TEACHER REPORT AND DIRECT ASSESSMENT OF PREFERENCES FOR IDENTIFYING REINFORCERS FOR YOUNG CHILDREN

CATHERINE A. COTE, RACHEL H. THOMPSON, GREGORY P. HANLEY, AND PAIGE M. MCKERCHAR

UNIVERSITY OF KANSAS

Teachers were asked to identify and rank 10 preferred stimuli for 9 toddlers, and a hierarchy of preference for these items was determined via a direct preference assessment. The reinforcing efficacy of the most highly preferred items identified by each method was evaluated concurrently in a reinforcer assessment. The reinforcer assessment showed that all stimuli identified as highly preferred via the direct preference assessment and teacher rankings functioned as reinforcers. The highest ranked stimuli in the direct assessment were more reinforcing than the teachers' top-ranked stimuli for 5 of 9 participants, whereas the teachers' top-ranked stimulus was more reinforcing than the highest ranked stimulus of the direct assessment for only 1 child. A strong positive correlation between rankings generated through the two assessments was identified for only 1 of the 9 participants. Despite poor correspondence between rankings generated through the teacher interview and direct preference assessment, results of the reinforcer assessment suggest that both methods are effective in identifying reinforcers for toddlers.

DESCRIPTORS: concurrent operants, direct assessment, indirect assessment, preference, reinforcer assessment, teacher report, toddlers

---

Early childhood education is heavily influenced by the guidelines for developmentally appropriate practice outlined by the National Association for the Education of Young Children (NAEYC; Bredekamp & Copple, 1997). According to these guidelines, curriculum is considered developmentally appropriate to the extent that it is individualized to address the needs and interests of each child. Thus, planned activities should be developed based on knowledge of the children's skill level and interests. There a number of resources (e.g., Bricker, Pretti-Frontczak, Johnson, & Straka, 2002; Hills, 1992) available to guide educators' selection of tools for assessing children's skills, but there has been relatively little focus on developing methods for assessing children's preferences.

The NAEYC also recognizes that it may be necessary to use a variety of teaching strategies to promote learning, including differential reinforcement (Bredekamp & Copple, 1997). Therefore, early childhood educators may implement reinforcement programs to promote acquisition of age-appropriate skills (e.g., continence, Simon & Thompson, 2006; social skills, Zanolli, Paden, & Cox, 1997). However, the literature says little about what strategies early childhood educators use to select reinforcers, and only a few behavior-analytic studies have described preference assessments developed for young children in early education (see, e.g., Hanley, Cammilleri, Tiger, & Ingvarsson, 2005; Reid, DiCarlo, Schepis, Hawkins, & Stricklin, 2003).

By contrast, extensive research has focused on identifying preferences of individuals with developmental disabilities (e.g., Carr, Nicholson, & Higbee, 2000; Conyers et al., 2002; DeLeon & Iwata, 1996; Fisher et al., 1992; Hanley, Iwata, Lindberg, & Conners, 2003; Roane, Vollmer, Ringdahl, & Marcus, 1998), and these

studies have collectively resulted in an efficient technology for identifying reinforcers. This technology was developed primarily to address the difficulties associated with determining preferences among individuals with limited language; thus, systematic preference assessments may also be useful in identifying preferences of young, typically developing children who may be unable to report their preferences accurately. However, the generality of this technology has not been thoroughly evaluated with respect to young children of typical development.

Given the limited resources available to early childhood educators, it appears that caregiver reports of preference (indirect assessments) rather than direct preference assessments are used to identify reinforcers. However, research conducted with individuals with developmental disabilities has found poor correspondence between results of caregiver reports of preference and direct preference assessments (Green et al., 1988; Reid, Everson, & Green, 1999; Windsor, Piche, & Locke, 1994). Perhaps a more important finding is that direct assessments of preference have been associated with superior predictive validity (i.e., direct assessments more effectively identify reinforcers relative to indirect assessments; Green et al.; Parsons & Reid, 1990). We sought to evaluate the generality of these findings by conducting similar comparisons of indirect and direct assessments of preferences with young children in center-based care.

We implemented two frequently used preference assessment procedures: a caregiver interview similar to that described by Fisher, Piazza, Bowman, and Amari (1996), and a paired-stimulus preference assessment (Fisher et al., 1992). The purposes were to determine if these methods would be effective in identifying reinforcers for young children and which would identify more potent reinforcers. More specifically, we compared rankings from each assessment to determine the extent of agreement

between the indirect and direct assessments for 9 young children of typical cognitive development. We then evaluated the reinforcing efficacy of top-ranked stimuli from each assessment type.

METHOD

Participants and Setting

Nine children who attended a university-based toddler classroom participated. All participants were between the ages of 18 and 29 months and were selected based on availability (i.e., regular attendance) and parental consent. Eight children were typically developing, and 1 child had been diagnosed with a severe left-sided physical impairment. All children followed some simple one-step instructions and made simple requests. Preference assessments were conducted in a small room (3 m by 2.4 m) adjacent to the children's classroom that was equipped with a one-way observation window.

Data Collection and Interobserver Agreement

Trained undergraduate and graduate students served as experimenters and data collectors. Observers manually recorded (i.e., with paper and pencil) selections made by the participants during the direct assessment of preferences. A selection was recorded if the participant touched an item within 5 s of the initial instruction, "pick one." A second observer simultaneously but independently collected data during an average of 67% of trials (range, 31% to 100% across participants). An agreement was defined as both observers recording the same response during a trial. Agreement scores were calculated by dividing the number of agreements plus disagreements and multiplying by 100%. Mean agreement for occurrence of selection responses was 100% for all participants.

During the reinforcer assessment, the frequency of switch pressing or duration of in-zone behavior was recorded within 10-s intervals

using handheld computers. A switch press was defined as the participant touching any part of the switch plate. In-zone behavior was recorded when the majority of the participant's body was inside a particular zone (1 m by 2 m) that was marked on the floor with masking tape. Interobserver agreement was assessed during a mean of 38% of sessions (range, 30% to 50% across participants). Agreement percentages were calculated by comparing observers' records on an interval-by-interval basis. The smaller number of responses (or duration of the response) in each interval was divided by the larger number; these fractions were then averaged across intervals and multiplied by 100% to obtain a percentage agreement score. The mean agreement across participants was 99% (range, 97% to 100%).

Procedure

Teacher interview. During a structured interview, three groups of three to four teachers were asked to generate a list of potential reinforcers for each participant. Prior to the interview, the selected teachers had worked with the participants in the toddler classroom during daily 3- or 4-hr shifts for a mean of 25 weeks (range, 6 to 43 weeks for individual teachers). The structured interview (available from the second author) was based on the Reinforcer Assessment for Individuals with Severe Disabilities (Fisher et al., 1996), but was modified for use with typically developing toddlers. Specifically, prompt questions about items that provided visual, olfactory, edible, and tactile stimulation were excluded; only questions about toys and auditory and social stimuli were included. During the interview, each group of teachers was asked by an experimenter to list a total of 10 preferred stimuli from toy, audible, and social domains. The group was then instructed to rank the 10 stimuli from most to least preferred; thus, rankings were based on teacher consensus. Consensus was achieved when all teachers verbally agreed on a rank for

Table 1

Stimuli Identified by Teachers and Used in Direct Assessments of Preference

Code	Stimuli
A	Balls (e.g., basketball, football)
B	Blocks (e.g., plastic, wooden)
C	Books (e.g., small, large, theme)
D	Bubbles
E	Coloring (e.g., paint, chalk, markers)
F	Dress-up items (e.g., hats, purses)
G	Food or dishes (e.g., pizza, bowls)
H	Mammal toys (e.g., babies, lions)
I	Musical Instruments (e.g., drum, bells)
J	Physical activities (e.g., dancing, peek-a-boo)
K	Physical interactions (e.g., hugs, tickles)
L	Play-Doh®
M	Popbeads
N	Puzzles
O	Shape sorter
P	Singing songs
Q	Tools or construction items (e.g., hammer, street cones)
R	Transportation toys (e.g., trains, trucks)
S	Dramatic play items (shopping basket, post office)

each item. The total duration of the interview ranged from approximately 10 to 20 min.

Direct preference assessment. The 10 items generated through the teacher interview (see Table 1) were then presented in pairs using the paired-stimulus preference assessment described by Fisher et al. (1992). Immediately prior to the direct preference assessment, exposure trials were conducted during which the participant was physically guided to select and manipulate each of the 10 items when presented singly. The 10 items were then presented in pairs, with each item paired once with every other item in a randomized sequence. Social stimuli (e.g., singing) were represented on cards (0.15 m by 0.15 m) (e.g., umbrella with raindrops for the raindrop song). During each trial, the two items were arranged 0.6 m apart in front of the participant. Each independent selection was followed by 15-s access to the selected item and continuous experimenter attention. Throughout all phases of the experiment, attention was delivered during stimulus presentations because participants were accustomed to a high level of adult interaction and typically did not display independent play skills. Simultaneous approach to both items was blocked by the experimenter.

If no selection was initiated within 5 s, the experimenter physically guided the participant to sample both items (no data were collected during this presentation). The two stimuli were then represented for an additional 5 s. If a selection was not made within 5 s, the stimuli were withdrawn and the next trial began.

Once the direct assessment was completed, the items were ranked from highest to lowest based on calculating the percentage of trials on which each stimulus was chosen (with 1 being assigned to the item with the highest selection percentage of trials). If the percentage of trials selected for two or more stimuli was equal, stimuli were ranked according to the outcome of trials on which the relevant stimuli were presented together. For example, if the percentage of trials on which a book and Play-doh® were selected was 80%, we examined the trial on which these two stimuli were compared with each other. If the book was selected over the Play-doh® on this trial, the book received the higher ranking. If distinct ranks could not be established through this method (e.g., the child did not choose either stimulus), the same rank was applied to each stimulus. For example, two items tied for third would each receive a rank of 3 and the next ranked stimulus would receive a rank of 5. The approximate duration of the direct preference assessments was 40 to 50 min per child, and each direct preference assessment was completed in one session.

Reinforcer assessment. A reinforcer assessment was conducted to determine the reinforcing effectiveness of the top-ranked items identified by teachers relative to the top-ranked stimuli based on the direct assessments of preference. The time between the preference assessments and reinforcer assessment varied (see Table 2) based on participant availability (e.g., some children were absent due to illnesses) and program schedules (e.g., a semester break interrupted the analysis for some children). For all participants, the reinforcing effectiveness of items ranked first by the teachers (TI1) and

Table 2
Days on Which Each Assessment Was Initiated, with All Teacher Interviews Beginning on Day 1

Child	Direct preference assessment	Reinforcer assessment
Alex	4	5
Alice	3	4
Kitty	5	28
Mack	4	6
Jack	7	8
Lance	2	5
Odella	13	34
Josh	5	22
Milton	6	25

the direct preference assessments (DA1) was assessed. For Milton, the first and second most highly ranked items were also evaluated because the teacher interview and direct preference assessment identified the same item (basketball) as most highly preferred. All sessions were 5 min in length. Prior to each session, the experimenter physically guided the participant to perform all available target responses and provided the designated consequences for 15 s. At the start of each session, the experimenter gave the verbal prompt, "pick one."

Milton and Josh were presented with three identical microswitches that were placed on the floor in front of them. Each switch press produced 15-s access to the corresponding stimulus and experimenter attention, or no programmed consequences (control). Items were randomly assigned to the three switches and placed approximately 0.3 m behind the appropriate microswitches for each session.

For the 7 remaining participants, the session room contained three experimental zones (1 m by 2 m) that were marked on the floor with tape. Two zones contained the highly preferred stimuli generated from the teacher interview and direct assessment (TI1 and DA1), and one zone was empty and served as the control. In-zone behavior resulted in continuous access to the corresponding stimuli and experimenter attention, or no programmed consequences (control). Stimuli were randomly assigned to zones prior to each session.

A concurrent-operant schedule was used to evaluate the reinforcing effectiveness of the items during the reinforcer assessment. During the first phase, three microswitches or zones were available concurrently, and two items (TI1 and DA1) were compared with a no-stimulus control. Once a reinforcement effect for one stimulus was consistently observed, that stimulus was eliminated from subsequent sessions to determine the reinforcing effectiveness of the remaining stimulus. Thus, during the second phase, only one item was compared with the no-stimulus control. For Lance and Odella, a second phase was not conducted because these participants allocated responding equally to each stimulus during the initial phase of the reinforcer assessment.

RESULTS

Figures 1, 2, and 3 show the rankings from both the teacher interview and direct assessment of preference. These graphs illustrate the differences in the relative rankings of specific items for each child. Large differences are depicted by relatively long vertical bars, whereas consistent rankings between assessments are depicted by overlapping data points or relatively short bars. The most consistent outcome between assessments was observed for Odella (Figure 3), for whom little distance was observed between data points for each item, signifying comparable preference rankings of stimuli across teacher report and direct preference assessment. By contrast, the most inconsistent outcomes across the two types of preference assessments were observed for Alice (Figure 1) and Jack (Figure 2), for whom consistently large differences were observed for items at the ends of each preference hierarchy. Unsystematic differences in assessment outcomes were observed for the remaining 6 children.

Spearman rank-order correlation coefficients comparing rankings of stimuli generated from the teacher interview and direct preference

assessment are also indicated in Figures 1, 2 and 3. Positive correlations were found for 4 of the 9 children. Of those children, only one strong positive correlation was evident ($r_s = .73$ for Odella). Strong negative correlations were evident with 2 participants ($r_s = -.89$ and $-.68$ for Alice and Jack, respectively); the more highly preferred the teachers rated an item for Alice and Jack, the less likely it was shown to be preferred during the direct assessments.

Figures 1, 2, and 3 also show the results of the reinforcer assessments. The most highly preferred item from the direct assessment (DA1) was relatively more reinforcing than the item ranked highest by teachers (TI1) for 5 children (Alex, Alice, Kitty, Mack and Jack; Figures 1 and 2). By contrast, the teacher interview item (TI1) was relatively more reinforcing than the item from the direct assessment (DA1) for 1 child (Josh, Figure 3), and items from both assessments (DA1 and TI1) were found to be similarly reinforcing for 2 children (Odella and Lance; Figures 2 and 3). Milton's teacher interview and direct preference assessment identified the same item (basketball) as most highly preferred, and this item was shown to be relatively more reinforcing compared to his second most preferred items from both assessments. Further analysis indicated that the second ranked teacher interview item (TI2) was relatively more reinforcing than the second ranked item from the direct assessment (DA2). Finally, when the more highly reinforcing item was removed from the reinforcer assessments for 7 children, the remaining items functioned as reinforcers for all participants.

DISCUSSION

We conducted teacher interviews and direct assessments to identify preferences of young children in an early education setting. Small groups of teachers identified and ranked 10 items; these same 10 items were then evaluated in a paired-stimulus preference assessment (Fisher et al., 1992). A comparison of the
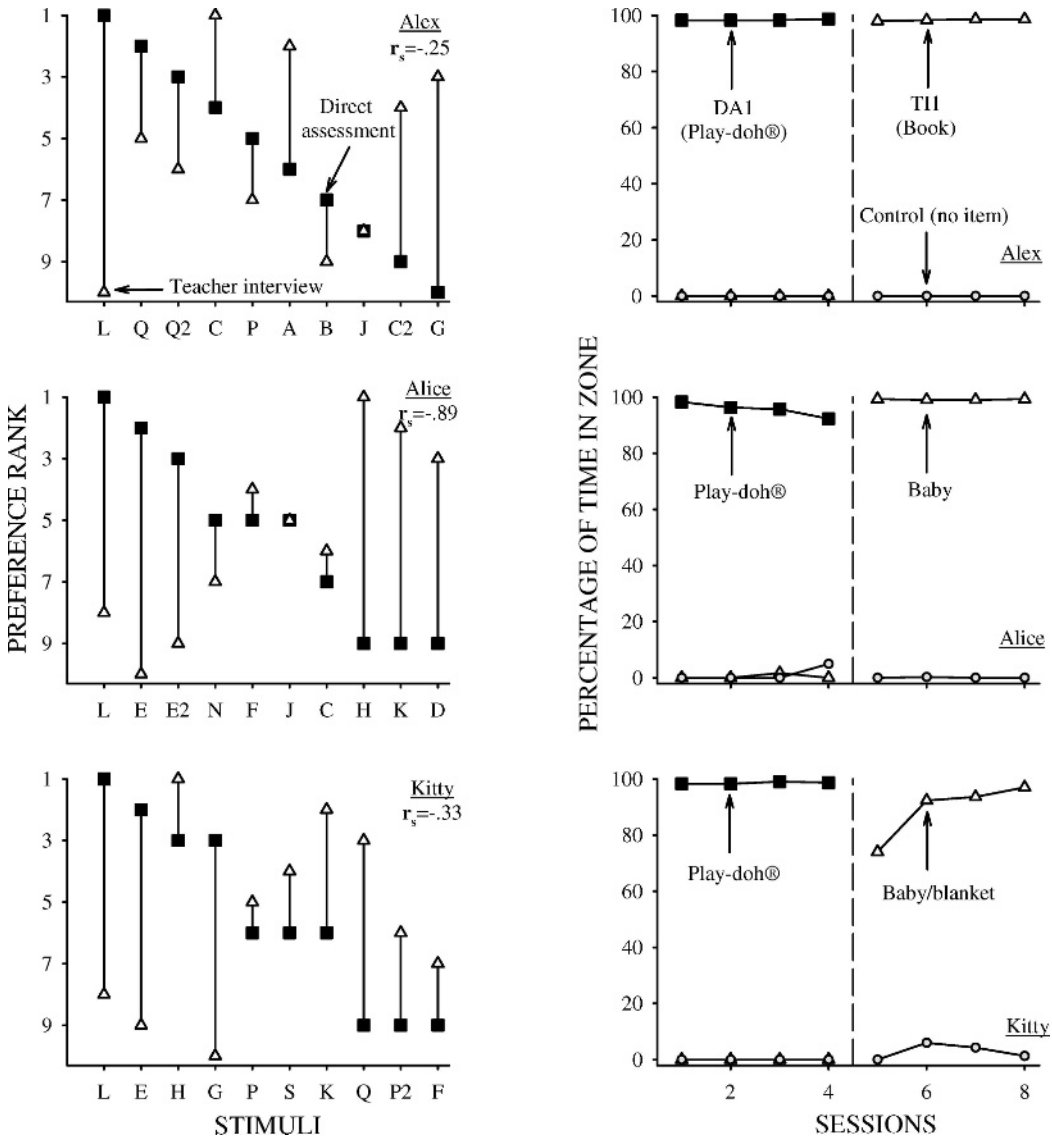
Figure 1.   Preference rank of stimuli from teacher interviews and direct preference assessments (left) and percentage of time in zone during the reinforcer assessments (right) for Alex (top), Alice (middle), and Kitty (bottom).

rankings resulting from these two assessments indicated poor correspondence (a strong positive correlation was observed for only 1 of 9 children). The items identified as most highly preferred in the direct assessment were more potent reinforcers than the items identified as most highly preferred via teacher report for 5 of 9 children, whereas the converse was shown for only 1 child. Nevertheless, all items identified as being highly preferred via both assessments functioned as reinforcers when competing sources of reinforcement were removed from the reinforcer assessment.

The limited agreement between our direct and indirect assessment outcomes is consistent with previous studies that have compared the results of caregiver report with direct preference assessments conducted with individuals with developmental disabilities (e.g., Green et al., 1988; Reid et al., 1999; Windsor et al., 1994).
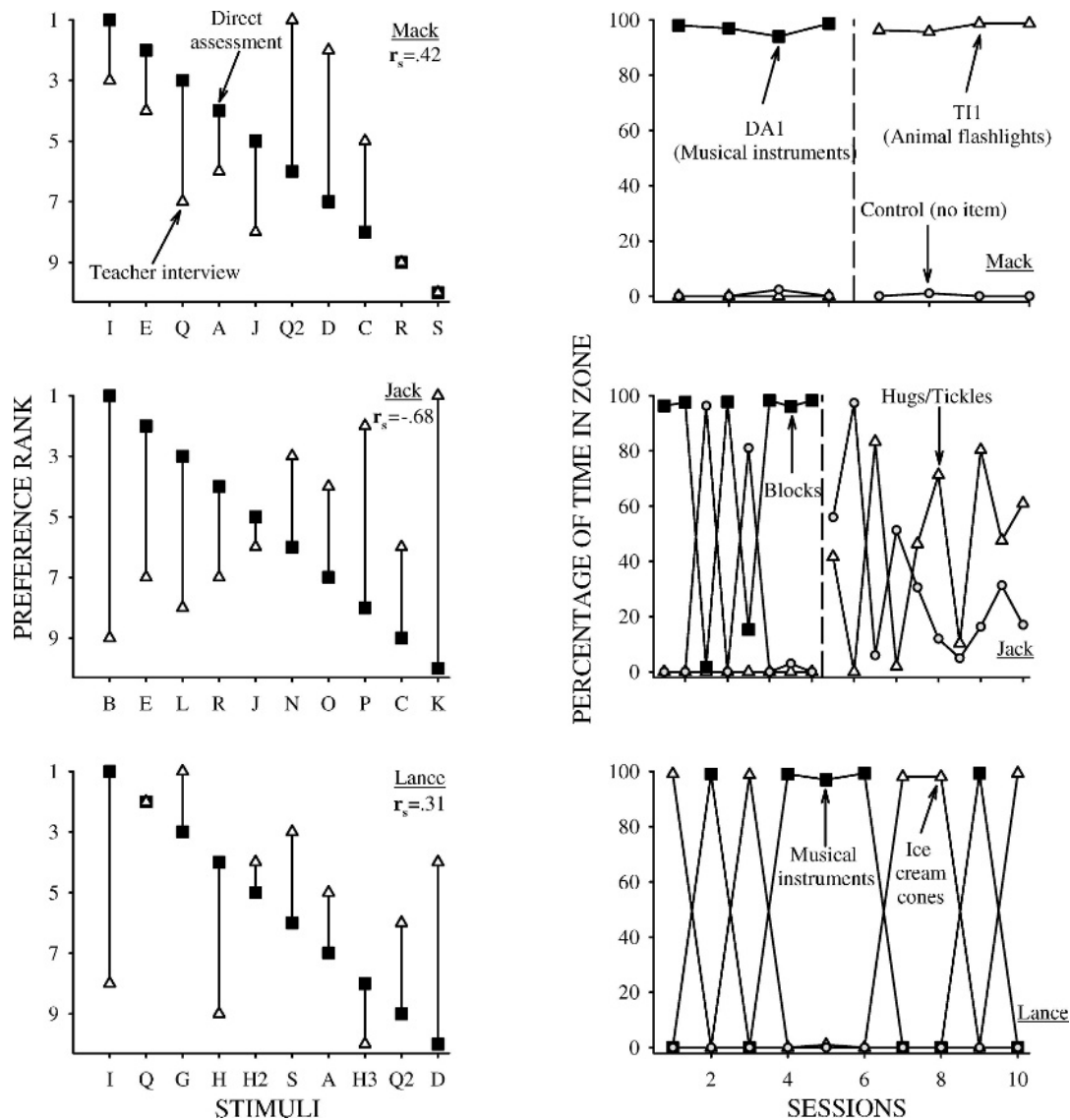
Figure 2.   Preference rank of stimuli from teacher interviews and direct preference assessments (left) and percentage of time in zone during the reinforcer assessments (right) for Mack (top), Jack (middle), and Lance (bottom).

The fact that the direct assessment identified more potent reinforcers relative to teacher interview alone is also consistent with previous research conducted with individuals with developmental disabilities (Fisher et al., 1996). However, all items identified by teachers as most highly preferred were shown to be effective reinforcers when compared with the no-stimulus control condition. Therefore, teacher report was effective in identifying

reinforcers for simple responses performed by typically developing toddlers.

It is also important to recognize that our direct assessment was informed by the results of the indirect assessment (i.e., the items presented during the direct preference assessment were generated through the teacher interview). Fisher et al. (1996) found that a direct preference assessment identified more potent reinforcers when it included stimuli generated by caregivers
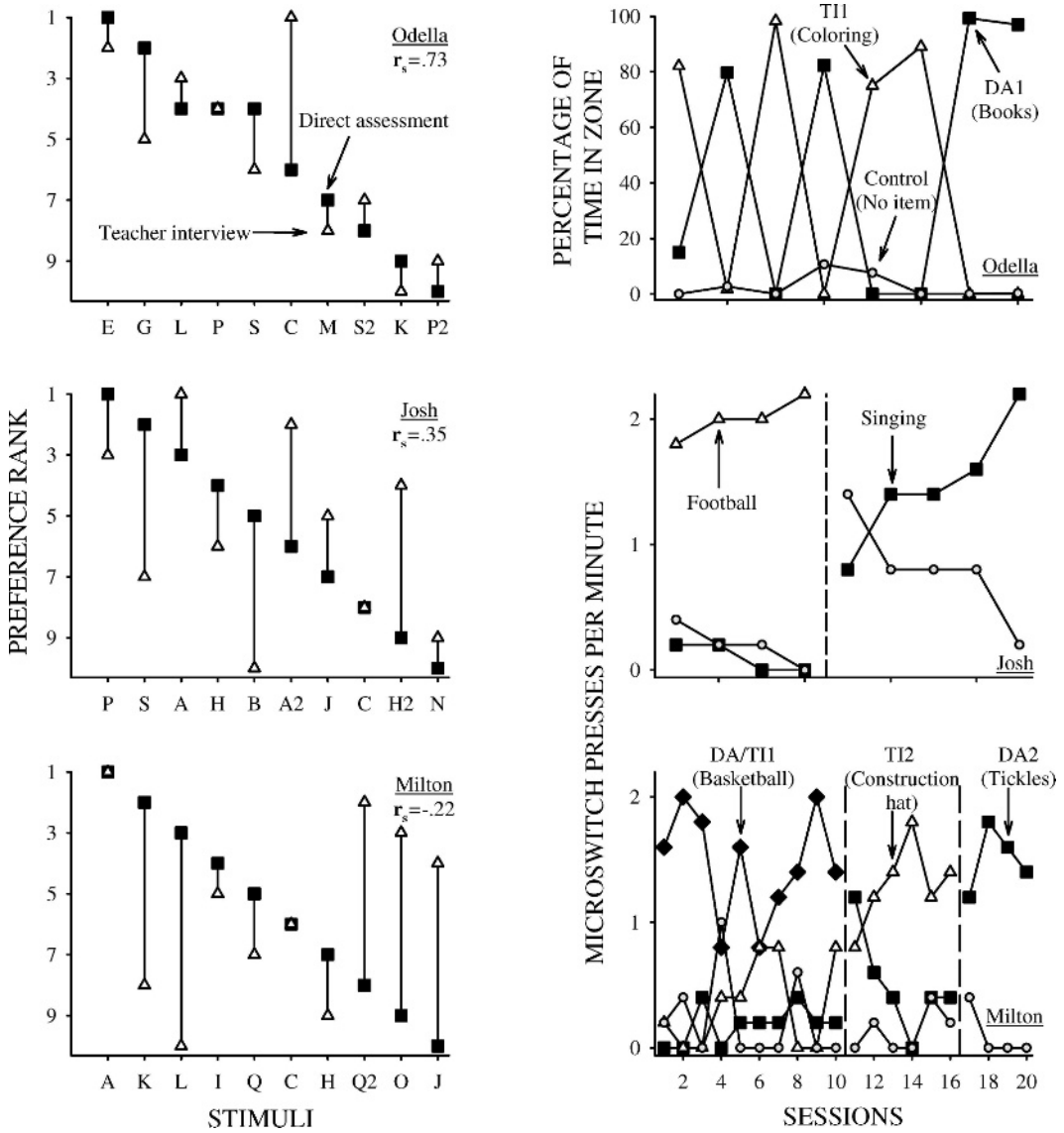
Figure 3. Preference rank of stimuli from teacher interviews and direct preference assessments (left) and percentage of time in zone and microswitch presses per minute during the reinforcer assessments (right) for Odella (top), Josh (middle), and Milton (bottom).

rather than a standard set of stimuli. Thus, in the current study, it is likely that the effectiveness of the direct preference assessment was, in fact, partially attributable to the teacher interview.

A possible limitation of the study is the varying amount of time between the teacher interview, direct preference assessment, and reinforcer assessment. It was necessary to conduct the teacher interview first, to generate stimuli to be presented in subsequent assessments. Therefore, relative to the direct preference assessment, the teacher interview was more temporally distant from the reinforcer assessment. Thus, it is possible that the superiority of the direct preference assessment may have resulted from changes in children's preferences (see Hanley, Iwata, & Roscoe, 2006; Zhou,

Iwata, Goff, & Shore, 2001). However, a comparison of the reinforcer assessment results and the length of time between assessments reveals that the top item from the teacher report was more likely to function as the most potent reinforcer if a relatively longer time interval passed between the teacher interview and the reinforcer assessment (as with Josh and Milton). Thus, it seems that the direct preference assessment was simply better at identifying potent reinforcers regardless of the length of delay between assessments.

Overall, our results suggest that a teacher interview and consensus may result in the identification of potent reinforcers for children, but a direct preference assessment informed by teacher interview is likely to identify items that are relatively more reinforcing compared to items identified by teacher report alone. It should be noted, however, that only the top-ranked stimuli from the indirect and direct assessments were evaluated in the reinforcer assessment. Therefore, this study does not provide information about the predictive validity of these assessment techniques across the hierarchy. Moreover, stimuli identified through preference assessments were presented contingent on simple, low-effort, arbitrary responses; therefore, additional research is necessary to determine whether the relative difference in reinforcing effectiveness evident during the reinforcer assessment would translate into meaningful differences in more effortful and complex classroom behavior (see Carr et al., 2000).

This study extends the literature on preference and reinforcer assessment by demonstrating the applicability of these methods to young children of typical cognitive development in an early education setting. In addition, results provide empirical support for the use of two methods of preference assessment—caregiver interview and direct preference assessment—that may be used to guide the development of individualized curriculum for young children.

Despite these preliminary findings indicating an advantage of direct over indirect assessments with typically developing children, there are a number of additional factors to consider when comparing the two assessment methods. Relative to the direct preference assessment, the teacher interview requires about one third less time, minimal expertise, and does not necessitate the purchase of all assessment items. In addition, the teacher interview was conducted during a regularly scheduled teachers' meeting and did not interfere with scheduled programming; the direct preference assessment required that each child be removed from the typical classroom routine. Thus, teacher interview may be preferred by early childhood educators, who are typically faced with limited resources and are motivated to minimize classroom disruptions. Future research should provide additional information to determine if the benefits of conducting a direct assessment of preference justify the additional time and effort required.

Future research should also explore alternative methods of assessing children's preferences that can be accomplished with minimal disruption of classroom routines. For example, Reid et al. (2003) conducted brief observations of children with developmental disabilities in inclusive settings to identify preferences for toys readily available in classrooms. Similarly, Hanley et al. (2005) recently demonstrated the validity and reliability of a time-sampling procedure for simultaneously determining preschoolers' preferences for ongoing classroom activities. The strength of these assessments is that they are both conducted during regularly scheduled free-play periods, with no disruption to the classroom routine. However, attempts to equate the availability of toys and to ensure similar teacher interactions with different activities are difficult when determining preferences during ongoing classroom routines. Therefore, future research should be devoted to developing preference assessment methods

that involve the direct observation and measurement of children's selection of, or engagement with, materials in early childhood classrooms, while both disruption to the classroom routine and the influence of uncontrolled variables are minimized. The goal of this research is the development of a valid preference assessment that would be easily implemented and readily adopted by early childhood educators.

REFERENCES

Bredekamp, S., & Copple, C. (1997). *Developmentally appropriate practice in early childhood programs* (rev. ed.). Washington, DC: National Association for the Education of Young Children.

Bricker, D., Pretti-Frontczak, K., Johnson, J., & Straka, E. (2002). *Assessment, evaluaton, and programming system for infants and children* (Vol. 1, 2nd ed.). Baltimore: Brookes.

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.

Conyers, C., Doole, A., Vause, T., Harapiak, S., Yu, D. C. T., & Martin, G. L. (2002). Predicting the relative efficacy of three presentation methods for assessing preferences of persons with developmental disabilities. *Journal of Applied Behavior Analysis, 35,* 49–58.

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.

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

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

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

Hanley, G. P., Cammilleri, A. P., Tiger, J. H., & Ingvarsson, E. T. (2005). *Towards a method for describing preschoolers' activity preferences.* Manuscript submitted for publication.

Hanley, G. P., Iwata, B. A., Lindberg, J. S., & Conners, J. (2003). Response restriction analysis: I. Assessment of activity preferences. *Journal of Applied Behavior Analysis, 36,* 47–58.

Hanley, G. P., Iwata, B. A., & Roscoe, E. M. (2006). Some determinants of changes in preference over time. *Journal of Applied Behavior Analysis, 39,* 189–202.

Hills, T. W. (1992). Reaching potentials through appropriate assessment. In S. Bredekamp & T. Rosengrant (Eds.), *Reaching potentials: Appropriate curriculum and assessment for young children* (Vol. 1, pp. 43–63). Washington, DC: National Association for the Education of Young Children.

Parsons, M. B., & Reid, D. H. (1990). Assessing food preferences among persons with profound mental retardation: Providing opportunities to make choices. *Journal of Applied Behavior Analysis, 23,* 183–196.

Reid, D. H., DiCarlo, C. F., Schepis, M. M., Hawkins, J., & Stricklin, S. B. (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.

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

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

Simon, J. L., & Thompson, R. H. (2006). The effects of undergarment type on the urinary continence of toddlers. *Journal of Applied Behavior Analysis, 39,* 363–368.

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

Zanolli, K. M., Paden, P., & Cox, K. (1997). Teaching prosocial behavior to typically developing toddlers. *Journal of Behavioral Education, 7,* 373–391.

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