JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2008, **41,** 25–37 NUMBER 1 (SPRING 2008)

DETERMINING INDIVIDUAL PRESCHOOLERS' PREFERENCES IN A GROUP ARRANGEMENT

STACY A. LAYER, GREGORY P. HANLEY, NICOLE A. HEAL, AND JEFFREY H. TIGER

UNIVERSITY OF KANSAS

This study sought to determine the accuracy of an assessment format in which selection outcomes were delayed and probabilistic; these are unavoidable features of an assessment designed to determine preferences of multiple children simultaneously. During the single arrangement, preference hierarchies were established by having a child repeatedly select from among several foods and by sequentially restricting preferred items from the array. After being taught the associations between colored cards and the same food items, group assessments were conducted with 3 children simultaneously, in which each child chose a card and all children received the food correlated with a randomly selected card from those that had just been selected. The group assessment appeared to be accurate and perhaps more efficient for determining preferences; thus, we posit that this arrangement is suitable for determining context preferences of multiple children simultaneously.

DESCRIPTORS: concurrent-chains arrangements, delay, preference assessments, preschool children, probabilistic outcomes

---

Preference assessments typically identify potentially reinforcing items, activities, or contexts through direct observation of a selection between two or more alternatives. These assessments are usually conducted with individuals for whom verbal report of preferences may be unreliable or impossible to obtain (e.g., young children and persons with severe developmental disabilities). Many studies have described procedures for assessing preferences among discrete items (e.g., foods, toys) to be used as reinforcers during intervention or instruction (e.g., DeLeon & Iwata, 1996; Fisher et al., 1992; Pace, Ivancic, Edwards, Iwata, & Page, 1985; Roane, Vollmer, Ringdahl, & Marcus, 1998). However, relatively few studies have described methods for determining preferences for important contexts or experiences, the results of which could produce lifestyle enhancements (e.g., Lohrmann-O'Rourke, Browder, & Brown, 2000; Risley, 1996) by allowing individuals who are typically unable to advocate for themselves to express their preferences for how their environment is to be arranged.

Hanley and colleagues have used various permutations of a concurrent-chains arrangement (Catania & Sagvolden, 1980) to determine individuals' preferences among contexts such as protracted activities (Hanley, Iwata, & Lindberg, 1999), function-based interventions (Hanley, Piazza, Fisher, Contrucci, & Maglieri, 1997; Hanley, Piazza, Fisher, & Maglieri, 2005), and effective instructional practices for preschoolers (Heal & Hanley, 2007; Tiger, Hanley, & Hernandez, 2006). These assessments all involved modified concurrent-chains arrangements in which children chose among discriminative stimuli that were correlated with various contexts in the initial link of a chain; children then experienced the correlated context in the terminal link of the chain. This process was then repeated for each child until a preference for a particular context was determined through visual inspection of selection patterns in the initial links.

Heal and Hanley (2007) introduced two procedural changes. First, the initially preferred

This research is based on a thesis by the first author in partial fulfillment of the requirements for the master's degree. Gregory Hanley is now at Western New England College, Nicole Heal is now at the May Center for Education and Neurorehabilitation, and Jeffrey Tiger is now at Louisiana State University.

Requests for reprints should be addressed to Gregory P. Hanley, Psychology Department, Western New England College, 1215 Wilbraham Road, Springfield, Massachusetts 01119 (e-mail: ghanley@wnec.edu).

doi: 10.1901/jaba.2008.41-25

context was removed and the assessment continued until only one context remained to establish a preference hierarchy among contexts. Second, assessments also were replicated by reintroducing all contexts and reapplying the same restriction criteria. These changes provided better information regarding children's preferences for contexts, but they also extended the time required to conduct the assessment.

The utility of determining children's preferences for behavior management strategies, group instructional techniques, or designs of classroom activities (e.g., free play, circle, or mealtime) is predicated on determining the individual preferences of all members of a classroom. Result from such assessments could inform individualized or alternating preferred classroom practices; conditions consistent with the preferences of the majority also could be based on assessment results. However, application of the context preference assessments, as they have been described most recently (Heal & Hanley, 2007; Heal, Hanley, & Layer, unpublished manuscript; Tiger, Hanley, & Heal, 2006), would require a prohibitive amount of time for a scientific investigation and certainly would preclude routine application as a practical means to determine children's preferences on a classwide basis. For example, 120 to 160 3-min trials were required to determine a context preference for each child in Heal and Hanley's study. If these procedures were extended to determine the context preferences of a classroom of 20 children and a similar amount of selection variability is assumed, more than 140 hr would be required to complete the classwide assessment.

Heal and Hanley (2007) noted that experiencing the contexts in the terminal links is the time-consuming aspect of the concurrent-chains procedures; by contrast, observing selections in the initial links—the actual measure of preference—requires little time. For instance, if one wanted to determine preschool children's preferences for different teacher responses to problem behavior or different free-play arrangements, it would be relatively easy to provide multiple children an opportunity to express a personal preference in the initial links of the assessment. Children could, for example, all touch one of multiple shapes when entering the classroom from outdoor play, with each shape corresponding to a different context. Because it is not likely that all children will select the same context on each day, the more difficult aspect of the procedures would be allowing multiple children to experience their selected context. For instance, it would not be feasible for classroom teachers to provide each child with a different behavior management strategy each day, and it would be impossible to allow multiple free-play arrangements to cooccur simultaneously due to the shared nature of these contexts.

Although Hanley et al. (1997) showed that differential reinforcement of selections (i.e., providing access to the contexts following selections) was critical for obtaining accurate descriptions of preferences, they did not demonstrate that this reinforcement needs to be continuous. It is possible that intermittent access to selected contexts may yield accurate descriptions of preference, and this intermittency would certainly make these assessments more efficient. For instance, assessments that scheduled access to selected contexts on every other selection would require half the time to conduct. If intermittent access to selected contexts yielded accurate assessments of preference, then it is possible that multiple children's preferences for various contexts could be determined simultaneously. As a means to this end, a type of lottery system could be adopted in which all children are provided with an opportunity to select their preferred context, but only one context, randomly selected from the chosen contexts, is experienced by the group. All children would have a chance of experiencing their selected context, but they would experience their selected context only intermittently.

With this group-oriented concurrent-chains arrangement, the selected outcomes are necessarily delayed because all children make selections before any context is experienced. The outcomes also are probabilistic in that each child's selected context will be experienced on only a proportion of trials. This probability would vary as a function of the number of contexts from which to select and the extent to which a particular child's selection was consistent with the majority of selections. Thus, although a group-oriented assessment might allow a determination of preferences for individual children in an efficient manner, it is unknown whether the delayed and probabilistic outcomes that are inherent in these procedures would compromise the accuracy of the assessment results.

Thus, the ultimate reason for conducting this study was to determine if group-oriented procedures could be applied to determine children's preferences for various contexts accurately. However, to conduct a comparative analysis of a group concurrent-chains arrangement and a typical arrangement for determining preferences for contexts (referred to hereafter as the single arrangement), it was necessary to simplify our analysis. We did this by determining preference hierarchies with individual children for edible items that could be consumed within seconds, as opposed to contexts that require minutes to experience, and comparing the hierarchies obtained with the single and group procedures. Our primary question was whether the delayed and probabilistic nature of the group arrangement imposed unwanted variability on children's preferences as identified using single arrangements.

METHOD

Participants and Setting

Participants included 14 children who were enrolled in an inclusive preschool classroom. All participants were between 3 years 2 months and 5 years 3 months old ($M = 4$ years 3 months) and typically developing. Participation was a function of the mutual availability of child and experimenter in addition to parental consent. Sessions took place in a small room in close proximity to the child's regular classroom. The room included a child-sized table and chairs. Each single assessment was completed during one session, and each group assessment was completed within either one or two sessions, with each session separated by less than 24 hr.

Dependent Variables and Data Collection and Analysis

The primary dependent measure was *selection,* defined as pointing to one item in an array of items. Although these children could vocalize their preferences, a motor response was chosen as the target behavior because motor responses are usually used in concurrent-chains assessments due the complicated nature of the experiences for which preferences are being assessed, and because it allowed a child's selections to remain covert during group arrangements when multiple children were present simultaneously. Depending on the assessment type, the items were either actual food items or small colored cards that were correlated with the same food items. Rankings for each food were calculated by dividing the total number of selections of a particular food by the total number of trials during which that same food was available, multiplying by 100%, and then ordering the foods from the highest to lowest percentage of selections (i.e., a 1 was assigned to the food with the highest selection percentage and a 4 was assigned to the food with the lowest selection percentage). Selection percentages were equal occasionally, but these ties were broken by assigning the food item chosen on the last trial the higher rank; thus, all ranks were absolute numbers. The last food chosen was assigned the higher rank because later choices followed the longer experience with the assessment items. Food rankings across single and group arrangements were examined for each child to determine if the group procedures imposed unwanted variability on preferences.

Interobserver agreement was measured during 100% of sessions for every assessment and child. Agreement was calculated by dividing the number of trials with selection agreements (observers recorded the same chosen food or colored card) by the number of trials and multiplying by 100%. Mean agreement was 97% (range, 15% to 100%) for all children in the single arrangement and 98% (range, 83% to 100%) for all children in the group arrangement. For the single session in which agreement was under 80%, an examination of the data revealed that the secondary observer scored two selections as occurring during the same trial in the beginning of the session, and this affected all trials for the remainder of the session.

Procedure

Single arrangement. Our single arrangement procedures were similar to those described by DeLeon and Iwata (1996) and Windsor, Piche, and Locke (1994) in that we conducted a multiple-stimulus preference assessment. The assessments were conducted with a single child, using the same four items for each child (jelly beans, chocolate peanut butter balls, cheese crackers, and raisins). These particular foods were cut into small pieces approximately the size of the raisin and were chosen because they each had a distinct flavor (e.g., sweet, nutty, salty, fruity). A cup of water was available continuously to the children in all sessions. A single child was present during this assessment. Prior to the start of the evaluation, a one-time sampling period was arranged in which the experimenter labeled each item as he or she placed it on a plate in front of the child, and then suggested that the child try the item (e.g., "This is a raisin; try it"). At the beginning of the initial assessment for each child (the Single 1 assessment), the experimenter presented an array of four items in a half circle on the table in front of the child so that all items were equidistant from each other and from the child. The experimenter instructed the child to point to his or her most preferred item (i.e., "Point to the one you like the best") and then delivered the selected item to the child; he or she had up to 30 s to consume the item. Any portion of the item left on the plate after 30 s was disposed of so that the presence of the food did not affect selections on the next trial. If a child did not make a selection, the experimenter would have waited 5 s and then delivered the prompt again; however, this did not occur for any of the children. If the child pointed to two items simultaneously or made an unclear selection by pointing between the two plates, the experimenter said, "I am not sure which one you wanted. Let's try that again. Point to the one you like the best." After a selection, the experimenter replaced the food and then rotated the items so that the last item on the child's right became the first item on the left, and the plates were arranged equidistant from each other again. A new trial then began in which the experimenter instructed the child to pick his or her most preferred item.

This was repeated until a preference was indicated according to one of the following restriction criteria: (a) The child chose the same food four times consecutively, (b) a particular food was selected four more times overall, (c) two foods were selected equally compared to each other and four more times overall, or (d) 20 total trials had been conducted with no other restriction criteria met. If a child met criteria (a), (b), or (c), that particular food was restricted and the child was presented with the remaining foods until the same criteria were met to obtain a rank of preference among the food items. If a child met criterion (d), the assessment ended. During each assessment, the experimenter delivered a brief statement of praise or social conversation unrelated to the child's choices (e.g., "You are sitting so nicely," or "It is so sunny, I can't wait to go outside later") following every other trial, and also briefly responded to social interactions initiated by the child at any time.

Exposure. It was necessary to pair the food items with arbitrary and distinct stimuli (i.e., 2.5 cm by 2.5 cm colored cards) for two reasons. First, during the group arrangement, each child's selection was placed in a box with the other children's selections, one was drawn at random, and all children received the item correlated with the card. It was more practical to use cards instead of foods for this portion of the procedures because placing food items in the box presents obvious problems with sanitation and food waste. In addition, the group arrangement was intended to simulate the concurrent-chains procedures used to evaluate preferences for contexts; therefore, in future applications of these methods, the cards will represent experiential contexts that cannot possibly be placed in box.

Therefore, prior to evaluating preferences in a group arrangement, we conducted exposure sessions to teach the children the correlations between the different foods and colored cards to be used as initial-link stimuli in the group arrangement. Four colors were chosen randomly to be correlated with each of the four foods presented in the single arrangement. Three children simultaneously participated in this procedure. During each trial, all children were instructed to point to one of four available colored cards (each child had his or her own set of four cards). After they pointed to the specified card, the experimenter provided brief praise and presented every child with the food that was correlated with that color. The child had access to the food for up to 30 s, and any portion of the item not left on the plate after 30 s was removed. This sequence of events was repeated for 20 trials, with each color and food association presented on five trials each in a random and counterbalanced order. Prior to the start of each subsequent session, the experimenter asked each child what happened when he or she pointed to each of the four cards. When the child was able to describe the contingencies correctly (e.g., the experimenter asked, "What do you get when you touch pink?"and the child responded, "peanut butter balls") following at least 24 hr since a previous exposure session, the exposure period was considered complete and the children moved into a group arrangement. The mean number of exposure sessions required for a child to learn the correlations was 2 (range, 1 to 5).

Group arrangement. Prior to making selections, 3 children were seated at a table and were provided with instructions to keep their selections a secret to avoid the possibility of one child's selection influencing another's. The experimenter attempted to group children so that as many children as possible within the group had differing foods ranked first according to the single arrangement, and one group participated at a time. All children made their selections in private; to accomplish this, the experimenter held up a folding poster-board barrier (1 m by 1 m) between the child choosing and the other children. An array of the four cards that had been paired with each of the foods in the exposure sessions was arranged in front of each child. Each child was instructed to point to the one that he or she liked best. After each selection, the colored card corresponding to each child's choice was placed inside a voting box, and this was repeated until each child had a card in the box. The experimenter then set aside the barrier, showed the box to all of the children, and drew one card out of the box. All children received the item correlated with the chosen card. This method allowed each child a chance of having his or her selected card chosen; however, this also resulted in food deliveries that were both delayed and probabilistic relative to the single arrangement. The probability of a child receiving the item that was selected was measured and is reported below. By contrast, delays between selecting and receiving the foods were not measured but were estimated to range between 1 to 2 min per trial.

Each card was replaced after every selection so that all children continued to select from the full array of four cards until a preference emerged. The different cards and their associated edible consequences were restricted using the same criteria described in the single arrangement. Once a preference hierarchy was established for a particular child, he or she was removed from the pool of children that could be grouped together for a session. To ensure groups of three for each session, children who completed the assessment were sometimes included in groups at the close of the overall evaluation, but their data were not used in the analyses.

In sum, the single and group arrangements were the same in that every child chose among the same four foods, and the same criteria for determining preferences and assigning ranks to food items were used. The few important differences between the two assessment formats were as follows. First, each child chose from an array of foods in the single arrangement, whereas each child chose a card that was correlated with a food in the group arrangement. Second, each child was guaranteed to receive every item that he or she chose in the single arrangement, whereas the number of times each child received his or her chosen item was probabilistic in the group arrangement. Third, there was a longer delay between selections and delivery of food in the group arrangement than in the single arrangement. Finally, each child was alone in the single arrangement, whereas 3 children were present in the group arrangement.

Experimental Design

Concurrent-chains designs allowed preferences to be determined within each assessment. In addition, an AABAB design was used to evaluate differences in selections as a function of the group arrangement. Initially, two single arrangements (Condition A) were conducted consecutively to determine if preferences were stable across repeated measures of the same assessment type before we attempted to determine if the group arrangement (Condition B) imposed unwanted variability on preferences. Next, the exposure condition was implemented to teach the card and food associations. The group arrangement was conducted next, followed by a reversal to a single arrangement, and finally a replication of the group arrangement. All 14 children experienced all conditions in a similar manner.

RESULTS

The primary research question regarding the imposition of variability via the group arrangement was addressed by evaluating changes in rank of each of the four foods across repeated measures of the same arrangements, and also repeated measures of the two different arrangements. Trial-by-trial data for 2 children are shown in Figure 1. These are examples of two common types of performance patterns. Al's data provide an example of consistent preference rankings across both single and group assessment formats. He selected a peanut butter ball on the first trial, followed by a jelly bean on the second trial. Al again selected peanut butter balls on the third and fourth trials, and then he selected a cheese cracker. He next selected peanut butter balls two more times. Peanut butter balls were then restricted because they had been selected four more times overall. After peanut butter balls were removed from the array, Al switched selections between jelly beans and cheese crackers until each were selected four more times than the raisin (which was never selected), and the end-of-session criterion had also been met. Al continued to show similar patterns of responding throughout the rest of the assessments, until the second group assessment when he chose each of the foods four times consecutively. Throughout every assessment, peanut butter balls were always the most preferred food, jelly beans second, cheese crackers third, and raisins last. Thus, it did not appear that the group arrangement altered the primary outcome of the preference assessment in any appreciable way; the same preference hierarchy resulted from all assessments. There was not a single participant for whom preferences were stable only across the

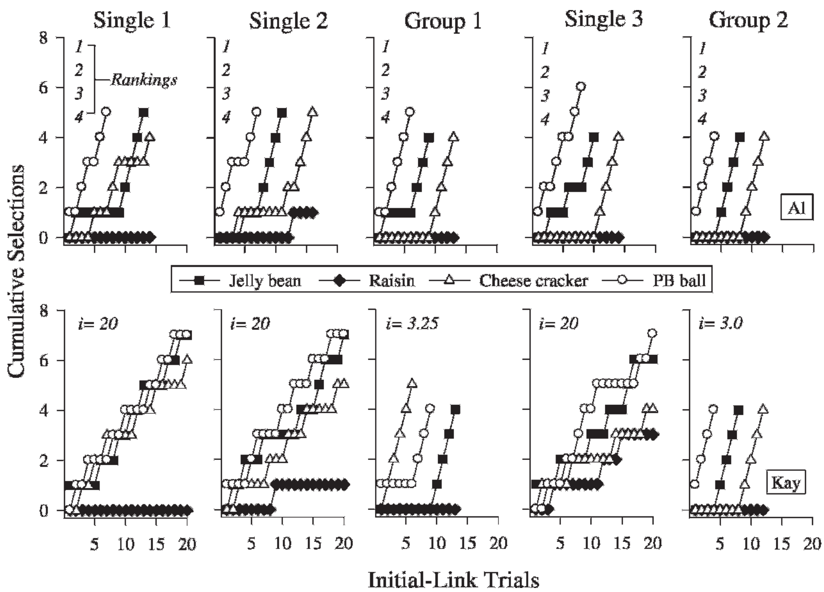

Figure 1. Cumulative selections across trials and relative rankings of food items are displayed for Al in the top panel. These data illustrate an example of a child whose preference hierarchy did not change during the entire assessment. The bottom panel shows cumulative selections across trials and indifference scores for Kay. Indifference scores describe the amount of selection variability within each assessment. These data illustrate an example of a child with highly indifferent responding during the single assessments and highly discriminated responding during the group assessments.

single arrangements; preferences were either stable across both assessment types or preferences were variable regardless of assessment type.

We also examined the data using a criterion for determining an agreement. According to this criterion, a ranking agreement occurred when all foods did not move more than one rank across the two assessments. When comparing the first and second single arrangement using this criterion, rank agreements occurred for 69% of children. A comparison of the first group assessment with the preceding single assessment revealed that rank agreements also occurred for 69% of children. Another way to examine rank consistency includes determining if the first ranked food remained the same across assessments. For single assessments, 57% of the children had their first ranked food remain the same across repeated measures of the single arrangement. A comparison of the first group assessment with the preceding single assessment revealed that the first ranked food remained the same for 57% of children. Considering that there was no difference between the percentage of children for whom rankings were generally consistent for these two comparisons (e.g., Single 1 to Single 2 and Single 2 to Group 1 on all rankings or just the top items), it seems unlikely that unwanted variability in food rankings was imposed by the group arrangement.

To further describe the extent of agreement within and across assessment types, a Spearman's correlation analysis was conducted on food rankings (Table 1). First, the rankings for all children for the initial assessment were compared to the rankings yielded by the second single assessment (Single 1/Single 2 in Table 1); this yielded a moderate, positive correlation of .60. Correlations of similar strength were found between all like assessment types ($\rho = .60$ for Single 1/Single 2, $\rho = .68$ for Single 2/Single 3, $\rho = .57$ for Single 1/Single 3, and $\rho = .72$ for Group 1/Group 2). These correlations were close to or higher than a generally acceptable test-retest reliability coefficient of .60 (Cicchetti

Table 1
Correlations and Days Expired Between Assessments

	Same arrangement type				Different arrangement type		
	Single 1/ Single 2	Single 2/ Single 3	Single 1/ Single 3	Group 1/ Group 2	Single 2/ Group 1	Group 1/ Single 3	Single 3/ Group 2
Correlations between assessments	.60	.68	.57	.72	.65	.74	.71
Mean number (range) of days between assessments	19 (1–34)	30 (7–60)	44 (8–68)	28 (7–70)	19 (6–47)	10 (1–49)	17 (3–69)

Note. All correlations between assessments are statistically significant at $p < .05$.

& Sparrow, 1981). These data showed that both single and group assessment types share a similar amount of consistency across applications. The correlations generated between the single and group assessments were also all above .60 ($\rho = .65$ for Single 2/Group 1, $\rho = .74$ for Group 1/Single 3, $\rho = .71$ for Single 3/Group 2). All of the above correlations are statistically significant ($p < .05$). By comparing the sets of correlations, it is apparent that the correlations within assessment types were of the same magnitude as those correlations between assessment types. Thus, these data are also consistent with the notion that the group assessment did not impose unwanted variability on the outcomes of the preference assessments.

Because the amount of time between assessments varied unsystematically, we determined whether the amount of time between assessments affected the correlations between assessments (see Table 1 for the mean number of days between assessments). A Pearson correlation analysis involving time between assessments and the rank correlations between the assessments showed that these variables were not correlated ($r = .04$), suggesting that variations in the rank correlations were probably not due to the differences in time between assessments.

One of the features of the group arrangement was that it was probabilistic that each child would receive the item that he or she preferred the most; traditional preference assessments guarantee that the child receives every item that he or she selects (i.e., the probability of receiving a selected item is 1). Table 2 shows that the probabilities that each child received the same item selected varied considerably across children. For example, the probability that Eva received the item that she selected in the second group arrangement was .83, whereas the probability that Kat received the item that she selected in the second group arrangement was .08 (Table 2). The mean probability of each child receiving the item that he or she chose in the group arrangement was .53 in the current study. Thus, children's selections resulted in access to their chosen item on about half the trials. To determine if the probabilistic outcomes influenced the extent of agreement between the single and group assessments, a Pearson correlation analysis between the mean probability that a child received selected items and the correlations between the single and

Table 2
Probability That a Child Received the Same Item Selected

	Assessment	
	Group 1	Group 2
Ed	.44	.43
Kay	.39	.67
Ben	.43	
Roy	.44	.75
Lea	.56	.50
Mia	.57	.69
Roz	.44	.42
Kat	.43	.08
Abe	.43	.53
Mel	.58	.53
Eva	.62	.83
Joy	.46	.58
Jim	.65	.53
Al	.77	.50
M	.52	.54

group arrangements was conducted. Our concern was that children with lower probabilities of receiving the items they selected (e.g., Kat) would be more likely to have inconsistent assessment outcomes. A weak, positive, and statistically insignificant correlation found between these variables ($r = .38, p > .05$) suggests that this was not the case. In addition, a scatter plot showed that this weak relation was highly dependent on two outliers. Thus, the probabilistic selection outcomes were likely not responsible for any disruptions in preference observed between single and group assessments.

Although preference hierarchies were not systematically different between single and group arrangements, the patterns in the data were different for each assessment type with respect to within-assessment variability. As we analyzed the data with respect to our primary question, it became apparent that children's selections were more highly discriminated (i.e., there was less within-session selection variability) during the group arrangements.

A metric to describe and quantify the amount of selection variability within and across assessments was derived from the data. We refer to this metric as an indifference score (i score) because indiscriminate responding would yield relatively high scores and discriminated responding would yield relatively lower scores. These scores were calculated by dividing the total number of trials in each assessment by the number of restriction criteria met plus one (one represented the end-of-session criterion). The i scores ranged from 3 (*highly discriminated*) to 20 (*highly indifferent*). For an example of a high i score, see Kay's first single assessment (Figure 1). Twenty trials were conducted without Kay meeting a single restriction criterion. Therefore, the i score was calculated by dividing the 20 total trials by 1, for a score of 20, the maximum score that could be achieved. By contrast, Kay's second group assessment was conducted for a total of 12 trials, and three restriction criteria were met. The i score was calculated by dividing 12 by 4 (restriction criteria met plus one) to obtain a score of 3. This is representative of the lowest possible i score.

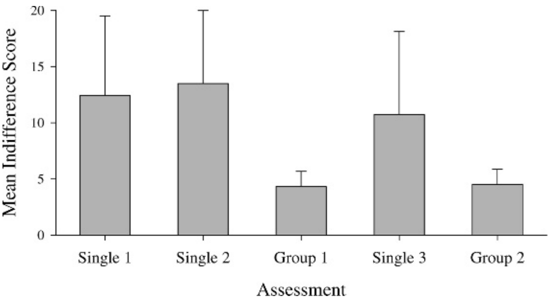

Figure 2. The mean indifference scores and standard deviations of those scores for all 14 children. Indifference scores refer to the amount of selection variability within an assessment.

Kay's data (Figure 1) also illustrate a common pattern of within-session variability observed within the single and group arrangements. During the first and second single assessments, Kay allocated a fairly equal number of selections between peanut butter balls, jelly beans, and cheese crackers. In the first group assessment, selections were more highly discriminated. When a third single assessment was conducted, she again frequently switched selections among all the foods. Finally, in the second group assessment, Kay engaged in the most discriminated responding by selecting peanut butter balls four times consecutively, followed by jelly beans four times consecutively and cheese crackers four times consecutively. Kay's data represent a strong example of the differences in variability between the single and group assessments. Her responding was indifferent during all single assessments with i scores of 20 (the maximum score), and highly discriminative responding was observed during the group assessments with scores of 3.3 and 3.0 (the minimum score), respectively, during each group assessment. These data show that the group assessment resulted in more discriminated selections than the single assessments for Kay.

Figure 2 shows the mean i scores across children for each of the five assessments. The

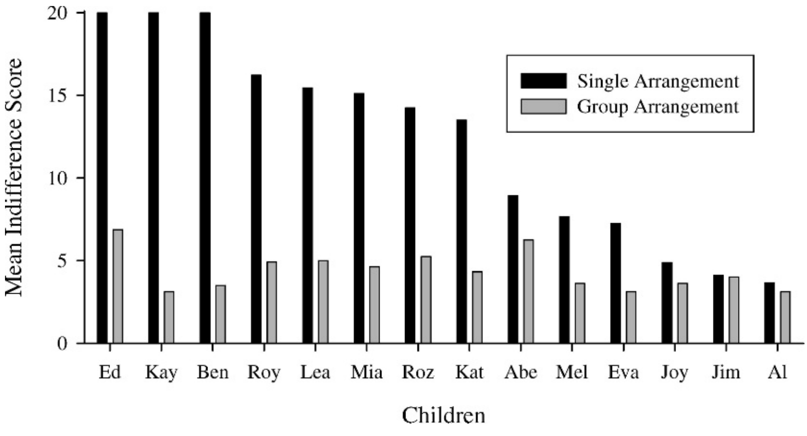

Figure 3. The mean indifference score by assessment type for each individual child.

data reveal that mean i scores were higher in each of the single arrangements than in the group arrangements. In addition, the standard deviation of scores was higher in each of the single arrangements than in the group arrangements. This effect was replicated in a reversal design for the grouped data. A different representation of the i scores is seen in Figure 3, which shows the mean i score by assessment type for each individual child. For all children, i scores in the single arrangements were either higher than or equal to i scores in the group arrangements. In addition, the range of group-arrangement i scores was minimal (3 to 7.5) relative to the range of single-arrangement i scores (3 to 20). To determine if a relation existed between the probability that a child received the same item he or she selected and i scores, a Pearson correlational analysis was conducted on these variables. A weak, negative correlation of −.28 resulted, indicating that selections were slightly more likely to reflect indifference when the child was less likely to receive his or her selected item.

DISCUSSION

This study sought to determine if group-oriented preference assessments, in which outcomes were both delayed and probabilistic, would impose unwanted variability on selections and disrupt the rank of preferences. The evidence against negative effects of the group arrangement were as follows: (a) There was no single example of stability in ranks being disrupted systematically by the group assessment procedures, (b) the consistency in rankings was similar for like and unlike assessment types, and (c) correlation coefficients were similar for like and unlike assessment types. In fact, for 11 of 14 participants, more highly discriminated responding, and thus more efficient determinations of preference, was associated with the group assessment. Thus, not only did the group arrangement not impose unwanted variability on selections during the identification of preferences, but the assessment actually minimized selection variability for 12 of 14 children and by a magnitude of over 50% for 10 of these 12 children.

However, there were some limitations of this study that are worth noting. First, because all assessments were conducted in the same order for all children, it is possible that the more highly discriminated performances during the group assessments may have been dependent on the fact that the single assessments routinely preceded them. Conducting the conditions in a BBABA design for some participants would have allowed the influence of an order effect to be detected, although not necessarily controlled (Hains & Baer, 1989). In addition, children selected among an array of foods in the single arrangement; in contrast, children selected from an array of colored cards that had been correlated with foods in the group arrangement in order to simulate a concurrent-chains arrangement. Tighter experimental control of the variables of delay and probability would have been achieved by using colored cards in the single arrangement as well (i.e., emulating concurrent-chains arrangements in both the group and single arrangements).

Although this study evaluated methods for assessing individual children's preferences in a group format using edible items, the ultimate purpose was to determine if the group arrangement may be applied to determine the preferences of groups of children who share the same environment. The simultaneous assessment of the individual preferences of multiple children is essential to the improvement of classroom practices because it will allow researchers and perhaps even teachers to evaluate multiple children's preferences for important classroom contexts such as teaching strategies, proactive and reactive behavior management strategies, reinforcement systems, scheduled activities, and other contexts that children often experience throughout their day. Teachers can choose from among a large number of classwide contexts or strategies for promoting a desirable classroom environment. In addition to the effectiveness of the strategy, it seems important to consider children's preferences between available strategies before adopting any particular one. Following the results of a group-oriented preference assessment, the context that is preferred by the majority of children could be implemented on a classwide basis. This democratic system could be supplemented by individualized programs based on child preference as resources allowed. The findings of this study are positive for the original purpose of this investigation and indicate that the group arrangement should be considered an acceptable method for assessing individual preferences for contexts in a group format.

Our study's outcome is ironic in that we were originally seeking a method that did not disrupt preferences in comparison to traditional preference assessments, and ended up identifying a method that resulted in an improvement over traditional procedures for most children. The fact that the assessment involving delayed and probabilistic outcomes represented an improvement over traditional assessments, in which selection outcomes were immediate and guaranteed, seems counterintuitive. However, Tustin (1994) and DeLeon, Iwata, Goh, and Worsdell (1997) reported results that are not unlike our own. For example, Deleon et al. found that as schedule requirements to earn reinforcers increased for 2 individuals with developmental disabilities, preferences for similar reinforcers became more discriminated. During their study, participants responded at similar levels for two items that were available in a concurrent-operants arrangement under equal and low schedule requirements (fixed-ratio [FR] 1 schedule). However, as response requirements increased for both items, participants allocated more responses to one item relative to the other. This indicated that at low schedule requirements, such as those in the single arrangement in the current study, no preference for one or the other item was apparent. As schedule requirements increased—this perhaps being similar to the group arrangement in the current study—consistent preferences emerged. The probabilistic nature of the group assessment in the current study essentially increased the response requirement necessary to obtain reinforcement (i.e., the preferred edible item). The mean probability that selections were reinforced with preferred foods was .53, indicating that the response requirement was essentially doubled. Therefore, preferences may have been more apparent sooner in the group arrangements because of the increased schedule requirement.

Additional support for the current findings comes from studies of behavioral economics in which preferences for two functionally similar items (referred to as substitutable reinforcers) are equal until the cost (in terms of responses required) of both items increases. For example, Madden, Smethells, Ewan, and Hursh (in press) found that at low FR schedule requirements, rats chose food and fat equally often, but as the FR schedule increased, food was selected more often. In relation to the current study, some of the participants did not appear to allocate selections for one particular food over the others in the single arrangement when all selections were reinforced immediately; however, when selections were reinforced intermittently in the group arrangement and the response requirement was essentially increased, selections were allocated toward one food over the others and preferences for a particular food emerged.

An alternative reason as to why the group arrangement resulted in more discriminated selections may be related to the differing densities of food consumption between the two assessments. Foods were consumed in rapid succession in the single arrangement. By contrast, the time between the consumption of individual foods was greater in the group arrangement. Thus, the delay imposed in the group arrangement may have established the reinforcing value of the most preferred foods more so than the shorter delay in the single assessment.

This was a methodological study designed to evaluate procedures for scaling up preference assessments for contexts. The results of context preference assessments are typically used to design environments that are preferred by the individual assessed. Although this study did use discrete items (food) to test the procedures, it was not designed to identify items that could potentially be used as reinforcers, and therefore no reinforcer assessments were conducted. However, the results of this study may have implications for assessing preferences for discrete items (edible items, stickers, toys) for individual children. Typical preference assessment results for discrete items are displayed in bar graphs that depict the percentage of selections per item (see Fisher et al., 1992, for an example). Items are usually ordered on the *x* axis of the graph in the order of highest to lowest percentage of selections. If the bars start high and end low, this pattern indicates differentiated responding, and a conclusive hierarchy of preferences is provided. However, if the bars are fairly equal in height (i.e., the distribution is flat), this indicates either indiscriminate responding or similar preference for all items (i.e., indifference). Conclusions regarding the most effective reinforcers are difficult to derive with the latter data pattern. Considering the results of the current study, it may be useful to impose delayed and probabilistic consequences during a subsequent preference assessment to obtain a more differentiated ranking of the items. This may be accomplished by providing the items on only half of the selections or providing access after variable but brief delays.

Because the delayed and probabilistic nature of the group assessment arose from the practical exigencies of scaling up concurrent-chains arrangements for simultaneously determining preferences of multiple children for contexts, our analysis does not permit independent analyses of the impact of delay and probability on preference assessment outcomes. Thus, it is unclear whether the improvements in the group arrangement were due to the delay, the probabilistic outcomes, or both. Because both are inherent in group-oriented preference assessments for contexts, identifying their isolated effects does not seem to be important. To address undifferentiated outcomes from more typical preference assessments involving discrete items, future research should, however, determine the independent contributions of delay and probability by conducting separate analyses of these variables. It may also be interesting to determine whether discriminative responding during preference assessments is enhanced by manipulating other parameters (e.g., effort) that influence responding in concurrent-operants arrangements.

REFERENCES

Catania, A. C., & Sagvolden, T. (1980). Preference for free choice over forced choice in pigeons. *Journal of the Experimental Analysis of Behavior*, *34*, 77–86.

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*, *86*, 127–137.

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis*, *29*, 519–532.

DeLeon, I. G., Iwata, B. A., Goh, H., & Worsdell, A. S. (1997). Emergence of reinforcer preference as a function of schedule requirements and stimulus similarity. *Journal of Applied Behavior Analysis*, *30*, 439–449.

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.

Hains, A. H., & Baer, D. M. (1989). Interaction effects in multielement designs: Inevitable, desirable, and ignorable. *Journal of Applied Behavior Analysis*, *22*, 57–69.

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

Hanley, G. P., Piazza, C. C., Fisher, W. W., Contrucci, S. A., & Maglieri, K. A. (1997). Evaluation of client preference for function-based treatment packages. *Journal of Applied Behavior Analysis*, *30*, 459–473.

Hanley, G. P., Piazza, C. C., Fisher, W. W., & Maglieri, K. A. (2005). On the effectiveness of and preference for punishment and extinction components of function-based interventions. *Journal of Applied Behavior Analysis*, *38*, 51–65.

Heal, N. A., & Hanley, G. P. (2007). Evaluation of preschool children's preferences for motivational systems during instruction. *Journal of Applied Behavior Analysis*, *40*, 249–261.

Heal, N. A., Hanley, G. P., & Layer, S. A. (2006). *Efficacy and acceptability of three teaching strategies that differ in amount of teacher directedness.* Unpublished manuscript.

Lohrmann-O'Rourke, S., Browder, D. M., & Brown, F. (2000). Guidelines for conducting socially valid systematic preference assessments. *Journal of the Association for Persons with Severe Handicaps*, *25*, 42–53.

Madden, G. J., Smethells, J. R., Ewan, E. E., & Hursh, S. R. (in press). Tests of behavioral-economic assessments of relative reinforcer efficacy: 1. Economic substitutes. *Journal of the Experimental Analysis of Behavior*.

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. *Journal of Applied Behavior Analysis*, *18*, 249–255.

Risley, T. (1996). Get a life! Positive behavioral intervention for challenging behavior through life arrangement and life coaching. In L. K. Koegel, R. L. Koegel, & G. Dunlap (Eds.), *Positive behavioral support: Including people with difficult behavior in the community* (pp. 425–437). Baltimore: Brookes.

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.

Tiger, J. H., Hanley, G. P., & Heal, N. A. (2006). The effectiveness of and preschoolers' preferences for variations of multiple-schedule arrangements. *Journal of Applied Behavior Analysis*, *39*, 475–488.

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

Tustin, R. D. (1994). Preference for reinforcers under varying schedule arrangements: A behavioral economics analysis. *Journal of Applied Behavior Analysis*, *27*, 597–606.

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

Received October 31, 2006
Final acceptance March 6, 2007
Action Editor, Mark Dixon