A COMPREHENSIVE EVALUATION OF REINFORCER IDENTIFICATION PROCESSES FOR PERSONS WITH PROFOUND MULTIPLE HANDICAPS

CAROLYN W. GREEN, DENNIS H. REID, VICKI S. CANIPE, AND SHIRLEY M. GARDNER

WESTERN CAROLINA CENTER, MORGANTON, NORTH CAROLINA

We evaluated comprehensively a preference assessment for identifying reinforcers for persons with profound multiple handicaps. Four experiments were conducted involving 18 individuals. Results of Experiment 1 replicated previous findings in that the assessment identified student preferences for respective stimuli, and caregiver opinion of preferences did not coincide with the systematic assessment. Results of Experiment 2 indicated highly preferred stimuli were likely to function as reinforcers in training programs, whereas stimuli not highly preferred did not function as reinforcers. Results of Experiment 3 suggested the 12 stimuli used in the assessment represented a comprehensive stimulus set for identifying preferences, although the utility of the set sometimes could be enhanced by caregiver opinion. Results of Experiment 4 indicated the assessment identified preferences likely to be maintained over time. Overall, results are discussed in terms of identifying limits and alternatives to a behavioral teaching technology when applied to persons with profound multiple handicaps.

DESCRIPTORS: profound multiple handicaps, severely handicapped, assessment, reinforcer, technology

An area of growing interest in applied behavior analysis is the identification of reinforcing stimuli for use in skill-training programs with persons who have severe handicaps. A major impetus for the increased interest is the recognition that such stimuli are usually critical to the success of training programs (Mason, McGee, Farmer-Dougan, & Risley, 1989; Pace, Ivancic, Edwards, Iwata, & Page, 1985), and it is often difficult to find reinforcers for this population (Wacker, Berg, Wiggins, Muldoon, & Cavanaugh, 1985). One group of persons with severe handicaps for whom it has been especially difficult to identify reinforcing items and events is individuals with profound mental and physical disabilities (e.g., Bailey & Meyerson, 1969; Zucker, D'Alonzo, McMullen, & Williams, 1980). These persons generally represent the most difficult-to-teach group of the entire population of individuals with severe handicaps (Landesman-Dwyer & Sackett, 1978). In fact, the frequent nonresponsiveness to behavioral training endeavors by individuals who have profound multiple disabilities has led several prominent professionals to question whether an effective behavioral technology currently exists to teach this population (Bailey, 1981; Ellis, 1981).

Recently, we have attempted to develop reinforcer identification processes for persons with profound multiple handicaps (Green et al., 1988) by extending the work on reinforcer identification previously successful with persons who have less serious disabilities (Pace et al., 1985). Results of our initial work were encouraging in regard to using stimulus preference assessments to predict what stimuli were likely to function as reinforcers in behavior-change programs. However, for some individuals we were not able to identify any highly preferred stimuli. Relatedly, we found no support for stimuli to function as reinforcers if the stimuli were not highly preferred, although we did not evaluate this type of stimuli thoroughly.

In light of the research findings to date, continued investigation seems warranted to evaluate more comprehensively the efficacy of stimulus preference assessments as a means of identifying reinforcers

We thank Jim Bodfish, Ed Konarski, and Jim Phillips for their comments on an earlier draft of the paper, and Carole Daves, Mary Carswell, and Mary Keller for their assistance in preparing the manuscript. Appreciation is also extended to the many educators at Western Carolina Center who assisted in implementing the assessment and training procedures.

Requests for reprints should be sent to Dennis H. Reid or Carolyn W. Green at Western Carolina Center, Morganton, North Carolina 28655.

for use with persons who have profound multiple handicaps. One particular reason for continued investigation is that this line of research in essence is beginning to test the limits of a behavior-change technology (cf. Bailey, 1981). That is, by identifying reinforcers for some individuals and not for others, and demonstrating adaptive behavior change only for individuals for whom a reinforcer has been identified, conclusions may be drawn regarding how a behavioral teaching technology based on positive reinforcement may be limited with some individuals (i.e., those for whom no reinforcers can be identified). Of course, before such conclusions are scientifically appropriate, a given technology should be evaluated comprehensively and should include replications of outcomes of investigations and examination of as many relevant variables as possible.

This investigation extended the comprehensiveness of the evaluation of preference assessment procedures as a method of identifying reinforcers for persons who have profound multiple handicaps. First (Experiment 1), we attempted to replicate the results of Green et al. (1988) by using a behavioral assessment process to identify stimulus preferences and by comparing results of the assessment with the more traditional caregiver opinion method of identifying preferences. Next (Experiment 2), we further attempted to replicate the results of Green et al. by evaluating whether stimuli assessed to be highly preferred would function as reinforcers in training programs. Experiment 2 also extended previous work by more thoroughly assessing whether stimuli not highly preferred have reinforcing value. Subsequently (Experiment 3), we evaluated the comprehensiveness of the initial pool of target stimuli used in previous research by examining whether preferences for stimuli in addition to the initial pool could be determined. Finally (Experiment 4), we assessed the durability of stimulus preferences.

GENERAL METHOD

Eighteen individuals participated in the four experiments, 6 of whom participated in two or more of the experiments. All participants lived in a public residential facility and received day treatment services through a youth or adult educational program staffed by certified special education teachers and teacher assistants. All experimental procedures were conducted by the educational staff under the experimenters' supervision.

Each participant (mean age 29 years; range, 14 to 55) had multiple handicaps, including mental retardation at the extreme lower end of the profound range (Grossman, 1983) and profound physical impairments. Each individual was nonambulatory and exhibited at least two other disabilities (e.g., visual and/or auditory impairment, seizure disorder, spastic diplegia, spastic quadriplegia, or hypertonicity). These individuals were selected for the study because, based on a review of the participants' records and recommendations from educational staff, each had a history of nonresponsiveness to training programs. In general, the participants displayed disabilities characteristic of individuals considered to have the most profoundly handicapping conditions (see Landesman-Dwyer & Sackett, 1978; Reid, Phillips, & Green, 1991, for a detailed description of this population) and were totally dependent on caregivers for survival.

The preference assessment and basic behavior-change procedures employed were the same as those described by Green et al. (1988). Hence, these procedures will be described briefly here except where procedural alterations occurred.

EXPERIMENT 1

Method

Participants and target stimuli. Six individuals participated in Experiment 1. Twelve stimuli were selected for preference assessments based on availability and ease of presentation, representation of a range of sensory input, and frequent (attempted) use as reinforcers with this population in previous research (Green et al., 1988). The stimuli were organized into four groups of three stimuli each. Stimulus presentation formats are presented in Table 1. Throughout all experiments, the students maintained their regular daily routine involving meals, treatment programs, and so forth, except for the addition of the experimental procedures.

Table 1

Presentation Format for Assessment of Stimulus Preferences

Stimulus	Presentation format
Hug	Assessor places both hands around upper arms/shoulders of student
Verbal interaction	Assessor talks to student for 3 s alternating to left, right, and front sides of student
Vibrator	Assessor strokes arm of student with vibrator for 3 s
Juice	For sighted student, assessor places cup of juice in visual field of student; for visually impaired, assessor places lip of cup to side of student's cheek
Pudding	For sighted student, assessor places spoon in visual field of student; for visually impaired, assessor places tip of spoon to lip of student
Rock music	Tape player with music on (increased volume for hearing impaired) presented to left, right, and front sides of student
Soft music	Same as with rock music
Tactile mitt	Assessor rubs student's arm with mitt
Light board	Assessor places light board in visual field of student; student's hand placed on switch that activates light board
Mechanical toy	Assessor activates toy for 3 s within visual field of student, student's hand/arm placed on toy
Hand-held toy	For sighted student, assessor presents toy by laying it on student's table top; for visually impaired, assessor touches toy to student's preferred hand
Hand clap	Assessor claps hand three times to the front, right, and left sides of student

Behavioral definitions and assessment procedure. The target behavior of *approach* was defined as the student making an apparent voluntary body movement toward the stimulus, maintaining contact with the stimulus for at least 3 s, exhibiting a positive facial expression, or making a positive vocalization within 8 s of the initial presentation of the stimulus. Observers monitored for a maximum of 5 s following each stimulus presentation and recorded the occurrence or nonoccurrence of approach behaviors. Reliability checks occurred during 23% of all assessment trials and involved all students. Reliability was calculated on a trial-by-trial basis for overall, occurrence, and nonoccurrence agreement percentages for approach behaviors by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100 (Bailey & Bostow, 1979). For each student, overall, occurrence, and nonoccurrence reliabilities each averaged (mean) at least 90%.

Experimental sessions. Each student was presented with each stimulus 36 times (30 assessment trials and six primer presentations). Three stimuli were presented each session in a counterbalanced fashion across sessions. A session began with a 5-s presentation of the stimulus (primer) in which the student was prompted to touch, taste, and/or look at the stimulus. Next, five trials were conducted. A trial began when the assessor presented a stimulus to the student as described in Table 1. If the student exhibited an approach behavior, the stimulus continued to be presented for an additional 5 s; otherwise the stimulus was removed and a new trial was initiated.

Staff opinion survey. Staff opinion of student preferences for each stimulus was assessed with a Likert-type survey, with scale values ranging from 5 (most preferred) to 1 (least preferred). Direct-care and professional personnel who worked with each student rated the student's apparent preference for each item. Three to 8 staff members completed a survey for each student (mean of 5 per student).

Results

Table 2 presents the percentage of approach behaviors to each stimulus averaged across all assessment sessions per student. Using an 80% criterion of approach behaviors as representing a highly preferred stimulus (Pace et al., 1985), all students except MG highly preferred one or more stimuli.

To compare results of the systematic assessment with the caregiver opinion assessment of preferences, the numerical value reported by staff on the Likert scale for each stimulus for each student was

Table 2
Mean Percentage of Approach Behaviors Across All Stimulus Presentations in Experiment 1

	Student					
Stimulus	SB % AP[a]	JT % AP	LM % AP	MR % AP	HC % AP	MG % AP
Hug	50	13	86	53	97	3
Vibrator	57	57	56	83	80	16
Verbal interaction	53	23	60	67	37	3
Mechanical toy	13	70	6	47	20	66
Clap	40	17	6	60	37	13
Tactile mitt	67	10	3	3	7	6
Juice	7	97	20	87	87	43
Rock music	17	30	3	37	60	46
Hand-held toy	87	70	20	7	60	26
Light board	50	97	7	88	40	63
Soft music	47	20	0	40	57	40
Pudding	67	97	0	100	100	63

[a] AP = approach behavior.

averaged across staff recordings. The 12 stimuli were then ranked for each student based on the average scores. Similarly, the 12 stimuli were ranked for each student according to the average percentage of approach behaviors across assessment sessions. A Spearman rank correlation coefficient was then calculated using the rankings on the systematic and opinion assessments. For 4 students, there was no statistically significant correlation between preferences based on the two rankings. For 2 students, the correlations were significant at the .10 and .05 level (r values of 0.49 and 0.58, respectively).

Discussion

Results of Experiment 1 support the results of Green et al. (1988) in three ways. First, the systematic assessment reliably indicated stimulus preferences among persons with profound multiple handicaps. Second, there are some individuals with profound multiple handicaps who do not appear to have a strong preference for any of the 12 stimuli evaluated. Third, caregiver opinion was not predictive of preferences for most students. For 2 students (for whom a significant correlation existed), the results are somewhat discrepant from those of Green et al., who found no significant correlations. However, given that a significant correlation was found for only 2 of 6 students, the results here (similar to those of Green et al.) suggest that caregiver opinion is not reliably predictive of stimulus preferences. The latter results highlight the importance of systematically evaluating stimulus preferences.

In light of the results of Experiment 1 regarding use of systematic assessment to identify stimulus preferences, Experiment 2 was designed to evaluate whether highly preferred stimuli would function as reinforcers in training programs. Experiment 2 also evaluated whether stimuli not highly preferred would function as reinforcers.

Experiment 2

Method

Participants and target stimuli. The 6 participants in Experiment 1 participated in Experiment 2. In addition, 3 other students with profound multiple handicaps participated. These 3 students were selected because they attended the same day treatment program as the other students and had not approached any stimulus on 80% or more of the trials when assessed using the same procedures described in Experiment 1.

For the students who highly preferred various stimuli, the stimuli were grouped into four categories (see *Results* concerning Student SB for an exception): high systematic/high opinion, high sys-

Table 3
Target Behaviors, Dependent Measures, and Stimuli Used Contingently with Students Showing a Stimulus Preference

Student	Target behavior	Dependent measure	Stimulus[a] High	Stimulus[a] Low
MR	Activate switch	Prompt level	Pudding	Hand toy
JT	Touch hand	Prompt level	Pudding	Mitt
SB	Activate switch	Prompt level	Candy	Mechanical toy
HC	Activate switch	Prompt level	Pudding Hug	
LM	Put object in container	Number of seconds	Hug	

[a] High = stimulus ranked among the four stimuli with highest level of approach behaviors on systematic assessment; low = stimulus ranked among the four stimuli with lowest level of approach behaviors.

tematic/low opinion, low systematic/high opinion, and low systematic/low opinion. A stimulus was considered to be high systematic if it ranked at least within the four most preferred stimuli on the systematic assessment for a given student and was low systematic if it ranked within the four least preferred. Similarly, high and low opinion meant the stimulus was among the four most preferred or least preferred, respectively, based on the opinion assessment ranking.

For the students who did not highly prefer any stimulus, the stimuli approached most frequently (the stimuli had to be approached on at least 50% of assessment trials) were used in training programs. Across all stimuli with these students, the average occurrence of approach behaviors to the stimuli used in the training programs was 63%, with no stimulus having been approached on more than 70% of the trials.

Behavioral definitions and measurement. A summary of the target behaviors and training programs for each student is presented in Table 3 (see also Table 4). All target behaviors were selected from individual program plans. For five training

Table 4
Results of Contingent Stimulus Applications in Training Programs with Students Who Demonstrated No Stimulus Preferences

Student	Target behavior	Experimental conditions (results)				
MH	Eye contact on request (percentage correct)	Baseline (51%)	Stimulus A rock music (42%)	Stimulus B verbal interaction (38%)	Stimulus C hand clap (53%)	
MG	Activate switch (prompt level)	Baseline (1.0)	Stimulus A light board (1.0)			
	Activate (simplified) switch (prompt level)	Baseline (1.4)	Stimulus B mechanical toy (1.2)	Stimulus C pudding (1.7)	Stimulus A light board (1.4)	
MF	Eye contact on request (percentage correct)	Baseline (45%)	Stimulus A verbal interaction (45%)	Stimulus B rock music (38%)		
NR	Activate switch on cue (percentage correct)	Baseline (0%)	Stimulus A flicker bulb (1%)	Stimulus B soft music (7%)	Stimulus C rock music (1%)	Stimulus D coffee (1%)
	Place object in container (amount of time)	Baseline (67 s)	Stimulus D coffee (79 s)	Baseline (65 s)	Stimulus D coffee (56 s)	

programs (Students SB, JT, MR, HC, and MG), the prompt level required for the student to perform the task was recorded per trial.

Reliability checks occurred during 26% of the training sessions, including during 27 of 30 experimental conditions across students. Across all sessions with reliability checks (750 trials), there were only 12 disagreements regarding the required prompt level.

For training programs designed to reduce the time required to complete a task (Students LM and NR), reliability checks occurred during 12% of all sessions, including during each experimental condition for each student. On 95% of the training trials with reliability checks (90 trials), there were no disagreements between observers regarding the recorded length of time. Finally, for programs designed to increase the percentage of trials with correct responses following a trainer's request (NR, MH, MF), reliability checks occurred during 24% of all sessions, including during each condition for each student, with only two disagreements regarding the occurrence of a correct response.

Experimental procedures: Baseline. During baseline, a graduated prompt sequence based on the locus of physical contact of the trainer's prompt (O'Brien, Bugle, & Azrin, 1972) was used to train targeted skills for Students SB, JT, MR, HC, and MG, as well as for one of Student NR's skills. For example, for Student SB's target behavior of activating an adaptive switch, the trainer began by saying, "(name), press switch." If the student did not comply within 5 s, the trainer repeated the instruction and partially guided the student by placing her hand on the student's elbow. If the student again failed to comply, the trainer repeated the instruction and increased the physical assistance by partially guiding the student at the wrist. The final level of prompt, if needed, was a verbal instruction with hand-over-hand full physical guidance by the trainer.

Instead of the graduated prompting sequence used with other students, Students LM and NR were given an initial cue to complete the task that was followed by an additional cue if the student was not attempting to complete the task after a designated period of time. The training programs for MF and MH were similar to those for LM and NR. Throughout all baseline sessions, no trainer-controlled consequence was provided following any student response. Training sessions consisted of either 5 or 10 discrete trials, depending on each student's individual program plan.

Experimental procedures: Contingency conditions. During each contingency condition, a stimulus from one of the above stimulus groups was applied for 3 to 5 s contingent on a designated level of performance. For Students SB, JT, MR, HC, and MG, the stimulus was provided when the student performed the task at the least intrusive prompt level required to evoke the behavior during baseline. As each student progressed to a less intrusive prompt, the criterion for stimulus presentation changed accordingly. Each training session began with a primer trial during which the trainer verbally cued and physically guided the desired behavior and then immediately presented the stimulus. Procedures for Students LM and NR were similar, except the presentation of the stimulus was dependent on the student performing the task in fewer seconds than his or her average time during the preceding three sessions. For the programs targeting increases in correct responses (NR, MH, and MF), the stimulus was presented after each correct response.

Experimental design. The experimental design was a sequential treatment design with an experimental reversal embedded within treatments. Specifically, following baseline, the reinforcing effects of a given stimulus were assessed. If behavior change occurred relative to baseline, then a reversal to the baseline condition was conducted to demonstrate functional control of the stimulus as a reinforcer. If behavior change was not apparent, another stimulus was selected and provided contingently. The latter process continued until (a) a behavior change occurred relative to the preceding condition, at which point a reversal to the preceding condition (or baseline) was conducted; or (b) all stimuli to be evaluated had been applied contingently. One exception was with Student LM, for whom three AB designs were implemented.

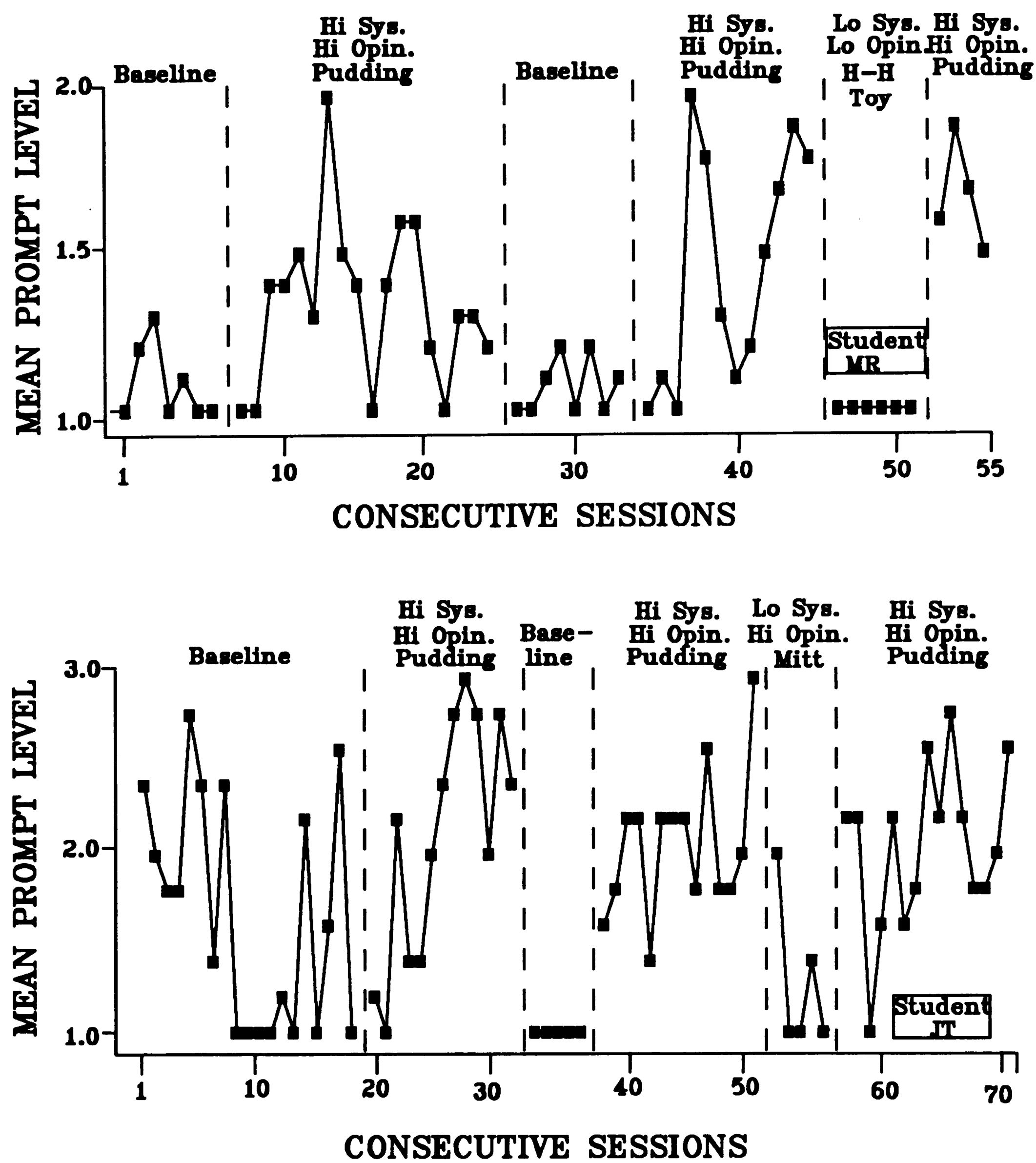


Figure 1. Mean prompt level (with 1.0 representing the most intrusive) required to evoke Student MR's (top panel) and Student JT's (bottom panel) completion of the target task for each session during each experimental condition.

Results

Performance of students with stimulus preferences. Results for Student MR (Figure 1) indicated a stimulus systematically assessed to be highly preferred (high systematic) functioned as a reinforcer, whereas a stimulus assessed not to be highly preferred (low systematic) did not. Specifically, when pudding, a systematically assessed, highly preferred stimulus (which was also ranked as highly preferred by staff), was applied contingently, the mean prompt level required to evoke MR's activation of an adaptive switch increased from a baseline mean of 1.07 to 1.33 (prompts were scored from low to high based on a most-to-least intrusive prompting sequence such that the higher the score, the more independence shown by the student). During the return to baseline, the mean level decreased to 1.07 and then increased to 1.45 when the highly pre-

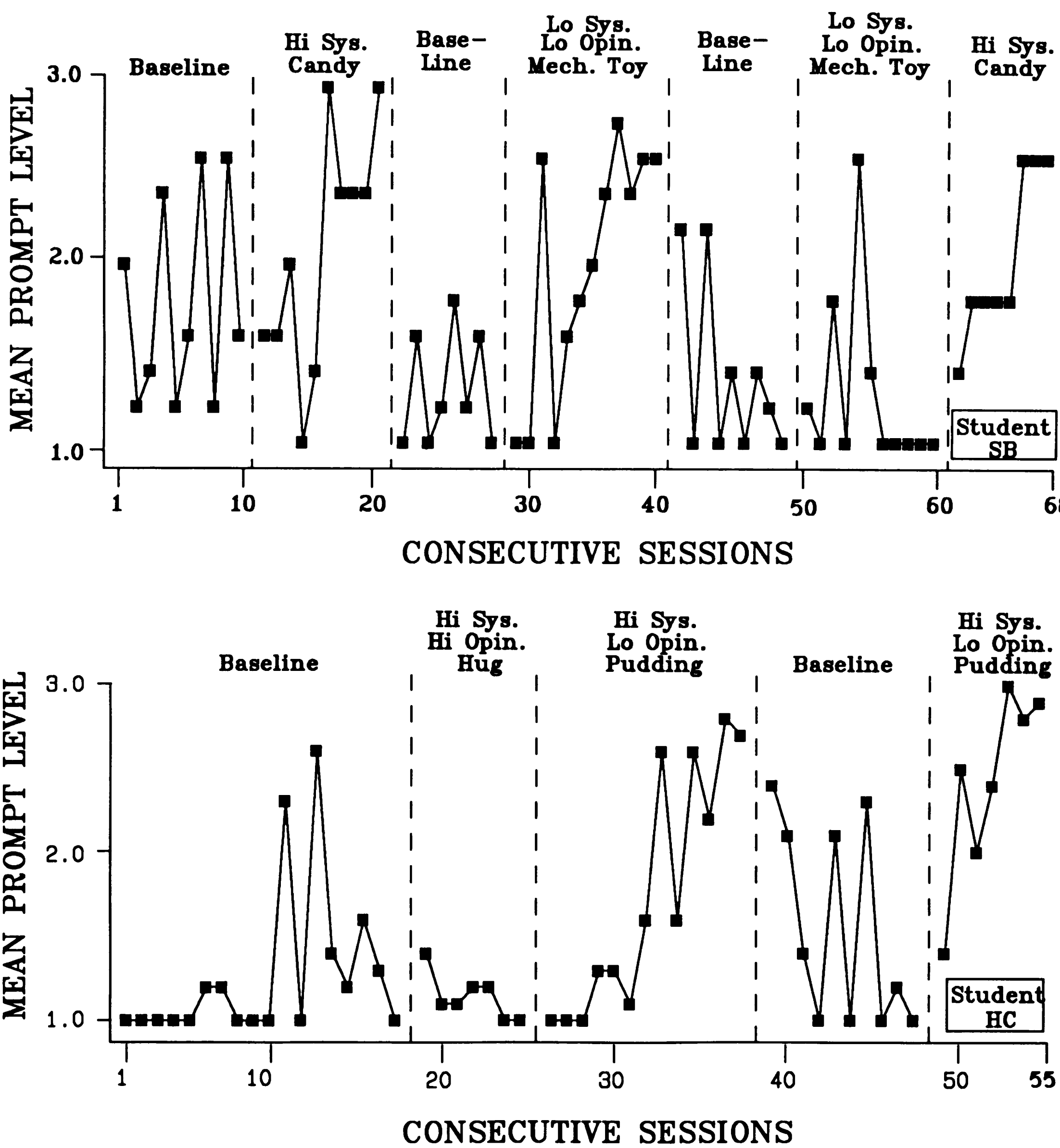


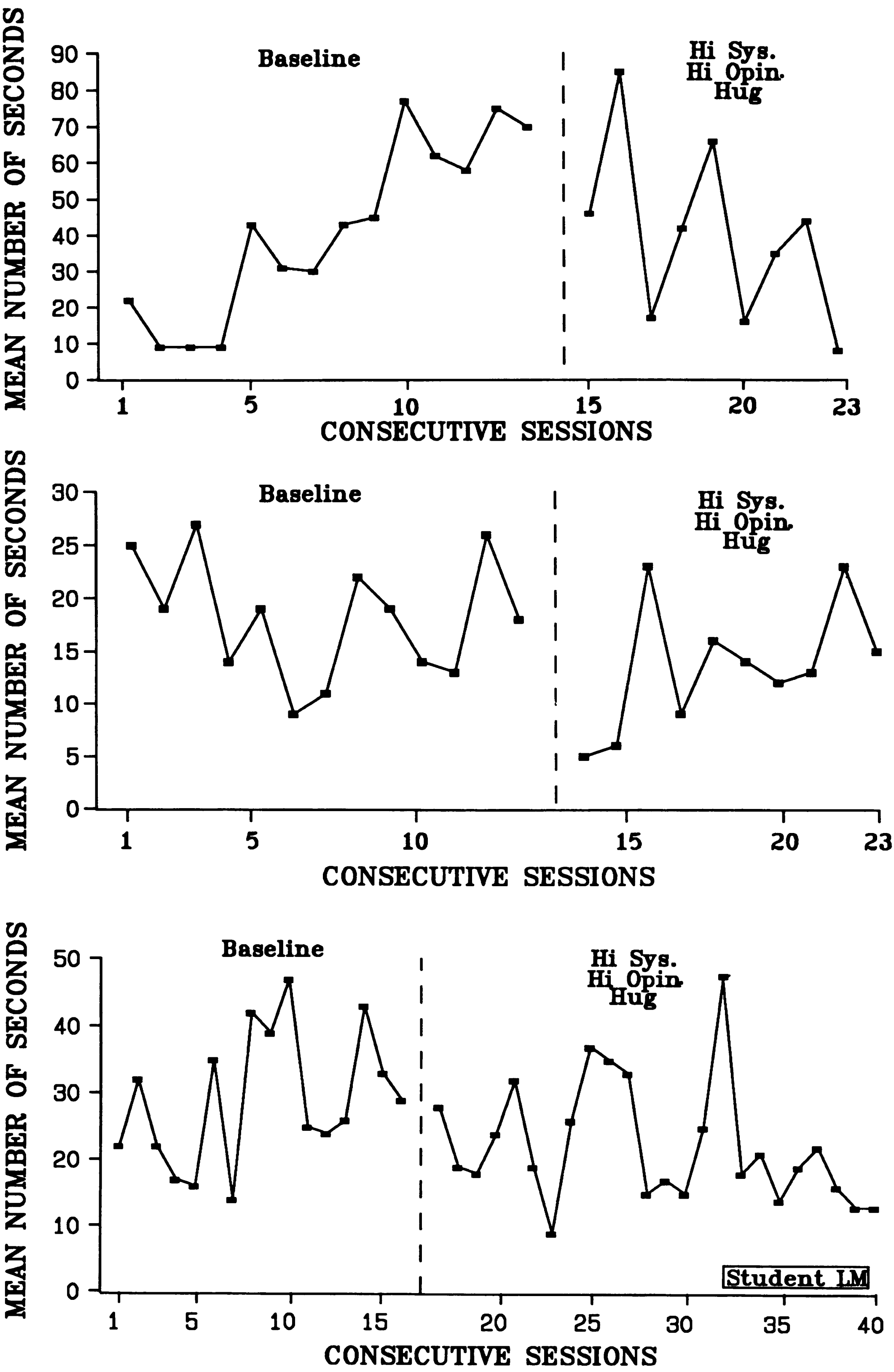
Figure 2. Mean prompt level (with 1.0 representing the most intrusive) required to evoke Student SB's (top panel) and Student HC's (bottom panel) completion of the target task for each session during each experimental condition.

ferred stimulus was reapplied. In contrast, when a hand-held toy, a stimulus that was not highly preferred (and was also ranked low by staff) was applied, the mean prompt level decreased to 1.0. Finally, when the highly preferred stimulus was again reapplied, the mean prompt level increased to 1.67.

Results for Student JT (Figure 1) were similar

→

Figure 3. Mean number of seconds required for Student LM to place an object in a container for each session during each experimental condition. Student LM was in a different wheelchair and prescribed body position for each program implementation represented in the top, middle, and bottom panels, respectively.


MEAN NUMBER OF SECONDS
Baseline
Hi Sys.
Hi Opin.
Hug
0
10
20
30
40
50
60
70
80
90
1
5
10
15
20
23
CONSECUTIVE SESSIONS
MEAN NUMBER OF SECONDS
Baseline
Hi Sys.
Hi Opin.
Hug
0
5
10
15
20
25
30
1
5
10
15
20
23
CONSECUTIVE SESSIONS
MEAN NUMBER OF SECONDS
Baseline
Hi Sys.
Hi Opin.
Hug
Student LM
0
10
20
30
40
50
1
5
10
15
20
25
30
35
40
CONSECUTIVE SESSIONS

to those for Student MR. The mean prompt level required to evoke JT's touching the trainer's hand increased relative to baseline (1.46 average for both baseline conditions) when pudding, a systematically assessed, highly preferred stimulus (which was also ranked high by staff), was applied contingently (2.09 average across conditions with the preferred stimulus). There appeared to be no reinforcing effects (mean of 1.28) of a tactile mitt, a stimulus systematically assessed not to be highly preferred (which was ranked high by staff).

For Student SB, a mean prompt level of 1.72 was required to evoke her switch activation during the first baseline (Figure 2), with an increase to 2.08 when a stimulus (candy) was applied that was systematically assessed to be highly preferred, and a decrease to 1.27 when baseline was reinstated. Staff opinion had not been assessed for the candy because it was selected based on SB's mother's recommendation (see Experiment 3). There was also an increase in prompt level (1.97) relative to baseline when a mechanical toy, a stimulus that was not highly preferred on the systematic assessment (and was ranked low by staff), was applied, with a decrease (1.37) when baseline was reimplemented. However, the increased prompt level was not replicated when this stimulus was reapplied (1.27). Subsequently, an increased prompt level was replicated (2.05) with the reintroduction of the highly preferred stimulus.

Results for Student HC (Figure 2) were somewhat discrepant in that a hug, a systematically assessed, highly preferred stimulus (which was also ranked high by staff), was not accompanied by an increased prompt level (1.14) relative to baseline (1.26). However, when a second stimulus was applied that was systematically assessed to be highly preferred (pudding, which was ranked low by staff), an increase occurred (1.75) with an increasing trend within the condition. There was a subsequent decrease (1.55) and decreasing trend when baseline was reinstituted and another increase (2.42) with an increasing trend when the stimulus was reapplied.

Relative to results for the 4 students just noted, results for Student LM provided rather minimal support for the reinforcing value of a stimulus that was systematically assessed to be highly preferred. On three occasions after a hug, a highly preferred stimulus (which was also ranked high by staff), was applied to reduce LM's time in completing a task, his wheelchair was modified by physical therapists to allow him to be in the most therapeutic body position. Hence, the highly preferred stimulus was applied and evaluated three separate times to correspond with each wheelchair and body position adaptation. During the first baseline the required time steadily increased (Figure 3), then steadily decreased when the highly preferred stimulus was applied contingent on LM completing the task in less time. With the second wheelchair/position (Figure 3), the baseline average of 18 s decreased to 13 s when the stimulus was applied. However, an increasing trend was apparent during the latter condition, at which point his chair/position was altered, thereby disallowing an opportunity to continue to evaluate the effects of the current stimulus condition. With the third chair/position, a decrease occurred from 29 s during baseline to 22 s during the contingent stimulus condition, with a slight decreasing trend during the latter condition.

Overall, at least one stimulus that was highly preferred on the systematic assessment was accompanied by rather clear behavior change for 4 students and a small amount of change with 1 when applied contingently on target behaviors. Among the applications of stimuli that were not highly preferred on the systematic assessment (regardless of the opinion ranking), no consistent behavior change occurred.

Performance of students who did not highly prefer any stimulus. For the 4 students who did not approach any stimulus on 80% or more of the trials of the preference assessments, no reinforcing effects of any stimulus were apparent. A typical behavior pattern of these students is represented by Student NR's data (Figure 4). Student NR's percentage of correct responses did not increase when any of four different stimuli were applied contingently (one stimulus was based on a second preference assessment in an attempt to find stimuli that evoked more approach responses; see Experiment

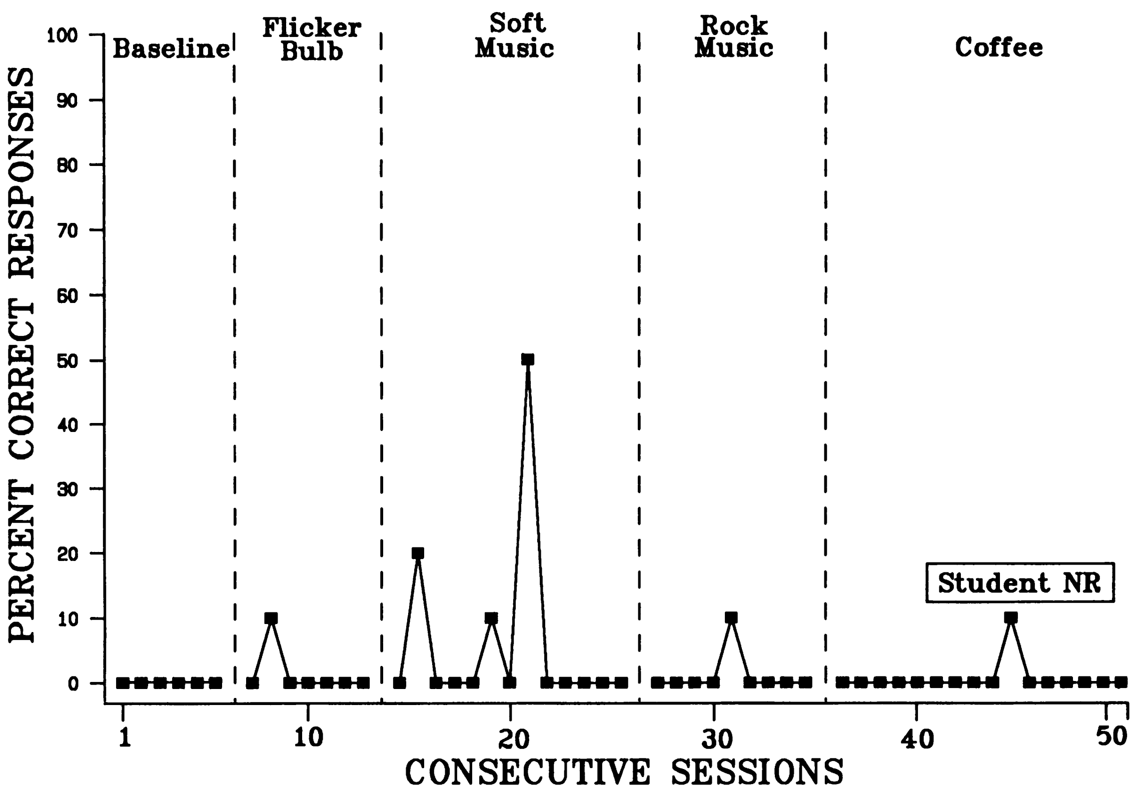


Figure 4. Percentage of correct responses for Student NR during her training program for each session during each experimental condition. The stimuli identified at the top indicate the respective stimuli applied contingently within each experimental condition (flicker bulb represents a modified light board).

3). When a different training program was later implemented with Student NR, no consistent behavior change was apparent either. Results for all students who did not highly prefer any stimulus are summarized in Table 4.

Discussion

As in Experiment 1, results of Experiment 2 support the results of Green et al. (1988) in several ways. First, stimuli assessed systematically to be highly preferred were likely to function as reinforcers when applied contingently in training programs. A highly preferred stimulus appeared to function to at least some degree as a reinforcer for all participants (although in both investigations not every stimulus that was highly preferred had reinforcing effects). This ability to predict stimuli likely to have reinforcing effects assumes heightened importance given the traditional difficulty noted earlier in finding reinforcers for this population.

A second way in which results support the findings of Green et al. (1988) is that there was no consistent evidence to indicate that stimuli not assessed systematically to be highly preferred would function as reinforcers. No consistent behavior change occurred with the three stimuli not highly preferred applied to 3 students who did highly prefer other stimuli. Also, with the six training programs (eight different stimulus presentations) involving 4 students who did not highly prefer any stimulus, no consistent behavior change was apparent (students who did not highly prefer any stimulus did not participate in training programs in the study of Green et al.).

In general, results of Experiments 1 and 2 support the importance of conducting systematic pref-

Table 5
Percentage of Approach Behaviors to Additional Stimuli in Experiment 3

Student	Target stimuli assessed	Percentage of approach behaviors
KJ	preferred staff member's voice	10
NR	coffee	67
	jewelry	3
	fingernail polish	20
SB	candy	100
LM	Pepsi	46
	ball play	13
JP	Pepsi	100
VT	coffee	50
	chocolate	43

erence assessments prior to using stimuli as reinforcing consequences in training programs with this difficult-to-teach population. However, several questions with the assessment process remained, including the degree to which the group of 12 stimuli represents a sufficient sample for formulating conclusions regarding which stimuli should be used as potential reinforcers. Experiment 3 was designed to address this question by evaluating potential preferences for other types of stimuli.

Experiment 3

Method

Participants. Each student ($N = 6$) had participated previously in a stimulus preference assessment as described in Experiment 1. Four of these students approached at least one stimulus on at least 80% of the previous assessment trials, whereas 2 did not approach any stimulus on 80% of the trials.

Stimuli, behavioral definitions, and assessment procedures. Behavioral definitions and assessment procedures were identical to those in Experiment 1 except that only the target stimuli were assessed. The stimuli assessed for preferences were selected based on discussions with the students' caregivers (teachers, teacher aides, direct-care staff, nurses, recreators, physical therapy assistants, and/or parents). Caregivers were questioned regarding what they thought a student really liked beyond what was included in the initial 12-item assessment. Based on the caregiver recommendations, one stimulus was assessed with 3 students, two were assessed with 2 students, and three with 1 student.

Reliability observations were conducted as in Experiment 1, during 31% of all assessment sessions, and involved each student and each stimulus. Averaged across all students and stimuli, occurrence, nonoccurrence, and overall reliabilities for approach behaviors each were at least 94%. For each student and each stimulus, reliability fell below 90% on only one occasion.

Results and Discussion

The target stimuli for each student and corresponding percentage of approach behaviors are presented in Table 5. One highly preferred stimulus was identified for 2 students. One of these students did not highly prefer any stimulus during the initial assessment of 12 stimuli, whereas the other student did highly prefer one stimulus.

Results indicated that for the most part (i.e., for 4 of 6 students), the initial pool of 12 stimuli established using the criteria described in Experiment 1 represented a comprehensive set of assessment stimuli for identifying highly preferred stimuli. Nevertheless, because for 2 students an additional highly preferred stimulus was identified based on caregiver opinion (and especially with regard to the student who had not highly preferred any of the 12 initial stimuli), the results suggest that it is desirable to include stimuli that caregivers identify as favorites within a preference assessment. Inclusion of such stimuli appears, at least in some cases, to enhance the probability of identifying a highly preferred stimulus and subsequently identifying a likely reinforcer to use in a training program (Experiment 2). However, because of the frequent inconsistency of caregiver opinion in terms of identifying specific stimuli that students will consistently approach when systematically assessed (Experiment 1), caregiver opinion should not be relied on exclusively.

In a sense, these results seem to contradict those of Experiment 1. That is, on the one hand, caregiver rankings of student preferences for the 12 stimuli in Experiment 1 did not coincide very closely with the results of the systematic assessments. On the other hand, when caregivers were questioned in Experiment 3 regarding favorite stimuli and those stimuli were systematically assessed, results of the assessment provided some support for caregiver opinion (i.e., 2 students approached a caregiver-selected stimulus on at least 80% of the trials). However, on closer scrutiny these results are not really contradictory because caregivers were responding to two different tasks. Specifically, caregivers may have been knowledgeable about a student's strongest preference in terms of his or her favorite stimulus but not very knowledgeable about preferences that may not be as strong and/or are more specific (i.e., involving a comparison among 12 stimuli). Some support for such an interpretation stems from research with ambulatory persons with profound mental retardation who are more skilled, suggesting that caregivers can identify favorite food items much better than they can identify specific client preferences when comparing multiple pairs of items (Parsons & Reid, 1990).

Results of Experiments 1 through 3 support the utility of the systematic assessment process for identifying highly preferred stimuli for persons with profound multiple handicaps and for identifying likely reinforcers. A question remaining with the process however, was the durability of the preferences. Experiment 4 addressed this question.

Experiment 4

Method

Participants. Twelve persons participated in Experiment 4, 4 of whom had participated in the Green et al. (1988) study.

Stimuli, behavioral definitions, and assessment procedures. The stimuli, behavioral definitions, and assessment procedures were the same as in Experiment 1. In total, 8 facility staff persons (a teacher and teacher aides) participated as assessors. However, the same staff member did not necessarily conduct both assessments for a given student. Eight of the students were assessed by 2 different staff members. The first and second assessments for each student were conducted across varying time periods, ranging from 4 to 28 months.

Table 6
Results of Correlational Analyses Between Preference Assessments in Experiment 4

Student	Amount of time between assessments	*r* value	Significance level
NR	12 months	0.31	NS
KJ	12 months	0.84	.001
TL	28 months	0.47	.10
MH	27 months	0.60	.025
DC	10 months	0.63	.025
JE	18 months	0.86	.001
MG	18 months	0.80	.005
JM	4 months	0.83	.001
JR	16 months	0.95	.001
BS	11 months	0.49	.10
HC	16 months	0.60	.025
PB	4 months	0.40	.10

Reliability observations were conducted as in Experiment 1, during 28% of all assessment sessions involving 7 students for the first assessment, and during 36% of assessment sessions involving 10 students for the second assessment (reliability checks were conducted during the first and/or second assessments for all students). For each assessment, reliability for occurrence, nonoccurrence, and overall agreement each averaged at least 93%. During each assessment for individual students, occurrence, nonoccurrence, and overall reliabilities never averaged below 86%.

Results

To compare results of the two assessments, the stimuli for each student were ranked according to the percentage of trials across assessment sessions on which the student approached each stimulus for each of the two assessments. A Spearman rank correlation coefficient was then calculated using the rankings on each of the two assessments for each student. As indicated in Table 6, relative preferences for the 12 stimuli remained quite consistent across assessments for most students. Specifically, there were statistically significant correlations between the two assessments for 11 students ($p <$

.025 for 8 students and $p < .10$ for 3 students). For 1 student there was no statistical correlation.

Although these correlations indicated that across the entire sample of 12 stimuli, approach behaviors were relatively consistent for most students across assessments, there were some notable changes regarding certain highly preferred stimuli. Specifically, there were 18 stimuli highly preferred on the first assessment (8 of the 12 students), and 12 of those were still highly preferred by the same students on the second assessment (after 4 to 18 months). However, there were also 12 stimuli highly preferred on only the second assessment (involving 7 students after 4 to 28 months). Perhaps most important, all students highly preferred at least one stimulus on either the first or second assessment.

GENERAL DISCUSSION

Overall, results of the four experiments provide rather consistent support for the use of the systematic assessment process for identifying reinforcing stimuli for persons with profound mental and physical disabilities. Highly preferred stimuli were identified for most of the participants (Experiment 1). Subsequently (Experiment 2), those stimuli usually functioned as reinforcers in skill-training programs. Results also suggested that caregiver opinion alone should not be relied on for identifying highly preferred stimuli with this population. Relatedly, the sample of 12 stimuli developed by Green et al. (1988) appeared to represent a relatively comprehensive set of stimuli to include in preference assessment protocols (Experiment 3), although the utility of the pool could be enhanced if needed in some cases (i.e., with individuals for whom no likely reinforcers are identified) by including other stimuli that caregivers recommend as favorites. Finally, the systematic assessment process identified preferences that for the most part were quite consistent over time in regard to relative differences across stimuli, although some important changes also occurred with certain highly preferred stimuli.

Based on the results just summarized, one means of increasing the probability that persons with profound multiple handicaps will achieve some success in behavioral training programs based on a positive reinforcement paradigm would be to conduct a systematic preference assessment. Such an assessment could be conducted prior to beginning a training program to increase the likelihood that stimuli used as consequences in the program would have reinforcing effects. By using highly preferred stimuli in this manner, it should be less likely that students will participate in training programs for extended time periods without any apparent behavior change (Bailey, 1981), the lack of change being due at least in part to the lack of a reinforcer for effecting behavior change. Further, for students who do not highly prefer any stimulus, periodic reassessments should be conducted to determine whether such preferences develop or change, as appeared to be the case in Experiment 4 (see Mason et al., 1989, for similar results with preferences among young children with autism). The practicality of using the preference assessment process should be enhanced when considering that procedures in the four experiments were conducted by the usual trainers of the students. Also, the procedures have been incorporated into the ongoing habilitation process conducted by other educators with other students in the school program in which the current students participated.

Although results of this study support the utility of the systematic assessment process, there nevertheless were some students (as in the initial Green et al., 1988, study) for whom no highly preferred stimuli were identified and, thus, no reinforcers were found. Of course, there is essentially an infinite array of stimuli that could be assessed for potential preferences; future research is warranted to identify other types of stimuli (e.g., vestibular stimulation) likely to be highly preferred. Nevertheless, for students who do not demonstrate any strong preference from month to month, a reasonable question seems to be whether a training approach based on positive reinforcement can be used to teach useful skills to these individuals. That is, if a training program is based on the principle of positive reinforcement and no positive reinforcers can be identified despite repeated attempts across numerous stimuli, then the evidence to date suggests such a training approach will not be effective. In short, teaching useful skills to this subpopulation of persons with profound multiple handicaps may be at least temporarily beyond the limits of our existing behavioral technology based on positive reinforcement.

If indeed, as some research evidence suggests, the handicaps of some individuals may be so profoundly debilitating that it is at least periodically and/or temporarily beyond the limits of our traditional (i.e., positive reinforcement) behavioral technology to provide effective teaching programs, the question arises as to what constitutes appropriate education and/or habilitation for these persons. One possibility would be to explore the application of other behavioral processes in training endeavors (Reid et al., 1991). For example, there has been very little work with classical conditioning paradigms among persons with profound multiple handicaps, although some early success was reported in this area (Rice, 1968). Applications of negative reinforcement processes (Iwata, 1987) could also be evaluated, perhaps using stimuli that individuals appear to avoid (Green et al., 1988) during preference assessments. Considering the latter possibility, however, it is not clear whether programs based on a student with profound handicaps escaping or avoiding an aversive situation in order to acquire a simple skill would be very acceptable to service providers and/or society in general. Future research could address both the efficacy and social acceptability of training programs based on a negative reinforcement principle.

In considering the application and evaluation of other behavioral processes in training programs with persons who have profound multiple handicaps, the criterion for a highly perferred stimulus as used in this study may warrant attention. Our criterion of 80% approach behaviors is based on previous research, most notably that of Pace et al. (1985). Although to date there has been no research support for stimuli to function as reinforcers if the stimuli are approached on less than 80% of assessment trials, it seems unlikely that every stimulus with reinforcing effects must be approached on exactly 80% or more of the trials (e.g., in some cases 75% approach behaviors may suffice). Hence, we recommend that the 80% criterion be viewed primarily as a guideline until further research is conducted. Somewhat relatedly, it may be that varying levels (less than 80%) of approach behaviors to different stimuli could have reinforcing effects in training programs using a response deprivation paradigm (Konarski, Johnson, Crowell, & Whitman, 1980). To date, the response deprivation hypothesis has received very little attention in applied research; however, such research seems warranted with persons who have profound multiple handicaps in light of the difficulty in identifying reinforcers with this population.

Another area that warrants research attention is the development of treatment programs that do not focus exclusively on skill acquisition per se yet attempt to enhance the quality of life of persons with profound multiple handicaps in other ways. Recently, suggestions regarding the development of such programs have been provided (Ivancic & Bailey, 1986; Reid et al., 1991). Given the outcome of this investigation, research seems warranted to explore further alternative treatment programs for persons with profound multiple handicaps and to determine objectively whether such programs do enhance quality of life.

REFERENCES

Bailey, J. S. (1981). Wanted: A rational search for the limiting conditions of habilitation in the retarded. *Analysis and Intervention in Development Disabilities,* **1**, 45–52.

Bailey, J. S., & Bostow, D. E. (1979). *Research methods in applied behavior analysis.* Tallahassee, FL: Copy Grafix.

Bailey, J., & Meyerson, L. (1969). Vibration as a reinforcer with a profoundly retarded child. *Journal of Applied Behavior Analysis,* **2**, 135–137.

Ellis, N. R. (1981). On training the mentally retarded. *Analysis and Intervention in Developmental Disabilities,* **1**, 99–108.

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.

Grossman, H. J. (Ed.). (1983). *Classification in mental retardation.* Washington, DC: American Association on Mental Deficiency.

Ivancic, M. T., & Bailey, J. S. (1986, September). *Assessing operant training potential.* Paper presented at the 6th Annual Meeting of the Florida Association for Behavioral Analysis, Orlando.

Iwata, B. A. (1987). Negative reinforcement in applied behavior analysis: An emerging technology. *Journal of Applied Behavior Analysis,* **20,** 361–378.

Konarski, E. A. Jr., Johnson, M. R., Crowell, C. R., & Whitman, T. L. (1980). Response deprivation and reinforcement in applied settings: A preliminary analysis. *Journal of Applied Behavior Analysis,* **13,** 595–609.

Landesman-Dwyer, S., & Sackett, G. P. (1978). Behavioral changes in nonambulatory, profoundly mentally retarded individuals. In C. E. Meyers (Ed.), *Quality of life in severely and profoundly mentally retarded people: Research foundations for improvement* (pp. 55–144). Washington, DC: American Association on Mental Deficiency.

Mason, S. A., McGee, G. G., Farmer-Dougan, V., & Risley, T. R. (1989). A practical strategy for ongoing reinforcer assessment. *Journal of Applied Behavior Analysis,* **22,** 171–179.

O'Brien, F., Bugle, C., & Azrin, N. H. (1972). Training and maintaining a retarded child's proper eating. *Journal of Applied Behavior Analysis,* **5,** 67–72.

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.

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

Reid, D. H., Phillips, J. F., & Green, C. W. (1991). Teaching persons with profound multiple handicaps: A review of the effects of behavioral research. *Journal of Applied Behavior Analysis,* **24,** 319–336.

Rice, H. K. (1968). Operant behavior in vegetative patients. III: Methodological considerations. *The Psychological Record,* **18,** 297–302.

Wacker, D. P., Berg, W. K., Wiggins, B., Muldoon, M., & Cavanaugh, J. (1985). Evaluation of reinforcer preferences for profoundly handicapped students. *Journal of Applied Behavior Analysis,* **18,** 173–178.

Zucker, S. H., D'Alonzo, B. J., McMullen, M. R., & Williams, R. L. (1980). Training eye-pointing behavior in a nonambulatory profoundly mentally retarded child using contingent vibratory stimulation. *Education and Training of the Mentally Retarded,* **15,** 4–7.

Received September 29, 1989
Initial editorial decision June 21, 1990
Revision received July 23, 1990
Final acceptance February 7, 1991
Action Editor, Terry J. Page