AN EVALUATION OF STIMULUS PRESENTATION ARRANGEMENTS ON CHILDREN’S ACQUISITION OF LISTENER BEHAVIOUR

by

Maria Turner, BCaBA
B.A., Simon Fraser University, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

The Faculty of Graduate and Postdoctoral Studies
(Special Education)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October, 2016

© Maria Turner, 2016
Abstract

Clinicians teach listener behaviour within the context early intervention for children with Autism Spectrum Disorder (ASD). Effective, evidence-based procedures are necessary to establish skills that children with ASD do not readily acquire through interactions with their caregivers and peers. In the present study, the effectiveness and efficiency of three different antecedent stimuli presentations (i.e., sample-first, sample-first-with-repetition, and comparison-first conditions) were compared using an adapted alternating treatments design. Participants were three children with ASD, aged 4- and 6-years old. The most efficient presentation varied across participants, and the results obtained with one efficiency measure did not always yield similar results to that obtained with the other efficiency measures. Implications for teaching listener behaviour in early intervention programming are addressed.
Preface

The Behavioural Research Ethics Board (BREB) from the University of British Columbia approved this project and associated method on May 13th, 2016. The certificate number is H16-00776.

I was the lead investigator for the experiment that this thesis contains. I was responsible for the data collection, analysis of data, and the majority of manuscript composition.
# Table of Contents

Abstract.................................................................................................................................................. ii  
Preface.................................................................................................................................................... iii  
Table of Contents ................................................................................................................................... iv  
List of Tables .......................................................................................................................................... iv  
List of Figures ......................................................................................................................................... vii  
Acknowledgements ............................................................................................................................... viii  

## 1 Introduction....................................................................................................................................... 1  
1.1 Conceptualization of Listener Behaviour ...................................................................................... 2  
1.2 Stimulus Control and Listener Behaviour ...................................................................................... 3  
1.3 Comparisons of Instructional Strategies for Teaching Listener Behaviour ................................. 4  
1.4 Consequence Comparisons ............................................................................................................ 5  
1.5 Antecedent Comparisons ................................................................................................................ 6  
1.6 Rationale for the Current Study ....................................................................................................... 9  

## 2 Method ............................................................................................................................................... 10  
2.1 Participants and Setting .................................................................................................................... 10  
2.2 Materials and Training Sets ............................................................................................................ 11  
2.3 Pre-experimental Assessments ........................................................................................................ 13  
2.4 Dependent Variables and Data Collection ...................................................................................... 15  
2.5 Interobserver Agreement ................................................................................................................ 18  
2.6 Procedural Fidelity ........................................................................................................................... 21  
2.7 Research Design .............................................................................................................................. 22  
2.8 Experimental Procedures ............................................................................................................... 22  
2.8.1 General Procedures .................................................................................................................... 24  
2.8.2 Baseline....................................................................................................................................... 24  
2.8.3 Intervention Comparison ............................................................................................................ 25  
2.8.4 Follow-up ................................................................................................................................... 26  
2.9 Participant-Specific Modifications .................................................................................................. 26  
2.9.1 Extinction for Prompted Responses ............................................................................................. 26  
2.9.4 Differential Observing Response ............................................................................................... 27  
2.9.5 Instructive Feedback .................................................................................................................. 28  

## 3 Results ................................................................................................................................................. 29  
3.1 Ron - Evaluation 1 ............................................................................................................................ 29  
3.2 Ron - Evaluation 2 ............................................................................................................................ 31  
3.4 Tom .................................................................................................................................................. 33  
3.5 Gary .................................................................................................................................................. 35  
3.6 Echolc Responses ............................................................................................................................ 36  
3.7 Early Responses ............................................................................................................................... 36  

## 4 Discussion .......................................................................................................................................... 37  
4.1 Relation Between Efficiency Measures ........................................................................................... 38  
4.2 Stimulus Control .............................................................................................................................. 39
4.3 Early Responses ................................................................................................................. 41
4.4 Comparison to Existing Research ...................................................................................... 42
4.5 Limitations .......................................................................................................................... 43
4.6 Clinical Implications .......................................................................................................... 44
4.7 Future Research .................................................................................................................. 46

5 Conclusion ............................................................................................................................ 47

References ............................................................................................................................... 49
List of Tables

Table 1. The stimuli taught using the sample-first, sample-first-with-repetition and comparison-first conditions during each evaluation across participants. .................................................. 12

Table 2. The percentage of sessions in which interobserver agreement (IOA) was assessed, the mean IOA scores, and the range of IOA scores across participants and dependent variables. .......................................................................................................................... 20

Table 3. The percentage of sessions in which procedural fidelity was assessed, the mean procedural fidelity scores, and the range of procedural fidelity scores across participants and measures.................................................................................................................. 23

Table 4. The percentage of trials with echoic responses across participants.............................................. 36

Table 5. The percentage of trials with early responses and touching the barrier (Gary only) across participants................................................................................................................................. 37
List of Figures

Figure 1. Results of the preference assessments across participants. .................................................. 16
Figure 2. Percentage of independent correct responding across baseline, intervention and follow
up conditions for Ron’s first and second evaluations. ................................................................. 30
Figure 3. Number of trials to mastery, cumulative duration of instruction to mastery, and average
duration per session across evaluations in the sample-first, sample-first-with-repetition and
comparison-first conditions. .............................................................................................. Error! Bookmark not defined.
Figure 4. The percentage of independent correct responses across baseline, intervention and
follow-up for Tom and Gary's evaluations. .................................................................................... 34
Acknowledgements

I wish to acknowledge the University of British Columbia, my professors, and my peers, for providing a supportive and stimulating learning environment throughout my Masters’ program. I will be forever changed from this experience. I would like to extend particular thanks to Dr. Pat Mirenda for the caring manner in which you have assisted me throughout this program; your support and guidance has been so appreciated.

I owe my deepest gratitude to Dr. Laura Grow. The experiences, skills, and knowledge that you have passed on to me have been invaluable. Your care in creating learning opportunities at every step of this program and your skillful support throughout each undertaking has allowed me to be successful in endeavours that I never would have known I could accomplish. I am ever grateful for your expert direction and guidance throughout this journey.

I also wish to acknowledge the support of Laura and my fellow lab mates in cultivating a climate of openness, cooperation, and mutual support within our lab that has so enriched my education. I am grateful for every project that I have had the pleasure of being involved in; all have been stimulating and joyful experiences.

I also acknowledge the support of Dr. Glen Davies, ABLE Developmental Clinic, and my supervisors Hilary McClinton and Kathryn Theroux. They have been supportive of this research project and instrumental in my professional development. Thank you for always having kind words of encouragement to share with me.

To my family, thank you for all you have done to help me achieve this goal; without your support I would not be here today.

To my mother, who never failed to stand by me, it is with your love and guidance that I have been able to achieve this accomplishment. You planted the seed of scientific inquiry with
me, starting with the cartoon that you kept in your office: a little girl looking through a magnifying glass, with the caption “When I grow up, I want to be a scientist just like Mommy.” You have always been a strong, powerful, and inspirational role model and I love you beyond words.

And finally, to Max, my wonderful, creative and inspiring partner-in-crime. I am ever grateful for your love and support that has been so instrumental to my success academically, musically, and beyond; I owe so much to you. You have only ever wished for me to realize my potential, and for that I am eternally grateful. For your comedic relief, bear hugs, listening ear, shoulder to cry on, and for always ensuring I have a good meal to eat at the end of a long day, thank you, times infinity, plus one.
1 Introduction

Children diagnosed with Autism Spectrum Disorder (ASD) have deficits in language and communication skills that typically require intervention (American Psychiatric Association, 2013). Effective and efficient instructional strategies are necessary for teaching skills that children with ASD do not readily acquire as a result of interactions with their caregivers, teachers, and peers. Large-scale studies of early intensive behavioural intervention (EIBI) have demonstrated significant improvements in social functioning, communication skills, and problem behaviour for children diagnosed with ASD (Lovaas, 1987; Howard, Sparkman, Cohen, Green, & Stanislaw, 2005; Sallows & Graupner, 2005). In the first large-scale study of EIBI, approximately half of the participants scored within a normal range of intellectual functioning and were included in general education classrooms with little to no additional support (Lovaas, 1987). Following the success of EIBI research, intervention protocols were manualized and published in EIBI curriculum books (e.g., Leaf & McEachin, 1999; Lovaas, 2003; Maurice, Green & Luce, 1996; Sundberg & Partington, 1998).

Instructors frequently use commercially available EIBI curricular manuals to guide the selection of instructional strategies for teaching new skills. However, interventions are recommended in EIBI manuals regardless of the empirical basis for their use. These strategies were shown to be effective as part of a treatment package, but have not benefited from systematic evaluation or have not been sufficiently examined in the literature. Additional research is needed to evaluate and compare specific behavioural interventions used to teach skills to optimize the instructional procedures and to maximize therapeutic gains from EIBI.
1.1 Conceptualization of Listener Behaviour

Instructors frequently teach listener behaviour in EIBI, particularly for learners without a fluent vocal verbal repertoire (Smith, 2001). Listener behaviour involves responding to another person’s verbal behaviour. For example, a caregiver says, “Give me the ball!” and the child reaches for a ball.

During listener training, instructors teach either simple discriminations or conditional discriminations (Green, 2001). A simple discrimination involves a three-term contingency: a discriminative stimulus (and S-delta), a response, and consequences for correct and incorrect responses. For example, an instructor says, “Come here” (i.e., the discriminative stimulus), the child walks to the instructor (response), and the instructor praises the child (consequence), but the child does not come when the instructor says, “Wave” (S-delta).

A conditional discrimination involves a four-term contingency: a sample stimulus, comparison stimuli, a response, and consequences for correct and incorrect responses (Green, 2001). A common type of conditional discrimination taught in EIBI is an auditory-visual conditional discrimination in which the sample stimuli are auditory and the comparison stimuli are visual. During an auditory-visual conditional discrimination, the sample stimulus alters the function of the comparison stimuli. The visual comparison stimuli include one discriminative stimulus and one or more S-deltas on any given trial. For example, an instructor delivers the auditory sample stimulus, “Boat,” and presents pictures of a chair, a boat, and a pencil as the visual comparison stimuli. The picture of the boat functions as the discriminative stimulus. That is, instructors would reinforce selection responses to the picture of the boat only if the instructor presents the auditory sample stimulus “Boat” immediately before the selection response. Instructors would not reinforce selection responses to the picture of the pencil and chair if the
instructor presents “Boat” immediately before the selection responses. Several terms are used to refer to auditory-visual conditional discriminations in clinical practice (e.g., listener behaviour, receptive labeling, receptive identification; see Grow & LeBlanc, 2013, for a more thorough list of the terms). For the purposes of this paper, we will refer to auditory-visual conditional discriminations as listener behaviour herein.

1.2 Stimulus Control and Listener Behaviour

One goal of listener behaviour programs is to gain stimulus control over selection responses. That is, the aim is for programmed or naturally occurring antecedent stimuli to exert control over the learner’s selection responses due to a history of reinforcement in the presence of the antecedent stimuli. Faulty stimulus control occurs when responses come under the control of irrelevant, or a restricted range of, antecedent stimuli rather than to the relevant, programmed stimuli (Grow & LeBlanc, 2013). Faulty stimulus control is common among learners with ASD, particularly when multiple antecedent stimuli are presented (Lovaas & Schreibman, 1971).

Faulty stimulus control may take the form of consistent patterns of responding to specific irrelevant characteristics of antecedent stimuli, known as error patterns. A side bias (i.e., position bias) occurs when a learner consistently selects the stimulus that is in a particular position in the array, regardless of the placement of the discriminative stimulus (Green, 2001). Win-stay responses, another form of error pattern, occur when a learner’s selection is influenced by the discriminative stimulus in the preceding trial (Lovaas, 2003). For example, a learner selects the horse on a given trial because the horse was the correct response in the immediately preceding trial. Early responses occur when the learner selects a stimulus before the delivery of the sample stimulus (Petursdottir & Aguilar, 2016), when none of stimuli in the comparison array function as a discriminative stimulus.
Entrenched faulty stimulus control is challenging to remediate, can hinder learning, and requires extensive resources to address (Green, 2001). Error patterns can generalize to other programs that include similar teaching preparations (Schilmoeller, Schilmoeller, Etzel, & LeBlanc, 1979). The careful arrangement of multiple antecedent stimuli, with attention to the desired stimulus control for listener behaviour relations, can help prevent faulty stimulus control from developing. It is important for practitioners to implement effective and efficient instructional strategies from the outset of listener behaviour instruction to avoid the development of faulty stimulus control.

1.3 **Comparisons of Instructional Strategies for Teaching Listener Behaviour**

Comparison research has increasingly focused on procedures used to establish skills, rather than eliminating problem behaviour (Shabani & Lam, 2013). Research comparing two or more instructional strategies for teaching the same skill may lead to the identification of more effective and/or efficient procedures (Shabani & Lam, 2013). Identifying which teaching procedures are most efficient has implications for the recipients of EIBI services through improved outcomes and reduced costs. Recent studies comparing two or more interventions for listener training have examined consequence as well as antecedent stimulus arrangements during listener training. Consequence arrangements that have been examined include error correction procedures for incorrect responses (e.g., Rodgers & Iwata, 1991) and reinforcement procedures for independent and prompted responses (e.g., Hausman, Ingvarsson, & Kahng, 2014). Antecedent stimulus arrangements that have been examined include within-stimulus prompting (Schreibman, 1975), instructional procedures (e.g., Grow, Carr, Kodak, Jostad, & Kisamore, 2011; Grow, Kodak, & Carr, 2014), and the order of presentation of antecedent stimuli in instruction (Petursdottir & Aguilar, 2016).
1.4 Consequence Comparisons

Researchers have compared different consequences for correct and incorrect responses during listener instruction to determine the relative effectiveness and efficiency of the consequence arrangements (e.g., Hausman et al., 2014; Rodgers & Iwata, 1991). Rodgers and Iwata compared a differential reinforcement only condition to an error correction procedure, where the trial was repeated until the participant emitted a correct response, and an avoidance condition, where non-relevant tasks were presented following incorrect responses. While all participants’ correct responses increased with the use of differential reinforcement with no scripted error correction procedure, three participants had the highest percentage of correct responses in the avoidance condition, and two of the participants performed best in the practice condition. While differential reinforcement alone was found to be sufficient to produce learning, this study demonstrated enhanced effects when specific error correction procedures were used, and the procedure that was most efficient varied by participant.

Hausman, Ingvarsson and Kahng (2014) compared the effects of reinforcement schedules on performance during the instruction of matching (i.e., visual-visual conditional discriminations) and spelling tasks. Acquisition was most rapid when reinforcers were delivered exclusively for accurate independent responding and prompted responses were placed on extinction. If participants learn more rapidly when only independent responses are reinforced, a greater number of behaviours can be added to an individual’s repertoire over time, thus adding to the social importance of EIBI (Hausman et al., 2014; Shabani & Lam, 2013). However, the authors note that it may be clinically useful to provide continuous reinforcement for both prompted and independent responses when a new skill is first introduced (Hausman et al., 2014). During instruction, stimulus control needs to be transferred from the prompt, which reliably
evokes the correct response, to the relevant antecedents, which do not yet evoke the correct response. Placing prompted responses on extinction could interfere with the transfer of stimulus control if prompted responses do not persist in the absence of reinforcement.

1.5 Antecedent Comparisons

When planning and selecting interventions to establish listener behaviour, instructors should carefully select the arrangement of antecedent conditions that best support the acquisition of appropriate stimulus control. This is most critical in arrangements that include multiple antecedents, such as listener behaviour skills, as individuals with ASD are likely to attend to only one of the two antecedent events (Grow et al., 2011; Schreibman, 1975). When methodologies that are most effective and efficient for establishing these skills are identified, they can then be recommended for best practice.

Schreibman (1975) examined the efficacy of within-stimulus and extra-stimulus prompting procedures for teaching visual and auditory discriminations to children with ASD. While extra-stimulus prompting did not result in acquisition for any of the relations, within-stimulus prompting was effective for the majority of sets. Within-stimulus prompting is likely more effective because it highlights the critical features of the discriminative stimulus to which the learner should attend. In contrast, extra-stimulus prompting involves the addition of a superfluous stimulus that is not a critical feature of the discriminative stimulus (Schreibman, 1975). When the irrelevant stimulus is removed, responding may return to baseline levels because the learner has not been taught to attend to the critical features of the discriminative stimulus.

Two instructional arrangements that have been examined include comparisons of the simple-conditional and conditional-only method of teaching listener behaviour (Grow et al.,
Grow and colleagues (2011) compared two commonly used methods for teaching listener behaviour: the conditional-only and simple-conditional methods. In the simple-conditional method, instructional targets were presented in a series of steps that gradually increased in complexity by targeting a particular stimulus in isolation, in a field of two, and in a field of three stimuli. Within each step, the target was massed trialed, where only one item functioned as the discriminative stimulus (i.e., simple discrimination), or randomly rotated, where the discriminative stimulus switched between the comparison stimuli depending on which sample stimulus was given (i.e., conditional discrimination). In the conditional-only method, targets were presented in a field of three and randomly rotated from the beginning of instruction. The conditional-only method resulted in fewer sessions to meet the mastery criterion, and the simple-conditional method was associated with patterns of faulty stimulus control (e.g., side bias). While the simple-conditional method is the arrangement presented in many early intervention manuals, it resulted in slower acquisition than the conditional-only method and was associated with the development of faulty stimulus control (Grow et al., 2011; Grow et al., 2014; Vedora & Grandelski, 2015).

Certain aspects of the arrangement of the simple-conditional method of instruction might have contributed to the development of faulty stimulus control. In the massed trial step of the simple-conditional method, only one stimulus from the comparison array functions as the discriminative stimulus. As the participant is not required to attend to the sample stimulus, this may have resulted in reinforcers being delivered for repeatedly selecting the same visual stimulus (Grow et al., 2011). In addition, one of the steps of the simple-conditional method requires the participant to respond in a field of two, where engaging in a side bias error pattern would be reinforced approximately 50% of the time (Grow & LeBlanc, 2013).
Petursdottir and Aguilar (2016) examined the efficiency of two antecedent stimulus presentations in teaching listener behaviour in an adapted, alternating treatments design combined with a multiple baseline design across stimulus sets. The conditions compared were a sample-first condition, where the sample stimulus is given before the presentation of the comparison-stimuli, or a comparison-first condition, where the comparison stimuli in the array are presented, then the sample stimulus is given. The materials were delivered using a computer-based presentation format, and the participants were three typically developing kindergarten students. The authors found that the sample-first presentation was as or more efficient than the comparison-first presentation across stimulus sets (Petursdottir & Aguilar, 2016). In addition to correct and incorrect responses, the authors measured the occurrence of early responding, or the selection of a stimulus from the comparison array before the delivery of the sample stimulus.

The comparison-first condition is uniquely vulnerable to participants engaging in early responses, as the materials are present in the field prior to the delivery of the auditory sample stimulus. When first presented, the stimuli in the array function as neutral stimuli. It is not until the sample stimulus is delivered that the stimuli in the array begin to function as either a discriminative stimulus or S-delta. Because the sample stimulus is critical in determining which stimulus in the comparison array functions as the discriminative stimulus, presenting the sample stimulus before the comparison array may increase its saliency. Then, when the discriminative stimulus and the S-deltas in the comparison array are presented, they will have a designated function rather than being presented as neutral stimuli.

The results of the Petursdottir and Aguilar (2016) study have important implications for the arrangement of antecedents in listener behaviour instruction in early intervention programming. The participants in this evaluation were typically developing children, and the
generality of the results to children diagnosed with ASD is unknown. In addition, teaching was conducted without the use of transfer-of-stimulus-control procedures and consisted of trial and error learning alone. Further research is necessary to assess the applicability of these findings to other populations and with instructional strategies (e.g., prompt fading procedures) that are typically used with individuals with ASD and other developmental disabilities.

1.6 Rationale for the Current Study

The temporal ordering of the sample and comparison stimuli can affect the acquisition of skills and the development of appropriate stimulus control during interventions for teaching listener behaviour (Petursdottir & Aguilar, 2016). As the sample stimulus alters the function of the visual stimuli in the comparison array, the stimuli are neutral when presented before the delivery of the sample stimulus. Enhancing the saliency of the sample stimulus by presenting it before the comparison stimuli may increase the likelihood that an individual will attend to the stimuli, and therefore may increase the efficiency of listener behaviour instruction.

Many EIBI curricular manuals recommend that instructions issued in discrete trial training should not be repeated to maintain instructional control and to teach “cleanly” (Maurice et al., 1996). However, due to the transient nature of the auditory sample stimulus in a listener behaviour trial, responding to a prompt in the absence of the sample stimulus changes the function of the participant’s response. If the sample stimulus is not present, the participant is only imitating the instructor’s prompt response, rather than engaging in a true conditional discrimination. Repeating the auditory sample stimulus every two seconds to maintain its saliency during early listener behaviour training may have a facilitative effect on the efficiency of instruction. Recommendations to repeat the auditory stimulus during listener behaviour
instruction have been made (Grow & Leblanc, 2013; Tarbox & Tarbox, 2009), but have not been examined empirically to date.

The purpose of the current study is to replicate and extend the study by Petursdottir and Aguilar (2016) to address the following research question.

1. Is there a difference in the relative effectiveness and efficiency of the sample-first, sample-first-with-repeated-sample, and comparison-first stimulus presentations in the instruction of listener behaviour for children diagnosed with ASD?

2 Method

2.1 Participants and Setting

Three participants were recruited for the study through local service agencies in the Vancouver, British Columbia area. Inclusion in the study required that each participant have: (a) a history of compliance with adult-delivered instructions, (b) a generalized imitation repertoire, (c) a visual-visual matching repertoire, and (d) a diagnosis of ASD by a qualified professional. Ethics approval from the UBC Behavioural Research Ethics Board was obtained prior to recruiting participants for the study.

Ron was a 4-year-old boy at the time of the study. He had been diagnosed with ASD by a licensed psychologist. He received early intensive behavioural intervention outside of this study, and had acquired a variety of listener behaviour skills in the past. He was able to identify stimuli by feature, function, and class, and was beginning to read. Ron had experience using a token economy system in his intervention program.

Tom was a 6-year-old boy at the time of the study. He had been diagnosed with ASD by a licensed psychologist. He received early intensive behavioural intervention outside of this study, and had acquired a variety of listener behaviour skills in the past, including identifying
stimuli by feature, function, and class. Tom had experience using a token economy system in his intervention program.

Gary was a 6-year old boy at the time of the study. He was diagnosed with ASD and episodic ataxia (type 2) by a licensed psychologist. Gary received early intensive behavioural intervention outside of the study. The clinical team reported difficulty with teaching listener behaviour in the past. Despite extensive training and modifications to his program on the part of the clinical team, Gary did not meet the mastery criterion for any trained sets for this program type. Gary’s clinical team reported a previous history of prompt dependency and error patterns during listener training.

All sessions were conducted in a quiet room with minimal distractions in the families’ homes. Ron’s sessions were conducted in the basement, and Tom and Gary’s sessions were conducted in the family’s living room. All rooms contained a child-sized table, two chairs, preferred toys and activities, and the materials required to conduct the sessions (e.g., data sheets, writing instrument, visual stimuli, timer).

2.2 Materials and Training Sets

The visual comparison stimuli for Ron and Tom consisted of picture cards 2.75 x 5” in size that were printed and laminated. For Gary, picture cards 4.75 x 3.25” in size were included. Country flags were printed on the entire card, and the remaining stimuli were clear, centred pictures printed on a plain white background.

A training set consisted of three auditory-visual stimulus pairs. Each training set was randomly assigned to one of three experimental conditions (i.e., nine stimulus pairs per comparison) except for the Greek letters stimuli, where the same stimuli taught in Ron’s second evaluation were assigned to counterbalanced conditions in Tom’s evaluation. The experimenter
equated the training sets to control for difficulty as a confounding variable. The training sets were equated to ensure equal complexity by the number of syllables, types of phonemes or consonant blends, initial consonants, redundant phonemes across words, and visual similarity of the stimuli through a logical analysis (Gast & Ledford, 2014).

The experimenter collaborated with at least one caregiver to identify possible listener behaviours to include in the training sets. As both Ron and Tom had acquired many listener behaviour skills before the start of the study, stimuli sets consisting of country flags and Greek letters were used for these evaluations to control for prior history. Gary had not yet acquired listener behaviours at the time of the study. Therefore, we selected common objects for his evaluation. See Table 1 for a summary of the stimuli taught across evaluations and conditions.

Table 1. The stimuli taught using the sample-first, sample-first-with-repetition and comparison-first conditions during each evaluation across participants.

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Sample-first</th>
<th>Sample-first-with-repetition</th>
<th>Comparison-first</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ron Eval 1</td>
<td>Senegal, Ghana,</td>
<td>Togo, China, Suriname</td>
<td>Samoa, Taiwan,</td>
</tr>
<tr>
<td></td>
<td>Cameroon</td>
<td></td>
<td>Liechtenstein</td>
</tr>
<tr>
<td>Ron Eval 2</td>
<td>sigma, pi, lambda</td>
<td>delta, phi, gamma</td>
<td>theta, psi, upsilon</td>
</tr>
<tr>
<td>Tom</td>
<td>gamma, delta, phi</td>
<td>psi, theta, upsilon</td>
<td>sigma, pi, lambda</td>
</tr>
<tr>
<td>Gary</td>
<td>scissors, hat, bike</td>
<td>boot, ice cream, keys</td>
<td>hammer, clock, truck</td>
</tr>
</tbody>
</table>

A physical barrier 9 ½ x 12” in size was used to prepare the visual stimuli out of sight of the participant. The use of the barrier allowed the experimenter to present the field of comparison stimuli at the proper time during the instructional trial (e.g., after the delivery of the sample stimulus in the sample-first condition, or before the delivery of the sample stimulus in the comparison-first condition).
2.3 Pre-experimental Assessments

The experimenter administered three pre-experimental assessments: (a) the Behavioral Language Assessment Form (BLAF; Sundberg & Partington, 1998), (b) the Reinforcer Assessment for Individuals with Severe Disabilities (RAISD; Fisher, Piazza, Bowman & Amari, 1996), and (c) a paired-choice preference assessment (Fisher et al., 1992). For Ron, the RAISD only identified five highly preferred stimuli that would function for the purpose of a preference assessment. Therefore, a 6-stimulus brief multiple-stimulus-without-replacement (MSWO; Carr, Nicolson & Higbee, 2000) preference assessment format was used for Ron instead of the paired-choice assessment.

Behavioral Language Assessment Form. The BLAF (Sundberg & Partington, 1998) is a questionnaire assessment tool that uses a rating scale to assess the following repertoires: a) cooperation, b) requesting, c) motor imitation, d) vocal play, e) vocal imitation, f) match-to-sample, g) receptive, h) labeling, i) receptive identification by feature, function and class, j) conversation, k) letters and numbers, and l) social interaction. Each repertoire has five possible scoring options. A score of one is considered a low score and is indicative of the absence of the target skill. A score of five is considered a high score suggesting that the individual has the skill in his/her repertoire. In the area of vocal play, a score of five would be obtained by someone who vocalizes frequently and says many understandable words and phrases, a score of one would be obtained by someone who is mute, and a score of three would be obtained by someone who vocalizes many speech sounds with varied notations (Sundberg & Partington, 1998).

The BLAF was used to assess each participant’s language repertoire and to determine if each participant met the inclusion criteria (i.e., compliance with adult-delivered instructions, a
generalized imitation repertoire, and a visual-visual matching repertoire). The assessment information was gathered through direct observation and an interview with caregivers.

Results from Ron and Tom’s BLAF indicated that they met the inclusion criteria for the study across all of required areas. Results from Gary’s BLAF indicated that he had met the inclusion criteria of the study in the areas of complying with adult-delivered instructions and visual-visual matching. While he had some imitation skills, he did not have a generalized imitation repertoire. The purpose of requiring that participants have a generalized imitation repertoire is to ensure that model prompting would be a successful prompt type. A prompting assessment was conducted with Gary to determine if a model prompt would evoke a selection response. Model prompting was effective for occasioning Gary’s selection responses; therefore, he was included in the study.

**Reinforcer assessment for individuals with severe disabilities.** The experimenter conducted the RAISD with at least one caregiver (Fisher et al., 1996). The purpose of the RAISD was to gather information about different classes of stimuli that may function as reinforcers (e.g., toys, attention). The stimuli endorsed by the participants’ caregivers were included in the paired-choice or MSWO preference assessments.

**Preference assessment.** The paired-choice preference assessment was used to identify a hierarchy of preferred stimuli that would be used as consequences for correct responses during intervention sessions (Fisher et al., 1992). The experimenter placed two stimuli in front of the participant and provided up to 5 s to make a selection response. The experimenter then provided 30 s of access to the selected stimulus. Each stimulus was paired with every other stimulus during the assessment. Data were collected on the participant’s selections to generate a hierarchy of preferred stimuli.
The brief MSWO preference assessment was used for the same purpose as the paired-choice preference assessment (Carr, Nicolson & Higbee, 2000; Fisher et al., 1992). The experimenter placed all stimuli in the field and asked the participant to, “Pick one.” The participant had 30 s of access to the stimulus selected, while the remaining stimuli in the array were rotated. The selected stimulus was not replaced in the array, and was no longer available during the assessment after being chosen. The array was presented three times. Data were collected on the participant’s selections to generate a hierarchy of preferred stimuli.

Figure 1 displays results of the preference assessment across participants. The top panel shows the results of Ron’s assessment; the second panel shows the results of Tom’s assessment, and the third and last panels show the results of Gary’s assessments across tangibles and across edibles.

2.4 Dependent Variables and Data Collection

A paper-and-pencil method was used to record participant and experimenter responses in real time. Sessions were videotaped for data collection purposes. The primary dependent variable was the percentage of independent correct responses in each session and the number of trials to reach the mastery criterion in each condition. The secondary dependent variables were the percentage of prompted correct responses, the cumulative duration of instructional time to meet mastery criterion, and the percentage of early responses.

Independent correct response. An independent correct response was defined as the participant selecting the correct picture within 5 s of an experimenter-delivered sample and comparison stimuli and without prompts. The percentage of independent correct responses for each training set was determined by dividing the number of independent correct responses in each session by the total number of trials (i.e., nine) and multiplying the proportion by 100.
Figure 1. Results of the brief MSWO and paired-stimulus preference assessments across participants.
**Prompted correct response.** A prompted correct response was defined as the participant touching the correct picture within 5 s of an experimenter-delivered model prompt. The percentage of prompted correct responses for each training set was determined by dividing the total number of independent correct responses by the total number of trials presented and multiplying the proportion by 100.

**Cumulative duration of instruction to mastery.** The observer recorded the duration of each session on a trial-by-trial basis using a stopwatch. The experimenter started the stopwatch immediately before delivering the antecedent of the first trial in the session. The experimenter stopped the stopwatch immediately after delivering the reinforcer in the last trial of the session. The cumulative duration of instruction to mastery was calculated by summing the duration times for each session in an experimental condition in seconds, then converting the final number to minutes. The duration of sessions conducted in baseline and follow-up conditions were not included in the calculation.

**Number of trials to mastery.** The total number of trials to mastery within each experimental condition was calculated by counting the number of teaching sessions that the experimenter conducted and multiplying the final count by 9. Sessions conducted in baseline and follow up were not included.

**Early response.** An early response was defined as the participant touching a picture in the comparison set before the presentation of a sample stimulus by the experimenter. Early responses were not scored as correct or incorrect, and were only possible during the comparison-first condition. The percentage of early responses in each session was calculated by dividing the number of early responses by the total number of trials and multiplying the proportion by 100. Early responses were measured only in the comparison-first condition.
**Touching the barrier.** Gary engaged in early responses in the comparison-first condition; he also attempted to touch the barrier in sample-first and sample-first-with-repetition trials, prior to the auditory sample being presented. A dependent variable, touching the barrier, was added to Gary’s evaluation. It was defined as briefly tapping the barrier for 1 s or less before the auditory sample stimulus was presented in sample-first and sample-first-with-repetition trials. The percentage of touching the barrier in each session was calculated by dividing the number of instances of touching the barrier by the total number of trials and multiplying the proportion by 100.

**Repetition of the sample stimulus.** A repetition of the sample stimulus was defined as the participant repeating the same auditory sample that the experimenter had given at any point during the trial. If the participant said one of the target sample stimuli that differed from the current trial or said the sample stimulus prior to the experimenter initiating that trial, a repetition was not scored. The percentage of repetitions in each session was calculated by dividing the number of repetitions by the total number of sessions and multiplying the proportion by 100.

**Selection response.** The observer recorded the first selection response made by the participant during each trial by circling the corresponding visual stimulus on the data collection sheet. Data on selection responses were used to conduct error analyses in the event of stalled progress.

### 2.5 Interobserver Agreement

The experimenter used behavioural skills training (i.e., instruction, modeling, and rehearsal with feedback; Reid & Parsons, 1995) to train an undergraduate-level research assistant to collect data for the study. Before beginning data collection, the research assistant achieved
80% agreement with the experimenter in each condition type for the dependent variables and experimenter procedural fidelity using mock videos of experimental sessions.

The experimenter and a research assistant independently collected data for the purposes of calculating interobserver agreement (IOA) during experimental sessions or through videotape observation of the experimental session. Two data collectors scored independent correct responses and prompted correct responses on a trial-by-trial basis. An agreement was scored if both observers recorded the occurrence or non-occurrence of the same dependent variable. A disagreement was scored if one observer recorded data that differed from the response scored by the second observer. Interobserver agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying the proportion by 100.

Table 2 shows the percentage of sessions in which IOA was collected. The mean percentage of agreement and the range for the percentage of agreement in each evaluation are reported across each dependent variable. Across all evaluations, at least 33.33% of sessions were scored for IOA and the mean agreement exceeded 94.8% for each dependent variable across evaluations.

The data collectors independently calculated the number of trials to mastery. An agreement was scored if the two observers had the same number. A disagreement was scored if one observer had a score that differed from the second observer. Interobserver agreement was calculated by dividing the smaller number by the larger number and multiplying the proportion by 100. Interobserver agreement for the number of trials to mastery was 100% across all evaluations.
Table 2. The percentage of sessions in which interobserver agreement (IOA) was assessed, the mean IOA scores, and the range of IOA scores across participants and dependent variables.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Independent Correct Response</th>
<th>Prompted Response</th>
<th>Error</th>
<th>Selection Response</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percentage (%) of sessions assessed</td>
<td>Mean IOA % (range)</td>
<td>Percentage (%) of sessions assessed</td>
<td>Mean IOA % (range)</td>
</tr>
<tr>
<td>Ron</td>
<td>33.33%</td>
<td>100% (100%)</td>
<td>33.33%</td>
<td>98.9% (89%-100%)</td>
</tr>
<tr>
<td>Eval 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ron</td>
<td>33.33%</td>
<td>100% (100%)</td>
<td>33.33%</td>
<td>100% (100%)</td>
</tr>
<tr>
<td>Eval 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tom</td>
<td>34.10%</td>
<td>97% (89%-100%)</td>
<td>34.10%</td>
<td>96.33% (89%-100%)</td>
</tr>
<tr>
<td>Gary</td>
<td>33.33%</td>
<td>100% (100%)</td>
<td>33.33%</td>
<td>100% (100%)</td>
</tr>
</tbody>
</table>
2.6 Procedural Fidelity

A trained data collector scored the experimenter’s accuracy with implementing the procedures in each experimental condition. Research assistants scored data from videotaped sessions. A procedural fidelity score was calculated by dividing the sum of the percentages of correctly implemented steps in each trial by the total number of trials. Procedural fidelity data were collected for the following intervention components: antecedent order, stimulus presentation, prompting, and consequences.

**Antecedent order.** Correct implementation for the antecedent order was defined as the experimenter presenting the sample and comparison stimuli in the order indicated for the current experimental condition (i.e., sample-first, sample-first-with-repetition, comparison-first).

**Stimulus presentation.** Correct implementation for the stimulus presentation was defined as the experimenter presenting the antecedent stimuli in accordance with the procedures in the protocol (i.e., correct auditory stimulus, correct visual stimuli, presenting the visual stimuli in the array in the order indicated on the data sheet).

**Prompting.** Correct implementation of the prompting procedure was defined as the experimenter implementing the prompting procedure as specified by the protocol (e.g., 0-s prompt delay, 5-s prompt delay, model prompting, etc.)

**Consequences.** Correct implementation of the consequences for the instructional trial was defined as the experimenter delivering the specified reinforcer within 3 s of an independent or a prompted corrected response (with the exception of Tom after reinforcement for prompted responses was removed), or responding appropriately to early responses or errors as scripted in the protocol.
Table 3 shows the procedural fidelity scores across measures for each evaluation. Procedural fidelity was scored for at least 25% of sessions across experimental conditions. The mean procedural fidelity score across all sessions was 98.96% for Ron’s first evaluation (range, 66% to 100%), 100% for Ron’s second evaluation, 99.74% for Tom (range, 89% to 100%), 99.38% for Gary (range, 89% to 100%).

2.7 Research Design

An adapted alternating treatments design (AATD) design was used to compare the effectiveness and efficiency of sample-first, sample-first-with-repetition, and comparison-first conditions (Sindelar, Rosenberg, & Wilson, 1985). The design was similar to that used by Petursdottir and Aguilar (2016). The AATD was selected because the research question is related to increasing non-reversible behaviour. While a parallel treatment design is appropriate, the AATD was considered a preferable design, as it requires significantly less time and resources for participants and the experimenter (Gast & Ledford, 2014).

2.8 Experimental Procedures

Based on participant availability, the experimenter conducted sessions three to five times per week in 30- to 45-min session blocks daily. Training sessions included three trials of each auditory-visual stimulus pair (i.e., 9-trial session). The sample and comparison stimuli were counterbalanced based on the rules described by Green (2001). Each visual stimulus was presented equally across all positions in the array and was never the discriminative stimulus during more than two consecutive trials. Each stimulus served as the discriminative stimulus three times in a 9-trial session. Across all conditions, the experimenter arranged the comparison stimuli in a single row in front of the participant, with a barrier in front of the stimuli to block the participants’ view before the start of the trial.
Table 3. The percentage of sessions in which procedural fidelity was assessed, the mean procedural fidelity scores, and the range of procedural fidelity scores across participants and measures.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Antecedent Order</th>
<th>Stimulus Presentation</th>
<th>Prompting Procedures</th>
<th>Consequences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percentage (%) of sessions assessed</td>
<td>Mean fidelity % (range)</td>
<td>Percentage (%) of sessions assessed</td>
<td>Mean fidelity % (range)</td>
</tr>
<tr>
<td>Ron Eval 1</td>
<td>33.33%</td>
<td>100% (100%)</td>
<td>33.33%</td>
<td>95% (66%-100%)</td>
</tr>
<tr>
<td>Ron Eval 2</td>
<td>26.66%</td>
<td>100% (100%)</td>
<td>26.66%</td>
<td>100% (100%)</td>
</tr>
<tr>
<td>Tom</td>
<td>31.81%</td>
<td>100% (100%)</td>
<td>31.81%</td>
<td>100% (100%)</td>
</tr>
<tr>
<td>Gary</td>
<td>33.33%</td>
<td>98% (89%-100%)</td>
<td>33.33%</td>
<td>100% (100%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.8.1 General Procedures

The sample-first, sample-first-with-repetition and comparison-first sets were presented with the following general procedures:

Sample-first condition. The experimenter initiated the trial by presenting the auditory sample stimulus and immediately removing the barrier to provide visual access to the comparison stimuli.

Comparison-first condition. The experimenter initiated the trial by removing the barrier to provide visual access to the comparison stimuli first and presented the auditory sample stimulus.

Sample-first-with-repetition condition. The experimenter initiated the instructional trial by giving the auditory sample stimulus and immediately removing the barrier to provide visual access to the comparison stimuli. The experimenter repeated the sample stimulus every 2 s until the participant engaged in a selection response, or until the experimenter delivered a model prompt as indicated by the research protocol.

2.8.2 Baseline

A baseline probe was conducted with each stimulus before intervention. The purpose of the baseline probe was to ensure that the participant did not already have the target stimuli in their repertoire. As baseline probes were conducted before the creation of the equated sets, baseline was conducted using a comparison-first presentation. The experimenter placed the items in the field and presented the auditory sample stimulus. The experimenter did not provide prompts or reinforcers for participant responses. Reinforcers were delivered intermittently (approximately every other trial) for general attending and appropriate behaviour (e.g., going to the table to start a session, attending to instructions). Targets were included in a training set if the
participant engaged in less-than-chance responding during baseline (e.g., 33% or below) or demonstrated a stimulus bias by selecting the target stimulus regardless of whether it was the discriminative stimulus or an S-delta on more than 66% of trials.

2.8.3 Intervention Comparison

The intervention comparison involved setting up the instructional trial in accordance with the procedures for the current experimental condition (e.g., sample-first, comparison-first) as well as implementing a constant prompt delay procedure and delivering reinforcers for prompted or independent correct responses.

The experimenter used a constant prompt delay to transfer stimulus control from a model prompt to the auditory sample stimulus and discriminative stimulus. Each evaluation began with a 0-s prompt delay, during which the experimenter presented a model prompt immediately after presenting the antecedent stimuli. Following two sessions at a 0-s prompt delay, the experimenter delivered the model prompt after 5 s without a response or immediately after an error, whichever event occurred first. The experimenter re-presented the auditory sample stimulus with each presentation of the model prompt. Regardless of the experimental condition, the mastery criterion was 100% independent correct responses across two consecutive sessions.

During Ron’s first evaluation, the implementation of intervention resulted in a rapid increase in independent correct responses. Therefore, in Ron’s Evaluation 2 and in Tom’s evaluation, the experimenter implemented the 0-s prompt delay for one session instead of two to increase the likelihood of detecting a difference between the experimental conditions.

The experimenter delivered praise and access to a reinforcer contingent on prompted or independent correct responses. In Ron and Tom’s evaluations, token reinforcers were delivered with each prompted or independent correct response, and both participants were able to trade in
nine tokens for a back-up reinforcer. In Gary’s evaluation, the experimenter provided access to a preferred tangible reinforcer for 30 s following prompted and independent correct responses. If the participant engaged in an error, the experimenter provided a model prompt immediately and re-presented the auditory sample stimulus.

If participants engaged in early responses, they were prompted to place their hands in a neutral position and the trial was re-presented. Gary began touching the barrier before the auditory sample stimulus had been delivered in the sample-first condition. This behaviour was topographically similar to the early response. An operational definition was created for this behaviour and data were collected similarly to that described for the early responses. Gary was also prompted to place his hands in a neutral position and the trial was re-presented.

2.8.4 Follow-up.

The procedures during the follow-up probe were identical to those described in baseline, except reinforcement was provided for independent correct responses. No prompting was provided during follow up. The experimenter conducted a follow-up probe 3 weeks following the mastery of each training set.

2.9 Participant-Specific Modifications

2.9.1 Extinction for Prompted Responses

Across participants and evaluations, reinforcement was delivered for independent and prompted correct responses. In Tom’s evaluation, a low and variable trend was observed in the data, and several behavioural patterns were present that indicated that Tom was not attending to the auditory or visual stimuli. For example, he would often respond quickly, without scanning the visual array in the 5-s prompt delay procedure. A modification was made to his protocol to remove reinforcement for prompted responses. At this point, the token requirement was reduced
so that five tokens instead of nine tokens were required to cash in for a back-up reinforcer, as the frequency of reinforcement decreased. Following this change, the latency to target responses increased and Tom began scanning the visual array more readily.

2.9.4 Differential Observing Response

A differential observing response (DOR) is a response an individual engages in before or during an instructional trial that results in sensory contact with the discriminative stimulus, where the response emitted is unique to the discriminative stimulus (Grow & LeBlanc, 2013). A DOR was introduced during Tom and Gary’s evaluations to address errors patterns observed during teaching. In Tom’s evaluation, the DOR was only implemented in the sample-first condition, as this was the only condition where faulty stimulus control was observed with this participant. The DOR for Tom’s evaluation involved the experimenter asking him to repeat the auditory sample stimulus that corresponded to the upcoming trial immediately before the trial was presented. For example, the experimenter said, “Say ‘Gamma,’” Tom repeated, “Gamma,” and the trial was initiated by presenting the sample stimulus (“Gamma,”) and removing the barrier to reveal the comparison array in the field. The particular echoic response targeted matched the stimulus that would function as the discriminative stimulus in the upcoming trial.

As Gary did not make progress with the listener behaviour training alone and did not have the necessary skills to engage in a DOR without specific teaching (e.g., did not have an echoic or sign language repertoire), experimental sessions were paused and DOR training was implemented. The purpose of the DOR training was to teach Gary to complete a unique motor movement in response to the auditory sample stimulus, to greater increase the likelihood of attending to the auditory component of the antecedent.
Training was conducted to teach Gary to engage in a manual sign in response to the experimenter saying the corresponding spoken word, in order to use this response as a DOR in the listener instruction. The manual signs were the same as the stimuli used in Gary’s evaluation (e.g., hammer, clock, etc.). During teaching, the experimenter said the targeted auditory stimulus (e.g., “Hammer,”) and then immediately provided a model and physical prompt at a 0-s prompt delay. Once two sessions were conducted at a 0-s prompt delay, the experimenter used a least-to-most prompt fading procedure to transfer stimulus control to the auditory stimulus alone. The experimenter first gave the instruction to engage in the manual sign; if this did not result in a correct response within 5 s, then the experimenter presented a model prompt. If the model prompt did not result in a correct response within 5 s, then a physical prompt was presented. The mastery criterion for the DOR training was two sessions at 100% independent correct responding.

Once established, the DOR would have been used in listener behaviour trials by requiring Gary to engage in the manual sign in response to the experimenter saying the corresponding word prior to each listener behaviour trial. For example, in the comparison-first condition, the experimenter would say “Boot,” Gary would engage in the manual sign for boot, and the trial would be initiated by removing the barrier to reveal the comparison array in the field and then presenting the sample stimulus, “Boot.” The particular manual sign targeted would be the one that would function as the discriminative stimulus in the upcoming trial.

2.9.5 Instructive Feedback

Instructive feedback is a procedure where extra information is added in an instructional trial, without placing a requirement on the participant to engage in any additional responses (Delmolino, Hansford, Bamond, Fiske & LaRue, 2013). For example, if a participant was being
taught to select boat, car, and train in response to the corresponding auditory sample stimulus, once the participant made a prompted or independent selection response, the experimenter could say, “Yes, that’s the boat, it floats on water.” The participant is not required to respond to the instructive feedback in any way.

As Tom’s independent correct responses did not increase significantly following the addition of the DOR alone, instructive feedback was added to Tom’s evaluation in the consequence portion of the teaching trials in the sample-first condition. In this condition, win-stay response patterns were observed with the targets “Delta” and “Gamma.” Following an independent or prompted correct response, the experimenter presented additional information by restating the sample stimulus and describing the physical features of the visual stimuli. For example, the experimenter said, “Yes, that’s Gamma, it has two lines,” or “Great! that’s Delta, it has three lines.” The purpose of the instructive feedback was to highlight the critical features that distinguish the two auditory-visual stimulus pairs in which faulty stimulus control was observed. Instructive feedback was not implemented with the third stimuli (Phi), which the participant selected correctly on 100% of trials during this phase of the study.

3 Results

3.1 Ron - Evaluation 1

The top panel of Figure 2 displays the percentage of independent correct responses during Evaluation 1 for Ron. During baseline, Ron engaged in 33% or below independent correct responding across stimulus sets. For the first two intervention sessions, prompting was implemented at a 0-s prompt delay. Therefore, there was no opportunity to engage in independent correct responses. In each condition, independent correct responses increased after the implementation of a 5-s prompt delay. The comparison-first stimulus set met the mastery
Figure 2. Percentage of independent correct responding across baseline, intervention and follow up conditions for Ron’s first and second evaluations.
criterion first, requiring 63 instructional trials in total. The sample-first with repetition and sample-first sets both required 90 trials to reach the mastery criterion. In the follow up probe, the sample-first and comparison-first sets maintained at 100%, and the sample-first-with-repetition set maintained at 89%. In the sample-first-with-repetition set, there was a deviation from the protocol in the follow-up condition in that the experimenter prompted following an error during one trial. However, this likely did not affect the results as results in this set were similar to those obtained in subsequent follow-up probes across evaluations.

Figure 3 depicts the efficiency measures across participants and evaluations, including the number of trials to mastery (top panel), the cumulative duration of instruction to mastery (middle panel) and the average duration of instructional sittings (bottom panel). In Ron’s first evaluation, the comparison-first condition was most efficient for the number of trials to mastery and cumulative duration of instruction to mastery, but the sample-first condition had the shortest average duration per sitting.

3.2 Ron - Evaluation 2

The bottom panel of Figure 2 displays the percentage of independent correct responses during Evaluation 2 for Ron. During baseline, Ron engaged in 33% or below independent correct responses across stimulus sets. For the first two intervention sessions, prompting was implemented at a 0-s prompt delay. Independent correct responses increased after the implementation of a 5-s prompt delay across conditions. The sample-first-with-repetition set met the mastery criterion first, requiring 54 instructional trials in total. The comparison-first set was the second stimulus set to reach the mastery criterion, requiring 63 instructional trials. The sample-first set required 126 trials to reach the mastery criterion. In the follow up probe, all stimuli sets maintained at 100% independent responding.
Figure 3. Efficiency measures across participants and evaluations. The number of trials to mastery is shown in the top panel, the cumulative duration of instruction to mastery is shown in the middle panel and the average duration of instructional sittings is shown in the bottom panel. The asterisks denote that the mastery criterion was not met for that evaluation.
Figure 3 depicts the efficiency measures across participants and evaluations, including the number of trials to mastery (top panel), the cumulative duration of instruction to mastery (middle panel) and the average duration of instructional sittings (bottom panel). In Ron’s second evaluation, the sample-first-with-repetition condition was most efficient in the measures of the number of trials to mastery, the cumulative duration of instruction to mastery, and the average duration per sitting.

3.4 Tom

The top panel of Figure 4 displays the percentage of independent correct responses during Tom’s evaluation. During baseline, Tom engaged in 44%, 11% and 0% independent correct responding across the comparison-first, sample-first-with-repetition, and sample-first stimulus sets, respectively. Tom scored above chance levels (44% independent) with the materials selected for the comparison-first set, where he demonstrated a stimulus bias by selecting the same visual stimulus in 89% of trials. He selected this stimulus both when it was the discriminative stimulus and when it was the S-delta. As this pattern of responding demonstrated a stimulus bias and faulty stimulus control, rather than above-chance levels of appropriate stimulus control, the stimulus set was included in the evaluation.

In the first session for each condition, prompting was implemented at a 0-s prompt delay. Across conditions, independent correct responses increased slightly after the implementation of a 5-s prompt delay. Following the removal of reinforcement for prompted responses, the sample-first-with-repetition and comparison-first stimulus set both met mastery criteria after 72 instructional trials. The sample-first set was associated with win-stay responses, and several further modifications to the instructional protocol were made (i.e., inclusion of a DOR and inclusion of instructive feedback) before the mastery criterion was met. The sample-first set
Figure 4. The percentage of independent correct responses across baseline, intervention and follow-up for Tom and Gary's evaluations.
required 225 trials to reach the mastery criterion. In the follow up probe, the sample-first set maintained at 100% independent responding, and the sample-first-with-repetition and comparison-first sets maintained at 89% independent responding.

Figure 3 depicts the efficiency measures across participants and evaluations, including the number of trials to mastery (top panel), the cumulative duration of instruction to mastery (middle panel) and the average duration of instructional sittings (bottom panel). In Tom’s evaluation, the sample-first-with-repetition and comparison-first conditions were tied for most efficient, as they were both mastered following an equal number of trials; however, the sample-first-with-repetition condition resulted in the shortest cumulative duration to mastery. The sample-first condition resulted in the shortest average duration per sitting.

3.5 Gary

The bottom panel of Figure 4 displays the percentage of independent correct responses during Gary’s evaluation. During baseline, Gary engaged in 0% independent correct responding for all stimulus sets, and did not engage in any selection responses. Over the course of intervention, independent correct responding did not increase above chance levels. Gary’s independent correct responses had a mean score of 21% (range: 0-55%) across sessions and conditions. A DOR was implemented to increase the saliency of the auditory sample stimulus in instruction by requiring Gary to complete a unique motor movement in response to each spoken word (i.e., signing the auditory sample stimulus). Despite training, Gary did not acquire the DOR. Across the fifteen DOR training sessions that were conducted, independent correct responses did not increase above 0%, and full physical prompting was required to achieve a response across all teaching trials. Gary demonstrated prompt dependency by reaching for the
experimenter’s hands during the independent and model prompt components of the least-to-most prompting procedure.

Gary did not meet the mastery criterion in any of the conditions. In Figure 3, the conditions in which a participant did not meet mastery is indicated by an asterisk. In Gary’s evaluation, the sample-first-with-repetition condition had the shortest average duration to conduct, and the comparison-first condition was the least efficient on average.

3.6 Echoic Responses

Table 4 displays the results for echoic responses across participants, evaluations and conditions. Echoic responses occurred when a participant repeated the auditory sample stimulus that the experimenter presented at any point during the instructional trial. Echoic responses occurred most frequently in the sample-first-with-repetition conditions, except for Ron’s second evaluation, where they occurred most often in the comparison-first condition.

Table 4. The percentage of trials with echoic responses across participants.

<table>
<thead>
<tr>
<th>Participant and DV</th>
<th>Sample First</th>
<th>Sample First With Repetition</th>
<th>Comparison First</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ron - Eval 1</td>
<td>3.3%</td>
<td>8.8%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Echoic Response</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ron - Eval 2</td>
<td>4.7%</td>
<td>3.7%</td>
<td>6.9%</td>
</tr>
<tr>
<td>Echoic Response</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tom</td>
<td>5.7%</td>
<td>8.3%</td>
<td>6.9%</td>
</tr>
<tr>
<td>Echoic Response</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gary</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

3.7 Early Responses

Table 5 displays the results for early responses across participants, evaluations and conditions. Early responses occurred when a participant selected a stimulus from the comparison
array before the sample stimulus had been presented, and were only possible in the comparison-first condition. In Ron’s first and second evaluations, early responses did not occur. In Tom’s evaluation, early responses occurred in 4.2% of the comparison-first trials. In Gary’s evaluation, early responses occurred in 13% of the comparison-first trials. During the sample-first condition, Gary engaged in responses to the barrier blocking visual access to the comparison stimuli. The dependent measure of touching the barrier was added in Gary’s evaluation. Gary touched the barrier in 4% of trials in the sample-first condition.

Table 5. The percentage of trials with early responses and touching the barrier (Gary only) across participants.

<table>
<thead>
<tr>
<th>Participant and DV</th>
<th>Sample First</th>
<th>Sample First With Repetition</th>
<th>Comparison First</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ron - Eval 1 Early Response</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Ron - Eval 2 Early Response</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Tom Early Response</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
</tr>
<tr>
<td>Gary Early Response</td>
<td>0%</td>
<td>0%</td>
<td>13%</td>
</tr>
<tr>
<td>Gary Touch Barrier</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

4 Discussion

The sample-first condition was more efficient than the comparison-first condition in a study conducted by Petursdottir and Aguilar (2016). Several EIBI manuals recommend presenting stimuli similar to that arranged in the comparison-first condition (Leaf & McEachin, 1999; Maurice, Green & Luce, 1996). The present study adds to the body of research on instructional strategies for teaching listener behaviour to learners with ASD. We evaluated the effects of sample-first, sample-first-with-repetition, and comparison-first conditions. Across
participants, the most efficient procedure varied. The comparison-first presentation resulted in mastery in the fewest number of trials in Ron’s first evaluation, but the sample-first-with-repetition condition was mastered in the fewest number of trials in Ron’s second evaluation. In Tom’s evaluation, the sample-first-with-repetition and comparison-first conditions met the mastery criterion in an equal number of trials, while the sample-first set took significantly longer. Gary did not master any of the training sets.

Similar to other studies that have compared instructional strategies, no condition was consistently found to be the most efficient across participants and evaluations (Shabani & Lam, 2013). Despite the mixed findings in the present study, the results are still useful to inform clinicians about procedural variations that may differentially affect skill acquisition (Shabani & Lam, 2013). The results of this study are informative in a number of areas that have implications for listener training. These areas include: the differing results obtained across efficiency measures within evaluations, the role of stimulus control, and the presence of early responses. The relation of the results to existing research, clinical recommendations for listener behaviour instruction, limitations of the current study and potential areas of future research are also addressed.

4.1 Relation Between Efficiency Measures

Two efficiency measures were evaluated in the study: the total number of trials to mastery and the total duration of instruction to mastery. In most cases, the number of trials to mastery and cumulative duration of instructional time produced similar findings. For example, if the sample-first condition resulted in mastery in the fewest number of trials in a given evaluation, the same condition was also identified as most efficient in the total duration of instructional time.
The average duration of instructional sessions often showed a differing pattern from these other two measures. In three out of four of the evaluations, the average duration per teaching session was highest and therefore least efficient in the comparison-first condition, and lowest in the sample-first condition, while the number of trials to mastery and total duration of instructional time indicated that a different condition was most efficient. In Ron’s first evaluation, the comparison-first condition was most efficient when the number of trials to mastery or cumulative duration of instruction were considered, but was indicated to be the second most efficient by the measure of the average duration of instructional time. Without a way to predict which condition will be most efficient, using a comparison-first presentation is more likely to result in longer teaching sessions on average and may be contraindicated.

4.2 Stimulus Control

Listener behaviour involves complex stimulus control where learners must differentially respond to a multiple antecedent stimulus arrangement. Since a learner must attend to both the auditory sample stimulus and the visual comparison stimuli, faulty stimulus control is commonly encountered when teaching this skill (Green, 2001; Grow et al., 2011; Grow & LeBlanc, 2013). The temporal ordering of the stimuli may affect the development of appropriate stimulus control due to how the stimuli function within the four-term contingency.

In the comparison-first condition, the stimuli in the array function as neutral stimuli when first presented, as there is no function-altering sample stimulus present to identify stimuli in the array as the discriminative stimulus or S-deltas. It was hypothesized that the sample-first presentation could enhance the saliency of the sample stimulus by presenting it before the additional stimuli that may compete for the learners’ attention. The inclusion of a sample-first-with-repetition procedure was hypothesized to further increase the likelihood that children attend
to the sample stimulus, as the instruction is repeated and can therefore maintain its’ saliency during the passing of time over the course of the trial. The sample-first-with-repetition condition was found to be most efficient in Ron’s second evaluation, and was found to result in mastery in a shorter duration of instruction despite requiring the same number of trials as the comparison-first condition in Tom’s evaluation.

Faulty stimulus control was noted for both Ron’s second evaluation and Tom’s evaluation in the sample-first condition: both participants engaged in win-stay responses, allocating responding to the visual stimulus that had functioned as the discriminative stimulus in the preceding trial. In Ron’s second evaluation, with additional teaching, proper stimulus control was achieved. In Tom’s evaluation, several modifications had to be made once he began engaging in win-stay responses in order to meet mastery. A DOR was included; when minimal progress was observed following that change, instructive feedback with the inclusion of a rule was also added. Gary did not acquire the appropriate stimulus control in any of the training sets.

Faulty stimulus control could have developed with these participants because the discriminations targeted in Ron’s second evaluation and Tom’s evaluation, consisting of Greek letters, involved subtle distinctions in the auditory and visual comparison stimuli. In each of these stimulus sets, there were at least two stimuli that ended in “a” (e.g., lambda, delta, sigma, etc.). For Ron, the error patterns occurred with Sigma and Lambda; for Tom, the error patterns occurred with Delta and Gamma. These error patterns occurred in the sample-first condition for both participants. The faulty stimulus control may have been due to the subtle distinction between the auditory sample stimuli within the training sets. Given that this pattern was not noted in the sample-first-with-repetition condition, which included an identical preparation as the
sample-first condition aside from the repetition of the sample stimulus, it is unlikely that the characteristics of the sample-first condition themselves resulted in the faulty stimulus control.

4.3 Early Responses

In their investigation of the effects of the order of antecedent stimuli during listener behaviour instruction, Petursdottir and Aguilar (2016) measured the occurrence of early responses. Early responses occur when the participant selects a visual stimulus from the comparison array before the auditory sample stimulus has been presented. The early responses were only measured in the comparison-first condition, as the order of stimulus presentation leaves it uniquely vulnerable to this response. Petursdottir and Aguilar reported that early responses occurred in a small percentage in of trials in the comparison-first condition in two out of the four evaluations in their study.

Early responses occurred in the present study in a relatively small proportion of trials as well, but occurred more frequently with Gary, who did not have prior success with listener behaviour instruction. Ron did not engage in early responses, and Tom engaged in a small percentage of early responses. As prior listener behaviour instruction likely occurred with a comparison-first presentation, it is possible that these responses had been placed on extinction for Ron in the past, and therefore he did not engage in them during our evaluation. Overall, a higher percentage of early responses were noted in the present evaluation than the percentages reported by Petursdottir and Aguilar (2016). It is possible that children with ASD have a higher propensity to engage in early responding than typically developing children due to a history with listener instruction involving a comparison-first preparation, or due to the limitations in language and communication skills that are a characteristic deficit of ASD.
Gary engaged in the highest percentage of early responses during the comparison-first condition, such that he selected a stimulus before the experimenter-delivered auditory sample stimulus. Gary sometimes engaged in multiple early responses within the same trial, which increased the duration of the trials. The average duration of teaching sessions during the comparison-first condition for Gary’s evaluation was longer than the other conditions. After early responses emerged in the comparison-first condition, Gary began engaging in a similar response in the sample-first condition. That is, he touched the visual barrier for the comparison stimuli before the experimenter-delivered auditory sample stimulus. This also elongated the trials in which this behaviour occurred.

4.4 Comparison to Existing Research

The present study is the first to replicate and extend the research by Petursdottir and Aguilar (2016). The present results differed from the findings in this study, where the sample-first condition was found to be as or more efficient than the comparison-first condition across participants. Several procedural variations could account for these differences.

In Petursdottir and Aguilar (2016), no prompting was provided, and learning occurred through differential reinforcement only: reinforcers were delivered for correct responses but no reinforcers were delivered for errors, and no prompting was provided. It is possible that the addition of prompting procedures increased the efficiency of teaching across all conditions, thus masking the differential effect of the order of antecedent stimulus presentation on learning. The use of an evidence based transfer-of-stimulus control procedure could have resulted in effective skill acquisition regardless of the order of the antecedent presentation.

Picture cards were also used to present the material in this study, rather than an automated presentation using a computer as in Petursdottir and Aguilar (2016). The physical
barrier used in this study allowed for the control of stimuli: the researcher could arrange the stimuli out of sight of the participant before revealing the array either before or after the presentation of the sample stimulus. However, it is possible that the antecedent stimuli were not isolated enough in the tabletop presentation to have differential effects on responding. As the tabletop presentation is likely common in early intervention programs, it is worth investigating whether a computer-based or tabletop stimulus presentation has an effect on students’ acquisition.

The present study used 9-trial sessions with three targets, where each target was targeted three times. Petursdottir and Aguilar (2016) used three stimuli sets, but provided four trials of each for a total of twelve teaching trials. There was also a third experimental condition in the present study (sample-first-with-repetition), resulting in three sets of materials being taught at one time, rather than two sets in Petursdottir and Aguilar. These differences could have increased the difficulty of the present study through an increased number of stimuli being taught over the course of the evaluation and a decreased number of trials to practice each target in each experimental session.

4.5 Limitations

The participants recruited to participate in this study are different in skill level and profile than the types of participants that most clinicians would be working with when first introducing listener behaviour. Ron and Tom both had prior experience with listener training, and therefore difficult-to-acquire stimulus sets using subtle discriminations were created for the study. Gary not only had prior experience with listener behaviour, but also had a history of faulty stimulus control.
Ron was the only participant who completed two evaluations, but differing results were noted across each. Several factors could account for these differences. While all efforts were made to equate the flag stimuli, upon further examination of the stimulus sets, the sets assigned to the sample-first-with-repetition and sample-first sets were visually similar, and both presented with subtle colour discriminations (e.g., a yellow star in the Cameroon flag versus a green star in the Senegal flag as the only major difference between the two stimuli), whereas the set used in the comparison-first condition had object-based discriminations (e.g., stars in the Samoa flag and a sun in the Taiwan flag). The similarity between the sample-first-with-repetition and sample-first sets could have affected the acquisition of correct stimulus control, as both sets were taught simultaneously and were visually similar. It is also possible that due to the participant’s prior learning history, object-based discriminations were easier to acquire than colour-based discriminations. If that was the case, then the stimulus sets were not properly equated in difficulty for this evaluation.

4.6 Clinical Implications

It is imperative that faulty stimulus control be avoided, as correcting error patterns can be a lengthy and difficult process (Grow & LeBlanc, 2013). Given the complexity of listener behaviour and the propensity of faulty stimulus control in the instruction of this skill (Grow et al., 2011), practitioners need to be aware of what the desired stimulus control in listener behaviour instruction is. Practitioners should also be aware of other available methods of arranging stimuli during listener behaviour instruction, as there was found to be variability in which procedure was most efficient for which participant. If one method is not successful for a particular student, clinicians should be able to select and implement alternative procedures that
may result in improved performance. The more resources clinicians have, the better able they are to individualize treatments to their clients and optimize outcomes.

In the present study, there was not a clear differentiation of a most efficient procedure across participants. However, within a clinical context, it may be reasonable to recommend the implementation of one stimulus presentation arrangement over another, depending on the characteristics of the learner. If a client has early responses in his/her repertoire or if early responses have not been placed on extinction in the past through prior instruction in this modality, it may be prudent for the clinician to use a sample-first procedure, to remove the opportunity to engage in this type of error pattern. If a participant has a long latency to respond during an instructional trial, a sample-first-with-repetition procedure could be included to maintain the saliency of the auditory sample stimulus during instruction.

Despite the mixed findings of the present study, it may be best to avoid using a comparison-first presentation for a number of reasons. The auditory sample stimulus alters the function of the comparison stimuli, highlighting one as the discriminative stimulus and the others as the s-deltas. By presenting the comparison stimuli in the absence of the auditory sample stimulus, the instructor would be presenting neutral stimuli at best. In addition, early responses occurred in the comparison-first condition in the present study and the study conducted by Petursdottir and Aguilar (2016). These responses are only possible in the comparison-first condition and lengthen the amount of time required to conduct trials. Gary's average duration per instructional session was significantly higher in the comparison-first condition compared to the other conditions due at least partially to the presence of early responses. Early responses were not seen at all in Ron’s evaluations and were seen less in Tom’s evaluation, likely due to these participants’ prior learning history with listener behaviour tasks. Given the inconclusiveness of
the present investigation and the findings of previous research that the sample-first condition was most efficient (Petursdottir & Aguilar, 2016), it may be best to conduct listener behaviour instruction using a sample-first presentation.

4.7 Future Research

Researchers should replicate the study with children that are starting listener behaviour training. Ideally participants should have the prerequisite repertoires for this skill (e.g., visual-visual matching, generalized imitation), but should have had no or limited prior experience with listener behaviour training, and should not have a history of faulty stimulus control with this skill. The ordering of antecedent stimuli may have a larger effect on acquisition of learners who are likely to engage in early responses, biased response patterns, or who are inexperienced with any of the presentations (i.e., have not had prior exposure to one of the procedures in the past), or the preparations of listener behaviour instruction (i.e., this is a new task for them).

Neither the present study, nor the study conducted by Petursdottir and Aguilar (2016), investigated a simultaneous stimulus presentation. Further comparisons could be conducted evaluating the effects of slight modifications to the antecedent presentations. One option would be to compare a simultaneous presentation with a sample-first presentation. It would be reasonable to assume that presenting the two antecedent events simultaneously may result in reduced duration of instruction, as long as the simultaneous presentation resulted in learning that was as or more efficient than a sample-first presentation. A simultaneous presentation may be preferable to a comparison-first presentation, as the comparison stimuli would not be presented in the absence of the sample stimulus if they are presented at the same time.

Other research could investigate whether or not the sample-first-with-repetition condition could be modified to facilitate the echoic response on the part of participants who have an echoic
repertoire, by requiring the participant to repeat the sample stimulus. The participant being prompted to repeat the sample stimulus during listener behaviour instruction could have a facilitative effect on the acquisition of the corresponding tact response, which may then emerge without additional training (i.e., naming; Horne & Lowe, 1996). Repeating the sample stimulus in listener behaviour instruction could also be facilitative for acquiring the listener behaviour sets, as a form of preserving the sample stimulus during the delay between the issuing of the sample stimulus and the participant selecting a target response (i.e., joint control; Causin, Albert, Carbone & Sweeny-Kerwin, 2013; Lowenkron, 1991/1998).

Finally, early intervention manuals only recommend that visual-visual matching, imitation, or auditory simple discriminations be taught prior to listener behaviour training (Leaf & McEachin, 1999; Maurice, Green & Luce, 1996). Gary did not successfully acquire the listener behaviour skills taught in this study, despite meeting most of these pre-requisites. Given that he also did not fully meet the inclusion criteria for the study, future research could examine which pre-requisite skills are essential and more likely to result in improved performance in listener behaviour instruction. That way, clinicians could ensure participants have prior to training in these skills prior to attempting listener behaviour instruction.

5 Conclusion

The present study replicated and extended the work of Petursdottir and Aguilar (2016) and investigated the effects of sample-first, sample-first-with-repetition, and comparison-first antecedent stimulus presentation arrangements on the acquisition of listener behaviour for three children diagnosed with ASD. Whereas in Petursdottir and Aguilar, the sample-first presentation was found to be most efficient across participants, the most efficient procedure in the current study varied across evaluations. In addition, measures comparing the number of trials to mastery,
cumulative duration of instruction to mastery, and average duration of instructional sittings did not always identify the same condition as most efficient within an evaluation. However, the comparison-first condition required the most time to conduct on average than the other two experimental conditions.

Data were also collected on the occurrence of early and echoic responses. Participants were most likely to engage in early responses in the comparison-first condition. Participants were most likely to repeat the auditory sample stimulus (echoic response) in the sample-first-with-repetition condition. The occurrence of early responses is likely to increase the amount of trials required to conduct listener behaviour trials.

Given the variability of results in the present evaluation, the complexity of the conceptual foundation of the stimulus control for listener behaviour, and the presence of early responses in the comparison first condition, it is recommended that practitioners use a sample-first stimulus presentation arrangement when teaching listener behaviour instruction. Further research is also required in this area to better guide the selection of efficient teaching procedures on the part of practitioners, to best address the skill deficits and provide the best hopes for improved outcomes for students with ASD.
References


doi:10.1002/jaba.16