# Does the Difficulty of the Training Tasks in an EIBI Program for Children With Autism Match the Learning Abilities of the Children?

by

Genevieve Roy-Wsiaki

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

University of Manitoba

Winnipeg

## Abstract

In many early intensive behavioural intervention (EIBI) programs, such as the St.Amant Autism Programs in Manitoba, the Assessment of Basic Language and Learning Skills-Revised (ABLLS-R) is used as an assessment guide, a curriculum guide, and a skills tracking system. The ABLLS-R lists a total of 544 potential training tasks. In the St.Amant Autism Programs, the Assessment of Basic Learning Abilities-Revised (ABLA-R) is also used to assess a child's ability to learn to perform 6 tasks, called levels, which are hierarchical in difficulty. Research has demonstrated that training tasks comparable to failed ABLA-R levels are very difficult to teach using standard prompting and reinforcement procedures, and training tasks that are mismatched to a client's highest-passed ABLA-R level result in more aberrant behaviours than matched tasks.

The purpose of this study was to evaluate whether the difficulty of the training tasks taught to children enrolled in the St.Amant Autism Early Learning Program matched the learning abilities of the children. In order to do so, five questions were examined based on the ABLLS-R, the ABLA-R, and archival data from the St.Amant Autism Programs. The results reveal several potentially important findings. First, observers who were knowledgeable about the ABLA-R reliably categorized 99 of the 544 ABLLS-R tasks into individual ABLA-R levels. Second, for a random sample of those 99 ABLLS-R tasks, Autism Consultants from the St.Amant Autism Programs averaged 90.5% agreement that those tasks were taught at the categorized ABLA-R levels. Third, across a sample of 14 children, 81% of their training tasks were mismatched to each child's highest-passed ABLA-R level. Fourth, for that sample of 14 children and across their 31 maladaptive behaviour assessments, 61% of the assessments had a score

representative of maladaptive behaviours. Fifth, for that same sample of children, their rates of acquisition of new training tasks were lower for mismatched tasks than for matched tasks.

These findings have important implications for potentially improving the services that children with ASD are receiving from the St.Amant Autism Programs and other EIBI programs.

Acknowledgements

First and foremost, I must express my sincere gratitude to the staff of the St.Amant Autism Programs and the St.Amant Research Centre, for their tremendous assistance and support. I would especially like to thank Drs. Kerri Walters, Carole Marion, and Toby Martin, as well as Mrs. Lindsay McCombe. I would also like to express my utmost appreciation for the following individuals who offered their time and efforts throughout this research: Lauren Kaminski, Morena Miljkovic, Sarah Rigby, Kari Webster, Maria Pongoski, and May Lee. I would like to acknowledge and thank my dissertation committee members, Drs. C. T. Yu, Joseph Pear, and Dennis Hrycaiko, for their invaluable guidance and feedback. Likewise, I would like to thank my advisor, Dr. Garry Martin, who ignited my passion for the science of behaviour analysis, who has shaped and positively reinforced my behaviour as a student and professional, and who has inspired me to do the same with my own students. Finally, a special thank you to my family, friends, and colleagues for their unwavering support and encouragements over the years. This research was supported by the University of Manitoba Graduate Fellowship and the Manitoba Graduate Scholarship.

Table of Contents

List of Tables

List of Figures

# Does the Difficulty of the Training Tasks in an EIBI Program for Children With Autism Match the Learning Abilities of the Children?

## Introduction

Children diagnosed with autism spectrum disorder (ASD) may demonstrate several symptoms, such as persistent deficits in social communication and social interaction, and restrictive and repetitive patterns of behaviour, interests, or activities (American Psychiatric Association, 2013). The symptoms may also cause clinically significant impairment in several areas of functioning (e.g., social or occupational). These symptoms must appear in the early developmental period, although the symptoms may not “fully manifest” until the children are older. The current description of ASD includes disorders that were previously diagnosed separately: autistic disorder, pervasive developmental disorder not otherwise specified (PDD-NOS), and Asperger syndrome (American Psychiatric Association, 2013).

Research has demonstrated that early intensive behavioural intervention (EIBI) is the most effective treatment for young children with ASD (Matson & Konst, 2013; Matson & Smith, 2008; Matson & Sturmey, 2011). EIBI is based on applied behaviour analysis (ABA), and includes the systematic application of learning principles and techniques to teach a wide variety of skills. This intensive intervention involves 30 to 40 hours per week of one-on-one instruction for a minimum of 2 years. Preschool children with ASD in EIBI programs have demonstrated significant gains, and many of them are indistinguishable from their peers by the age of 6. This is in part due to the use of curriculum guides in EIBI programs that are designed to teach skills in a hierarchical manner (least-to-most difficult), to potentially achieve the skill level of typically

developing children of their age (e.g., Matson & Konst). One such curriculum is the Assessment of Basic Language and Learning Skills-Revised (ABLLS-R; Partington, 2006). The ABLLS-R is used as an assessment, a curriculum guide, and a skills tracking system for children who present with delays in their acquisition of language skills.

Another assessment that might facilitate the selection of ABLLS-R skills to target for each individual child is the Assessment of Basic Learning Abilities (ABLA; Kerr, Meyerson, & Flora, 1977). The ABLA assesses a testee's ability to learn six discriminations, referred to as levels. Research has demonstrated that: (a) training tasks can be analyzed according to the highest ABLA level needed to perform them (DeWiele & Martin, 1996); (b) the ABLA performance of persons with intellectual disabilities (ID) can be used to match the learning ability of participants with the ABLA difficulty of training tasks (Martin, Thorsteinsson, Yu, Martin, & Vause, 2008); and (c) that the ABLA difficulty of training tasks that are matched to a client's highest-passed ABLA level results in fewer aberrant behaviours than tasks that are mismatched to that client's highest-passed ABLA level (Vause et al., 2000; Vause, Martin, & Yu, 1999). Research has also indicated that the pass-fail performance of testees on ABLA Level 5 is very similar to the pass-fail performance of those testees on ABLA Level 6, and that Level 5 provides relatively little additional information regarding a testee's abilities (Martin & Yu, 2000). Additional research has indicated that a useful replacement for the original ABLA Level 5 is a two-choice visual-visual nonidentity matching (VVNM) discrimination task (Sakko, Martin, Vause, Martin, & Yu, 2004). The ABLA was recently modified to include the replacement Level 5 task, and is now called the

Assessment of Basic Learning Abilities-Revised, or ABLA-R (DeWiele, Martin, Martin, Yu, & Thomson, 2011).

Several studies have also demonstrated that the ABLA performance of children with ASD can be used to match their learning ability to the ABLA difficulty of training tasks (e.g., Murphy, Martin, & Yu, 2014; Schwartzman et al., 2009). The question arises, in an EIBI program for children with ASD, does the ABLA-R level of difficulty of training tasks selected from the ABLLS-R match the learning ability of the children based on their ABLA-R performance? If not, are increased maladaptive behaviour scores found for those children? And do those children demonstrate lower rates of acquisition of new tasks? In this research, I examined these questions. My thesis research was conducted with staff and children's archived data from the St.Amant Autism Early Learning Program, an EIBI program in Winnipeg, Manitoba. First, I assessed whether ABLLS-R tasks could be reliably categorized according to the highest ABLA-R level needed to perform them. Second, for those ABLLS-R tasks that could be so categorized, I evaluated whether those tasks were actually taught at the individual ABLA-R levels into which they were categorized. Third, for a sample of children in the St.Amant Autism Early Learning Program, I compared the ABLA-R difficulty level of training tasks selected from their ABLLS-R curricula to the ABLA-R performance of those children, in order to identify the proportions of tasks that were matched as well as mismatched to the children's highest-passed ABLA-R levels. Fourth, I compared the children's proportions of mismatched tasks to their maladaptive behaviour scores, to assess whether increased maladaptive behaviour scores were found for children who were presented with larger proportions of mismatched tasks. Fifth, for that same sample of children, I compared the

proportions of mismatched tasks to their rates of acquisition of new tasks, to identify whether those children demonstrated lower rates of task acquisition when presented with larger numbers of mismatched tasks.

## Literature Review

### Behavioural Assessment

Behavioural assessment has been defined as "the collection and analysis of information and data in order to (a) identify and describe target behaviors, (b) identify possible causes of the behavior, (c) guide the selection of an appropriate behavioral treatment, and (d) evaluate treatment outcome" (Martin & Pear, 2015, p. 190). Behavioural assessment procedures have been used successfully to develop and evaluate treatments for improving target behaviours for persons with ID (e.g., Eldevik, Jahr, Eikeseth, Hastings, & Hughes, 2010; Feldman, Condillac, Tough, Hunt, & Griffiths, 2002; Feldman & Griffiths, 1997; Passey & Feldman, 2004; Wacker, Berg, Harding, & Cooper-Brown, 2011). Behavioural assessment procedures have also been used to improve treatment of children with ASD (e.g., Eldevik, Hastings, Jahr, & Hughes, 2012; Feldman et al., 2012; Kodak & Grow, 2011). As stated in the introduction, this research focused on two behavioural assessment procedures, the ABLLS-R and the ABLA-R, that have been used to improve the treatment of children with ASD. Before describing these two procedures in detail, I will review diagnostic criteria and behavioural treatment procedures for children with ASD.

### Autism

ASD is a severe and chronic developmental disorder that is typically diagnosed between the ages of 18 to 36 months. Over the last 25 years the prevalence of autism has increased; the current prevalence of autism is estimated to be 1 in 68 (Centers for Disease Control and Prevention, 2014). This increase in prevalence could be due to various factors such as environmental changes, broader diagnostic criteria for autism, increased efforts in diagnosing autism, and more parents looking for a diagnosis due to increased awareness and government funding for autism treatment services.

Children with ASD have deficits in social communication as well as social interaction, and present with restricted and repetitive patterns of behaviour, interests, or activities (American Psychiatric Association, 2013). Deficits in social communication and social interaction may present as: (a) lack of social or emotional reciprocity such as reduced sharing of enjoyment, interests, or achievements with others or failure to respond to social interactions; (b) deficits in nonverbal communicative behaviours for the purpose of social interaction, such as a lack of eye-to-eye gaze and facial expression, or abnormal body postures; and (c) deficits in developing, maintaining, and understanding relationships, such as absence of interest in peers or failure to develop age-appropriate peer relationships. Children with ASD also show restricted and repetitive patterns of behaviour, interests or activities. These may present as: (a) stereotyped or repetitive motor movements, use of objects, or speech (e.g., lining up toys or flipping objects); (b) insistence on sameness, inflexible adherence to routines, or ritualized patterns of verbal and/or nonverbal behaviour (e.g., distress related to small changes, or a need to eat the same foods every day); (c) highly restricted, fixated interests that are abnormal in intensity or focus (e.g., strong attachment to unusual objects); and (d) hyper- or

hyporeactivity to sensory input or unusual interests in sensory aspects of the environment (e.g., excessive smelling or touching of different objects, or fascination with lights or movement). Deficits in these areas must manifest during the early developmental period, although they may not fully manifest until the demands exceed the children's abilities (American Psychiatric Association, 2013). EIBI has shown to be effective in teaching the skills mentioned above; however, the effectiveness of treatment varies across individuals (e.g., Eikeseth, 2009; Eikeseth, Smith, Jahr, & Eldevik, 2002; Feldman, 2004; Howlin, Magiati, & Charman, 2009; Lovaas, 1987; Matson & Smith, 2008; McEachin, Smith, & Lovaas, 1993; Perry & Condillac, 2003; Rogers & Vismara, 2008; Virués-Ortega, 2010; Wright, 2014).

**EIBI**

EIBI for children with ASD is based on behavioural principles and procedures, such as discrimination and contingency training, reinforcement, shaping, chaining, fading, and extinction. These are typically incorporated into Discrete Trials Teaching (DTT), a teaching approach that consists of presenting multiple individual teaching trials, each one lasting 5 to 20 seconds. DTT trials typically involve a set of 12 to 20 trials, with a brief pause between each trial. Each individual teaching trial typically consists of an antecedent, a response, and a consequence (Eldevik et al., 2013; Fazzio & Martin, 2011). When the instructor issues an instruction (antecedent), a specific cue will be given to the child to emit a specific behaviour (response), and once the child has responded correctly he or she will receive immediate reinforcement, such as praise or preferred activities (consequence). The skills taught are broken down into components in order to first teach basic skills and progressively build upon a child's skill set. For example, the instructor

might start by teaching the child to label one preferred item, then two, then three, and continue building the child's repertoire in this manner. Once the child has learned to label items, he or she could learn how to request those items, and this skill could be broken down into components: First, learning to request when the item is present (in front of the child), then learning to request when the item is not present (somewhere else in the house).

Lovaas (1987) conducted one of the first outcome studies of EIBI with young children with autism using an intensive long-term treatment procedure. The children were assigned to either a control group or a treatment group. The treatment group received 40 hours per week of one-to-one treatment, whereas the control group received 10 hours or less per week of one-to-one treatment. Treatment was provided for 2 or more years. The procedures used were based on behavioural principles. Therapists also used time-outs and contingent physical aversives (e.g., a loud "No" or slap on the thigh) when targeting undesirable behaviours. Parents were actively involved in learning and implementing the treatment procedures. Treatment goals consisted of reducing self-stimulatory and aggressive behaviours, building compliance, and teaching various skills such as imitation, appropriate toy play, expressive and abstract language, interactive play with peers, expression of emotions, observational learning, and other preacademic skills. Results demonstrated that the treatment group made significantly more progress than the control group with regards to educational placement and intelligence quotient (IQ) scores. For instance, 47% of the children with autism in the treatment group attained normal intellectual and educational functioning, as opposed to 2% of the children in the control group. A follow-up study by McEachin et al. (1993) was conducted with the same two

groups of children a few years later. Results demonstrated that the treatment group maintained its gains in comparison to the control group. Furthermore, those children from the treatment group who had attained the best results in the initial study were now indistinguishable from their typically developing peers on intelligence and adaptive behaviour tests. These results suggested that EIBI for young children with autism was effective and produced long-term gains.

Many years later, Matson and Smith (2008) provided an overview of EIBI studies and reviews found in the literature. Across six review papers, conclusions regarding the effectiveness of EIBI varied from being widely accepted as the best empirically validated intervention, to questioning its long-term effects. Overall, factors such as methodology, age to begin treatment, children's IQ scores, language levels, type of behavioural treatment, components used, intensity of treatment, location of treatment, hours of treatment provided per week, and who was providing treatment were all selected as areas that needed further examination. In terms of specific client characteristics that may affect the treatment variables listed above, severity of symptoms, IQ score, and comorbid psychopathologies were identified as most important. Matson (2007) noted that aside from measuring IQ, adaptive behaviours, and communication, measures of core ASD symptoms and challenging behaviours are also necessary. Additionally, further assessments of intervention side effects, parent satisfaction, and factors maintaining challenging behaviours would also provide relevant information. One of the studies included in the review (Ruble & McGrew, 2007) conducted a survey with parents of children with ASD. Parents had rated in-home behaviour therapy as providing the best outcomes, and reported best effects when the therapy was provided for children at a

younger age. Finally, according to Matson and Smith, the large number of EIBI studies that have been conducted does suggest that not only is this treatment effective, it is also the only effective treatment to date for young children with ASD. Furthermore, research also suggests that EIBI provided within community settings such as large publicly funded programs in Canada (e.g., Perry et al., 2008) or mainstream preschool settings (e.g., Eldevik et al., 2012) can be effective alternatives to small model EIBI programs.

In recent years, several meta-analyses have been conducted to address limitations such as the ones described by Matson and Smith (2008). Reichow and Wolery (2009) conducted a three-part synthesis that included a meta-analysis of EIBI studies for children with ASD. The meta-analysis included an analysis of the mean change effect size for IQ scores, and results demonstrated a large effect size, suggesting that EIBI can be an effective treatment for children with ASD. Eldevik et al. (2009) conducted a replication and extension of Reichow and Wolery's study in order to address methodological limitations. Eldevik et al. also found that EIBI demonstrated large to moderate effect sizes for change in IQ scores and adaptive behaviour composites (communication, daily living skills, socialization, motor skills) when this group was compared to no intervention controls and eclectic services (i.e., treatment as usual for children with ASD). However, Eldevik et al. used a standardized mean difference effect size, rather than the mean change effect size used by Reichow and Wolery. Overall, these results support EIBI as the best available treatment for children with ASD, although both studies noted that the results should be interpreted with caution as several limitations were identified. These included amongst others, the small number of studies included in the meta-analyses, lack of random assignment within studies, lack of comparisons with other interventions, and

potential differences in treatment variables (e.g., frequency of supervision and training). Both studies concluded that these limitations should be addressed in future studies, to continuously improve upon and advance EIBI treatment for children with ASD.

Similarly, Virués-Ortega (2010) conducted a meta-analytical study that involved quality assessment, sensitivity analysis, meta-regression, dose-response meta-analysis and meta-analysis of studies of different metrics, in order to address the limitations of previous studies evaluating the effectiveness of early ABA intervention for children with ASD. According to Virués-Ortega, these limitations include the varying methods, designs, treatment features and quality standards of published studies. Results demonstrated that long-term comprehensive ABA treatment produced medium to large positive effects for intellectual functioning, language development, acquisition of daily living skills and social functioning for children with ASD.

In 2011, Peters-Scheffer, Didden, Korzilius, and Sturmey conducted a meta-analysis to examine the effectiveness of EIBI studies for young children with ASD, focusing on full scale verbal and performance IQ scores and adaptive behaviour. The study included EIBI groups and comparison control groups that consisted of other treatments or treatment as usual. Outcomes demonstrated that the EIBI groups outperformed the control groups on full scale and nonverbal IQ scores, receptive and expressive language, and composite adaptive behaviour. These results, as well as identified limitations, are consistent with the aforementioned findings.

In addition, Perry et al. (2011) examined four possible predictors of outcomes for children enrolled in a large publicly funded program (see Perry et al., 2008). The predictors of outcome included age at entry, IQ score, adaptive functioning, and autism

severity. Data demonstrated that IQ score at intake was clearly the greatest predictor of outcomes. As well, age at entry was an important predictor, indicating that significantly younger children with higher rates of cognitive and adaptive functioning, as well as milder autism ratings at entry, achieved greater success in the program. Another study by Eldevik et al. (2010) revealed that for children receiving behavioural intervention, their IQ scores and adaptive behaviour at intake were predictors of gains made in adaptive behaviour, while the intensity of the intervention was also a predictor for gains made in IQ scores and adaptive behaviour.

More recently, Matson & Konst (2013) reviewed 22 existing EIBI treatment studies to further assess the implementation of follow-up measures used to evaluate long-term effects of early interventions. Of the 22 studies, 19 actually reported post-test data and only three studies reported follow-up information. Overall, results demonstrated that although EIBI is the most effective treatment for children with ASD, long-term maintenance of gains has yet to be confirmed.

Currently, EIBI programs typically include 30 to 40 hours of one-on-one therapy per week, as well as a number of hours of parental participation per week. It is typically recommended that children with ASD be placed in an EIBI program before they reach the age of 4 years. A number of government-funded programs are now providing EIBI for children with ASD, and in Canada these programs are available in all of the 10 provinces (Martin & Pear, 2015). However, in such programs, if a child is taught tasks that are above or below his or her ability level (i.e., mismatched to the child's ability level), he or she may not benefit fully from the training that is provided. Currently in EIBI programs

for children with ASD, training tasks are commonly selected from curriculum guides, such as the ABLLS-R.

**Description of the ABLLS-R Curriculum**

The ABLLS-R (Partington, 2006) provides an assessment, a curriculum guide, and a skills tracking system for children who demonstrate delays in acquiring language skills. The ABLLS-R contains four categories of skills (Basic Learner Skills Assessment, Academic Skills Assessment, Self-Help Skills Assessment, and Motor Skills Assessment) that list a total of 25 skill areas. The Basic Learner Skills category contains the following skills: cooperation and reinforcer effectiveness, visual performance, receptive language, imitation, vocal imitation, requests, labeling, intraverbals, spontaneous vocalizations, syntax and grammar, play and leisure, social interaction, group instruction, follow classroom routine, and generalized responding. The Academic Skills category contains: reading, math, writing, and spelling. The Self-Help Skills category contains: dressing, eating, grooming, and toileting. Finally, the Motor Skills category contains: gross motor and fine motor. Also, each of the 25 skill areas listed above are broken down into tasks that are more or less organized hierarchically in terms of difficulty. Some examples of the tasks are presented in Table 1. Once the child has been assessed on each skill, staff will be able to create a program based on the skills that the child has already gained and the skills that the child could not yet perform. As the child progresses, staff will be able to track what the child has learned, and what following skill and task needs to be targeted.

**The ABLA and Persons With ID**

Another assessment that may be useful in guiding the selection of tasks is the ABLA (Kerr et al., 1977). The ABLA is commonly used for persons with ID.

Table 1

*Example of Tasks From the ABLLS-R Identified for the Cooperation and Reinforcer Effectiveness Skill (A)*

| Task | Task name | Task objective | Question | Examples |
|---|---|---|---|---|
| A1 | Take reinforcer when offered | When offered a known reinforcing item or activity, the student will take/use the item or activity. | When you hold out and offer a known reinforcer, will the student take the reinforcer? | M & M taken and eaten |
| A2 | Take a reinforcer from two choices of items | When offered one reinforcing item or activity and another non-reinforcing item or activity, the student will select the reinforcing item or activity. | When you hold out and offer a reinforcer and a non-reinforcing item, will the student take the reinforcer? | M & M vs. shoe, will take M & M |
| A3 | Look at a non-reinforcing item | Student will look and track changes in location of a non-reinforcing item presented by an instructor. | If you hold up a non-reinforcing item, will the student look at it and watch it as you move it to different locations in front of the student? | When you hold up a shoe and ask the student to look at the shoe, student will look at it and watch it as you move it to a variety of positions in front of him (e.g., up/ down/ left/ right) |
| A4 | Take common object when offered | When offered a common object, the student will take the item. | When you hold out and offer an item, will the student take the object? | When you hold out a shoe, student will take it |
| A5 | Approaches when a response is required for reinforcement | When a specific tangible reinforcer is available and there is a clearly identified requirement for a known response to be performed, the student will approach the instructor and perform the required response (i.e., the student knows that there is a reinforcer available but that he will need to approach the instructor and do a simple task to get the reinforcer). | When one of the child's reinforcers is available but the child must engage in a known response to receive the reinforcer, does the student approach and perform the response being requested? Do reinforcers maintain their value when simple responses (beyond approach and taking) are required? | Instructor has raisins that the student enjoys (and student is hungry), the student walks across the room and performs an imitative response to get a raisin from the instructor |

*Note*. Reprinted with permission from "The assessment of basic language and learning skills: An assessment, curriculum guide, and skills tracking system for children with autism or other developmental disabilities (ABLLS-R protocol)" by J. W. Partington, 2006, Pleasant Hill, CA: Behavior Analysts, Inc., p. 1.

Researchers have hypothesized that for people with ID, the deficits in learning certain tasks may be a function of deficits in learning the prerequisite auditory, visual, and motor discriminations. The ABLA is a useful tool by which these discriminations are assessed. It is a dynamic assessment during which a tester, using standardized prompting and reinforcement procedures, attempts to teach a testee a simple imitation and five two-choice discriminations, called levels. The ABLA assesses the ease or difficulty with which an individual can learn to reliably perform each level. These levels were selected by Kerr et al. because one or more of them appeared to be required for a client to readily learn a large number of self-care, academic, prevocational, and vocational tasks in training programs.

The levels of the ABLA include: Level 1, a simple imitation; Level 2, a two-choice position discrimination; Level 3, a two-choice visual discrimination; Level 4, a two-choice visual quasi-identity match-to-sample discrimination; Level 5, a two-choice auditory discrimination; and Level 6, a two-choice auditory-visual combined discrimination (see Table 2). Prior to testing a particular ABLA level, a demonstration, a guided trial, and an opportunity for an independent response at that level occur. Following a correct independent response, testing of that level begins. Correct responses during testing are reinforced with praise and an edible, and incorrect responses are followed by an error correction procedure, which consists of a demonstration, a guided trial, and an opportunity for an independent response. Testing of a level continues until the individual meets the pass criterion of eight consecutive correct responses, or the fail criterion of eight cumulative errors. Correct responses or errors on assisted trials (e.g., a demonstration or a guided trial) do not count towards the pass or fail criteria. The pass

Table 2

*A Description of the ABLA Levels and the Types of Discriminations Required*

| **ABLA Level** | **Test Task** | **Everyday Examples** |
|---|---|---|
| **Level 1**<br><br>**Imitation** | When given a piece of foam, can the student imitate the teacher placing the foam into a container? | Children playing Follow-the-Leader. |
| **Level 2**<br><br>**Position Discrimination** | When presented with a yellow can and a red box in a stable position, can a student consistently place a piece of foam into the container on the left? | Turning on the cold (vs. the hot) water tap. |
| **Level 3**<br><br>**Visual Discrimination** | When presented with a yellow can and a red box in randomly alternated left-right positions, can a student consistently place a piece of foam into the can? | Locating own printed name on blackboard when placed in different positions with other names. |
| **Level 4**<br><br>**Visual Identity Match-to-Sample Discrimination** | When presented with a yellow can and a red box in randomly alternated left-right positions, can a student consistently place a small yellow cylinder into the can, and a small red cube into the box? | Sorting socks into pairs. |
| **Level 5**<br><br>**Auditory Discrimination** | When presented with a yellow can and a red box (in fixed positions), can a student consistently place a piece of foam in the appropriate container when the teacher randomly says, "red box" or "yellow can"? | Responding appropriately to requests such as, "fork" vs. "spoon," when both are in a consistent location on either side of a plate. |
| **Level 6**<br><br>**Auditory-Visual Combined Discrimination** | When presented with a yellow can and a red box in randomly alternated left-right positions, can a student consistently place a piece of foam into the correct container when the teacher requests either "red box" or "yellow can"? | Responding appropriately to requests such as, "pass the salt" vs. "pass the pepper" when the salt and pepper shakers are in different places on the table from meal to meal. |

*Note*. Reprinted with permission from "The kerr meyerson assessment of basic learning abilities revised: A self-instructional manual (second edition)" by L. DeWiele, G. L. Martin, T. Martin, C. T. Yu, and K. Thomson, 2011, Winnipeg, MB: St.Amant Research Centre, p. 2 & 4.

criterion of eight cumulative correct responses was chosen because the probability that eight consecutive correct responses will occur by chance in a two-choice discrimination, assuming successive responses are independent, is quite low.

Research has shown that the levels of the ABLA are hierarchical in terms of difficulty (Kerr et al., 1977; Martin, Yu, Quinn, & Patterson, 1983; Wacker, Steil, & Greenebaum, 1983). Failed ABLA levels are difficult to teach using standard prompting and reinforcement procedures and may require hundreds of trials before the discrimination is learned, if it is learned at all (Meyerson, 1977; Stubbings & Martin, 1995, 1998; Wacker, Kerr, & Carroll, 1983; Witt & Wacker, 1981; Yu & Martin, 1986). Training tasks can be reliably analyzed according to the highest ABLA level needed to complete them (Stubbings & Martin, 1995). If the ABLA level of a training task is chosen above the client's highest-passed level, he or she may not be able to learn the task even following several hundred trials of reinforced practice. If the ABLA level of a training task is chosen at or below the client's highest-passed ABLA level, he or she is typically able to learn the task very quickly. When a client passes a level (e.g., Level 3) and fails all higher levels, the client is said to be at the highest-passed level (e.g., Level 3). Therefore, the client's ABLA level has been found to be predictive of the types of tasks that will be readily learned (e.g., simple imitation tasks or match-to-sample tasks).

Martin et al. (2008) reviewed studies that examined performance of participants with ID on the ABLA in order to predict their learning performance on (a) a variety of simple imitations and two-choice discriminations (e.g., Thorsteinsson et al., 2007), (b) three-choice and four-choice discriminations (e.g., Wacker et al., 1983), (c) the relative efficacy of three presentation modes for assessing preferences (e.g., DeVries et al., 2005),

(d) compliance of adults with ID and children with and without ID (e.g., LaForce & Feldman, 2000), and (e) participants' ability to learn to respond to the spoken names of pictures of common objects (e.g., Verbeke, Martin, Yu, & Martin, 2007). These studies demonstrated that the predictive validity of the ABLA has been very high. Matching the ABLA level of training tasks with a client's current ABLA level is important for both clients and staff, in that training tasks matched to a client's ABLA level result in fewer challenging behaviors than tasks that are mismatched to that client's ABLA level (Vause et al., 2000; Vause et al., 1999).

The ABLA has also shown to be a better indicator of a client's learning ability than experienced staff and parents. Stubbings and Martin (1998) asked staff to judge which tasks particular clients would easily master. These judgments were compared thereafter with predictions based on each client's ABLA level. Results indicated that even though each staff member had been working with each client for at least eight months, clients' ABLA performance was significantly more accurate in predicting which tasks clients would learn quickly. These results were replicated by Thorsteinsson et al. (2007).

As indicated earlier, research has also indicated that the pass-fail performance of testees on ABLA Level 5 is very similar to the pass-fail performance of those testees on ABLA Level 6, therefore Level 5 provides little additional information regarding a testee's abilities (Martin & Yu, 2000). Based on additional research (Sakko et al., 2004), a new ABLA Level 5 was developed as a two-choice visual-visual nonidentity matching (VVNM) discrimination task (see Table 3). For this task, an individual is asked to match something to something else in the environment even though the two items are not similar on any dimension (e.g., placing a cup with a saucer). More recently, the ABLA

Table 3

*The New Level 5 of the ABLA-R and an Example of an Everyday Task*

| **New ABLA Level 5** | **Test Task** | **Everyday Examples** |
|---|---|---|
| **Level 5**<br><br>**Visual Non-Identity Match-to-Sample Discrimination** | When presented with a yellow can and a red box in randomly alternated left-right positions, can a student consistently place a purple-colored piece of wood shaped like the word *Can* into the can, and a piece of silver-colored wood shaped like the word *BOX* into the box? | Placing a cup with a saucer. |

*Note*. Reprinted with permission from "The kerr meyerson assessment of basic learning abilities revised: A self-instructional manual (second edition)" by L. DeWiele, G. L. Martin, T. Martin, C. T. Yu, and K. Thomson, 2011, Winnipeg, MB: St.Amant Research Centre, p. 4.

was modified to include this new Level 5 task, and is now called the ABLA-R (DeWiele et al., 2011). All other levels have remained as described in the original ABLA, and the six levels have maintained their hierarchical order from least-to-most difficult. The ABLA-R continues to be used to assess the ease or difficulty with which an individual can learn to reliably perform each level, and just as it was stated in previous ABLA literature (e.g., Feldman, 2000), the ABLA-R remains a valid assessment of an individual's level of discrimination ability.

**The ABLA and Children With ASD**

There is a limited amount of research regarding individuals with ASD and the ABLA. Ward and Yu (2000) tested 32 children (ages 3 to 9 years; 20 with ASD) on the ABLA, and all but one child displayed the hierarchical pass-fail patterns on the ABLA consistent with that reported in previous research for persons with ID. These results demonstrated that children diagnosed with ASD follow the same progression on the ABLA as typically developing children (Casey & Kerr, 1977) and persons diagnosed with ID. As well, Ward and Yu found that individuals with ASD who passed ABLA Levels 5 and 6 communicated using two or more words, while individuals who failed Levels 5 and 6 communicated using single words or signs.

Condillac (2003) evaluated the predictive validity of various ABLA-related materials, using two sets of tasks. The A-tasks included similar containers and materials to the ABLA but with different colors, and the B-tasks included everyday materials. The study involved 46 participants with a dual diagnosis of mental retardation (MR) and autistic disorder (eight of whom had a diagnosis of PDD-NOS). Results showed an average prediction accuracy of 86% across both sets of tasks, with an average accuracy of

95% on A-tasks, and an average accuracy of 78% on B-tasks. More specifically, with the six everyday B-tasks, predictions were validated for 72% of the participants at Level 1, 57% of the participants at Level 2, 80% of the participants at Level 3, 85% of the participants at Level 4, and 87% of the participants at Level 6.

In another study, Schwartzman et al. (2009) assessed the predictive validity of the ABLA with 16 children with ASD, eight who performed at Level 4 and eight who performed at Level 6. Twenty criterion tasks were selected and analyzed according to the highest-passed ABLA level needed to perform them. Predictions were then made on whether the children would learn each of the criterion tasks based on their ABLA performance, and a parent of each child also predicted the child's learning performance on the tasks. Results demonstrated that 94% of the predictions based on ABLA performance were confirmed, and the ABLA was significantly more accurate for predicting a child's performance than were the parents. Similar results were found by Murphy et al. (2014) for children with ASD at ABLA Levels 2 and 3.

Viel et al. (2011) examined whether children with ASD who passed ABLA Level 6 (group 1) would more readily learn object naming than children who failed ABLA Level 6 (group 2). Results demonstrated that mastery of ABLA Level 6 did in fact predict object-naming performance. In other words, participants who passed ABLA Level 6 mastered a larger array of new object names (tacts) and in significantly fewer trials than participants who failed ABLA Level 6. This study was a replication of the Verbeke, Martin, Thorsteinsson, Murphy, and Yu (2009) study, conducted with individuals with ID.

These findings are valuable to therapists and workers who provide treatment services to children with ASD. If staff can correctly identify the ABLA difficulty level of training tasks and their clients' highest-passed ABLA levels, they will be able to select the most appropriate training tasks for each child, and in turn the children will learn more effectively and in fewer trials.

**Use of the ABLA to Determine if the Difficulty of Training Tasks Matches the Learning Ability of Clients in a Training Program**

Thus far, only two studies have been published that used the ABLA to determine if the ABLA difficulty of training tasks for clients in a training program matched the ABLA level of those clients, and those studies were conducted with persons with ID. As indicated previously, reference to a client's ABLA level refers to the client's highest-passed ABLA level.

DeWiele and Martin (1996) examined the use of the ABLA in a large residential training facility for persons with ID. First, they examined whether experts could reliably categorize clients' daily tasks according to the ABLA levels. DeWiele and Martin randomly selected a number of tasks from each training department and residence within the facility that would be rated by the experts. For each task a description was included, which listed the materials needed, the position of the materials, the prompts given, and the correct response. Two experts received the task descriptions and an overview of the ABLA, with instructions to read the description and to score each task according to the highest ABLA level that would be required to complete the task. Interobserver reliability (IOR) was assessed based on both sets of scores. Results demonstrated 68.6% agreement

across the six ABLA levels, with most agreements for the categorization of Level 1 tasks, and least agreements for the categorization of Level 5 tasks.

Second, DeWiele and Martin (1996) assessed the ABLA levels of training tasks typically presented to the clients at the training facility. Tasks that were categorized with 100% agreement were selected to provide a sample of tasks typically presented to clients in the training facility. About half of the tasks presented to clients only required an ABLA Level 2 discrimination.

Third, DeWiele and Martin (1996) tested a random sample of 10% of the residents on the ABLA. IORs were conducted for 44% of the tests. As the residents had a diagnostic range from borderline to profound ID, of the total sample, a number of clients were not testable for various reasons. For the remainder of the sample, ABLA levels were obtained. Results indicated that a large number of clients passed Level 6; however, the majority of tasks presented to these clients were of Level 2 difficulty. Therefore, these results suggested that the ABLA difficulty of the training tasks for these clients did not match their highest-passed ABLA level.

Vause et al. (2000) assessed the frequency of aberrant behaviour of 13 individuals with ID who were enrolled in three adult training classrooms at St.Amant, a large training facility for individuals with ID. First, Vause et al. administered the ABLA to each participant in order to obtain an ABLA level for each individual. Within this group, ABLA levels ranged from Level 2 to Level 6, with one Level 2, five Level 3s, six Level 4s, and one Level 6. Second, training tasks presented to the participants across the three classrooms were obtained and analyzed according to the highest ABLA level required to complete each task. The categorization of tasks demonstrated difficulty levels ranging

from ABLA Level 2 to Level 4. Third, the participants' aberrant behaviours were observed to produce a list of behaviours and definitions for each participant. The definitions were based on the Aberrant Behavior Checklist (ABC; Aman & Singh, 1986). Examples of aberrant behaviours included repetitive speech, rocking, restlessness, waving objects, and banging or tapping.

During the first stage of the study, Vause et al. (2000) recorded the ABLA level of training tasks presented to the participants by their instructors, and recorded instances of aberrant behaviour occurring while the participants were engaged in those tasks. During the second stage of the study, Vause et al. trained the instructors on the ABLA, and then prompted the instructors to match the ABLA difficulty level of the training tasks to the ABLA level of the participants. Vause et al. used an A-B design across three instructors to identify whether there was an increase in the presentation of matched training tasks before and after receiving ABLA training and prompting.

Results indicated that 9 of the 13 participants engaged in more aberrant behaviours when presented with training tasks that did not match their ABLA levels, and fewer aberrant behaviors when presented with tasks that were a match. Overall, 87% of training tasks were below the participants' ABLA levels. Additionally, before training and prompting the three instructors on the ABLA, 16%, 32%, and 12% of tasks presented to participants in each of the three classrooms matched their ABLA levels. After training and prompting, the percentage of matched tasks presented to participants increased to 77%, 79%, and 79% in each classroom. Finally, after teaching staff about matching the ABLA difficulty of training tasks to the ABLA level of participants within each classroom, mean aberrant behaviour decreased from 47% to 30%, 45% to 40%, and 57%

to 45%. The effect found for percentage of aberrant behaviours was not significant; however, Vause et al. (2000) suggested that this might have been due to a limited sample size. These findings support previous research suggesting that presenting training tasks matched to a client's ABLA level will result in less aberrant behaviour than the presentation of mismatched tasks (Vause et al., 1999), and that this finding occurs for mismatched training tasks that are above or below the participants' ABLA levels. Therefore, it is important for staff to ensure that many of the training tasks are matched to a client's ABLA-R level.

**The St.Amant Autism Early Learning Program for Children With ASD**

The St.Amant Autism Early Learning Program in Winnipeg, Manitoba, offers up to 3 years of EIBI services to children with ASD, and currently serves approximately 91 children with ASD across Manitoba. The program is intended for children under the age of 5 who have been diagnosed with Autistic Disorder, Asperger Syndrome, Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS), or ASD according to the DSM-IV-TR (American Psychiatric Association, 2000), and more recently, the revised diagnosis of ASD in the DSM-5 (American Psychiatric Association, 2013). The program provides two models of service, the comprehensive model and the focused model. The comprehensive model of service includes 31 hours per week of one-on-one instruction sessions and 5 hours per week of parent-led teaching for approximately 18 to 25 teaching goals. The focused model includes 15.5 hours per week of one-on-one instruction sessions and 5 hours per week of parent-led teaching for approximately 10 to 15 teaching goals. In general, each child is assigned a team, consisting of an Autism Consultant, an Autism Senior Tutor, and an Autism Tutor. The Autism Consultant is responsible for

conducting assessments and developing training programs for a number of adaptive and academic skills, as well as behavioural intervention programs to address maladaptive behaviours. The Autism Senior Tutor is responsible for assisting the Autism Consultant in training staff and parents on administering these various programs, and entering all program-related data into specific databases. The Autism Tutor is responsible for conducting daily one-on-one instruction sessions with the children, which consist of implementing and collecting trial-by-trial data for all of the programs. Skills are generally broken down into tasks, and tasks can also be broken down into a number of specific exemplars. A training program for a particular task usually involves a number of exemplars that a child must learn in order for the overall task to be considered acquired or mastered. For example, a task that involves matching objects of the same colour may be taught by breaking down each colour or each pair of objects into separate exemplars that the child must learn. Instructions are based on ABA principles and procedures, and used to teach multiple important skills, increase appropriate and desirable behaviours, and decrease inappropriate and challenging behaviours. Multiple assessments (e.g., the ABLLS-R and the ABLA-R) are conducted on a 6-month or yearly basis for all children in order to measure various skills as well as challenging behaviours. All assessment results for each client are saved in the St.Amant Autism Programs' archives, an electronic database, to ensure the data is stored in a secure and confidential manner. This archival database also allows researchers to conduct studies pertaining to large community-based early intervention programs and their effectiveness, and EIBI-related outcomes for children with autism.

For instance, Wright (2014) conducted a retrospective analysis of the archival

data from children who participated in the St.Amant Autism Early Learning Program. Assessment data was obtained for the following yearly assessment intervals: Intake, Year 1, and Year 2. Several variables were evaluated, such as language skills, adaptive functioning, autism symptom severity, cognitive functioning, and rate of development. In addition, specific outcome predictors were examined, including whether intake variables (e.g., age at entry, cognitive level, and language skills) influenced the children's outcomes, and whether "higher functioning" children demonstrated more gains when compared to "lower functioning" children. The data indicated that for 100 children, statistically and clinically significant reduction in autism symptom severity was achieved, in addition to statistically and clinically significant gains in language skills, adaptive behaviour skills, cognitive level, and rate of development following one year of EIBI treatment. For the smaller sample of 50 children with data for all of intake, Year 1, and Year 2 assessment intervals, statistically and clinically significant decreases in autism symptom severity were identified following two years of EIBI treatment, with statistically and clinically significant gains in language skills, adaptive behaviour skills, and rate of development. When comparing "higher functioning" to "lower functioning" children, the former demonstrated statistically and clinically significant gains, surpassing the "lower functioning" group on each standardized outcome measure, and was more likely to reach "typical" levels of functioning on these measures following two years of EIBI treatment. Furthermore, language and cognitive functioning at intake were the strongest predictors of outcome at Year 1.

## Statement of the Problem

Thus far, there is limited research on outcomes of early intensive intervention

offered within large publicly funded programs. However, recent studies have emphasized the need for such research, given that large community-based programs are becoming more and more common (e.g., Cohen, Amerine-Dickens, & Smith, 2006; Perry et al., 2008). For example, most EIBI programs in Canada are based on larger models, and are also to some extent, funded by the government. Therefore, not only is there a need for research within these programs to ensure children with ASD are receiving the most effective service, but evaluating the effectiveness of these programs will also provide some degree of accountability. Most importantly, this research may reveal important factors that need to be improved upon or addressed in order to provide better EIBI treatment for children with ASD.

In addition, and as indicated earlier, there are only two published studies that have used the ABLA to determine if the ABLA difficulty of training tasks for clients with ID in a training program matched the ABLA level of those clients. The results of these studies suggested that the ABLA difficulty of the clients' training tasks did not match their ability level (DeWiele and Martin, 1996), and that presenting training tasks matched to a client's ability level would result in less aberrant behaviour than the presentation of mismatched tasks (Vause et al., 2000). Research is needed to evaluate whether similar results may be found within large community based EIBI programs for children with ASD.

The purpose of this research was to evaluate whether the difficulty of the training tasks taught to children enrolled in the St.Amant Autism Early Learning Program matched the learning abilities of the children. In order to do so, I examined several questions. *First*, can observers who are knowledgeable about the ABLA-R reliably

categorize each of the 544 tasks of the ABLLS-R into an ABLA-R level per task? Based on DeWiele and Martin (1996), I predicted that observers familiar with the ABLA-R would be able to reliably categorize a number of ABLLS-R tasks into individual ABLA-R levels. *Second*, for the ABLLS-R tasks that were categorized into ABLA-R levels, do Autism Consultants from the St.Amant Autism Programs typically teach each of those ABLLS-R tasks at the ABLA-R levels into which they have been categorized (e.g., is an ABLLS-R task categorized at ABLA-R Level 4 typically taught as a Level 4 task)? As staff from the Autism Programs refer to the ABLLS-R as a skills tracking system and curriculum guide, I predicted that Autism Consultants do teach those tasks as described in the ABLLS-R, and therefore at the ABLA-R levels into which they were categorized. *Third*, based on archived data for a sample of children from the St.Amant Autism Early Learning Program, to what extent did the ABLA-R difficulty level of ABLLS-R training tasks match the learning abilities of the children as assessed by the ABLA-R? As DeWiele and Martin (1996) found with their participants, I predicted that the ABLA-R difficulty of ABLLS-R tasks taught to the children would not closely match the highest-passed ABLA-R levels of the children (referred to as their ABLA-R levels). In other words, a large proportion of training tasks would be mismatched to the children's ABLA-R levels. *Fourth*, for that sample of children, was there a relation between the proportions of training tasks mismatched to the children's ABLA-R levels and their maladaptive behavior scores? As Vause et al. (2000) found with individuals with ID, I predicted that more aberrant behaviours would be found for children with ASD who were presented with a larger number of mismatched tasks as compared to children with tasks matched to their ABLA-R levels. *Fifth*, for that same sample of children, was there correspondence

between the children's proportions of mismatched tasks and their rates of acquisition of new training tasks? I predicted that lower rates of acquisition would be found for children with ASD who were presented with a larger number of tasks mismatched to their ABLA-R levels.

## Study 1: Can ABLLS-R Tasks Be Reliably Categorized Into ABLA-R Levels?

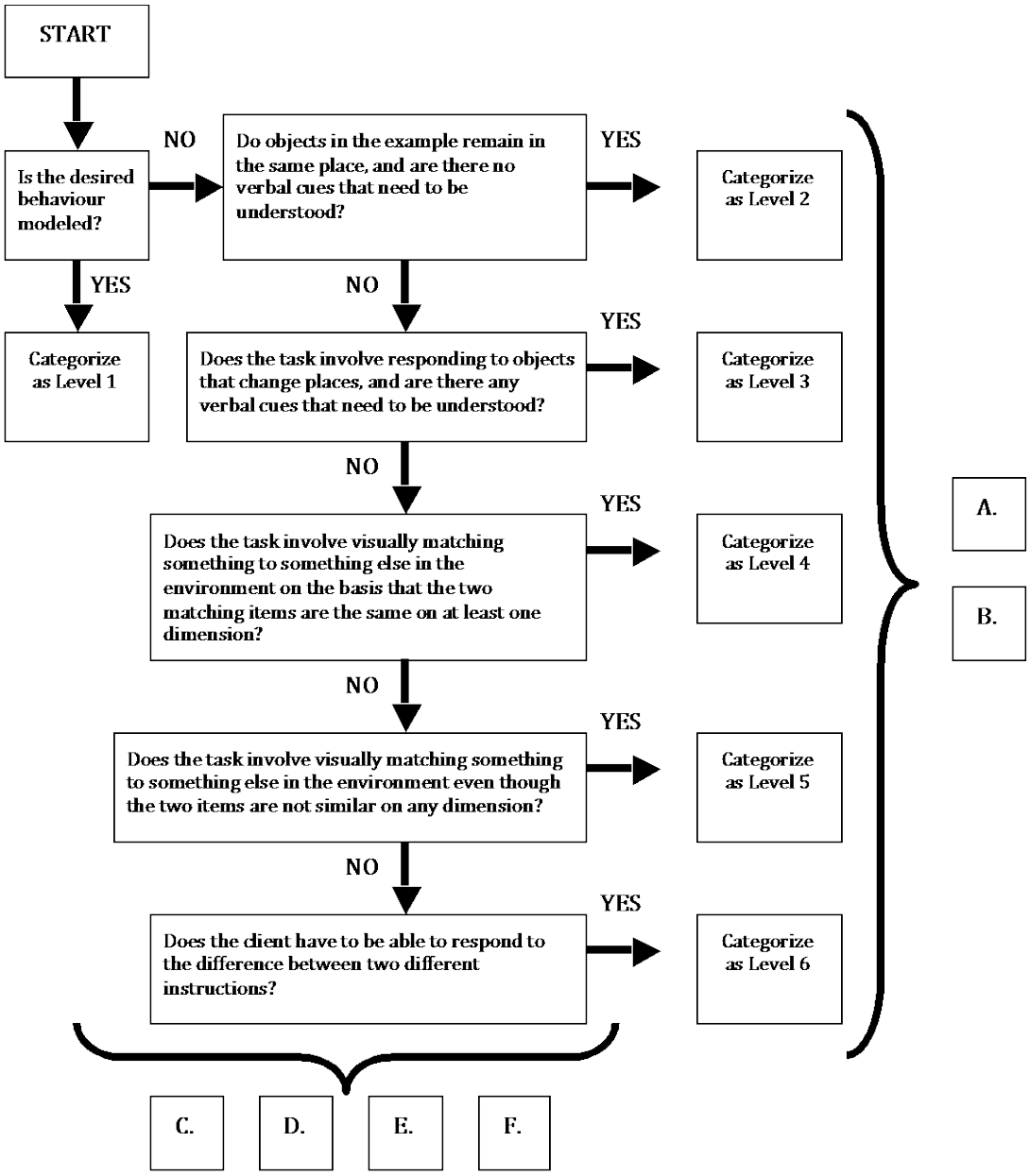
In this study, I examined the following question: Can observers who are knowledgeable about the ABLA-R reliably categorize each of the 544 tasks of the ABLLS-R into an ABLA-R level? In order to answer this question, four different phases were required.

In general, each phase was initiated to further develop the materials needed for the categorizations, to add important information that would facilitate the categorization process, and to provide additional training for the observers who executed the categorizations. The objective of every new phase was for the observers to more reliably categorize the ABLLS-R tasks; meaning that following each phase, a review of the latest categorizations took place to analyze the process, the categorization materials, and whether potential errors may have occurred due to the materials (e.g., terms were not clear, or information was missing) or due to insufficient training. Across each phase, I was able to eliminate ABLLS-R tasks that did not closely match any of the ABLA-R levels, eliminate ABLLS-R tasks that could not be reliably categorized into one of the ABLA-R levels, and finally, identify ABLLS-R tasks that could be reliably categorized according to the highest ABLA-R level needed to perform them. This study concluded once a satisfactory reliability score was obtained for all of the ABLLS-R tasks that were categorized into specific ABLA-R levels.

**Phase 1**

**Procedure**. I first reviewed guidelines (DeWiele et al., 2011) for categorizing training tasks into the highest ABLA-R level needed to perform them. In addition, all ABLLS-R tasks were reviewed to identify tasks that did not meet the guidelines identified by DeWiele et al., as well as tasks that included multiple choices, verbal responses, and more complex responses when compared to the six ABLA-R levels. This information was used to create a flowchart (see Figure 1) for categorizing ABLLS-R tasks into the ABLA-R levels needed to perform them. An observer, who had recently completed her honours thesis involving ABLA-R categorizations of training tasks taught to individuals with ID, and I used this flowchart to independently categorize all 544 ABLLS-R tasks across the 25 skill areas.

**Results.** Once the observer and I had independently categorized all 544 tasks, I compared the ABLA-R levels that we assigned to the tasks in order to identify the number of agreements and disagreements. A categorization was defined as an agreement if both of us assigned the same ABLA-R level to a task; otherwise, the categorization was defined as a disagreement. Percent agreement for categorization of ABLLS-R tasks was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100% (Martin & Pear, 2015). A predetermined IOR score of 80% agreement was identified as the passing criterion. In addition, a task was labeled as categorizable if it was found to closely approximate one of the six ABLA-R levels; otherwise, the task was labeled as non-categorizable. Results demonstrated agreement on 82% of tasks, which included agreements on tasks identified as categorizable and non-categorizable. A total of 277 ABLLS-R tasks were identified as




**A.** If the task matches one of these categorizations, but involves an array of more than two objects, categorize as **Level ___, Multiple-object discrimination**

**B.** If the task matches one of these categorizations, but requires a more complex single imitative or discriminative response, categorize as **Level ___, Complex ABLA Level**

**C.** If instruction-following is not required to complete the task, categorize as **Non-Categorizable**

**D.** If a sequence of steps is required to complete the task, categorize as **Non-Categorizable**

**E.** If a verbal response is required to complete the task, categorize as **Non-Categorizable**

**F.** If the task does not involve a discriminative or imitative response, categorize as **Non-Categorizable**

*Figure 1.* Flowchart for categorization of ABLLS-R tasks into the highest ABLA-R level needed to perform them (Phase 1).

non-categorizable, as they did not meet the guidelines outlined in the flowchart, leaving 267 tasks identified as categorizable or potentially categorizable. However, when calculating agreement for categorizable tasks alone, the observer and I demonstrated agreement on only 61% of tasks. I also found that for approximately 30 tasks, two potential categorizations had been assigned to a single task. In other words, it appeared that more than one description could match certain tasks (e.g. a multiple-object discrimination as well as a complex ABLA-R level, see Figure 1). Based on these outcomes, it appeared that additional clarification of the terms used for categorization (e.g., complex ABLA-R level) was required to facilitate the categorization process, and increase potential agreement between observers. In addition, re-categorization of the tasks was required to assess whether these changes improved reliability between observers.

**Phase 2**

**Procedure.** To improve the flowchart, I identified within our categorizations those ABLLS-R tasks that resulted in a disagreement as well as tasks that were given more than one categorization. Based on this data, I was able to either eliminate or modify particular terms used for categorization that appeared ambiguous. For instance, it became apparent that the terms *complex* and *instruction-following* were too vague, and lead to confusion when categorizing tasks. In addition, discussions relating to the flowchart indicated that this type of model might not necessarily display the guidelines in a clear and effective manner. As a result, new forms were developed to include newly revised categorization rules (see Figure 2), as well as a task categorization form (see Figure 3 and Figure 4) that contained descriptions and pertinent questions relating to the ABLA-R

**Required materials:**

a. Categorization rules
b. Task categorization forms
c. ABLLS-R manual

**Categorization Rules**

1. Use the guidelines below to categorize each task.

2. Read the task description.

3. Categorize each task as one of the following (a task may only be categorized within a single category. In other words, a task may not belong to two different categories):

    a. If the task's antecedent is not clear, mark as **Non-Categorizable**

    b. If the task does not involve a discriminative or imitative response, mark as **Non-Categorizable**

    c. A discriminative response is required, but it does not match any of the ABLA-R discriminations, mark as **Non-Categorizable**

    d. The task can be categorized as an **ABLA-R level**

4. If the task can be categorized as an ABLA-R level, review the ABLA-R table. If the task matches one of the ABLA-R categorizations in the table, **check or complete all the boxes that apply for that level.**

*Figure 2*. Categorization rules given to the observers for them to read and follow as they categorize the ABLLS-R tasks (Phase 2).

**ABLLS-R task:_____**

| **Task categorization (check one of the following):** | | |
|---|---|---|
| a | Antecedent is not clear | |
| b | Does not involve a discriminative or imitative response | |
| c | A discriminative response is required, but does not match the ABLA-R | |
| d | The task can be categorized as an ABLA-R level * | |

* If you marked the task as **d**, review and complete the ABLA-R table found on the reverse side. To complete the table, first check the box for one of the ABLA-R levels. Then, for that level, check one of the two options for each of the four columns that are applicable to the task.

*Figure 3*. First side of two-sided categorization form given to the observers for them to complete for each assigned ABLLS-R task (Phase 2).

| ABLA-R description | ABLA-R level | Antecedents and/or instructions | # of objects displayed on a trial | # of responses per trial | More than one response must occur in a sequence? |
|---|---|---|---|---|---|
| Is the desired behaviour modeled? *(imitation)* | Level 1 ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | All responses at categorized level ☐ |
| | | More complex than ABLA-R ☐ | More than ABLA-R ☐ | More than ABLA-R ☐ | Only first response at categorized level ☐ |
| Do objects in the example remain in the same place? *(position discrimination)* | Level 2 ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | All responses at categorized level ☐ |
| | | More complex than ABLA-R ☐ | More than ABLA-R ☐ | More than ABLA-R ☐ | Only first response at categorized level ☐ |
| Does the task involve responding to objects that change places? *(visual discrimination)* | Level 3 ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | All responses at categorized level ☐ |
| | | More complex than ABLA-R ☐ | More than ABLA-R ☐ | More than ABLA-R ☐ | Only first response at categorized level ☐ |
| Does the task involve visually matching something to something else in the environment on the basis that the two matching items are the same on at least one dimension? *(match-to-sample)* | Level 4 ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | All responses at categorized level ☐ |
| | | More complex than ABLA-R ☐ | More than ABLA-R ☐ | More than ABLA-R ☐ | Only first response at categorized level ☐ |
| Does the task involve visually matching something to something else in the environment even though the two items are not similar on any dimension? *(visual-visual non-identity match)* | Level 5 ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | All responses at categorized level ☐ |
| | | More complex than ABLA-R ☐ | More than ABLA-R ☐ | More than ABLA-R ☐ | Only first response at categorized level ☐ |
| Does the client have to be able to respond to the difference between two instructions? Does the task also involve identifying objects? *(auditory-visual discrimination)* | Level 6 ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | All responses at categorized level ☐ |
| | | More complex than ABLA-R ☐ | More than ABLA-R ☐ | More than ABLA-R ☐ | Only first response at categorized level ☐ |

*Figure 4.* Second side of two-sided categorization form given to the observers for them to complete for each assigned ABLLS-R task (Phase 2).

(e.g., are the antecedents and/or instructions comparable to, or more complex than, the ABLA-R?). The questions were designed to help the observers categorize the tasks, and to provide additional information that would help me understand each observer's categorizations. These new forms required an individual to read and follow the guidelines, and use the descriptions to answer all relevant questions. A single form (combination of Figures 3 and 4) was to be completed per ABLLS-R task assigned for categorization. All ABLLS-R tasks were included in the re-categorization, minus those tasks that were agreed upon in Phase 1 as non-categorizable (i.e., tasks that both the observer and I agreed did not match any of the ABLA-R levels). Therefore, the remaining 267 tasks required re-categorization with the new forms.

In order to re-categorize the ABLLS-R tasks into the highest ABLA-R level needed to perform them, and to assess the agreement scores between observers, three new observers were recruited. These observers were three undergraduate students who were registered in a Supervised Field Study course, who had taken psychology courses in the past, and who were looking to gain experience to eventually complete their honours degree and apply to graduate studies. Prior to the observers independently categorizing ABLLS-R tasks, a training session was conducted to provide an overview of both assessments (i.e., the ABLA-R and the ABLLS-R), and to discuss the forms and categorization procedure. In addition, training was provided on how to categorize ABLLS-R tasks according to the highest ABLA-R level required to perform them. Eleven training tasks were randomly selected from the 267 tasks across the 25 ABLLS-R skill areas, to be categorized by the three observers using the new guidelines and categorization form. Once any two of the three observers demonstrated 100% agreement

across two tasks, training was completed. This criterion was reached within six tasks. As these six tasks were used for training purposes, they were not included in the calculation of the final reliability score. The 261 tasks were then divided between the three observers, such that every task was assigned to two observers in order to compare the categorizations and obtain a reliability score. Consequently, each observer was assigned 174 tasks to categorize.

**Results.** Once the three observers independently categorized their designated tasks, I compared the ABLA-R levels assigned to each task by two of the observers to identify the number of agreements and disagreements, and obtain a reliability score. Results for the 261 tasks demonstrated that an agreement was reached for 70% of tasks, which included tasks identified as categorizable and non-categorizable. When looking specifically at the 79 tasks agreed categorizable (i.e., tasks that could be categorized according to an ABLA-R level), results demonstrated a reliability score of 77% for agreements on the ABLA-R level of those tasks. Based on this data, it appeared that agreements for categorizable tasks improved by approximately 16% (from the Phase 1 score of 61% agreement). However, disagreements on whether a task was considered categorizable or non-categorizable were still found for 30% of the tasks. In addition, when looking specifically at the non-categorizable tasks, disagreements were found for 69% of those tasks. In other words, the individuals demonstrated disagreement when assigning those tasks to a specific category (see Table 4 for a summary of these results). Therefore, although agreements for categorizable tasks improved, the reliability score did not meet the predetermined requirement of a minimum reliability score of 80%. Based on this information, it was determined that additional categorization training may be

Table 4

*Phase 2 Categorization Results for 261 ABLLS-R Tasks*

| Categorizations | Number of tasks | Percentage of tasks |
|---|---|---|
| Agreed categorizable | 79 | 30% |
| Agreements on ABLA-R level | 61 | 77% |
| Disagreements on ABLA-R level | 18 | 23% |
| Agreed non-categorizable | 104 | 40% |
| Agreements on category of non-categorizable | 32 | 31% |
| Disagreements on category of non-categorizable | 72 | 69% |
| Disagreements on categorization | 78 | 30% |

*Note.* Interobserver reliability (IOR) scores or percent agreement for categorization of ABLLS-R tasks was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100% (Martin & Pear, 2015).

beneficial, and further revisions to the forms were required.

**Phase 3**

**Procedure.** A third phase of categorization was initiated. For this phase, tasks to be re-categorized included tasks for which there was a disagreement on whether to identify them as categorizable or non-categorizable (78 tasks), and tasks for which there was a disagreement on the ABLA-R level determined to be the best match (18 tasks). As was done previously, tasks agreed upon as non-categorizable were excluded. Therefore, a total of 96 tasks were included in this phase. In order to obtain a valid reliability score, each task was re-assigned in such a way that it was only assigned to the observer who had not yet categorized it. In other words, rather than assigning one task to two observers so as to compare their categorizations and obtain a reliability score (as was done in Phase 2), each task was only assigned to the observer who had not yet categorized that task. Training was also provided to the three observers, using the 96 tasks. Training sessions consisted of reviewing each task and discussing it until the observers reached an agreement as to its categorization. However, to ensure that the observers did not review tasks that they had not yet categorized, they were paired for sessions according to the categorizations they had already conducted. Therefore, during training, each pair of observers only reviewed the tasks that they had categorized in the previous phase. During training sessions, I also prompted the observers to identify any difficulties that they may have experienced with the forms, and modifications were made based on their comments and suggestions. This included revised categorization rules (see Figure 5), a new form containing ABLA-R level descriptions and the types of discriminations associated with each level (see Figure 6), and a revised categorization form (see Figure 7 and Figure 8).

**Required materials:**

a. Categorization rules
b. Additional ABLA-R description
c. Task categorization forms
d. ABLLS-R manual

**Categorization Rules**

1. Use the guidelines below to categorize each task.

2. Read the task description (i.e., objective, question and highlighted criterion).

3. Identify within the description the antecedent and the response.

4. Categorize each task as one of the following (a task may only be categorized within a single category. In other words, a task may not belong to two different categories):

    a. If the task's antecedent is not clear, mark as **Non-Categorizable**

    b. If the task does not involve a discriminative or imitative response, mark as **Non-Categorizable**

    c. A discriminative response is required, but it does not match any of the ABLA-R discriminations, mark as **Non-Categorizable**

    d. The task can be categorized as an **ABLA-R level**

5. If the task can be categorized as an ABLA-R level, review the ABLA-R table. If the task matches one of the ABLA-R categorizations in the table, **check or complete all the boxes that apply for that level.**

**Notes:**

- Avoid looking at the examples provided, as they are not necessarily representative of the task objective.
- Avoid making assumptions about a task (e.g., assuming that there is a verbal instruction provided, assuming the number of objects presented).

*Figure 5.* Categorization rules given to the observers for them to read and follow as they categorize the ABLLS-R tasks (Phase 3). The shaded sections highlight new and/or revised information based on modifications that were made following Phase 2.

| **ABLA-R Level** | **Type of Discrimination** |
|---|---|
| 1) Imitation: A tester puts an object into a container and asks the client to do likewise. | A simple imitation |
| 2) Position Discrimination: When a red box and a yellow can are presented in a fixed position, a client is required to consistently place a piece of beige foam in the container on the left when the tester says, "Put it in." | A simultaneous visual discrimination with position, color, shape and size as relevant cues |
| 3) Visual Discrimination: When a red box and a yellow can are randomly presented in left-right positions, a client is required to consistently place a piece of beige foam in the yellow can when the tester says, "Put it in." | A simultaneous visual discrimination with color, shape and size as relevant cues |
| 4) Match-to-Sample Discrimination: A client demonstrates Level 4, if when allowed to view a yellow can and a red box in randomly alternating left-right positions, and is presented randomly with a yellow cylinder and a red cube, he/she consistently places a yellow cylinder in the yellow can and a red cube in the red box. | A conditional visual-visual identity discrimination with color, shape and size as relevant cues |
| 5) Visual-Visual Non-identity Match: A client demonstrates Level 5, if when allowed to view a yellow can and a red box in randomly alternating left-right positions, and when presented randomly with words 'Can' and 'BOX', he/she consistently places 'Can' in the yellow can and 'BOX' in the red box. | A conditional visual-visual nonidentity discrimination; color, shape, and size are not relevant cues |
| 6) Auditory-Visual Discrimination: The same as Level 5, except that the left-right positions of the containers are randomly alternated. | A conditional auditory-visual nonidentity discrimination, and with only color, shape and size as relevant visual cues |

*Figure 6.* ABLA-R level description form given to the observers for them to refer to as they categorize the ABLLS-R tasks (Phase 3). Adapted with permission from "Overview of research on the assessment of basic learning abilities test" by G. L. Martin and D. C. T., Yu, 2000, *Journal on Developmental Disabilities, 7,* p. 14-15.

**ABLLS-R task:_____**

| **Task categorization (check one of the following):** | | |
|---|---|---|
| a | Antecedent (i.e., what happens before the response) is not clear | |
| b | Does not involve a discriminative or imitative response | |
| c | A discriminative response is required, but does not match the ABLA-R (if in doubt, check ABLA-R descriptions found on the reverse side) | |
| d | The task can be categorized as an ABLA-R level * | |

* If you categorized the task as **d**, review and complete the ABLA-R table found on the reverse side. To complete the table, first check the box for one of the ABLA-R levels. Then, for that level, check one of the two options for each of the four columns that are applicable to the task.

*Figure 7.* First side of two-sided categorization form given to the observers for them to complete for each assigned ABLLS-R task (Phase 3). The shaded sections highlight new and/or revised information based on modifications that were made following Phase 2.

| ABLA-R description | ABLA-R level | Antecedents and/or instructions | # of objects displayed on a trial (skip if N/A or less than ABLA-R) | # of responses per trial | More than one response must occur in a sequence? (skip if N/A) |
|---|---|---|---|---|---|
| Is the desired behaviour modeled before the response occurs? Does the task involve an imitation? *(imitation)* | Level 1 ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | All responses at categorized level ☐ |
| | | More complex than ABLA-R ☐ | More than ABLA-R ☐ | More than ABLA-R ☐ | Only first response at categorized level ☐ |
| Does the task involve responding to objects that remain in the same place? *(position discrimination)* | Level 2 ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | All responses at categorized level ☐ |
| | | More complex than ABLA-R ☐ | More than ABLA-R ☐ | More than ABLA-R ☐ | Only first response at categorized level ☐ |
| Does the task involve responding to objects that change places? *(visual discrimination)* | Level 3 ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | All responses at categorized level ☐ |
| | | More complex than ABLA-R ☐ | More than ABLA-R ☐ | More than ABLA-R ☐ | Only first response at categorized level ☐ |
| Does the task involve visually matching something to something else in the environment on the basis that the two matching items are the same on at least one dimension? What is being matched? *(match-to-sample)* | Level 4 ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | All responses at categorized level ☐ |
| | | More complex than ABLA-R ☐ | More than ABLA-R ☐ | More than ABLA-R ☐ | Only first response at categorized level ☐ |
| Does the task involve visually matching something to something else in the environment even though the two items are not similar on any dimension? What is being matched? *(visual-visual non-identity match)* | Level 5 ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | All responses at categorized level ☐ |
| | | More complex than ABLA-R ☐ | More than ABLA-R ☐ | More than ABLA-R ☐ | Only first response at categorized level ☐ |
| Does the task involve an auditory discrimination as well as a visual discrimination? What is the auditory cue? What is the visual cue? *(auditory-visual discrimination)* | Level 6 ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | Comparable to ABLA-R ☐ | All responses at categorized level ☐ |
| | | More complex than ABLA-R ☐ | More than ABLA-R ☐ | More than ABLA-R ☐ | Only first response at categorized level ☐ |

*Figure 8.* Second side of two-sided categorization form given to the observers for them to complete for each assigned ABLLS-R task (Phase 3). The shaded sections highlight new and/or revised information based on modifications that were made following Phase 2.

Once training was finished, the observers were provided with their new task assignments and new forms.

**Results.** Once completed, I compared each new categorization to the previous Phase 2 categorizations, in order to identify the number of agreements and disagreements, and to calculate a new reliability score. For each of the 96 tasks, an agreement occurred if the new categorization matched either of the Phase 2 categorizations**.** A disagreement occurred if the new categorization did not match either of the Phase 2 categorizations. Results for the 96 tasks demonstrated that an agreement was reached for 88% of tasks, which included tasks identified as categorizable and non-categorizable. A reliability score of 56% was found for agreements on the ABLA-R levels of the categorizable tasks, with 15 new agreements on ABLA-R levels, while another 12 tasks were identified as categorizable with disagreements on the ABLA-R levels. In addition, 11 of the 96 tasks demonstrated a disagreement with the previous Phase 2 categorizations, such that two observers from the Phase 2 categorizations identified those 11 tasks as categorizable, but the observer from the Phase 3 categorizations identified those 11 tasks as non-categorizable. Given that the Phase 2 categorizations had identified those tasks as categorizable with disagreements on the ABLA-R level, they were left as such. Therefore, Phase 3 eliminated any disagreements on whether a task was considered categorizable or non-categorizable. In other words, all 96 tasks were agreed upon as either categorizable or non-categorizable (see Table 5 for a summary of these results).

Across Phase 2 and Phase 3, I identified a cumulative total of 99 tasks for which there was agreement that the tasks were categorizable, including 76 tasks with agreed ABLA-R levels. As a result, there was a cumulative total of 77% agreement on the ABLA-

Table 5

*Phase 3 Categorization Results for 96 ABLLS-R Tasks*

| Categorizations | Number of tasks | Percentage of tasks |
|---|---|---|
| Agreed categorizable | 27 | 28% |
| Agreements on ABLA-R level | 15 | 56% |
| Disagreements on ABLA-R level | 12 | 44% |
| Agreed non-categorizable | 58 | 60% |
| Agreements on category of non-categorizable | 32 | 55% |
| Disagreements on category of non-categorizable | 26 | 45% |
| Disagreements with Phase 2 categorizations | 11 | 11% |

*Note.* Interobserver reliability (IOR) scores or percent agreement for categorization of ABLLS-R tasks was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100% (Martin & Pear, 2015).

R levels of those 99 tasks (see Table 6). However, the requirement of a minimum reliability score of 80% had not been met, and some reliability scores were much lower than anticipated. It was suggested that a typical IOR measure, such as calculating percent agreement between observers, may not be the best measure of reliability given that the observers were ranking each categorizable task according to their level of difficulty, which corresponded to one of six ABLA-R levels. Consequently, a different reliability measure was tested to evaluate the reliability of the categorizations.

**Phase 4**

**Procedure.** The reliability of the 99 tasks that observers agreed were categorizable was tested using Cohen's weighted Kappa (Cohen, 1968). This measure was selected for two main reasons: First, it allows a measure of agreement between two or more observers, while taking into account agreements that may occur by chance. Second, the weighted kappa can be used to assign a linear weighting to the different categories, taking into account that one category (or in this case, ABLA-R level) may be more difficult than another. Thus, each ABLA-R level was rated at a value from one to six, using an ordinal scale (i.e., $1 < 2 < 3 < 4 < 5 < 6$). By assigning a linear weighting to the different levels, the test also took into consideration the extent of the disagreement. For instance, a disagreement for a task categorized by one observer as a Level 1 and by another observer as a Level 2 may be less significant than a disagreement for a task categorized by one observer as a Level 1 and by another observer as a Level 6.

**Results.** The test produced a kappa coefficient of 0.7601, with a 95% confidence interval ranging between 0.6114 and 0.9087. According to interpretation guidelines first proposed by Landis and Koch (1977), the coefficient of 0.7601 demonstrated substantial

Table 6

*Cumulative Totals of Phase 2 and Phase 3 Categorizations for 261 ABLLS-R Tasks*

| Categorizations | Number of tasks | Percentage of tasks |
|---|---|---|
| Agreed categorizable | 99 | 38% |
| Agreements on ABLA-R level | 76 | 77% |
| Disagreements on ABLA-R level | 23 | 23% |
| Agreed non-categorizable | 162 | 62% |
| Agreements on category of non-categorizable | 64 | 40% |
| Disagreements on category of non-categorizable | 98 | 60% |
| Disagreements on categorization | 0 | 0% |

*Note.* Interobserver reliability (IOR) scores or percent agreement for categorization of ABLLS-R tasks was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100% (Martin & Pear, 2015).

agreement for the 99 tasks agreed categorizable. Given this information, it was determined that for each of the 23 tasks for which there was still no agreement on the ABLA-R levels, I would randomly select one of the two observers' categorizations and assign it as the ABLA-R level. Consequently, all 99 tasks agreed categorizable were reliably assigned a corresponding ABLA-R level.

**Results Across Phases**

Based on the reliability assessments conducted in each of the four phases, results demonstrated that a number of ABLLS-R tasks could be reliably categorized according to the highest ABLA-R level needed to perform them. This study began with 544 ABLLS-R tasks, although this number was reduced to 538 tasks due to the six tasks used for training in Phase 2. When evaluating all of the 538 ABLLS-R tasks, observers agreed that a total of 439 tasks were non-categorizable, meaning that those particular tasks did not fit any of the predetermined ABLA-R guidelines. Observers agreed that a total of 99 tasks were categorizable, and none of the tasks remained as disagreements between a non-categorizable and a categorizable task (see Figure 9). Furthermore, the 99 ABLLS-R tasks that observers agreed were categorizable were all reliably categorized according to the highest ABLA-R level needed to perform them, with 35 tasks categorized as Level 1, five tasks categorized as Level 2, three tasks categorized as Level 3, 24 tasks categorized as Level 4, 13 tasks categorized as Level 5, and 19 tasks categorized as Level 6 (see Figure 10). Table 7 lists the 99 ABLLS-R tasks categorized according to ABLA-R levels. With these 99 tasks, I proceeded with Study 2.

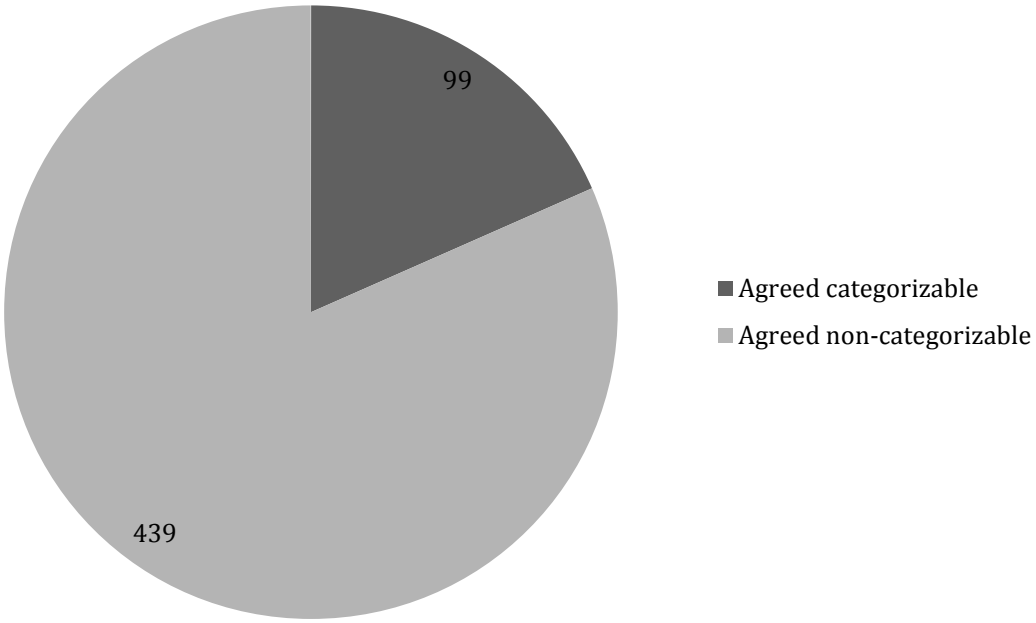



*Figure 9.* Proportions of categorizations for 538 ABLLS-R tasks.

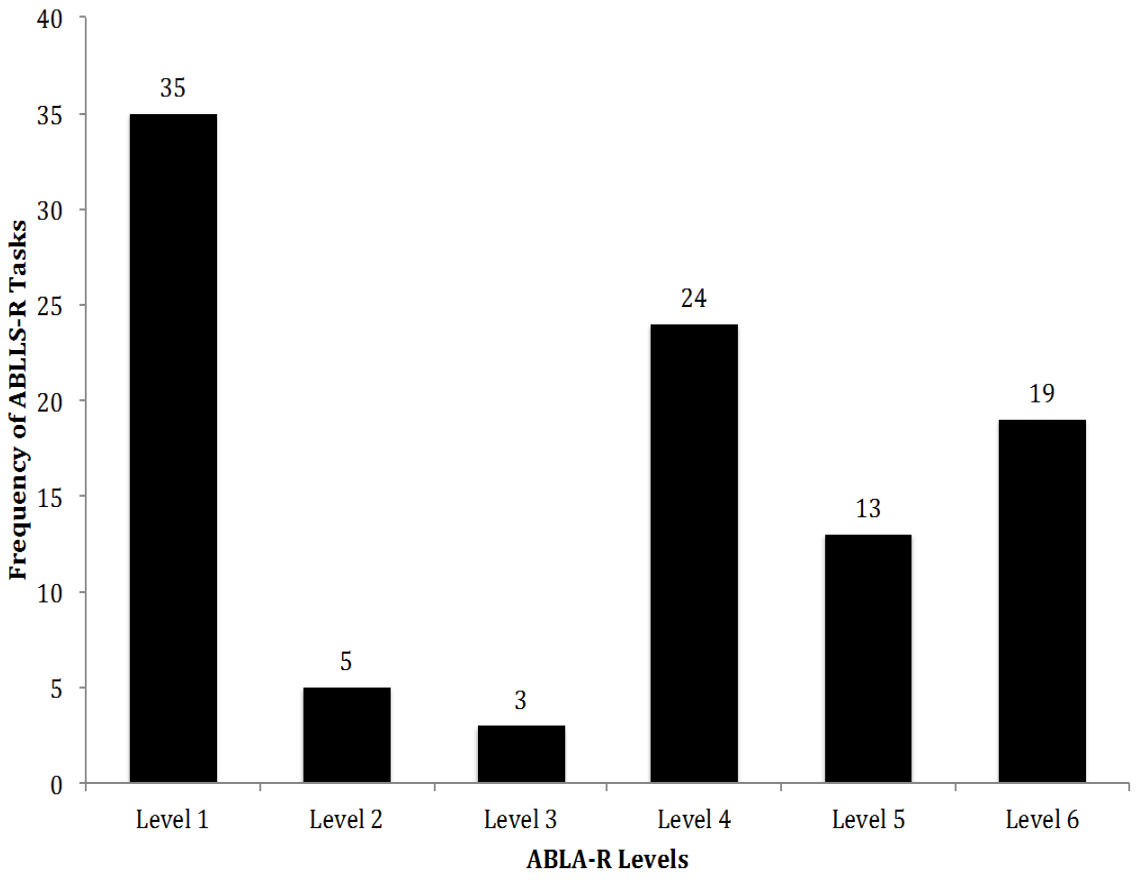



*Figure 10*. Frequency of ABLLS-R tasks with substantial agreement on the ABLA-R level based on Kappa.

Table 7

*ABLLS-R Task Categorizations With Substantial Agreement on the ABLA-R Level of Each Task Based on Kappa*

| ABLA-R level | Categorized ABLLS-R tasks | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Level 1 | B20 | D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8 | D9 | D10 | D11 |
| | D12 | D13 | D14 | D15 | D16 | D17 | D18 | D19 | D20 | D21 | D22 | D23 |
| | D24 | D26 | D27 | L8 | L9 | N6 | R1 | R3 | T3 | Z26 | Z28 | |
| Level 2 | B12 | C10 | C11 | Z2 | Z6 | | | | | | | |
| Level 3 | A3 | B21 | Q9 | | | | | | | | | |
| Level 4 | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 | B13 | B15 | C55 |
| | Q1 | Q2 | Q6 | Q8 | T2 | T4 | Z3 | Z4 | Z5 | Z7 | Z20 | Z22 |
| Level 5 | B10 | B11 | B16 | B17 | B18 | B19 | B23 | B24 | C24 | C36 | C56 | Q5 |
| | Z23 | | | | | | | | | | | |
| Level 6 | C12 | C13 | C14 | C16 | C17 | C32 | C35 | C40 | C45 | C47 | C48 | |
| | N8 | Q3 | R9 | R10 | R11 | R12 | R13 | R15 | | | | |

## Study 2: For ABLLS-R Tasks That Can Be Categorized as ABLA-R Levels, Are They Taught at Their Respective ABLA-R Levels by Training Staff in the St.Amant Autism Programs?

In this study, I examined the following question: For the 99 ABLLS-R tasks that matched ABLA-R levels, do Autism Consultants from the St.Amant Autism Programs typically teach those ABLLS-R tasks at the ABLA-R levels into which they have been categorized (e.g., is an ABLLS-R task categorized at ABLA-R Level 4 typically taught as a Level 4 task)? To address this question I created the Fidelity of Training Programs Survey, to be completed by the Autism Consultants, to determine if they taught the ABLLS-R tasks at the ABLA-R levels into which the tasks were categorized. Considering that the Autism Consultants have extremely busy caseloads, and that it would have taken many hours for them to consider how they developed programs to teach all of the 99 tasks that matched ABLA-R levels, my survey asked them to assess their teaching procedures for 12 of the 99 tasks, involving two tasks from each of the six ABLA-R levels.

### Procedure

From the 99 categorized ABLLS-R tasks identified in Study 1, I randomly selected two ABLLS-R tasks from those categorized at ABLA-R Level 1, two tasks from those categorized at Level 2, and so on for the ABLLS-R tasks at each ABLA-R level, for a total of 12 tasks. I then created a survey question for each task, with the objective of evaluating whether Autism Consultants who work for the St.Amant Autism Programs develop training programs based on the guidelines in the ABLLS-R manual. Each question prompted the Autism Consultants to review an ABLLS-R task that was categorized into an ABLA-R level, and to answer “Yes” or “No”, as to whether they developed their programs based on

the descriptions (or guidelines) provided. The descriptions for each ABLLS-R task (i.e., task objective, question, example) were obtained from the ABLLS-R manual that is used by the Autism Consultants as a curriculum guide. For a sample of two of the 12 questions, see Figure 11. The full version of the survey is in Appendix A.

The survey was administered at St.Amant, during an Autism Programs' Team Meeting. On that day, 16 of the 18 Autism Consultants were present. All Autism Consultants in attendance received a copy of the survey. A copy of the form containing ABLA-R level descriptions and discriminations associated with each level (see Figure 6, p. 41) was also attached to the survey, as an additional reference for the Autism Consultants to review when completing the survey. At the beginning of the meeting, I reviewed the project description and the consent and survey forms with the Autism Consultants, and left the forms with an Administrative Assistant. I then left the room. Autism Consultants who wished to complete the survey and who had provided their consent to participate in this study were allowed to complete the survey during the Team Meeting. As it was estimated that the survey would take approximately fifteen minutes to complete, Autism Consultants were given 15 minutes to answer the survey. During this time, the supervisors also excused themselves from the room until all surveys had been returned, given that participation was voluntary and anonymous. Completed survey forms were given to the Administrative Assistant who remained in the room. Autism Consultants who did not wish to participate could return their survey forms to the Administrative Assistant at any point during the 15-minute period. The Administrative Assistant then forwarded all of the surveys to the St.Amant Research Coordinator's office, where I collected the surveys.

**Results**

**Fidelity of Training Programs Survey**

**BACKGROUND: In an earlier phase of my study, reliability assessments were conducted to categorize ABLLS-R tasks into ABLA-R levels. We were successful in categorizing 99 ABLLS-R tasks.**

**PURPOSE: The purpose of this survey is to assess whether Autism Consultants develop training programs based on the guidelines in the ABLLS-R manual.**

**INSTRUCTIONS: Please read each question and ABLLS-R task description carefully. For each question, check either YES or NO.**

2. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 1 task? Meaning that the desired behaviour is modeled before the response occurs, and involves an imitation.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **D4** | Imitation of leg and foot movements | Upon request, student will imitate a gross motor activity involving foot and leg movements. | Will the student imitate a gross motor action involving foot and leg movements when asked to "Do this"? | Stomp foot |

8. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 4 task? Meaning that the task involves visually matching something to something else in the environment on the basis that the two matching items are the same on at least one physical dimension.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **Q8** | Match individual letters to letters on word card | The student will be able to match individual letters to the letters on cards with single 5 letters words. | Can the student match individual letters to the letters on cards with single 5 letter words? | Given a word card with the word "train", the student will match individual letter cards to the letters on the word card |

*Figure 11*. The instructions and two questions (#2 and #8) from the Fidelity of Training Programs Survey that contains 12 questions. Task descriptions adapted with permission from "The assessment of basic language and learning skills: An assessment, curriculum guide, and skills tracking system for children with autism or other developmental disabilities (ABLLS-R protocol)" by J. W. Partington, 2006, Pleasant Hill, CA: Behavior Analysts, Inc.

A total of 14 surveys were completed, and two surveys were returned incomplete. For each question on the survey, a checked “Yes” was interpreted as an agreement with the question, and a checked “No” was interpreted as a disagreement with the question. Percent agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100% (Martin & Pear, 2015). Summarized data demonstrated an average of 90.5% agreement across Autism Consultants, with percent agreement ranging from 75% to 100% (see Table 8). These results indicated that in general, training programs were developed based on the guidelines in the ABLLS-R manual, and more specifically, there was an average of 90.5% agreement that the categorized ABLLS-R tasks were taught at their respective ABLA-R levels. With this information, I was able to begin Study 3.

## Study 3: An Analysis of Three Questions Based on Archived Data From the St.Amant Autism Early Learning Program

Staff from the St.Amant Autism Programs generally administer a number of assessments approximately every six months or yearly, starting at intake, to measure various outcomes such as learning ability, level of independence with adaptive behaviours, frequency and severity of maladaptive behaviours, and skill acquisition. In this study, I examined three questions. For a sample of children with ASD for whom archived data was obtained from the St.Amant Autism Early Learning Program: (A) What percentage of training tasks were taught at, below, and above each child’s highest-passed ABLA-R level? (B) What was the relationship between the proportions of matched versus mismatched tasks and maladaptive behaviour scores? And finally, (C) what was the relationship between the proportions of matched versus mismatched tasks and rates of task acquisition?

Table 8

*Individual and Summarized Results of Survey Completed by Autism Consultants From the St.Amant Autism Programs*

| | Survey question | | | | | | | | | | | | Agreement | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Participant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Number | Percent |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 12 | 100% |
| 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 12 | 100% |
| 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 12 | 100% |
| 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 12 | 100% |
| 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 12 | 100% |
| 6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 11 | 92% |
| 7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 11 | 92% |
| 8 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 11 | 92% |
| 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 11 | 92% |
| 10 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 10 | 83% |
| 11 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 10 | 83% |
| 12 | 1 | 1 | 1 | 1 | | 1 | 1 | | 1 | 1 | 1 | 1 | 10 | 83% |
| 13 | 1 | 1 | | | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 9 | 75% |
| 14 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 9 | 75% |
| Total average | | | | | | | | | | | | | 11 | 90.5% |

*Note*. 1 = “Yes” or agreement, and indicates that the ABLLS-R task described in the survey question was taught at its categorized ABLA-R level; 0 = “No” or disagreement, and indicates that the ABLLS-R task described in the survey question was not taught at its categorized ABLA-R level; blank = no response.

**Participants**

Existing client data was obtained from the archival database for 23 children who received 1 to 3 years of services between 2007 and 2013 from the St.Amant Autism Early Learning Program. However, nine children were excluded from the study due to missing data that was required to conduct the necessary comparisons (as described previously). Therefore, the sample for all three parts of Study 3 consisted of 14 children who met the inclusion criteria. Inclusion criteria for each child involved available archived data for all specified assessments, namely the ABLA-R and the Scales of Independent Behavior-Revised (SIB-R; Bruininks, Woodcock, Weatherman, & Hill, 1996), as well as archived data from the ABLLS-R training task databases, all of which had to be available for any given assessment period. For example, when I examined Year 1 assessment data for a child, scores were required for the ABLA-R and the SIB-R, and corresponding ABLLS-R training task data was required within that same time period. Furthermore, given that certain assessment periods rarely contained all of the required data (e.g., Intake, Year 2.5, Year 3), assessment periods selected for this study were limited to Year 1, Year 1.5, and Year 2, such that the children who had all required data for any of these three assessment periods met the inclusion criteria. It should be noted that for the purposes of this study, personal information such as age, gender, specific diagnosis, etc. was not required for data analysis. Consequently, none of this information was made available throughout data collection. To maintain anonymity and confidentiality, staff from the St.Amant Autism Programs identified clients within the archival database for whom the necessary data was present, and that data was saved in a separate database for me to retrieve and analyze.

## Part A: What Percentage of Training Tasks Were Taught At, Below, and Above Each Child's ABLA-R Level?

### Measures: The ABLA-R and the ABLLS-R.

***ABLA-R.*** As described in the literature review, the ABLA-R (DeWiele et al., 2011) assesses a client's basic learning ability; in other words, the ease or difficulty with which an individual can learn to reliably perform a simple imitation and five two-choice discriminations, organized in hierarchical levels. In order to compare each child's reliably categorized ABLLS-R training tasks to his or her respective ability level, the ABLA-R was selected as the measure of basic learning ability, although the children's original test scores were based on the ABLA.

***ABLLS-R.*** As described in the literature review, the ABLLS-R (Partington, 2006) provides an assessment, a curriculum guide, and a skills tracking system for children who demonstrate delays in acquiring language skills. The ABLLS-R contains four categories of skills (Basic Learner Skills Assessment, Academic Skills Assessment, Self-Help Skills Assessment, and Motor Skills Assessment) that list a total of 25 skill areas. The 25 ABLLS-R skills, and the tasks broken down within each skill, are described in the current ABLLS-R manual. As demonstrated in Study 2, Autism Consultants typically developed training programs based on the ABLLS-R guidelines, thus, each child's training tasks typically reflected the skills and tasks outlined in the ABLLS-R manual. Each child's training task data is typically collected and then entered in a database, and when the child exits the St.Amant Autism Early Learning Program, the database is archived. For this study, I analyzed each child's archived ABLLS-R training task data, as it was entered in their respective databases. The ABLLS-R training task data involved tasks as well as

their respective exemplars, which were either mastered (i.e., the task or exemplar was learned) or in progress (i.e., the child was learning that particular task or exemplar).

**Procedure.**

For all 14 children who met the inclusion criteria, I analyzed each of their archived data for any given assessment period in the following manner. First, I compared each child's ABLA-R assessments to his or her ABLLS-R training task database, to ensure the dates for a particular assessment period corresponded. As I wanted to analyze their assessment scores as well as their training tasks within the same time period, I selected training task data that corresponded to tasks that were taught within 1 month prior to the assessment dates. Typically, the evaluations for a specific assessment period (e.g., Year 1) are conducted within 1 to 2 months prior to the child's assessment report due date. Therefore, all assessment scores and training task data corresponded approximately to the same 1-month period.

Second, I identified within the 1 month of training task data, those tasks that corresponded to the 99 ABLLS-R tasks that were reliably categorized according to the highest ABLA-R level needed to perform them, in Study 1. This provided a list of training tasks and their respective ABLA-R levels. The list of tasks also included each task's exemplars (mastered or in progress). For any given task, the exemplars were considered to be the same ABLA-R level as the task itself, because the exemplars involve variations of the same task. However, for the purposes of this study, exemplars were considered individual tasks.

Third, I compared this list to the ABLA-R score for that assessment period, to identify the training tasks that were *at* the child's highest-passed ABLA-R level, *below*

the child's ABLA-R level, and *above* the child's ABLA-R level. A task at the child's ABLA-R level signified that the task's categorized ABLA-R level was identical to the child's highest-passed ABLA-R level (e.g., a task categorized as a Level 4 and a child whose highest-passed level was Level 4), and was therefore labeled a "match." A task below the child's ABLA-R level signified that the task's categorized ABLA-R level was at a level below the child's highest-passed ABLA-R level (e.g., a task categorized as a Level 4 and a child whose highest-passed level was either Level 5 or Level 6), and was therefore labeled a "mismatch." A task above the child's ABLA-R level signified that the task's categorized ABLA-R level was at a level above the child's highest-passed ABLA-R level (e.g., a task categorized as a Level 4 and a child whose highest-passed level was either Level 1, Level 2, or Level 3), and was also labeled a "mismatch."

Fourth, I determined the percentage of tasks taught at, below, and above each child's ABLA-R level by calculating the number of tasks matched or mismatched, dividing by the total number of tasks, and multiplying by 100%. For specific mismatched percentages, I divided the number of tasks mismatched below by the total number of mismatched tasks, and I divided the number of tasks mismatched above by the total number of mismatched tasks. This procedure was repeated for every assessment period selected for analysis, across the sample of 14 children. See Appendix B for the corresponding data collection sheet.

**Results.**

As presented in Table 9, data was analyzed for a total of 31 assessment periods, with three children who each had one available assessment period, and 11 children who each had either two or three available assessment periods. ABLA-R levels varied across

Table 9

*Percentage of Tasks At, Below, and Above Each Child's ABLA-R Level*

| Participant | Assessment period | ABLA-R level | Number of tasks | Percentage of tasks | | | |
|---|---|---|---|---|---|---|---|
| | | | | Matched | Mismatched | Mismatched below | Mismatched above |
| 1 | 1 | 6 | 38 | 0% | 100% | 100% | 0% |
| | 2 | 6 | 14 | 43% | 57% | 100% | 0% |
| 2 | 1 | 3 | 19 | 0% | 100% | 58% | 42% |
| | 1.5 | 6 | 50 | 0% | 100% | 100% | 0% |
| | 2 | 6 | 49 | 22% | 78% | 100% | 0% |
| 3 | 1 | 4 | 20 | 50% | 50% | 30% | 70% |
| 4 | 1 | 2 | 17 | 0% | 100% | 29% | 71% |
| | 1.5 | 2 | 19 | 0% | 100% | 58% | 42% |
| | 2 | 3 | 21 | 5% | 95% | 20% | 80% |
| 5 | 1 | 1 | 22 | 68% | 32% | 0% | 100% |
| | 2 | 3 | 35 | 3% | 97% | 71% | 29% |
| 6 | 1 | 4 | 84 | 49% | 51% | 65% | 35% |
| | 1.5 | 4 | 35 | 11% | 89% | 55% | 45% |
| | 2 | 6 | 55 | 5% | 95% | 100% | 0% |
| 7 | 1 | 4 | 74 | 1% | 99% | 25% | 75% |
| | 2 | 6 | 36 | 17% | 83% | 100% | 0% |
| 8 | 1 | 3 | 58 | 0% | 100% | 36% | 64% |
| | 1.5 | 3 | 26 | 0% | 100% | 50% | 50% |
| | 2 | 3 | 26 | 0% | 100% | 69% | 31% |
| 9 | 1 | 4 | 31 | 39% | 61% | 32% | 68% |
| | 1.5 | 6 | 33 | 0% | 100% | 100% | 0% |
| | 2 | 6 | 10 | 30% | 70% | 100% | 0% |
| 10 | 1 | 4 | 42 | 12% | 88% | 51% | 49% |
| | 1.5 | 4 | 24 | 21% | 79% | 58% | 42% |
| | 2 | 4 | 29 | 14% | 86% | 96% | 4% |
| 11 | 1.5 | 4 | 6 | 0% | 100% | 50% | 50% |
| 12 | 1 | 4 | 18 | 22% | 78% | 79% | 21% |
| 13 | 1 | 4 | 42 | 52% | 48% | 80% | 20% |
| | 2 | 4 | 38 | 45% | 55% | 90% | 10% |
| 14 | 1 | 4 | 23 | 52% | 48% | 27% | 73% |
| | 2 | 4 | 12 | 42% | 58% | 57% | 43% |
| Total | | | 1006 | | | | |
| Average | | | 32 | 19% | 81% | 64% | 36% |

*Note*. The children's original test scores were based on the ABLA.

assessment periods, with a total of one assessment with a child at Level 1, two assessments with children at Level 2, six assessments with children at Level 3, fourteen assessments with children at Level 4, zero assessments with children at Level 5, and eight assessments with children at Level 6 (see Figure 12). Note that as research demonstrated that ABLA Level 5 added little additional information (Martin & Yu, 2000), and that most children who passed Level 5 could also pass Level 6, staff from the St.Amant Autism Programs did not test Level 5 when administering the ABLA during an assessment period. In other words, if a child was able to pass Level 4, the testee would skip Level 5 and proceed with testing the child on Level 6. Also included in the analysis was a total of 1006 tasks (including each task's respective exemplars), used to calculate the percentage of tasks at, below, and above each child's ABLA-R level. The number of tasks per child ranged from 6 to 84, with an average of 32 tasks per child. This number varied according to the number of reliably categorized training tasks that were identified from the list of tasks taught within 1 month prior to the assessment dates.

Overall, results demonstrated that an average of 19% of tasks were matched and 81% of tasks were mismatched to the children's ABLA-R levels. Of those mismatched tasks, 64% were mismatched below the children's ABLA-R levels, and 36% were mismatched above the children's ABLA-R levels (see Figure 13). Table 9 presents each child's specific percentages.

**Part B: What Was the Relationship Between the Proportions of Matched Versus Mismatched Tasks and Maladaptive Behaviour Scores?**

**Measure: The SIB-R.**

***SIB-R.*** Children's maladaptive behaviour scores reflected the scores of the

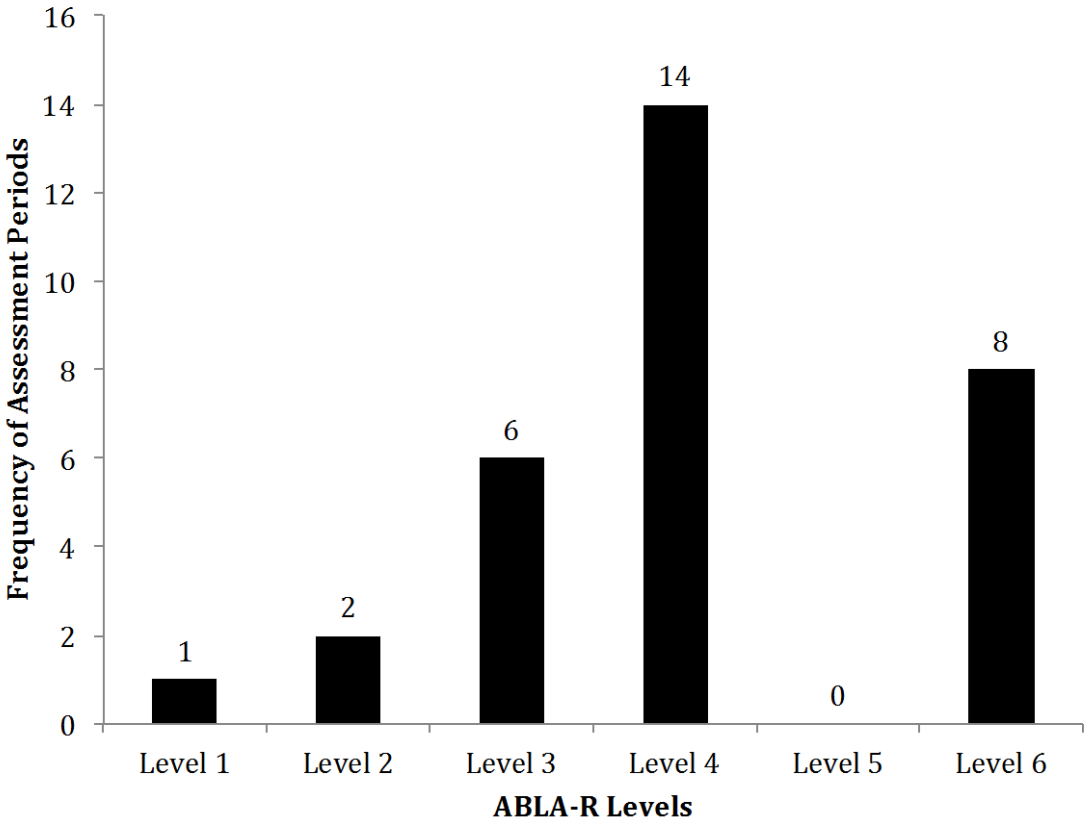



*Figure 12*. Total frequency of ABLA-R levels across assessment periods. Note that the children's original test scores were based on the ABLA.

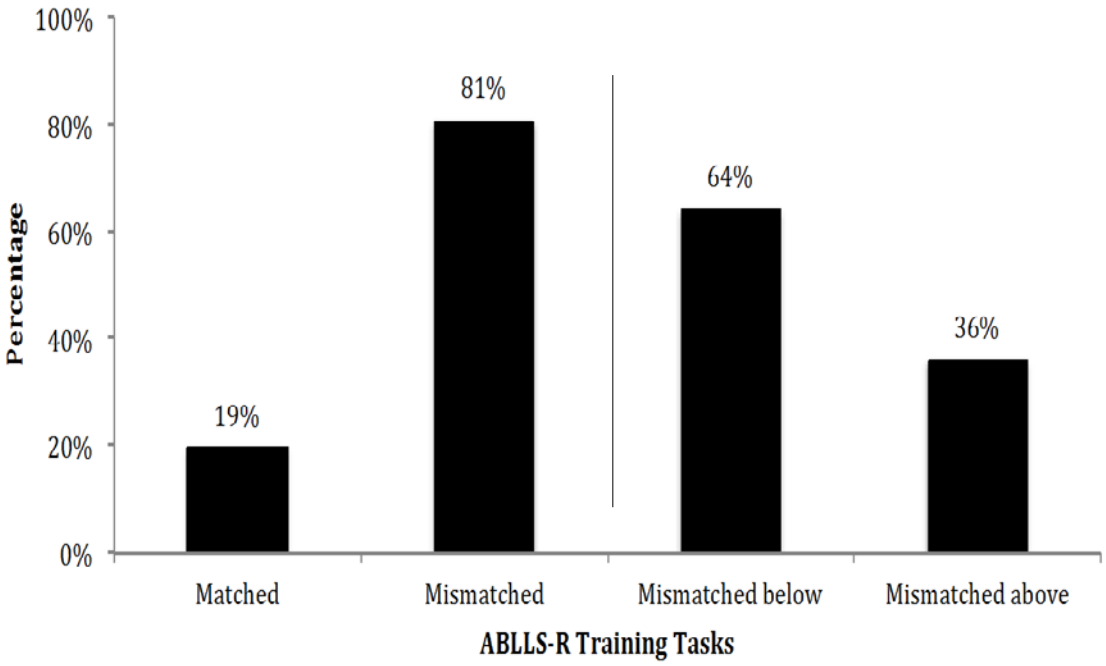



*Figure 13*. Average percentages of tasks matched, mismatched, mismatched below, and mismatched above children's highest-passed ABLA-R levels.

SIB-R. The SIB-R is a comprehensive, norm-referenced assessment of 14 areas of adaptive behaviours and 8 areas of maladaptive behaviours. It can be administered by means of three different forms (Full Scale, Short Form, Early Development Form), and all of the forms allow the tester to assess the client's challenging behaviours. However, it is important to note that this assessment only offers an overview of a child's present strengths and challenges, as evaluated by the person who is being interviewed (e.g., parents, other caregivers). As such, the results represent the interviewee's perspective regarding the child's level of independence and level of frequency and severity of challenging (maladaptive) behaviours. When assessing the maladaptive behaviours, the severity and frequency of the behaviours are measured. The eight maladaptive areas include: hurtful to self, hurtful to others, destructive to property, disruptive behaviour, unusual or repetitive habits, socially offensive behaviour, withdrawal or inattentive behaviour, and uncooperative behaviour. These eight areas are grouped to provide three maladaptive behaviour indexes/sub-scores: The Internalized Index (hurtful to self, unusual or repetitive habits, withdrawal or inattentive behaviour), the Asocial Index (socially offensive behaviour, uncooperative behaviour), and the Externalized Index (hurtful to others, destructive to property, disruptive behaviour). The General Maladaptive Index (GMI) is a composite score derived from the three sub-scores, and for the purposes of this study, the composite score was chosen as the measure of maladaptive behaviour for each child.

**Procedure.**

For all 14 children and across their 31 assessment periods, I obtained an SIB-R GMI score. I then referred to the SIB-R Maladaptive Behavior Indexes Profile to assign a

corresponding label to each score (i.e., Normal, Marginally serious, Moderately serious, Serious, or Very serious). For instance, a score of -10 and above is considered normal, a score between -11 and -20 is considered marginally serious, a score of -21 to -30 is considered moderately serious, a score between -31 and -40 is considered serious, and a score of -41 and below is considered very serious. Note that any score other than what was considered normal was recognized as a score representative of maladaptive behaviours, regardless of whether it was considered marginal or serious. Finally, I compared the children's maladaptive behaviour scores to their proportions of matched and mismatched tasks from Part A.

**Results.**

Across 31 SIB-R assessments, 12 assessments (39%) demonstrated normal behaviour scores and 19 assessments (61%) demonstrated maladaptive behaviour scores. More specifically, nine assessments (6 children) demonstrated marginally serious scores, three assessments (3 children) demonstrated moderately serious scores, five assessments (4 children) demonstrated serious scores, and two assessments (2 children) demonstrated very serious scores (see Figure 14). A Pearson product-moment correlation was tested between average percentages of matched tasks and average maladaptive behaviour scores across 12 children (for whom a percentage of tasks matched their ABLA-R level). Each child's average percentage of tasks or average maladaptive behaviour score was calculated by averaging the data across his or her respective assessment periods. The result ($r$ = -.436 , $p$ = .157) demonstrated that the relationship between children's average proportions of matched tasks and maladaptive behaviour scores was not statistically significant at a $p$ value of .05. The same test was conducted between average percentages

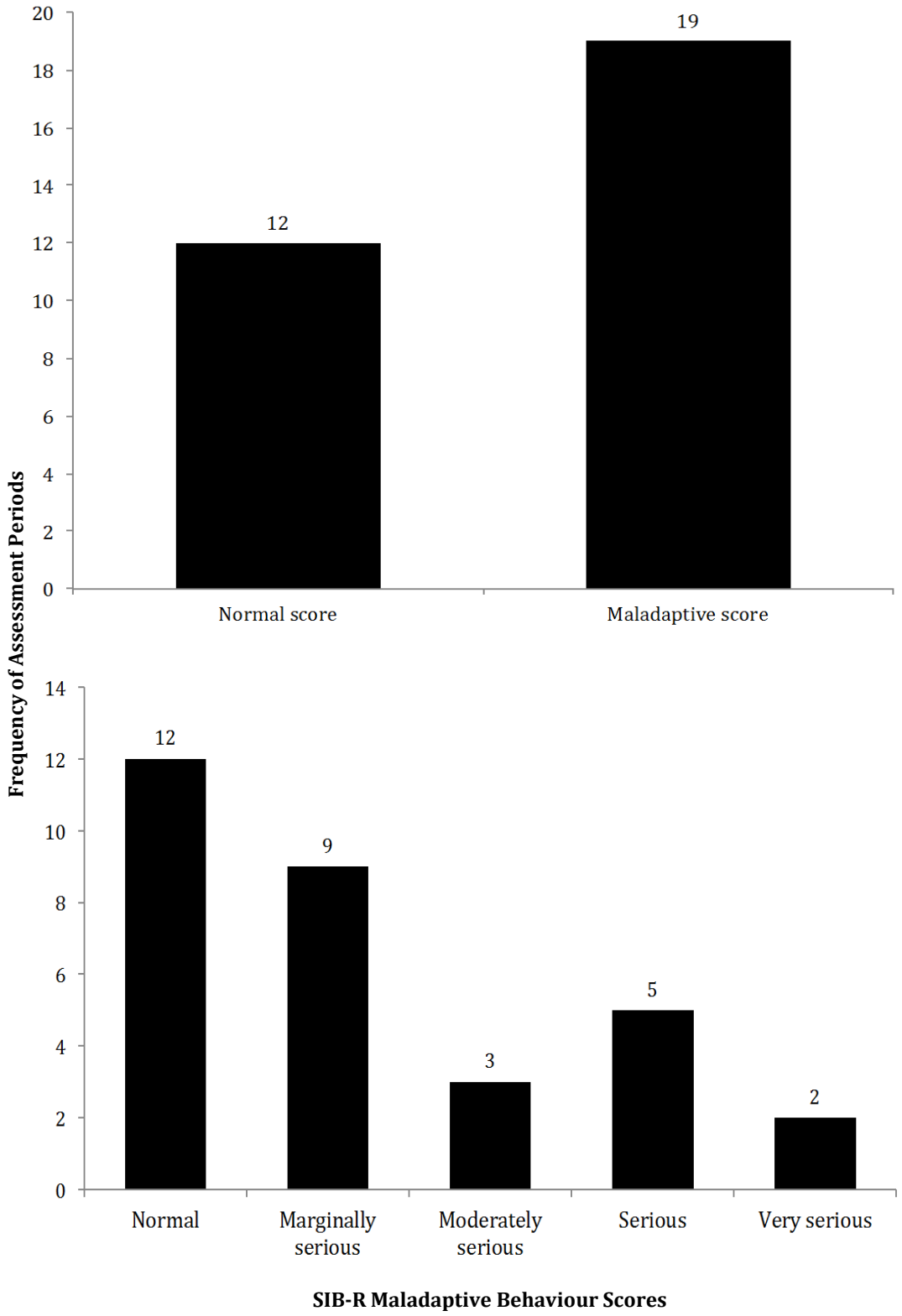



*Figure 14*. Frequency of assessment periods with normal and maladaptive behaviour scores.

of mismatched tasks and average maladaptive behaviour scores across all 14 children. The result ($r$ = -.047, $p$ = .873) demonstrated that the relationship between children's average proportions of mismatched tasks and maladaptive behaviour scores was not statistically significant at a $p$ value of .05. Both outcomes may be due to sample size restrictions. Nevertheless, more maladaptive behaviour scores were found for children in this sample, and overall, these children were presented with larger numbers of mismatched tasks versus tasks matched to their learning ability. See Table 10 for each child's respective maladaptive behaviour scores.

## Part C: What Was the Relationship Between the Proportions of Matched Versus Mismatched Tasks and Rates of Task Acquisition?

### Measure.

Trial-by-trial data is consistently collected for ABLLS-R training tasks taught during instruction sessions. Consequently, the third question was examined by using a trials-to-criterion measure.

An average rate of training task acquisition was obtained by calculating, for each child, the total number of trials required to meet the mastery criterion (i.e., criterion that demonstrates understanding of the task) divided by the total number of tasks. Note that across children and training programs, the mastery criterion may vary and is determined by the Autism Consultants to reflect the task's level of difficulty based on the child's current level of ability. Nevertheless, it provides an overall trials-to-criterion measure.

### Procedure.

For all 14 children and across 30 assessment periods (one child did not master any tasks during a particular 1-month period), I obtained trials-to-criterion data for a total of

Table 10

*Summary of Percentage of Matched and Mismatched Tasks With SIB-R GMI Scores and Profiles Across Assessment Periods*

| Participant | Assessment period | Percentage of tasks | | | | SIB-R GMI score | Index profile |
|---|---|---|---|---|---|---|---|
| | | Matched | Mismatched | Mismatched below | Mismatched above | | |
| 1 | 1 | 0% | 100% | 100% | 0% | -4 | Normal |
| | 2 | 43% | 57% | 100% | 0% | -21 | Moderately serious |
| 2 | 1 | 0% | 100% | 58% | 42% | -4 | Normal |
| | 1.5 | 0% | 100% | 100% | 0% | -3 | Normal |
| | 2 | 22% | 78% | 100% | 0% | -4 | Normal |
| 3 | 1 | 50% | 50% | 30% | 70% | -14 | Marginally serious |
| 5 | 1 | 0% | 100% | 29% | 71% | -11 | Marginally serious |
| | 1.5 | 0% | 100% | 58% | 42% | -26 | Moderately serious |
| | 2 | 5% | 95% | 20% | 80% | -33 | Serious |
| 6 | 1 | 68% | 32% | 0% | 100% | -32 | Serious |
| | 2 | 3% | 97% | 71% | 29% | -24 | Moderately serious |
| 7 | 1 | 49% | 51% | 65% | 35% | -18 | Marginally serious |
| | 1.5 | 11% | 89% | 55% | 45% | -17 | Marginally serious |
| | 2 | 5% | 95% | 100% | 0% | -15 | Marginally serious |
| 8 | 1 | 1% | 99% | 25% | 75% | 1 | Normal |
| | 2 | 17% | 83% | 100% | 0% | -6 | Normal |
| 11 | 1 | 0% | 100% | 36% | 64% | -35 | Serious |
| | 1.5 | 0% | 100% | 50% | 50% | -9 | Normal |
| | 2 | 0% | 100% | 69% | 31% | -32 | Serious |
| 13 | 1 | 39% | 61% | 32% | 68% | -16 | Marginally serious |
| | 1.5 | 0% | 100% | 100% | 0% | -12 | Marginally serious |
| | 2 | 30% | 70% | 100% | 0% | -8 | Normal |
| 14 | 1 | 12% | 88% | 51% | 49% | -6 | Normal |
| | 1.5 | 21% | 79% | 58% | 42% | -9 | Normal |
| | 2 | 14% | 86% | 96% | 4% | -10 | Normal |
| 15 | 1.5 | 0% | 100% | 50% | 50% | -52 | Very serious |
| 17 | 1 | 22% | 78% | 79% | 21% | -18 | Marginally serious |
| 18 | 1 | 52% | 48% | 80% | 20% | -2 | Normal |
| | 2 | 45% | 55% | 90% | 10% | -20 | Marginally serious |
| 19 | 1 | 52% | 48% | 27% | 73% | -42 | Very serious |
| | 2 | 42% | 58% | 57% | 43% | -38 | Serious |
| Average | | 19% | 81% | 64% | 36% | -17 | Marginally serious |

*Note*. GMI = General Maladaptive Index. The maladaptive behaviour index values are interpreted in the SIB-R as follows: +10 to -10 = Normal; -11 to -20 = Marginally serious; -21 to -30 = Moderately serious; -31 to -40 = Serious; -41 and below = Very serious.

462 mastered tasks (including each task's respective exemplars) corresponding to the 99 ABLLS-R tasks from Study 1. With this trials-to-criterion data, I calculated the number of trials required to achieve the mastery criterion for each task. For each child, I then proceeded in calculating the average number of trials required to achieve the mastery criterion for tasks that matched the child's ABLA-R level, and tasks that were a mismatch below and above the child's ABLA-R level. The average number of trials was calculated by dividing the total number of trials, corresponding to tasks matched, mismatched below, or mismatched above the child's ABLA-R level, by the total number of tasks, and multiplying by 100%. This provided individual rates of acquisition, in addition to a total average rate of acquisition across all children. See Appendix C for the corresponding data sheet.

**Results.**

Table 11 presents the children's individual rates of acquisition for tasks matched, mismatched below, and mismatched above their highest-passed ABLA-R levels. Overall, results demonstrated that children required an average of 25 trials to master tasks matched to their ABLA-R level, and 42 trials to master tasks mismatched to their ABLA-R level. More specifically, children required an average of 47 trials to master tasks mismatched below their ABLA-R level, and 34 trials to master tasks mismatched above their ABLA-R level (see Figure 15). As shown in Table 11, the availability of data corresponding to tasks matched, mismatched below, and mismatched above was variable. For instance, not all children had tasks that matched their ABLA-R level, for a given assessment period. Thus, the average rate of acquisition for matched tasks was based on data for 9 children and 12 assessment periods, while the average rate of acquisition for

Table 11

*Average Rates of Acquisition for Tasks Matched, Mismatched Below, and Mismatched Above the Children's ABLA-R Levels*

| Participant | Assessment period | Average rate of acquisition | | |
|---|---|---|---|---|
| | | Tasks matched to ABLA level | Tasks below ABLA level | Tasks above ABLA level |
| 1 | 1 | | 18 | |
| | 2 | 8 | 33 | |
| 2 | 1 | | 22 | 31 |
| | 1.5 | | | |
| | 2 | 6 | 74 | |
| 3 | 1 | 26 | 29 | |
| 4 | 1 | | 4 | 3 |
| | 1.5 | | 25 | 60 |
| | 2 | | | 39 |
| 5 | 1 | | | 31 |
| | 2 | | 178 | |
| 6 | 1 | 3 | 15 | 39 |
| | 1.5 | | 38 | 168 |
| | 2 | | 30 | |
| 7 | 1 | | 38 | 11 |
| | 2 | | 29 | |
| 8 | 1 | | 208 | 15 |
| | 1.5 | | 100 | |
| | 2 | | 95 | 19 |
| 9 | 1 | 14 | 32 | 43 |
| | 1.5 | | 30 | |
| | 2 | | 29 | |
| 10 | 1 | 73 | 52 | 9 |
| | 1.5 | 16 | 70 | 67 |
| | 2 | 113 | 64 | |
| 11 | 1.5 | | 85 | 46 |
| 12 | 1 | 96 | 61 | |
| 13 | 1 | 44 | 15 | |
| | 2 | 79 | 80 | 5 |
| 14 | 1 | 20 | | 1 |
| | 2 | | 48 | 485 |

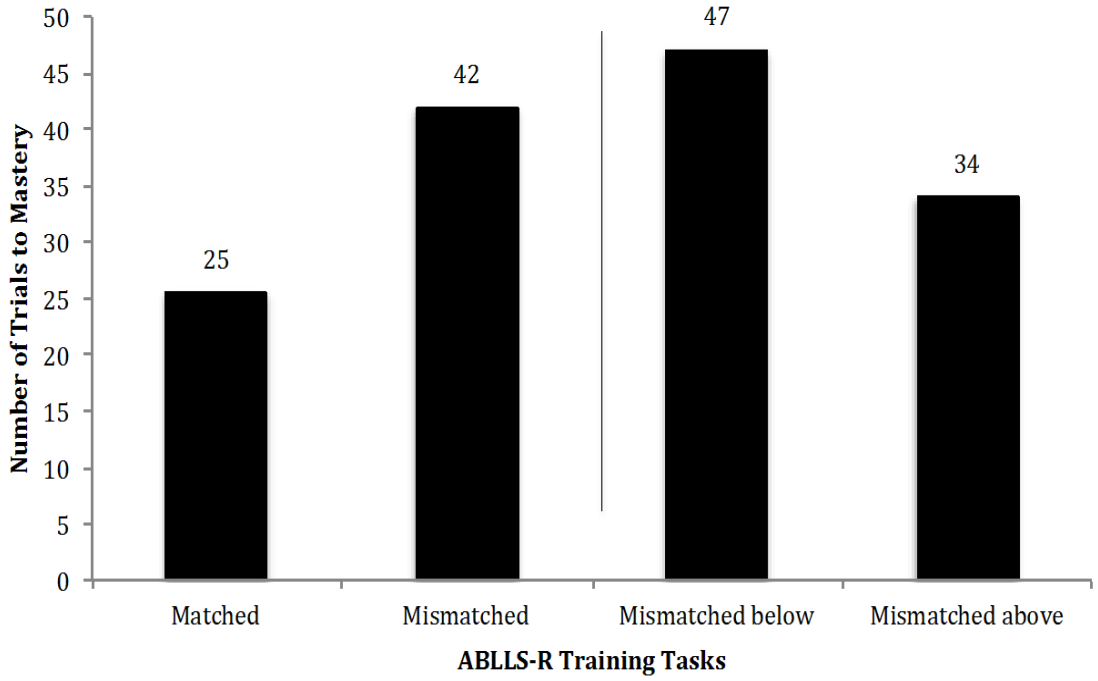



*Figure 15*. Average rates of acquisition for tasks matched, mismatched, mismatched below, and mismatched above the children's highest-passed ABLA-R levels. Average rates were calculated based on trials-to-criterion data for 462 ABLLS-R training tasks.

mismatched tasks was based on data for all 14 children and 30 assessment periods. A paired-samples *t*-test was conducted to compare average rates of acquisition for matched versus mismatched tasks across 9 children (children who had rates for both matched and mismatched tasks). Results demonstrated that there was no significant difference in rates of acquisition between matched tasks ($M$ = 32.77, $SD$ = 32.65) and mismatched tasks ($M$ = 52.73, $SD$ = 36.95); $t(8) = 1.267$, $p = 0.241$. However, based on the average rates of acquisition reported above, lower rates of acquisition were found for mismatched tasks, and overall, children in this sample were presented with a larger number of tasks that were mismatched to their ability levels.

## Discussion

The purpose of this study was to evaluate whether the difficulty of the training tasks taught to children enrolled in the St.Amant Autism Early Learning Program matched the learning abilities of the children. In order to do so, five questions were examined across three different studies.

In Study 1, I examined whether observers who are knowledgeable about the ABLA-R could reliably categorize each of the 544 tasks of the ABLLS-R into an ABLA-R level per task. My hypothesis stated that observers familiar with the ABLA-R would be able to reliably categorize a number of ABLLS-R tasks into an ABLA-R level per task. In fact, observers agreed that a total of 99 tasks were categorizable, 439 tasks were non-categorizable in that those particular tasks did not fit any of the predetermined ABLA-R guidelines, and none of the tasks remained as disagreements between a non-categorizable and a categorizable task. Of the 99 ABLLS-R tasks that observers agreed were categorizable, all were reliably categorized according to the highest ABLA-R level

needed to perform them, with 35 tasks categorized as Level 1, five tasks categorized as Level 2, three tasks categorized as Level 3, 24 tasks categorized as Level 4, 13 tasks categorized as Level 5, and 19 tasks categorized as Level 6. These categorizations were confirmed reliable with substantial agreement based on Kappa.

My hypothesis for Study 1 was based on a study conducted by DeWiele and Martin (1996), who evaluated whether individuals familiar with the ABLA could reliably categorize 194 training tasks according to the ABLA levels needed to perform them. Training tasks consisted of tasks typically presented to individuals with ID at a residential training facility. Results demonstrated that overall categorization agreement across tasks was 68.6% (which corresponded to a total of 133 tasks), with the number of tasks on which the individuals agreed ranging from 2 to 61 per ABLA level. Study 1 presents a few notable differences when compared to DeWiele and Martin's study. For instance, the sample of training tasks in Study 1 consisted of 544 specific ABLLS-R tasks typically taught to children with ASD who were enrolled in a large government-funded program. In addition, observers used the ABLA-R and not the original ABLA to categorize the ABLLS-R training tasks. Note that the only difference between the ABLA and the ABLA-R is that the latter has the new Level 5 task (VVNM). Finally, rather than determining interobserver reliability based on percent agreement, I used Cohen's weighted Kappa as the measure of agreement between the observers, given that they were required to categorize each task into one of six hierarchically-ordered ABLA-R levels.

Study 1 represents the first attempt to reliably categorize the 544 ABLLS-R training tasks according to an ABLA-R level per task. The finding that 99 tasks were agreed upon by observers as categorizable represents a contribution of practical

significance, given that EIBI programs often use the ABLLS-R (or other similar resources) as a skills tracking system and curriculum guide, and some use the ABLA-R as a learning assessment tool. When comparing the list of 99 categorized ABLLS-R tasks to the ABLA-R level descriptions, staff working with children with ASD may find it easier to develop programs that are appropriate for the children's learning ability levels, and thereby reduce potential frustration or number of trials required to master particular tasks or skills. Future research might examine a replication and extension of this study, recruiting observers who have demonstrated experience with the ABLA-R and the ABLLS-R to conduct the categorizations. Future research might also examine the use of additional ABLA-R levels (e.g., auditory-auditory identity matching (AAIM)) to categorize the ABLLS-R tasks, such that more tasks may be reliably categorized into ABLA-R levels.

In Study 2, I examined whether Autism Consultants from the St.Amant Autism Programs typically taught a random sample of those 99 ABLLS-R tasks at the ABLA-R levels into which they were categorized. My hypothesis stated that Autism Consultants did in fact teach those tasks as described in the ABLLS-R, and therefore at the ABLA-R levels into which they were categorized, considering that staff from the Autism Programs typically refer to the ABLLS-R as a skills tracking system and curriculum guide. The results indicated that training programs were typically developed based on the guidelines in the ABLLS-R manual. More specifically, there was an average agreement of 90.5% that the categorized ABLLS-R tasks were taught at their respective ABLA-R levels. This study contributes important information with regards to EIBI and evaluations of

procedural fidelity, as it is often difficult in large EIBI meta-analysis studies to confirm that treatment and other procedures are delivered as described.

Study 2 proceeded with a survey that was completed by Autism Consultants from the St.Amant Autism Programs. To ensure that most Autism Consultants would complete and return their survey, I randomly selected two tasks from those categorized at each of the six ABLA-R levels, and created a simple 12-question survey that could be answered by checking either “Yes” or “No”. The survey was administered during a meeting where all Autism Consultants who were present, and who wished to complete the survey, would be given a period of time to do so. With this method, 14 of the 16 Autism Consultants present at the meeting completed and returned their surveys. Although the Autism Consultants showed very high agreement that the 12 ABLLS-R tasks on the survey were taught at their respective ABLA-R levels, future research might examine whether Autism Consultants follow the ABLLS-R guidelines for all 99 categorized tasks. Future research might also conduct a more thorough procedural fidelity evaluation, by obtaining and comparing specific training programs to their corresponding ABLLS-R task descriptions. Furthermore, future research might examine the extent to which Autism Tutors follow program procedures as written by the Autism Consultants.

The results of Study 2 were critical in order to proceed to Study 3. If results had demonstrated weak agreement between staff, that would have implied that Autism Consultants did not generally refer to the ABLLS-R manual in the same manner to develop training programs, and it would not have been feasible to compare the children’s ABLLS-R training tasks to their respective highest-passed ABLA-R levels. In Study 3, I examined three different questions using archived data for a sample of 14 children from

the St.Amant Autism Early Learning Program. First, Part A of this study examined to what extent the ABLA-R difficulty level of ABLLS-R training tasks matched the learning abilities of the children as assessed by the ABLA-R. As DeWiele and Martin (1996) found with their participants, my hypothesis stated that the ABLA-R difficulty of ABLLS-R tasks taught to the children would not closely match the highest-passed ABLA-R level of each child; meaning that a large proportion of training tasks would be mismatched to each child's ABLA-R level. As predicted, results demonstrated that an average of 19% of tasks were matched and 81% of tasks were mismatched to the children's ABLA-R levels. Of those mismatched tasks, 64% were mismatched below the children's ABLA-R levels, and 36% were mismatched above the children's ABLA-R levels. To obtain these results, a total of 31 assessment periods and 1006 tasks (counting each task's respective exemplars) were included in the analysis to determine the percentage of tasks at, below, and above each child's ABLA-R level.

It should be noted that 29 of the 1006 tasks included in the analysis were labeled as "MOD," "M," or "PRO" (e.g., C17MOD). These labels indicated that the task (e.g., C17) had been modified or extended to some degree. Due to the fact that Study 2 demonstrated that Autism Consultants generally follow the ABLLS-R guidelines to develop their training programs, and that it was not possible to review these specific training programs to examine the extent to which they had been modified from the ABLLS-R descriptions, the 29 tasks were also included and analyzed as original ABLLS-R tasks with their assigned ABLA-R level categorizations.

Part A of Study 3 reveals potentially important information related to the difficulty level of training tasks presented to children with ASD enrolled in a large

government-funded program. Children in these types of programs generally obtain services for a limited amount of years, therefore it is crucial that staff be able to quickly identify and develop training programs that are appropriate for a child's ability level, at any given point in time. Doing so may increase the effectiveness of instruction sessions, and in turn, children may learn at a faster rate a larger number of skills that are required for daily functioning and integration into school and other occupational areas. In addition, given that previous research (e.g., Vause et al., 1999) has demonstrated that the presentation of a larger number of tasks mismatched to a child's highest-passed ABLA level results in increased levels of aberrant behaviours, this may also negatively impact the children's success in these programs. Future research might examine whether additional staff training on the use of particular assessments (e.g., the ABLLS-R and the ABLA-R) may improve the development of training programs to ensure a better match to the children's ability levels.

Part B of Study 3 examined whether there was a relation between the proportions of training tasks mismatched to the children's highest-passed ABLA-R levels and their maladaptive behavior scores. My hypothesis stated that more aberrant behaviours would be found for children with ASD who were presented with a larger number of mismatched tasks as compared to children with tasks matched to their ABLA-R levels. A total of 31 SIB-R assessments were included in the analysis, and compared to each child's respective percentages of matched and mismatched tasks. Overall results demonstrated that 12 assessments (39%) demonstrated normal behaviour scores and 19 assessments (61%) demonstrated maladaptive behaviour scores. More specifically, nine assessments (6 children) demonstrated marginally serious scores, three assessments (3 children)

demonstrated moderately serious scores, five assessments (4 children) demonstrated serious scores, and two assessments (2 children) demonstrated very serious scores. Despite the lack of a statistically significant relationship between the children's proportions of mismatched training tasks and their maladaptive behaviour scores, 61% of scores were indicative of maladaptive behaviours for children in this sample, and overall, these children were presented with larger numbers of mismatched tasks versus tasks matched to their learning ability, as demonstrated in Part A of Study 3.

My hypothesis was based on a study by Vause et al. (2000), who assessed the frequency of aberrant behaviour of 13 individuals who were enrolled in three classrooms in a large training facility for individuals with ID. During the first stage of the study, Vause et al. recorded the ABLA level of training tasks presented to the participants by their instructors, and recorded instances of aberrant behaviour occurring while the participants were engaged in those tasks. Results indicated that 9 of the 13 participants engaged in more aberrant behaviours when presented with training tasks that did not match their ABLA levels, and fewer aberrant behaviors when presented with tasks that were a match. After teaching staff about matching the ABLA difficulty of training tasks to the highest-passed ABLA level of participants within each classroom, mean aberrant behaviour in the three classrooms decreased from 47% to 30%, 45% to 40%, and 57% to 45%. The effect found by Vause et al. was not significant, although they report that this may have been due to a limited sample size. Overall these results were consistent with previous research indicating that presenting training tasks matched to a client's ability level will result in less aberrant behaviours than the presentation of mismatched tasks.

Part B of Study 3 differs from Vause et al. (2000) in that my data analysis was based on archived data for a sample of children with ASD from the St.Amant Autism Early Learning Program, rather than direct observations and training with staff who taught individuals with ID. Also, Vause et al. directly observed the participants' aberrant behaviours, rather than obtaining preexisting behaviour scores. These differences concur with a limitation of this study; that is, it was not possible to determine with certainty that the SIB-R scores indicative of maladaptive behaviours were directly related to the presentation of mismatched tasks. First, there was no statistically significant relationship found between the maladaptive behaviour scores and the proportions of tasks mismatched to the children's ABLA-R levels. This may also have been due to the small sample size. Second, given that the maladaptive behaviour scores were based on archived data, there were no direct observations to demonstrate that the presentation of mismatched tasks caused the increase in maladaptive behaviours for these children. In other words, the maladaptive behaviours present during those assessment periods may have been caused by a number of other potential variables (e.g., reinforcement of the maladaptive behaviours by other people in the child's environment, potential health issues, lack of motivation during one-on-one instruction sessions, etc.). It is also important to note that because parents or legal guardians complete the SIB-R based on their own perceptions of their child's behaviour, the results may lack accuracy and they may be influenced by other events or situations that may have occurred near the assessment date. Consequently, future studies may look at conducting direct observations if feasible, or identifying a more objective measure of maladaptive behaviour (e.g., a functional analysis or functional assessment conducted within a particular time period).

Nevertheless, the results obtained in Part B provide important information pertaining to children in this sample. Overall, 11 of the 14 children demonstrated a varying degree of maladaptive behaviour, and these behaviours may very well interfere with the children's ability to learn and acquire valuable skills.

Part C of Study 3 examined whether there was correspondence between the children's proportions of mismatched tasks and their rates of acquisition of new training tasks. My hypothesis stated that lower rates of acquisition would be found for children with ASD who were presented with a larger number of tasks mismatched to their highest-passed ABLA-R levels. Results demonstrated that children required an average of 25 trials to master tasks matched to their ABLA-R level, and 42 trials to master tasks mismatched to their ABLA-R level. More specifically, children required an average of 47 trials to master tasks mismatched below their ABLA-R level, and 34 trials to master tasks mismatched above their ABLA-R level; perhaps this particular difference was due to boredom or lack of motivation with the larger proportions of tasks that are considered too easy (note that an average of 64% of mismatched tasks were mismatched below the children's ABLA-R levels). There was no statistically significant difference in rates of acquisition between matched and mismatched tasks, although this may have been due to the limited sample of data corresponding to tasks matched, mismatched below, and mismatched above for any given assessment period. Despite this limitation, lower rates of acquisition were found for mismatched tasks. These results contribute beneficial information to the EIBI literature, and more specifically, to staff working with children enrolled in large EIBI programs, such that it is important to ensure that children are

presented with appropriate training tasks that may be learned in fewer trials while decreasing levels of maladaptive behaviours.

As a whole, this research offers a number of contributions to the current EIBI literature. In addition to the aforementioned contributions, it is important to note that this research is the first to examine these questions with children with ASD enrolled in a large EIBI government-funded program. As mentioned in the literature review, many EIBI programs are becoming larger community-based programs, and in Canada, many of these programs are government-funded. Thus, this study has practical significance. With the tools and resources that have been developed through this study (e.g., a list of ABLLS-R tasks categorized into ABLA-R levels, a procedural fidelity survey, and various data sheets), other government-funded and privately funded EIBI programs may be able to further examine the effectiveness of their services, and if needed, improve the development and delivery of the training programs presented to children with ASD. Although the sample size of Study 3 was limited, the majority of the children in this study demonstrated the predicted results, suggesting that it may be possible to generalize these results to other children enrolled in EIBI programs. Given the increased prevalence of ASD, the limited amount of funding, and the increasing waitlists for families with children with ASD, these studies contribute valuable information that may benefit both service providers and families receiving services. Other strengths of these studies include that all of the data was obtained from one EIBI program (rather than multiple EIBI programs often used in meta-analysis studies), and that for a number of children, data was obtained at six-month or yearly intervals across two or three assessment periods.

In addition to these strengths and contributions, a number of limitations were also identified and described within the three studies. Another limitation to Study 3 involved restrictions pertaining to data access and available resources. Originally, it was discussed that children who were currently enrolled in the St.Amant Autism Programs would be recruited for this study. However, as recruitment with families of children with ASD enrolled in the Autism Programs can prove to be limiting due to the small pool of families and the large number of research studies occurring simultaneously, it was agreed that the Autism Program's archival database would provide a larger sample of available data. Subsequently, it was noted that children's data entered in the archival database also included personal information (e.g., names, phone numbers, addresses) that understandably implied requiring ethical approval to retrieve and analyze the necessary data. Given the concerns brought forth by the Research Ethics Board (REB) regarding access to this information without specific client consent from each child for whom data was saved in the archival database, it was suggested that an alternative procedure be employed to access the necessary information. Consequently, staff from the Autism Programs volunteered their time and efforts to filter through the archival database, identify children for whom all necessary data was present, scan all of the data for potential personal information, eliminate this information as it was not required for the purposes of Study 3, and finally, save the data into a separate database for which I had access. However, this procedure was time-consuming and required the assistance of many individuals. Further research might examine other potential alternatives that would allow direct access to such data, and therefore increase accessibility to larger samples of potential available data.

Another limitation is that, due to the inclusion criteria related to the children's assessment and ABLLS-R data (as described earlier), potential participants or potential assessment periods that could have been included in Study 3 were eliminated. Also, due to missing data for tasks matched and mismatched to the children's ABLA-R levels, some analyses were conducted with a smaller sample of data. Consequently, it may be that statistical significance was not achieved due to the smaller size of available assessment data. In addition, although data was obtained across two or three assessment periods for a number of children, the sample size precluded any comparisons across these periods. If future studies are able to obtain and analyze a larger sample of data across consecutive assessment periods (e.g., Year 1, Year 1.5, and Year 2), this may provide interesting comparison results that could be used in addressing these particular questions.

Despite these limitations, these three studies and their results offer significant contributions to the current EIBI literature. Overall results demonstrated that: a) observers who were knowledgeable about the ABLA-R reliably categorized 99 of the ABLLS-R tasks into individual ABLA-R levels; b) Autism Consultants from the St.Amant Autism Programs averaged 90.5% agreement that a random sample of the 99 tasks were taught at the categorized ABLA-R levels; c) for a sample of 14 children from the St.Amant Autism Early Learning Programs, an average of 81% of their training tasks were mismatched to each child's highest-passed ABLA-R level; d) for the SIB-R assessments of those children, 61% of the assessments had scores representative of maladaptive behaviours; and e) for that same sample of children, they averaged 25 trials to master tasks matched to their individual highest-passed ABLA-R levels, and 42 trials to master tasks mismatched to their ABLA-R levels. The results of these studies expand

the current research on children with ASD, the ABLA-R and the ABLLS-R, with the objective of improving individualized training procedures and curricula currently used in EIBI programs for children with ASD.

References

Aman, M. G., & Singh, N. N. (1986). *Aberrant behavioral checklist manual.* East Aurora, NY: Slosson Educational Publications.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Publishing.

Bruininks, R. H., Woodcock, R. W., Weatherman, R. F., & Hill, B. K. (1996). *Scales of independent behavior-revised: Comprehensive manual*. Chicago, IL: Riverside Publishing Company.

Casey, L., & Kerr, N. (1977). Auditory-visual discrimination and language prediction. [Monograph]. *Rehabilitation Psychology, 24*(3), 137-155. doi:http://dx.doi.org.proxy2.lib.umanitoba.ca/10.1037/h0090912

Centers for Disease Control and Prevention, Autism and Developmental Disabilities Monitoring Network. (2014). *Prevalence of autism spectrum disorder among children aged 8 years* (MMWR No. 63-SS02). Retrieved from http://www.cdc.gov/mmwr/preview/mmwrhtml/ss6302a1.htm?s_cid=ss6302a1_w

Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early intensive behavioral treatment: Replication of the UCLA model in a community setting. *Developmental and Behavioral Pediatrics, 27*, 145-155. doi:http://dx.doi.org/10.1097/00004703-200604002-00013

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*(4), 213-220. doi:http://dx.doi.org/10.1037/h0026256

Condillac, R. A. (2003). *Assessment of discrimination skills in individuals with autism: Validity of the assessment of basic learning abilities test* (Doctoral dissertation). Available from PsycINFO. (620266928; 2003-95012-149). Retrieved from http://search.proquest.com/docview/620266928?accountid=14569

DeVries, C., Yu, C. T., Sakko, G., Wirth, K. M., Walters, K. L., Marion, C., . . . MacLean Jr., W. E. (2005). Predicting the relative efficacy of verbal, pictorial, and tangible stimuli for assessing preferences of leisure activities. *American Journal on Mental Retardation, 110*(2), 145-154. Retrieved from http://search.proquest.com/doc view/620718433?accountid=14569

DeWiele, L., & Martin, G. L. (1996). Can the ABLA test help staff match training tasks to the abilities of developmentally disabled trainees? *International Journal of Practical Approaches to Disability, 20*, 7-11.

DeWiele, L., Martin, G. L., Martin, T., Yu, C. T., & Thomson, K. (2011). *The kerr meyerson assessment of basic learning abilities revised: A self-instructional manual* (2nd ed.). Winnipeg, MB: St.Amant Research Centre. Retrieved from http://stamant.ca/wp-content/uploads/2013/02/ABLA-R-self-instructional-manual-20140630.pdf

Eikeseth, S. (2009). Outcome of comprehensive psycho-educational interventions for young children with autism. *Research in Developmental Disabilities, 30*(1), 158-178. doi:http://dx.doi.org/10.1016/j.ridd.2008.02.003

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2002). Intensive behavioral treatment at school for 4- to 7-year-old children with autism: A 1-year comparison controlled study. *Behavior Modification*, *26*(1), 49-68. doi:http://dx.doi.org/10.1177/0145445502026001004

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child & Adolescent Psychology*, *38*(3), 439-450. doi:http://dx.doi.org/10.1080/15374410902851739

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2010). Using participant data to extend the evidence base for intensive behavioral intervention for children with autism. *American Journal on Intellectual and Developmental Disabilities*, *115*(5), 381-405. doi:http://dx.doi.org/10.1352/1944-7558-115.5.381

Eldevik, S., Hastings, R. P., Jahr, E., & Hughes, J. C. (2012). Outcomes of behavioral intervention for children with autism in mainstream pre-school settings. *Journal of Autism and Developmental Disorders*, *42*(2), 210-220. doi:http://dx.doi.org/10.1007/s10803-011-1234-9

Eldevik, S., Jahr, E., Eikeseth, S., Hastings, R. P., & Hughes, C. J. (2010). Cognitive and adaptive behavior outcomes of behavioral intervention for young children with intellectual disability. *Behavior Modification*, *34*(1), 16-34. doi:http://dx.doi.org/10.1177/0145445509351961

Eldevik, S., Ondire, I., Hughes, J. C., Grindle, C. F., Randell, T., & Remington, B. (2013). Effects of computer simulation training on in vivo discrete trial teaching.

*Journal of Autism and Developmental Disorders, 43*(3), 569-578. doi:http://dx.doi.org/10.1007/s10803-012-1593-x

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

Feldman, M. A. (2000). Special issue on the assessment of basic learning abilities test: Introductory comments. *Journal on Developmental Disabilities, 7*(2).

Feldman, M. A. (Ed.). (2004). *Early intervention: The essential readings*. Malden, MA: Blackwell Publishing. Retrieved from http://search.proquest.com/docview/620276608?accountid=14569

Feldman, M. A., Condillac, R. A., Tough, S., Hunt, S., & Griffiths, D. (2002). Effectiveness of community positive behavioral intervention for persons with developmental disabilities and severe behavior disorders. *Behavior Therapy, 33*(3), 377-398. doi:http://dx.doi.org/10.1016/S0005-7894(02)80034-X

Feldman, M. A., & Griffiths, D. (1997). Comprehensive assessment of severe behavior problems. In N. N. Singh (Ed.), *Treatment of severe behavior problems: Models and methods in developmental disabilities* (pp. 23-48). Belmont, CA: Thomson Brooks/Cole Publishing Co. Retrieved from http://search.proquest.com/docview/619142979?accountid=14569

Feldman, M. A., Ward, R. A., Savona, D., Regehr, K., Parker, K., Hudson, M., . . . Holden, J. J. A. (2012). Development and initial validation of a parent report measure of the behavioral development of infants at risk for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(1), 13-22.

doi:http://dx.doi.org/10.1007/s10803-011-1208-y

Howlin, P., Magiati, I., & Charman, T. (2009). Systematic review of early intensive behavioral interventions for children with autism. *American Journal on Intellectual and Developmental Disabilities, 114*(1), 23–41. Retrieved from http://search.proquest.com/docview/621978710?accountid=14569

Kerr, N., Meyerson, L., & Flora, J. A. (1977). The measurement of motor, visual, and auditory discrimination skills [Monograph]. *Rehabilitation Psychology, 24*(3), 95-112. doi:http://dx.doi.org.proxy2.lib.umanitoba.ca/10.1037/h0090912

Kodak, T., & Grow, L. L. (2011). Behavioral treatment of autism. In Fisher, W. W., Piazza, C. C., & Roane, H. S. (Eds.), *Handbook of applied behavior analysis.* (pp. 402-416). New York, NY: Guilford Press. Retrieved from http://search.proquest.com/docview/964198923?accountid=14569

LaForce, J. C., & Feldman, M. A. (2000). Role of discrimination ability in the cooperative behavior of persons with developmental disabilities. *Journal on Developmental Disabilities, 7*(2), 156-170. Retrieved from http://search.proquest.com/docview/619832035?accountid=14569

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. Retrieved from http://www.jstor.org/stable/2529310?seq=1#page_scan_tab_contents

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*(1), 3-9. doi:http://dx.doi.org/10.1037/0022-006X.55.1.3

Martin, G. L., & Pear, J. J. (2015). *Behavior modification: What it is and how to do it* (10th ed.). Upper Saddle River, NJ: 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 developmental disabilities: A review. *Behavior Modification, 32*(2), 228-247. doi:http://dx.doi.org/10.1177/0145445507309022

Martin, G. L., & Yu, D. C. T. (2000). Overview of research on the assessment of basic learning abilities test. *Journal on Developmental Disabilities, 7*(2), 14-15. Retrieved from http://search.proquest.com/docview/619827702?accountid=14569

Martin, G. L., Yu, D., Quinn, G., & Patterson, S. (1983). Measurement and training of AVC discrimination skills: Independent confirmation and extension. *Rehabilitation Psychology, 28*(4), 231-237. doi:http://dx.doi.org/10.1037/h0090974

Matson, J. L. (2007). Current status of differential diagnosis for children with autism spectrum disorders. *Research in Developmental Disabilities, 28*(2), 109-118. doi:http://dx.doi.org/10.1016/j.ridd.2005.07.005

Matson, J. L., & Konst, M. J. (2013). What is the evidence for long term effects of early autism interventions? *Research in Autism Spectrum Disorders, 7*(3), 475-479. doi:http://dx.doi.org/10.1016/j.rasd.2012.11.005

Matson, J. L., & Smith, K. R. M. (2008). Current status of intensive behavioral interventions for young children with autism and PDD-NOS. *Research in Autism Spectrum Disorder, 2*(1), 60-74. doi:http://dx.doi.org/10.1016/j.rasd.2007.03.003

Matson, J. L., & Sturmey, P. (Eds.). (2011). *International handbook of autism and*

*pervasive developmental disorders*. New York, NY: Springer Science + Business Media. doi:http://dx.doi.org/10.1007/978-1-4419-8065-6

McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation, 97*(4), 359-372. Retrieved from http://search.proquest.com/docview/618295967?accountid=14569

Meyerson, L. (1977). AVC behavior and attempts to modify it [Monograph]. *Rehabilitation Psychology, 24*(3), 119-122. doi:http://dx.doi.org.proxy2.lib.umanitoba.ca/10.1037/h0090912

Murphy, C., Martin, G. L., & Yu, C. T. (2014). The predictive validity of the assessment of basic learning abilities versus parents' predictions with children with autism. *Education and Training in Autism and Developmental Disabilities, 49*(4), 601-611. Retrieved from http://search.proquest.com/docview/1642634282?accountid=14569

Partington, J. W. (2006). *The assessment of basic language and learning skills-revised.* Pleasant Hill, CA: Behavior analysts, Inc. Retrieved from https://www.partingtonbehavioranalysts.com/page/ablls-r-25.html

Passey, J., & Feldman, M. (2004). Descriptive analysis of parent-child interactions in young children with or at risk for developmental delay. *Behavioral Interventions, 19*(4), 233-246. doi:http://dx.doi.org/10.1002/bin.164

Perry, A., & Condillac, R. (2003). *Evidence-based practices for children and adolescents with autism spectrum disorders: Review of the literature and practice guide.*

Children's Mental Health Ontario. Retrieved from http://www.kidsmental health.ca/documents/EBP_autism.pdf

Perry, A., Cummings, A., Geier, J. D., Freeman, N. L., Hughes, S., LaRose, L., . . . Williams, J. (2008). Effectiveness of intensive behavioral intervention in a large, community-based program. *Research in Autism Spectrum Disorders, 2*(4), 621-642. doi:http://dx.doi.org/10.1016/j.rasd.2008.01.002

Perry, A., Cummings, A., Geier, J. D., Freeman, N. L., Hughes, S., Managhan, T., . . . Williams, J. (2011). Predictors of outcome for children receiving intensive behavioral intervention in a large, community-based program. *Research in Autism Spectrum Disorders, 5*(1), 592-603. doi:http://dx.doi.org/10.1016/j.rasd.2010.07.003

Peters-Scheffer, N., Didden, R., Korzilius, H., & Sturmey, P. (2011). A meta-analytic study on the effectiveness of comprehensive ABA-based early intervention programs for children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 5*(1), 60-69. doi:http://dx.doi.org/10.1016/j.rasd.2010.03.011

Reichow, B., & Wolery, M. (2009). Comprehensive synthesis of early intensive behavioral interventions for young children with autism based on the UCLA young autism project model. *Journal of Autism and Developmental Disorders, 39*(1), 23-41. doi:http://dx.doi.org/10.1007/s10803-008-0596-0

Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatment for early autism. *Journal of Clinical Child & Adolescent Psychology, 37*(1), 8-38. doi:http://dx.doi.org/10.1080/15374410701817808

Ruble, L. A., & McGrew, J. H. (2007). Community service outcomes for families and children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 1*(4), 360-372. doi:http://dx.doi.org/10.1016/j.rasd.2007.01.002

Sakko, G., Martin, T. L., Vause, T., Martin, G. L., & Yu, C. T. (2004). Visual-visual nonidentity matching assessment: A worthwhile addition to the assessment of basic learning abilities test. *American Journal on Mental Retardation, 109*(1), 44-52. doi:2.0.CO;2" TARGET="_blank">http://dx.doi.org/10.1352/0895-8017(2004)109<44:VNMAAW>2.0.CO;2

Schwartzman, L. J. V., Vause, T., Martin, G. L., Yu, C. T., Campbell, L., Danbrook, M., & Feldman, M. (2009). Predicting the learning ability of children with autism: The assessment of basic learning abilities test versus parents' predictions. *Education and Training on Developmental Disabilities, 44*(2), 271-279.

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*(5), 473-484. Retrieved from http://search .proquest.com/docview/619305949?accountid=14569

Thorsteinsson, J. R., Martin, G. L., Yu, C. T., 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 on Mental Retardation, 112*(2), 130-139. doi:http://dx.doi.org /10.1352/0895-8017(2007)112[130:PLAOPW]2.0.CO;2

Vause, T., Martin, G. L., Cornick, A., Harapiak, S., Chong, I., Yu, D. C. T., & Garinger, J. (2000). Training task assignments and aberrant behavior of persons with developmental disabilities. *Journal of Developmental Disabilities, 7*(2), 37-53. Retrieved from http://search.proquest.com/docview/619832164?accountid=14569

Vause, T., Martin, G. L., & Yu, D. (1999). Aberrant behaviour of persons with developmental disabilities as a function of the characteristics of training tasks. *International Journal of Rehabilitation Research, 22*, 321-325. Retrieved from http://journals.lww.com/intjrehabilres/Citation/1999/12000/Aberrant_behaviour_of_persons_with_developmental.10.aspx

Verbeke, A. K., Martin, G. L., Thorsteinsson, J. R., Murphy, C., & Yu, C. T. (2009). Does mastery of ABLA level 6 make it easier for individuals with developmental disabilities to learn to name objects? *Journal of Behavioral Education, 18*(3), 229-244. doi:http://dx.doi.org/10.1007/s10864-009-9087-4

Verbeke, A. K., Martin, G. L., Yu, C. T., & Martin, T. L. (2007). Does ABLA test performance predict picture name recognition with persons with severe developmental disabilities? *The Analysis of Verbal Behavior, 23*, 35-39. Retrieved from http://search.proquest.com/docview/621745760?accountid=14569

Viel, J., Wightman, J., Marion, C., Jeanson, B., Martin, G., 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*(4), 1370-1377. doi:http://dx.doi.org/10.1016/j.rasd.2011.01.019

Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose-response meta-analysis of multiple outcomes. *Clinical Psychology Review, 30*(4), 387-399. doi:http://dx.doi.org/10.1016/j.cpr.2010.01.008

Wacker, D. P., Berg, W. K., Harding, J. W., & Cooper-Brown, L. (2011). Functional and structural approaches to behavioral assessment of problem behavior. In Fisher, W. W., Piazza, C. C., & Roane, H. S. (Eds.), *Handbook of applied behavior analysis.* (pp. 165-181) New York, NY: Guilford Press. Retrieved from http://search.proquest.com/docview/964200943?accountid=14569

Wacker, D. P., Kerr, N. J., & Carroll, J. L. (1983). Discrimination skill as a predictor of prevocational performance of institutionalized mentally retarded clients. *Rehabilitation Psychology, 28*(1), 45-59. doi:http://dx.doi.org/10.1037/h0091005

Wacker, D. P., Steil, D. A., & Greenebaum, F. T. (1983). Assessment of discrimination skills of multiply-handicapped preschoolers and prediction of classroom task performance. *Journal of the Association for the Severely Handicapped, 8*, 65-78. Retrieved from http://eric.ed.gov/?id=EJ289959

Ward, R., & Yu, D. (2000). Bridging the gap between visual and auditory discrimination learning in children with autism and severe developmental disabilities. *Journal on Developmental Disabilities, 7*(2), 142-155. Retrieved from http://search.proquest.com/docview/619831998?accountid=14569

Witt, J. C., & Wacker, D. P. (1981). Teaching children to respond to auditory directives: An evolution of two procedures. *Behavior Research of Severe Developmental Disabilities, 2*, 175-189.

Wright, L. M. (2014). *Intensive behavioral intervention in a community-based program for children with autism: A retrospective effectiveness study* (Doctoral dissertation). Available from PsycINFO. (1642631668; 2014-99240-048). Retrieved from http://search.proquest.com/docview/1642631668?accountid =14569

Yu, D., & Martin, G. L. (1986). Comparison of two procedures to teach visual discriminations to severely mentally handicapped persons. *Journal of Practical Approaches to Developmental Handicap, 10,* 7-12.

Appendix A

## Fidelity of Training Programs Survey

**BACKGROUND: In an earlier phase of my study, reliability assessments were conducted to categorize ABLLS-R tasks into ABLA-R levels. We were successful in categorizing 99 ABLLS-R tasks.**

**PURPOSE: The purpose of this survey is to assess whether Autism Consultants develop training programs based on the guidelines in the ABLLS-R manual.**

**INSTRUCTIONS: Please read each question and ABLLS-R task description carefully. For each question, check either YES or NO.**

1. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 1 task? Meaning that the desired behaviour is modeled before the response occurs, and involves an imitation.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **B20** | Delayed replication of a sequence | After having been shown a model of items in a specific sequence and then having the model removed, the student will be able to replicate the sequence. | Can the student replicate a sequence of items after having been shown a model of items in a specific sequence and then having the model removed? | After showing a pattern of three items and then remove the display, the student replicates the pattern. |

2. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 1 task? Meaning that the desired behaviour is modeled before the response occurs, and involves an imitation.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **D4** | Imitation of leg and foot movements | Upon request, student will imitate a gross motor activity involving foot and leg movements. | Will the student imitate a gross motor action involving foot and leg movements when asked to “Do this”? | Stomp foot |

3. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 2 task? Meaning that the task involves responding to objects that remain in the same place or position.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **B12** | Block designs from picture | When given a block design card, student will place blocks to make the design shown on the card. | Can the student do block designs from looking at a picture of a design. | |

4. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 2 task? Meaning that the task involves responding to objects that remain in the same place or position.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **Z2** | Places objects in a form box | The student will be able to place objects of various shapes in a form box. | Can the student place objects of various shapes in a form box? | |

5. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 3 task? Meaning that the task involves responding to objects that change places or positions.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **A3** | Look at a non-reinforcing item | Student will look at and track changes in location of a non-reinforcing item presented by an instructor. | If you hold up a non-reinforcing item, will the student look at it and watch it as you move it to different locations in front of the student? | When you hold up a shoe and ask the student to look at the shoe, student will look at it and watch it as you move it to a variety of positions in front of him (e.g.,up/down/left/right) |

6. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 3 task? Meaning that the task involves responding to objects that change places or positions.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **B21** | Delayed finding a sample | When shown a specific item and it is removed, and five items are then presented (one being the same as the original item), the student will be able to find the initial item. | Can the student find an item which had previously been shown to the student after it has been removed, and then presented in an array of three items (one being the same as the original item)? | After showing the student a picture of a cat, hide the picture, wait a few seconds, present pictures of a dog, cat, and bird and say "Find the same" or "Find it" |

7. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 4 task? Meaning that the task involves visually matching something to something else in the environment on the basis that the two matching items are the same on at least one physical dimension.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **B5** | Match identical pictures to sample | When given a picture, the student will match to an identical picture in an array of three pictures. | Can the student match pictures to an identical picture presented in an array of three pictures? | |

8. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 4 task? Meaning that the task involves visually matching something to something else in the environment on the basis that the two matching items are the same on at least one physical dimension.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **Q8** | Match individual letters to letters on word card | The student will be able to match individual letters to the letters on cards with single 5 letters words. | Can the student match individual letters to the letters on cards with single 5 letter words? | Given a word card with the word "train", the student will match individual letter cards to the letters on the word card |

9. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 5 task? Meaning that the task involves visually matching something to something else in the environment even though the two items are not similar on any physical dimension.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **B17** | Sort by function | When provided with a sample of two items that have different functions and shown a demonstration of matching another item that has the same function, the student will be able to sort additional pictures that have the same functions with the correct sample items. | When provided with a model of items that are used for a certain function, can the student sort pictures of items into groups of items that have the same function (e.g., scissors with ones you cut with, pen with things you write with)? | When shown a pen and a bottle of bubbles and the placement of a crayon on the pen and a candle on the bubbles, student continues to place other pictures on things you write with and things you blow on the correct piles. |

10. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 5 task? Meaning that the task involves visually matching something to something else in the environment even though the two items are not similar on any physical dimension.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **C36** | Select associated pictures | Given an array of items or pictures and asked, "What goes with this?" the student will be able to select the item which is used with the item being shown (e.g., select a bat when shown a picture of a ball). | When given a display of objects or pictures and shown a different item and asked, "what goes with this?" can the student select an item from the display which is used with the item being presented (e.g., select a bat when shown a picture of a ball). | Bat and a ball; cup and juice; socks and shoes; knife, fork and a spoon; shirt, hat and pants. |

11. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 6 task? Meaning that the task involves an auditory discrimination as well as a visual discrimination.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **R9** | "More" | The student will be able to receptively identify and label examples of the word "more". | Can the student receptively identify and label examples of the word "more"? | |

12. For the ABLLS-R task below, do you typically develop a program to teach the task as outlined in the description below, and therefore as an ABLA-R Level 6 task? Meaning that the task involves an auditory discrimination as well as a visual discrimination.

YES _____ NO _____

| Task | Task Name | Task Objective | Question | Example |
|---|---|---|---|---|
| **N8** | Gets & returns own materials | The student will follow instructions to get and return his own educational materials. | Can the student get and put away his own educational activities? | Student puts his workbook into his storage area when told to do so at the end of a task |

Appendix B

## Calculation of the Percentage of Tasks Matched, Mismatched Below, and Mismatched Above the ABLA-R Level

**Participant #: ____**

**Year 1** ABLA-R Level:____

1) Percentage of tasks matched to ABLA-R level*= _____

2) Percentage of tasks mismatched to ABLA-R level*= _____

a) Percentage of mismatched tasks below ABLA-R level*=_____

b) Percentage of mismatched tasks above ABLA-R level*=_____

**Year 1.5** ABLA-R Level:____

1) Percentage of tasks matched to ABLA-R level= _____

2) Percentage of tasks mismatched to ABLA-R level= _____

a) Percentage of mismatched tasks below ABLA-R level=_____

b) Percentage of mismatched tasks above ABLA-R level=_____

**Year 2** ABLA-R Level:____

1) Percentage of tasks matched to ABLA-R level= _____

2) Percentage of tasks mismatched to ABLA-R level= _____

a) Percentage of mismatched tasks below ABLA-R level=_____

b) Percentage of mismatched tasks above ABLA-R level=_____

*Percentages are calculated based on the following:
# matched tasks/ total # of tasks
# mismatched tasks/ total # of tasks
# mismatched below/ total # of mismatch
# mismatched above/ total # of mismatch

** To calculate the total number of tasks, include and review data entered up to 1 month before the assessment date

Appendix C

## Calculation of Rates of Training Task Acquisition for Tasks Matched, Mismatched Below, and Mismatched Above the ABLA-R Level

Participant #: _____ ABLA-R level: _____ Assessment period: _____

We will identify reliably categorized tasks that were mastered during the 1-month assessment period (1 month up to the date of the assessments). We will categorize each task as either a task matched, mismatched below or mismatched above the ABLA-R level. Rate of acquisition will then be calculated by counting the number of trials required to master each task and dividing by the total number of tasks.

| Mastered tasks **matched** to ABLA-R level | Rate of acquisition (# of trials/# of tasks) | Mastered tasks **below** ABLA-R level | Rate of acquisition (# of trials/# of tasks) | Mastered tasks **above** ABLA-R level | Rate of acquisition (# of trials/# of tasks) |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |