IN VIVO VERSUS SIMULATION TRAINING: AN INTERACTIONAL ANALYSIS OF RANGE AND TYPE OF TRAINING EXEMPLARS

NANCY A. NEEF, JEFFREY LENSBOWER, IDIL HOCKERSMITH, VALERIE DePALMA, AND KEVIN GRAY

DEVEREUX INSTITUTE OF CLINICAL TRAINING AND RESEARCH AND
WASHINGTON COUNTY ASSOCIATION FOR RETARDED CITIZENS

We analyzed the role of the range of variation in training exemplars as a contextual variable influencing the effects of in vivo versus simulation training in producing generalized responding. Four mentally retarded adults received single case instruction, followed by general case instruction, on washing machine and dryer use; one task was taught using actual appliances (in vivo) and the other using simulation. In vivo and simulation training were counterbalanced across the two tasks for the 2 subject pairs, using a within-subjects Latin square design. With both paradigms, more errors were made after single case than after general case instruction during probe sessions with untrained washing machines and dryers. These results suggest that generalization errors were affected by the range of training exemplars and not by the use of simulated versus natural training stimuli. Although both general case simulation and general case in vivo training facilitated generalized performance of laundry skills, an analysis of training time and costs indicated that the former approach was more efficient. The study illustrates a methodology for studying complex interactions and guiding decisions on the optimal use of instructional alternatives.

DESCRIPTORS: simulation, general case training, methodology, generalization, laundry skills

With the increased focus on community skills for developmentally disabled persons, considerable controversy has emerged over the use of simulation versus in vivo training (Nietupski, Hamre-Nietupski, Clancy, & Veerhusen, 1986). Simulation training procedures, in which controlling stimuli or their approximations are presented outside their normal stimulus context, have been used to teach developmentally disabled individuals a variety of independent living skills, including pedestrian (Page, Iwata, & Neef, 1976), public transportation (Neef, Iwata, & Page, 1978), clothing selection (Nutter & Reid, 1978), restaurant (van den Pol et al., 1981), telephone (Horner, Williams, & Stevely, 1987), leisure (Giangreco, 1983), and purchasing skills (Haring, Kennedy, Adams, & Pitts-Conway, 1987; Nietupski, Welch, & Wacker, 1983). In other investigations, however, skills acquired through simulation training have failed to generalize to natural community environments (Coon,

Vogelsberg, & Williams, 1981; Freagon & Rotatori, 1982; Marchetti, McCartney, Drain, Hooper, & Dix, 1983; McDonnell, Horner, & Williams, 1984; Morrow & Bates, 1987). Such findings have led some authors to argue for in vivo instead of classroom-based instruction (Belmore & Brown, 1978; Brown et al., 1983; Ford & Mirenda, 1984; Hill, Wehman, & Horst, 1982; McDonnell et al., 1984; Vogelsberg, Williams, & Bellamy, 1982). To determine the procedural alternative best suited for a particular set of circumstances, information is needed on the conditions required for optimal application of those procedures (Johnston, 1988; McDonnell et al., 1984; Van Houten, 1987). Because training conducted exclusively in the community can be expensive, inconvenient, time consuming, and, in some cases, hazardous, it is important to identify variables that enhance the effectiveness of simulation training in producing generalized responding.

The effectiveness of in vivo training in producing generalized responding in handicapped learners has been shown to be affected by the extent to which the range of stimulus and response variation in the natural environment or instructional universe is represented in the training examples. General case

The authors gratefully acknowledge the assistance of Gregg Macmann, Melissa Hensel, Suzanne Meyers, Lee Michael, Robert DeHaven, and Guy F. Collins. Requests for reprints should be sent to Nancy A. Neef, The Devereux Foundation, Institute of Clinical Training and Research, 19 South Waterloo Road, Box 400, Devon, Pennsylvania 19333-0400.

instruction, involving the use of multiple training examples that sample the range of stimulus and response variations, has resulted in generalized responding across untrained exemplars in performance of a vocational skill (Horner & McDonald, 1982) and use of vending machines (Sprague & Horner, 1984) by moderately and severely mentally retarded high school students. When training exemplars were not selected based on the relevant characteristics that differentiated members within the defined stimulus class, generalization errors were found to be functionally related to the restricted range of training stimuli.

In one of the few studies to examine general case training in a simulated setting, Horner *et al.* (1987) found that the procedure was effective in teaching generalized telephone answering skills to 4 moderately to severely retarded high school students; however, because simulation training did not incorporate dialing on a vertical plane, placing calls did not generalize to a pay telephone until *in vivo* training was also conducted. Thus, as with *in vivo* training, the effectiveness of simulation training in producing generalized responding might be affected by the extent to which variations among members of the stimulus class are represented in the training exemplars.

The issues raised by previous research present design challenges. From a traditional perspective, the questions of substantive interest might be approached through a multivariate analysis of variance (MANOVA), or complex variation thereof (e.g., Cook & Campbell, 1979; Huitema, 1980), with two main effects (*in vivo* vs. simulation training; single vs. general case exemplars) across multiple dependent measures (e.g., acquisition, generalization, error patterns during acquisition and generalization, and training costs). Within this traditional design framework, correlational methods provide a flexible but nonanalytical (Baer, 1977; Michael, 1974) methodology for the study of complex interactions (e.g., Cohen & Cohen, 1983). There are also practical limitations to large-group designs (e.g., funding to support large-scale investigations across the range of relevant variables) that

necessitate the use of alternative single case methodologies for the analysis of complex interaction effects (Baer, Wolf, & Risley, 1987).

We conducted a within-subjects analysis of the interaction between the range of stimulus variation in training exemplars and type of instructional materials. Specifically, we compared the effects of simulation versus *in vivo* training of laundry skills, using single and general case exemplars, on measures of generalization and types of errors with 4 mentally retarded adults. We also conducted an analysis of cost and training time in order to evaluate the practicality of simulation training.

METHOD

Participants and Settings

The 4 participants had resided in their respective community residential facilities for 4 years and had lived either with their families or in similar community-based residences prior to that time. All were reported to have received previous instruction in laundry skills as part of their habilitation plan but had made minimal progress.

Client 1 was a 41-year-old male who was diagnosed as moderately mentally retarded on the basis of adaptive functioning, with cognitive skills in the mildly retarded range ($IQ = 68$ as measured on the WAIS). He was also diagnosed as behavior disordered, for which medication (Mellaril and Mebaral) had been prescribed. Client 2 was a 64-year-old male who was diagnosed as severely mentally retarded ($IQ = 32$). Clients 3 and 4 were females, aged 31 and 55 years, who were also diagnosed as severely mentally retarded with measured IQs of 20 and 26, respectively. Clients 3 and 4 were reported to engage frequently in perseverative behavior.

In vivo training and probe sessions were conducted in the laundry room of the clients' residences, community group homes, day habilitation center, and in a local laundromat. Simulation training was conducted adjacent to the laundry room in the day habilitation center and at the clients' residences.

Tasks

The target tasks were use of washing machines and dryers. The specific requirements for operating a washing machine were that, when presented with (a) a washing machine, (b) a basket of presorted laundry, (c) quarters, nickels, and dimes (for coin-operated machines), (d) laundry detergent and pre-marked measuring cup, and (e) a request to wash the clothes with the specified temperature, cycle, and load setting, the client was to place the clothes in an empty washing machine, add the detergent, select the proper settings, activate the machine, and place all clothes in the basket upon completion of the cycle. The specific requirements for operating a dryer were that, when presented with (a) a dryer, (b) a basket of recently washed (damp) clothes, (c) quarters, dimes, and nickels (for coin-operated machines), and (d) a request to dry the clothes at a specified temperature, cycle, and/or time setting, the client was to place the clothes in an empty dryer, clean the lint trap if necessary, select the proper settings, activate the machine, and remove the clothes when dry. These behaviors were further task analyzed, yielding 11 and 9 steps for washing machine and dryer use, respectively.

Stimuli

Stimuli (see Table 1) were selected based on an analysis of the types of washing machines and dryers that clients were most likely to encounter in the community (i.e., at the community residential facilities, the day habilitation center, and the local laundromat). Of the seven different washing machines and seven dryers surveyed, four or five of each were selected for training according to a general case analysis of the minimum number of machines representing the maximum stimulus variation among the target set (Horner, Sprague, & Wilcox, 1982). The remaining washers and dryers served as generalization probe stimuli. The machines used during general case training also served as generalization probe stimuli after single case training.

A simulated washing machine and dryer for each actual counterpart were designed from photographs

of the above stimuli. To make the simulated machines as similar as possible to their actual counterparts, small parts (e.g., knobs and switches) from discarded machines were used when available, and attempts were made to approximate the kinesthetic and visual properties of the *in vivo* stimuli (e.g., some temperature buttons were made from indented styrofoam that depressed slightly when pushed and were covered with aluminum foil to simulate chrome). To make the simulated machines as economical as possible, inexpensive materials were used in their construction (e.g., cardboard, styrofoam, and used parts). The simulated machines also were designed to be convenient to transport by using lightweight (e.g., cardboard, pressboard paneling, and styrofoam) and collapsible materials and by using a minimum number of components (e.g., the same box served as the base for each of the machines). Finally, the simulated machines were designed to be efficient by using the minimum number of components possible to represent maximum variations among machines and by using interchangeable components (e.g., top panels that could be slid in and out of grooves on the body of the machine).

A cardboard furniture box (100 cm by 65 cm) with openings (33 cm by 52 cm) and sealable lids for front and top loading machines initially served as the base. This was subsequently replaced with more durable pressboard paneling; four panels slid into slots formed by two base pieces. A frame that rested on the top of the side panel struts within the box permitted insertion of an agitator from an inoperative washing machine. Slits with pockets were inserted in the box to contain removable lint screens. Four washer and three dryer removable panels representing combinations of stimulus features (e.g., temperature setting, load size) and their locations on actual machines were constructed of cardboard covered with shiny paper and inexpensive hardware. Each panel was 19 cm by 61 cm and could be slid along the grooves of a top and bottom styrofoam base with a cardboard back. A patch of transparent red acetate was used to cover a styrofoam rectangular button on one of the panels,

Table 1
Stimulus and Response Variations Among Single Case Training, General Case Training, and Generalization Probe Dryers

	Dryers						
	1	2	3	4	5	6	7
Steps	GC Training, Ss 1-4	GC Training, Ss 1-4	Gen Probe, Ss 1-4	SC Training, Ss 1 and 3; Gen Probe, Ss 2 and 4	SC Training, S2; GC Training, Ss 1, 3, and 4	SC Training, S4; Gen Probe, Ss 1-3	GC Training, Ss 1-4
Open lid Stimulus	Left front handle	Right front han- dle	Top center inden- tation	Top center inden- tation	Top center inden- tation	Left front handle	Left bottom front pedal and left front handle Press, Pull
Response	Pull	Pull	Lift	Lift	Lift	Pull	
Clean lint trap Stimulus	None	Curved screen on door rim	Oblong screen back right Lift, Pull	Curved screen on door rim Pull	Oblong screen back right Lift, Pull	Oblong screen back right Lift, Pull	Oblong screen back right Lift, Pull
Response	None	Pull					
Set temperature Stimulus	Left panel buttons	Right panel switch	Center panel dial	Center panel dial	Center panel dial	Center panel dial	Center panel dial
Response	High High-cool down Low	Cotton/high Perm press Delicate/low	Fluff air Knits Gentle/low Perm press Heavy/high Turn	Fluff air Knits Gentle/low Perm press Heavy/high Turn	Fluff air Knits Gentle/low Perm press Heavy/high Turn	Fluff air Knits Gentle/low Perm press Heavy/high Turn	Fluff air Knits Gentle/low Perm press Heavy/high Turn
Set cycle Stimulus	Left panel buttons	Right panel switch	Left panel dial	Left panel dial	Left panel dial	Left panel dial	Left panel dial
Response	Cotton Perm press Knit Push	Perm press Normal Flip	Auto perm press Timed Dry Auto Dry Turn	Auto perm press Timed Dry Auto Dry Turn	Auto perm press Timed Dry Auto Dry Turn	Auto perm press Timed Dry Auto Dry Turn	Auto perm press Timed Dry Auto Dry Turn

Table 1
(Continued)

Steps	Dryers						
	1	2	3	4	5	6	7
	GC Training, Ss 1-4	GC Training, Ss 1-4	Gen Probe, Ss 1-4	SC Training, Ss 1 and 3; Gen Probe, Ss 2 and 4	SC Training, S2; GC Training, Ss 1, 3, and 4	SC Training, S4; Gen Probe, Ss 1-3	GC Training, Ss 1-4
Set time							
Stimulus	None	Right bottom dial Off-45	Left panel dial Very dry-cool down 60-off Very dry-less dry Off Turn	Left panel dial Very dry-cool down 60-off Very dry-less dry Off Turn	Left panel dial Very dry-cool down 60-off Very dry-less dry Off Turn	Left panel dial Very dry-cool down 60-off Very dry-less dry Off Turn	Left panel dial Very dry-cool down 60-off Very dry-less dry Off Turn
Response	None	Turn					
Activate							
Stimulus	Coin slot	Time dial	Right panel but- ton Push	Right panel but- ton Push	Cycle dial	Time dial	Time dial
Response	Insert coins	Turn	Push	Push	Push	Push	Push
Remove clothes							
Stimulus	Right panel light out	Machine stops	Buzzer	Buzzer	Buzzer	Buzzer	Buzzer

indicating that the machine had been activated. Coin-operated activators were made from shallow boxes (25 cm by 10 cm) containing a thin piece of styrofoam with circular indentations for coins. A cardboard tab with the operation cost was attached to the front of the slide box. A tape recording of a washing machine and dryer in use was made to signal activation and completion of the washer/dryer cycles.

Training Procedures

Training sessions were conducted three to four times per week for each client. Each session consisted of washing or drying one basket of clothes, which varied along dimensions such as load size and type (e.g., white vs. dark). The clients were trained individually using procedures consistent with the guidelines outlined by Bellamy, Horner, and Inman (1979). Each session began with a request to the client to wash or dry the clothes using a particular dial setting. Subsequent steps performed correctly were followed by praise or a pat on the back. Steps performed incorrectly were followed by the trainer saying "stop," with both errors and lack of responding (within 5 s) followed by the trainer modeling and describing the correct response. After correcting the response, the trainer restored the conditions to those in effect during the previous step (if possible) and instructed the client, "Now you show me." If the response was again performed incorrectly, the trainer interrupted it, restored the conditions preceding the incorrect response, and verbally instructed the client in the correct response while presenting a gestural and, if necessary, physical prompt. Each completion of the task comprised a session, after which the client was presented with coffee or a soft drink. Training continued in this way until the client performed the task with no more than one error on two consecutive sessions.

Probes

Probes on all washing machines and dryers were conducted before training and after criterion had been met. Each probe consisted of attempted completion of the task on one machine. Probes were conducted across experimental phases to assess gen-

eralization of performance under nonreinforcement conditions and to untrained machines. During each probe session, the client was given the requisite materials and asked to perform the task. If the client made an error that precluded performance of the next response in the chain, his or her attention was diverted while the trainer completed the step. No feedback was provided for performance, but the client was given coffee or a soft drink at the completion of the probe.

Experimental Design and Conditions

In vivo and simulation training were counter-balanced across two tasks (washing machine and dryer) for the 2 client pairs. Following baseline probes, Clients 1 and 2 received in vivo training on one washing machine and simulation training on one dryer (single case training). The machine targeted for this training differed for each client. Clients 1 and 2 then received general case in vivo training for washing machines and general case simulation training for dryers. The machines targeted for this condition, when combined with the one used for single case training, sampled the range of stimulus and response variation defined for each step of the task analysis. The machines in the general case condition were trained on a rotating basis across sessions until criterion was met on each one. Thus, the general case consisted of the combined stimuli from both training conditions. The two conditions were differentiated by assessing generalization after criterion was reached on the first and last machines trained. The tasks targeted for in vivo versus simulation training were reversed for the second pair of clients; Clients 3 and 4 received in vivo single case training followed by in vivo general case training on dryers concurrently with simulation single case training followed by simulation general case training on the washing machines.

This single case variation of a Latin square design was selected to permit comparison of in vivo versus simulation training both within and between subjects across tasks, as well as comparison of single case versus general case instruction within each of the two training paradigms. Although the latter comparison did not control for the number of ma-

chines trained per se, the results of the study by Sprague and Horner (1984) indicated that this was not a functional variable. Given that training on one machine (single case) could not follow training on more than one machine (general case), order effects were inherent.

Data Collection and Reliability

Integrity of training procedures. In 19% of the training sessions across all clients, conditions, and tasks, one of the experimenters recorded whether the trainer implemented the procedure as specified on each task step. A step was scored as correct if the trainer provided social reinforcement after a correct response by the trainee and the appropriate prompt after an incorrect response. The mean percentage of correct trainer behaviors was 99% (range, 93% to 100%).

Training performance. During each training session, data were collected by the trainer on the client's performance of each step of the task analysis. A step was scored as correct if it was performed independently within 5 s. One of the experimenters simultaneously but independently recorded responses during 31% of the training sessions for all clients within each training condition for both tasks. Interobserver agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100. Mean agreement scores for each client on both tasks ranged from 96% to 100%.

Probe performance. Data were collected on the percentage of correct responses on trained and untrained machines during probes and on the types of errors that were made. For each step of the task analysis, the observer recorded the response as correct or, if incorrect, the type of error made. Incorrect responses were categorized as (a) no response, a failure to initiate a response within 5 s of the instruction or the preceding response; (b) direction error, a response other than that specified in the verbal direction but to the appropriate stimulus; (c) stimulus error, a response made to an incorrect stimulus; (d) topography error, using the wrong topography with the correct stimulus; (e) sequence error, a response performed in an inappropriate

order; and (f) incomplete response, such as failing to remove all clothes from the machine. The percentage of errors that occurred within each category was calculated in order to examine more precisely the effects of type (simulation vs. in vivo) and range (single vs. general case) of training on response patterns and generalization errors.

Independent observations were conducted by the trainer and one of the experimenters for 21% of the probe tasks across all clients, in each condition, and for both washing machines and dryers. Agreements were defined as both observers recording the same error on a step or recording the step as performed correctly. Reliability was calculated by dividing the number of agreements by agreements plus disagreements and multiplying by 100. Mean scores for both occurrence and nonoccurrence were 96% or higher, with ranges from 83% to 100%.

RESULTS

Training

Table 2 shows each client's percentage of correct probe responses on the training washing machines and dryers across experimental conditions. Although the posttraining improvements were substantial in most cases, 100% correct responding occurred only for Client 1 on the single case machines. Despite some within-subject variability, the mean increase in the percentage of correct responses was similar for washing machines and dryers and for single case and general case training (range, 36% to 41%).

Generalization

Figure 1 shows the mean and range percentage of correct responses on untrained washing machines and dryers across experimental conditions. Similar trends were evident across clients, tasks (washing machines and dryers), and type of stimulus (actual and simulated), irrespective of the training paradigm (simulation or in vivo). In all cases except one (dryers for Client 1), performance on untrained simulated and actual washing machines and dryers increased following single case training. The mean

Table 2
Percentage of Correct Responses on Pre- and Posttraining Probes of Training Washing Machines and Dryers

	Washing machine(s)				Dryer(s)			
	Single case		General case		Single case		General case	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Client 1	70	100	60	92	83	100	83	95
Client 2	50	82	51	94	50	90	43	90
Client 3	33	81	47	81	38	88	43	88
Client 4	22	69	42	87	38	67	24	85

percentage increase on actual machines was 31% for washing machines and 20% for dryers. However, the highest percentage of correct responses on untrained machines occurred following general case instruction, ranging from 81% (Client 2, dryers) to 92% (Client 1, dryers), with an average increase relative to single case training of 12% for washing machines and 18% for dryers.

Table 3 shows the distribution of incorrect re-

sponses on untrained exemplars across error categories for each experimental condition. Most of the errors during baseline were “no responses,” and direction errors accounted for the highest proportion of probe errors. Stimulus and response (topography) errors were rarely made, and occurred least often, after general case training. No consistent differences in errors occurred for simulated versus actual machines.

Figure 1. Percentage of correct responses (mean and range) on probes with untrained actual and simulated washing machines and dryers across experimental conditions for Clients 1 through 4.

Table 3
Percentage of Incorrect Responses Within Error Categories on Generalization Probes Across Experimental Conditions

	No response	Direction error	Stimulus error	Topography error	Sequence error	Incomplete response
Washing machines						
Baseline						
Actual	48	30	7	4	6	5
Simulated	30	32	25	5	6	2
Single case training						
Actual	10	54	26	4	2	4
Simulated	6	46	24	11	6	8
General case training						
Actual	11	80	0	0	3	6
Simulated	8	90	2	0	2	0
Dryers						
Baseline						
Actual	59	19	11	6	3	3
Simulated	55	14	16	5	4	7
Single case training						
Actual	19	56	9	5	5	7
Simulated	8	45	20	12	6	10
General case training						
Actual	22	42	7	0	15	15
Simulated	20	51	7	0	17	5

Cost Analysis

Clients 1 and 2 met training criterion after 3.4 and 13.6 hr of instruction on simulated dryers and after 20 hr and 36.1 hr of instruction on actual washing machines. Clients 3 and 4 met training criterion after 7.2 hr and 8.2 hr of instruction on simulated washing machines and after 43 hr and 36 hr of instruction on actual dryers. The mean cost per client was \$221 for simulation training and \$478 for in vivo training. Simulation training costs consisted of \$419 for trainer salary (\$12.91 per hour times total training time), \$445 for the training stimuli (film and development, construction materials, and labor), and \$21 for travel. In vivo training costs consisted of \$1,744 for trainer salary, \$35 for the training stimuli (for coin-operated machines), and \$134 for travel.

DISCUSSION

The purpose of this study, to evaluate the multiple effects of interventions, is consistent with the recent call for analyses of the contextual variables

influencing the behavioral effects of procedures (Hains & Baer, 1989; Johnston, 1988; Van Houten, 1987). We identified a set of independent variables that do and do not appear to influence the effectiveness of simulation training in producing generalization of a functional skill, and we used a methodology that involved multiple dependent measures to guide decisions on the use of alternative training paradigms. Thus, our analysis involved multiple independent and dependent variables and their interactions within and between clients.

We conducted this analysis by using a within-subjects variation of a counterbalanced Latin square design. Thus, the study illustrates an alternative to the use of the multielement design suggested by Hains and Baer (1989) for the study of complex interactions. As they pointed out, traditional uses of single-subject designs preclude the study of interactions between variables that is necessary to determine the conditions affecting the generality and dependability of technologies. Although our design did not permit the demonstration of experimental control in skill acquisition, our purpose was

to respond to the more pressing challenge of assessing the interaction of variables related to generalization. The results suggest that generalization errors were affected by the range of training exemplars and not by the use of simulated versus natural training stimuli; with both simulation and in vivo paradigms, more errors occurred on untrained machines after single case training than after general case training.

This finding is supported by the error analysis. After being trained on one washing machine or dryer, a number of stimulus and generalization errors occurred with the untrained machines (e.g., inappropriately applying the learned response of pulling a dial with a machine activated by pushing a lever), whereas very few topography or stimulus errors occurred following general case training. Instead, the majority of errors during this condition were the result of no response (suggesting a problem with fluency) or the wrong temperature, load, or cycle setting (reflecting difficulties in discriminating among printed words); these deficits might be addressed through a more stringent training criterion or the use of permanent-product cues such as picture prompts (e.g., Thompson, Braam, & Fuqua, 1982).

Although general case instruction was shown to enhance the effectiveness of simulation training in producing generalization to other exemplars, generalization to other environments may be problematic if distractors operating in community settings lessen attention to, and control by, relevant stimuli. In the study by Morrow and Bates (1987), for example, laundry skills acquired by severely handicapped students in the classroom did not generalize during probes in community laundromats, even to those washing machines that were represented fully by simulated (cardboard replica) or natural (actual machine) training stimuli. Similarly, McDonnell *et al.* (1984) reported that classroom simulation training of purchasing skills, using stimuli that maximally approximated criterion exemplars, did not result in generalization to community grocery stores and that errors resulted primarily from failure to orient to the cash register. Thus, instructional settings (classroom- vs. community-based) may rep-

resent "marker variables" (Hains & Baer, 1989) that have effects independent of the training materials (simulated vs. actual). This problem was not observed in the present study, possibly because training was conducted in the environments in which generalization to other machines was assessed.

In addition to behavioral effectiveness, practicality and administrative requirements are important considerations in the choice of procedures (Hopkins, 1987; Johnston, 1988). Although both general case simulation and general case in vivo training facilitated generalized performance of clothes-washing and clothes-drying skills, the former approach was more cost efficient in this situation. First, the simulation training materials (the model, detergent, and coins) were reusable and thus were nonrecurring expenses. Second, in vivo training required waiting for the completion of an entire wash or dry cycle, whereas it was possible to abbreviate this period under simulated conditions (Cuvo, Jacobi, & Sipko, 1981). Third, it was possible to conduct general case simulation training in the same location, whereas in vivo training required time-consuming travel to multiple training sites. The down time associated with in vivo training, however, might be used to target other skills (e.g., teaching bus riding during travel to the training sites) necessary for community living and employment.

The results of this investigation indicate that the use of general case programming with simulation training may increase the probability of generalization to exemplars in the natural environment and may offer an efficient and cost-effective alternative to the exclusive use of in vivo training. It is likely that the similarity of the stimulus features on the simulated model to their actual counterparts, as well as the subjects' previous or simultaneous exposure to actual washing machines and dryers in natural environments, contributed to the effectiveness of simulation training in producing generalized responding. Thus, simulation training might be most appropriate as an adjunct to in vivo training. Indeed, the combination of simulation plus in vivo training has been reported to be more effective than either approach alone (McDonnell & Horner, 1985;

McDonnell et al., 1984). For example, simulation training might be used initially to minimize the required amount, time, and cost of subsequent instruction in the community. Thereafter, simulation training could be used to provide massed-trial remedial training of specific behaviors that were performed incorrectly during in vivo instruction.

The most efficacious application of simulation and in vivo training is likely to be affected by variables such as the nature of the task, as well as by the training goals (e.g., mastery of a constant, circumscribed set of stimuli vs. responsiveness to a relatively large number of stimulus variations). The selection of training procedures is often guided by multiple considerations, including prevailing treatment philosophies. Treatment decisions that are made in the scientific rather than the political arena are likely to yield the most satisfactory outcomes. Continued investigation of the conditions under which alternative procedures are likely to be effective may offer opposing factions a common goal and a larger treatment arsenal.

REFERENCES

- Baer, D. M. (1977). "Perhaps it would be better not to know everything." *Journal of Applied Behavior Analysis*, **10**, 167-172.
- Baer, D. M., Wolf, M., & Risley, T. R. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, **20**, 313-327.
- Bellamy, G. T., Horner, R. H., & Inman, D. (1979). *Vocational habilitation of severely retarded adults: A direct service technology*. Baltimore: University Park.
- Belmore, K., & Brown, L. (1978). A job skill inventory strategy designed for severely handicapped potential workers. In N. Haring & D. Bricker (Eds.), *Teaching the severely handicapped* (Vol. 3). Columbus, OH: Special Press.
- Brown, L., Nisbet, J., Ford, A., Sweer, M., Shiraga, B., York, J., & Loomis, R. (1983). The critical need for nonschool instruction in educational programs for severely handicapped students. *Journal of the Association for the Severely Handicapped*, **8**, 71-77.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation*. Boston, MA: Houghton Mifflin.
- Coon, M. E., Vogelsberg, R. T., & Williams, W. (1981). Effects of classroom public transportation instruction on generalization to the natural environment. *Journal of the Association for the Severely Handicapped*, **6**, 46-53.
- Cuvo, A. J., Jacobi, L., & Sipko, R. (1981). Teaching laundry skills to mentally retarded students. *Education and Training of the Mentally Retarded*, **16**, 54-64.
- Ford, A., & Mirenda, P. (1984). Community instruction: A natural cues and corrections decision model. *Journal of the Association for the Severely Handicapped*, **9**, 79-87.
- Freagon, S., & Rotatori, A. F. (1982). Comparing natural and artificial environments in training self-care skills to group home residents. *Journal of the Association for the Severely Handicapped*, **7**, 73-86.
- Giangreco, M. F. (1983). Teaching basic photography skills to a severely handicapped young adult using simulated materials. *Journal of the Association for the Severely Handicapped*, **8**, 43-49.
- Hains, A. H., & Baer, D. M. (1989). Interaction effects in multielement designs: Inevitable, desirable, and ignorable. *Journal of Applied Behavior Analysis*, **22**, 57-69.
- Haring, T. G., Kennedy, C. H., Adams, M. J., & Pitts-Conway, V. (1987). Teaching generalization of purchasing skills across community settings to autistic youth using videotape modeling. *Journal of Applied Behavior Analysis*, **20**, 89-96.
- Hill, J. W., Wehman, P., & Horst, G. (1982). Toward generalization of appropriate leisure and social behaviors in severely handicapped youth: Pinball machine use. *Journal of the Association for the Severely Handicapped*, **6**, 38-44.
- Hopkins, B. L. (1987). Comments on the future of applied behavior analysis. *Journal of Applied Behavior Analysis*, **20**, 339-346.
- Horner, R. H., & McDonald, R. S. (1982). Comparison of single instance and general case instruction in teaching a generalized vocational skill. *Journal of the Association for the Severely Handicapped*, **8**, 7-20.
- Horner, R. H., Sprague, T., & Wilcox, B. (1982). General case programming for community activities. In B. Wilcox & G. T. Bellamy (Eds.), *Design of high school programs for severely handicapped students* (pp. 61-98). Baltimore: Paul H. Brookes.
- Horner, R. H., Williams, J. A., & Stevely, J. D. (1987). Acquisition of generalized telephone use by students with severe mental retardation. *Research in Developmental Disabilities*, **8**, 229-247.
- Huitema, B. E. (1980). *The analysis of covariance and alternatives*. New York: John Wiley.
- Johnston, J. M. (1988). Strategic and tactical limits of comparison studies. *Behavior Analyst*, **11**, 1-9.
- Marchetti, A., McCartney, J., Drain, S., Hooper, M., & Dix, J. (1983). Pedestrian skills training for mentally retarded adults: Comparison of training in two settings. *Mental Retardation*, **21**, 107-110.
- McDonnell, J. J., & Horner, R. H. (1985). Effects of in vivo versus simulation-plus-in vivo training on the acquisition and generalization of grocery item selection by high school students with severe handicaps. *Analysis and*

- Intervention in Developmental Disabilities*, 5, 323-343.
- McDonnell, J., Horner, R. H., & Williams, J. (1984). Comparison of three strategies for teaching generalized grocery purchasing to high school students with severe handicaps. *Journal of the Association for Persons with Severe Handicaps*, 9, 123-133.
- Michael, J. (1974). Statistical inference for individual organism research: Mixed blessing or curse? *Journal of Applied Behavior Analysis*, 7, 647-653.
- Morrow, S. A., & Bates, P. E. (1987). The effectiveness of three sets of school-based instructional materials and community training on the acquisition and generalization of community laundry skills by students with severe handicaps. *Research in Developmental Disabilities*, 8, 113-136.
- Neef, N. A., Iwata, B. A., & Page, T. J. (1978). Public transportation training: In vivo versus classroom instruction. *Journal of Applied Behavior Analysis*, 11, 331-344.
- Nietupski, J., Hamre-Nietupski, S., Clancy, P., & Veerhusen, K. (1986). Guidelines for making simulation an effective adjunct to in vivo community instruction. *Journal of the Association for Persons with Severe Handicaps*, 11, 12-18.
- Nietupski, J., Welch, J., & Wacker, D. (1983). Acquisition, maintenance, and transfer of grocery item purchasing skills by moderately and severely handicapped students. *Education and Training of the Mentally Retarded*, 18, 279-286.
- Nutter, D., & Reid, D. (1978). Teaching retarded women a clothing selection skill using community norms. *Journal of Applied Behavior Analysis*, 11, 475-487.
- Page, T. J., Iwata, B. A., & Neef, N. A. (1976). Teaching pedestrian skills to retarded persons: Generalization from the classroom to the natural environment. *Journal of Applied Behavior Analysis*, 9, 433-444.
- Sprague, J. R., & Horner, R. H. (1984). The effects of single instance, multiple instance, and general case training on generalized vending machine use by moderately and severely handicapped students. *Journal of Applied Behavior Analysis*, 17, 273-278.
- Thompson, T. J., Braam, S. J., & Fuqua, R. W. (1982). Training and generalization of laundry skills: A multiple probe evaluation with handicapped persons. *Journal of Applied Behavior Analysis*, 15, 177-182.
- van den Pol, R. A., Iwata, B. A., Ivancic, M. T., Page, T. J., Neef, N. A., & Whitley, F. P. (1981). Teaching the handicapped to eat in public places: Acquisition, generalization, and maintenance of restaurant skills. *Journal of Applied Behavior Analysis*, 14, 61-69.
- Van Houten, R. (1987). Comparing treatment techniques: A cautionary note. *Journal of Applied Behavior Analysis*, 20, 109-110.
- Vogelsberg, R. T., Williams, W., & Bellamy, G. T. (1982). Preparation for independent living. In B. Wilcox & G. T. Bellamy (Eds.), *Design of high school programs for severely handicapped students* (pp. 153-173). Baltimore: Paul H. Brookes.

Received August 2, 1989

Initial editorial decision October 5, 1989

Revision received March 20, 1990

Final acceptance May 16, 1990

Action Editor, David P. Wacker