Journal on Developmental Disabilities
Le journal sur les handicaps du développement
JoDD
Volume 22, Number 2, 2016

Effects of a Self-Instructional Manual, Computer-Aided Personalized System of Instruction, and Demonstration Videos on Declarative and Procedural Knowledge Acquisition of the Assessment of Basic Learning Abilities

Authors

Lei Hu,
Joseph J. Pear

Department of Psychology,
University of Manitoba,
Winnipeg MB

Correspondence

umhul@myumanitoba.ca
Joseph.Pear@umanitoba.ca

Keywords

Assessment of Basic Learning Abilities (ABLA), self-instructional manual (SIM), computer-aided personalized system of instruction (CAPSI), demonstration videos, teaching university students

Abstract

We evaluated the effects of three training components to teaching 12 university students the Assessment of Basic Learning Abilities *(ABLA; an instrument that measures an individual's ability to learn some basic behavioural functions). The three components involved the students (a) studying the ABLA using a self-instructional manual (SIM), (b) working on the manual combined with passing unit tests delivered via a computer-aided personalized system of instruction (CAPSI) program, and (c) watching demonstration videos. A multiple baseline design across the two training conditions was used to evaluate the effects of the components. Eleven students after receiving CAPSI training, as opposed to only three students after receiving SIM alone training, scored 85% accuracy or higher on declarative knowledge performance. Watching demonstration videos about the ABLA after the SIM and CAPSI training increased procedural knowledge performance for 10 students. The present study is one of the first to compare training effects of three components on teaching a behavioural assessment. The practical implications of training procedures were discussed.*

Practitioners working with individuals with autism or related disorders frequently encounter difficulty in determining what tasks should be taught to whom. To facilitate teaching, they should know the individuals' learning abilities and design training tasks accordingly (Martin & Yu, 2000). The Assessment of Basic Learning Abilities (ABLA) test – originally known as the Auditory Visual Combined Discrimination test (Kerr, Meyerson, & Flora, 1977) – was developed to address this difficulty. It measures how rapidly a testee can learn to perform some tasks, which are thought to reflect most of the activities found in daily life. During the administration of the ABLA test, a tester uses standard prompting and reinforcement procedures to teach one simple imitation and five two-choice visual and auditory discrimination tasks to the testee (DeWiele & Martin, 1998; DeWiele, Martin, Martin, Yu, & Thomson, 2010). The six separate tasks are hierarchically ordered in difficulty and therefore are referred to as measuring an individual's functioning at six ABLA levels. Research during the past two decades indicates that ABLA is a robust and reliable instrument to match the learning abilities of individuals with various levels of developmental disabilities (Yu, Martin, & Williams, 1989; Martin & Yu, 2000; Martin, Thorsteinsson, Yu, Martin, & Vause, 2008; Viel et al., 2011).

Administering the ABLA test with high fidelity requires a high quality of training. Training practitioners the repertoire of concepts, principles, rules, and facts, and the repertoire of procedures of administering the test involves training in both declarative knowledge (knowing *that*) and establishment of proficiency in procedural knowledge (knowing *how*) of the ABLA (Sternberg, 1998). The most common training strategies for behavioural techniques, including the ABLA, consist of complex direct instruction, typically through trainers conveying knowledge, modeling, role-playing, and providing feedback (Roscoe & Fisher, 2008; Sarokoff & Sturmey, 2004). However, direct instruction is labour intensive and difficult to carry out within a stable teaching structure (e.g., a highly experienced trainer needs to reliably deliver the knowledge and procedures to different trainees). With an increasing trend in the prevalence of diagnosed autism (Matson & Kozlowski, 2011) and a high demand of qualified staff in this field (Test, Flowers, Hewitt, Solow, & Taylor, 2003), direct instruction is becoming more costly. An effective, alternative training approach is needed.

To fulfill the need, some researchers have been applying learning principles to develop self-instructional manuals (SIMs) of behavioural techniques to facilitate training declarative and procedural knowledge (e.g., discrete-trial teaching, Fazzio & Martin, 2011; ABLA, DeWiele & Martin, 1998; DeWiele et al., 2010). The salient features of the SIMs include the following: (a) selected study materials are presented in small portions; (b) each small portion is accompanied by study questions; and (c) SIM users are asked to proceed to each successive portion only after mastering the current one (by demonstrating 100% accuracy on answering the study questions of the current one). Recent studies regarding the evaluation of SIMs indicate that the manuals can be effective in promoting knowledge development and increasing implementation accuracy (Fazzio, Martin, Arnal, & Yu, 2009; Wightman, Boris, Thomson, Martin, Fazzio, & Yu, 2012).

DeWiele, Martin, and Garinger (2000) compared the effectiveness of a SIM for the ABLA test (DeWiele & Martin, 1998) with providing its original description (Kerr, Meyerson, & Flora, 1977), which was the best available information package at the time. In DeWiele et al.'s first experiment, 21 undergraduate psychology students were randomly assigned to learn either the SIM or the original description and then, prior to a formal administration test, provided with an opportunity to practice what they had learned with confederates playing the role of an individual with a developmental disability. The authors found that the participants who studied the SIM produced better performances in obtaining declarative knowledge than those who studied the original description, in terms of accurately completing a comprehension exam about the ABLA test, a speed exam to gauge speed and accuracy of responding to questions about conducting the test, and a classification exam to assess ability to clarify training tasks in accordance with hierarchical levels of the test. With respect to procedural knowledge acquisition, the participants who studied the SIM produced superior results over their counterparts in administrating the ABLA test with the confederates.

In a subsequent experiment, DeWiele et al. (2000) evaluated a revised SIM for the ABLA test in an environment in which the participants were direct-care service providers in a residential training facility and administered the test to assigned clients from the facility. The participants were asked to study the SIM and to attempt to achieve mastery (90% accuracy) on the comprehension, speed, and classification exams specified in the prior experiment. Failure to reach the criterion led to restudying the SIM and retaking the exams. The participants then were required to practice administering the ABLA test to each other, with one of them role-playing a client with a developmental disability. Results indicated that, compared to those who studied the original description in the preceding experiment, staff who were trained with the SIM achieved better results in administering the ABLA test to real clients in a shorter period of time. In addition, based on the judgments of experts (i.e., professionals in the field of developmental disabilities), important clinically significant differences favouring the use of the SIM on the length of study, practice time of participants, and results obtained on the exams were observed.

Since, along with studying the SIM and answering study questions, the participants were required to self-practice or to role-play with each other, the effectiveness of the manual would likely be diminished in an environment in which the participants' behaviours on self-evaluating of study questions and self-practicing are not monitored. In addition, the effectiveness of the SIM is the reliance on the assumption that the learners will adhere to the mastery-before-proceeding-to-the-next-unit contingency. These limitations may hamper dissemination of the SIM. Computer-aided personalized system of instruction (CAPSI) offers a practical solution.

CAPSI is conceptualized as a teaching-learning process that involves the use of Internet-connected computers in the mediation of student-instructor and student-student interactions and the evaluation of learning (Pear, Schnerch, Silva, Svenningsen, & Lambert, 2011). It has adopted some features of personalized system of instruction (PSI, developed by Keller, 1968: a mastery-based method of teaching based upon learning principles). Five defining features distinguish PSI from direct instruction or lecture-based training: (a) the instructional material is presented in a written form and is broken down into small units; (b) students must demonstrate mastery of a given unit by passing a test of that unit; (c) students learn through the units in sequence at their own pace; (d) proctors, either former students or current students who have passed early units, grade and provide feedback to students regarding unit test performance; and (e) instructors may use occasional lectures as the main means of motivating interest, rather than as the source of imparting knowledge.

Following the tenets of PSI, a CAPSI course focuses on small units of information (e.g., one section or chapter) on selected textual materials (e.g., textbooks or manuals). Students are instructed to download a unit of content with assigned study questions from CAPSI, study the unit thoroughly, and respond to a number of study questions (e.g., 10) randomly sampled by the system. Feedback on each question (i.e., each correct answer is followed by a praise statement, and each incorrect answer followed by a presentation of corrective information) is provided either by the system immediately or by a human being (i.e., an instructor, mentor, or peer-reviewer) within 24 hours in written form. The mastery criterion of each unit is pre-determined. Students who reach the criterion of a given unit are allowed to proceed to the next unit. By contrast, students who fail to meet the criterion are prohibited by the system from proceeding. They are required to restudy the unit and rewrite a new test for the unit until the mastery criterion is met. Previous studies have indicated that CAPSI is effective in teaching university courses (Springer & Pear, 2008), enhancing critical thinking (Svenningsen & Pear, 2011), and developing declarative or verbal and procedural knowledge about specific behavioural techniques (Zaragoza Scherman, 2015).

Since the content of the SIMs has been sequentially separated into small units or sections accompanied by study questions, it seems that the SIMs are suitable textual materials for training individuals using CAPSI. For instance, Hu, Pear, and Yu (2012) evaluated a training package to teach the ABLA to three university students. The package included a SIM for the ABLA (DeWiele & Martin, 1998), five mastery-based unit tests corresponding to five ABLA levels (the original level 5 was deleted from the 1998 manual because research has shown that it overlaps with level 6), and five demonstration videos (one video for each level, which was accessible after passing a test for that level). The training was delivered via CAPSI, in which, for each level, students had to sequentially read a unit from the SIM, write randomly sampled study questions on a mastery basis (i.e., answer at least nine out of 10 correctly), and watch a demonstration video after passing the test for that level. The study questions required short answers that were automatically marked by the system. Unsuccessful attempts led to restudying the unit and rewriting the test. The training procedure ended when the student passed the last unit test and watched the videos for the last ABLA level. Acquisition of declarative and procedural knowledge of the ABLA were evaluated in a multiple baseline design across students. Results showed that the training package, as a whole, consisting of the SIM combined with unit tests and videos delivered via CAPSI was effective in developing declarative knowledge and in teaching the students to conduct the ABLA to a simulated client role-playing an individual with autism.

In summary, previous training interventions using the ABLA SIM share similarities in that participants were asked to read the SIM and attempted to either achieve mastery on all study questions (DeWiele et al., 2000) or pass mastery-based unit tests with sampled questions delivered via CAPSI (Hu et al., 2012). In addition, supplementary training components, consisting of either practicing with someone role-playing a client or watching demonstration videos, were also included in the training. Thus, the intervention effects of using the SIM combined with only passing mastery-based unit tests delivered via CAPSI remains unclear. Moreover, the effect of studying the SIM, along with study questions, has never been compared with CAPSI training. Finally, the effect of watching videos as a supplementary component has not been studied.

The primary purpose of the present study was to compare the effects of a SIM and a SIM combined with passing mastery-based unit tests delivered via CAPSI on participants' acquisition of declarative and procedural knowledge. Another purpose of the study was to determine whether students would improve their procedural performance further by viewing demonstration videos (Catania, Almeida, Liu-Constant, & DiGennaro Reed, 2009). Considering that CAPSI is highly efficient with regard to the utilization of human resources and can be used to train SIM users online, it was expected that the results of this study would indicate that combining the SIM with CAPSI would be more effective than the SIM alone for teaching declarative and procedural knowledge about the ABLA. It was also expected that the use of videos would be highly effective in teaching procedural knowledge.

Method

Participants

Recruitment posters were posted on information bulletin boards in key buildings at our university. Twelve university students (six males and six females) participated in the study, which was approved by the Psychology/Sociology Research Ethics Board of our university. According to the participants' self-report, they had not previously read any content related to the ABLA, had not had experience working in any behavioural intervention program, had not previously used CAPSI, and had Internet access. They had very diverse academic backgrounds; half of them took psychology courses before the study; and two of them were graduate students enrolled in a Master's program. The demographic information is provided in Table 1. Each participant received a total of 65 Canadian dollars for participating.

Table 1. Participants' Demographic Information

ID	*Gender*	*Age Range*	*Highest Level of Education*	*University Majors*	*University Minors*
P1	Female	26–30	2nd year Master's	Family Social Sciences	N/A
P2	Male	16–20	3rd year undergraduate	Psychology	N/A
P3	Female	21–25	2nd year undergraduate	Psychology	N/A
P4	Male	16–20	3rd year undergraduate	Psychology	Spanish
P5	Male	26–30	2nd year Master's	City Planning	N/A
P6	Male	16–20	2nd year undergraduate	Microbiology	N/A
P7	Male	21–25	1st year undergraduate	Not decided yet	N/A
P8	Female	21–25	1st year undergraduate	Engineering	N/A
P9	Male	21–25	2nd year undergraduate	Engineering	N/A
P10	Female	36+	Postgraduate	Nursing	English/ Sociology
P11	Female	16–20	1st year undergraduate	Biochemistry	N/A
P12	Female	16–20	1st year undergraduate	Not decided yet	N/A

Note: N/A = Not Applicable.

Materials

The training materials consisted of the ABLA SIM (2nd edition, DeWiele et al., 2010), the web-based CAPSI program, and demonstration videos. The CAPSI program in this study was only used to present unit tests and automatically mark answers on the unit tests.

The six levels of the ABLA SIM were combined into three sets with approximately equal length. Each set of contents to be taught included an introduction, which described basic concepts and general guidelines for using the ABLA, and two levels of the ABLA that were 3 levels apart (although the ABLA levels were presented in order of difficulty in the SIM from levels 1 to 6, there is no research evidence indicating any difference in difficulty for testers to learn to administer). The systematic selection of the contents tended to average any difference in difficulty in administering the ABLA across the three sets. Set A consisted of an introduction, the simple imitation task (Level 1), and the visual identity match-to-sample discrimination (Level 4); Set B consisted of an introduction, the position discrimination (Level 2), and the visual non-identity match-to-sample discrimination (Level 5); Set C consisted of an introduction, the visual discrimination (Level 3), and the auditory-visual combined discrimination (Level 6). The introduction section, which was taken from the SIM, was similar to all three sets of contents to provide general information and guidance for the two levels that were learned.

Materials for administrating the ABLA included two containers (viz., a yellow can and a red box) and five manipulanda (viz., a piece of foam, a cube, a cylinder, a purple piece of wood with the word "Can" carved on it, and a silver piece of wood with the word "BOX" carved on it). A video camera and a tripod were used to record testing sessions for retrospective scoring.

For each phase, a written test with 10 fill-in-the-blank questions was used to measure declarative knowledge acquisition about the ABLA. An application test involving conducting 12 trials was used to measure procedural knowledge acquisition, i.e., how accurately participants were able to implement the ABLA test on two pre-determined levels (i.e., Levels 1 and 4 in Set A, Levels 2 and 5 in Set B, or Levels 3 and 6 in Set C). For each trial, a 20- to 33-component checklist (the number of components varied at levels being tested), called the ABLA tester evaluation form (Martin, Martin, Yu, Thomson, & DeWiele, 2011), was used to evaluate accuracy with which participants implemented the ABLA test. An anonymous training feedback and evaluation survey consisting of 11 items was given to participants to measure their evaluations of the training components.

Setting

The training setting of the first three phases could be anywhere the participants chose (e.g., home). Therefore, their learning behaviours were not supervised. However, when participants were asked to use CAPSI, they were required to have a computer or hand-held device (e.g., BlackBerry, iPhone, iPad) connecting to the Internet. The training of the last phase occurred in a research room equipped with an Internet enabled computer at our university so that the participants could access demonstration videos. The written tests and application tests of each phase were conducted in a testing room at the university.

Independent Variables

The training conditions to be compared, consisting of either the SIM alone or the SIM plus CAPSI, constituted the independent variable. Both conditions involved the participant reading designated sets of contents from the SIM and responding to study questions corresponding to three units (i.e., a unit for an introduction section and two units for detailed information covering the two levels of the ABLA) of each set. Moreover, in the last phase of the experiment, the participants viewed videos demonstrating correct procedures and common mistakes made in the administration of six ABLA levels.

In the SIM condition, participants were required to study a designated set of contents from the SIM, answer study questions corresponding to three units, and check their responses against answer keys included in the set of contents. In the CAPSI condition, participants were required to study a set of contents and study questions included in the set, with the omission of the answer keys, and write three mastery-based unit tests delivered via CAPSI. Each mastery-based test consisted of 10 study questions, with the answers being marked automatically by the CAPSI program. The study questions for each unit of each set of contents were identical for both conditions; however, participants in the CAPSI condition were required to demonstrate mastery of 10 questions randomly sampled from the pool of questions for each unit. There were 46, 25, 18, 20, 17, 18, and 16 study questions for the introduction unit and units of Levels 1 to 6, respectively.

The total number of questions (i.e., study questions plus questions reserved for written tests) across all units was 190. Most of the study questions required one-word (e.g., *Following each correct response, you should provide **praise** to the student*), multiple-word (e.g., *The containers involved in the testing of Level 6, Auditory-Visual, are the **BOX** and the **Can***), and short-phrase (e.g., *The verbal prompt for Level 1 is "**where does it go**?"*) answers. A minority of the questions were of the true-false (e.g., *An incorrect response for Level 1 is defined as placement of the object anywhere other than in the container. **True/False***. The statement is True) and two-choice (e.g., *The position of the containers during the testing of Level 2, Position Discrimination, are **stable/alternate** from one trial to the next.* The word "stable" is the correct choice) types.

Dependent Variables and Data Collection

The dependent variables were (a) declarative knowledge of the ABLA; (b) accuracy of administering the ABLA levels to the first author playing the role of a client with autism; and (c) the participants' subjective evaluations of the training conditions and components. The first two variables were measured in all phases. The third was measured only at the end of the last phase.

Declarative knowledge of the ABLA was assessed by written tests (described previously – Materials and Setting sections), which were marked by a research assistant using a standardized answer key. The questions in the written tests did not overlap with the study questions during training. Procedural knowledge – that is, accuracy of conducting the ABLA levels – was evaluated by application tests (described previously – Materials and Setting sections), which were rated by the first author and the research assistant using the 20- to 33-component behavioural checklist. The research assistant was blind to the experimental condition each participant was in.

Research Design

A multiple-baseline design across two training conditions was used. Each participant was exposed to all three sets of contents, one set for each phase with three phases in total. Different orders of the sets were used to counterbalance any order effect across the participants. Considering that participants were to learn the entire SIM by studying all three sets (A, B, and C), there were six possible orders in which they could experience these sets: ABC, ACB, BAC, BCA, CAB, and CBA. Twelve participants were randomly assigned to two groups, six participants in each group, and each order was randomly assigned to one of the participants in each group. The SIM condition was always introduced first and thus served as baseline. Participants in Group 1 received training under the SIM condition once followed by training under the CAPSI condition twice, and participants in Group 2 received training under the SIM condition twice followed by training under the CAPSI condition once. Finally, in the fourth phase, participants from both groups watched demonstration videos. The methodology is illustrated in Table 2.

Procedure

Phase I: Training under the SIM condition. Participants were asked to read the assigned set of contents (see Table 2 for the specific set of contents being assigned to each participant), answer study questions corresponding to the three units of the set, and check their responses against answer keys that were provided. When they felt ready to be tested, participants were instructed to make an appointment with the first author to take a written test and to administer the two ABLA levels they studied. Based on the participants' self-report, the training process took a mean of 2 hours (range: 1.75 to 3 hours).

Table 2. Research Design

ID	*Phase I*	*Phase II*	*Phase III*	*Phase IV*
P1	SIM (Set A)	CAPSI (Set B)	CAPSI (Set C)	Videos (All levels)
P2	SIM (Set A)	CAPSI (Set C)	CAPSI (Set B)	Videos (All levels)
P3	SIM (Set B)	CAPSI (Set A)	CAPSI (Set C)	Videos (All levels)
P4	SIM (Set B)	CAPSI (Set C)	CAPSI (Set A)	Videos (All levels)
P5	SIM (Set C)	CAPSI (Set A)	CAPSI (Set B)	Videos (All levels)
P6	SIM (Set C)	CAPSI (Set B)	CAPSI (Set A)	Videos (All levels)
P7	SIM (Set A)	SIM (Set B)	CAPSI (Set C)	Videos (All levels)
P8	SIM (Set A)	SIM (Set C)	CAPSI (Set B)	Videos (All levels)
P9	SIM (Set B)	SIM (Set A)	CAPSI (Set C)	Videos (All levels)
P10	SIM (Set B)	SIM (Set C)	CAPSI (Set A)	Videos (All levels)
P11	SIM (Set C)	SIM (Set A)	CAPSI (Set B)	Videos (All levels)
P12	SIM (Set C)	SIM (Set B)	CAPSI (Set A)	Videos (All levels)

Note: Multiple baseline design across training conditions in which two groups of six participants move across conditions in a semi-staggered manner, and three sets of contents (A, B, and C) from the SIM counterbalance any order effect across participants.

Phase I: Post-training measurements. The post-training measurements, which consisted of a declarative knowledge and a procedural knowledge test, were conducted in a mean of 6.6 days (range: 2 to 8 days) after the commencement of training in Phase I. The written (declarative knowledge) test, including 10 fill-in-the-blank questions, had a 10-minute limit. The questions in the test were novel to the participants, but closely followed the format and difficulty of the study questions from the set of contents. The written tests were only on material from the relevant section of the SIM, not on the introduction that was included in each set. During the application (procedural knowledge) test, each participant was asked to conduct 12 trials of the ABLA on the two studied levels (6 trials for each level), which were administered in a hierarchical order, with a brief break (30 seconds) between levels. For example, when Set A (including ABLA levels 1 and 4) was being administered, the participants who studied the set were always required to finish implementing 6 trials of Level 1 before starting Level 4. Note that the application test of each set of contents included two sessions, with each ABLA level presented as a session (the lower level was presented first). Each participant was asked to let the experimenter know when he or she finished a trial and was going to move on to the next trial. The application test was completed in a mean of 8.5 minutes (range: 5.5 to 21 minutes). All sessions were videotaped for retrospective scoring. Participants did not receive any further training or feedback on their performance during and after the written and application test.

Phase II: Training under either the SIM or the CAPSI condition. Six participants remained in the SIM condition while the others commenced training under the CAPSI condition (as shown in Table 2). All participants started out studying a different set of contents from the SIM. The participants in the SIM condition received the training process described previously. The participants in the CAPSI condition were asked to read the assigned set of contents sequentially on a unit-by-unit basis and to access a unit test. For each set, three tests corresponded to the three units: an introduction, an ABLA level from 1 to

3, and another ABLA level from 4 to 6. To pass a unit, the participant was required to correctly respond to at least nine out of 10 fill-in-the-blank and single-choice questions within 15 minutes.

The CAPSI program automatically marked the test and immediately provided feedback to all questions. Each correct answer was followed by a praise statement (e.g., “good work”) on the computer screen, and each incorrect answer was followed with a presentation of all acceptable answers. If a participant met the mastery criterion, he or she would be complimented (e.g., “congratulations”) with a result of a “pass” and could proceed to the next unit. If a participant scored below 90% correct, the program would notify the participant that he or she was required to restudy the material and write another test on the unit no sooner than 15 minutes after the failed attempt. The training under the CAPSI condition for phase II ended after the participant successfully passed all three tests in that phase. The training process with the involvement of CAPSI took a mean of 2.5 hours (range: 2 to 3.5 hours).

Phase II: Post-training measurements. The post-training measurements were similar to Phase I measurements and were conducted in a mean of 7.3 days (range: 3 to 12 days) after the commencement of training in Phase II. The participants were asked to complete a declarative knowledge test with a novel set of 10 questions and to administer a procedural knowledge test with 12 trials on the two ABLA levels that they just studied. The two tests were delivered in the same manner as described above.

Phase III: Training under the SIM plus CAPSI condition. All participants studied the remaining set of contents specific to them and accessed the CAPSI program for testing as described above for Phase II. The training process in this phase took a mean of 2 hours (range: 1.5 to 3.5 hours).

Phase III: Post-training measurements. The post-training measurements were conducted in a mean of 6.3 days (range: 2 to 9 days) after the commencement of training in Phase III. Similar to the measurements in the Phases I and II, the participants were, again, asked to finish a written test on a novel set of questions and to conduct 12 trials on the two ABLA levels they studied.

Phase IV: Watching demonstration videos. After participants completed the post-training measurements in the previous phase, they were asked to make an appointment to meet the first author to watch six demonstration videos. The videos were stored in a laptop and were played by media player software installed on the computer. Six videos showed actors (psychology graduate students in the field of applied behaviour analysis) demonstrating the correct procedures and common errors for assessing the administration of the ABLA levels, one video per level. All videos took approximately 40 minutes to watch and played only once.

Phase IV: Post-video measurements. The post-video measurements were conducted in a mean of 3.8 days (range: 1 to 9 days) after the completion of the measurement in Phase III. Like the measurements conducted in the previous phases, a declarative knowledge and a procedural knowledge test were given. However, these measurements were cumulative: i.e., a novel set of 10 questions was used to test declarative knowledge covered by the entire SIM, and 12 trials selected from all six levels of the ABLA test (2 trials per level), without replacement, were used to test procedural knowledge. Data presentation of the application test differed from previous phases in that two sessions in this phase presented participants’ performance on the ABLA levels being learned under the SIM and the CAPSI condition (the averaged data for the levels being studied under the SIM condition always presented first).

Interobserver Agreement and Procedural Integrity Checks

Interobserver agreement (IOA) on the videotaped sessions was assessed as follows: The research assistant randomly sampled and viewed a mean of 45% (range: 35% to 50%) of the sessions across all participants. The first author and research assistant independently recorded either the occurrence or nonoccurrence of the participant’s specific behaviours on each trial of the sampled sessions on a component-by-component basis using the checklist described above. An agreement was defined as both raters jointly scoring a component on the checklist as correct, incorrect, or not applicable for the level being tested. A disagreement was defined as a discrepancy between the two raters in scoring a component on the checklist. IOA per participant was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by100% (Martin & Pear, 2011, p. 266). Mean and range agreements across 12 participants is reported in Table 3.

Table 3. Interobserver Agreement (IOA) Across Participants

ID	*Mean*	*Range*
P1	94%	85%–99%
P2	94%	89%–99%
P3	92%	88%–95%
P4	100%	
P5	100%	
P6	96%	89%–98%
P7	98%	96%–99%
P8	96%	94%–98%
P9	95%	89%–100%
P10	97%	92%–100%
P11	89%	75%*–96%
P12	86%	80%–96%

* *The low IOA occurred in Phase I in which the participant did not explicitly inform the experimenter the completion of each trial.*

For procedural integrity, when scoring the sessions the research assistant also recorded whether or not the first author's behaviours were performed in accordance with a script on a trial-by-trial basis. The script described the first author's planned (i.e., correct or incorrect) response on each trial of the application measurements. Its purpose was to ensure that all of participants encountered the same response across trials. The mean procedural integrity score was 95% (range: 83% to 100%) based on 65% randomly sampled sessions across all participants.

Results

Table 4 shows participants' mean declarative and procedural knowledge performance among three sets of the ABLA for the first three training phases. The mean difference of declarative knowledge among the three sets was not significant, $F(2, 33) = 0.68$, $p > .05$. The mean difference of procedural knowledge among the three sets also was not significant, $F(2, 33) = 0.12$, $p > .05$. One-way ANOVA post hoc multiple comparisons indicated no significant differences between the three sets of the declarative and procedural knowledge tests. Table 5 presents participants' mean procedural knowledge performance including the data of the last phase on all six ABLA levels. An ANOVA test did not indicate any significant difference on implementing the levels, $F(5, 66) = 0.25$, $p > .05$. These data suggest that although the levels presented in order of difficulty for testees, the difference in difficulty may not be experienced by the testers by sets and levels.

Figure 1 (on page 74) shows performance (% correct) on declarative knowledge and procedural knowledge tests across training conditions for each participant. Arrows for each participant indicate the highest performance on the two different tests. The 85% accuracy criterion is used in the present study to evaluate whether a participant has reached mastery of the ABLA.

Table 4. Mean Declarative and Procedural Knowledge Performance Among Three Sets of the ABLA for the First Three Training Phases

	Declarative Knowledge[a]		*Procedural Knowledge*[b]	
Test Contents	*M (%)*	*SD*	*M (%)*	*SD*
Set A (Levels 1 & 4)	68.4	30.9	57.8	29.7
Set B (Levels 2 & 5)	78.3	10	60.1	22.1
Set C (Levels 3 & 6)	77.5	22.6	63.1	27.4

Note: [a] *Multiple comparisons using Tamhane's T2 test (equal variances not assumed);* [b] *multiple comparisons using Fisher's LSD test (equal variance assumed).*

Table 5. Participants' Mean Procedural Knowledge Performance on Six ABLA Level

	ABLA Levels					
Statistics	*Level 1*	*Level 2*	*Level 3*	*Level 4*	*Level 5*	*Level 6*
M (%)	52.8	60	64.5	61.6	60.3	62
SD	29.6	24.9	29	30.2	21.4	26.7

Table 6. Mean Declarative Knowledge Performance (% Correct Answers) for Two Groups of Participants Across Phases

Groups	*Phase I SIM*		*Phase II CAPSI*		*Phase III CAPSI*		*Phase IV Demo Videos*	
	M (%)	SD	M (%)	SD	M (%)	SD	M (%)	SD
Group 1	55.2	28.7	80.3	18.7	89.2	11.9	78.4	15.4
	Phase I SIM		*Phase II SIM*		*Phase III CAPSI*		*Phase IV Demo Videos*	
Group 2	54.3	23.4	76.7	8.0	92.7	5.9	77.7	11.0

This criterion is chosen because of its clinical significance in practical settings (e.g., DeWiele et al., 2000). For the declarative knowledge test, all participants reached their highest performance in the CAPSI condition. Three of six participants (participants 1 to 6 who received CAPSI training twice) in Group 1 reached their highest performance immediately after completing the first CASPI training in Phase II. All six participants (participants 7 to 12 who received CAPSI training only once) in Group 2 reached their highest performance immediately after the CAPSI training in Phase III. Moreover, only one participant (participant 5) in Phase I and two participants (participants 7 and 9) in Phase II scored greater than 85% accuracy after the SIM training. Finally, watching demonstration videos did not improve participants' declarative knowledge performance.

For the procedural knowledge test, participants' performance was averaged in each session (as shown in Figure 1). Ten of 12 participants reached their maximum application performance after watching the videos in Phase IV. The remaining two participants (participants 1 and 5) reached their best performance (99% accuracy) after the completion CAPSI training in Phase II or III, limiting the room for further improvement. Moreover, no participant scored greater than 85% accuracy after the SIM training in Phase I. Comparing performance in Phase II in which participants were trained in one of the two training conditions, three of six participants (participants 1, 4, and 5) in the CAPSI condition in Group 1 scored greater than 85% accuracy while only one of six participant (participant 10) in the SIM condition in Group 2 met this criterion. When all participants were trained in the CAPSI condition in Phase III, only one more participant (participant 9) than those in Phase II achieved the specified criterion. Interestingly, even though all 12 participants were asked to administer all levels of the ABLA in Phase IV, eight of them scored greater than 85% accuracy after watching the videos. This was twice the number of participants in Phase II that scored greater than 85% accuracy.

Figure 2 (on page 75) presents the mean percentage of correct performance on declarative knowledge (top panel) and procedural knowledge (bottom panel) tests across training conditions for the two groups of participants. Mean declarative knowledge increased across the first three phases regardless of training conditions. Table 6 shows that mean declarative performance for the two groups was approximately

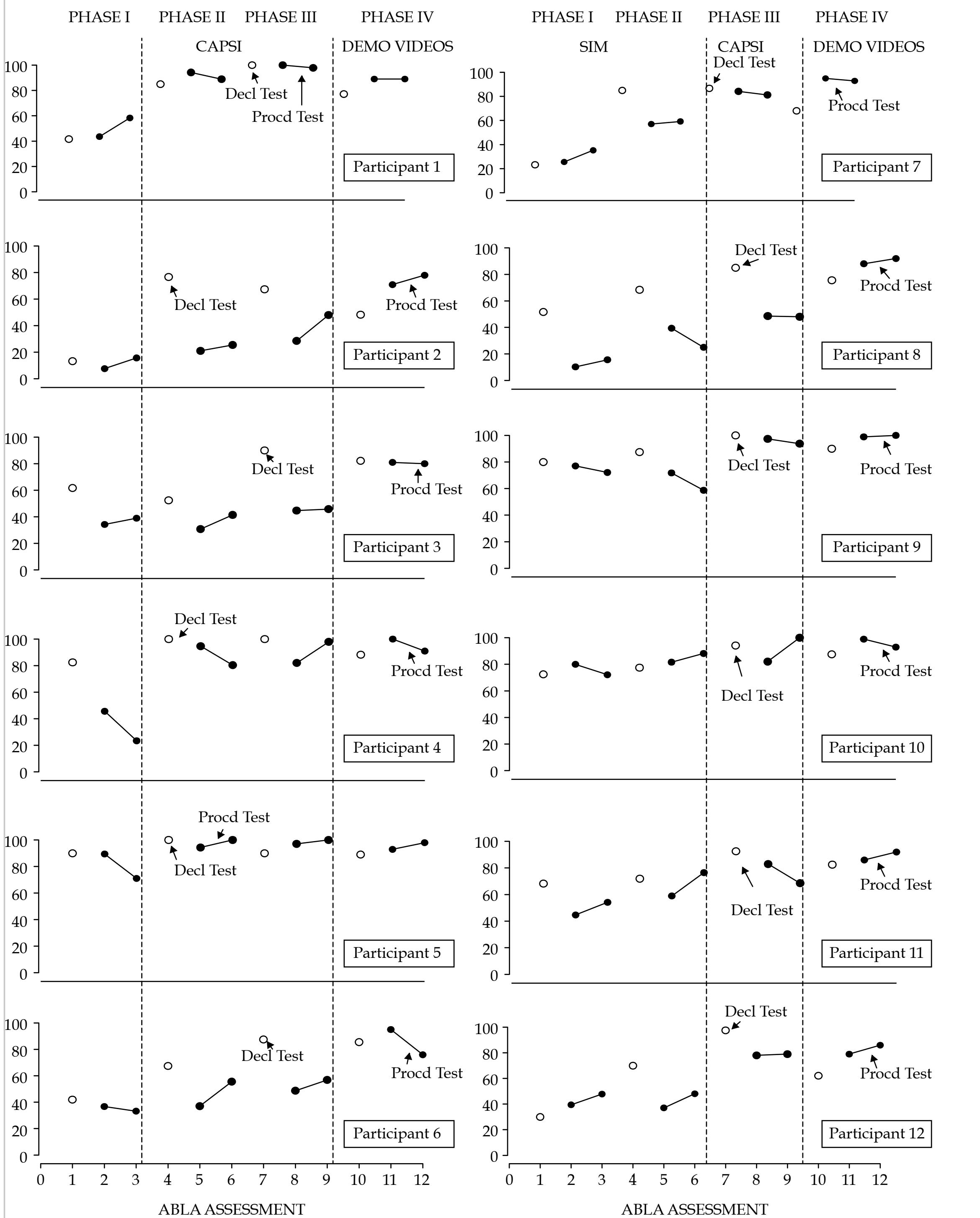
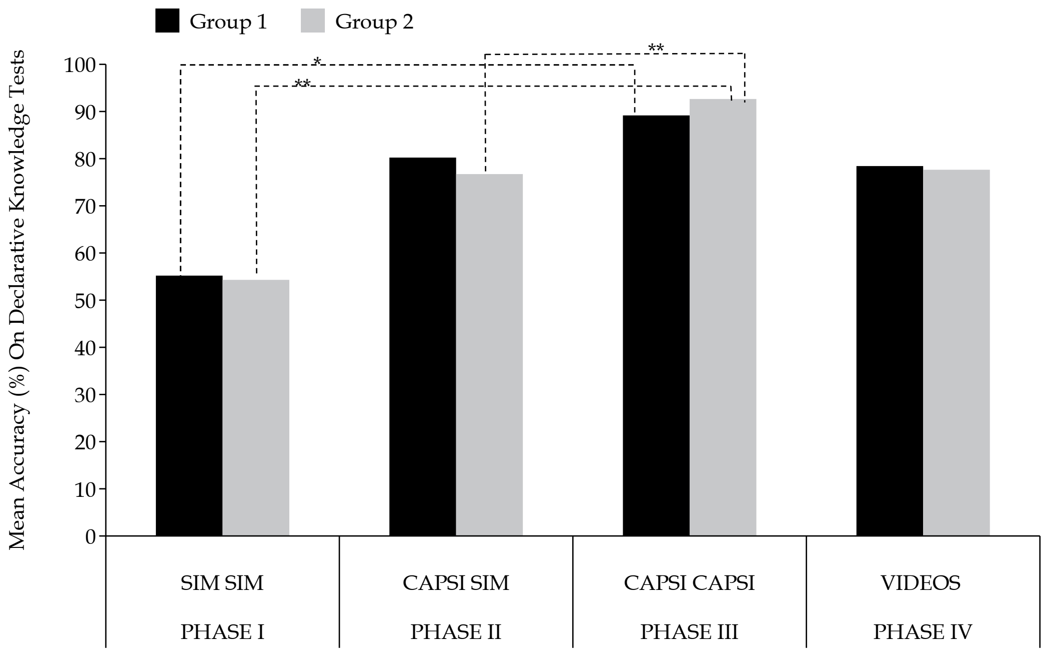


Figure 1. *Accuracy (% correct) on declarative knowledge and procedural knowledge tests across training conditions for each participant. Arrows for each participant indicate his or her highest performance on the two different tests.*



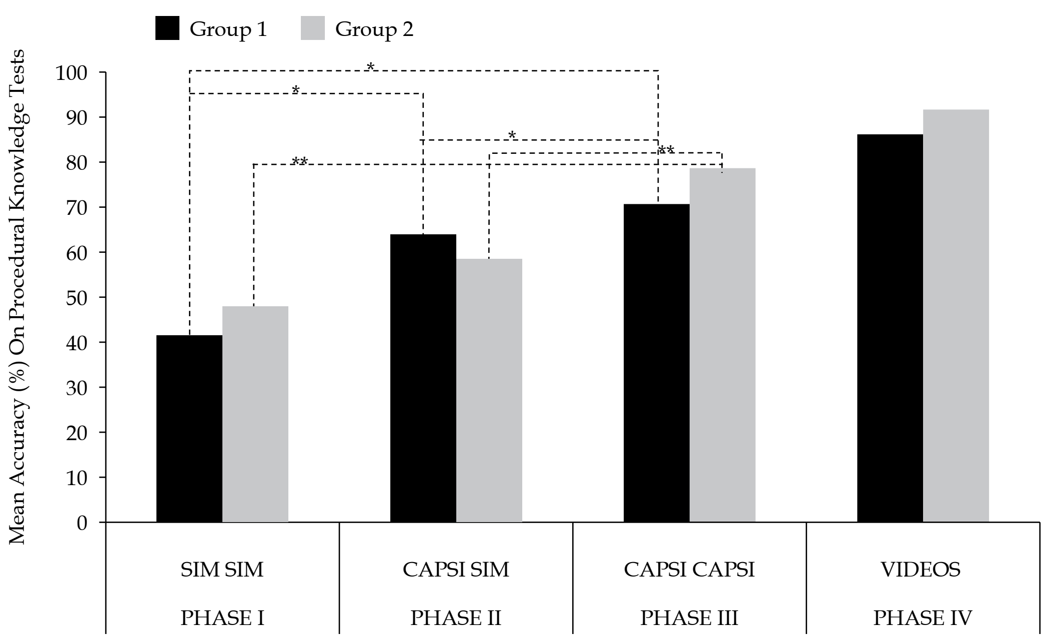


Figure 2. Mean accuracy (% correct) on declarative knowledge and procedural knowledge tests across training conditions for two groups of participants

*Note: * p < .05; ** p < .01*

equal after SIM training in Phase I, substantially increased in Phase II regardless of training conditions, then maximized after CAPSI training in Phase III, and slightly decreased after watching the videos in Phase IV. Although the between-group comparison in Phase II showed 3.6% difference in accuracy, the superiority over CAPSI training was not statistically significant: $t(10) = 0.43$, $p > .05$, two-tailed. However, within-group comparisons indicated that, compared to the 8.9% improvement ($t[5] = 1.12$, $p > .05$, two-tailed) for those in Group 1 who continuously received CAPSI training in Phase III, participants in Group 2 who just commenced the CAPSI training produced a significantly better outcome (16% improvement; $t[5] = 4.53$, $p < .01$, two-tailed).

Table 7 shows that mean procedural performance for the two groups also increased across phases regardless of training conditions. Group 2 started with higher performance than Group 1 after SIM training in Phase I. The two groups improved in procedural knowledge regardless of training conditions in Phase II. The performance continuously increased after CAPSI training in Phase III and reached its maximum after participants watched the videos in Phase IV. The between-group comparison in Phase II showed 5.5% difference in accuracy, but the advantage over CAPSI training was not statistically significant: $t(10) = 0.36$, $p > .05$, two-tailed. However, within-group comparisons indicated that, in contrast to the 10.6% improvement ($t[5] = 1.86$, $p > .05$, two-tailed) for those in Group 1 who remained in SIM training in Phase II, participants who switched to CAPSI training showed a significantly larger improvement (22.4%; $t[5] = 2.76$, $p < .05$, two-tailed). Similarly, participants in Group 2 improved significantly from Phases II to III when they finished CAPSI training (20.1% improvement; $t[5] = 4.09$, $p < .01$, two-tailed). Finally, both SIM and CAPSI training could not maximize procedural performance, which hence needs further training with the demonstration videos.

Pearson's correlation coefficient revealed that the two different tests were significantly and strongly correlated with each other, $r(48) = 0.66$, $p = .00$, two-tailed, suggesting that the participants, who had better performance on answering questions about the ABLA, were more likely to accurately administer the assessment, and vice versa.

Finally, on the training feedback and evaluation survey twice as many participants viewed the CAPSI training as being extremely helpful (six participants) as viewed the SIM training as being extremely helpful (three participants) with regard to the declarative knowledge tests. Three times as many participants perceived the videos as being extremely helpful (9 participants) as perceived either the SIM (three participants) or CAPSI training (three participants) as being extremely helpful with regard to the procedural knowledge test.

Discussion

Few studies have investigated the effectiveness of a training approach in teaching behavioural techniques to individuals under a completely unmonitored environment (Hu et al., 2012). To our knowledge, no studies have gone even further to compare the training effects of two self-instructional approaches on facilitating

Table 7. Mean Procedural Knowledge Performance (% Correct) for Two Groups of Participants Across Phases

Groups	*Phase I SIM*		*Phase II CAPSI*		*Phase III CAPSI*		*Phase IV Demo Videos*	
	M (%)	SD	M (%)	SD	M (%)	SD	M (%)	SD
Group 1	41.6	22.8	64	31.8	70.7	28.1	86.2	8.1
	Phase I SIM		*Phase II SIM*		*Phase III CAPSI*		*Phase IV Demo Videos*	
Group 2	47.9	24.7	58.5	18.8	78.7	16.6	91.7	6.1

declarative and procedural knowledge acquisition. Moreover, the effects of watching demonstration videos have received much recent attention and have been shown to be highly effective in teaching procedural knowledge (Catania et al., 2009). The present study compared two self-instructional approaches (i.e., SIM and CAPSI training conditions) on teaching the ABLA and also evaluated the effects of watching videos as supplemental training to the two approaches.

In this study, performance for each participant across phases on writing declarative knowledge tests suggests that: (a) studying the manual with all study questions can be effective, considering that three participants scored 85% accuracy or higher after SIM training; (b) CAPSI training (i.e., studying the manual and passing three mastery-unit tests) appears to be more beneficial as 11 participants scored 85% or higher after CAPSI training and as substantial improvements (15% or higher) occurred immediately after the first phase of CAPSI training for eight participants; (c) even though the helpfulness of the videos was reported, it seems that watching them provided a better model for procedural behaviour than for verbal behaviour; and (d) practice effects existed across the phases even though strong attempts were made to preclude them, indicating strong generalization across the teaching procedures at all ABLA levels. The between-group comparison in Phase II did not show significant difference favouring CAPSI training; however, the within-group comparison indicated that, compared to their performance after SIM training in Phase II, participants in Group 2 produced significant better outcome after CAPSI training. Taken altogether, the results may suggest the effectiveness of CAPSI on facilitating declarative knowledge acquisition.

Performance for each participant across phases on conducting procedural knowledge tests suggests that: (a) SIM training alone may produce suboptimal results as no participant achieved a high level of accuracy (85% or higher) in this condition; (b) although CAPSI training appeared to also show superiority on teaching administering the ABLA levels, participants' performance only reached a moderate level, suggesting the helpfulness of adding other training components (e.g., watching videos); (c) watching the demonstration videos further improved procedural performance to a high level, indicating the usefulness of adding the videos (thus extending the findings of Catania et al., 2009); and (d) practice effects appeared to have been present across the phases. Similar to participants' performance on the declarative knowledge tests, the between-group comparison did not differentiate procedural knowledge for the two groups of participants in Phase II; however, the within-group comparison suggests the superiority of CAPSI training. In addition, the results of the last phase indicate the usefulness of using the demonstration videos to teach administering the ABLA.

In summary, studying the SIM combined with passing unit tests delivered through CAPSI (CAPSI condition) appeared to be more effective than studying the SIM and study questions of the manual (SIM condition) in declarative knowledge acquisition. However, it appeared to be no more effective than the SIM condition in procedural knowledge acquisition, indicating the need of other training components. The addition of watching videos modeling correct procedures and presenting common errors in the administration of the ABLA did appear to be highly effective in teaching procedures for conducting the assessment. In addition, the declarative and procedural knowledge tests were complementary and tended to assess participants' abilities from different perspectives. However, keeping in mind that correlation does not necessarily imply causation, the strong positive correlation between the two tests suggests that, for professional development, the more effective a practitioner learns declarative knowledge about the ABLA, the more precise he or she would implement the assessment.

The results of this leave some questions to be answered in future studies. First, we do not know why some participants' performance continuously increased regardless of the training conditions. This may have been due to repeated exposure to similar material (e.g., introduction section from the SIM) across the first three phases for all participants. However, none of the tests contained materials from the introduction sections. Second, the between-group comparisons did not show significant differences. This may due to the limited sample size of each group. Third, the experiment did not include a generalization phase in which participants could apply learned skills to individuals whose basic learning abilities need to be assessed. Thus, we do not know whether the CAPSI training and demonstration videos would be effective for individuals working with real clients. Fourth, related to the first limitation, we do not know what the results would have been if the CAPSI condition rather than the SIM condition had served as the baseline. Finally, we do not know whether providing feedback after the tests would have improved accuracy on subsequent tests.

There are a number of possibilities for future research. First, in order to make a clear comparison, an independent groups design consisting of a SIM alone and a SIM plus CAPSI group might be needed. Or, as an alternative, a pre-test measure could be included in each phase to compare with the performance after training of that phase. Second, replications with other SIMs and with participants who might not be such efficient learners (as university students tend to be) would be beneficial. Third, research is needed to compare the effects of successfully passing three versus more mastery-based unit tests for each CAPSI training phase. Finally, participants in this study only viewed the videos once. Future research could evaluate the impact of the number of the videos viewed on the administration of the ABLA.

The findings of this study have implications for training individuals on behavioural techniques including the ABLA. SIMs combined with mastery-based unit tests and demonstration videos delivered via CAPSI could be an effective and low-cost training approach to teach both declarative knowledge and application of behavioural procedures. This comports with the findings of Hu et al. (2012) who used a training package involving CAPSI to facilitate teaching the ABLA. Because the participants both in this study and in the study by Hu et al. were trained in an unmonitored environment in which they studied at any location they chose and at their own pace and wrote unit tests online, a self-instructional approach involving the use of CAPSI would appear to be highly efficient. Because of the increasing demand for trained practitioners working with individuals with autism and related disorders, CAPSI with embedded demonstration videos holds promise to be an economical tool for conducting instruction and in maintaining a structure for staff training.

Key Messages From This Article

People with disabilities. You deserve to have your basic abilities accurately assessed by qualified assessors, using the assessment of basic learning abilities (ABLA). In addition, you deserve to have your abilities strengthened and extended with appropriate tasks based on the results of the assessment.

Professionals. Being able to provide the ABLA test with high fidelity to people with disabilities requires an effective and efficient training method.

Policymakers. The utilization of a computer-aided personalized system of instruction (CAPSI) program would be effective for training if a self-instructional manual and demonstration videos can be incorporated into the program.

References

Catania, C. N., Almeida, D., Liu-Constant, B., & DiGennaro Reed, F. D. (2009). Video modeling to train staff to implement discrete-trial instruction. *Journal of Applied Behavior Analysis*, *42*, 387–392.

DeWiele, L. A., & Martin, G. L. (1998). *The Kerr-Meyerson assessment of basic learning abilities: A self-instructional manual*. Unpublished Manuscript, Department of Psychology, University of Manitoba, Winnipeg, Canada.

DeWiele, L. A., Martin, G. L., & Garinger, J. (2000). Field testing a self-instructional manual for the ABLA test. *Journal on Developmental Disabilities, 7*(2), 93–108.

DeWiele, L., Martin, G., Martin, T., Yu, C.T., & Thomson, K. (2010). *The Kerr-Meyerson assessment of basic learning abilities revised: A self-instructional manual* (2nd ed.). St.Amant Research Center. Retrieved from http://stamantresearch.ca/abla/

Fazzio, D., & Martin, G. L. (2011). *Discrete-trials teaching with children with autism. A self-instructional manual*. Winnipeg, MB: Hugo Science Press.

Fazzio, D., & Martin, G. L., Arnal, L., & Yu, C.T. (2009). Instructing university students to conduct discrete-trials teaching with children with autism. *Research in Autism Spectrum Disorders*, *3*, 57–66.

Hu, L., Pear, J. J., & Yu, C.T. (2012). Teaching university students knowledge and implementation of the assessment of basic learning abilities. *Journal on Developmental Disabilities*, *18*(1), 12–19.

Keller, F. S. (1968). Good-bye teacher. *Journal of Applied Behavior Analysis*, *1*, 79–89.

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., Marin, T., Yu, D., Thomson, K., & Dewiele, L. (2011). *The ABLA-R Tester Evaluation Form*. Retrieved from: https://stamant.ca/research/abla/

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

Martin, G. L., Thorsteinsson, J. R., Yu, C.T., Martin, T., & 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 on research of the assessment of basic learning abilities test. *Journal on Developmental Disabilities*, *7*(2), 10–36.

Matson, J. L., & Kozlowski, A. M. (2011). The increasing prevalence of autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*, 418–425.

Pear, J. J., Schnerch, G. J., Silva, K. M., Svenningsen, L., & Lambert, J. (2011). Web-based computer-aided personalized system of instruction. In W. Buskist & J. E. Groccia (Eds.), *New directions for teaching and learning. Vol. 128: Evidence-based teaching* (pp. 85–94). San Francisco, CA: Jossey-Bass.

Roscoe, E. M., & Fisher, W. W. (2008). Evaluation of an efficient method for training staff to implement stimulus preference assessments. *Journal of Applied Behavior Analysis*, *41*, 249–254.

Sarokoff, R. A., & Sturmey, P. (2004). The effects of behavioral skills training on staff implementation of discrete-trial teaching. *Journal of Applied Behavior Analysis*, *37*, 535–538.

Springer, C. R., & Pear, J. J. (2008). Performance measures in courses using computer-aided personalized system of instruction. *Computers & Education*, *51*, 829–835.

Sternberg, R. J. (1998). Abilities are forms of developing expertise. *Educational Researcher*, *27*, 11–20.

Svenningsen, L., & Pear, J. J. (2011). Effects of computer-aided personalized system of instruction in developing knowledge and critical thinking in blended learning courses. *The Behavior Analyst Today, 12*, 33–39.

Test, D. W., Flowers, C., Hewitt, A., Solow, J., & Taylor, S. J. (2003). Statewide study of the direct support staff workforce. *Mental Retardation*, *41*, 276–285.

Viel, J., Wightman, J., Marion, C., Jeanson, B., Martin, G. L., Yu, D., & Verbeke, A. (2011). Does mastery of ABLA Level 6 make it easier for children with autism to learn to name objects? *Research in Autism Spectrum Disorders*, *5*, 1370–1377.

Wightman, J., Boris, A., Thomson, K., Martin, G., Fazzio, D., & Yu, C.T. (2012). Evaluation of a self-instruction package for teaching tutors to conduct discrete-trials teaching with children with autism. *Journal on Developmental Disabilities*, *18*(3), 33–43.

Yu, D., Martin, G., & Williams, W. L. (1989). Expanded assessment for discrimination learning with the mentally retarded persons: A practical strategy for research and training. *American Journal on Mental Retardation*, *94*, 161–169.

Zaragoza Scherman, A., Thomson, K., Boris, A., Dodson, L., Pear, J. J., & Martin, G. (2015). Online training of discrete-trials teaching for educating children with autism spectrum disorders: A preliminary study. *Journal on Developmental Disabilities*, *21*(1), 22–33.