Graduate Theses and Dissertations   Graduate School

6-20-2003

Evaluation Of A Presentation And Measurement Method For Assessing Activity Preference

Tara L. Lieblein
University of South Florida

Follow this and additional works at: https://scholarcommons.usf.edu/etd

Part of the American Studies Commons

Evaluation Of A Presentation And Measurement Method For Assessing Activity

Preference

by

Tara L. Lieblein

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Arts
Applied Behavior Analysis Program
College of Graduate Studies
University of South Florida

Major Professor: Pamela G. Osnes, Ph.D.
Glen Dunlap, Ph.D.
Iser G. DeLeon, Ph.D.

Date of Approval:
June 20, 2003

Table of Contents

List of Tables

List of Figures

An Empirical Assessment of Activity Preference: A Presentation and
Measurement Method for Assessing Preference of Non-tangible Stimuli

Tara L. Lieblein

ABSTRACT

Much research has focused on the development of methods of measuring
preference for stimuli. These methods have shown to be an accurate and valid
way to identify potential reinforcers. However, these methods have only been
conducted with tangible stimuli and have not been extended to non-tangible
stimuli or activities, potentially because these types of stimuli are not appropriate
for current preference assessment presentation methodologies. This study used
a single stimulus presentation preference assessment to identify preferred
activities for two adults with developmental disabilities. Two measures (duration
of engagement and indices of happiness) were collected to identify preferred
stimuli. For both participants, there were differences in happiness measure
between activities. The engagement measure only produced differentiated
results for one participant. Reinforcer assessments were conducted to determine
if the measures of preference were able to identify high preference stimuli that
functioned as more effective reinforcers more than stimuli identified as low
preference. Both participants exhibited high rates of an arbitrary response during
all conditions of the reinforcer assessment. Therefore, the reinforcer
assessments did not validate the results of the preference assessments.

Chapter One

Literature Review

Reinforcement is one of the most fundamental concepts in the science of human behavior (Skinner, 1953) that we know as behavior analysis. It is also a vital part of the function-based treatment of behavior disorders, especially among individuals with developmental disabilities. The use of reinforcement as a component of interventions for problem behaviors is frequently used because it is relatively easy and effective (Ivancic, 2000) and not as intrusive as punishment. It is vital to use stimuli that are highly preferred and function as reinforcers if these interventions are to be effective. As a result, much research has focused on developing technologies for the assessment of preference of stimuli and the validity of these assessments in identifying reinforcers. Additionally, with the growing focus on person-centered planning and improving quality of life for individuals with disabilities (Koegel, Koegel, & Dunlap, 1996) in applied behavior analysis, it is important for practitioners to have valid and reliable ways to assess preference for a variety of stimuli that can be incorporated into an individual's daily schedule. A systematic approach typically involves a stimulus preference assessment and a reinforcer assessment. A stimulus preference assessment attempts to assess preference and predict the reinforcing effects of stimuli while

a reinforcer assessment tests the reinforcing effects of stimuli by measuring

changes in rate of behavior (Fisher, Piazza, Bowman, & Amari, 1996).

Verbal Report and Surveys

The use of surveys to identify potential reinforcers is a quick and easy

method that is commonly used. Most surveys require caregivers to report on

preferred items for people with developmental disabilities and limited

communication. However, the little research that has been done to evaluate the

validity of caregiver verbal-report (Green et al., 1988; Green et al., 1991) and

self-report methods (Northup, George, Jones, Broussard, & Vollmer, 1996) of

identifying reinforcers have not shown significant correspondence with

systematic assessments conducted with the same stimuli. Northup et al. (1996)

found that a modified child reinforcement survey did not result in accurately

identifying reinforcers compared to a reinforcer assessment when used with

verbal children. Northup (2000) replicated this study with a larger sample of

individuals with the same results. The Green et al. studies (1988, 1991) showed

that caregiver reports did not accurately predict preferences for children with

multiple severe handicaps. In another study with nonverbal children, a structured

interview called the Reinforcer Assessment for Individuals with Severe

Disabilities (RAISD) identified reinforcers more effectively than a less structured

list when used in combination with a preference assessment (Fisher, Piazza,

Bowman, & Amari, 1996). In the RAISD, caregivers identify preferred stimuli for

individuals with developmental disabilities in several categories (e.g., visual,

edible, auditory, etc.) that are assumed to facilitate the process by providing

prompts to the reporter. The interview also requires the caregiver to rank the stimuli according to preference. However, this interview was designed to supplement a paired preference assessment (a systematic assessment that requires the individual to choose an item from various pairs of items presented) in order to provide accurate results and was not intended to be used individually. Therefore, it does not reduce the time or effort required to accurately identify reinforcers.

Systematic Preference Assessments

Much research has explored various ways to empirically assess stimulus preference in individuals with developmental disabilities who have limited communication. Research on stimulus preference assessment methodology typically falls into one of three general categories: single stimulus presentation (items are presented individually and responses to items scored one at a time over several trials), multiple stimulus presentation (items are presented all at once and which items are responded to is recorded), and paired stimulus, or choice, presentation (items presented in pairs and items chosen are recorded) (Ivancic, 2000). Each method comes with its own benefits and limitations that will receive further elaboration. From an applied perspective, an important consideration in using stimulus preference assessments is ease of use because preferences can vary across time (Green, Reid, Canipe, & Gardner, 1991; Zhou, Iwata, Goff, & Shore, 2001) resulting in a need for assessments to be conducted frequently (DeLeon, Fisher, Rodriguez-Catter, Maglieri, Herman, & Marhefka, 2001). Recently, the research on preference assessments has focused on

developing methods that are more time-efficient while still accurate in identifying preferred stimuli and potential reinforcers (Carr, Nicholson, & Higbee, 2000; DeLeon & Iwata, 1996; Roane, Vollmer, Ringdahl, & Marcus, 1998).

Single Stimulus Presentation. Single stimulus (SS) presentation preference assessments (Pace, Ivancic, Edwards, Iwata, & Page, 1985; Green, Reid, Canipe, & Gardner, 1991) involve presenting one stimulus at a time to the individual for a brief period of time. Preference is assessed by comparing approach responding to each of the items (Pace et al., 1985) or duration of interaction (Hagopian, Rush, Lewin, & Long, 2001) over several trials. Percent of trials with approach responding or percent duration of interaction is calculated for each item. This type of presentation is easy to implement and time-efficient. Additionally, single stimulus presentation may be more appropriate to use with individuals who have difficulty making choices, a skill that is required in multiple and paired stimulus presentation methods. However, single presentation may provide an overestimate of stimulus preference (Fisher, Piazza, Bowman, Hagopian, Owens, & Slevin, 1992) because some individuals may interact with any item placed in front of them, even if it is not preferred.

Multiple Stimulus Presentation. A multiple stimulus (MS) presentation (Windsor, Piche, & Locke, 1994; DeLeon & Iwata, 1996) involves presenting an assortment of items at the same time and allowing the individual to choose one item from the array. The earliest version of this method involved replacing an item after it was chosen so that all items are present for each trial (Windsor et al., 1994). In this study, this method identified preferences but not as consistently as

3

a paired stimulus (PS) assessment. Items identified by the PS assessment as preferred were not identified in the MS assessment. This may have been due in part to the continuous presence of the most highly preferred items in the MS assessment (DeLeon & Iwata, 1996). However, the MS assessment did take less time to complete than the paired stimulus assessment.

In an effort to develop a method that incorporated the benefits of both methods, DeLeon and Iwata (1996) used a multiple stimulus without replacement (MSWO) method to assess preferences of stimuli. An array of items was presented simultaneously as in the MS assessment. However, after an item was chosen, it was removed from the group and the individual chose another item. This continued until all items were selected or the individual no longer selected an item. This process was repeated several times and a selection percentage was calculated by dividing the number of times an item was chosen by the number of trials it was available. These percentages were then used to rank-order the items by preference. This method requires the individual to discriminate between stimuli, which may yield more differential responding and give a more sensitive measure of preference than the single stimulus presentation. However, the multiple presentation method sometimes offers undifferentiated results when data are collected on approach responding.   More effective results using duration-based data were produced for the same stimuli used when presented in a single stimulus format. (DeLeon, Iwata, Conners, & Wallace, 1999).

Paired Stimulus Presentation. Finally, when using a paired (or choice) stimulus (PS) presentation (Fisher et al. 1992; Piazza, Fisher, Hagopian, Bowman, & Toole, 1996), stimuli are presented in pairs and the individual chooses one. Each stimulus is presented with every other stimulus and data are collected on approach responding. A preference hierarchy is developed based on the percent of approach responding for each item. This method seems to be the most accurate method for predicting reinforcer effectiveness but is the most time-consuming, taking more than twice the amount of time to administer than a multiple stimulus presentation format (DeLeon & Iwata, 1996). Several variations of the PS methodology have been examined, such as modification for individuals with visual impairments (Paclawskyj & Vollmer, 1995) and the use of verbal and pictorial representations of stimuli (Cohen-Almeida, Graff, & Ahearn, 2000; Northup, George, Jones, Broussard, & Vollmer, 1996; Parsons, Harper, Jensen, & Reid, 1997). The results of these studies will be discussed in more detail below.

Alternative Presentations. The free-operant method of assessing stimulus preference has also been investigated (Ringdahl, Vollmer, Marcus, & Roane, 1997; Roane, Vollmer, Ringdahl, & Marcus, 1998). This presentation method is similar to the MS assessment in that an array of stimuli is present throughout the assessment. However, in the free-operant method, the individual is allowed to freely interact with any stimulus for any amount of time. The stimuli are neither repeatedly presented nor removed and the individual is not required to choose any item. Duration of interaction with each stimulus serves as a measurement of

preference. This method may be of benefit for individuals who exhibit problem behaviors to gain access to tangible items and who may engage in those behaviors during a SS, MS, or PS assessment.

As stated previously, several studies have examined the utility of using verbal or pictorial representations of stimuli instead of actual tangible stimuli. The efficacy of use of pictures in preference assessments has shown limitations. The Northup et al. (1996) study showed that both verbal and pictorial PS assessments accurately identified high- and low-preference items with verbal children diagnosed with ADHD. While this method may be applicable to this population, the same success has not been seen for individuals with developmental disabilities and receptive and expressive language deficits. With the latter population, verbal (Cohen-Almeida et al., 2000) and pictorial (Parsons et al., 1997) PS assessments have not yielded as much differential responding and accurate results as tangible PS assessments conducted with the same stimuli. Additionally, Higbee, Carr, and Harrison (1999) found that pictures used in an MSWO assessment did not identify potent reinforcers as successfully as a tangible MSWO assessment conducted with the same stimuli.

Limitations to Preference Assessment Methodology. There are limitations to the current technology for identifying preferences and reinforcers. Most of the research conducted using these different methods was done using tangible items, such as toys or food. Additionally, the stimuli were delivered contingent on approach responding or the individual was allowed to engage with the stimulus when duration measures were used. This may limit their utility in that they may

6

not be appropriate for assessing preferences for non-tangible stimuli and activities to which immediate access is difficult to provide (i.e., going to the store, riding a bicycle). Finally, because of the lack of support for using pictorial and verbal PS methods with individuals with developmental disabilities and language deficits, it is possible that stimuli that could have functioned as potent reinforcers were not included in preference assessment research because they did not lend themselves to the presentation method.

Using activities as stimuli, Hanley, Iwata, and Lindberg (1999) found that differential responding during a PS assessment only occurred when the individual was given contingent access to the chosen activity and did not occur when the individual was not given access to the chosen activity. This procedure would be quite a lengthy and difficult process to carry out with non-tangible stimuli, such as going to the park. Work task preferences have been identified using a PS assessment (Reid, Parsons, & Green, 1998). However, all work tasks assessed had associated tangible materials that were presented to the participants to represent each task. In a study of instructional activity preference, Foster-Johnson, Ferro, and Dunlap (1994) used a rating scale based on student interaction with the activity, including manipulation of materials, resistance to removal of materials, and movement toward materials when they were moved away. This method was useful in that activities did not have to be presented simultaneously. Also, measures of preference other than approach responding were used. This measurement method could be useful for identifying preference for activities that are difficult to provide direct access to contingent on approach

7

responding. However, this method still requires the manipulation of tangible materials by the experimenter in order to measure preference.

Reinforcer Assessments

After preferred stimuli are identified, a reinforcer assessment is often conducted to evaluate the reinforcing value of a stimulus (Piazza et al., 1996; Hagopian et al., 2001). To evaluate reinforcer effectiveness, Piazza et al. (1996) compared the rates of an unprompted behavior when followed by a high-preference, low-preference, and no stimulus (control) using a multielement design. If a high-preference stimulus serves as a reinforcer for an individual, the rate of the unprompted behavior it follows will be higher than a middle-preference stimulus or no stimulus. DeLeon and Iwata (1996) used a reversal design to measure reinforcement effects of stimuli by delivering stimuli contingent on a target response. Stimuli were identified as reinforcers when contingent delivery of the stimuli produced higher rates of responding than baseline and baseline levels returned during reversal phases. These assessments can be even more time-consuming than a preference assessment because they utilize a single-subject design to measure behavior change (Ivancic, 2000). Therefore, it is important that researchers use reinforcer assessments to determine the validity of the various preference assessment methods (Piazza et al. 1996) so that they may be used efficiently in application and practice without the use of a reinforcer assessment.

Measurement

Preference assessments have typically utilized data collected on two types of responding to measure preference; approach responding (Pace et al., 1985; Fisher et al., 1992) and duration of interaction (DeLeon et al., 1999; Hagopian et al., 2001). While approach responding typically yields differentiated responding during MS assessments, DeLeon et al. (1999) showed that in cases where individuals displayed undifferentiated approach responding, duration measures yielded differentiated results. Hagopian et al. (2001) also showed that duration of engagement was an effective measure of preference in a SS assessment. In a study by Derby et al. (1995), the authors compared the results of a PS assessment using approach responding and latency to the first aberrant response. The study showed that latency to the first aberrant response was effective in predicting reinforcer effectiveness.

Additional research may also present potential alternate measurements of preference. For instance, Green and Reid (1996) operationally defined and measured happiness in individuals with disabilities. They defined happiness as any facial expression or vocalization typically considered to be an indicator of happiness among people without disabilities including smiling, laughing, and yelling while smiling. Unhappiness was defined as any facial expression or vocalization typically considered to be an indicator of unhappiness among people without disabilities such as frowning, grimacing, crying, and yelling without smiling. Results showed that these definitions were successful in identifying indices of happiness across and within participants and that indices of happiness

9

increased in the presence of most preferred stimuli as compared to least preferred stimuli (as determined by a SS preference assessment). Indices of unhappiness increased when least preferred stimuli were presented as compared to most preferred stimuli. Further studies replicated and extended these findings to increase happiness and decrease unhappiness during treatment of individuals with developmental disabilities (Green, Gardner, & Reid, 1997; Green & Reid, 1999).

In an effort to develop a data-based method for prescribing empirically derived treatment packages for severe problem behavior (aggression, self-injury, etc.), Fisher, Piazza, Bowman, Hagopian, and Langdon (1994) conducted a stimulus avoidance assessment based on the SS preference assessment developed by Pace et al. (1985). During the stimulus avoidance assessment frequency of negative vocalizations, avoidant movements, and positive vocalizations were used to measure nonpreference or preference, respectively, for several potential punishment procedures. Results showed that measurement of these target responses accurately predicted the punishing effects of these procedures during a punisher assessment. These same methods were also used to develop empirically derived treatment packages for pica (Fisher, Piazza, Bowman, Kurtz, Sherer, & Lachman, 1994). Although the measurements used in these studies were not used specifically to measure preference during preference assessments, they have been validated as being associated with preferred and aversive stimuli and could logically serve as measurements of preference during preference assessments.

The current state of preference assessment and reinforcer identification research allows for accurate and efficient assessment of tangible stimuli and activities for individuals with developmental disabilities and language deficits. One limitation of the current technology is that no research has been conducted to assess preference for non-tangible stimuli and activities that are not applicable to the current presentation formats. There are several benefits of being able to assess preference of non-tangible stimuli for individuals with developmental disabilities. First, person-centered approaches to developing support plans for these individuals would be able to incorporate data-supported activity preferences into plans. Previous research has shown that preferred activities identified by person-centered planning processes are not always preferred according to systematic assessments (Green, Middleton, & Reid, 2000; Reid, Everson, & Green, 1999). Second, having access to preferences beyond tangible stimuli can greatly increase quality of life for individuals with disabilities by providing more opportunities to interact socially with others and learn complex skills that preferred activities may require (Koegel, Koegel, & Dunlap, 1996). Finally, being able to assess preference for non-tangible stimuli will allow practitioners to utilize a greater range of reinforcers in treating individuals with developmental disabilities. The goal of this study was to evaluate an empirical method for assessing preference of non-tangible stimuli or stimuli that cannot be practically presented in a MS or PS format for individuals with developmental disabilities who have limited receptive and expressive language.

11

Chapter Two

Method

Participants and Setting

The participants for this study were two adults diagnosed with profound

mental retardation. Both participants lived in a residential facility for adults with

developmental disabilities. The same adults participated in both phases of the

study. Dan was a 45-year old male diagnosed with profound mental retardation.

He was non-verbal and used a few gestures and facial expressions to

communicate. George was a 35-year old male diagnosed with profound mental

retardation. He was mainly non-verbal and used gestures and a few one-word

expressions to communicate.

Sessions were conducted in the natural environment in which the activity

occurred. For example, if going to the park was being assessed, sessions were

conducted in the park. Sessions were ten minutes in length.

Stimulus Selection

Stimuli used in the preference and reinforcer assessments were selected

based on the report of residential staff and supervisors who were familiar with the

participants. Caregivers were asked to identify activities that are often available

for the participants to engage in. Of the identified activities, there were six

activities for Dan and four activities for George that did not have any associated

materials or were not practical to be presented in a multiple or paired stimulus format were selected for use in the study. Additionally, a control condition (no activity) and a work activity were also included in the assessment for both participants as a comparison with the other activities. The purpose of the control condition was to provide a baseline level of happiness for each participant as a comparison for the activity conditions. The purpose of the work task condition was to measure happiness and engagement while the participants were engaged in a repetitive and routine task that was not considered a leisure activity. Therefore, it could be determined how long the participants would engage in this task as compared to an activity that "preferred". Also, a comparison of indices happiness could be made between leisure activities included in the assessment and a work task that was not a leisure activity.

Dan's selected activities were riding a golf cart, going to the canteen, taking a walk, sitting in church, playing basketball, and riding a bike. George's selected activities were riding a golf cart, sitting on a swing outside, taking a walk, and playing basketball. Sitting in church and going to the canteen were also activities identified for George. However, after conducting the preference assessment with Dan, it was not practically possible to assess the reinforcing effects of the activities and validate all of the results for the preference assessment because of the format of the reinforcer assessment. Therefore, these activities were not included for George.

Preference Assessment

     Preference Assessment Trials. Preference for the selected activities was assessed using a single stimulus presentation format. An activity trial consisted of a 10 min opportunity for the resident to engage in an activity. At the beginning of the trial, the resident and data collector(s) were in the setting where the activity took place. The trial began when the experimenter said to the resident, "You can (activity) if you want. You don't have to and you can stop any time you want." Each set of activity trials was presented in a random order within the set. This order was determined by writing each activity on a piece of paper and drawing one at a time out of a box until all activities had been selected. The set was repeated in another random order until four trials of each activity were conducted. The random orders of all four sets were determined prior to conducting the preference assessment. Each trial was 10 min in length.

     Trials within a set were conducted in as small a time span as possible. If possible, all trials in a set were conducted in the same day with as little time between sessions as possible. Additionally, sets were conducted in as short a period of time as possible. For instance, if a set was finished in one day and there was time to conduct more trials, the next set was begun. However, the participants attended a work site for four hours a day (9:00 a.m. to 11:00 a.m. and 2:00 p.m. to 4:00 p.m.) and sessions were not conducted during these times. Most sessions were conducted from 11:30 a.m. to 12:30 p.m. and 1:00 p.m. to 2:00 p.m. The trials for sitting in church could only be conducted twice a week, Wednesday evenings and Sunday mornings, because services were only offered

at these times. The trials before and after church trials were conducted the days prior to and after the days when church services were offered.

Target Behaviors and Data Collection. Data were collected via paper and pencil using a 10 s partial-interval data collection system. There were 60 continuous intervals during each 10 min trial. During each trial, data were collected on occurrence of engagement in the activity and indices of happiness for each 10 s interval. If an indication of happiness occurred at any time during an interval, a plus sign (+) was recorded for that interval. If there was no indication of happiness during the interval, a minus sign (-) was recorded. If the participant engaged in the activity at any point during the interval, a check mark (√) was recorded. If the participant did not engage in the activity at any point during the interval, a slash mark (/) was recorded.

Due to the diversity of the activities being assessed, operational definitions of engagement in an activity (and control and work tasks) were specific to that activity and are presented in Table 1. During the control condition, participants sat in a room with a table and chairs and blank walls. During the work task, the participants sorted paper by separating each sheet and placing it in a bin. These sessions took place at a work site with tables and chairs where there were other residents of the center were sorting paper as well. Indices of happiness were defined as they were by Green and Reid (1996). Happiness was defined as any facial expression or vocalization typically considered to be an indicator of happiness among people without disabilities including smiling, laughing, and yelling while smiling.

15

Table 1. Activities and definitions for engagement for each participant

Activity	Operational Definition for Engagement
Dan	
Riding a golf cart	Seated on cart with seat belt on. If attempts to remove seat belt or stand, cart will be stopped
Going to the canteen	Walking in the direction of the canteen, is physically inside the doorway of the canteen, or walking back toward his home after being inside the canteen
Taking a walk	Walking within 3' of data collector and not moving away
Church	Seated in a pew and body oriented toward the front of the church
Playing basketball	Actively manipulating the ball or retrieving it after throwing it
Riding a bike	Seated on bike and moving pedals
Sorting paper	Sitting in chair at the table with hands touching the work materials or waiting for more work materials from staff
Control	Is physically inside the doorway of the session room
George	
Riding a golf cart	Seated on cart with seat belt on, if attempts to remove seat belt or stand, cart will be stopped
Playing basketball	Actively manipulating the ball or retrieving it after throwing it
Taking a walk	Walking within 3' of data collector and not moving away
Sitting on swing	Buttocks touching the swing
Sorting paper	Sitting in chair at the table with hands touching the work materials or waiting for more work materials from staff
Control	Physically inside the doorway of the session room

Training of observers involved verbal instruction and demonstration of the

data collection system by the principal investigator of this study. Verbal

instruction involved explanation of the data collection sheet, the operational

definitions, and the codes used for scoring. Observers first watched the principal

investigator collect data and then practiced collecting data on indices of

happiness and duration of engagement with the principal investigator on a

resident of the center, who was not a participant in the study, engaging in an

activity similar to the ones being assessed in the study. Percent agreement

between the principal investigator and the observers being trained was calculated on each of the practice sessions. Practice data collection sessions were conducted with each observer until there were three consecutive sessions of 80% or greater agreement.

During the preference assessment, observers recorded the presence or absence of engagement and happiness for each 10 s interval. An average of the percent of intervals with engagement for the four trials for each activity were used to rank order the activities from most (highest percent of intervals) to least (lowest percent of intervals) preferred. A mean percentage of intervals with indices of happiness for the four trials for each activity was used to rank order the activities from most (highest percent of intervals) to least (lowest percent of intervals) preferred. The engagement rank order and the happiness rank order were compared to determine if both measurements yielded similar results.

Interobserver agreement was collected on 31.3% of all trials for Dan and 12.5% of all trials for George by a trained independent observer. Agreement was calculated separately for the happiness and the engagement measure. Interobserver agreement was calculated by dividing the number of interval agreements by the number of interval agreements plus interval disagreements. Interobserver agreement was calculated for both the engagement and the happiness measure. Independent data collection was assured by observers who were located at least three feet away from each other during data collection. This distance allowed no observation of the other observer's data sheet. The mean interobserver agreement for Dan was 85.4% (range 78.2-98.3%) for the

17

happiness measure and 96.8% (range 81.7-100%) for the engagement measure. The mean interobserver agreement for George was 85.3% (range 72.2-96.7%) for the happiness measure and 100% for the engagement measure.

Reinforcer Assessment

Activity and Target Response Selection. A reinforcer assessment was conducted to validate the results of the preference assessment. The activities selected for the reinforcer assessment were selected based on the results of the two measures of the preference assessment. For Dan, all activities except sitting in church and going to the canteen were assessed in the reinforcer assessment. This included the control condition and sorting paper (work task). For George, sitting on the swing, taking a walk, playing basketball, and the control condition were assessed in the reinforcer assessment.

Target responses to be measured for reinforcement effects were selected based on the participants' repertoires and abilities. Simple tasks were selected in order to eliminate extraneous variables that may affect reinforcer effectiveness. Responses that did not take a long time to complete were selected in order to eliminate a long delay between trials. If responses were lengthy to complete, the response could have become aversive by delaying the participants' access to reinforcement. Additionally, more complex tasks that require discrimination or a chain of behaviors could lead to more errors while emitting the response, adding a confounding variable to the reinforcer assessment. If participants attempted to emit the response but made an error, they would not receive access to the reinforcer. This could have reduced the

18

true effectiveness of the reinforcer being assessed. Finally, these types of stimuli have not previously been included in a reinforcer assessment. Simple tasks were used to evaluate a new methodology that can be extended to more complex tasks once utility has been established. The target response for Dan was assembling a nut and a bolt. The target response for George was folding a piece of paper in half.

Data Collection. Data were collected on the occurrence or non-occurrence of an independent response for each trial of a session via a paper and pencil data collection system. A percentage of trials with a response was calculated for each session by dividing the number of trials with a response by the total number of trials (10) and multiplying that number by 100. Interobserver agreement was calculated by dividing the number of trial agreements by the number of trial agreements plus trial disagreements and multiplying by 100. Interobserver agreement was collected on 41.7% of the sessions for Dan and 50.0% of the sessions for George by a trained independent observer. Mean interobserver agreement for both participants was 100%. Independent data collection was assured by observers being at least three feet away from each other during data collection.

Pre-assessment Training. Training trials were conducted to ensure that the participants could emit their respective target responses independently and without assistance. Training trials took place in a 10'X12' session room equipped with a one-way mirror and a table and chairs. Sessions consisted of 10 trials. The trial began when the experimenter placed the response materials in front of

the participant. Each trial consisted of an opportunity to emit the response

independently (a response emitted without any assistance within 5 s of the

materials being presented), followed by a verbal prompt, model prompt, and

physical guidance if the response was not emitted within 5 s of the previous

prompt. During the pre-assessment training, the participant had 5 s to emit the

response independently before prompting began so that the training was

consistent with the actual reinforcer assessment sessions. A verbal prompt

consisted of the experimenter telling the participant to emit the response (i.e.,

"George, fold the paper."). A model prompt consisted of the experimenter telling

the participant to emit the response while demonstrating it (i.e., "George, fold the

paper like this"). A physical prompt consisted of the experimenter physically

guiding the participant to emit the response with the least amount of guidance

necessary. Training continued until the response was emitted independently on

80% of the trials across two consecutive sessions.

Procedures. A multielement design was used to assess activities

identified as high, middle, and low preference as compared to a no stimulus

(baseline) condition from the results of each rank order (engagement and

happiness). One set of activities was conducted in a random order and repeated

until the data were differentiated and stable or until it was clear that differential

responding would not occur.

Reinforcer assessment sessions took place in the setting in which the

activity being assessed usually occurs. At the beginning of the session, the

participant was as close to the setting as possible without actually engaging in

the activity. For example, if the activity was riding on a golf cart, the participant was standing next to the golf cart but not sitting on it. Sessions consisted of ten trials. A trial consisted of the response materials being presented to the participant. Sessions began when the response materials were presented for the first trial and ended when there was no response within 5 s of the materials being presented on the last trial, after the response was emitted on the last trial, or at the end of the 1 min of access to the activity being assessed. Data collectors recorded "response" or "no response" for each trial. A response was recorded when the participant independently emitted the response within 5 s of the materials being presented. One verbal prompt was given before the beginning of the first trial: "George, fold the paper." No other prompts were given during the session.

Baseline. During baseline sessions (or control condition), the target response materials were presented to the participant. If the participant emitted the target response within 5 s of the materials being presented, the materials were removed for 5 s and then placed back on the table for the beginning of the next trial. If no response was emitted within 5 s of the materials being presented, the materials were removed for 5 s and then placed back on the table for the next trial.

Reinforcement. During reinforcement sessions, the materials were presented to the participant. If the participant emitted the target response within 5 s of the materials being presented, he was given 1 min of access to the activity being assessed. If no response was emitted within 5 s of the materials being

21

presented, the materials were removed for 5 s and then placed back on the table for the next trial.

A percentage of trials with a response was calculated for each session by dividing the number of trials with a response by the total number of trials (10) and multiplying that number by 100. Interobserver agreement was calculated by dividing the number of trial agreements by the number of trial agreements plus trial disagreements and multiplying by 100.

Chapter Three

Results

Preference Assessment

The mean percentage of intervals with happiness and with engagement were the two measures of preference used to create a happiness and an engagement rank order of preference. These data are presented in Table 2 for both participants.

Table 2. Rank orders from happiness and engagement measures

Rank	Dan Happiness	Dan Engagement	George Happiness	George Engagement
1	Church	*Church	Golf cart	*Golf cart
2	Canteen	*Walk	Swing	*Swing
3	Walk	*Golf cart	Walk	*Control
4	Ball	*Control	Ball	Walk
5	Bike	*Sorting paper	Control	Sorting paper
6	Golf cart	Canteen	Sort paper	Ball
7	Control	Bike		
8	Sort paper	Ball		

* Mean of 100% of intervals with engagement, all ranked 1

The happiness measure produced more differentiated results than the engagement measure for Dan, as seen in Figure 1. Sitting in church and going to the canteen had the highest mean percentage of intervals with happiness, followed by taking a walk, playing ball, riding a bike, and riding the golf cart. The control condition and sorting paper had the lowest mean percentage of intervals with happiness. The mean percentage of intervals with happiness for the control condition and the work task were 11.3% and 10.4%.

The engagement measure for Dan did not produce differentiated results, as the happiness measure had done. There was a mean of 100% of intervals with engagement for most activities, including the control condition and work task, except for going to the canteen, which averaged 97.9%. Only one of the trials for canteen had less than 100% of intervals with engagement (91.7%), as shown in Figure 3. Playing basketball and riding a bike had lower mean percentages of interval with engagement, which averaged 82.5% and 95.0%, respectively.

Given the small differences in the happiness measure between activities, four activities and the work task were selected for assessment in the reinforcer assessment with the goal of validating the rank order produced by the happiness measure. Sitting in church and going to the canteen were not included because these activities were not able to be practically presented in the reinforcer assessment format.
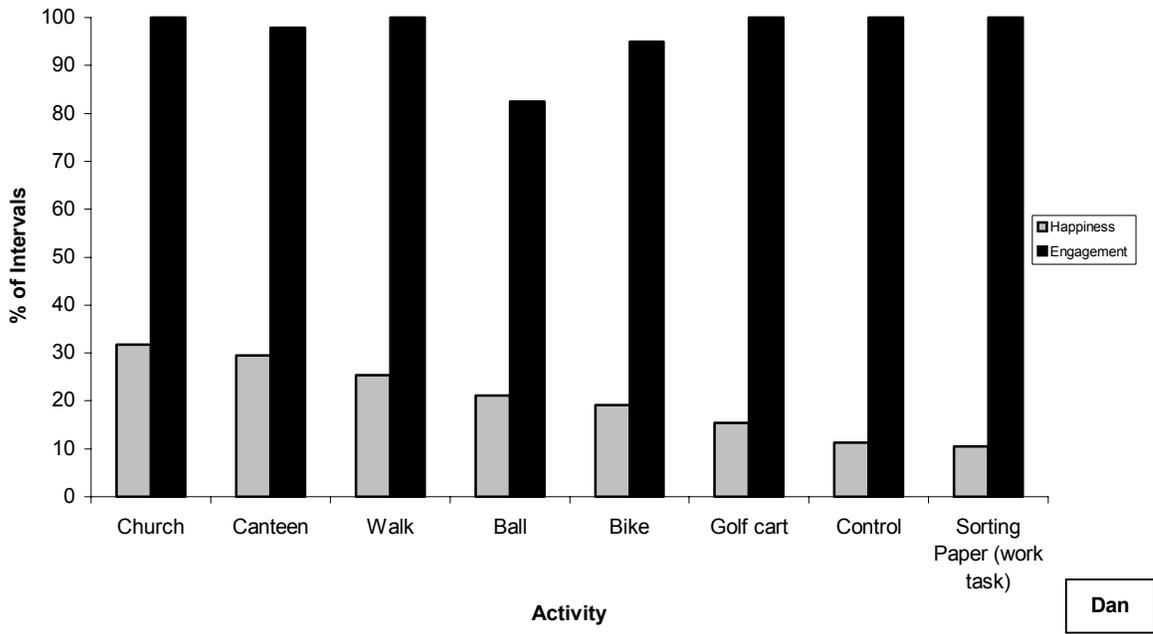
Figure 1. Mean percent of intervals with happiness and engagement for Dan
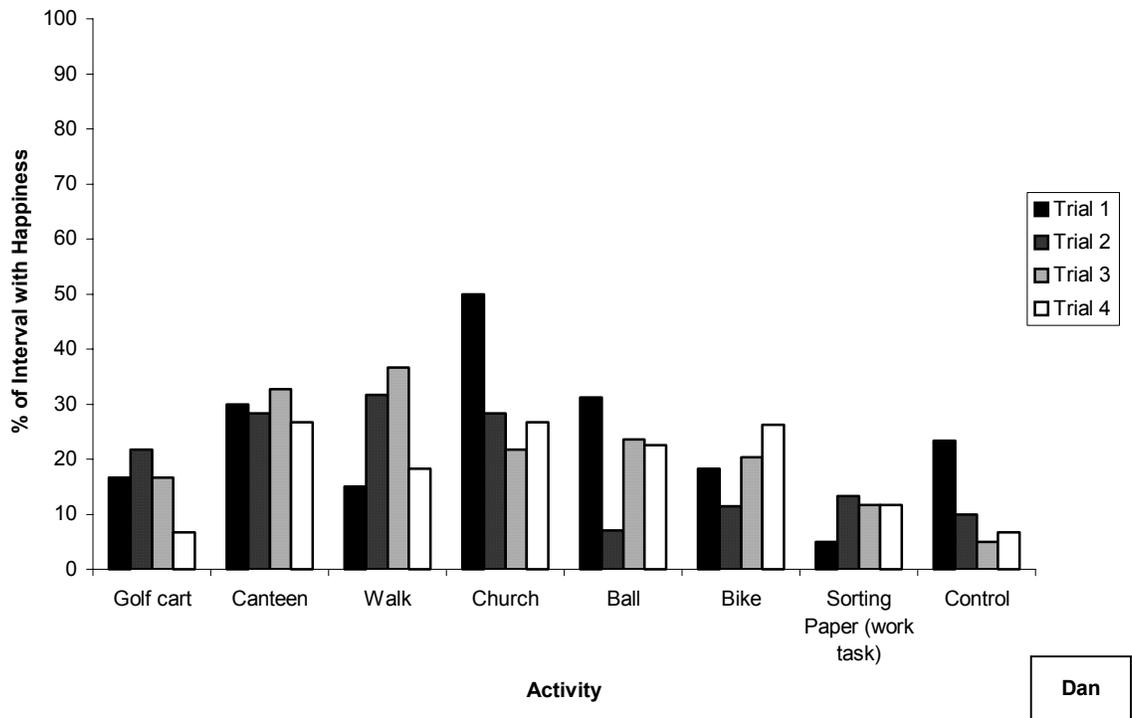

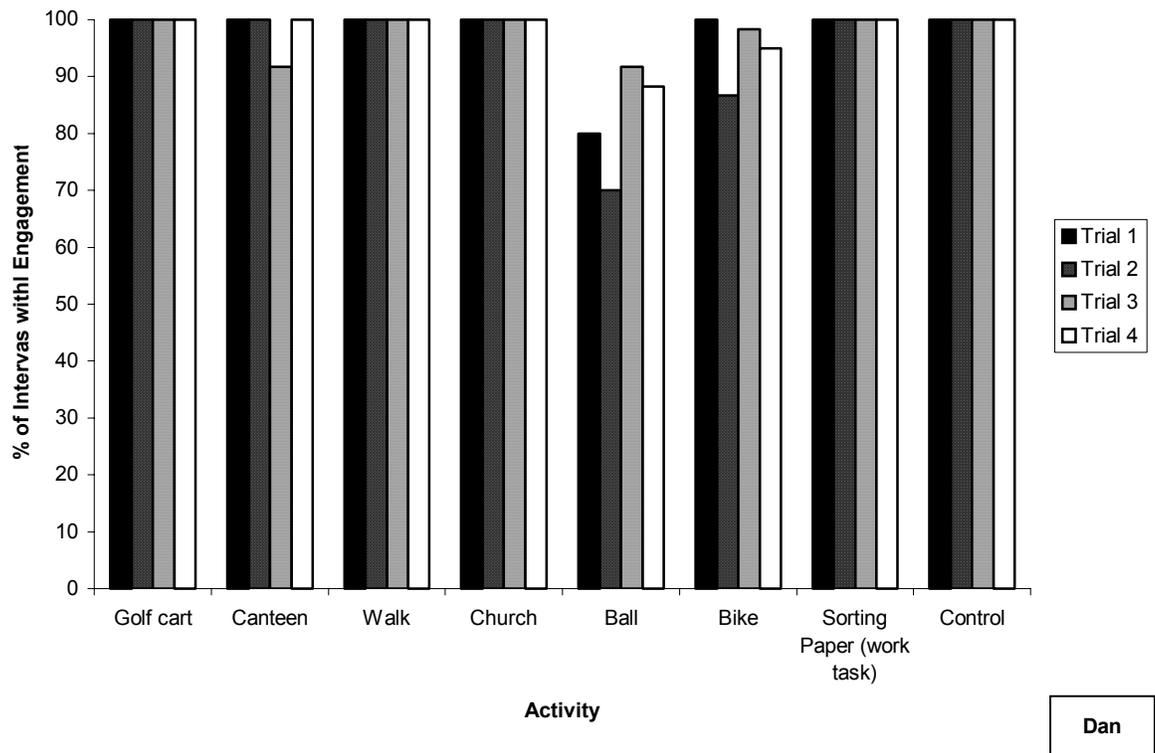Figure 2. Percent of intervals with happiness per trial for Dan

Figure 3. Percent of intervals with engagement per trial for Dan

George displayed less overall happiness than Dan during the preference assessment but his happiness data showed some differentiation. The rank order produced by the happiness measure from most to least happiness is rising a golf cart, sitting on a swing, taking a walk, playing basketball, control, and sorting paper (work task). All the activities assessed (riding a golf cart, sitting in a swing, taking a walk, and playing basketball) had higher mean percentages of intervals with happiness than the control condition or the work activity (see Figure 4). There was little difference between the means for the activities. However, in examining the data across trials (see Figure 5), there was only one trial for

playing basketball where there was a high percentage of intervals with happiness (33.3%). This trial's percentage was particularly high because there were only three 10-second intervals where George was engaged in the activity and he displayed an incident of happiness during one of these intervals, making the percent 33.3%. This trial significantly inflates the mean because the two of trials had no intervals with happiness and one trial had 5.6%. Therefore, without this trial, the mean percentage of intervals with happiness was 1.9%, which is lower than the control condition.

The engagement measure produced similar results. Riding a golf cart and sitting on a swing had 100% of intervals with engagement, followed by taking a walk (80.0%), sorting paper (61.3%), and playing basketball (20.8%). The control condition averaged 100% of intervals with engagement.

As can be seen in Table 2, the rank orders produced by the two measures yielded similar results. The two highest activities were the same, riding a golf cart and sitting on a swing. Walk was identified as a middle preference activity and playing basketball as a low preference activity by both rank orders. Given the results of the two measures for George, three activities were selected for the reinforcer assessment. Sitting on a swing (high preference), taking a walk (middle preference), and playing basketball (low preference) were assessed in addition to the control condition.
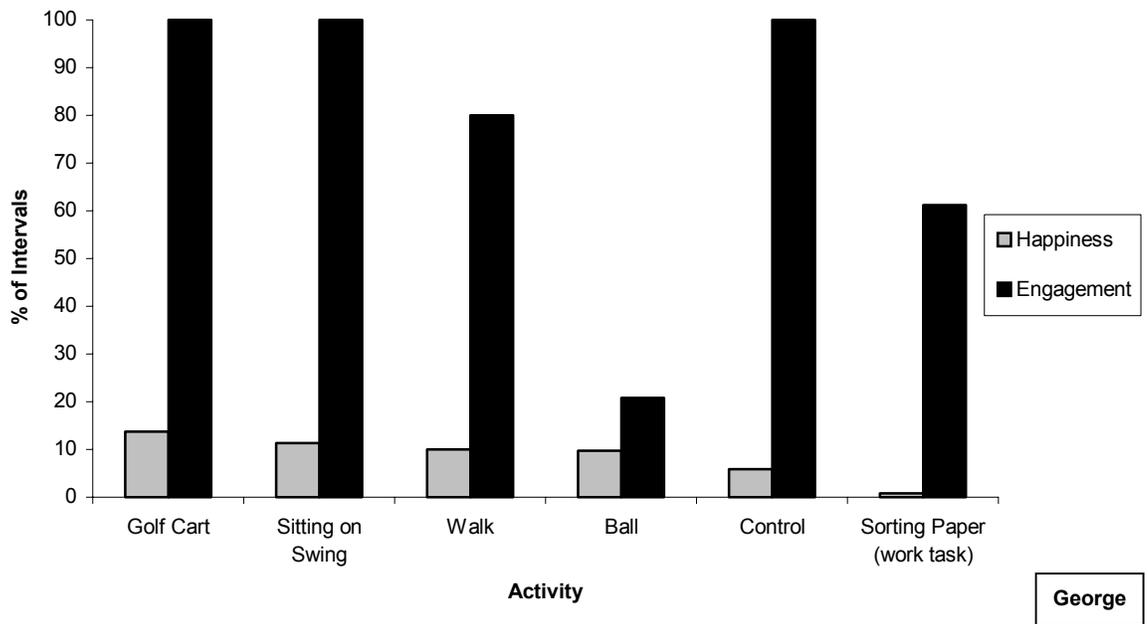
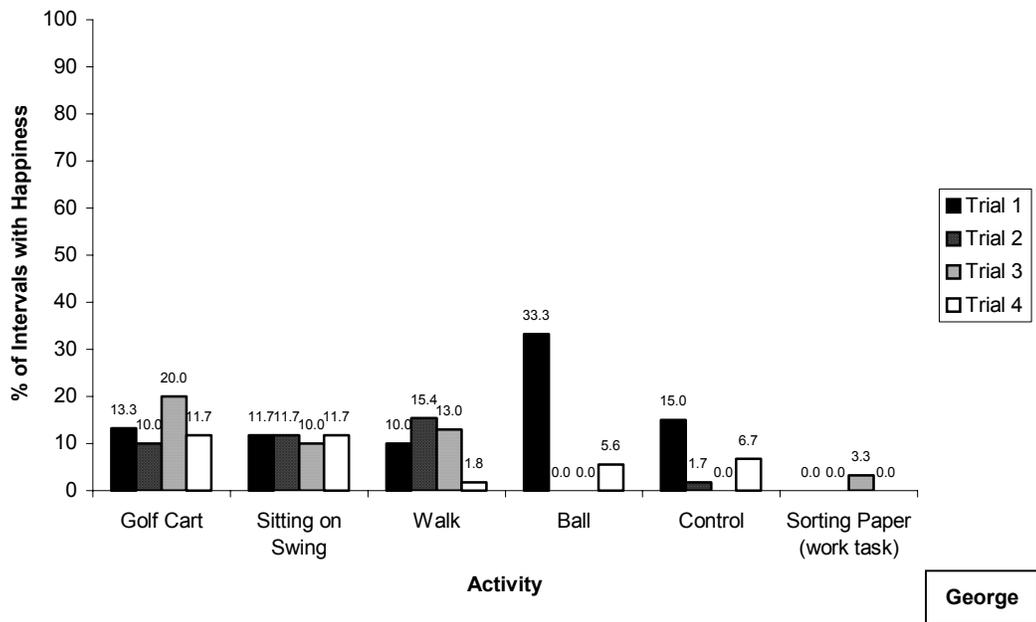Figure 4. Mean percent of intervals with happiness and engagement for George

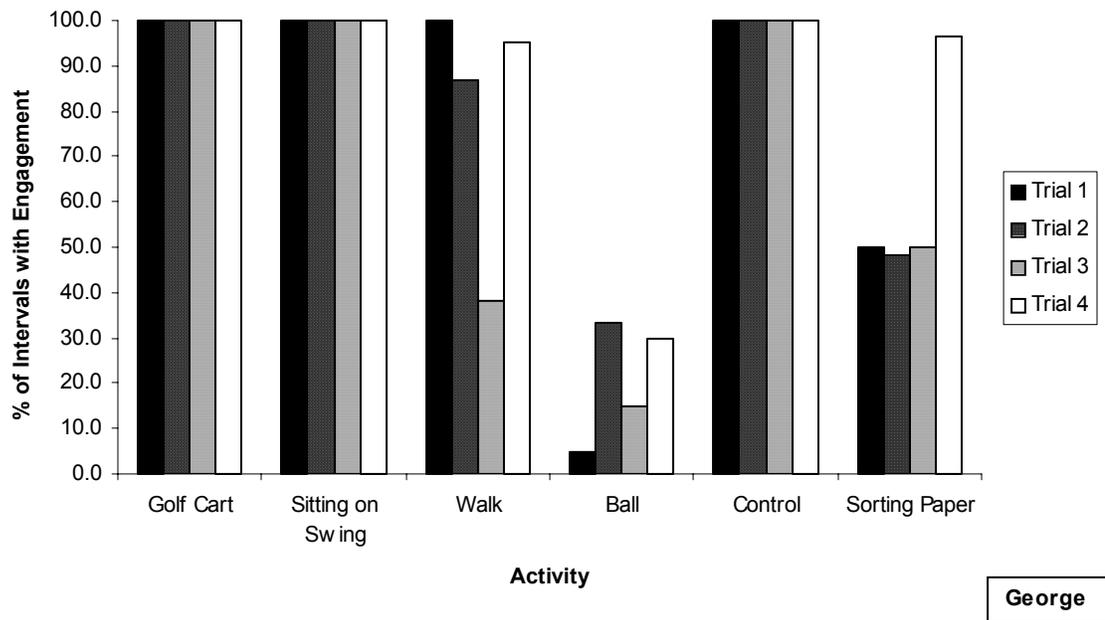Figure 5. Percent of intervals with happiness per trial for George

Figure 6. Percent of intervals with engagement per trial for George

Reinforcer Assessment

The results of the preference assessment could not be validated by the results of the reinforcer assessment for either participant. Dan responded independently with the target response (assembling a nut and bolt) on all trials across all activities, including the control condition for two sets of conditions. George also responded independently with his target response (folding a piece of paper in half) on all trials across all activities and the control condition for two sets of conditions. Therefore, it could not be determined if the activities that were high preference as indicated by the happiness and/or the engagement measure functioned as more effective reinforcers than those indicated as low preference.
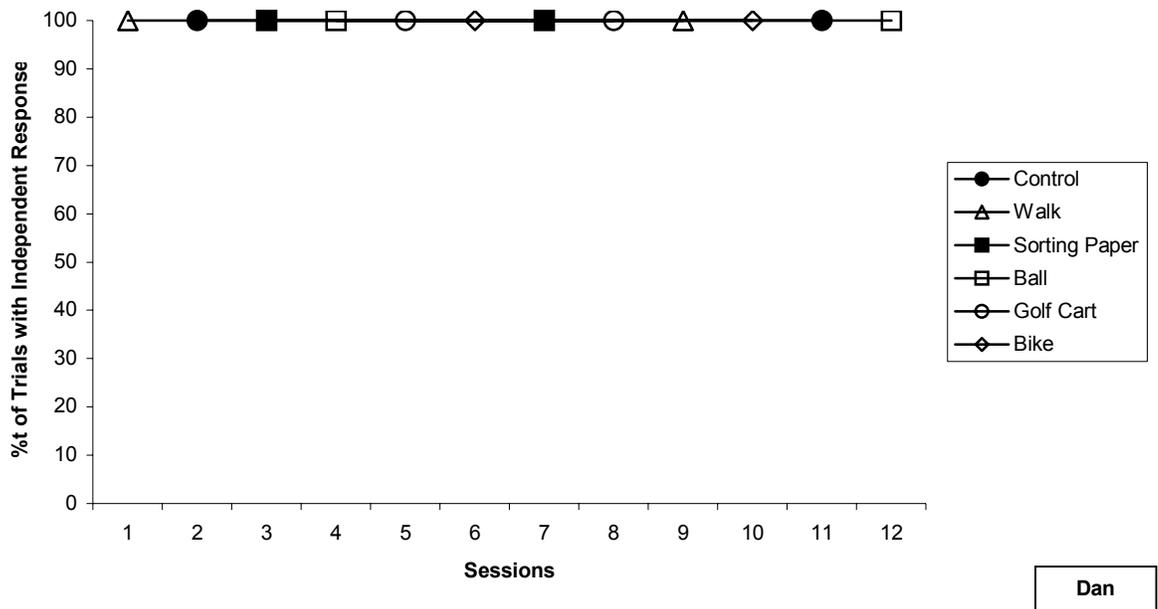
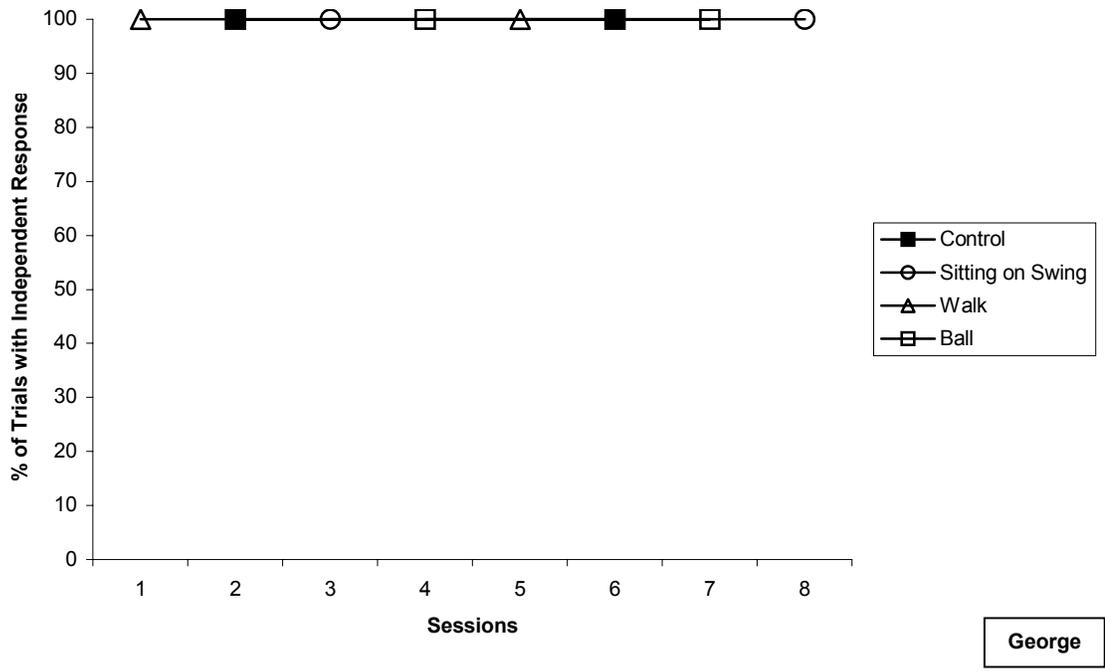Figure 7. Reinforcer assessment results for Dan

Figure 8. Reinforcer assessment results for George

Chapter Four

Discussion

The evaluation of this method for assessing preference for activities and stimuli that cannot be practically presented in multiple or paired stimulus format showed tentative potential. The happiness measure produced differential results for both participants. Some activities had higher percentages of intervals with happiness than others and all activities being assessed had higher percentages of interval with happiness than the work task and the control condition. The inclusion of the control and work task conditions was successful in identifying a baseline level of happiness with which to compare the activities being assessed. This suggests that the activities assessed were more preferred than not engaging in an activity and engaging in a work task. The happiness data produced a clear hierarchy. However, caution should be taken in making inferences of preference based on those results due to the small differences between activities, especially those in the middle ranks of the hierarchy. Stronger conclusions can be drawn about preferences between activities at the high and low ends of the hierarchy, such as sitting in church (ranked 1) and riding the golf cart (ranked 6) for Dan. The 16.2% difference in percentage of intervals with happiness suggests that sitting in church is more preferred than riding a golf cart. However, the same confidence cannot be given for playing ball (ranked 4)

as more preferred than riding a bike (ranked 5) for Dan, where there was only a 2.0% difference in the happiness measure.

The engagement measure only produced differentiated results for one participant, George. The results of his engagement measure suggest that riding a golf cart and sitting on a swing were the highest preference activities, taking a walk was a middle preference activity, and playing basketball was a low preference activity.

Dan engaged in all activities on almost all of the trials. Playing basketball had the lowest percentage of intervals with engagement. Although it may appear that playing basketball was less preferred than the other activities, the lower percentage of intervals with engagement may have been a result of the operational definition. The operational definition for engaging in playing basketball was actively manipulating the ball or retrieving it after throwing it. During the basketball trials, Dan would stand in front of the basket while holding the ball and make loud vocalizations for about 30 seconds approximately two times during each trial. During these times he was holding the ball but not manipulating it as the definition states. Therefore, he was not considered engaged in the activity and the percent of intervals was lower. However, he would remain in the area of the basket and continue playing basketball for the duration of the trial. It is possible that if the operational definition had been different, he would have had a mean near 100% of intervals with engagement. This example highlights one of the potential problems with the engagement measure with these types of stimuli. Operational definitions have to be specific

to the stimuli. However, the definitions for engaging in different activities may require the specification of different actions involving various levels of physical effort. For example, the main action required for engaging in riding the golf cart was sitting. The main action for playing basketball was manipulating the ball, which requires more exertion than sitting. These differences may make comparisons between these activities difficult. Despite this limitation, interaction or engagement has been shown to be an effective measure of preference with tangible stimuli. It may serve as a more effective measure of preference for activities if the activities have similar features, such as setting, physical requirements, etc. than the activities assessed in this study.

Another limitation to the engagement measure was the comparison control condition. Both participants engaged in the control condition for 100% of the intervals across all trials. The control condition consisted of sitting in a chair at a table in a small room for ten minutes. George spent a lot of time sitting and waiting in his daily routine and had likely sat at a table and waited for longer than ten minutes on many occasions. It is possible that sitting in the room for the entire session length was not a reflection of his preference for the activity but simply a situation that occurs daily and a behavior in which he frequently engages normally. Additionally, the results for the work task, sorting paper, may have been influenced by George's daily routine. Sorting paper is a task that George engaged in daily to earn money. Therefore, there was a history of tangible reinforcement for engaging in the activity, which may not make it a good comparison activity for the engagement measure. Dan also spent a lot of time

sitting and waiting and sorted paper to earn money. The same conclusions can be drawn for his data in the control condition and work task. The control and the work conditions were included primarily to serve as comparison conditions for the happiness measure to see how much happiness was exhibited by the participants while they were not engaging in the test activities.

The results of this study suggest that indices of happiness can be useful and convenient measures of preference, especially for the type of stimuli used in this study. These data can be collected in a variety of settings and do not require the assessor to manipulate associated materials that certain stimuli may not possess. Additionally, these data can be collected while the participant is actually interacting with the stimuli or engaging in the activity and does not require the stimulus or activity to be delivered contingent on a selection response. The utility of indices of happiness may have potential as a measure of preference in systematic preference assessment. Unfortunately this potential could not be confirmed by the reinforcer assessment. Green and Reid (1996) were able to show that higher indices of happiness occurred when in the presence of higher preference stimuli than in the presence of lower preference stimuli. The present study demonstrated that for the participants, higher indices of happiness were associated with activities that occurred during leisure time (taking a walk, playing basketball) than while the participants were not engaged in an activity or in a work task. Further studies might examine the utility of indices of happiness as actual measures of preference for tangible stimuli by measuring indices of happiness while participants are interacting with tangible

stimuli. This would bolster the use of indices of happiness as measures of preference.

One of the most obvious limitations to this study was the failure of the reinforcer assessment to validate the results of the preference assessment. Both participants emitted the target response on all trials across all activities. A possible explanation for these data lies in the environment of the center at which they live. Both participants have lived at the center since they were children. Typically the resident to staff ratio is high and most residents receive little individual attention and reinforcement for compliance with requests. Requests are made of Dan and George often and they usually comply immediately. If they do not comply, staff often continually prompt in a stern voice until compliance occurs or they are physically guided to complete a request. It is possible that continuous prompting and physical guidance are aversive to Dan and George and their high frequency of compliance is a result of negative reinforcement. This combined with the lack of positive reinforcement for compliance may explain why Dan and George completed the task immediately without prompting, regardless of the activity (or lack of activity) that followed.

If the study had been conducted with participants without this type of history, the results may have been different. The results of the preference assessment may have been validated by differential responding during the reinforcer assessment. Future research might replicate this study with different participants, such as adults with developmental disabilities living in group homes or with family or children with developmental disabilities, in order to account for

this limitation. Additionally, reinforcer assessments might test the effects of potential reinforcers in supporting and increasing more complex behaviors.

Finally, other reinforcer assessment formats might be used to measure reinforcer effects. It is possible that had a different experimental design been used, differential responding may have occurred. For example, if a reversal design had been used where stable responding for one activity was established before testing another activity, the participants may have better discriminated between activities. Additionally, reversing back to the control condition in between activities may have also aided in differential responding.

This study assessed preferences for activities and stimuli that have not been included in systematic preference assessments in the past because they were not practical to present in paired or multiple stimulus format. Until now, these types of stimuli have been absent in the extensive preference assessment literature. In practice, often assumptions are made that the activities available to people with developmental disabilities are preferred without any supporting evidence of their preferences. A preference assessment of this type would help ensure that opportunities are given to engage in preferred activities. Additionally, the development and refinement of a method to assess preference for these types of stimuli would assist practitioners in the identification of many more reinforcers to use in the function-based treatment of problem behaviors. Regardless of the next steps taken, the development of a methodology appropriate to these activities would undoubtedly benefit people with

developmental disabilities by identifying a wider range of reinforcers and

preferred activities that would increase their quality of life.

References

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.

Cohen-Almeida, D., Graff, R. B., & Ahern, W. H. (2000). A comparison of verbal and tangible stimulus preference assessments. Journal of Applied Behavior Analysis, 33, 329-334.

DeLeon, I. G., Fisher, W. W., Rodriguez-Catter, V. R., Maglieri, K., Herman, K., & Marhefka, J. (2001). Examination of relative reinforcement effects of stimuli identified through pretreatment and daily brief preference assessments. Journal of Applied Behavior Analysis, 34, 463-473.

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.

DeLeon, I. G., Iwata, B. A., Conners, J. & Wallace, M. D. (1999). Examination of ambiguous stimulus preferences with duration-based measures. Journal of Applied Behavior Analysis, 32, 111-114.

Derby, K. M., Wacker, D. P., Andelman, M., Berg, W., Drew, J., Asmus, J., Prouty, A., & Laffey, P. (1995). Two measures of preference during forced-choice assessments. Journal of Applied Behavior Analysis, 28, 345-346.

Fisher, W. W., Piazza, C. C., Bowman, L., & Amari, A. (1996). Integrating caregiver report with a systematic choice assessment to enhance reinforcer identification. American Journal on Mental Retardation, 101, 15-25.

Fisher, W. W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., & Langdon, N. A. (1994). Empirically derived consequences: A data-based method for prescribing treatments for destructive behavior. Research in Developmental Disabilities, 15(2), 133-149.

Fisher, W. 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.

Fisher, W. W., Piazza, C. C., Bowman, L. G, Kurtz, P. F., Sherer, M. R. & Lachman, S. R. (1994). A preliminary evaluation of empirically derived consequences for the treatment of pica. Journal of Applied Behavior Analysis, 27, 447-457.

Fisher, W. W., Piazza, C. C., Hagopian, L. P., Bowman, L. G., & Toole, L. (1996). Using a choice assessment to predict reinforcer effectiveness. Journal of Applied Behavior Analysis, 29, 1-9.

Foster-Johnson, L., Ferro, J., & Dunlap, G. (1994). Preferred curricular activities and reduced problem behaviors in students with intellectual disabilities. Journal of Applied Behavior Analysis, 27, 493-504.

Green, C. W., Gardner, S. M., & Reid, D. H. (1997). Increasing indices of happiness among people with profound multiple disabilities: A program replication and component analysis. Journal of Applied Behavior Analysis, 30, 217-228.

Green, C. W., Middleton, S. G. & Reid, D. H. (2000). Embedded evaluation of preferences sampled from person-centered plans for people with profound multiple disabilities. Journal of Applied Behavior Analysis, 33, 639-642.

Green, C. W. & Reid, D. H. (1996). Defining, validating, and increasing indices of happiness among people with profound multiple disabilities. Journal of Applied Behavior Analysis, 29, 67-78.

Green, C. W. & Reid, D. H. (1999). Reducing indices of unhappiness among individuals with profound multiple disabilities during therapeutic exercise routines. Journal of Applied Behavior Analysis, 32, 137-147.

Green, C. W., Reid, D. H., White, L. K., Halford, R. C., Brittain, D. P., & Gardner, S. M. (1998). Identifying reinforcers for persons with profound handicaps: Staff opinion versus systematic assessment of preferences. Journal of Applied Behavior Analysis, 21, 31-43.

Green, C. W., Reid, D. H., Canipe, 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.

Hagopian, L. P., Rush, K. S., Lewin, A. B., & Long, E. S. (2001). Evaluating the predictive validity of a single stimulus engagement preference assessment. Journal of Applied Behavior Analysis, 34, 475-485.

Hanley, G. P., Iwata, B. A, & Lindberg. J. S. (1999). Analysis of activity preferences as a function of differential consequences. <u>Journal of Applied Behavior Analysis, 32</u>, 419-435.

Higbee, T. S., Carr, J. E., & Harrison, C. D. (1999). The effects of pictorial versus tangible stimuli in stimulus preference assessments. <u>Research in Developmental Disabilities, 20</u>(1), 63-72.

Ivancic, M. T. (2000). Stimulus preference and reinforcer assessment application. In J. Austin & Carr, J. E. (Eds.), <u>Handbook of Applied Behavior Analysis</u> (pp. 19-38). Reno, NV: Context Press.

Koegel, L. K., Koegel, R. L., & Dunlap, G. (1996). <u>Positive behavioral support: Including people with difficult behavior in the community</u>. Baltimore: Brookes.

Northup, J. (2000). Further evaluation of the accuracy of reinforcer surveys: A systematic replication. <u>Journal of Applied Behavior Analysis, 33</u>, 335-338.

Northup, J., George, T., Jones, K., Broussard, C. & Vollmer, T. (1996). A comparison of reinforcer assessment methods: The utility of verbal and pictorial choice procedures. <u>Journal of Applied Behavior Analysis, 29</u>, 201-212.

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

Paclawskyj, T. R. & Vollmer, T. R. (1995). Reinforcer assessment for children with developmental disabilities and visual impairments. <u>Journal of Applied Behavior Analysis, 28</u>, 219-224.

Parsons, M. B., Harper, V. N., Jensen, J. M., & Reid, D. H. (1997). Assisting older adults with severe disabilities in expressing leisure preferences: A protocol for determining choice-making skills. <u>Research in Developmental Disabilities, 18</u>(2), 113-126.

Piazza, C. C., Fisher, W. W., Hagopian, L. P., Bowman, L. G., & Toole, L. (1996). Using a choice assessment to predict reinforcer effectiveness. <u>Journal of Applied Behavior Analysis, 29</u>, 1-9.

Reid, D. H., Everson, J. M., & Green, C. W. (1999). A systematic evaluation of preferences identified through person-centered planning for people with profound multiple disabilities. <u>Journal of Applied Behavior Analysis, 32</u>, 467-477.

Reid, D. H., Parsons, M. B. & Green, C. W. (1998). Identifying work preferences among individuals with severe multiple disabilities prior to beginning supported work. <u>Journal of Applied Behavior Analysis, 31</u>, 281-285.

Ringdahl, J. E., Vollmer, T. R., Marcus, B. A., & Roane, H. S. (1997). An analogue evaluation of environmental enrichment: The role of stimulus preference. <u>Journal of Applied Behavior Analysis, 30</u>, 203-216.

Roane, H. S., Vollmer, T. R., Ringdahl, J. E., & Marcus, B. A. (1998). Evaluation of a brief stimulus preference assessment. <u>Journal of Applied Behavior Analysis, 31</u>, 605-620.

Skinner, B. F. (1953). <u>Science and Human Behavior</u>. New York: The Free Press.

Windsor, J., Piche, L. M. & Locke, P. A. (1994). Preference testing: A comparison of two presentation methods. <u>Research in Developmental Disabilities, 15</u>(6), 439-455.

Zhou, L., Iwata, B. A., Goff, G. A., & Shore, B. A. (2001). Longitudinal analysis of leisure-item preferences. <u>Journal of Applied Behavior Analysis, 34</u>, 179-184.