Education and Training in Developmental Disabilities, 2004, 39(3), 253–264
© Division on Developmental Disabilities

Comparison of Two Approaches for Identifying Reinforcers in Teaching Figure Coloring to Students with Down Syndrome

Dilek Erbas, Arzu Ozen, and Cimen Acar
Anadolu University

Abstract: The purpose of the present study was to extend previous research on reinforcer assessment by comparing effectiveness of stimuli identified by two preference procedures on teaching figure coloring to three children with disabilities in Turkey; and to find out what special education teachers think about social validation of the two preference assessment procedures. In Phase 1, preferences of three students with Down syndrome were identified by using two assessment procedures. In Phase 2, reinforcing effects of the most often selected stimuli both from the stimulus preference and paired-stimulus preference assessments were evaluated. In Phase 3, special education teachers' opinions about the two preference assessment procedures were determined. Results reveal that both preference assessment procedures were effective in identifying reinforcing stimuli for all three students with Down syndrome. Moreover, both procedures were found to be acceptable to some extent; however, results showed that the paired stimulus was more acceptable than the stimulus preference procedure.

Several studies that aimed to identify the reinforcing stimuli to be used in the education of individuals with disabilities have been conducted over the last two decades (e.g., DeLeon & Iwata, 1996; Fisher, Piazza, Bowman, Hagopian, Owens, & Slevin, 1992; and Pace, Ivancic, Edwards, Iwata, & Page, 1985). A review of literature yields that various procedures have been utilized to determine the stimuli. Stimulus preference (Pace et al) and paired stimulus preference assessments (Fisher et al.) are the most frequently used stimuli identifying procedures.

Literature indicates that techniques that are going to be used to identify the preferred reinforcers should consist of two steps (e.g., Fisher et al., 1992; and Pace et al., 1985). The first step is to try to determine if a variety of stimuli are preferred by the student or not; in the second step, the goal is to figure out if the identified stimuli serve as reinforcers or not. Pace et al. developed a procedure to test if the stimuli functioned as reinforcers. In the first phase of this study, preference assessment was conducted by providing each of 16 stimuli ten times. All approach responses of subjects to each stimulus were recorded. If the subject displayed approach response to the stimulus within 5 s, s/he was allowed to enjoy the stimulus for 5 s. On the contrary, if the subject did not show an approach response for the given stimulus, then s/he was verbally prompted to touch the stimulus. At the end of assessment procedure, stimuli were recorded as preferred or non-preferred if the given stimuli were attractive to the subjects among 80% and 50% or less of all the presentations respectively. Results of this procedure revealed that preference assessment was efficient in determining the preferred stimuli for all subjects. In the second phase of this study, preferred stimuli were assessed for their efficacy as reinforcers. Outcomes of this phase indicated that contingent presentation of the preferred stimuli increased subjects' responses compared to baseline and non-preferred conditions. Conducting this study, Pace et al. developed a simple but effective method to identify reinforcers for individuals with profound developmental delays. Also, this procedure was easy, time effi-

The authors are grateful to Gönül Kircaali-Iftar and Elif Tekin Iftar for their insightful reviews and feedback. The study was presented as a poster session at the Annual Convention for Behavior Analysis in Toronto, Canada in May 2002. Correspondence concerning this article should be addressed to Dilek Erbas, Education, Research, and Training Center for the Speech and Language Disordered, Anadolu University, Eskisehir, TURKEY, 26470. E-mail: dderbas@anadolu.edu.tr

cient (2 hrs per subject), and economical (Pace et al.).

Fisher et al. (1992) extended the findings of Pace et al. (1985) by comparing results of a single stimulus and a choice presentation format. In this study, 16 stimuli were paired and then presented during assessment. Subject's approach response to either stimulus was recorded. The subject was allowed to engage with the stimulus to which they displayed an approach response for five seconds. The subject was stopped if they approached the two stimuli simultaneously or was verbally prompted upon displaying no approach response for any of the stimuli for five seconds. If the subject did not show any interest in any of the stimuli, then the stimuli pair was removed and the next pair was presented. A concurrent operant paradigm was used to compare the reinforcing effectiveness of the stimuli in the second phase of the study. Outcomes stated that all subjects spent more time in the square or the chair when the trial consisted of the high-high stimuli. Therefore, paired evaluation of the stimuli revealed greater differentiation among the stimuli and foresaw which stimulus would work as a reinforcer more accurately when assessed in a concurrent operant arrangement.

Different variations of the procedure developed by Pace et al. (1985) and Fisher et al. (1992) have been developed and tested for their efficacy in several studies (e.g., DeLeon & Iwata, 1996; Roane, Vollmer, Ringdahl, & Marcus, 1998; and Windsor, Piche, & Locke, 1994). Furthermore, effectiveness of identified reinforcers was determined by using simple operant free behavior (e.g., in square or in seat behavior, switch pressing, key pecking) in most of the studies (e.g., Fisher et al.; and Pace et al.). Of course, this does not necessarily mean that the stimulus determined as preferred by making use of the concurrent operant paradigm with a simple free operant response will have the same function as a reinforcer in more complicated and clinically pertinent responses. Conversely, the stimulus found to be non-or-less preferred will be no good as a reinforcer when evaluated on a single operant paradigm (Piazza, Fisher, Hagopian, Bowman, & Toole, 1996). Therefore, the aim of this study is to judge which stimuli, identified by using two different reinforcer assessment methods, are more effective in teaching consecutive behaviors by extending the studies utilizing the concurrent operant paradigm. Furthermore, acceptability of assessments and interventions has become a topic of interest to applied behavior analysts under the term of social validation. However, no single study was found on the opinions and ideas of practitioners about the social validation of reinforcer assessment procedures. Yet, reinforcers are the most frequently used procedures in the training of individuals with disabilities. Thus, it is extremely valuable to know the opinions of the practitioners about such a frequently utilized procedure. In addition, it is hoped that striving to identify reinforcers in Turkey will be a way of evaluating efficacy of reinforcer identifying procedures under Turkey's conditions. Therefore, the purpose of this study was twofold: first, to extend previous research on reinforcer assessment by comparing effectiveness of stimuli identified by two methods on teaching figure coloring to three children with disabilities in Turkey; and second, to determine opinions of special education teachers about two preference assessment procedures.

Phase 1: Preference Assessments

Method

The purpose of Phase 1 was to identify participants' preferences by using two assessment procedures (paired-stimulus and stimulus preference assessment procedures).

Participants

Participants were three preschool children in a self-contained classroom at a school for students with disabilities. Their diagnoses were Down syndrome. Participants demonstrated at least one voluntary measurable behavior such as touching, looking at, and holding—used to indicate approach to various stimuli. All participants had the necessary gross and fine skills for coloring. Physical examinations demonstrated that both auditory and visual skills of participants were within normal limits. Before starting the study, parents of participants were informed about the study and written

permission was obtained for their child's participation. Real names of the participants were replaced with pseudonyms during reporting.

Murat was a 5-year-old boy who was ambulatory and nonverbal, but he was learning several basic signs, such as "go," and "up." He responded to one step simple commands (e.g., "look", "write", "come"). He showed excitement by laughing and clapping. He could turn his head toward the sound when his name was called. He could not take care of daily living skills such as toileting, dressing, or taking a shower independently. He was working on feeding himself with finger foods. He had a few leisure skills, such as watching TV and listening to music.

Can was a 6-year-old boy. He was ambulatory, and had well-developed receptive and expressive language skills.

Mert was a 5-year-old boy. He understood two-step commands that occurred frequently in his daily routine. He was ambulatory, able to follow simple commands, and communicated through gestures and some functional words (i.e., water, food, come, go). Mert and Can could meet their daily needs independently.

Setting

Sessions were conducted in the subjects' self-contained classroom measuring 3 m by 3 m, with an observation mirror on one of the walls. The room contained six tables and ten chairs facing each other and a variety of toys and educational materials. Individually identified stimuli were in plastic containers that could not be seen by participants.

All sessions were videotaped. A certified special education teacher conducted all sessions with the participants. Independent data collectors, working as teaching assistants at Anadolu University, collected inter-observer agreement and procedural reliability data.

All assessment sessions lasted between 10 and 15 min depending on responsiveness of participants. Sessions were conducted two times a day, two to four days a week, as individual schedules permitted.

Stimuli

Ten of the same edible and leisure (i.e., toys) stimuli were selected for each participant. Criteria for the selection of stimuli were availability within their environment, age–appropriateness, and ease of presentation. Stimuli selected according to results of the interviews with and observations of each participant were a ball, child's telephone, car, baby doll, child's iron, pretend/role play, biscuit, chocolate, legos, and a truck. Stimuli used during the assessments were not delivered to participants in other training programs in the participants' school.

Data Collection and Reliability

The recording system selected for this phase was discrete categorization (Kazdin, 1982). It was used because selection behaviors had discrete beginnings and endings, and could be classified as either performed or not performed. The experimenter recorded whether a selection or non-selection response to the stimulus occurred in each trial. Selection was defined as the participant moving toward the stimulus with hand or arm, touching it, or holding it with both hands or one hand within 10 s. Non-selection was defined as the absence of holding, touching within 10 seconds, pushing the stimuli away, and turning his/her head away from the stimulus. Data were collected on selection responses for each stimulus in each trial during paired-stimulus preference and stimulus preference assessments. Percentages of selection response trials were calculated for each stimulus by dividing the number of selected stimuli trials by number of trials in which that stimulus was presented and multiplying by 100 (Hagopian, Rush, Lewin, & Long, 2001).

Reliability was conducted across paired stimulus preference and stimulus preference assessments by using the point-by-point method. A second observer collected data independently during an average of 40% of the trials. If both observers recorded a behavior as nonselection or as a selection during the same trial, an agreement was recorded. On the other hand, if one observer recorded a behavior as selection and the other did not, a disagreement was recorded, and vice versa. Inter-observer agreement was calculated by dividing agreements by agreement plus disagreements and multiplying by 100 (Kazdin, 1982). Mean percentage of agreement across subjects was

99.8% (range 97.8%-100%) during stimulus preference assessment, and 97.4% (range 96.9%-100%) during paired stimulus preference method.

Procedural reliability data were also collected to make sure that the two procedures—used to identify the preferred stimuli of the participants—were administered by the experimenter equally and reliably. To collect procedural reliability data, 40% of the videos used to record both assessment procedures were selected randomly. Procedural reliability data were collected by using a Procedural Reliability Data Record Sheet that contained all steps of both procedures. Watching the videos, observers recorded which steps the experimenter fulfilled and which one(s) the experimenter missed. Then, procedural reliability was calculated by dividing the number of occurrences for each step across all sessions by the number of occurrences plus non-occurrences, and multiplying by 100 (Billingsley, White, & Munson, 1980). Mean percent agreement on interobserver reliability was 98% (range 97.60%-100%) across all subjects during paired-stimulus preference assessment, and 96% (range 95.54%-99.75%) across all subjects during stimulus preference assessment. Mean percent agreement on procedural reliability was 100% across all subjects during both preference assessments.

Procedure

Prior to Phase 1, interviews and direct observations were conducted with each participant to identify potential stimuli. Using an interview form, the participants' teacher and parents were queried about what he liked, what objects he preferred, and what activities he enjoyed getting involved in. The form consisted of four domains: edibles, tactiles, socials, and activities. Each domain had at least 10 items. Following the interview, parents were asked to rank the stimuli generated from the interview form according to their predictions regarding child preference. Based on responses from the interview, 10-15 potentially preferred stimuli were identified for each subject (see Table 1).

Following the interviews, direct observations of participants were conducted while participants were in the classroom. Participants were observed during 10 min sessions at various times for two days. During the observation time, he was engaged with different items. Results of observations showed that 10 potentially-preferred stimuli could be used. Eight of the stimuli were the same items identified by the parents during interviews (see Table 2).

TABLE 1

List of Potential Reinforcers for Participants Based on Interview Results

Participants	*Preferred Stimuli*
Murat	Bubbles, singing, pencil, music, star, fruit, juice, popcorn, puzzle, pizza, candy, chocolate, wagon, mirror, baby doll.
Can	Kosh ball, bingo, crayons, banana, play dough, magazine, soda, doritos, musical toy, ball, turtle, chocolate.
Mert	Book, bingo, puzzle, plastic turtle, X-men, pizza, soda, chocolate, play dough, pencil, computer, baby doll.

In the first phase of the study, paired-stimulus preference and stimulus preference assessments were held. In the paired-stimulus method, preference for 10 stimuli was assessed using procedures described by Fisher et al. (1992). During the paired-stimulus method, each item was paired with all other items at least once. Afterwards, the participant was presented with randomized pairs of items 5 cm apart from each other and on a table approximately 0.5 m in front of the participant. After the participant's attention was secured, he was asked to select the one he wanted. When the participant touched or consumed one of the two items within 10 s every time it was presented, the response was recorded as a selection and the other stimulus was removed. Then, the participant was allowed to access the stimulus for 30 s. Attempts to approach, touch or hold both stimuli were immediately stopped. If the participant did not approach either stimulus within 10 s, the experimenter removed both stimuli, recorded it as a non-selection, and initiated a new trial. The experimenter waited 15 s between presentations of different stimuli.

TABLE 2

List of Potential Reinforcers for Participants Based on Direct Observations Results

Participants	*Preferred Stimuli*
Murat	Singing, fruit, juice, popcorn, puzzle, pizza, candy, chocolate, baby doll, child's phone, ball, truck.
Can	Bingo, crayons, banana, soda, doritos, ball, turtle, chocolate, car, truck, legos, biscuit, baby doll.
Mert	Book, bingo, puzzle, X-men, soda, chocolate, play dough, computer, legos, pretend play.

In the stimulus preference assessment, the same 10 stimuli presented during paired stimulus method were also assessed using stimulus preference assessment based on procedures described by Pace et al. (1985). During the study, the same 10 stimuli were individually presented ten times each in counterbalanced order. In each session, one stimulus was placed on a table approximately 0.5 meters in front of the participant. If the participant approached the item, this resulted in access to the stimulus for 5 s. If the participant did not approach a stimulus after 5 s, the experimenter removed the stimuli. For example, a session would begin by placing a child's phone in front of the participant. If the participant approached, touched or held the child's phone, the child's phone was activated for 5 s. If there was no approach, touch or holding, the experimenter removed the stimuli and initiated a new session. The experimenter waited 15 s between presentations of different stimuli.

The preference hierarchy from both paired-stimulus preference and stimulus preference assessments was developed by ranking stimuli based on average percentage of trials during which selection response occurred across the trials. Stimuli selected in 85% or more of the trials were used in Phase 2.

Results and Discussion

Figure 1 shows results obtained from the two preference assessment procedures as percentage of trials on which selection responses were observed for each participant. All participants were very responsive for 60% or more of the stimuli under paired–stimulus preference assessment, except for Murat. However, they were generally unresponsive, selecting only half of the assessment stimuli in 50% or less of the trials during stimulus preference assessment. But, Murat's selection responses were higher during stimulus preference assessment than paired stimulus preference. In other words, Mert, Can, and Murat, displayed a high level of selection for six, five, and four of ten stimuli respectively during paired stimulus preference assessment. Edible stimulus was found to be the most often selected preference for all three subjects during both of the preference assessment procedures.

Based on results of Phase 1, it is possible to conclude that both assessment procedures were able to identify preferred stimuli for all participants. As demonstrated in Phase 1, all participants differentially selected the stimuli. Data also support previous research that suggested that patterns of responding were idiosyncratic across the participants (e.g., Fisher et al., 1992; and Pace et al. 1985). For example, a baby doll could serve as a reinforcer for participant one, but not for the other participants. Furthermore, participants preferred the same stimuli equally in both assessment procedures. Paired stimulus preference assessment also identified several stimuli as preferred that were not identified through stimulus preference. These results show that preference responses of the participants increased if they were given choices, except for Murat. In other words, paired stimulus preference expanded the selection behaviors of the participants compared to stimulus preference. Furthermore, personal control and dignity were developed by the participants because they were given opportunity to make decisions with the paired stimulus preference assessment (Dunlap, DePerczel, Clarke, Wilson, White, & Gomez, 1994).

Hagopian et al. (2001) stated that the main disadvantage of the stimulus preference procedure is that some individuals may approach all of the stimuli presented, which restricts the range of relative preference identified via the assessment. Results of the current study are consistent with this study, and reveal that Murat displayed the same pattern of responding

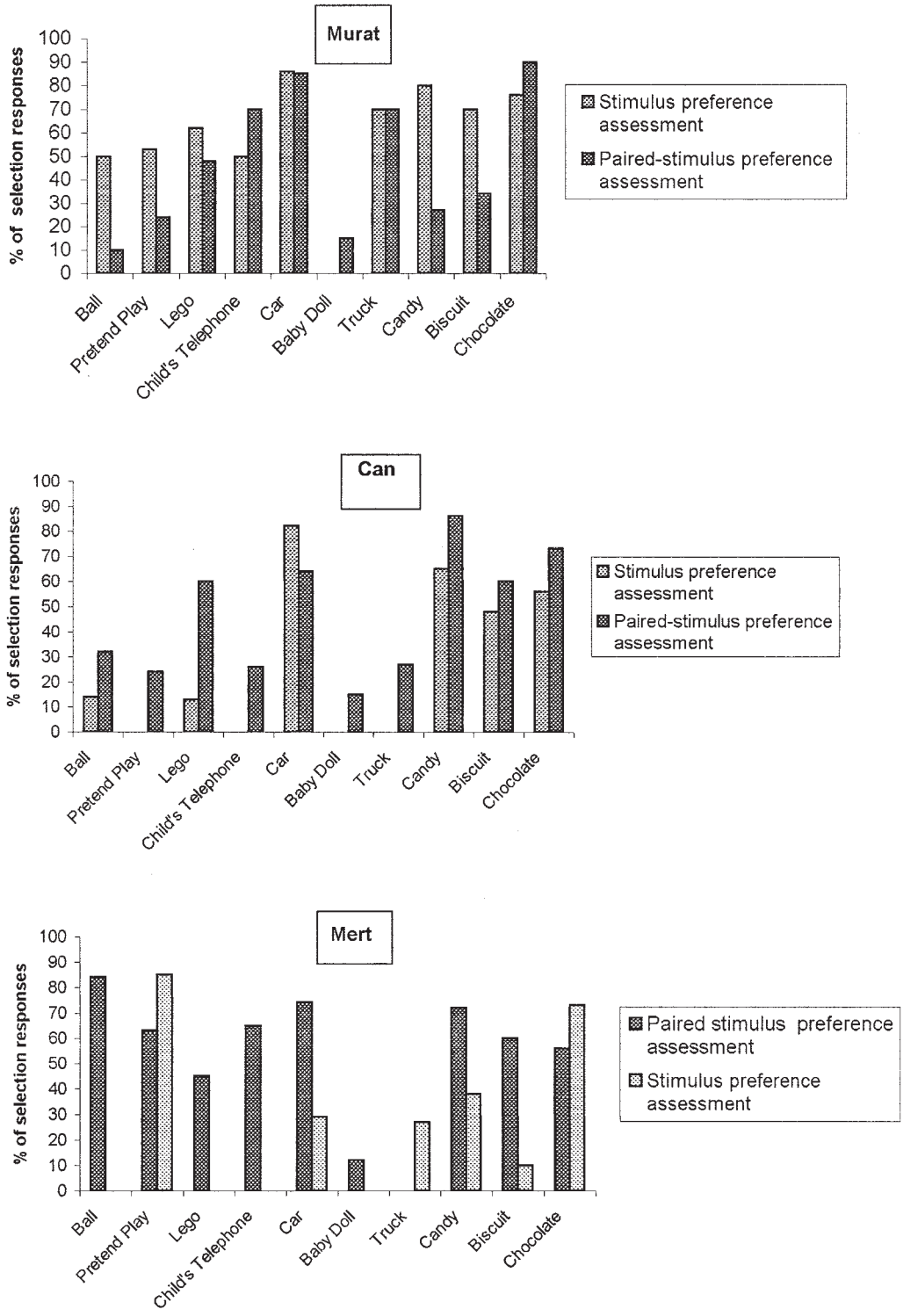


Figure 1. Percentage of selection responses during paired stimulus and stimulus preference assessments for Murat, Can, and Mert.

when being administered the stimulus preference method.

The most suitable stimuli for participants were identified by interviewing parents and teachers in the first phase of the first step of this study. When the stimuli determined according to the observation and interview results were compared with the ones identified by using both preference assessment procedures, no relation could be found between them. Similar to the results obtained by Green, Reid, Canipe, and Gardner (1991) and Green, Reid, White, Halford, Brittain, and Gardner (1988), results of the current study show that the stimuli ranking identified according to the outcomes of interviews with parents and direct observations of participants are not the same as the ranking obtained after conducting any of the stimuli identifying methods. Therefore, caregivers nominated items have been shown not to correlate with empirically identified preferred items (Green et al., 1988, 1991).

Preference assessment is a critical step although the data reported by Fisher et al. (1992), Green et al. (1988), and Pace et al. (1985) do indicate that preference does not equate with reinforcer function. Therefore, reinforcing effects of the stimuli must be evaluated in the second step of preference assessment.

Phase 2: Reinforcer Assessment

The purpose of this phase was to evaluate reinforcing effects of the most often selected stimuli both from the stimulus preference and paired-stimulus assessment procedures.

Method

Participants and Setting

Participants, setting, and session schedules were the same as in Phase 1.

Task

Target response for each participant was to color a simple figure (e.g., a picture of a rabbit). A coloring task was selected for several reasons. First, a coloring task was functional and age-appropriate for the participants. Second, it was suitable for their motor abilities. Third, it was derived from participant-instructional objectives. Fourth, participants already had some steps of coloring task in their repertoire; therefore, it was easy to differentiate reinforcer effects. If a more complex skill that they did not have in their repertoires had been chosen, it would have been difficult to differentiate between skill deficits and reinforcers ineffectiveness (Piazza et al., 1996). The task of picture coloring was broken down into smaller components (see Table 3).

Experimental Design

Adapted alternating treatments design was used to evaluate effectiveness of both reinforcer assessment procedures to teach figure coloring task. During baseline sessions, data were collected at least three times from all participants regarding their figure coloring skills. After baseline was over, two different tests conditions were conducted. In one test condition, the most highly preferred stimuli from paired-stimulus assessment were presented in response to participant's correct response. In the other test condition, the most highly preferred stimuli from stimulus preference assessment were presented. Participants were exposed to these conditions randomly and order of presentation of these conditions was alternated from one session to another. Thirty minute breaks were given between each condition to help participants discriminate conditions.

Each condition was implemented once a day with each participant, in the morning and

TABLE 3

Task Analysis

Figure Coloring
1. Picking up markers or paint brush from the table.
2. Holding the markers or paint brush with hand.
3. Moving hand and arm with backward and forward movements until all area is colored while keeping marker in contact with the paper.
4. Staying in the line of the figure.
5. Putting the markers or paint brush back on the table.

afternoon. Sessions for each condition were conducted in the same manner with the teacher, materials, pictures, and time constant throughout the study. The only difference between the two sessions involved the stimuli identified from paired-stimulus and stimulus preference assessment that was presented during the session. During each condition, percentage of correct responses for figure coloring was recorded. Predictive validity of the preference assessment procedure could be determined if correct responses were high during the presentation of the most preferred stimulus identified using paired stimulus assessment compared to the presentation of the most preferred stimulus identified by stimulus preference assessment.

Data Collection and Reliability

Video recording was used to gather data. Videotapes were then analyzed. During this phase, plus (+) and minus (−) signs were used respectively for correct and incorrect responses. Correct response was defined as independent completion of each step of the task analysis. Incorrect response was defined as the participant completing the step with a prompt or the participant not showing any response. Percentage of correct steps was computed by dividing the number of correct responses by the total number of responses. The result was multiplied by 100 and graphed.

Reliability was conducted across baseline and both test conditions by using the point-by-point method. Two observers collected the data simultaneously but independently. Interobserver agreement was calculated by dividing agreements by agreement plus disagreements and multiplying by 100 (Kazdin, 1982). Mean percent agreement across participants for correct responses was 94.87% (range, 92.30%-96.45%) during baseline, 96.5% for correct responses, 97.8% (range 91.6%-to 100) during stimulus preference condition (the first test condition), and 98.2% (range 93.6%-100%) during paired-stimulus preference condition (the second test condition). Then, procedural reliability was calculated by dividing number of occurrences for each step across all sessions by number of occurrences plus non-occurrences, and multiplying by 100 (Billingsley, et al., 1980). Mean percent agreement on procedural reliability was 100% for baseline and both test conditions across all participants.

Baseline sessions. The task of coloring a simple figure was broken down into smaller components of each skill (see Table 3). Baseline data were collected on all three participants by using each step of the task analysis. During baseline sessions, (a) materials (paper, and painting materials) were placed in front of the participant on a table and verbal cue was given (e.g., "color the picture"), (b) 5 s were allowed for the participant to perform the first step of the task analysis, (c) if the participant responded correctly a "+" was recorded, and the next step of the task analysis was assessed, (d) if the participant responded incorrectly, a "−" was recorded and the step was completed for the participant, and assessment continued with the next step. The same procedures were followed through the task analysis steps. No reinforcement or prompting was given during baseline sessions.

Test conditions. Prior to each condition, the participant was given verbal instructions about contingencies (e.g., "If you color the picture, you'll get a candy"). In the first condition, the experimenter gave the paper and painting materials to the participant and gave the verbal cue "color the picture." The experimenter allowed 3 s to complete the first step of the task analysis. If the participant performed the first step correctly, a plus (+) was recorded in the data sheet, and the stimuli identified from paired-stimulus preference assessment were given to the participant. If the participant did not complete the first step of the task analysis within 10 s, physical and verbal prompts were provided (i.e., hand over assistance to the participant by coloring the picture while saying "pick up markers from the table") to complete the required step of the task analysis. Then, the most preferred stimuli identified from paired-stimulus preference assessment with a verbal reinforcer were given to the participant. The same procedures were used through the task analysis steps.

In the second condition, procedures were the same as in the first condition, except that the most highly preferred stimuli identified from stimulus preference assessment were provided in response to the participant's correct response.

Results and Discussion

Effects of preferred stimuli from both preference assessment procedures on each participant's performance of figure coloring are shown in Figure 2. Results show that both preference assessment procedures were effective in identifying reinforcing stimuli for three students with Down syndrome since their performance increased during test conditions relative to baseline condition. However, baseline performances of the subjects were generally high (M = 52.55; range of individual means 50% to 54.33%) since participants had the coloring skills in their repertoires. All participants' correct responses increased following the administration of preferred stimulus identified from both preference assessment procedures (M = 89.39, range of individual means 82% to 93.57%). During the first test condition in which the most preferred stimuli identified from paired stimulus preference assessment were delivered, the participants' performances were high (M = 92.79, range of individual means 89% to 93.57%). During the second test condition in which the most preferred stimuli identified from stimulus preference assessment were delivered, the participants' performances were high as well (M = 86, range of individual means 82% to 89%). A comparison of individual performances across test conditions showed that all of their performances exceeded 92.79% under the first condition in which the stimuli identified from paired stimulus assessment were used.

The purpose of this phase was to evaluate the reinforcing effects of the most highly selected stimuli identified from both the stimulus preference assessment and paired-stimulus assessment. Since correct responses were high during first test condition compared to second test condition, it could be concluded that the stimuli identified from paired stimulus assessment were more effective in teaching figure coloring to three participants with Down syndrome compared to the stimuli identified from stimulus preference assessment. Since sessions were conducted before lunch for Murat and Can, it may be logical to state that deprivation increased the effectiveness of edible stimuli (candy and chocolate).

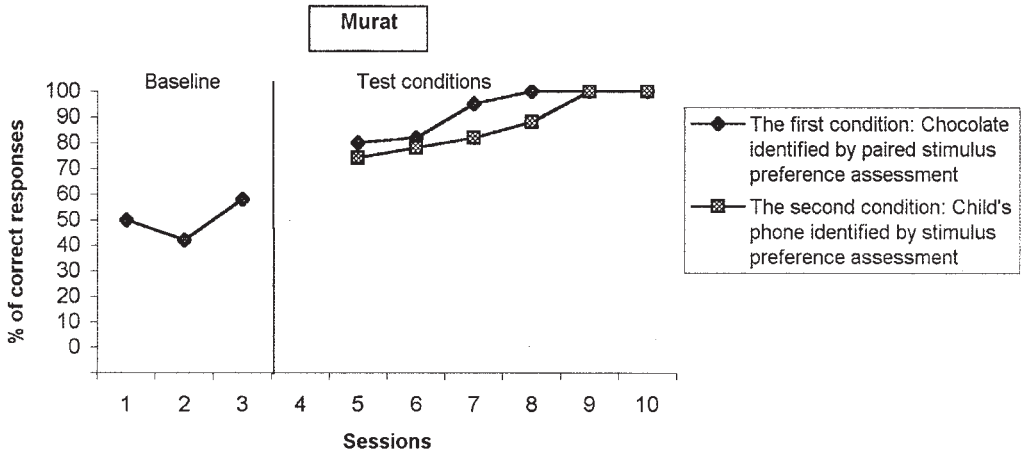
Several limitations need to be acknowledged in the current study. That the participants displayed more correct responses in the first test condition can be attributed to the stimuli they had been given to fulfill the related steps of figure coloring in the second test condition. In other words, the stimuli given in the second test condition to make the participants display correct responses related to coloring skills might have run the risk of affecting the responses the participants produced in the first condition; and this may be the most significant limitation of the present study. Although we tried to counter balance test conditions for this study, it may be concluded that carry over effect (one of the limitations of alternative treatments design) was effect in the second phase of the current study.

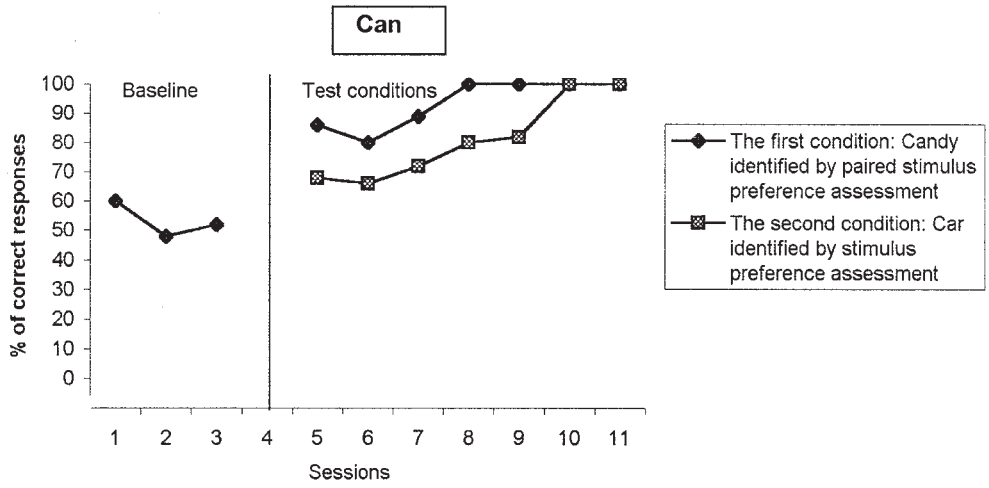
The current study also has a potential limitation regarding use of edible stimuli in the preference assessment techniques. The study was conducted in a short time, and edible stimuli were given as small pieces to impede the occurrence of satiation. Therefore, no satiation or eating routine was observed. If it had taken a long time to complete the study, the behavior changes participants displayed in Phase 1 would not have been this high. Thus, prolongation of the study would require a constant evaluation to determine whether the identified stimuli were still preferred or not.

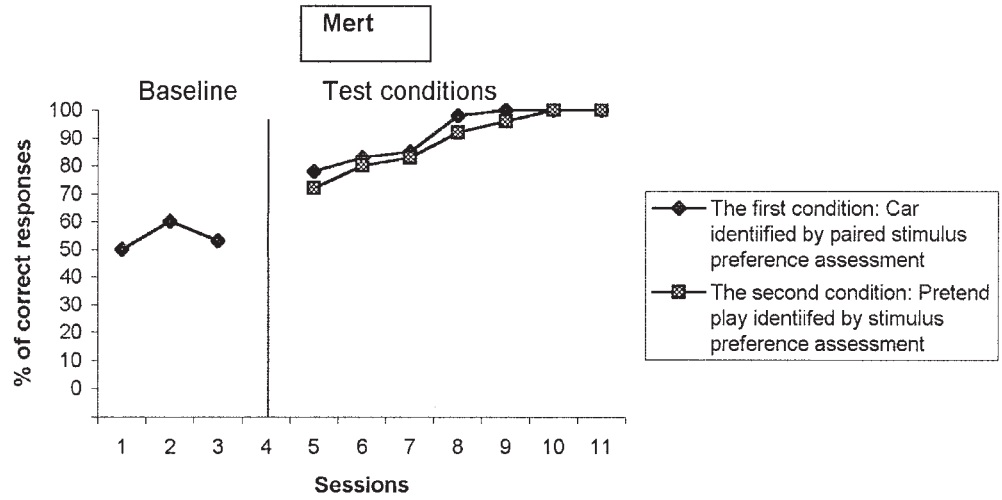
Use of verbal reinforcers together with the preferred stimuli in both test conditions might be another reason that data gathered in test conditions of the second phase were high. Since the preferred stimuli were not the sole cause of participants showing correct responses during teaching of coloring skills, the verbal reinforcers' potential influence must be taken into consideration when evaluating the current study.

In spite of these limitations, results of the study have some implications for practice and future research. Based on findings of the study, it may be suggested that teachers and practitioners should stop depending on their subjective opinion and start holding systematic preference assessments to determine individual preferences of their students in Turkey. Furthermore, they should also test if the students' preferences would serve as reinforcers or not.

A second contribution of the current study is application of preference identification pro-


Murat
Baseline
Test conditions
% of correct responses
100
90
80
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10
Sessions
The first condition: Chocolate identified by paired stimulus preference assessment
The second condition: Child's phone identified by stimulus preference assessment


Can
Baseline
Test conditions
% of correct responses
100
90
80
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11
Sessions
The first condition: Candy identified by paired stimulus preference assessment
The second condition: Car identified by stimulus preference assessment


Mert
Baseline
Test conditions
% of correct responses
100
90
80
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11
Sessions
The first condition: Car identiiified by paired stimulus preference assessment
The second condition: Pretend play identiifed by stimulus preference assessment

cedures to children with Down syndrome in a natural context (i.e., in their school and during the instructional routines). Considering the criticisms claiming that the target skills were simple and arbitrary in the previous studies (Hagopian et al., 2001; Piazza et al., 1996), the experimenter selected a skill (dependent variable) that typically developing preschool children could produce, and that was in the daily instructional lives of the participants in the current study.

Finally, it is concluded that both of the preference assessment procedures were effective to determine reinforcers that are used to improve the quality of education provided in institutions serving children with disabilities in Turkey.

The target task of this study was already present in the subjects' repertoire, and this suggests that it is possible to conduct future research assessing effectiveness of the preferred stimuli as reinforcers by choosing a new and functional skill that is not in the participants' repertoires.

Evaluating whether participants display any problem behavior during preference identification procedures, and which procedures cause more problem behaviors, are other topics of interest for future studies.

Phase Three: Social Validation of Preference Assessment Procedures

The purpose of this phase was to find out what special education teachers participating in the study think and feel about the two preference identification procedures.

Participants

Four special education teachers participated in Phase 3, one conducted the current study and the other three worked in the same school where the current study was conducted. They have been working with students with developmental disabilities for 10 to 14 years. None of them has attended any theoretical or applied workshops about preference assessment procedures, nor have they any knowledge concerning the procedures in question.

Data Collection

A social validation form was used to find out the opinions of the special education teachers who conducted research in this present study. The social validation form consisted of 18 questions. The first nine and the following nine questions of the scale aim to find out how the special education teachers perceive paired stimulus preference assessment and stimulus preference assessment, respectively. The fifth and eighth items of the scale were adapted from another scale developed by Witt and Elliot (1985) to judge acceptability of the behavior changing techniques applied to problem behaviors.

The teacher administering this study was asked to fill out the social validation form at the end of the study. The other three special education teachers watched 20% of the videos of Phase 1 and 2 and filled out the social validation form. Then, answers they gave were analyzed by calculating frequency of the participants' responses for each item on the form.

Results and Discussion

Findings of the social validity analyses revealed that all four special education teachers stated that they could make use of both the stimulus preference and paired stimulus preference procedures in their classrooms. Moreover, they also added that they might advise other teachers to use one of these techniques. However, teachers said that characteristics of the students and the time limitations of the person who would to administer the preference assessment procedure would dictate which method would be used. Therefore, the availability of more procedures to be used to assess preferences and reinforcers is important.

There is enough evidence providing empirical support for the efficacy of both procedures in the literature about preference as-

Figure 2. Percentage of correct responses during baseline, the first test and the second test conditions for Murat, Can, and Mert.

sessment, and teachers working in Turkey should be informed about these procedures. Nevertheless, results of the current study present the ideas of four special education teachers. Other special education teachers should be encouraged to contribute their opinions about the procedures to confirm these results. Despite the fact that both procedures were found to be acceptable to some extent, results showed that paired stimulus was more acceptable than stimulus preference. Further research could be done to shed light on factors affecting the teachers' preferences. These studies might help researchers to make necessary and effective modifications on the procedures to expand their acceptability. Such research can foster the development of procedures that are not only effective but also likely to be applied by the relevant people.

References

Billingsley, F. F., White, O. R., & Munson, R. (1980). Procedural reliability: A rationale and an example. *Behavioral Assessment, 2*, 229–241.

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

Dunlap, G., DePerczel, M., Clarke, S., Wilson, S. W., White, R., & Gomez, A. (1994). Choice making to promote adaptive behavior for students with emotional and behavioral challenges. *Journal of Applied Behavior Analysis, 27*, 505–518.

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

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 systematic assessment of preferences. *Journal of Applied Behavior Analysis, 21*, 31–43.

Kazdin, A. E. (1982). *Single-case research design*. New York: Oxford Press.

Hagopian, L. P., Rush, K. S., Lewin, A. B., & Long, E. S. (2001). Evaluating the predictive validity of a single stimulus engagement preference assessment. *Journal of Applied Behavior Analysis, 34*, 475–486.

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.

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

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.

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

Witt, J. C., & Elliot, S. N.(1985). Acceptability of classroom management strategies. In T. R. Kratchowill (Ed.), *Advances in school psychology* (pp. 251–288). Hillsdale, NJ: Lawrence Erlbaum.

Received: 30 April 2003
Initial Acceptance: 6 June 2003
Final Acceptance: 10 September 2003