# Evaluation of Mixed Multiple- and Paired-Stimulus Preference Assessments with Adults with Developmental Disabilities

by

**Carly Thiessen**

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

**Master of Arts**

Department of Psychology

University of Manitoba

Winnipeg, Manitoba, Canada

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

# Evaluation of Mixed Multiple- and Paired-Stimulus Preference Assessments with Adults with Developmental Disabilities

By

**Carly Thiessen**

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

**Master of Arts**

Abstract

Direct preference assessment is an effective method for identifying preferred stimuli and reinforcers for people with severe and profound developmental disabilities with limited or no communication skills. The multiple-stimulus without replacement (MSWO) assessment procedure has been found to be as effective as and more time efficient than the paired-stimulus (PS) procedure. However, individuals with severe and profound developmental disabilities may be unable to discriminate and respond reliably to multiple-stimulus assessments, which typically involve 6-8 stimuli. Therefore, it would be beneficial to evaluate whether a multiple-stimulus assessment with fewer stimuli would be as effective and time efficient as the PS assessment. The present study evaluated a mixed multiple- and paired-stimulus (MS/PS) preference assessment procedure that involved presenting 4 stimuli simultaneously and compared its effectiveness with the MSWO procedure that involved presenting 8 stimuli and with the PS procedure. The most preferred item identified by each preference assessment procedure was then tested for its reinforcing effectiveness. The results showed that both the MS/PS and MSWO procedures were positively and moderately correlated with the PS procedure (average *tau* = 0.53 and 0.57, respectively). The average number of trials to complete the assessment was 29 for the MS/PS procedure, 20 for the MSWO procedure, and 56 for the PS procedure. The most preferred stimuli identified by each procedure functioned as positive reinforcers during reinforcer evaluation. The MS/PS procedure did not produce the anticipated benefits. The results may be accounted for by the participants' ability to discriminate more than 4 stimuli concurrently.

Acknowledgements

I thank the participants for their cooperation and the families' support throughout the study. I thank Gareth Davies, Lindsay Arnal, and Alejandra Zaragoza Scherman for their assistance in reliability assessments and data collection. I thank my advisor, Dr. C. T. Yu, and my committee members, Dr. Garry Martin and Dr. Dennis Hrycaiko, for their support and guidance in completing this project. This research was supported in part by a University of Manitoba Department of Psychology Fellowship and by grant MOP77604 from the Canadian Institutes for Health Research.

Lastly, I thank my family and friends for their encouragement, optimism, support, and understanding. Thank you to my Mom for always being there and listening, Liam and Alexi for being my ultimate supporters and my escape from stress, Dad for the confidence in me, and Alex, for everything.

Table of Contents

List of Tables and Figures

# Evaluation of Mixed Multiple- and Paired-Stimulus Preference Assessments with Adults with Developmental Disabilities

For persons with severe and profound developmental disabilities with little or no verbal communication, determining their preferences can be challenging for caregivers. An effective method to differentiate preferences and identify reinforcers for this population is through direct preference assessments (Hagopian, Long, & Rush, 2004). A direct preference assessment involves systematically presenting various choices and observing an individual's approach responses or engagement with those choices. One area of investigation has focused on comparing the effectiveness and efficiency of presenting one, two, or more choices concurrently during preference assessments. The paired-stimulus (PS) presentation procedure (presenting two choices at a time) has been found to be more effective than the single-stimulus presentation procedure (presenting one stimulus at a time) in differentiating preferences. The multiple-stimulus without replacement (MSWO) procedure, with 6-8 choices, has been shown to be more efficient and at least as effective relative to the PS procedure. However, recent findings have shown that not all individuals with developmental disabilities are capable of responding to multiple-stimulus preference assessments due to the large number of items (Davies, Thiessen, Yu, Martin, & Martin, 2009). Therefore, the current research examined the effectiveness and efficiency of a mixed multiple- and paired-stimulus (MS/PS) presentation in order to reduce the stimulus array size.

## *Preference Assessment and Reinforcer Evaluation*

Preference is defined as a pattern of responding to stimuli in choice situations (Martin, Yu, Martin, & Fazzio, 2006). For people with developmental disabilities, preference assessments involve presenting choices systematically and measuring their approach or selection responses (e.g., pointing to or reaching for a stimulus). In examining the effectiveness of preference assessments, researchers have asked two main questions: does the assessment identify a preferred stimulus, and is the identified preferred stimulus a positive reinforcer? A preferred stimulus is typically defined as one that is chosen on at least 80% of the assessment trials in which it is presented relative to other available choices (Pace, Ivancic, Edwards, Iwata, & Page, 1985). To determine if a preferred stimulus functions as a positive reinforcer, a reinforcer evaluation is carried out using the identified stimuli. This usually involves measuring the frequency of a target behaviour (e.g., pressing a micro-switch) during a baseline phase, in which no programmed consequences follow the occurrence of the target behaviour, and during a reinforcement phase, in which the behaviour is followed by the presentation of the stimulus being evaluated. The baseline and reinforcement phases are usually repeated in a reversal design. An increase in responding during the reinforcement phases and a decrease in responding during the baseline phases are necessary to conclude that the stimulus is a positive reinforcer (e.g., Lee, Yu, Martin, & Martin, in press).

*Number of Stimuli during Preference Assessments*

The number of stimuli available on each trial during preference assessments affects its effectiveness and efficiency. Research has investigated several values: presenting on each trial 1 stimulus (single-stimulus), 2 stimuli (paired-stimulus), and 6 to 8 stimuli (multiple-stimulus). To assess the preferences for an array of stimuli, the single-stimulus procedure consists of presenting each stimulus of the array individually and in random order for a fixed number of trials. The participant is given the opportunity to approach each stimulus and the stimulus that has been approached most frequently is considered the most preferred (Pace et al., 1985).

The paired-stimulus or PS procedure involves presenting stimuli in pairs on each trial, and giving the participant an opportunity to choose one of the stimuli (Fisher et al., 1992). Each stimulus is paired with every other stimulus in the array twice and presented in counterbalanced positions. Again, the stimulus that is chosen on the highest percentage of trials based on its availability is the most preferred stimulus.

In the multiple-stimulus *with* replacement (MSW) procedure, all of the stimuli in the array are presented concurrently on each trial and the client is given the opportunity to choose one. Windsor, Piche and Locke (1994) assessed the preference rankings of 6 stimuli with 8 adults with severe to profound developmental disabilities using MSW procedure and the PS procedure. The participants were described as being able "to discriminate and visually scan without difficulty" although no formal discrimination assessments were conducted. They found that similar preferences were identified with both procedures, and the same items were identified as most preferred for 6 of the 8 participants. The PS procedure, however, produced slightly more consistent preference rankings across successive administrations. Rank correlations (Kendall's Tau) between the PS and MSW procedures averaged .75 across the 8 participants, ranging from .48 to 1. The MSW procedure was more efficient, requiring a total of 50 trials while the PS procedure required 150 trials to administer. Reinforcer evaluation of the preferred stimuli identified was not conducted in this study.

DeLeon and Iwata (1996) compared an MS *without* replacement (MSWO) procedure to the PS and MSW procedures. This procedure was similar to the MSW procedure except that a selected stimulus was not presented on subsequent trials. The discrimination skills of the participants were not reported. Across 7 adults with profound developmental disabilities and using an array of 7 stimuli, the PS and MSWO procedures produced similar preference rankings, with moderate to high correlations (mean $r = .72$). The MSW procedure resulted in slightly lower correlations (mean $r = .61$) and produced more unselected items than the other two procedures. The MSW and MSWO procedures required comparable times to administer, and less than half the time that the PS procedure required. The reinforcing effectiveness of the most preferred stimuli identified were not tested. However, the researchers evaluated the reinforcing effectiveness of the stimuli that were never selected during the MSW procedure, but had been selected in the other two assessments. For 3 of the 4 participants, those stimuli maintained higher rates of responding during reinforcement phases than in baseline, which suggested that the MSW procedure was more likely than the other procedures to miss identifying reinforcers.

Carr, Nicholson and Higbee (2000) evaluated the MSWO procedure with 8 stimuli and they also reduced the number of assessment sessions from five (DeLeon & Iwata, 1996) to three, resulting in a total of 21 trials to administer the assessment. Three children diagnosed with autism participated. The discrimination skills of the children were not reported. Reinforcer evaluations were conducted for all children following the assessments. Target responses included low frequency behaviours from each child's curriculum. Three stimuli with different preference rankings (high, first ranked; medium, fourth or fifth ranked; and low, eighth ranked) were evaluated for each child. For all children, the low preference stimuli produced small or no increase in responding during the reinforcement phase relative to baseline. The medium preference stimuli increased responding slightly in two children, but had no effect for the remaining child. The high preference stimuli produced the highest response rates during reinforcement relative to baseline and relative to the low and medium preference stimuli for all children.

Graff and Ciccone (2002) evaluated the MSWO procedure with 15 individuals with developmental disabilities. Each assessment included 7 trials, and was conducted 7 times, resulting in a total of 49 trials. Twelve of the 15 individuals completed a second set of preference assessments 6 months later, resulting in a total of 27 data sets for analysis. The authors completed post hoc analyses of their data to determine whether the highest ranked stimuli identified in 7 sessions could have been identified in fewer sessions and/or fewer trials per session. They determined that when the number of sessions was reduced from 7 to 5, the top ranked stimuli were the same for 22 of the 27 data sets. When the number of sessions was further reduced to 3, the top ranked stimuli were the same in 19 of the 27 data sets. When the number of trials was reduced from 6 to 3, the top ranked stimuli were the same in 25 of the 27 data sets. Four of the 15 individuals participated in reinforcer assessments following the preference assessments, where the target response was a button press. An ABAB replication-reversal design was used for the reinforcer assessments, during which no programmed consequence was provided for responding during baseline phases and the top-ranked stimulus was delivered on a variable-ratio schedule contingent on the target response during reinforcement phases. The top-ranked stimuli functioned as positive reinforcers for all four participants.

Davies et al. (2009) compared the PS and MSWO procedures to assess the preferences for 6 food items with 9 adults with developmental disabilities. They found that the stimulus preference rankings from the MSWO and PS procedures were positively correlated for all participants (mean *tau* = .72). They also found that four participants did not select the most-preferred stimuli identified by the PS assessments as their most-preferred stimuli during the MSWO assessments. Preferred-stimulus selection percentages for those participants generally increased during MSWO as the number of stimuli decreased across trials, suggesting possible difficulties in discrimination involving multiple stimuli.

*Statement of the Problem*

Overall, previous research has shown that the MSWO preference assessment procedure is as effective as the PS procedure in differentiating preferences and more efficient than PS assessment (e.g., DeLeon & Iwata, 1996). However, for the MSWO procedure to be effective, the person being assessed must be able to perform a simultaneous visual discrimination involving an array of 6-8 stimuli, which could be challenging for individuals with developmental disabilities (Davies et al., 2009). Given the importance of conducting direct preference assessments, the considerable number of trials needed for the PS procedure, and recognizing that individuals with developmental disabilities may have difficulty discriminating among 6 or more stimuli, the present research examined an alternative preference assessment procedure. This procedure combined MSWO involving smaller arrays and PS procedures to determine the preference ranking among stimuli in the arrays (hereafter referred to as the MS/PS procedure). This procedure was compared to the standard MSWO and PS procedures. In addition, the reinforcing effectiveness of the most preferred stimuli identified in each procedure were evaluated.

## Method

*Participants and Setting*

Five adults with moderate to profound developmental disabilities were recruited from St. Amant, a residential and community agency for people with developmental disabilities (see Table 1). Participants were also assessed on the *Assessment of Basic Learning Abilities* test (ABLA; Kerr, Meyerson, & Flora, 1977; Vause, Yu, & Martin, 2007).

Table 1

*Participants' Characteristics*

| Participant | Age (Yrs) | Sex | Mental Retardation[a] | SIB-R[b] Broad Independence | SIB-R Social/ Communication | ABLA[c] Discrimination Passed |
|---|---|---|---|---|---|---|
| 1 | 42 | F | Profound | 19 months | 10 months | Auditory-visual |
| 2 | 38 | M | Severe | 43 months | 30 months | Matching-to-sample |
| 3 | 52 | F | Moderate | 21 months | 22 months | Auditory-visual |
| 4 | 47 | F | Severe | 8 months | below 3 months | Simple visual |
| 5 | 42 | M | Severe | 21 months | 10 months | Matching-to-sample |

[a] According to agency records

[b] *Scales of Independent Behaviors-Revised* (Bruininks, Woodcock, Weatherman, & Hill, 1996)

[c] *Assessment of Basic Learning Abilities* (Kerr et al., 1977)

All research sessions were conducted in an assessment room at the St. Amant Research Centre. The tester and a participant sat at a table facing each other during all sessions. An observer was present during some sessions to conduct reliability checks. The University of Manitoba Psychology/Social Research Ethics Board approved the study and written informed consent from each participant's legal decision-maker was obtained before the study commenced.

*Instruments and Materials*

Standard ABLA test materials included a round yellow can, a square red box with diagonal black stripes, a small white/beige piece of foam, a small yellow cylinder and a small red cube (DeWiele & Martin, 1998). Various food items were used for preference assessments. A round micro-switch (6 cm in diameter) was used for the reinforcer assessments. The switch produced an audible "click" when 2-3 g of force was applied.

*Procedure*

*Paired-stimulus preference assessment.* The PS assessment was completed for each participant to identify 8 food stimuli, which were used in subsequent phases of the study. The PS preference assessment involved presenting two items on each trial, and asking the participant to choose one. The chosen item was given to the participant immediately for consumption before the next trial was presented. Each item was paired with every other item, and each pairing was presented twice with the items in counterbalanced positions, which yielded 56 trials for 8 stimuli. The stimulus pairings were presented in a random order.

During the paired-stimulus preference assessments (and all other subsequent preference assessments described below), an item was considered chosen if the participant emitted an approach response, which was defined as either touching or pointing to the item within 8 seconds following the presentation of the stimuli and the instruction to choose. After each trial, the experimenter (and the observer if present) recorded which stimulus had been selected, if any. If an approach response did not occur within 8 s, the trial would be recorded as "no selection" and the next trial was presented. The preference value for each food item was computed as the percentage of trials it was selected out of the number of trials it was presented.

The PS assessment was conducted until at least a high preference stimulus (selected on 80% or more of the trials), a low preference stimulus (selected on not more than 20% of the trials), and 6 other stimuli with preferences between the most and least preferred stimuli were identified.

*Multiple-stimulus without replacement assessment.* The MSWO procedure was conducted for the 8 food items identified by the PS assessment. Each assessment consisted of 3 sessions of up to 7 trials each, totalling a maximum of 21 trials. During the first trial of each session, all 8 items were presented in a row in quasi-randomized positions on the table in front of the participant, who was asked to choose one. A selected item was not replaced on the next trial. On each subsequent trial, the item that was positioned on the far left of the previous trial was moved to the far right and the rest of the items were shifted one position to the left. A session ended if no item was chosen on a trial or after the 7th trial (when only one item remained on the table).

*Multiple/paired-stimulus assessment.* In the MS/PS assessment, the 8 food items were split randomly into two arrays of 4 stimuli each (e.g., arrays 1 and 2). Each 4-stimulus array was then assessed using the MS procedure as described above (up to 3 trials per session). A maximum of 18 trials were needed to complete the MS assessments for both arrays. The next step involved a PS assessment procedure using one item from each array to determine the relative preference among the 8 items. Stimulus pairs were formed and assessed in the order shown in Table 2 to achieve maximum efficiency, starting with the top ranked item from one array being paired with the bottom ranked item from the other array. For example, if the items were ranked A > B > C > D (where “>" means “ranked above”) for array 1 and E > F > G > H for array 2, I would pair items E and D for the PS assessment (Step 1 of Table 2). This required an additional 2 trials (to counterbalance positions). If the results showed that D > E or D = E, I would be able to derive the rank order of the 8 items to be A > B > C > D > E > F > G > H or A > B > C > D/E > F > G > H, respectively (see Table 2). However, if E > D, I would pair E with C, the next higher ranked item above D, for a PS assessment, and so on. Depending on item E’s ranking relative to the items in array 1, items F, G and H might be paired with the stimuli from array 1 to determine their rankings. Table 2 illustrates all possible pairings that might occur, which could require a maximum of 32 trials. Adding these trials to the initial 18 MS trials, the MS/PS procedure would require a maximum of 50 trials.

*Design.* Each participant received the three preference assessments procedures in an ABC design, where A refers to PS, B refers to MSWO, and C refers to MS/PS. Participants 1, 3 and 5 received the ABC sequence and Participants 2 and 4 received the ACB sequence. The rankings produced by the MSWO and MS/PS assessments were compared to the A phase (PS assessment). The A phase was repeated for Participant 3 after the MS/PS assessment to evaluate preference stability because her data suggested a possible preference shift.

*Reinforcer Assessment*

After completing all preference assessments, the most preferred items identified for each participant in the PS, MSWO, and MS/PS assessments were evaluated for their

Table 2

*Paired Stimulus Algorithm of Comparison Steps*

| Step | | Stimuli | Results | | Proceed to | Overall Derived Ranking |
|---|---|---|---|---|---|---|
| 1.00 | | E vs. D | D>E | | Finish | A B C D E F G H |
| | | | D=E | | Finish | A B C D/E F G H |
| | | | D<E | | Step 2.00 | |
| 2.00 | | E vs. C | C>E or C=E | | Step 2.10 | |
| | | | C<E | | Step 3.00 | |
| | 2.10 | F vs. D | | D>F | Finish | A B C E D F G H |
| | | | | D=F | Finish | A B C E D/F G H |
| | | | | D<F | Step 2.11 | |
| | 2.11 | G vs. D | | D>G | Finish | A B C E F D G H |
| | | | | D=G | Finish | A B C E F D/G H |
| | | | | D<G | Step 2.12 | |
| | 2.12 | H vs. D | | D>H | Finish | A B C E F G D H |
| | | | | D=H | Finish | A B C E F G D/H |
| | | | | D<H | Finish | A B C E F G H D |
| 3.00 | | E vs. B | B>E or B=E | | Step 3.10 | |
| | | | B<E | | Step 4.00 | |
| | 3.10 | F vs. D | | D>F | Finish | A B E C D F G H |
| | | | | D=F | Finish | A B E C D/F G H |
| | | | | D<F | Step 3.11 | |
| | 3.11 | G vs. D | | D>G | Step 3.20 | |
| | | | | D=G | Step 3.20 | |
| | | | | D<G | Step 3.12 | |
| | 3.12 | H vs. D | | D>H | Step 3.20 | |
| | | | | D=H | Step 3.20 | |
| | | | | D<H | Step 3.20 | |
| | 3.20 | F vs. C | | C>F | Finish | A B E C F (DGH) |
| | | | | C=F | Finish | A B E C/F (DGH) |
| | | | | C<F | Step 3.21 | |
| | 3.21 | G vs. C | | C>G | Finish | A B E F C G (DH) |
| | | | | C=G | Finish | A B E F C/G (DH) |
| | | | | C<G | Step 3.22 | |
| | 3.22 | H vs. C | | C>H | Finish | A B E F G C (HD) |
| | | | | C=H | Finish | A B E F G C/H D |
| | | | | C<H | Finish | A B E F G H C D |

Table 2 continued

| Step | Stimuli | Results | Proceed to | Overall Derived Ranking |
|---|---|---|---|---|
| 4.00 | E vs. A | A>E or A=E | Step 4.10 | |
| | | A<E | Step 5.10 | |
| 4.10 | F vs. D | D>F | Finish | A E B C D F G H |
| | | D=F | Finish | A E B C D/F G H |
| | | D<F | Step 4.11 | |
| 4.11 | G vs. D | D>G | Step 4.20 | |
| | | D=G | Step 4.20 | |
| | | D<G | Step 4.12 | |
| 4.12 | H vs. D | D>H | Step 4.20 | |
| | | D=H | Step 4.20 | |
| | | D<H | Step 4.20 | |
| 4.20 | F vs. C | C>F | Finish | A E B C F (DGH) |
| | | C=F | Finish | A E B C/F (DGH) |
| | | C<F | Step 4.21 | |
| 4.21 | G vs. C | C>G | Step 4.30 | |
| | | C=G | Step 4.30 | |
| | | C<G | Step 4.22 | |
| 4.22 | H vs. C | C>H | Step 4.30 | |
| | | C=H | Step 4.30 | |
| | | C<H | Step 4.30 | |
| 4.30 | F vs. B | B>F | Finish | A E B F (CDGH) |
| | | B=F | Finish | A E B/F (CDGH) |
| | | B<F | Step 4.31 | |
| 4.31 | G vs. B | B>G | Finish | A E F B G (CDH) |
| | | B=G | Finish | A E F B/G (CDH) |
| | | B<G | Step 4.32 | |
| 4.32 | H vs. B | B>H | Finish | A E F G B H C D |
| | | B=H | Finish | A E F G B/H C D |
| | | B<H | Finish | A E F G H B C D |

Table 2 continued

| Step | Stimuli | Results | Proceed to | Overall Derived Ranking |
|---|---|---|---|---|
| 5.00 | Known: | E > A | | |
| 5.10 | F vs. D | D>F | Finish | E A B C D F G H |
| | | D=F | Finish | E A B C D/F G H |
| | | D<F | Step 5.11 | |
| 5.11 | G vs. D | D>G | Step 5.20 | |
| | | D=G | Step 5.20 | |
| | | D<G | Step 5.12 | |
| 5.12 | H vs. D | D>H | Step 5.20 | |
| | | D=H | Step 5.20 | |
| | | D<H | Step 5.20 | |
| 5.20 | F vs. C | C>F | Finish | E A B C F (DGH) |
| | | C=F | Finish | E A B C/F (DGH) |
| | | C<F | Step 5.21 | |
| 5.21 | G vs. C | C>G | Step 5.30 | |
| | | C=G | Step 5.30 | |
| | | C<G | Step 5.22 | |
| 5.22 | H vs. C | C>H | Step 5.30 | |
| | | C=H | Step 5.30 | |
| | | C<H | Step 5.30 | |
| 5.30 | F vs. B | B>F | Finish | E A B F (CDGH) |
| | | B=F | Finish | E A B/F (CDGH) |
| | | B<F | Step 5.31 | |
| 5.31 | G vs. B | B>G | Step 5.40 | |
| | | B=G | Step 5.40 | |
| | | B<G | Step 5.32 | |
| 5.32 | H vs. B | B>H | Step 5.40 | |
| | | B=H | Step 5.40 | |
| | | B<H | Step 5.40 | |
| 5.40 | F vs. A | A>F | Finish | E A F (BCDGH) |
| | | A=F | Finish | E A/F (BCDGH) |
| | | A<F | Step 5.41 | |
| 5.41 | G vs. A | A>G | Finish | E F A G (BCDH) |
| | | A=G | Finish | E F A/G (BCDH) |
| | | A<G | Step 5.42 | |
| 5.42 | H vs. A | A>H | Finish | E F G A H B C D |
| | | A=H | Finish | E F G A/H B C D |
| | | A<H | Finish | E F G H A B C D |

*Note*. The algorithm is written assuming that rankings from initial MS stimulus arrays (from most to least preferred) was Array 1: A B C D; Array 2: E F G H. Each stimulus pairing indicates two assessment trials to counterbalance stimuli positions. In Overall Derived Ranking column, letters from left to right indicate order of preference from most to least preferred. / = tie for rank spot between the two stimuli. ( ) = letters inside may be ranked in a number of ways; actual ranking would have been determined in an earlier step.

reinforcing effectiveness in a combined alternating-conditions and ABAB design. All sessions were conducted with the experimenter seated across a table from the participant, with the micro-switch placed in front of the participant. Pressing the micro-switch to produce an audible “click” was the target response in all conditions.

During a baseline phase, pressing the switch received no programmed consequences. I modelled and guided the participant to press the switch once at the beginning of each session, and then gave the participant an instruction to “press the switch”. Approximately once per minute for the remainder of the session, I re-issued the instruction to “press the switch”. Also once per minute, I praised the participant for a desirable response other than pressing the button (e.g., sitting nicely, paying attention, etc.). Each session lasted 3 minutes. The time and occurrence of each switch press was recorded. Number of switch presses per minute was computed for each session. Once the rate of responding across baseline sessions was stable, a reinforcement phase began.

During a reinforcement phase, the most preferred food item from each of the three assessments was presented following every switch press. One item was used as the consequence per session and all items were presented for one session each before they were repeated. The order that the items were presented was counterbalanced across every block of two sessions.

During each session in a reinforcement phase, I began the session by modelling and instructing the participant to press the switch, similar to a baseline session. After each switch press, I immediately removed the switch from the table and presented the food item to the participant for consumption. After the participant had finished consuming the food item, I returned the switch to the table. The time and occurrence of each switch press was recorded. A stop watch was used to ensure that the participant had 3 minutes for responding (excluding time taken to administer and consume the food item). Rate of responding per minute for each session was computed.

*Reliability and Procedural Integrity Checks*

Interobserver reliability checks occurred on 50% of all preference assessment sessions and procedural integrity checks occurred on 44% of the preference assessment sessions. During interobserver reliability checks, the observer independently recorded the participant's selection on each trial. To compute percent agreement, a trial was defined as an agreement if both the observer and the experimenter recorded the same chosen item; otherwise, it was a disagreement. Percent agreement for a session was calculated by dividing the number of agreements by the sum of agreements and disagreements, and multiplying the quotient by 100% (Martin & Pear, 2007). Percent agreement during preference assessments was always 100%.

For procedural integrity, the observer evaluated the experimenter using a checklist of steps to be followed on each trial (e.g., the correct stimuli were presented, stimuli were presented in the correct positions, the correct instruction to choose was given, and the selected item was given to the participant within 2 s). A trial was scored as correct if all steps were carried out correctly. The percentage of trials delivered correctly per session was 100% during all observed sessions.

Interobserver reliability and procedural integrity checks were also conducted for 40% of the reinforcer assessment sessions. For interobserver reliability checks, the observer independently recorded the time of each switch press by the participant. The session (3-minute responding time) was divided into 10-s intervals, and the responses recorded by the experimenter and by the observer that fell within each interval were compared. Percent agreement was calculated using a partial interval, occurrence and nonoccurrence agreement method. An interval was scored as an agreement if both the experimenter and the observer had recorded no response or both had recorded at least 1 response; otherwise, it was scored as a disagreement. The percentage of agreement intervals was computed for each session. Mean percent agreement during reinforcer assessments was 95.4%, range = 83 - 100%.

For procedural reliability, the observer evaluated the experimenter's behaviours using a checklist to be followed during a session (e.g., the initial demonstration and the initial instruction were completed correctly at the beginning of the session, the micro-switch was removed after each response during reinforcement, and the correct consequence was provided within 2 s following a switch press). The percentage of correct steps carried out per opportunity was computed for each session. Procedural integrity during reinforcer assessments was always 100%.

## Results

Figure 1 shows the percentage of trials each food stimulus was selected during PS, MSWO, and MS/PS preference assessments for each participant (P1 through P5). Stimuli are ordered from most to least preferred on the horizontal axis according to the PS assessment results. For P3 who received two PS assessments, the first was used for this comparison. The MSWO procedure identified the same most preferred stimulus as the PS procedure for P2 and P4. The most preferred stimulus from the PS assessment was ranked second for P3 and P5 and third for P1 by the MSWO procedure. The MSWO procedure identified the same least preferred stimulus as the PS procedure for three participants (P1, P2, and P5). The least preferred stimulus from the PS assessment was ranked sixth for P4 and fourth for P3.

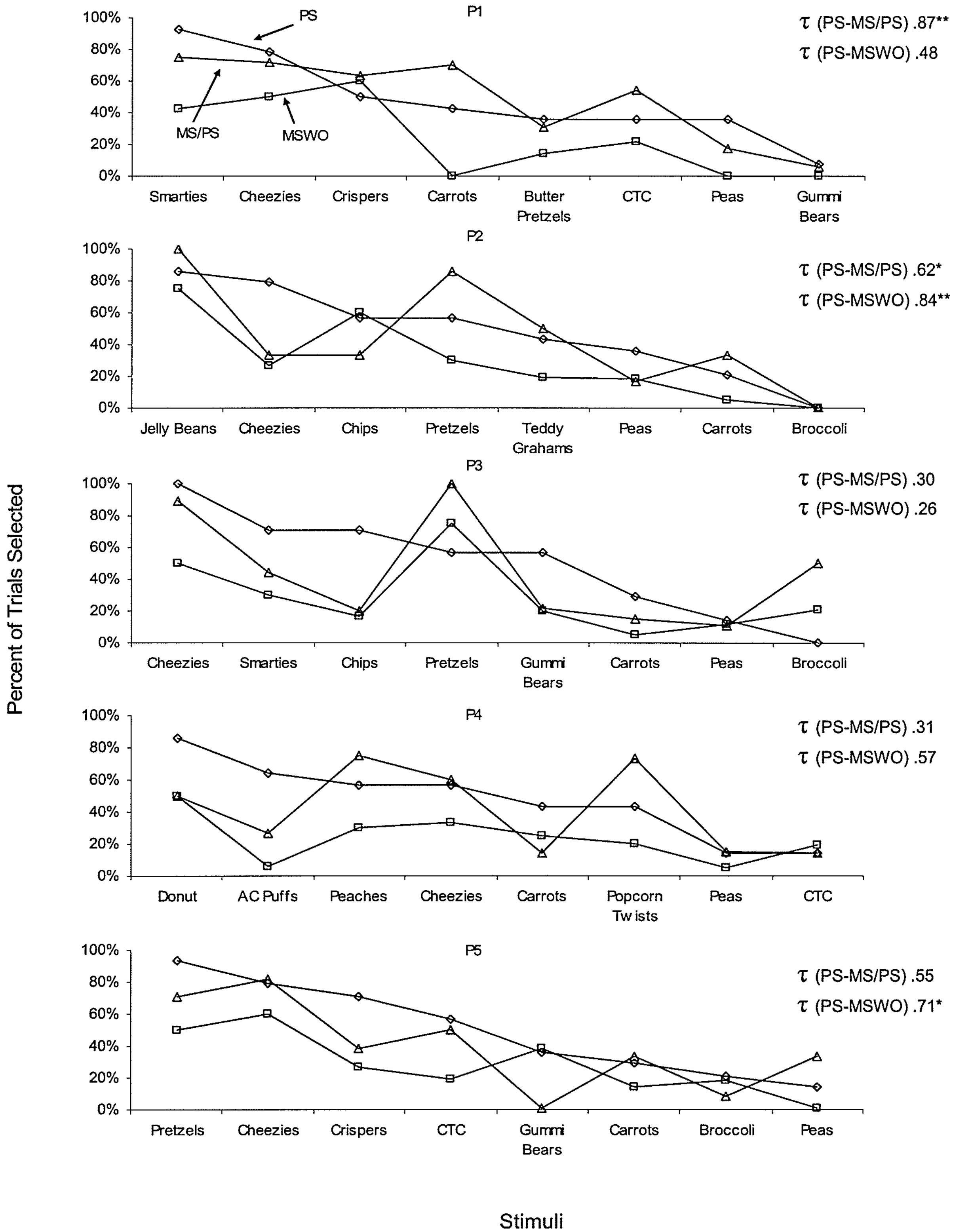


*Figure 1.* Percentage of trials each food stimulus was selected during PS, MSWO, and MS/PS preference assessments for each participant. Kendall's Tau correlations are shown in each graph. Significant correlations are noted by * ($p < .05$) and ** ($p < .01$). CTC refers to cinnamon toast crunch and AC Puff refers to apple cinnamon puffs.

The MS/PS procedure identified the same most preferred stimulus as the PS procedure for P1 and P2. The most preferred stimulus from the PS assessment was ranked second for P3 and P5, and fourth for P4. The MS/PS procedure identified the same least preferred stimulus as the PS procedure for three participants (P1, P2, and P4). The least preferred stimulus from the PS assessment was ranked third for P3 and fifth for P5.

Correlation coefficients (Kendall's tau) between PS and the MSWO and MS/PS procedures are shown in each participant's graph in Figure 1. Stimulus preference ranks from the MSWO and PS procedures were positively correlated for all participants (mean *Tau* = .57; range, .26 – .84), and significantly correlated ($p = .05$, 2-tailed) for P2 and P5. Stimulus preference ranks from the MS/PS and PS procedures were positively correlated also for all participants (mean *Tau* = .53; range, .30 – .87), and significantly correlated ($p = .05$, 2-tailed) for P1 and P2. The lowest correlations for both MSWO and MS/PS procedures were observed in P3. Therefore, the PS assessment was repeated to evaluate whether P3's preference had changed since the first assessment. The results indicated that her preferences had remained relatively stable (*tau* between the two PS administrations was .89). The three top ranked stimuli retained their rankings in both PS assessments, whereas the lowest three ranked stimuli switched positions in the second administration.

The MSWO procedure required an average of 20 trials to complete, ranging from 18 to 21 trials across participants. The MS/PS procedure required an average of 29 trials to complete, ranging from 24 to 36 trials across participants. The PS procedure required the most trials to complete for all participants. Four participants required the minimum 56 trials to establish a preference hierarchy using the PS procedure, and P3 required 112 trials to establish a hierarchy.

Figure 2 shows the number of switch presses per minute per session during reinforcer evaluation. The most preferred stimulus identified in each procedure was evaluated. The rate of responding increased during reinforcement over baseline phases for all participants. All stimuli evaluated were approximately equally effective.

To examine whether the number of choices available had affected performance during preference assessment, I examined the percentage of trials that the most-preferred stimulus was selected on trials with a different number of available choices, and calculated the proportion it deviated from chance. The procedure for calculating this deviation involved the following steps. First, trials with the same number of choices were aggregated (e.g., trials with 8 choices, trials with 7 choices, and so on). Second, the percentage of trials that the most-preferred stimulus was selected was computed (e.g., 66% if the participant had selected the most-preferred stimulus on two of the three 8-choice trials). Third, because the probability of selecting the most-preferred stimulus by chance differed based on the number of available choices (e.g., 12.5% on 8-choice trials and 50% on 2-choice trials), a comparison of the percentages that the most-preferred stimuli were selected would not have been appropriate (i.e., a 66% selection of the most preferred stimulus is 53.5% above chance on 8-choice trials, but only 16% above chance on 2-choice trials). Therefore, selection of the most-preferred stimulus was expressed as the "proportion of deviation from chance" using the following formula:

(% most-preferred stimulus selected – chance %) / (100 – chance %)

For example, a 66% selection of the most preferred stimulus would yield a .61 deviation from chance on 8-choice trials ([66 – 12.5] / [100 – 12.5]) and only .32 deviation from chance on 2-choice trials ([66 – 50] / [100 – 50]. The proportion of deviation from

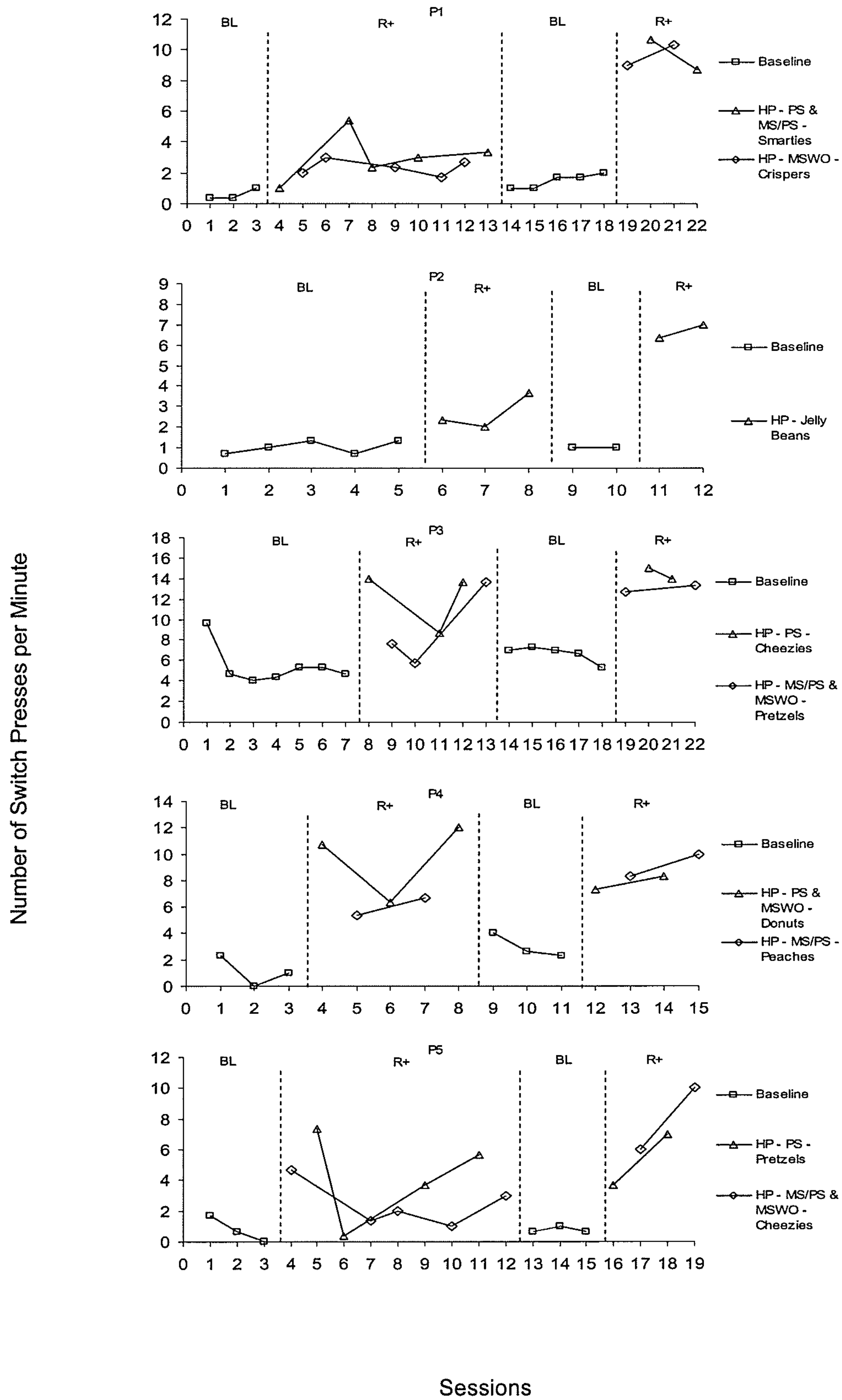


*Figure 2.* Number of switch presses per minute for each participant in baseline and reinforcement sessions. The stimuli evaluated are indicated in the legend of each graph.

chance, therefore, could range from –1 (maximum deviation possible below chance) to +1 (maximum deviation possible above chance), with 0 being equal to chance.

Figure 3 shows the proportion of deviation from chance for the most-preferred stimulus selected across trials with different number of choices in the MSWO assessment for each participant. All five participants appeared to be able to select their most-preferred stimulus well above chance on at least one of the multiple stimulus options with 4 or more choices. For example, P1 selected the most-preferred stimulus at +1 (maximum deviation above chance) on 6-choice trials; P2 showed a deviation from chance of +.62 and +1 on 8-choice and 5-choice trials, respectively; P3 showed a deviation from chance of +.60 on 6-choice trials; P4 showed a deviation from chance of +1 on 7-choice trials; and lastly, P5 showed a deviation from chance of +1 on 6-choice trials.

## Discussion

Although both the MSWO and MS/PS procedures identified the same most-preferred stimulus as the PS procedure for only two of five participants, preference rankings of stimuli were correlated positively with the PS assessments. In addition, the most-preferred stimuli selected in both procedures, even in cases where they were selected on less than 80% of the trials, functioned as positive reinforcers for all participants. This finding is consistent with previous reports that less preferred stimuli may function as positive reinforcers (e.g., Lee et al., in press; Roscoe, Iwata & Kahng, 1999; Taravella, Lerman, Contrucci, & Roane, 2000).

The MS/PS procedure was created in anticipation of participants not being able to select the most-preferred stimulus on trials involving more than 4 stimuli. I hypothesized that such individuals would benefit from this procedure over the MSWO procedure.

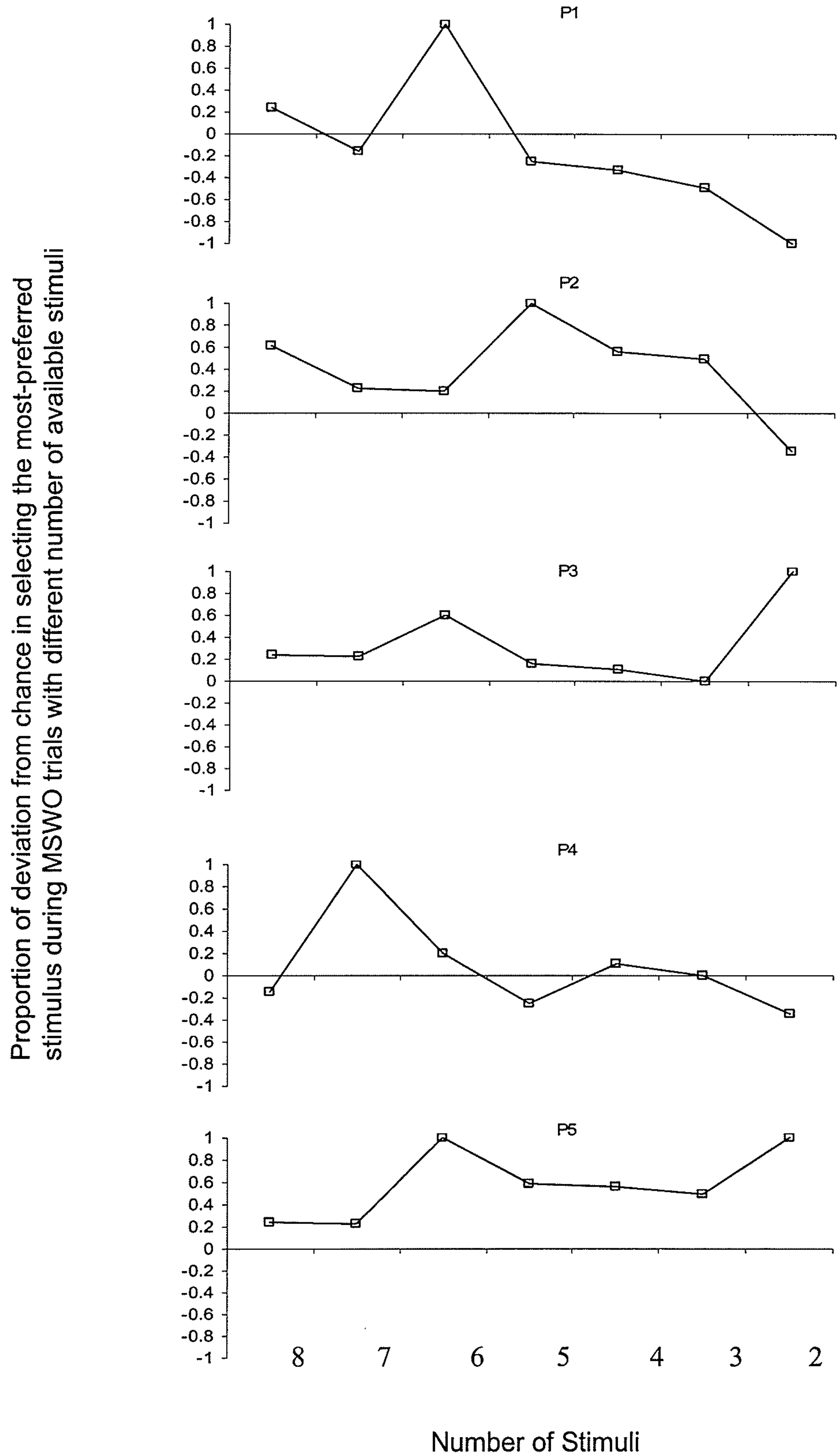


*Figure 3.* Proportion of deviation from chance in selecting the most-preferred stimulus during MSWO trials with different number of available stimuli.

However, the results in Figure 3 suggest that all participants were able to discriminate on trials with 5 or more choices. If so, the benefits of the MS/PS procedure would have been limited for these participants. Additional participants with less proficient discrimination skills (i.e., able to discriminate between 4 or fewer stimuli simultaneously) may provide a stronger evaluation of the effects of MS/PS procedure.

Perhaps not surprisingly, participants' performances on the ABLA did not predict performance on discriminations involving more than two stimuli. The ABLA test measures a person's ability to *learn* to make certain two choice discriminations over a number of trials which include standard prompting, reinforcement and correction procedures. Multiple stimulus preference assessments do not involve standard prompting or correction procedures, but instead test a person's ability to scan an array of stimuli and select their most preferred item. Therefore, a different kind of test is needed to predict performance on multiple stimulus assessments. Research is needed to develop a practical and yet effective procedure to predict the ability to perform discrimination based on the number of available stimuli.

In summary, for individuals who have demonstrated that they can discriminate multiple stimuli at least sometimes, all three procedures performed equally well on identifying positive reinforcers. However, the MSWO procedure was the most efficient, requiring the fewest trials to complete the assessment, and probably the simplest to administer among the three procedures.

# References


Bruininks, R. H., Woodcock, R. W., Weatherman, R. F., Hill, B. K. (1996). *Scales of Independent Behavior-Revised Comprehensive Manual.* Itasca,IL: Riverside Publishing.

Carr, J. E., Nicolson, A. C., & Higbee, T. S. (2000). Evaluation of a brief multiple-stimulus preference assessment in a naturalistic context. *Journal of Applied Behavior Analysis, 33,* 353-357.

Davies, G., Thiessen, C., Yu, C.T., Martin, T., & Martin, G. (2009). Evaluation of multiple-stimulus preference assessment with adults with developmental disabilities. *American Journal on Intellectual and Developmental Disabilities.* Manuscript submitted for publication.

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29,* 519-533.

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

Fisher, W. W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25,* 491-498.

Graff, R.B., & Ciccone, F.J. (2002). A *post hoc* analysis of multiple-stimulus preference assessment results. *Behavioral Interventions, 17,* 85-92.

Hagopian, L.P., Long, E. S., & Rush, K.E. (2004). Preference assessment procedures for individuals with developmental disabilities. *Behavior Modification, 28,* 668-677.

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

Lee, M. S. H., Yu, C. T., Martin, T. L., & Martin, G. L. (in press). On the relation between reinforcer efficacy and preference. *Journal of Applied Behavior Analysis.*

Martin, G., & Pear, J. (2007). *Behavior modification: What it is and how to do it* ($8^{th}$ ed.). Upper Saddle River, NJ: Prentice Hall.

Martin, T. L., Yu, C. T., Martin, G. L., & Fazzio, D. (2006). On choice, preference, and preference for choice. *The Behavior Analyst Today, 7(2),* 234-241.

Pace, G. M., Ivancic, M. T., Edwards, G. L., Iwata, B. A., & Page, T. A. (1985). Assessment of preference and reinforcer value with profoundly retarded individuals. *Journal of Applied Behavior Analysis, 18,* 249-255.

Roscoe, E. M., Iwata, B. A., & Kahng, S. W. (1999). Relative versus absolute reinforcement effects: Implications for preference assessments. *Journal of Applied Behavior Analysis, 32,* 479-493.

Taravella, C. C., Lerman, D. C., Contrucci, S. A., & Roane, H. S. (2000). Further evaluation of low-ranked items in stimulus-choice preference assessments. *Journal of Applied Behavior Analysis, 33*, 105-108.

Vause, T., Yu, C.T., & Martin, G.L. (2007). The Assessment of Basic Learning Abilities test for persons with intellectual disability: A valuable clinical tool. *Journal of Applied Research in Intellectual Disabilities, 20*, 483-489.

Windsor, J., Piche, L. M., & Locke, P. A. (1994). Preference testing: A comparison of two presentation methods. *Research in Developmental Disabilities, 15,* 439-455.