TEACHING RETARDED WOMEN A CLOTHING SELECTION SKILL USING COMMUNITY NORMS

DEBORAH NUTTER AND DENNIS H. REID¹

COLIN ANDERSON CENTER, ST. MARYS, WEST VIRGINIA AND
NORTHERN INDIANA STATE HOSPITAL & DEVELOPMENTAL
DISABILITIES CENTER, SOUTH BEND, INDIANA

Five institutionalized retarded women were taught a clothing selection skill to coincide with popular fashion. Normative data concerning popular styles of color-coordinated clothing were collected through observations of women's apparel in a local community. Using a puzzle simulation of a woman with alternative pieces of colored clothing, color coordination training was conducted using modelling, instructions, practice, praise, and feedback to teach popular selections of color combinations. A generalization training component with actual clothing was also included. Training was accompanied by large increases in percentage of popular color selections with the puzzle and actual clothing. The increases maintained over a seven- to 14-week followup period. Time efficiency measures supported the use of the simulation procedure as opposed to total reliance on actual clothing. The results underscore the utility of local norms in community preparation programs and the efficacy of instructional procedures for severely and profoundly retarded persons as part of their participation in the deinstitutionalization movement.

DESCRIPTORS: clothing selection skills, training, normative data, generalization, simulation training, community preparation, retardates

The national focus on deinstitutionalization and mainstreaming of retarded persons has recently been an area of concern in behavior analysis research. Behavioral technologies have been used in the preparation of developmentally delayed persons for community settings by teaching a variety of skills including money usage (Lowe and Cuvo, 1976), mobility skills (Page, Iwata, and Neef, 1976; Neef, Iwata, and Page, 1978), cooking (Bellamy and Clark, 1977), leisure skills (Johnson and Bailey, 1977), house-keeping (Bauman and Iwata, 1977), and telephone usage (Leff, 1975). Despite the success of such programs, serious concern exists over the

adequacy of training of many retarded persons leaving institutions and the subsequent success of deinstitutionalization in general (Gollay, 1977; Santiestevan, 1976). Continued research on methods of developing the wide array of skills necessary to move from an institution to the mainstream of community living should help alleviate the inadequacies in training.

An area that has not been addressed in the behavioral literature on community preparation is the clothing a retarded individual wears and his/her related physical appearance. The importance of clothing on a handicapped person's lifestyle has been emphasized for decades (Newton, 1976.) Similarly, the significance of a normal style of dress (*e.g.*, garment color) by developmentally disabled persons has been recognized (Gay Apparels, *Note 1*). Particular concerns have been expressed over the development of clothing selection skills by retarded persons (O'Connor, 1976; President's Committee on Mental Retardation, 1973; Tobias and Cortazzo, 1963) and the need to include training in dressing style as preparation for community involve-

¹Appreciation is expressed to the administration of Colin Anderson Center for its support of program activities. Special appreciation is extended to Paul Gault for his assistance in preparing the training apparatus, and to Hilda Gault, Dan Hursh, and Rick Van den Pol for assistance in preparing the final manuscript. Reprints may be obtained from Dennis H. Reid, Northern Indiana State Hospital and Developmental Disabilities Center, South Bend, Indiana 46634.

ment (Thornley, 1976; White, 1976). The importance of these skills is also reflected in federal accreditation standards, which require training in clothing selection in residential facilities (Joint Commission on Accreditation of Hospitals, 1971, sec. 2.3.2.1). Finally, an informal questionnaire survey of 15 administrators from public residential facilities serving retarded persons in five states indicated that 94% believed developmentally delayed individuals should dress in accordance with popular fashion, and 100% agreed that retarded persons should not "stand out" due to unusual color combinations of clothing.² If institutionalized retarded persons are to become part of the community mainstream and not stand out from the rest of the population, the development of skills to select clothing that coincides with popular fashion is warranted. However, despite the noted importance of a retarded person developing these competencies, methods of teaching clothing selection skills have not been systematically developed, described, and evaluated.

A conceptual analysis of the procedures necessary to teach skills commensurate with popular opinion reveals a basic difference from the procedures to teach community skills previously mentioned. Processes involved in programs to develop street crossing and use of public transportation, for example, rely heavily on a task analysis of the necessary behaviors. While clothing selection also involves task analysis to a certain degree, an additional component of major significance is the determination of what clothing variables reflect popular fashion. Before a retarded person can be taught to select clothing to coincide with what other persons in the community usually select, normative data describing the community fashions are necessary. If these data are not readily available, methods to acquire

the information need to be built into teaching programs.

The use of normative data in programs with retarded persons is not new to applied behavior analysis. For instance, normative data have been used to evaluate training in feeding skills (Azrin and Armstrong, 1973; O'Brien and Azrin, 1972). A recent discussion of normative measures in relation to the social validity of applied behavior analysis programs indicates the need for continued research in this area (Kazdin, 1977). Similarly, the significance of normative data in community preparation procedures in particular has been well described (Heiny and Stachowiak, 1976; Kennett, 1977), although supporting research has been limited.

The primary purpose of the present investigation was to evaluate a method for teaching institutionalized, developmentally delayed women to select color-coordinated clothing in accordance with popular fashion based on local norms. A component was also included to evaluate the time efficiency of the simulation teaching process relative to processes more directly related to the natural environment.

A second purpose was to extend the literature on community preparation skills to more seriously retarded individuals. With relatively few exceptions (*e.g.*, Bellamy and Clark, 1977), teaching programs for community involvement have been reported with individuals displaying deficiencies above the severe range. However, the deinstitutionalization movement has also been initiated on a more modest basis with severely and profoundly retarded persons (Berkson and Landesman-Dwyer, 1977; Close, 1977), and teaching procedures for these individuals are needed.

METHOD

Participants and Setting

Five female clients of a state residential facility for developmentally disabled persons participated. The residents ranged in age from 21 to 36 yr, with a mean of 26, and each had resided

²The survey was conducted during a seminar of the Management Training Program, Center for Developmental and Learning Disorders, University of Alabama in Birmingham. Gratitude is expressed to Bill Garove and his staff for their cooperation in conducting the survey.

in the institution for at least 10 yr before the study. Results of traditional psychological evaluations reported their intellectual skills to fall within the severe or profound range of retardation (mean IQ of 25, range 22 to 29). All participants spoke in short sentences and participated in simple conversations. Each person was ambulatory and performed basic self-care tasks independently. Discussions with ward attendants indicated that each woman could verbally identify basic colors. None of the residents displayed serious disruptive or abusive behaviors and all were candidates for a community-based group home, according to information in their files. Another resident had been selected for the project but was excluded immediately before baseline when she was transferred to the group home.

Several criteria were involved in selecting these particular participants. First, each person was recommended by direct care staff to be in need of training in appropriate clothing selection when questioned by the experimenter. Second, only those persons referred by the ward staff who completely dressed themselves each day were selected. Finally, each participant had to exhibit the majority of the skills required for dressing, as measured by the Behavioral Characteristics Progression Checklist (1973). Each client selected regularly performed over 70% of the dressing items on the checklist and each experienced difficulty on the items relating to clothing selection. When selected, the purpose of the program was described to each woman, along with the explanation that they could terminate the project at any time if they desired to do so. No participant chose to terminate throughout the study.

All sessions were conducted in a training room near the participants' living area. It was furnished with a rectangular table, two chairs, and a clothing rack. The trainee's chair was placed at one end of the table with its back against a one-way window leading to an observation room. The instructor's chair was adjacent to the client's and was also visible through the observation window.

Training Apparatus

A wooden puzzle depicting a nude female when intact was designed by the experimenters and served as the primary training device. It was constructed by tracing a commercially available Barbie cutout doll (available from Mattel, Inc.) onto a sheet of plywood (0.80 cm thick) and then cutting around the outline with a jigsaw. The resulting wooden doll, standing 35.6 cm tall was then cut into nine puzzle pieces. The pieces were obtained by dividing the body at the neck, shoulders, waist, thighs, and ankles. The body was painted in flesh tones and facial features were painted by a volunteer from the local community. The puzzle outline was completed by nailing the flat frame from which the doll had been cut to a piece of plywood (37.5 cm by 18.4 cm by 0.80 cm). Hence, one could complete the puzzle by placing the pieces into the indented outline of the woman on the plywood backing.

Simulated clothing for the doll was provided by covering pieces of mattress foam, 0.95 cm thick, with cloth cut from small pieces of fabric. Sleeveless shirts and slacks were constructed by removing the appropriate puzzle pieces and cutting the foam to those dimensions. Selected scraps of fabric were then glued to the foam. Wearing apparel could be placed on the doll by putting the cloth-covered foam puzzle pieces into place for the legs (pants) and torso (blouse).

Fabrics used with the puzzle included a sample of each solid color to be taught as part of a popular color combination between pants and blouse during training. Also included were samples of solid colored clothing, which were included under unpopular color combinations. A total of 15 different color categories for blouses was used and 14 different color categories for slacks. Within each color category there were at least two different shades of the color represented, such as light and dark orange fabrics to represent the orange category.

Normalization Sample

To determine what popular color combinations consisted of, 649 women were observed in

the community where the residents would most likely be moving. Approximately 75 % of all observations were conducted in a large indoor shopping mall, with additional observations in a local restaurant and on the sidewalk of a downtown street. Approximately 12.5 hr of observation time were necessary to complete the sample.

Observers. Observations on the normalization sample were conducted by the senior experimenter, with reliability observations conducted by an undergraduate student assigned to the project as part of a practicum experience, two graduate students serving an internship at the institution, or the other experimenter. The observers were trained before actual data collection through instructions regarding color categories and scoring procedures, questions and answers with the experimenter, practice observations, and feedback.

Observation procedures. During actual data collection, observers were seated several feet apart so that individual observation sheets could not be viewed across observers in order to promote independent recording. Each was equipped with data sheets and a pen with instructions to record the color of the upper and lower garment of each woman who walked in front of where they were seated, with two exceptions. First, the female had to be, in the observer's judgement, at least a teenager or older. Second, the woman had to be wearing two separate, visible garments, excluding coats and hats. In the case of more than one female passerby at the same time, or if one observer questioned whether a female was at least a teenager, the primary observer designated with a head nod or gestures who was to be scored. If two upper garments were visible (e.g., a shirt with a sweater over it), the outermost item was rated. An upper garment was defined as a shirt, blouse, sweater, or blazer. Slacks, shorts, skirts, or jumpers were defined as lower garments. Each clothing item was first scored as either nonsolid or solid. The nonsolid category included all clothing of at least two distinguishable fabric colors. For a sleeved upper garment to be recorded as nonsolid, it had to dis-

play two colors on both the torso portion and the sleeves, whereas sleeveless ones had to display two colors on both front and back. All other items were scored as solid.

After the observer scored the upper and lower garments as either solid or nonsolid, the color of each solid piece was recorded under one of 15 categories: white, brown, tan, black, gray, red, pink, burgundy, yellow, orange, rust, blue, light blue, green, or purple. The 15 colors had been selected by the experimenters as being representative of the majority of colors worn, based on previous observations of women in the community and consultation with a fashion design instructor at a state university. Limiting the number of possible colors provided more structure to the tasks of observing and training the clients (as opposed to using an almost infinite set of color descriptors to represent different color shades) while still permitting color discrimination. Nonsolid color combinations were not included in the teaching process.

The rationale for teaching a clothing selection skill dealing with only solid-colored upper garments with solid-colored lower garments as opposed to teaching other possible combinations (*i.e.*, solid upper/nonsolid lower, nonsolid/solid, and nonsolid/nonsolid) was as follows. First, it was believed that using one combination would be more feasible, considering the available resources for conducting the project, yet the purpose of evaluating the teaching method could still be accomplished. Second, the solid/solid combination was the most frequently occurring combination within the normalization sample. This combination occurred during 47% of the sample, whereas nonsolid/solid averaged 40%, solid/nonsolid 10%, and nonsolid/nonsolid 3%. Finally, pilot observations indicated that nonsolid colored garments were extremely difficult to achieve reliability among observers as to the dominant color.

Reliability. Reliability checks were conducted by two observers recording simultaneously but independently on all observations of the normal-

ization sample. Calculations were initially conducted for the occurrences of each solid color category and the occurrences of each combination of a solid colored upper garment with a solid colored lower garment. Reliability was defined as the number of agreements for each occurrence divided by the number of agreements plus disagreements. For reliability of solid color categories, an agreement was scored only if both observers recorded the same category for a specific piece of clothing worn by one individual. A disagreement was scored if the reliability observer recorded a category not recorded by the other observer or if the reliability observer did not record a category scored by the other. Hence, if one observer scored red and one recorded brown for a particular garment, then a disagreement was calculated for each of those categories. For reliability of solid/solid combinations, an agreement was scored only if both observers recorded a solid/solid combination for the outfit of a particular individual. A disagreement was scored if either observer recorded a solid/solid combination and the other observer recorded any other combination (solid/nonsolid, nonsolid/solid, or nonsolid/nonsolid).

Overall reliability averaged across all solid colors was 0.87 (mean), with a range of 1.00 for the most reliable color to 0.65 for the least reliable. Twelve of the 15 categories were above 0.82. The least reliable category, "light blue", was most frequently confused with "blue". Reliability for the occurrence of solid lower and upper garment combinations was 0.93.

Response Definitions

Response definitions were derived from a matrix of all possible solid color combinations in the normative sample. The matrix summarized the observed frequency of each possible solid color combination of upper and lower garments. Except for combinations including "blue" lower garments, any two colors that occurred together for at least 1% of the total solid/solid combinations were designated popular color combina-

tions and those that never occurred in combination were called unpopular. Combinations occurring between 0 and 1% of the time were labelled neutral. For combinations involving "blue" lower garments, popular matches were defined as those occurring at least 5% of the total time, while unpopular matches were combinations observed 1% of the time or less. The separate criteria were established for "blue" lower garments because of their disproportionately high frequency of occurrence. If the same criteria were used for "blue" as for the other colors, and "blues" frequency data were included with that for all other colors, essentially only "blue" would be popular for a lower garment and thereby limit the amount of color coordinating skills that could be obtained. The resulting solid color combinations included 27 popular (Table 1) and 137 unpopular matches. The 27 popular combinations represented 58% of all solid/solid color combinations recorded by the observers.

In addition to collecting reliability on the occurrence of each individual color as described previously, reliability was also calculated for the occurrence of each specific combination of colors that constituted a popular match. Again, reliability was defined as agreements divided by agreements plus disagreements, where an agreement was counted when both observers recorded identical colors for the upper garment and identical colors for the lower garment for a specific outfit observed. A disagreement was recorded if one observer recorded a combination defined as a popular combination and the other observer recorded a different color upper garment and/or a different color lower garment for that specific individual. For the 27 popular color combinations, reliability averaged a mean 0.86, range 0.44 to 1.00. For 21 of the combinations there was one or less disagreement. The combination with the lowest reliability, "blue/light blue", was most frequently confused with "blue/blue". More detailed information is presented in summary form in Table 1.

Table 1
Summary of Popular Color Combinations and Interobserver Reliabilities

Color			Color		
Upper Garment	Agreements		Upper Garment	Agreements	
Lower Garment	Agreements + Disagreements	Reliability	Lower Garment	Agreements + Disagreements	Reliability
White/gray	5/7	0.71	Gray/gray	3/3	1.00
White/white	5/6	0.83	Red/black	3/4	0.75
White/brown	6/6	1.00	Red/red	5/7	0.71
White/black	8/8	1.00	Pink/pink	9/9	1.00
White/orange	3/4	0.75	Burgundy/burgundy	3/3	1.00
White/rust	2/3	0.67	Yellow/green	3/3	1.00
White/blue	20/26	0.77	Orange/orange	4/5	0.80
White/green	10/10	1.00	Rust/rust	8/9	0.89
White/purple	3/3	1.00	Blue/blue	24/30	0.80
Brown/brown	6/7	0.86	Blue/light blue	4/9	0.44
Brown/tan	4/4	1.00	Light blue/white	3/3	1.00
Tan/brown	3/4	0.75	Light blue/light blue	4/4	1.00
Tan/tan	5/6	0.83	Green/green	12/13	0.92
Black/black	5/7	0.71			

Trainer

One trainer (experimenter) conducted all baseline, training, generalization, and test sessions. She had received undergraduate training in behavior analysis and had been employed at the facility as a bachelor's level psychologist for approximately six months before the study. She had participated in previous training programs to teach self-help and social skills and had participated as a behavioral observer in a previous research project.

Observation Procedures and Reliabilities

During the initial sessions, an observer scored all resident responses. However, after several sessions, this was discontinued, as it was determined that the trainer could record all responses without interfering with her other training responsibilities. For reliability purposes, observers made periodic recordings with the trainer as to which color the client chose for each trial at least once during each experimental condition for each client.

The observers were seated behind a one-way window viewing the training room. The trainer was not always aware of the presence of reliability observers. Each observer scored which one

of four alternative colors was selected by the participants in response to each stimulus item presented. Persons participating as reliability observers included a graduate student serving an internship at the institution, a secretarial employee who volunteered to participate in the program, and the second experimenter. A brief explanation of the data sheet was made before actual observations. The reliability observers (excluding the experimenters) were not told which color combinations were considered popular.

Reliability for resident choices was evaluated by comparing the color recorded by the reliability observer for each response with that recorded by the trainer. There were no disagreements between the trainer and reliability observer during the entire study.

Experimental Procedures

Baseline. Participants were individually escorted to the training room and immediately seated at the table. The puzzle apparatus, with torso and hips/legs removed, was placed on the table directly in front of the resident. During the first baseline session, the trainer explained that she would be using the puzzle apparatus to teach the client how to match clothing colors accord-

ing to popular fashion. After the first session for each woman, no explanation was presented and trials began after an initial greeting from the trainer.

The first three baseline sessions consisted of 27 trials each (one for each popular color combination previously defined), and were conducted on weekdays with not more than one daily session for each participant. In each case, the trainer inserted a pair of slacks into the puzzle and simultaneously presented four colored shirts as alternative responses. Participants were then asked to insert the colored shirt that would look best with the slacks. The four alternatives included one shirt that would make a popular color combination with the stimulus slacks and three shirts that would present an unpopular combination. To aid the trainer in determining which blouse to present with each pair of pants, five separate lists of 27 trials were established before baseline. On each list, the color of the stimulus pants was listed in random order for each trial, followed by the four shirt colors to present to the participant as possible responses. The three shirts that would form an unpopular combination with the designated pair of pants were randomly selected from all colors that would form an unpopular match. The particular list used during a given session was determined randomly. Hence, the three colors of the shirts that would result in an unpopular color combination varied randomly across sessions.

To eliminate from the baseline trials those popular color combinations that the residents already chose on a regular basis, the results from the first three baseline sessions were reviewed. Any of the 27 popular color combinations that a resident correctly selected on each of the first three baseline sessions were excluded from subsequent baseline sessions and eventual training. The number of resulting combinations for the remainder of baseline and eventual training for the five participants totalled 18 for Cathy, 15 for Ruth, 15 for Kathryn, 16 for Michelle, and 16 for Ellen.

Within each baseline session, participants were

praised intermittently for following trainer instructions and attending to the task (e.g., "Ruth, you are working very hard! Nice job"; "Keep up the hard work, Ellen."). However, no specific feedback or praise was presented regarding accuracy of performance, and participants' questions were answered with general statements such as, "If you continue to try this hard, you will do fine."

After each trial, the trainer (or observer) immediately recorded the resident's response as to which colored shirt was selected. After each session, the trainer reviewed the selection recordings and scored whether each combination was popular or unpopular by comparing the color of the shirt selected and the color of the pants in the puzzle with a master list of popular and unpopular color combinations (based on criteria described previously). For reliability evaluation regarding the trainer's accuracy of scoring the responses by comparing them to the master list, a second observer independently scored the responses from eight sessions, including at least one session for each client. There were no disagreements between the observer and the trainer regarding the accuracy of comparing the client choices to the master list.

Color-coordination training. Color-coordination training, consisting of modelling, instructions, practice, feedback, and praise was introduced during the second phase of the study. Residents met individually with the trainer as during baseline. The trainer inserted a pair of slacks and a shirt of one popular color combination in the puzzle while simultaneously labelling the two colors and informing the resident that it was a color outfit worn by many people. The clothes were then removed and the resident was encouraged to practise what the trainer had modelled. Following replacement of the clothes, the trainer praised the resident for complying (e.g., "Nice job, that's just what I wanted you to do."). The shirt was again removed and was presented simultaneously with three unpopular alternatives for the participant's consideration. The resident was then asked to select the shirt

that would look best with the slacks. Correct responses were enthusiastically praised (e.g., "Nice work, that's a popular combination!"). Incorrect responses were followed by verbal feedback as to the inaccuracy of the selection (e.g., "Those colors are not worn together by many people.") and then retrained according to the entire procedure just described. Following three consecutive correct responses (with three unpopular alternatives available) for one popular combination, another color selection was taught using the same procedure. The number of popular combinations taught each session varied according to the client's rate of progress. At least one new combination was introduced each session, and never more than five. Training sessions lasted 20 to 30 min each and were conducted on weekdays with not more than one training session per day for each client. An average of 14.20 training sessions (range 10 to 20) was conducted per resident.

Beginning with the second training session for each participant, review trials were conducted at the start of the session for color combinations previously trained. Each of the combinations taught the previous session was reviewed, as well as two or three combinations taught in previous sessions. Incorrect responses were corrected by the trainer (e.g., "That is not the popular combination we worked on yesterday, let's practice it again."), and those color combinations were retrained according to the original training procedure before new color combinations were introduced. Also, in an attempt to promote generalization from the puzzle to real clothing, reference was made during the review process to actual clothes. The trainer mentioned to the resident that she should think about her clothes while completing the puzzle. In addition, samples of donated clothing were presented to the resident to compare to the puzzle combination. For instance, if the puzzle combination consisted of brown slacks and a white blouse, the trainer asked the resident to pick up a brown pair of slacks and a white blouse. The resident was instructed to verbalize the colors of pants and

blouse while holding them together. The trainer then provided feedback saying that those clothes were like the color pieces in the puzzle and, hence, would make a popular clothing selection. References to actual clothing were provided once per review session for each color combination reviewed during that session.

Test sessions. To evaluate the effects of training and generalization from reinforced training trials to nonreinforced trials, test sessions were conducted over all combinations immediately following each training session. Procedures for test sessions were identical to those of baseline sessions, with no trainer feedback or praise presented regarding the accuracy of resident performance. A criterion for terminating training was established before the study. A minimum mean of 80% correct trials averaged across five consecutive test sessions was the criterion chosen.

Generalization. To determine if changes in color-coordinating skills demonstrated with the puzzle were accompanied by changes using actual clothing, generalization probes were conducted several times for each participant during baseline and color-coordination training conditions. Clothing loaned by staff members was obtained for each solid color used during the study as well as a large number of nonsolids. All pieces of clothing used were independently agreed upon by two observers as representative of a specific color category before use for generalization measures. For each probe, 12 solid shirts and eight solid slacks were randomly selected from the total clothing pool and were interspersed on a clothing rack with eight nonsolid shirts and one nonsolid pair of slacks also randomly selected. The inclusion of solid colors with nonsolids was conducted to approximate normal clothing groupings. For recording purposes, each garment was identified by a small numbered tag on the inside of the material.

Each participant was individually escorted to the training room and asked by the trainer to select outfits (pants and tops) from the clothing rack in colors that would make a popular combination. As each outfit was chosen, it was hung

on a hook in the wall, and after a maximum of four selections was made, all garments were returned to the clothing rack. For the first probe during baseline, all residents chose 15 outfits (three groups of four selections and one group of three). For all other probes, clients chose 10 outfits (two groups of four selections and one of two selections). Within each group of selections, residents could not pick the same pair of clothes. However, across groups of selections, residents were allowed to select the same pair as in the previous four selections. No feedback was provided about the quality of participant selections, but praise was intermittently provided for attending to the task of selecting clothes.

Scoring of resident performance was based on definitions derived from the normative data. If both items were of solid colors, the responses were judged popular or unpopular according to previously specified criteria. If the pair of solid colored garments selected did not meet the popular or unpopular criteria, a neutral pairing was scored. If one garment was a nonsolid, then the nonsolid category was marked. Resident performance on the generalization task was evaluated by computing the percentage of popular combinations out of all popular and unpopular selections.

Reliability measures for generalization responses were conducted during each experimental condition for each participant. The reliability observer was stationed in the observation room until four outfits (or the final group of outfits) were selected. At that point, the observer walked into the training room, scored the resident's selections by the identification/number tag and returned to the observation room. On every occasion, the recordings by the reliability observer as to which items were selected agreed with the recordings of the trainer.

Followup. Followup measures were conducted with each participant, ranging from seven to 14 weeks following termination of color-coordination training. Procedures used during followup sessions were identical to those for generalization probes.

Experimental Design

A multiple baseline across participants was employed. Following an initial baseline, color-coordination training was sequentially implemented for each participant, as reflected in Figure 1.

Time Efficiency Measures

One rationale for relying heavily on the puzzle apparatus to teach popular color coordination, as opposed to using actual clothing exclusively, was that the former approach would be more time efficient. To evaluate this hypothesis, several time measures were included. The trainer (or observer) recorded the amount of time necessary to complete each trial following the initial instructions for a baseline session (puzzle) and a generalization session (clothes) for each client. In addition, the trainer timed trials for each participant in one session during which they were provided the same task as with the puzzle only using actual clothes. That is, the trainer laid a pair of solid-color slacks on the table and presented four alternative blouses for the resident to select a color combination. The trainer began timing with a stopwatch at the beginning of the first trial and terminated at the end of the last trial.

RESULTS

The effects of color-coordination training on percentage of popular color combinations with the puzzle apparatus are displayed in Figure 1. During baseline, each resident demonstrated relatively low percentages of popular combinations, averaging (mean) 28% for Cathy, 27% for Ruth, 19% for Kathryn, 51% for Michelle, and 23% for Ellen. Immediate increases occurred as color-coordination training was implemented. All residents increased their mean level of popular selections during test sessions by at least 37 percentage points while color coordination training was in effect. Mean percentages during the training condition were 70, 81, 70, 88, and 80%

Fig. 1. Per cent of popular color combinations selected by each participant during baseline, test, and generalization sessions. Test sessions followed color-coordination training sessions and were identical to baseline sessions. The followup sessions occurred at the specified intervals after color-coordination training ended.

for the five clients, respectively. All residents achieved the previously defined criterion for termination of training of at least a mean 80% popular combinations averaged over five consecutive test sessions.

Results of the generalization measures indicated that the increases in percentage of popular selections with the puzzle apparatus were paralleled by increases in popular selections with actual clothing (Figure 1). All five clients demonstrated a mean increase of at least 32 percentage points over baseline averages with solid colored garments while color coordination training was in effect. Cathy increased from a mean of 0% during baseline to a mean of 47% during training, Ruth from 22 to 87%, Kathryn from 22 to 65%, and Michelle from 29 to 88%. The increase for Ellen was less observable, changing from a mean 37% during baseline to 69% during training, with more overlap in the individual session results between baseline and training than the other four clients.

Followup measures conducted with actual clothing indicated that the increase during color-coordination training generally maintained after training terminated (Figure 1). The followup results for each client were at least 23 percentage points higher in popular selections than the baseline mean for generalization sessions (average of 57% higher for all clients). For three residents, the followup results reflected a greater percentage of popular selections than the mean for generalization sessions during color-coordination training.

Time Efficiency Measures

The mean amount of time to conduct a trial using the puzzle apparatus with each client was less than the time required with trials involving actual clothing. Trials using only the puzzle averaged 31 sec each for completion; trials involving the selection of one shirt from a choice of four averaged 67 sec for the five clients. Trials that included selecting one pair of slacks and one shirt from the clothing rack averaged 52 sec for completion.

DISCUSSION

The results indicated that profoundly and severely retarded persons could be taught a community preparation skill to coincide with popular fashion. All residents increased their selection of popular, color-coordinated clothing with the puzzle when the training procedure was implemented. Notable increases in popular selections were also apparent with four participants with actual clothing during nonreinforced trials, indicating that generalization did occur. The followup results, ranging from seven to 14 weeks after training terminated, demonstrated that the increases in color-coordination skills with actual clothing were not of brief duration. A less apparent increase in the generalization measure occurred with the fifth participant. The smaller changes may have been due to an initially higher set of clothing selection skills. During baseline, her popular selections with actual clothing were more variable than the other participants'.

The reliance on normative data through observations of community residents represents an addition to the previously reported literature on preparation of retarded persons for mainstreaming. Other community involvement skills may be taught using similar approaches based on local norms, such as social manners, colloquial speech, leisure pursuits (*e.g.*, dance styles or popular community sports), and a variety of other appearance variables (hair style, make-up fashions, other clothing variables, *etc.*). However, when evaluating local customs, a number of considerations must be dealt with. For example, a popular "criterion" must be established. In this investigation, an occurrence criterion for each popular color combination of at least 1% of the total solid/solid sample was used. This was established after determining that an extremely wide diversity exists in color combinations. While 1% of the total solid/solid combinations appears to be an infrequent definition of popular, increasing the occurrence criterion would have greatly reduced the possible solid color combinations meeting the criterion. Future research seems

warranted to look at other ways of determining popular criteria, provided the methods are feasible for institutions to implement on local levels with available resources. For instance, stressing the nonselection of unpopular combinations may be more productive. Other considerations are the permanency of the norms over time (*e.g.*, seasonal changes), establishing general rules of popular behavior as opposed to teaching specific cases (*e.g.*, teaching the "unpopularity" of combining upper and lower garments of different color designs such as stripes and checks), and teaching retarded persons methods to determine for themselves what is popular. However, the latter strategy may be most effectively employed with persons with less serious disabilities than the severely and profoundly retarded persons who participated in this study.

Of particular concern when establishing popular criteria is the reliability of observations of the community samples. In the present project, reliabilities for specific colors, for occurrences of popular solid/solid color combinations, and for the general category of solid/solid occurrences regardless of color, averaged above 0.85. While some color combinations fell below customary standards, this did not present a major problem, as the main disagreements were with highly similar colors, such as blue and light blue, which would generally not make a grossly unpopular match if intermixed with each other when paired with other colored garments. Future research should consider combining these categories into one category.

Descriptions of the benefits of simulation procedures with community preparation programs have been previously reported (*e.g.*, Neef *et al.*, 1978; Page *et al.*, 1976). The present investigation relied heavily, although not exclusively, on a simulation procedure, as actual clothes were used in review sessions in an attempt to "train to generalize" (Stokes and Baer, 1977). The time efficiency data support the Neef *et al.*, 1978, results with less seriously retarded persons, in that the simulation procedure required less time than procedures using actual clothes. Additional con-

siderations that support the simulation approach here are the additional expenses involved in using a wide variety of actual clothes (*e.g.*, purchasing new clothes) and the storage and retrieval logistics (*e.g.*, amount of space required, prevention of theft, *etc.*).

The initial plans of the investigation included the use of the clients' personal clothing in part of the training procedures. These plans were abandoned due to the lack of a wide diversity of garment colors among each client's wardrobe and the interference this process would have caused with the daily living area routine, such as the laundry procedure, the storage of clothes in each participant's closet or chest drawers by direct care staff, *etc.* However, to help the clients incorporate the skills acquired during the investigation into their daily routine, new ward procedures were implemented after color-coordination training and before followup. A staff member (experimenter) met with the clients in the early evening hours on weekdays, during which time they selected their outfit for the next day. If the outfit included two solid-colored garments of a popular match, the staff member praised the selection. Unpopular or neutral selections were not praised regarding the specific colors, although praise was provided for attending to the task of selecting clothes. The experimenter eventually terminated her participation, but the living-area staff continued to structure the evening so that the women selected which clothes they would wear the next day. These events probably contributed to the successful followup results noted. Also, the clients were later provided with funds from the facility to purchase new clothing outfits to help expand their wardrobe. The experimenter accompanied them on the shopping trips and observed that no unpopular color combinations were selected, although some persons selected nonsolid garments in addition to the popular solid combinations.

As noted previously, most community preparation programs have involved persons with retardation levels above the severe range. The results of this investigation suggest that similar

programs can be successful with more seriously retarded individuals. Considering the relatively modest but ongoing deinstitutionalization of profoundly and severely retarded persons into structured community living arrangements, future research similar to the present study is warranted to help prepare these persons for their new life styles.

REFERENCE NOTE

1. Gay Apparels Co., Inc. *Special catalog for home care clients*. Cotter, Arkansas: Author, undated.

REFERENCES

- Azrin, N. H. and Armstrong, P. M. The "mini-meal"—A method for teaching eating skills to the profoundly retarded. *Mental Retardation*, 1973, 11, 9-13.
- Bauman, K. E. and Iwata, B. A. Maintenance of independent housekeeping skills using scheduling plus self-recording procedures. *Behavior Therapy*, 1977, 8, 554-560.
- Behavioral characteristics progression*. Palo Alto, California: VORT Corporation, 1973.
- Bellamy, G. T. and Clark, G. Picture recipe cards as an approach to teaching severely and profoundly retarded adults to cook. *Education and Training of the Mentally Retarded*, 1977, 12, 69-73.
- Berkson, G. and Landesman-Dwyer, S. Behavioral research on severe and profound mental retardation (1955-1974). *American Journal of Mental Deficiency*, 1977, 81, 428-454.
- Close, D. W. Community living for severely and profoundly retarded adults: A group home study. *Education and Training of the Mentally Retarded*, 1977, 12, 256-262.
- Gollay, E. Deinstitutionalized mentally retarded people: A close look. *Education and Training of the Mentally Retarded*, 1977, 12, 137-144.
- Heiny, R. W. and Stachowiak, R. J. REAL matching of socialization activities with individual needs. *Mental Retardation*, 1976, 14, 12-15.
- Johnson, M. S. and Bailey, J. S. The modification of leisure behavior in a half-way house for retarded women. *Journal of Applied Behavior Analysis*, 1977, 10, 273-282.
- Joint Commission on Accreditation of Hospitals. *Standards for residential facilities for the mentally retarded*. Chicago: Author, 1971.
- Kazdin, A. E. Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification*, 1977, 1, 427-452.
- Kennett, K. F. The family behavior profile: An initial report. *Mental Retardation*, 1977, 15, 36-40.
- Leff, R. B. Teaching TMR children and adults to dial the telephone. *Mental Retardation*, 1975, 13, 9-12.
- Lowe, M. L. and Cuvo, A. J. Teaching coin summation to the mentally retarded. *Journal of Applied Behavior Analysis*, 1976, 9, 483-489.
- Neef, N. A., Iwata, B. A., and Page, T. J. Public transportation training: In vivo versus classroom instruction. *Journal of Applied Behavior Analysis*, 1978, 11, 331-344.
- Newton, A. Clothing: A positive part of the rehabilitation process. *Journal of Rehabilitation*, 1976, 42, 18-22.
- O'Brien, F. and Azrin, N. H. Developing proper mealtime behaviors of the institutionalized retarded. *Journal of Applied Behavior Analysis*, 1972, 5, 389-399.
- O'Connor, G. Home is a good place: A national perspective of community residential facilities for developmentally disabled persons. *Monograph of the American Association of Mental Deficiency*, 1976, No. 2.
- Page, T. J., Iwata, B. A., and Neef, N. A. Teaching pedestrian skills to retarded persons: generalization from the classroom to the natural environment. *Journal of Applied Behavior Analysis*, 1976, 9, 433-444.
- President's Committee on Mental Retardation. *MR 72 Islands of Excellence* (Dew Publication No. 15 73-7). Washington, D.C., 1973.
- Santistevan, H. *Out of their beds and into the streets*. Washington, D.C.: American Federation of State, County, and Municipal Employees, 1976.
- Stokes, T. F. and Baer, D. M. An implicit technology of generalization. *Journal of Applied Behavior Analysis*, 1977, 10, 349-367.
- Thornley, M. L. A national trend: The impact of educating and training formerly institutionalized children and adults in the community. *Education and Training of the Mentally Retarded*, 1976, 11, 152-155.
- Tobias, J. and Cortazzo, A. D. Training severely retarded adults for greater independence in community living. *The Training School Bulletin*, 1963, 60, 23-37.
- White, R. Education of the retarded: A point of view. *Education and Training of the Mentally Retarded*, 1976, 11, 295.

Received 12 December 1977.

(Final Acceptance 5 May 1978.)