Evaluation of Preference Assessment Procedures for Use with Infants and Toddlers

Karena S. Rush, Ph.D., Bruce P. Mortenson, Ph.D., and Sarah E. Birch, M.S. Ed

Abstract

The current practice of preference assessment offers a variety of approaches to determine an individual's preference for specific items. However, the majority of the research has been completed on school-age children or individuals with significant cognitive or behavioral deficits. Further, studies on preference assessments have not adequately addressed the toddler population and have almost completely omitted the infant population. The current paper reviews the literature on preference assessments inclusive of both direct and indirect assessment methods. A comparison of methods follows, detailing the advantages and disadvantages of the techniques for use with very young children. The paper concludes with practical considerations for future research endeavors in the area of preference assessments with infants and toddlers.
Keywords: Preference Assessment, Reinforcer Assessment, Infant Assessment

Behavior intervention techniques which employ positive reinforcement are effective for promoting skill acquisition and reducing maladaptive behaviors among a wide variety of populations (Akin-Little, Eckert, Lovett & Little, 2004; Ellison, 1997; Lannie & McCurdy, 2007). For such strategies to work, effective reinforcers must be identified. The most effective means for identifying reinforcers is through the use of systematic preference and reinforcer assessments (Hagopian, Long, & Rush, 2004; Logan & Gast, 2001). Preference assessments are techniques designed to identify items that are preferred by an individual and thus may serve as reinforcers. However, preference assessments do not actually assess the reinforcing value of an item. A reinforcer assessment may be used following a preference assessment to determine whether or not a preferred item functions as a reinforcer. Reinforcer assessments are primarily utilized in research to assess the validity of a preference assessment approach, whereas preference assessments are widely used in both research and clinical practice (Hagopian et al., 2004; Logan & Gast, 2001).

Current research indicates that "reinforcers" are widely used with the infant and toddler population. For example, legislative and policy making bodies clearly advocate for early intervention inclusive of children as young as 18 months or younger who exhibit delays in adaptive behavior, communication, and/or social-emotional skills (Chao, Bryan, Burstein,& Cevriye, 2006; Individuals with Disabilities Education Improvement Act, 2004; Gresham, Beebe-Frankenberger, & MacMillan. 1999; Webster-Stratton & Reid, 2003). These children may be exposed to intensive intervention programs which often include significant reinforcement components. In addition, young children with feeding disorders typically participate in intensive feeding interventions that require the identification of preferred foods or items to increase food consumption (Piazza, Patel, Gulotta, Sevin, & Layer, 2003). Young children who exhibit early onset severe behaviors disorders (such as self-injurious behavior) are also placed in intensive behavioral programs designed to decrease severe problem behaviors while increasing appropriate communicative responses (Kurtz et al., 2003). Finally, typically functioning infants and toddlers are frequently exposed to reinforcement procedures. For example, reinforcement is often used to teach infants sign language (Thompson, Cotnoir-Bichelman, McKerchar, Tate, & Dancho, 2007; Thompson, McKerchnar, & Dancho, 2004), to enhance toilet training programs (Luxim, Christophersen, & Purvis, 1997), and to increase desirable behaviors in the preschool and daycare settings (Hanley, Cammilleri, Tiger & Ingvarsson, 2007; Layer, Hanley, Heal, & Tiger, 2008). Surprisingly, while the use of reinforcement is reported in many studies examining the effectiveness of intervention techniques with

these populations, there is very little research on the validity of the assessment techniques used prior to the development of the intervention. It appears that many decisions are based on clinical judgment rather than empirical evidence. Without the knowledge or understanding of the idiosyncratic nature of the child's preference, we risk designing misaligned interventions during a critical period for the child (Kennedy, 2002).

A large body of research describing valid and reliable techniques to effectively determine reinforcers does exist for other populations such as the developmentally disabled (Hagopian et al., 2004; Logan & Gast, 2001). In fact, the majority of preference assessment techniques currently in use were originally designed for this population (e.g. DeLeon & Iwata, 1996; Fisher et al, 1992; Green et al., 1988; Roane, Vollmer, Ringdahl & Marcus, 1998; Windsor, Piche, & Locke, 1994). Identifying potential reinforcers for the general education population has also received significant attention (e.g. Daly et al., 2009; Damon, Riley-Tillman, & Fiorello, 2008; Northrup, George, Jones, Broussard, & Vollmer, 1996; Resetar & Noell, 2008; Schanding, Tingstrom, & Sterling-Turner, 2009). Given that working with very young children presents a unique set of challenges such as limitations in communication, short attention spans and rapidly changing interests, it would be important to consider such factors when developing effective assessments for use with this population.

The purpose of this study is to review current literature on preference assessments that may be utilized when working with very young children with and without disabilities. Both direct and indirect preference assessment procedures will be examined and techniques for modifying current assessments for use with this population will be discussed. Lastly directions for future research will be presented.

Indirect Preference Assessment Techniques

Interviews. Interviews are one of the most common assessment techniques used for identifying potential reinforcers. Interviews can be structured or unstructured and are typically conducted with one or more care providers including parents, teachers, and direct care staff. In an unstructured interview, the care provider is asked to name items that the child likes (e.g. Piazza et al., 2003; Thompson, Cotnoir-Bichelman, McKerchar, Tate, & Dancho, 2007). Sometimes it is difficult for parents or teachers to identify items, particularly when they have a very young child with developmental delays. In these instances, a structured interview might allow the examiner to elicit more information to assess potential reinforcers. Given the lack of structured interviews designed specifically for the infant and toddler population existing interview schedules may be adapted for use. One interview that may be modified for use with this population is the Reinforcer Assessment for Individuals with Severe Disabilities (RAISD; Fisher, Piazza, Bowman, & Amari, 1996). The RAISD was originally developed as an assessment tool for individuals with severe developmental disabilities. The interviewer asks questions about potentially reinforcing items including tactile, auditory, visual, and olfactory stimulating items as well as edible items and social stimuli (Fisher et al., 1996). Recently, this interview was modified for use with typically developing toddlers (Cote, Thompson, Hanley & McKerchar, 2007). In this study, the examiners used questions pertaining only to auditory and social stimuli. Cote and colleagues found that items identified during the interview as preferred consequently served as reinforcers when assessed in a reinforcer assessment. Preferences identified through the interview were then compared to results obtained from a direct preference assessment using the paired-stimulus technique described by Fisher et al. (1992). The results suggested that both assessment methods identified items that functioned as reinforcers. However, the items identified as preferred in the direct assessment served as more potent reinforcers when reinforcer assessments were conducted (Cote et al., 2007).

Surveys. Two types of surveys exist in the literature for assessing preference, the standardized survey and the clinician generated survey. To date, standardized surveys have been developed for use with individuals with developmental disabilities and school age children (Fantuzzo, Rohrbeck,

Hightower, & Work, 1991; Matson et al., 1999). However, none are currently available for the infant and toddler population. Clinician or researcher generated preference assessment surveys are much more common. On clinician generated preference assessment surveys, a list is typically generated of potential reinforcers and the participant or caregiver is asked to identify preferred items. The person completing the form may be asked to check off all preferred items, rank the items in order of preference, state how much they or their child likes an item, or pick between two items.

Northrup and colleagues evaluated two different types of clinician generated preference assessment surveys (Northrup, 2000; Northrup et al., 1996). The first survey, a modified version of the Child Reinforcement Survey (CRS; Fantuzzo et al., 1991), included a list of items representing the following categories- edible items, peer attention, teacher attention, activities, tangible items, and escape. In these studies, examiners read a list of items and asked the child to tell them if they liked the item "not at all," "a little" or "a lot." Total accuracy, when compared to results of the reinforce assessments, was approximately 57%. This suggests that the surveys were no better than chance in predicting reinforcers. The researchers concluded that surveys may actually do a better job identifying which items would *not* serve as reinforcers.

Northrup and colleagues also evaluated a survey conducted in the forced-choice format (Northrup, Jones, Broussard, & George, 1995; Northrup et al., 1996). In this format, the child was asked "would you rather play with toy 1 or toy 2." The results from this survey resulted in a total accuracy of 70% when compared to the results of the reinforcer assessment. The benefit of the verbal forced-choice assessment was that it only took 2-3 minutes to administer and appeared to produce more reliable and valid results when compared to the traditional survey format. However, these studies were conducted with school age children (ages 5-9) with ADHD thus it would be important to determine at what age a child may serve as a reliable informant as there is significant evidence that suggest a poor correspondence between verbal self-report and subsequent behavior for young children (Northrup, 2000; Risley & Hart, 1968). A more likely procedure when working with young children would be having the parent or teacher serve as the informant.

Didden and de Moor (2004) conducted a similar study with parents and teachers of toddlers with and without developmental delays. Parents and teachers completed a forced-choice survey in which they were instructed to identify which of two items the child would prefer. Each item was paired with every other item. Items were then placed in rank order. Results indicated that the pair-wise survey did not significantly correspond with a direct assessment measure of student preference. Reinforcer assessments were not conducted so it is unknown if those items that were identified as potential reinforcers in the survey actually served as reinforcers.

The Didden and de Moor (2004) findings were consistent with other research on indirect preference assessment measures. Specifically, Green, Reid, Canipe and Gardner (1991) developed a Likert-type measure in which they asked direct care staff to measure the preference of specific items for their clients. They found low correspondence between items identified as preferred by staff members and items identified as preferred using a direct assessment technique. Additionally, the authors found that only items identified as preferred through the use of direct assessment subsequently functioned as effective reinforcers for the participants. Parson and Reid (1990) developed a 5-point rating scale of edible items and asked caregivers to rate the preference of each item. This study evidenced low correspondence between items identified as preferred by staff members and those identified as preferred during direct assessment.

Overall, the majority of the research suggests that indirect measures are not as accurate as direct measures in identifying items that may serve as reinforcers (Cote et al., 2007; Didden & de Moor, 2004; Green et al., 1991; Parson & Reid, 1990; Windsor et al., 1994). They often result in multiple items being

identified as potential reinforcers, when in practice not all of the items identified as preferred served as reinforcers when assessed via a direct reinforcer assessment. As imperfect as the indirect approaches are, these methods do have utility. First, they provide a starting point for direct measures. Given the lack of communicative abilities often inherent in the population served, survey and interview approaches are needed to narrow the possible field of reinforcers. The lack of accuracy on the caregivers' part to identify sustainable reinforcers does not diminish the importance of their input and participation. In addition, when time or resources are limited, indirect measures have been found to identify items that may serve as reinforcers even if they are not as potent as reinforcers identified via direct measures.

Direct Preference Assessment Techniques

Direct preference assessment techniques are considered the most valid means of identifying potential reinforcers (Hagopian et al., 2004). Below is a description of techniques frequently used and a review of the literature on the psychometric properties of each procedure.

Forced-Choice Assessment. The most widely used direct assessment technique is the forced-choice or paired-choice assessment (Fisher et al., 1992). In this assessment, the individual is shown items in pairs and asked to pick one. Each item is paired with every other item. Based on the number of times an item is chosen, a hierarchy is developed. There is a large body of research on the effectiveness of this procedure in identifying potential reinforcers for individuals with developmental disabilities (Fisher et al., 1992; Hagopian et al., 2004; Haynes, Derby, McLaughlin, & Weber, 2002). In addition, there is research expanding its use to children in the general educational classrooms (Damon et al., 2008; Northrup et al., 1995; Schanding et al., 2009). Recently, research has emerged extending its use to young children. Cote et al. (2007) conducted forced-choice preference assessments with typically developing toddlers (ages 18-29 months). Cote and colleagues found that the forced-choice assessment method resulted in high levels of predictive validity. Specifically, items identified as highly preferred in the forced-choice assessment also served as reinforcers in the reinforcer assessment. Similar results were found in a study conducted with an infant (Rush, Kurtz, Leiblein, & Chin, 2005). In this study, forced-choice preference assessments and reinforcer assessments were conducted with a 13-month-old boy with a history of prenatal drug exposure. A modification was made to the procedure described in Fisher et al. (1992) to make it more appropriate for a very young child. Specifically, 8 rather than 12-16 items were used in the assessment thus allowing for the assessment to be shorter in duration. Results suggested that the infant's pattern of responding was stable across two forced-choice preference assessments administered one week apart and that those items identified as highly preferred served as reinforcers during a reinforcer assessment.

Didden and de Moor (2004) also conducted forced-choice preference assessments with young children. Of the twenty toddlers who served as participants, ten had mild developmental and physical disabilities and ten were non-disabled. Results from this study indicated that the forced-choice preference assessment was more effective in developing a rank ordering of preference for the group of developmentally delayed toddlers than for the typically developing toddlers. Differences between groups of toddlers rather than individual differences were examined making direct comparison of outcomes among these reported studies difficult. Didden and de Moor examined a specific group of toys and evaluated whether or not the group of participants were consistent in which toys were chosen as preferred. Results suggested that there was not a consistent pattern of responding for the group of typical toddlers, however, the preference assessment may have produced significant rankings at an individual level. In addition, reinforcer assessments were not conducted in this study; therefore it is difficult to hypothesize the validity of the assessment procedure for either the disabled or typically developing toddlers. Taken together, the data presented in the studies described above suggest the need for further investigation of this technique with young children with and without disabilities. Specifically, methodological modifications such as having a caregiver present, manipulating the type of toys used, and manipulating the number of items presented need to be further investigated.

Single Stimulus Assessment. In the single stimulus assessment, items are presented to the individual one at a time and then a rater assesses whether or not the child approaches the item (Pace, Ivancic, Edwards, Iwata, & Page, 1985). Percentage of trials in which each item is approached is calculated to determine preference ratings. In their initial study, Pace et al. presented 16 items, 10 times each across eight assessment sessions (in each session, four items were presented five times each). Based on data from reinforcer assessments, the items identified as preferred served as reinforcers. However, upon further evaluation of the procedure, Fisher et al (1992) noted that the procedure may result in an increase in false positive responding. That is, not all of the items that were identified as preferred were equally effective when evaluated via a reinforcer assessment. While the forced-choice assessment may produce a more valid ranking, the single stimulus procedure has been recommended for use with lower functioning individuals because this assessment does not require the individual to make a choice between items (Hagopian et al., 2004). The fact that choice skills are not needed for this procedure make it potentially useful for very young children as well, yet to date, there is no research on the utility of the single stimulus procedure with this population. When determining whether or not the single stimulus procedure would be appropriate, Thomson, Czarnecki, Martin, Yu, and Martin (2007) suggest to first determine if the individual is able to make a choice from a field of two or more items. If they are able to make choices, then one can choose among a variety of preference assessment options. However, if they are not able to choose from an array of items, then it would be most appropriate to use the single stimulus procedure.

Single Stimulus Engagement (SSE). In the single stimulus procedure described above, data are collected on whether or not the child approached the item. Another way to measure preference is to collect data on engagement. That is, how long did a child engage with the item? The single stimulus engagement procedure was designed to examine engagement rather than approach in an assessment similar to the single stimulus procedure technique. It was hypothesized that this modification would improve the ability to rank order items that are delivered individually (DeLeon, Iwata, Conners, & Wallace, 1999). In the SSE procedure, the individual is presented each item individually for a two-minute period (across five sessions for a total duration of 10 minutes). Data are collected on the percentage of time the participant manipulates each item (based on the 10 minute total). In the original study, seven tangible items were presented, thus the assessment took approximately 70 minutes per client (DeLeon et al., 1999). Results suggested that assessments that examined how long a child played with an item produced more differentiated results than assessments that just examined whether or not a child approached an item. However, the duration based assessment, like the approach based assessment, also resulted in a large number of items being identified as preferred (due to the child engaging with many of the items that were evaluated). Hagopian, Rush, Lewin, and Long (2001) also evaluated an SSE procedure in which 8-13 items were presented for a two-minute period across three assessment sessions (total duration 6 minutes). This assessment method was compared to the forced choice method. The forced choice method produced more differentiated rankings, however, the items identified as preferred in the SSE procedure also served as reinforcers when assessed via a reinforcer assessment. The main concern with this assessment method is that the individual may approach all of the stimuli thus resulting in false positives and the inability to identify the most preferred items. Studies on this method have been conducted with children and adults with developmental disabilities. Thus, to date, the effectiveness of this technique with the infant and toddler population is unknown. However, as with the single stimulus procedure, the single stimulus engagement procedure does not require the ability to make choices, thus it may be a useful technique for infants or toddlers who are unable to choose among items.

Multiple Stimulus with replacement (MS). In the multiple stimulus with replacement procedure, multiple items are presented to the participant simultaneously (Windsor et al., 1994). Once an item has been consumed, it is replaced prior to the next trial so that all items are presented in each trial. Results suggest that one benefit of the MS procedure is that it takes less time than the forced-choice procedure.

While both the forced choice and MS procedures resulted in identifying similar items as preferred, a limitation of MS procedure is that it resulted in less stable stimulus preference rankings when compared to the forced choice procedure. As noted with other procedures, data supporting the MS procedure with the infant and toddler populations is largely untested.

The Multiple Stimulus Procedure Without Replacement (MSWO). The Multiple Stimulus procedure was modified such that once the item was consumed it was not replaced (DeLeon, & Iwata, 1996). The MSWO procedure is similar to the MS procedure with regards to all other components, the exception being the absence of a replacement procedure. Before the assessment begins, the participant is given a sample of each edible item and 30seconds access to each leisure item. Items are placed in front of participant in a straight line with the prompt to pick one item. The participant has 30 seconds to select an item; once an item is selected the individual has 30 seconds access to the item. At the end of the 30 seconds, the item is removed from the assessment. Items are then rotated (taking the item from far left and placing it on the far right and then moving each item to control for positional effects). Once the items have been rotated the participant is told to choose another item. This procedure continues until all items have been selected or 30 seconds passed without an item being selected.

While the MSWO method produces valid results, it can also be lengthy depending on the number of items in the assessment. In an attempt to identify a more efficient method for identifying potential reinforcers, Layer et al. (2008) modified the procedure by assessing three preschool children simultaneously. The authors found that the group format was an efficient method for identifying valid reinforcers even with a delay between the selection and receipt of the reinforcer (the delivery of the reinforcer occurred after all children had chosen their reinforcer thus wait time ranged from an estimated 1-2 minutes). The children in this study were 3-5 years old. The efficacy of this modification with younger children is unknown and would be important to assess as infants and toddlers may not be able to wait for the delivery of the reinforcer. For purposes of assessing infants, one modification would be to reduce the number of items assessed. This would be important not only to decrease the length of the assessment but also because very young children may not have the ability to scan a large number of items. Thus, in addition to reducing the total number of items, one would also need to expose the child to each item prior to the beginning of the assessment (e.g. have the child engage with each item for 30 seconds in a pre-assessment session). Another modification to the MSWO procedure may be to decrease the number of stimulus presentations. For example, Carr, Nicholson & Higbee (2000) decreased the number of stimulus presentations from five to three. He found that the shortened presentation method produced stable and valid results for autistic children as young as two-years-old. Also, Carr et al. suggested that it may be possible to obtain similar results with presenting the items just once. Correlations were conducted between the data from the first presentation of the item and the data summarized for all three presentations and found that there was a high correspondence. That is, the percentage of approach for each item in the first session was similar to the average percentage of approach across all three sessions. This would allow for a brief assessment which would be of obvious benefit when assessing young children. A final modification Carr and colleagues made to this procedure was that the assessment was conducted in the natural environment. This modification may allow for a more accurate assessment of the reinforcing value of an item in that the item is being assessed in the actual setting in which it will be used. This should enhance the generalizability of the assessment results.

Brief Stimulus Preference Assessment/Free Operant Assessment. The brief stimulus preference assessment was developed to account for some of the criticism related to the length of other assessment methods (Roane et al., 1998). The free operant assessment is a five minute assessment during which the child has access to items arranged in a circle on the table. Prior to starting the assessment, the therapist walks the child around a table and has them manipulate each item to ensure the child is aware of each. Once the assessment begins the child is free to manipulate any of the items presented. Data are collected using a 10 second partial interval recording procedure which is then converted into the total percent of

intervals in which the items were manipulated. Roane et al. assessed the effectiveness of the procedure on individuals with developmental disabilities ranging from 3 years old to adulthood. They reported strong predictive validity based on results from reinforcer assessments. While this does not allow for a discrete ranking of the items it does have advantages to some of the other procedures. First, it is brief, which is an advantage when working with young children. Second, the items are never removed during the assessment, thus may result in fewer problem behaviors than assessments in which the items are removed from the child. As noted by Roane et al., a limitation of the free operant assessment includes the increased likelihood of satiation during the procedure. Because a child can have continuous access, they may lose interest in the activity if it is used immediately following the preference assessment as a reinforcer. This may be a particular concern for young children who may satiate at a faster rate than older children. However, to date, there are no studies examining this procedure with very young children and conclusions drawn on generalizability remain largely hypothetical.

Ortiz and Carr (2000) compared free operant and restricted operant procedures using 4-7 year old children with developmental disabilities as their participants. They found that the free operant assessment was only able to identify the most preferred items because the children never chose all the items (i.e. they often only played with 2-3 items) whereas the MSWO method produced a ranking of items. However, it should be noted that the highly preferred items identified by both procedures were similar and resulted in a sustainable reinforcer for each child. A benefit of this study is that it compared two well developed preference assessment methods within the natural environment and used functional behaviors in the reinforcer assessment. This would be important because often times the behaviors that are used in the reinforcer assessment have limited practical utility (Kennedy, 2002). Thus, an item may serve as a reinforcer when the child needs to emit a low effort response such as pressing a lever or standing in a square to access the item but if asked to produced an effortful response, the preferred item may no longer serve as a reinforcer. Given the time consuming and artificial nature of many of the preference assessments, recent investigations on the identification of potential reinforcers have moved into the natural setting using naturalistic observation techniques.

Observations. Reid, DiCarlo, Schepis, Hawkins, and Stricklin (2003) developed an observational assessment to determine toy preference. Young children (27 months- 45 months old) with developmental delays served as participants. Reid and colleagues compared varying length (e.g., 5, 10, 15 minute) observation assessments and found that five 5-min observations produced reliable and valid results. Observations took place during free play times in class once per day for five minutes. Data were collected on which toys were played with via a 15-second interval system across the 5 minute observation. Percentage of intervals in which an item was played with was then calculated to determine a ranking of preferred items. The benefit of this procedure is that it takes place in the natural environment without disrupting the day to day functioning of the classroom. While the children in this study had developmental disabilities, future research should look at children without disabilities.

Hanley and colleagues (2007) developed a similar observational procedure to assess preference for events naturally occurring in the preschool classroom environment. Hanley et al. used a momentary time sampling procedure to simultaneously observe multiple preschool children in a classroom that served children of typical and atypical development. The researchers originally compared 30, 60, 90, and 120 second momentary time sampling intervals and found that the 90-second interval system was more preferred by the raters and produced results similar to the longer observations. The procedure allows for up to 20 children to be observed at one time. A paper and pencil data collection system was used in which each child's name was listed with a time and a place to record data. Stopwatches were used to determine when to record data for a specific child. Data were collected on in-zone/engagement behavior. Percentage of time in-zone was then calculated based on the number of intervals a child was in a specific zone divided by the total number of observation intervals.

The benefit of the two observation systems described above is that they are more time efficient than other preference assessment procedures (e.g. more than one child can be observed at a time over a brief period), they are conducted in the natural setting (thus theoretically it should enhance generalization), and permit the identification of reinforcers already present in the natural environment.

Informal observations have also been used to identify reinforcers. Specifically, a teacher observes what the child is currently playing with and then uses it as a reinforcer to get the child to complete nonpreferred tasks. Unfortunately, there are no data on the accuracy, reliability, or validity of this method. However, this method is widely used in the school setting therefore systematic evaluations of the effectiveness of this technique are greatly needed.

Practical Considerations

When conducting assessments with very young children, there are many variables to consider in identifying the most appropriate method for a specific child. The wide use of methods being implemented without empirical support suggest that many assume measures found to be effective for individuals with a specific mental age are therefore appropriate for use with children of the same chronological age. However, recent research has shown that such as assumption in not accurate. Regarding preferences or reinforcing properties of various items, Didden and de Moor (2004) reported that significant differences exist between children with developmental disabilities and their typically developing peers, even if they are functioning at the same mental age. Their work indicates that these groups demonstrate specific differences in their preferences. Developmentally disabled toddlers preferred dynamic toys while typical peers did not show a similar preference. They also found differences in the effectiveness of specific assessment methods (e.g. they found the forced-choice assessment to be effective in rank ordering preferred items for a group of toddlers with developmental disabilities but not for a group of typically developing peers). These findings indicate that we cannot automatically generalize the validity of assessments designed for individuals with developmental disabilities to individuals without disabilities. Below is a description of specific variables to consider when identifying a preference assessment for a very young child.

Child Variables

Choice Making Skills. First, the child's ability to scan a field of items and make choices should be determined prior to conducting a preference assessment. If a child is not capable of scanning multiple items or is not able to choose among items, a single stimulus approach or a naturalistic observation would most likely be the assessment method of choice. However, if the child is able to scan and choose among two or more items, then a forced-choice or multiple stimulus assessment may be used. Typically, imitation and choice-making abilities are assessed indirectly. However, a more formal assessment may provide more accurate information. For example, the Assessment of Basic Learning Abilities Test (ABLA) has been used to assess whether or not a child can engage in simple imitation and two-choice discrimination tasks (Kerr, Meyerson, & Flora, 1977; Martin, Thorsteinsson, Yu, Martin, & Vause, 2008). On the ABLA, there are six levels in the assessment including: imitation, position discrimination, visual discrimination, match-to-sample, auditory discrimination, and auditory-visual combined discrimination. The ABLA has been found to be effective in predicting how individual's perform on future imitation and two-choice tasks (Martin et al., 2008). To date, this research has been conducted on individuals with developmental disabilities, children with pervasive developmental disabilities, and typically developing children (Martin et al., 2008). However, the youngest children assessed with this measure have been four years old. Therefore, the validity of the measure in very young children is not known. While this appears to be a promising measure, further research is needed before it can be recommended for use with infants and toddlers.

Attention Span. The child's attention span should also be taken into consideration. When working with infants, it is important to consider the length of the assessment. Some assessments have been specifically developed to be brief (e.g. free operant assessment & observations), while others can be modified to be shorter in duration (e.g. forced choice & MSWO). In general, assessments ranging from 5-15 minutes may be appropriate. However, children with particularly short attention spans may be best assessed using the briefest assessments such as the free operant procedure or through observation techniques (Hanley et al., 2007; Reid et al., 2003).

Problem Behaviors. The child's level of problem behaviors and ability to handle transitions or changes in the environment should be considered. Children who engage in problem behaviors that have a tangible function may do best with assessment methods that do not involve the removal of items. For example, Roane et al. (1998) reported a lower frequency of problem behavior using the free-operant procedure than other methods where items are removed throughout the assessment. Children who have difficulty with transitions or changes in the environment may also benefit from a procedure in which there is little disruption. Specifically, the multiple stimulus, free-operant or naturalistic assessments may be best for this population. Naturalistic assessments allow for the least disruption to the child's environment and may therefore produce the most valid results for children who experience distress under such conditions. A summary of these considerations can be found in Table 1.

Table 1
Measures Appropriate Based on Skill Level

Skill	Skill Level Adequate	Skill Level Deficient
Choice Making	**FC, FOA, MS, MSWO**	**SS, SSE**
Attention Span	**FC, MS, MSWO**	**FOA, NO**
Problem Behaviors	**SS, SSE, PC, MS, MSWO**	**FOA, NO**

Note. Table 1 represents which measures are most appropriate when specific skills are adequate or deficient. FC= Forced choice, FOA= free operant, MS= multiple stimulus, MSWO= multiple stimulus without replacement, SS= single stimulus, SSE= single stimulus engagement, NO= naturalistic observation

Environmental Variables

Type of preferred items. Research conducted by Kurtz and colleagues (2003) provides evidence that items preferred by very young children may be unique to this population. Data from the assessment and treatments of 30 infants and toddlers (ranging from 10 months- 4 .11 years old) with severe SIB suggested that almost half of these participants engaged in inappropriate behaviors to gain access to parent attention or preferred items. Thus, Kurtz et al. concluded that social reinforcers such as toys and parental attention (e.g. verbal attention in the form of soothing statements and use of motherese; and physical attention in the form of being held, rocked, or hugged) are potent reinforcers for infants and toddlers. While attention is often assessed as a potential reinforcer, the type of attention would be important to consider with very young children. Specifically, hugs and rocking may be more reinforcing than playing, and giving praise using a high pitched motherese voice may be more reinforcing than praise given in a more typical prosody.

Establishing Operations. The effects of establishing operations on the reinforcing value of a preferred item have long been established and may be even more important to examine in young children given their short attention spans and the transient nature of reinforcers (Berg et al., 2000; Carr, Nicholson, & Higbee, 2000; Green, Reid, Canipe, & Gardner, 1991; Kurtz et al., 2003). Kurtz et al. (2003) reported

that a large amount of a young child's day is spent engaging in play and interacting with caregivers thus the reinforcing value of activities might vary significantly across the day as satiation and deprivation vary.

Berg and colleagues (2000) systematically evaluated the effects of satiation and found that pre-session exposure to attention altered the effectiveness of attention as a reinforcer. The participants in this study were 22-50 months old. It would be interesting to see if this effect is similar for younger children. Given the significance of parental attention for infants and toddlers, it would be important to assess this population as well.

McAdams et al. (2005) evaluated the effects of satiation and deprivation on the results of forced-choice preference assessments for three individuals with developmental disabilities and three typically developing preschoolers. McAdams et al. found that deprivation and satiation have significant effects on which items were chosen during the preference assessment. However, for some participants, the effects may not have been larger than what is often seen due to natural fluctuation. Earlier studies have also suggested that satiation can have a significant influence on the reinforcing value of preferred items (DeLeon et al., 2001; Rush et al., 2005). DeLeon and colleagues found that when one item was used in a 30 minute treatment session, the item ceased to serve as a reinforcer as early as 10 minutes into the session. However, when the experimenters rotated between two reinforcer sets every 10 minutes, the reinforcing value of the items was maintained across the treatment session. The participant in the DeLeon study was an 11 year old girl with developmental disabilities. Rush et al. (2005) found similar results with a 13-month-old. The authors reported that the reinforcing effects of the item identified as the most preferred item diminished across the reinforcer assessment sessions. Specifically, the infant was engaged in the task associated with the highly preferred item for 90% of intervals during the 1st session of the reinforcer assessment but only 45% of the time during the 3rd session (three 10 minute sessions were conducted back to back). On the second day of the assessment, the infant was engaged in the task associated with the highly preferred item for 70% of the intervals during the 1st session and only 30% of the intervals during the 3rd session. These fluctuations were also attributed to satiation. Based on these findings, one recommendation is to rotate among multiple reinforcers (DeLeon et al., 2001; Rush et al., 2005). However, it is also recommended that more research be conducted with infants and toddlers on the effectiveness of constant versus varied or multiple reinforcers (Rush et al., 2005).

Generalizability. One area that must be considered is the degree to which the responses of infants and toddlers are influenced by the environment within which they are assessed. One reason why this variable is important to consider is that many assessments take place outside a child's everyday environment. Young children may be particularly susceptible to being influenced by the environment in which the assessment takes place. For example, where and with whom the assessment is conducted may influence the results. Bruzek & Thompson (2007) found that observing a peer play with an item altered the reinforcing value of a preferred item. Specifically, they found that when a child observed a peer playing with an item they had nominated as low to moderately preferred, the reinforcing quality of the item was altered in a subsequent reinforcer assessment. Specifically, the child allocated more responses to the moderately preferred item than they had originally. This is an important variable to consider as most of our assessments take place outside of the natural environment and thus may have limited generalizability (Kennedy, 2002).

Hanley et al. (2007) also suggests that subtle changes such as where a book is read may alter the reinforcing value of an item. For example, reading a book while laying on beanbag in the classroom may be more reinforcing than sitting in chair in experimental setting. Haynes (2002) reported similar concerns. In this study, it was found that while the child identified similar items as highly preferred in preference assessments conducted by a parent or novel therapist, what the child ultimately chose as a reinforcer during treatment depended on with whom the child was working. When the parent served as the therapist, the child chose to work for immediate reinforcers. However, when a novel person served as

the therapist, the child chose to work for delayed reinforcers. Thus it appears that who serves as the therapist may serve as an establishing operation (increasing or decreasing the reinforcing quality of an item) or as a discriminative stimulus (signaling the availability of a specific reinforcer). It may be that the child finds a certain item more reinforcing with specific individuals or they may have learned that certain individuals do not follow through with delayed reinforcers and therefore are more likely to pick something available immediately.

Since novel therapists are often the ones conducting assessments, these results may not generalize to the natural setting or to other individuals. Again, the research on infants and toddlers is limited. Kurtz et al. (2003) suggest training parents to serve as therapists as they will be the ones serving as the change agents. In addition, because many children show distress when separated from their parents at this age, the results of the assessments may be invalid. While Hanley et al. (2007) used preschoolers; the participant in the Haynes (2002) study was a school aged boy with ADHD. Thus, more research is needed with the younger population.

Task Demands. A final consideration is to examine the task for which the preferred item needs to serve as a reinforcer. Type of task, task duration, and preference for the task all influence the reinforcing value of the preferred item. Specifically, items that are identified as preferred may not serve as reinforcers (Logan & Gast, 2001). Therefore, variables that influence the effectiveness of a potential reinforcer need to be evaluated. For example, Thompson and colleagues (2004; 2007) found that parent and teacher report were adequate in identifying items that served as reinforcers when teaching typically developing infants sign language. These items appeared to serve as reinforcers in that increases in the target operant response were observed. However, it may be that this task required minimal response effort thus it was not necessary to identify a highly potent reinforcer. Yet, when attempting to get a young child with a feeding disorder to take a bite of a nonpreferred food or when attempting to get a child with autism to sit and comply with tasks presented during an intensive early intervention program, a more potent reinforcer may be needed as these tasks may be viewed by the individual as aversive and requiring a high level of effort. In this situation, it would be necessary to use a more direct measure of preference. More research is needed on such variables and how they may influence preference assessment results for very young children with and without disabilities.

Conclusion

Based on the limited research currently available on the use of preference assessments with infants and toddlers, it is recommended that indirect assessments not be used in isolation. Direct and observational assessments produce more reliable and valid results and can be easily modified for use with very young children. Child variables to consider when choosing an assessment method include the child's ability to scan and choose among multiple items, attention span, engagement in problem behaviors and ability to transition and tolerate changes in environment. Environmental variables to consider include the task for which the item must serve as a reinforcer, the effects of establishing operations and discriminative stimuli on the assessment process, and the generalizability of the results.

There are a large number of issues that warrant further investigation. First, few preference assessment procedures have been empirically evaluated with very young children and almost no research exists with the infant population; thus more research examining the effectiveness of each assessment method is needed. In addition, the child and environmental variables described above have limited empirical support and warrant further investigation. Without such investigations, it is difficult to say what types of specific modifications may be necessary for this population. Given the current emphasis on early identification and intervention, this research is essential for developing empirically derived assessment methods for very young children.

References

Akin-Little, K., Eckert, T., Lovett, B., & Little, S. (2004). Extrinsic reinforcement in the classroom: Bribery or best practice. *School Psychology Review, 33,* 344-362.

Berg, W., Peck, S., Wacker, D., Harding, J., McComas, J., Richman, D., & Brown, K. (2000). The effects of pre-session exposure to attention on the results of assessments of attention as a reinforcer. *Journal of Applied Behavior Analysis, 33,* 463-477.

Bruzek, J. & Thompson, R. (2007). Antecedent effects of observing peer play. *Journal of Applied Behavior Analysis, 40,* 327-331.

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

Chao, P., Bryan, T., Burstein, K. & Ergul, C. (2006). Family-centered intervention for young children at-risk for language and behavior problems. *Education Journal, 34,* 147-153.

Cote, C., Thompson, R., Hanley, G., & McKerchar, P. (2007). Teacher report and direct assessment of preferences for identifying reinforcers for young children. *Journal of Applied Behavior Analysis, 40,* 157-166.

Daly, E., Wells, N., Swanger-Gange, M., Carr, J., Kunz, G, & Taylor, A. (2009). Evaluation of the multiple-stimulus without replacement preference assessment using activities as stimuli. *Journal of Applied Behavior Analysis, 42,* 563-574.

Damon, S., Riley-Tillman, T., & Fiorello, C. (2008). Comparing methods of identifying reinforcing stimuli in school consultation. *Journal of Educational and Psychological Consultation,* 18, 31-53.

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

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

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

Didden, R. & de Moor J. (2004). Preference assessment in toddlers with mild developmental and physical disabilities: A comparative study. *Journal of Developmental and Physical Disabilities, 16,* 107-116.

Ellison, D. (1997). Compliance training decreases maladaptive behaviors in two people with developmental disabilities: Report of five years' treatment. *Behavioral Interventions, 12,* 183-194.

Fantuzzo, J., Rohrbeck, C., Hightower, A., & Work, W. (1991). Teacher's use of and children's preferences of rewards in elementary school. *Psychology in the Schools, 28,* 175-181.

Fisher, W., Piazza, C., Bowman, L., Hagopian, L., Owens, J., & 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., Piazza, C., Bowman, L., & Amari, A. (1996). Integrating caregiver report with a systematic choice assessment. *American Journal on Mental Retardation, 101,* 15-25.

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

Green, C., Reid, D., White, L., Halford, R., Brittain, D. & Gardner, S. (1988). Identifying reinforcers for persons with profound handicaps: Staff opinion versus systematic assessment of preferences. *Journal of Applied Behavior Analysis, 21,* 31-43.

Gresham, F. M., Beebe-Frankenberger, M. E., & MacMillan, D. L. (1999). A selective review of treatments for children with autism: Description and methodological considerations. *School Psychology Review, 28,* 559-575.

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

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

Hanley, G., Cammilleri, A., Tiger, J., Ingvarrson, E. (2007). A method for describing preschoolers activity preferences. *Journal of Applied Behavior Analysis, 40,* 603-618.

Haynes, R., Derby, M., McLaughlin, T., & Weber, K. (2002). A comparison of forced-choice preference assessment procedures using a parent and novel therapist. *Journal of Positive Behavioral Interventions, 4,* 176-181.

Individuals with Disabilities Education Improvement Act. Pub. L. No. 108-446. Stat. 2647 (2004).

Kennedy, C. (2002). Toward a socially valid understanding of problem behavior. *Education and Treatment of Children, 25,* 142-153.

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

Kurtz, P., Chin, M., Huete, J., Tarbox, R., O'Connor, J., Paclawskyj, T., & Rush, K. (2003). Functional analysis and treatment of self-injurious behavior in young children: A summary of 30 cases. *Journal of Applied Behavior Analysis, 36,* 205-219.

Lannie, A. & McCurdy, B. (2007). Preventing disruptive behavior in the urban classroom: Effects of the Good Behavior Game on student and teacher behavior. *Education and Treatment of Children, 30,* 85-98.

Layer, S., Hanley, G., Heal, N., & Tiger, J. (2008). Determining individual preschoolers' preferences in a group arrangement. *Journal of Applied Behavior Analysis, 41,* 25-37.

Logan, K. & Gast, D. (2001). Conducting preference assessments and reinforcer testing for individuals

with profound multiple disabilities: Issues and procedures. *Exceptionality, 9*, 123-134.

Luxem, M., Christophersen, E., & Purvis, P. (1997). Behavioral-medical treatment of pediatric toileting refusal. *Journal of Developmental and Behavioral Pediatrics*, *18*, 34-41.

McAdams, D., Klatt, K., Koffarnus, M., Dicesare, A., Solberg, K., Welch, C., & Murphy S. (2005). The effects of establishing operations on preferences for tangible items. *Journal of Applied Behavior Analysis, 38*, 107-110.

Martin, G., Thorsteinsson, J., Yu, C., Martin, T., & Vause, T. (2008). The assessment of basic learning abilities test for predicting learning of persons with intellectual disabilities. *Behavior Modification, 32*, 228-247.

Matson, J., Bielecki, J., Mayville, E., Smalls, Y., Bamburg, J., & Baglio, C. (1999). The development of a reinforce choice assessment scale for persons with severe and profound mental retardation. *Research in Developmental Disabilities, 20*, 379-384.

Northrup, J. (2000). Further evaluation of the accuracy of reinforcer surveys: A systematic replication. *Journal of Applied Behavior Analysis*, *33*, 335-338.

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

Northrup, J., Jones, K., Broussard, C., & George, T. (1995). A preliminary comparison of reinforcer assessment methods for children with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis*, *28*, 99-100.

Ortiz, K & Carr, J. (2000). Multiple-stimulus preference assessments: A comparison of free-operant and restricted-operant formats. *Behavioral Interventions, 15*, 345-353.

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

Parson, M. & Reid, D. (1990). Assessing food preferences among persons with profound mental retardation: Providing opportunities to make choices. *Journal of Applied Behavior Analysis, 23*, 183-195.

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

Piazza, C., Patel, M., Gulotta, C., Sevin, B., & Layer, S. (2003). On the relative contributions of positive reinforcement and escape extinction in the treatment of food refusal. *Journal of Applied Behavior Analysis*, *36*, 309-324.

Reid, D., DiCarlo, C., Schepis, M., Hawkins, J., & Stricklin, S. (2003). Observational assessment of toy preferences among young children with disabilities in inclusive settings: Efficiency analysis and comparison with staff opinion. *Behavior Modification, 27*, 233-250.

Resetar, J. & Noell, G. (2008). Evaluating preference assessments for use in the general education population. *Journal of Applied Behavior Analysis, 41*, 447-451.

Risely, T. & Hart, B. (1968). Developmental correspondence between the nonverbal and verbal behavior of preschool children. *Journal of Applied Behavior Analysis,1*, 267-281.

Roane, H., Vollmer, T., Ringdahl, J., & Marcus, B. (1998). Evaluation of a brief stimulus preference assessment. *Journal of Applied Behavior Analysis*, *31*, 605-620.

Rush, K., Kurtz, P., Leiblein, T., & Chin, M. (2005). The utility of a paired-choice preference assessment in predicting reinforcer effectiveness for an infant. *Journal of Early and Intensive Behavior Intervention*, *2*, 247-250.

Schanding, G., Tingstrom, D., & Sterling-Turner, H. (2009). Evaluation of stimulus preference assessment methods with general education students. *Psychology in the Schools*, *46*, 89-99.

Thomson, K., Czarnecki, D., Martin, T., Yu, C. & Martin, G. (2007). Predicting optimal preference assessment methods for individuals with developmental disabilities. *Education and Training in Developmental Disabilities*, *42*, 107-114.

Thompson, R., Cotnoir-Bichelman, N., McKerchar, P., Tate, T. & Dancho, K. (2007). Enhancing early communication through infant sign training. *Journal of Applied Behavior Analysis, 40*, 15-23.

Thompson, R., McKerchnar, P., & Dancho, K. (2004). The effects of delayed physical prompts and reinforcement on infant sign language acquisition. *Journal of Applied Behavior Analysis, 37*, 379-383.

Tiger, J., Hanley, G., & Hernandez, E. (2006). An evaluation of the value of choice with preschool children. *Journal of Applied Behavior Analysis, 39*, 1-16.

Webster-Stratton, C. & Reid, J. (2003). Treating conduct problems and strengthening social and emotional competence in young children: The dina dinosaur treatment program. *Journal of Emotional and Behavioral Disorders, 11*, 130-143.

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

Author Contact Information:

Karena S. Rush, Ph.D.
Millersville University
Department of Psychology
Millersville, PA 17551
krush@millersville.edu

Bruce P. Mortenson, Ph.D.
Towson University
Department of Psychology
Towson, MD 21252
bmortenson@towson.edu