The Clinical Utility of Two Reinforcement Preference Assessment Techniques: A Comparison of Duration of Assessment and Identification of Functional Reinforcers

Traci Lanner, Brandon Nichols, Sean Field, Jennifer Hanson & Thomas Zane

Abstract

In order to maximize the learning of skills, it is crucial that the most powerful reinforcement be used. Research describes several different methods of reinforcement preference assessments, including forced choice, free-operant, and multiple stimulus array procedures. Researchers have also tested many variables to determine their potential impact on reinforcer identification including setting events, duration of exposure to tested stimuli, and differing schedules of delivery. From a clinical perspective, a major question is, what is the most efficient preference assessment that will produce the most valid identification of potential reinforcers? The purpose of this study was to compare multiple stimulus and forced-choice formats along two dimensions-duration to complete assessment and the identification of stimuli verified to function as reinforcers. Individuals diagnosed with developmental disabilities/autism served as participants. Experimenters conducted two types of preference assessments per participant – multiple stimulus array without replacement, and forced-choice, using the same items. Once the most preferred stimuli were identified in each format, experimenters then tested the reinforcing power by making each contingent upon a task that the participants had not yet learned. The study found that the forced choice method was faster to administer and that all tested items actually functioned as reinforcers.
Keywords: autism; preference assessment; reinforcement

Educators have relied on the use of positive reinforcement techniques for many years to modify human behavior (e.g., Cooper, Heron, & Heward, 2007). One particular challenge in working with individuals who exhibit developmental disabilities such as autism is that of selecting effective reinforcers. Caregiver interviews are frequently utilized but are not necessarily accurate predictors of reinforcers (e.g., Green, Reid, Canipe, & Gardner, 1991). As an alternative, Pace, Martin, Ivancic, Edwards, Iwata, and Page (1985) described a systematic preference assessment. Participants were presented with various stimuli, one at a time, and their approach or non-approach of the item was measured. Items which were approached were found to be more reinforcing than non-approached items when utilized in a behavior change program. Since then, several variations have been developed, including forced choice/paired stimulus (FC/PS; Fisher, Piazza, Bowman, Hagopian, Owens, & Slevin, 1992; Mason, McGee, Farmer-Dougan, & Risley, 1989), multiple stimulus presentations with replacement (MSW; Windsor, Piche, & Locke (1994), and the Multiple Stimulus Without Replacement (MSWO; DeLeon & Iwata 1996). In the PS approach, two potentially reinforcing items are presented at the same time and the participant is asked to "pick one." All items are compared to one another and their position on the table is controlled. In the MSW approach, all items are available to the participant at the same time. Once a selection is made and the participant accesses the item, it is placed back in the lineup for a second selection. Position in the line is controlled for by rotating the items after each selection. The MSWO procedure is similar except that once an item has been chosen, it is removed from the lineup. It is assumed that those items which are chosen first will function more effectively as reinforcers.

In addition to being able to accurately identify reinforcing items, it is critical to applied clinicians that preference assessment techniques are able to be carried out in the most efficient manner, with regard to duration of procedure. In order to better serve clients, it would be helpful to know which preference assessment techniques yield the most accurate prediction of reinforcers in the least amount of administration time. DeLeon and Iwata (1996) explored this topic by comparing the paired stimulus (PS) presentation, with the multiple stimuli with and without replacement presentations (MSW and MSWO).

They found that the MSW was the fastest to administer but that the MSWO and the PS "identified more stimuli that are at least minimally reinforcing than does the MS procedure" (p. 530). Thus, the purpose of the current research was to compare administration time of the MSWO and the PS assessments and their ability to accurately identify reinforcers given a sorting task.

Method

Participants and Setting

Four individuals with a primary diagnosis of autism and varying levels of mental retardation served as participants. They lived in a private residential facility for individuals with developmental disabilities. These four males (ranging in ages, 14-20 years) followed simple one-step directions and were generally compliant throughout the study. One participant communicated using vocal speech, while the other three communicated via simple gestures and limited use of picture symbols. All participants required assistance with daily living skills such as tooth brushing and bathing. They were all able to feed themselves if food was cut up to the typical size.

The study was conducted at the residential school facility that the participants attended. Sessions were held in the school building in a room measuring 21' x 12'. This room had one window, eight chairs, and a large table.

Phase 1

Purpose and dependent measure

The purpose of phase 1 was to conduct preference assessments to determine length of time to complete and the relative rankings obtained. Each participant completed both paired choice (PC) and multiple stimuli without replacement (MSWO) systematic preference assessments. The dependent variables were the ranking of the potential reinforcers and the duration of time to administer each assessment.

Materials

Prior to conducting the preference assessment, staff from the classroom and residence provided a list of the top five items they thought were preferred by the individual. The researchers cross-referenced the lists from each location and in one case, the testers randomly determined a fifth item, as no staff were able to produce a total of five for one individual. The investigators reviewed discrepancies and determined which items were to be used (see Table 1 for the items chosen). The researchers presented all edible items as portions smaller than a half-inch in diameter in a small paper bowl. They also presented all liquids in an 8oz plastic cup and were measured out to a predetermined level. The investigators presented all tangible items (e.g., book, DVD, etc.) directly on the table. When a participant selected the DVD or CD, the researchers placed it in a portable DVD or CD player and started it for the participants, allowing them to watch or hear it for the 30 seconds.

Table 1

Items used as stimuli on preference assessments

Subject	1	2	3	4	5
Matty	Book	Apple	Cookie	CD	Movie
Jason	Keyboard	Chocolate	Lotion	Book	Pickle
Cody	Movie	CD	Chocolate	Cracker	Water
David	Doritos	Funyons	Rice Cake	Cheetos	Cookie

The experimenters used two preference assessment strategies reported in the previous literature. The PS and MSWO were administered in random order. For each participant, the researchers presented a total of 5 MSWO sessions and one PS consisting of 20 trials. The 20 trial PS was only administered once in comparison to the 5 MSWO sessions to ensure an equal number of item presentations across the two conditions.

Procedure

MSWO

For each Multiple Stimulus Without Replacement (MSWO) session, the experimenter brought one participant into the testing room at a time and asked him to sit down across the table from the investigator. The tester then placed one of each of the five items on the table in a row parallel to the participant, randomly ordered and one half inch apart from the next. The tester then directed the participant to "pick one." The participant had 5 seconds to respond or a "no response" was recorded. When the participant touched or picked up an item, he had 15 seconds to consume edible items or 30 seconds to engage with the leisure item. While the participant was engaged with the item, the tester shifted the row so that the item farthest to the right of the participant was moved to the other end of the row and other items were shifted one location to the right, removing from view the selected item once the participant was finished with it. Once the row had been shifted, the tester scored which item was selected. Following the reinforcement period, the tester again repeated the procedure above until all of the items had been selected or there had been no response from the participant for 3 consecutive directions.

PS

The participant sat at the table across from the tester who then presented two of the items, while varying the order of presentation. Over the course of the assessment, the tester counter-balanced the placement of the items to ensure that each item had the same number of presentations with all other stimuli and was presented on the participant's left and right side an equal number of times. The researcher placed the items six inches apart from each other on the table and then asked the participant to "pick one". Following a response, the participant had 15 seconds to consume an edible or 30 seconds to engage with a leisure item. Immediately following the participants' response, the tester removed the item that was not selected. While the participant was engaging with the selected item, the tester scored which response was made. After the designated time, the experimenter told the participant, "time's up," and presented the next pair of items. This was repeated until all 20 presentations were completed. If no choice was made after 5 seconds, the trial ended.

Results

The paired stimulus preference assessment was shorter in duration than the MSWO for all participants. Durations for the paired stimulus ranged from 287s to 621s (M = 454.5 s). For the MSWO procedure, total duration for all 5 sessions ranged from 337s to 742s (M = 621.5s; see Figure 1).

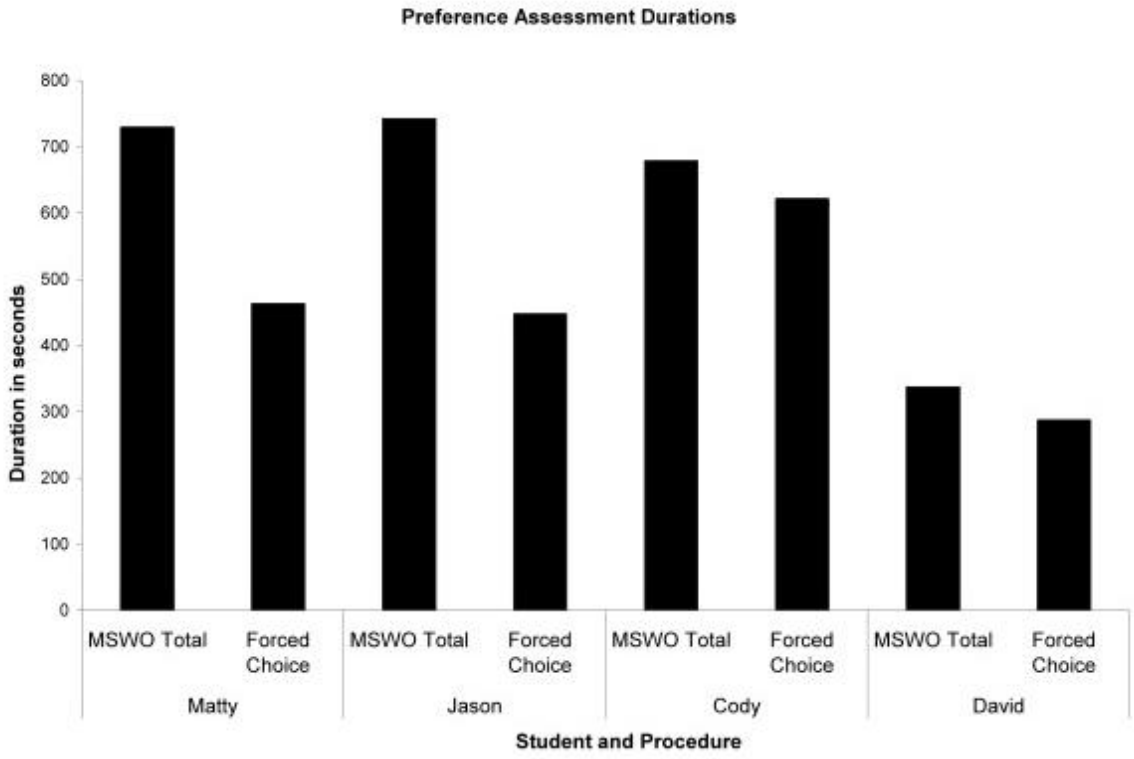

Figure 1. Duration in seconds to administer preference assessments across students.

The results of both preference assessments showed a great deal of variability in the rankings (see Table 2). For three of the four participants, the highest and lowest ranked items were the exact same in both assessments. With only one participant were the items ranked in the same order on both assessments.

Table 2

Ranking of items on preference assessments

	MSWO RANK	PS RANK
Cody	Chocolate	Chocolate
	Cheezits	Movie/Music
	Music	
	Movie	Cheezits
	Water	Water
Jason	Chocolate	Pickles
	Pickles	Chocolate
	Book	Lotion
	Lotion	Book/Keyboard
	Keyboard	
David	Cookie	Cookie
	Cheetos	Cheetos
	Rice Cake	RiceCake
	Funyons	Funyons
	Doritos	Doritos
Matty	Cookie	Cookie
	Apple	Apple
	Movie/Music	DVD/Book
	Book	Music

Phase II

Purpose and Dependent Measure

The purpose of this phase was to use the stimuli identified in the preference assessment in a test of reinforcer functioning – that is, the researchers assessed the extent to which the stimuli identified actually functioned as reinforcers by increasing the strength of responses that they followed. Following a baseline period, each participant completed a sorting task, where the tester contingently presented each preference assessment stimulus for a correct response. The dependent measure was the rate of sorting envelopes per minute.

Materials

The investigators used all materials (Table 1) and the same room for testing in this phase. Additional materials present during each session were three 8 1/2 " X 11" plastic trays, 33 letterhead envelopes, 21 magazines, and 33 manila envelopes. The experimenters arranged these trays parallel to one another and 1.5 feet from where the participant sat during the session. Directly in front of the

participant was a stack of all the envelopes and magazines excluding three (one of each kind), which served as labels for the correct tray for the sorting task. The testers timed each session with a stopwatch and used a clock on the room wall to track the time allotted for accessing reinforcement.

Procedure

Baseline

During baseline conditions, the experimenter brought a single participant into the testing room and sat at the end of the table in front of the stack of envelopes and magazines. The tester sat directly to the participant's left and told him to "sort." Following the instruction, the tester started the stopwatch. The task involved sorting 3 styles of envelope/brochures into the appropriate trays. Incorrect responses were not corrected, however, following 30 seconds of no correct responses, the tester modeled a response in which he/she sorted one of each of the items into their appropriate trays. Following the model the tester again stated "sort", and this procedure was repeated until either at least one correct response had been made or the session ended. Following a correct response, the tester remained silent for the remainder of the session. All baseline sessions were five minutes in duration and when the session was finished, the experimenter counted the number of envelopes or magazines in their appropriate trays, minus any modeled responses. Then s/he calculated a rate of response per minute by dividing the number of correct responses by duration in minutes.

Test of effectiveness of stimuli as reinforcer

The experimenter conducted these sessions exactly the same as the baseline sessions, with the following changes. Following a correct response, the experimenter gave the participant the preferred item prescribed for that condition (edible items for 10 seconds, leisure items for 15 seconds). During the reinforcement period, the experimenter stopped the stopwatch recording the five minute session duration to ensure that time spent during reinforcement was not recorded as part of the free operant sorting time. Following the reinforcement period, the tester told the participant "time's up" and the item was removed from his/her possession. The tester then restarted the hand timer. The experimenter calculated rate in each experimental session. If a participant sorted all of the available materials, the tester stopped the time and the number of correct responses was recorded. S/he then replaced the materials and the session continued by re-starting the timer.

Return to Baseline

After stability was achieved during the phase in which reinforcement effectiveness was tested, there was a return to baseline phase, utilizing the exact procedures described above in baseline.

Reliability

Interobserver agreement (IOA) was assessed for each phase using a second observer. For IOA of phase one, the second observer independently ranked items on the preference assessment and compared to the testers on 25% of all sessions. The number of agreements was divided by the total number of disagreements plus agreements multiplied by 100%. Agreement was 100%. For IOA of phase two, the second observer independently counted the total number of envelopes sorted and compared them to the tester's total number and rate of sorting on a total of 20% of the sessions. Again, there was perfect agreement.

Procedural reliability was also assessed for each phase. During phase one, the second observer used a 13-step checklist to independently note whether each step of the protocol was correctly followed by the tester. The second observer then calculated a percentage of the protocol steps conducted appropriately. This was completed for 25% of sessions and the mean was 93% (range 85-100%). In

phase two, procedural reliability was calculated in the same manner described above on 20% of the total sessions. The reliability observer also used a second stopwatch to note whether all time limits were observed per protocol. The mean procedural reliability was 98% (range 83-100%).

Results

During the initial baseline phase, rates of responding were very low for all participants (see Figures 2-5). The average rate of sorting responses during baseline across all participants was 0.15 items per minute (range, 0 to 0.3).

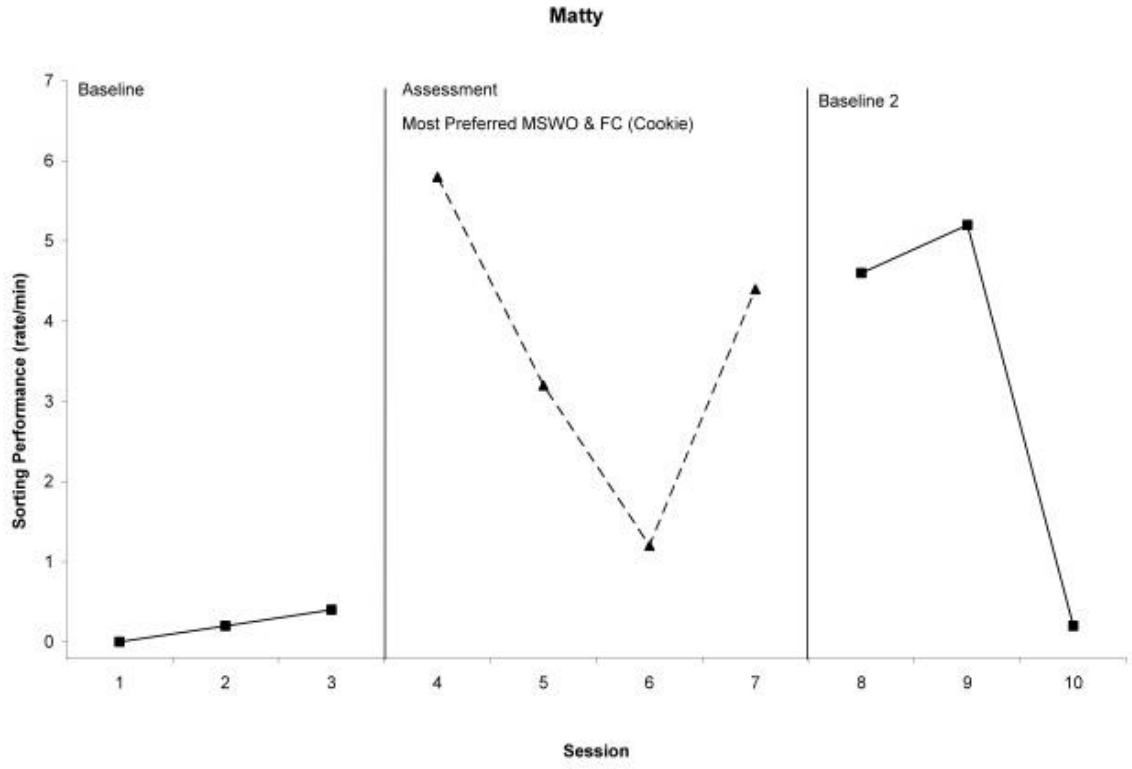

Figure 2 Response rate per minute during the sorting task under baseline and assessment conditions. Assessment condition shows the results of the most preferred item(s) as determined by the MSWO and FC preference assessment.

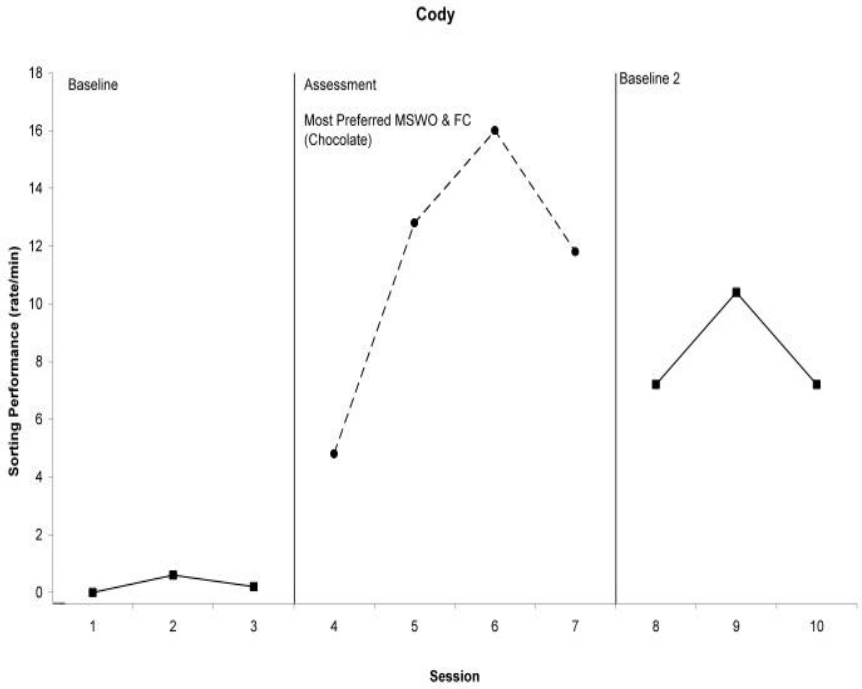

Figure 3 Response rate per minute during the sorting task under baseline and assessment conditions. Assessment condition shows the results of the most preferred item(s) as determined by the MSWO and FC preference assessment.

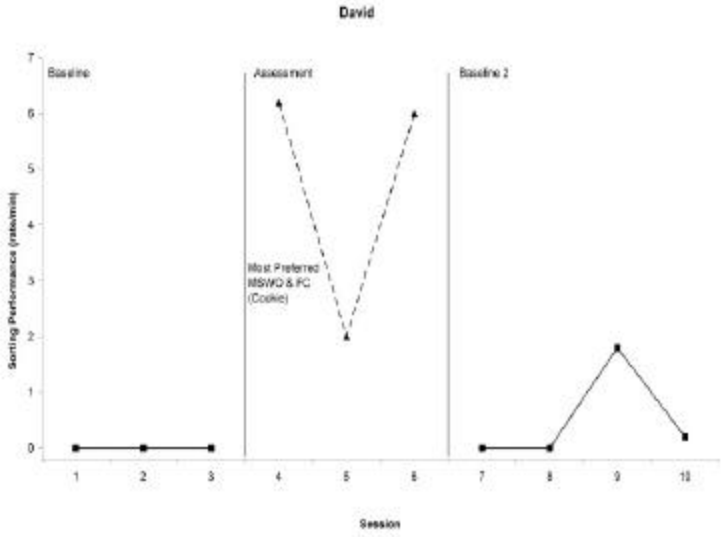

Figure 4 Response rate per minute during the sorting task, under baseline and assessment conditions. Assessment condition shows the results of the most preferred item(s) as determined by the MSWO and FC preference assessment.

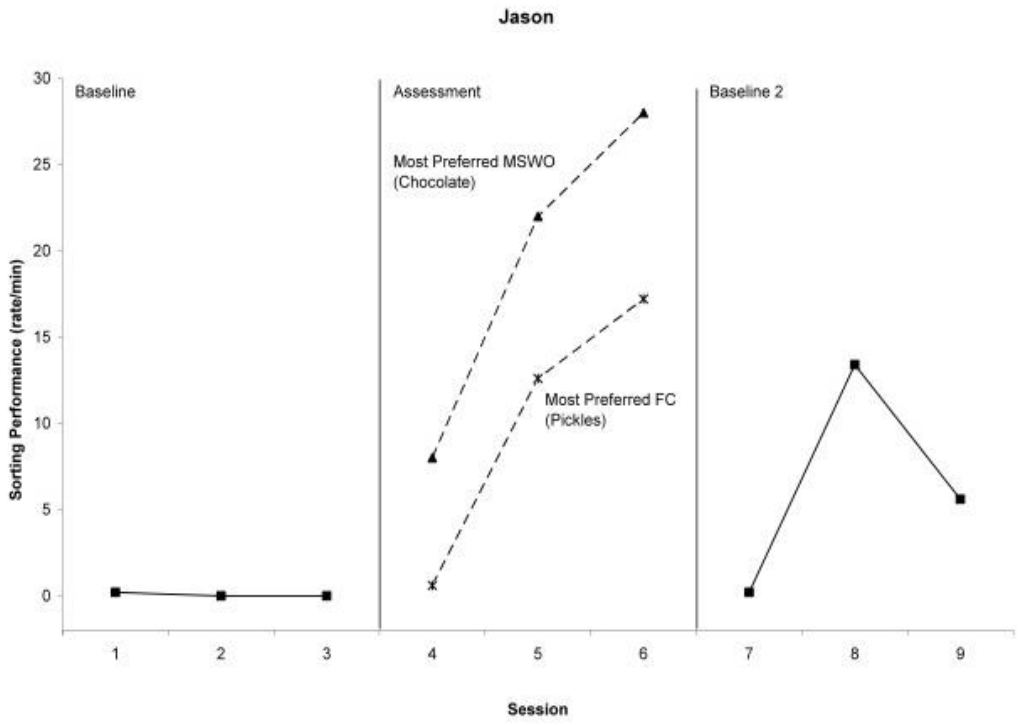

Figure 5 Response rate per minute during the sorting task, under baseline and assessment conditions. Assessment condition shows the results of the most preferred item(s) as determined by the MSWO and FC preference assessment.

Once the selected stimuli were made contingent upon the sorting response, rates of responding across all participants increased. The same pattern was observed - each item tested functioned as a reinforcing stimulus by increasing the rate of sorting over baseline conditions.

For three of the four participants, the two preference assessment strategies produced the same stimulus as highest rated. For Matty, both preference assessments indicated that cookie was his most preferred item. Using this item contingently produced higher rates of sorting than without it, as evidenced by the return to baseline levels in the Baseline 2 condition. Cody's most preferred item, as determined by both preference assessments, was chocolate. When presented contingently upon sorting, a clear gain was seen in this response. When Baseline 2 conditions were implemented, the rate of sorting envelopes decreased. During the preference assessment, David selected cookie more than any other item. David increased the rate of sorting during the experimental phase when cookie was made contingent, and approached levels of the original baseline condition in the Baseline 2 condition.

However, Jason's results were different, in that the two preference assessments produced different stimuli that were ranked the highest. The PS method showed pickles to be the most preferred item, while chocolate was found to be most preferred in the MSWO procedures. Nevertheless, using the two different stimuli produced increases in sorting over baseline procedures. The return to baseline conditions (in Baseline 2) increases support of the hypothesis that the two tested stimuli functioned as positive reinforcers.

Discussion

This research compared two preference assessments, Multiple Stimulus without Replacement and Paired Stimulus, in terms of duration to administer and accuracy in predicting and ranking reinforcers. Five stimuli were evaluated for each participant. After completion of both preference assessments, the stimuli were tested as to their functionality as reinforcement during a free operant sorting task. The Paired Stimulus Preference Assessment was found to have a shorter duration to administer than the MSWO assessment. For three out of the four participants, the highest and lowest ranked items were the same on both preference assessment formats. All stimuli, implemented contingently on the sorting task, increased the participants' response rates above baseline.

DeLeon and Iwata (1996) found that the duration of the MSWO procedure was less than that of the Paired Stimulus (PS) preference assessment. However, the current study found the PS assessment to be less time consuming. This may be due to the unequal amount of trials performed in each preference assessment in the current research. The MSWO contained 25 trials while the PS contained 20 trials, as previously explained. A future question would be whether MSWO, conducted in fewer trials to shorten the duration, would still predict stimuli that will actually function as positive reinforcers.

The current study replicated Pace et al.'s (1985) findings of preferred stimuli increasing the rate of behavior. All current participants performed above baseline when an F1 reinforcement schedule was implemented. Knowing that both preference assessments predicted the same stimulus as highest ranked for three of the four participants suggests that either method would be appropriate to use when assessing an individual's preferences. Furthermore, each top-ranked stimulus functioned as a contingent reinforcer; that is, when used contingently upon a sorting response, it increased the rate of sorting.

The generality of these findings may be compromised due to potential confounding variables or limitations to the study. One limitation resulted from Matty's baseline not being continued until stable. The intervention was implemented after 3 data points, though baseline should have been continued due to the increasing trend. It is possible that his baseline performance would have been higher had the condition continued. This would impact the comparison of his results in intervention to the baseline condition. The researchers also did not take into account the time of day the sessions were held. Setting events such as meal or leisure times (free access to preferred stimuli) may have influenced the participants' motivations to work for reinforcement. Another limitation of this study was that preferences of the participants were based on staff observation. Though all stimuli did function as reinforcers, it is not clear as to whether other stimuli may have functioned as stronger reinforcers. Future research should include a control item to provide comparison as to whether it is any contingency that reinforces the behavior or if it is the result of the chosen "preferred" items.

In the future, researchers should include more duration comparisons of varying preference assessments that present the same number of trials. The present research differed from DeLeon and Iwata's findings for duration of assessments. Assessment of potentially influential variables should also be analyzed more clearly as well. These variables may include type of task, task difficulty, setting events, and stimuli chosen. Leisure time (free access to preferred stimuli) and meal times may affect the participants' motivation for reinforcement. Future research should also explore the relationship between rankings of reinforcing stimuli and task difficulty levels. That is, would a highly ranked stimulus functioning as a positive reinforcement effectively alter a behavior that is considered to be highly aversive? Another possible question for future research relates to the generality of the current findings with regards to stimuli selected by means other than formal preference assessments. Would stimuli selected by staff or caregiver report, or that participants engaged with frequently in a free operant context, be shown to function as reinforcers in the reinforcement test? This would be a different question than addressed in the current study, in which the stimuli tested were a result of formal preference assessment strategies.

In sum, these findings suggest that there may possibly be time savings based upon the type of preference assessment that is conducted. However, generally speaking, any type of preference assessment can identify stimuli that will actually function as positive reinforcers. From a clinical perspective, one needs a quick but yet accurate assessment of potential reinforcers. The results of this study lend confidence to the idea that genuine reinforcers can be obtained through this process and that they can be tested quickly to determine what will function as contingent motivators.

References

Cooper, J.O, Heron, T.E, & Heward, W.L. (2007). Applied Behavior Analysis. Upper Saddle River, NJ: Merril.

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

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

Green, C.W., Reid D. H., Canpie, V.S., & Gardner, S.M. (1991). A comprehensive evaluation of reinforcer identification processes for persons with profound multiple handicaps. *Journal of Applied Behavior Analysis, 24,* 537-552.

Mason, S.A., McGee, G.C. , Farmer Dougan, V., & Risley, T.R. (1989). A practical strategy for ongoing reinforcer assessment, *Journal of Applied Behavior Analysis, 22,* 171-179.

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

Windsor, J., Piche, L.M., & Locke, P.A. (1994). Preference testing: A comparison of two presentation methods.

Contact Information:

Traci Lanner
Springbrook
2705 State Highway 28
Oneonta, NY 13820

Thomas Zane
Center for Applied Behavior Analysis