the participant was required to place a piece of foam in the correct container, or to point to the correct object. The verbal prompts consisted of naming the target stimulus, varying the pitch and speed of each verbal prompt, similar to the standard ABLA 2-choice discrimination. For example, "red box" was said quickly and at a high pitch, "yellow can" was said slowly and at a low pitch, and "blue triangle" was said at a normal pitch and speed.

The pass criterion for the three-choice tasks was five consecutive correct responses; the fail criterion was five cumulative incorrect responses. These criteria were selected such that the probability of passing the 3-choice task by chance (0.02), with independent responses across trials, was as close as possible to that of the 2-choice task (0.03).

Reliability assessments. Interobserver reliability checks were conducted on 64% of all test sessions across participants. During interobserver reliability checks, the examiner and an observer independently recorded the participant's responses for each trial. An agreement was scored if the examiner and observer recorded the same response for a trial, and a disagreement was scored if the examiner and observer recorded different responses for a trial. Calculation of percent agreement consisted of dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100% (Martin & Pear, 2003). Percent agreements had a mean of 98%, with a range of 75 to 100%.

Procedural integrity checks were conducted on 62% of all test sessions across participants. During procedural integrity checks, an observer used a checklist to record on each trial whether the containers/stimuli were in the correct position, correct instructions were given, correct demonstration and guided trials were provided following an error, and reinforcement was provided following for correct responses. A trial was considered correct if the examiner made no errors. The percentage of correctly administered trials per session averaged 99% across sessions, with a range of 77 to 100%.

Results

The nine participants at ABLA levels 2, 3, and 4 passed a mean of 1.78 (SD=1.2) 3-choice tasks at their ABLA level, and a mean of 0.11 (SD=0.33) 3-choice tasks above their ABLA level. Only one participant passed a task above his ABLA level. Data for each participant on 3-choice tasks at and above his/her ALBA level are presented in Table 2. A paired t-test showed

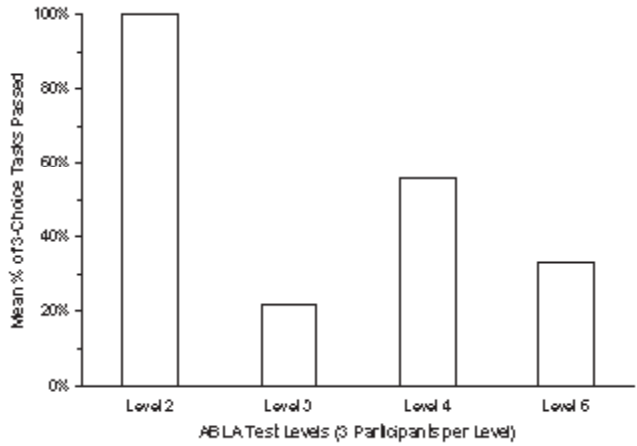
that the difference in number of tasks passed at versus above the participants' ABLA levels was statistically significant ($t(8) = 4.47$, $p=.001$, one-tailed). The mean percent of 3-choice tasks passed at the participants' ABLA Levels are shown in Figure 1 for each ABLA level.

Table 2. Three-choice tasks passed and failed by participants at each ABLA level

ABLA Level	*Participant*		*At*			*Above*	
2	1	1	2	3	1	2	3
		P	P	P	P	F	F
	2	P	P	P	F	F	F
	3	P	P	P	F	F	F
3	4	P	F	F	F	F	F
	5	F	F	F	F	F	F
	6	P	F	F	F	F	F
4	7	F	F	P	F	F	F
	8	P	F	F	F	F	F
	9	P	P	P	F	F	F
6	10	P	F	F	N/A	N/A	N/A
	11	P	F	F	N/A	N/A	N/A
	12	P	F	F	N/A	N/A	N/A

Note. ABLA = Assessment of Basic Learning Abilities; At = Three-choice tasks based on the highest ABLA level the participant can pass; Above = Three-choice tasks based on the lowest ABLA level the participant fails. The three-choice tasks are described in Table 1.

Figure 1. Mean percent of three-choice tasks passed by participants at their ABLA test levels.



Each of the three Level 6 participants passed one of the three 3-choice tasks at his/her ABLA level. For all three Level 6 participants, the task passed was the ABLA extension task, which consisted of the original ABLA test materials (i.e., red box and yellow can) and an additional blue triangular prism-shaped container (Task 1 at Level 6 in Table 1).

Discussion

The standard two-choice ABLA test successfully predicted performance on three-choice tasks involving similar discriminations. Specifically, participants performed significantly better on three-choice discriminations at the highest ABLA level they passed than on three-choice discriminations at the first ABLA level that they failed. At least two features of this result merit further investigation.

First, the performance difference between three-choice tasks at and three-choice tasks above a participant's ABLA level depended partly on the nature of the tasks. The three-choice discrimination assessment included an extension task that consisted of the existing ABLA materials (i.e., red box and yellow can) and an additional blue triangular prism-shaped container, as well as two analogue tasks that consisted of everyday items. Participants performed better at their ABLA level on the extension task (83% pass rate) than on the analogue tasks (37.5% pass rate). All participants had been previously tested using the two-choice ABLA, so it is possible that the increased performance on the extension task was due to familiarity with the task materials. For the extension task it was necessary for only one new container (i.e., the blue triangular prism-shaped container) to acquire appropriate stimulus control, whereas for the analogue tasks two or three new containers or stimuli needed to control responding appropriately. Although our task materials were chosen from previous predictive validity studies involving two-choice prevocational tasks (Stubbings & Martin, 1995, 1998), a staff questionnaire to ascertain each participant's familiarity with various analogue task materials may aid in choosing analogue task items in future studies.

Second, participants at Level 2 passed all three-choice tasks at Level 2, and their mean performance on three-choice predictive tasks at their ABLA level was higher than the mean performance of participants at levels 3, 4, and 6 on predictive tasks at their respective ABLA levels. Significance testing of the differences among participant levels was not appropriate due to the small sample sizes (three participants per level), but if this finding is robust, it may suggest factors that influence performance at each ABLA level.

For example, it may be that at Level 2, the intended S-delta stimuli (i.e., stimuli or containers that are the incorrect choice for the required response) have less control over the discriminations than do the S-delta stimuli at other levels. Accordingly, the number of S-delta stimuli would make little difference for Level 2 participants, but may make a larger difference for Level 3 participants, underscoring the impact of selection versus rejection on the predictions made with a two-choice assessment (Boelens, 2002; Carrigan & Sidman, 1992).

The use of different pass criteria for the two-choice and three-choice discrimination tasks warrants discussion. The standard ABLA test has a pass criterion of eight consecutive correct responses and a fail criterion of eight cumulative errors. This translates to a probability of 0.03 that an individual will pass the test by chance, with independent responses. This probability is obtained by likening random responses on each trial to tossing either heads (H) or tails (T) with a fair coin. The test (or game) consists of generating from 1 to 8 sequences of up to eight tosses. Any T result terminates that sequence, and the next sequence (if any) begins. The test is passed only by obtaining an HHHHHHHH sequence. It is failed by obtaining eight sequences that end in a T. For example, if the student "tosses" HHHT, the sequence terminates, the student accrues one "strike" against her, and a new sequence begins. The probability of any given sequence yielding a winning outcome is $(1/2)^8$, or 1/256. The probability of a non-winning sequence is therefore 255/256. The probability of losing the game is the probability of tossing 8 consecutive non-winning sequences, or $(255/256)^8$. The probability of winning is therefore $1 - (255/256)^8$, or about 0.03. For a three-choice test, a pass criterion of five consecutive correct responses and a fail criterion of five cumulative errors produces a probability of passing that is very close to that of the standard ABLA test, translating to a probability of 0.02 that an individual will pass the test by chance, with independent responses. This makes the pass/fail criterion of the three-choice predictive validity tasks slightly more stringent than the two-choice ABLA discrimination.

The results of the present study are consistent with the "Christmas Tree" model of discrimination ability proposed by Kerr (1977). In this model, the 6 discrimination levels of the ABLA test are the branches of the tree, and the tasks at each level are arranged along these branches, with easier tasks being closer to the trunk and more difficult tasks further out on the branches. Tasks that are at the same level of discrimination (e.g., Level 4 tasks), but that differ in complexity (horizontal ordering of tasks) are easier to learn than tasks at higher levels (vertical ordering of tasks). For example, a Level 4 task that requires matching along two dimensions (e.g., colour and shape)

may be easier than a task with only one dimension. Similarly, three-choice tasks might be more difficult than two-choice tasks, but a three-choice discrimination at the individual's current ABLA level may be easier to learn than a task that requires a two-choice discrimination above that person's level. Both of these propositions received support from this study's results. The potential value of the model is that learning tasks that are "closer to the trunk" at a particular level may facilitate rapid learning of other tasks on that branch (Kerr, 1977).

It is interesting to speculate about why some of the three-choice tasks were failed and some were passed within the same ABLA level. We have already discussed the possibility that some three-choice tasks may have been more familiar than others, and that some three-choice tasks (e.g., Level 4 tasks) may require matching on fewer dimensions. Because of the small sample size of the current study, an examination of the performance of individual participants on the tasks did not provide insight into these issues (see Table 2).

The experimental design for this type of predictive validity study required that participants initially be tested on the two-choice ABLA test. The two-choice and three-choice tasks at and above a participant's ABLA level were then presented in random order. Moreover, past studies have shown that failed ABLA levels are very difficult to teach (Martin & Yu, 2000). Thus it is unlikely that practice effects were a factor in participants' task performance on the three-choice tasks.

Many choices in everyday life involve more than two options. Although the ABLA does not include three-choice tasks, our results show that it predicts three-choice task performance, and that this may be particularly true for tasks at ABLA Level 2. Future studies involving three-choice discriminations in daily living activities (e.g., eating, dressing, grooming) and other domains would broaden the applicability of the ABLA. This study contributes to the literature by strengthening the ABLA's predictive validity, and by providing a new direction for future research involving multiple choices and everyday tasks.

Acknowlegements

We thank the participants for their cooperation throughout the study and Aynsley Verbeke, Kerri Walters, Deborah Baschuk, Kendra Thomson, Kailyn Stobbe, Sara Spevack, and Jennifer Thorsteinsson for their assistance on reliability assessments. This research was supported by grant MT6353 from the Canadian Institutes of Health Research

References

Boelens, H. (2002). Studying stimulus equivalence: Defense of the two-choice procedure. *The Psychological Record, 52*, 305-314.

Carrigan, P. F., & Sidman, M. (1992). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 58*, 183-204.

DeWiele, L. A., & Martin, G. L. (1998). *The Kerr-Meyerson Assessment of Basic Learning Abilities: A self-instructional manual.* (Available from Dr. G. L. Martin, Psychology Department, University of Manitoba, Winnipeg, MB, Canada, R3T 2M6)

Kerr, N. (1977). Implication of results for theory and practice: The "Christmas tree" model. *Rehabilitation Psychology, 24*(3), 133-136.

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

Martin, G. L., & Pear, J. J. (2003). *Behavior modification: What it is and how to do it.* Upper Saddle River, NJ: Prentice Hall.

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

Martin, G. L., Yu, C. T., & Vause, T. (2004). Assessment of Basic Learning Abilities test: Recent research and future directions. In W. L. Williams (Ed.), *Advances in developmental disabilities: Etiology, assessment and intervention.* Reno: Context Press.

Stubbings, V., & Martin, G. L. (1995). The ABLA test for predicting performance of developmentally disabled persons on prevocational training tasks. *International Journal of Practical Approaches to Disability, 19*, 12-17.

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.

Correspondence

Garry L. Martin
University of Manitoba
129 St. Paul's College
Winnipeg, MB
R3T 2M6

gmartin@cc.umanitoba.ca