PROGRAMMING THE GENERALIZATION OF A GREETING RESPONSE IN FOUR RETARDED CHILDREN[1]

TREVOR F. STOKES, DONALD M. BAER[2], AND ROBERT L. JACKSON

UNIVERSITY OF WESTERN AUSTRALIA, UNIVERSITY OF KANSAS, AND
PYRTON TRAINING CENTER

Reinforcement techniques of prompting and shaping were employed to develop hand-waving, a useful social greeting response, in four institutionalized retarded subjects. A multiple-baseline design across subjects demonstrated the reliable functioning of the training procedures. Specifically, it showed that training and maintenance of the greeting response by one experimenter was not usually sufficient for generalization of the response to the more than 20 other members of the institution staff who had not participated in the training of the response. However, high levels of generalization to staff members were recorded for three subjects over periods ranging from one to six months after a second experimenter trained and maintained the response in conjunction with the first experimenter. The fourth subject, although never receiving training by a second experimenter, showed similar results following a second training by the first experimenter.

The usual need for generalization of therapeutic behavior change is widely accepted, but it is not always realized that generalization does not automatically occur simply because a behavior change has been accomplished. Thus, the need to program generalization, rather than expect it, is a point requiring both emphasis (Baer, Wolf, and Risley, 1968; Lovaas, Koegel, Simmons, and Long, 1973; Patterson, McNeal, Hawkins, and Phelps, 1967) and effective techniques.

Unfortunately, only a few research studies establish useful techniques for programming generalization across experimenters. Redd (1970) and Redd and Birnbrauer (1969) examined some conditions under which retarded children's cooperative play will generalize to the presence of training and nontraining experimenters. Corte, Wolf, and Locke (1971) and Lovaas and Simmons (1969) analyzed the generalization across experimenters of the punishment of self-destructive behavior by retarded children. In general, they found that training by two or three experimenters was required to promote generalization to a few other experimenters not involved in training. Similarly, Garcia (1974) taught a conversational speech form to nonverbal retarded children; generalization to a third experimenter was exhibited only after the subjects had experienced specialized training by two other experimenters.

Another prominent example of this sort of finding was the development, maintenance, and programmed generalization of a greeting response, "Hello" in hospitalized adult schizophrenics by Kale, Kaye, Whelan, and Hopkins (1968). In that study, prompts and cigarette reinforcers increased the rate of greetings, which were discriminated to the experimenter's ap-

proach. However, when little or no generalization was found with a new "test" person not involved in the training, Kale et al. employed five more trainers (all at once) to reinforce the response, which was sufficient to promote response generalization to the "test" person. However, this study allowed a possible confounding of time and further reinforced trials with the generalization-programming procedures, as evidenced by possible upward trends in the generalization data before the use of the five additional trainers.

The present study employed a multiple-baseline design across children who experienced varying amounts of time and trials before encountering the experimental treatment. This design controlled for coincidental changes in the greeting behavior of the subjects by showing that the experimental variables effected the desired behavior changes only when these procedures were applied to each subject, rather than at any prior times. Furthermore, the present study examined the use of additional trainers to promote generalization more independently of time and trials factors: training by additional experimenters was not introduced until the percentages of response by the subjects had stabilized or displayed a clear downward trend. In addition, the range of generalization across experimenters was examined more extensively than by Kale et al. (1968): it was asked whether fewer than five additional trainers might be adequate to promote generalization and whether certain individuals facilitate generalization more than others.

METHOD

Subjects

Four retarded children, all permanent residents of a state institution,[3] were classified as severely or profoundly retarded. David, Bruce, and Wayne were males aged 11, 13, and 13 yr respectively; Kerry was a 10-yr-old girl. These

---

[3]Pyrton Training Centre, Eden Hill, Western Australia.

subjects were chosen on the basis of having no critical physical disability (e.g., blindness) and no useful form of social greeting behavior.

All of the subjects lived in the same dormitory with about 12 other children. The dormitory was organized, in part, according to operant conditioning principles. Thus, while serving as subjects in this study, the children were also involved in the continuing programs of the dormitory. However, none of these was concerned with the development of social greeting behavior.

Response Measurement

The hand-wave was chosen as a common example of a useful social greeting. A response was scored as a wave if it met the following criteria: (1) the hand, which had to be empty, was raised above elbow level; and (2) there were at least two back-and-forth motions, either of a single arm from the shoulder or elbow, or a single hand from the wrist, mainly in the vertical plane, which did not contact any part of the subject's body or any other person or object.

Two specific exclusions were made: pointing responses in which a single finger was outstretched, and fending-off responses resembling the gesture of a police officer halting traffic.

All responses were classified in one of three ways: (1) spontaneous response: a correct response emitted within 10 sec of the trainer or prober and the subject coming within 3 ft (0.9 m) of one another; or (2) prompted response: a correct response emitted only after some visual and/or manipulative prompt by the trainer, such as holding the reinforcer in front of the subject, or taking his arm and moving it through the motions of the defined wave response; or (3) incorrect response: any response within the specified 10 sec that did not meet the definition requirements.

The reliability of each prober's and each trainer's recording of greeting responses was assessed by one of the trainers, who independently scored the response of the subject to the trainer or prober making the contact. Later comparison of all such response classifications gave a measure

of reliability (as a percentage of agreements). Reliability checks were made on at least one probe day for each subject during every probe. In addition, an unobtrusive reliability estimate was made by a trainer (E1) scoring the response of a subject without the prober knowing this was being done. The trainer noted the prober's name, and the time, place, and response of the subject during a contact, and only when he was out of sight of the prober did he write these down. Later comparison of these response recordings gave an unobtrusive measure of reliability.

The reliability of the probers' and the trainers' recording of greeting responses was assessed on 266 contacts; responses were scored as spontaneous on 107 contacts. Reliability ranged from 98% to 100% for the four subjects, with a mean of 99%. The unobtrusive measure of reliability was based on 52 contacts and was 96%.

Design

The experimental design was a multiple-baseline or "sequential analysis" design across subjects (Baer and Sherman, 1970). In general, the study proceeded by a series of alternations of training periods and probes for any generalization resulting from that training. Initially, all probers tested for any existing tendency of subjects to give greetings to them. Then, one trainer, E1, taught the first subject to give responses to him. All probers then probed this child and all others to see if generalization had occurred. If generalization had not occurred or was transitory, one additional trainer, E2, also taught the subject (concurrent with the ongoing maintenance by the first trainer), and another probe for generalization was made of all children by all probers. This pattern of incremental additions of one more trainer after each probe was to continue until satisfactory generalization occurred. When one subject's response was considered to have generalized, E1 then began teaching the next subject. This cycle was continued until the responses of all subjects were considered to have generalized.

Training and Probe Contacts

Training of the greeting response and probes of its resultant generalization both consisted of a number of contacts with a subject. Contacts were initiated by the trainers or probers under the following conditions. The subject was approached within 3 ft (0.9 m) at times when he was standing or sitting, but not lying or running. The trainer or prober always stood within the subject's field of vision. A contact lasted for 10 sec, or until a greeting response was given, whichever was first. There was at least a 15-min break between successive contacts by the prober, and there was a break between contacts by successive probers. The probers were instructed that for 5 min before a contact with a subject, they should remain out of the subject's sight, or, if the subject was in the courtyard, to have been at least 15 yd (13.5 m) from the subject for the 15 min preceding the contact.

Training

Training was conducted by the senior author (E1), working as an assistant in a dormitory, and by the staff psychologist (E2) consulting with that dormitory.

Training contacts were made in four settings. Initial sessions (usually the first training day for the subject) were carried out in a small room, where a large number of responses were prompted for initial shaping of the greeting response. Following this, training contacts were made in a dormitory corridor, playroom, and courtyard, with approximately equal frequency each day. Approximately 20 training contacts were made on any one day. During the first few days of training, prompts were used liberally, and then were gradually decreased in number. After a few days of training, as the greeting response developed, prompts were not used at all.

During the initial training day, there were no time limits governing the frequency of contacts. Training usually commenced with breaks between contacts of a few seconds to a few minutes. These intervals were gradually increased until at

least a 15-min break between contacts was required, such conditions commencing on the same day as the use of prompts was discontinued. Setting restrictions, as outlined above for probers, were also specified as of that day. A day on which prompts had to be given did not count as a "day" in defining the criterion of sufficient training (described next).

The criterion of sufficient training was defined as: (1) two consecutive days of the calendar on which the percentage of spontaneous greeting responses was 85% or higher; or (2) three successive days on which the percentage of spontaneous greeting responses was 85% or higher, if there was an intervening break in training of at least one calendar day between the first and second criterion day, or before the next probe was to begin. When this criterion had been met, training contacts were reduced in frequency to a maintenance schedule (as described below) and probes for generalization were made.

During probes for generalization, training was maintained by the trainer(s), each of whom averaged three to six contacts per day with the subject on about three-quarters of the days on which that subject was probed. Also, after each subject's response had generalized thoroughly, his training was continued on a maintenance schedule even during the training of other subjects: on at least one-quarter of the experimental days when the current subject-in-training was trained, the greetings of previously trained subjects were maintained by primary and social reinforcement three to six times each day.

During training, the experimental consequences of a spontaneous or prompted greeting response were M&Ms or potato chips, depending on the subject's known preference, and a social stimulus such as a smile, the words "Hello (name)" and often physical contact such as a pat on the head. These were administered on a continuous schedule, i.e., after each response.

Probes

During those days when generalization was being assessed, as few as four or as many as 14 probers, unsystematically sampled from an available pool of at most 23 staff members on that day, might be recording a subject's greeting responses to them. On the average, about eight probers recorded greeting responses on any one day. These probers were all members of the training staff of the dormitory, known professionally as Training Assistants. The Training Assistants were directly responsible for the well-being, activities, and training of the children in the dormitory. These assistants happened to be familiar with operant procedures and were experienced in behavioral data collection. As is usual in such institutions, several assistants resigned their positions while the study was in progress, whereupon new assistants commenced work in the dormitory in their places. As these new assistants began working in the dormitory, they also participated in the study as probers, i.e., they assessed each subject's generalization of the greeting response to them.

A probe for generalization consisted of each prober making a number of contacts with the four subjects. Every other day, each prober was to make three contacts with each of two subjects. On the alternate days, the other two subjects were probed. Probes were conducted over at least eight days, and often more. Thus, each subject was probed on at least four days of each probe (period) of the study, each day consisting of a total of approximately 20 contacts. If, after an eight-day probe, the trend of the generalization data was unclear or unsettled for any of the subjects, then the probes were continued (with all subjects, two subjects on each alternate day) until the direction and/or level of the generalization was judged to be clear.

During probes, a contact was made, and if a spontaneous greeting response was emitted, it was answered only with a smile, "Hello (name)" and often physical contact. This was the usual social consequence of a greeting in that dormitory before the study. The prober then went on about his normal activities. If the subject did not wave, the prober moved on after 10 sec had elapsed. No prompts were given during these
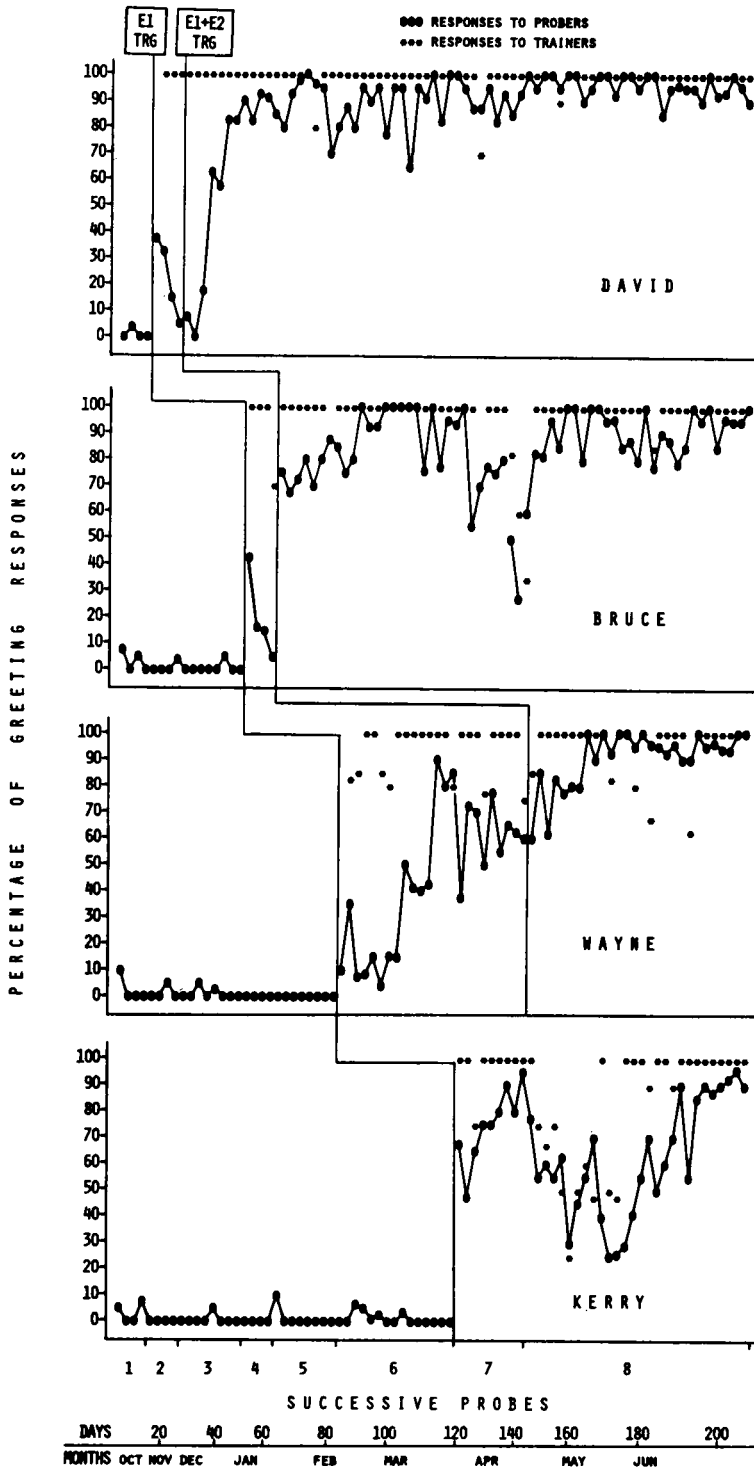
Fig. 1. Generalization of greetings to probers, and maintenance by trainers, as a percentage of contacts made, across the four subjects of the study. (In the case of Bruce, two probe points just subsequent to Day 140 are not connected to the other points. These represent two probes conducted during a three-day period of increased drug administration aimed at controlling disruptive behavior, as discussed in the Results section.)

contacts. If the subject walked away from the prober before waving, he was not followed; the prober remained facing in the same direction and stayed for the 10 sec. If the subject returned within the 10 sec and gave a greeting response, this was counted as a spontaneous response and was answered. If the subject did not return, an incorrect response was recorded.

Probes were usually conducted daily, except when the subject being trained went home for the weekend, or E2 had off-duty days. Contacts were made throughout the day during the normal activities within the dormitory. The probers were asked to make one contact per subject per probe day in each of the corridor, playroom, and courtyard settings, and to make some contacts in the morning and some in the afternoon. Thus, all assistants on duty during a particular probe day made three contacts with each of the two children being probed on that day. Immediately after the contact, results were recorded by the probers on special forms provided for that purpose. Since at various times during the day the assistants approached the subject as a normal part of the dormitory routine, greetings that occurred under these circumstances were met with a smile and "Hello (name)", but were not recorded.

Whether the probers carried out the assigned number of contacts was also monitored: this efficiency measure was the total number of contacts made, considered as a percentage of the total number possible on that day if all the probers on that day had completed all of their assigned contacts. Over all subjects and probes, the probers completed 7161 contacts of the 8553 requested, an efficiency score of 83%.

RESULTS

Probes

Figure 1 shows for each subject the generalization of greeting responses to all probers and the maintenance of the response by trainers as a percentage of their total respective contacts. Figure 2, a related but less sensitive measure of

the generalization, shows for each subject the range of the greeting responses across probers as a percentage of probers receiving at least one greeting response during contacts made by them on that day. In Figure 2, the range of the response to the trainers, E1 and E2, was not plotted because it was always 100%.

Both Figure 1 and Figure 2 show that after near-zero baselines of generalization had been established for all subjects, each successive behavioral intervention by E1 led to a prompt and significant, but not necessarily durable, increase in generalization.

David evidenced only a transitory generalization after training by E1. After training by E2 as well, his level of generalization rose to be consistently above 80%. Thus, training and maintenance by two trainers was associated with the level of generalization exhibited over the next six months.

Bruce showed a similar (but longer) baseline, transitory generalization after training by E1, and generalization consistently above 80% after training by E1 and E2. Thus, again, training and maintenance by two trainers was associated with the maintenance of a high level of generalization over a four-month period. One significant deviation may be noted. During Days 142 to 144 (Probe 7), the daily introduction of 2 mg of a sedative drug (anetensol) apparently produced a marked reduction in response in all areas of activity. Reduction of this dosage on Day 145 was followed by increased response in all areas.

Wayne's training and maintenance by E1 over an extended period of time was associated with an increase in generalization to an apparently stable 61% average during Probe 7. This generalization also increased to a level well over 80% after training and maintenance by E2 was introduced.

Kerry, after introduction of training by E1, quickly generalized to a level above 80% (Probe 7). However, the first six days of Probe 8 showed a marked lowering of these percentages to *both* the trainer and probers, due mainly to a shift in the topography of her responses, such

that they often fell short of the definition of a spontaneous wave. After a further period of intensive training by E1 (begun on Day 164) aimed at correcting this response-topography shift, her generalization increased to the previous high level and remained there. Thus, training and maintenance by one trainer, E1, over an extended period of time, but with a second phase of intensive, remedial shaping required, yielded as much generalization over the last two weeks of the study as the other subjects had evidenced only after training and maintenance by both E1 and E2.

At the completion of the study, the four subjects were greeting a large number of probers, even though they were receiving no more consequences from the probers for each of these responses than they had received before training. Across the days of each probe, generalization of the greeting response was measured to approximately 15 different probers, on the average. However, a more significant fact is that during the eight months of the study, the continued high levels of generalization were measured to a slowly changing group of probers: a total of 27 probers participated in the study. David, during the six months when his generalization was consistently in the 80% to 100% range, responded to 23 different probers, Bruce to 23 probers during five months, Wayne to 22 probers during two months, and Kerry to 20 probers during one month. During these times, high percentages of response generalization were seen immediately to each new prober as he entered the study.

During all probes, with all subjects, a mean of eight probers completed a mean of 22 contacts on each day with each of the subjects being probed on that day.

Training

Percentages of spontaneous and prompted greeting responses to trainers during the seven training periods are given in Figure 3. In general, all training was effective: only four to 10 training days were required to reach the specified criterion.

Range of Generalization across Probers

Differences in generalization to different probers were examined during transitional probes. Transitional probes were those probes during which greeting responses initially increased from a near-zero to a higher level of generalization, *i.e.*, Probes 2 and 3 for David, Probes 4 and 5 for Bruce, Probe 6 for Wayne, and Probe 7 for Kerry. For this comparison, only transitional probe data were appropriate: greeting percentages during earlier and later probes showed too little variance to allow comparison of the probers' effectiveness in evoking waves.

Generalization to individual probers by David during his transitional probes ranged from 20% to 72% across the 14 different probers involved; for Bruce, the rates ranged from 36% to 88% across 13 probers; for Wayne, 12% to 68% across 14 probers; and for Kerry, 44% to 100% across 13 probers.

Probers were ranked in order of their probability of receiving greeting responses from each child during transitional probes. Only those probers who probed at least three subjects at least 12 times each during the transitional probes were considered. Ten probers were so qualified for David and Kerry, and 11 probers for Bruce and Wayne. Spearman rank-order correlation coefficients were then computed for each possible pair of subjects, *i.e.*, a correlation between the rank orders of those probers who qualified for both subjects of the pair. Such correlations, if high and positive, would indicate consistent effectiveness or ineffectiveness by the probers in evoking waves from the two subjects of each pair. The obtained correlation coefficients were:

$$r\text{D.B} = 0.47$$
$$r\text{D.W} = -0.63$$
$$r\text{D.K} = 0.14$$
$$r\text{B.W} = -0.25$$
$$r\text{B.K} = -0.17$$
$$r\text{W.K} = 0.37.$$

None of these was very large or statistically significant (0.05 level, one-tailed test). Thus, the
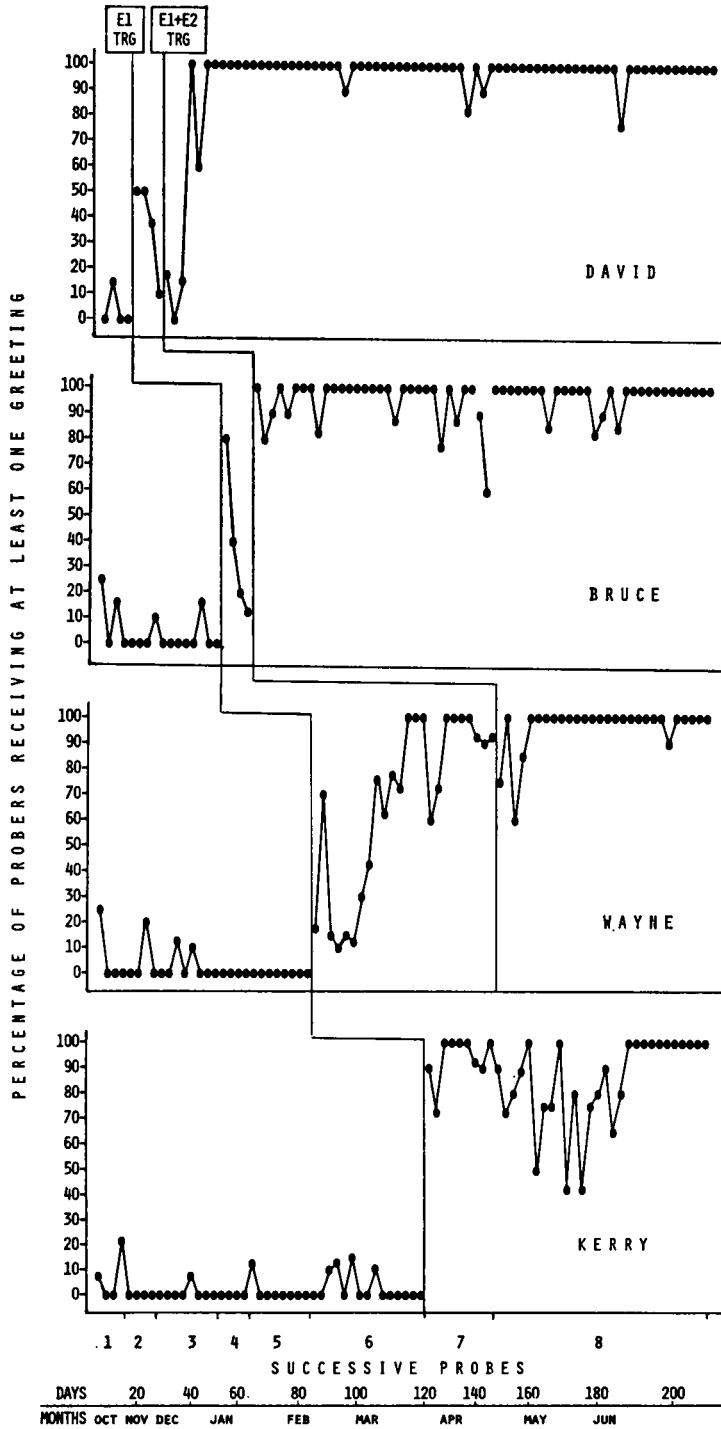
Fig. 2. Range of the generalization of greetings across probers, as a percentage of probers receiving at least one greeting on that day, across the four subjects of the study. (In the case of Bruce, two probe points just subsequent to Day 140 are not connected to the other points. These represent two probes conducted during a three-day period of increased drug administration aimed at controlling disruptive behavior, as discussed in the Results section.)
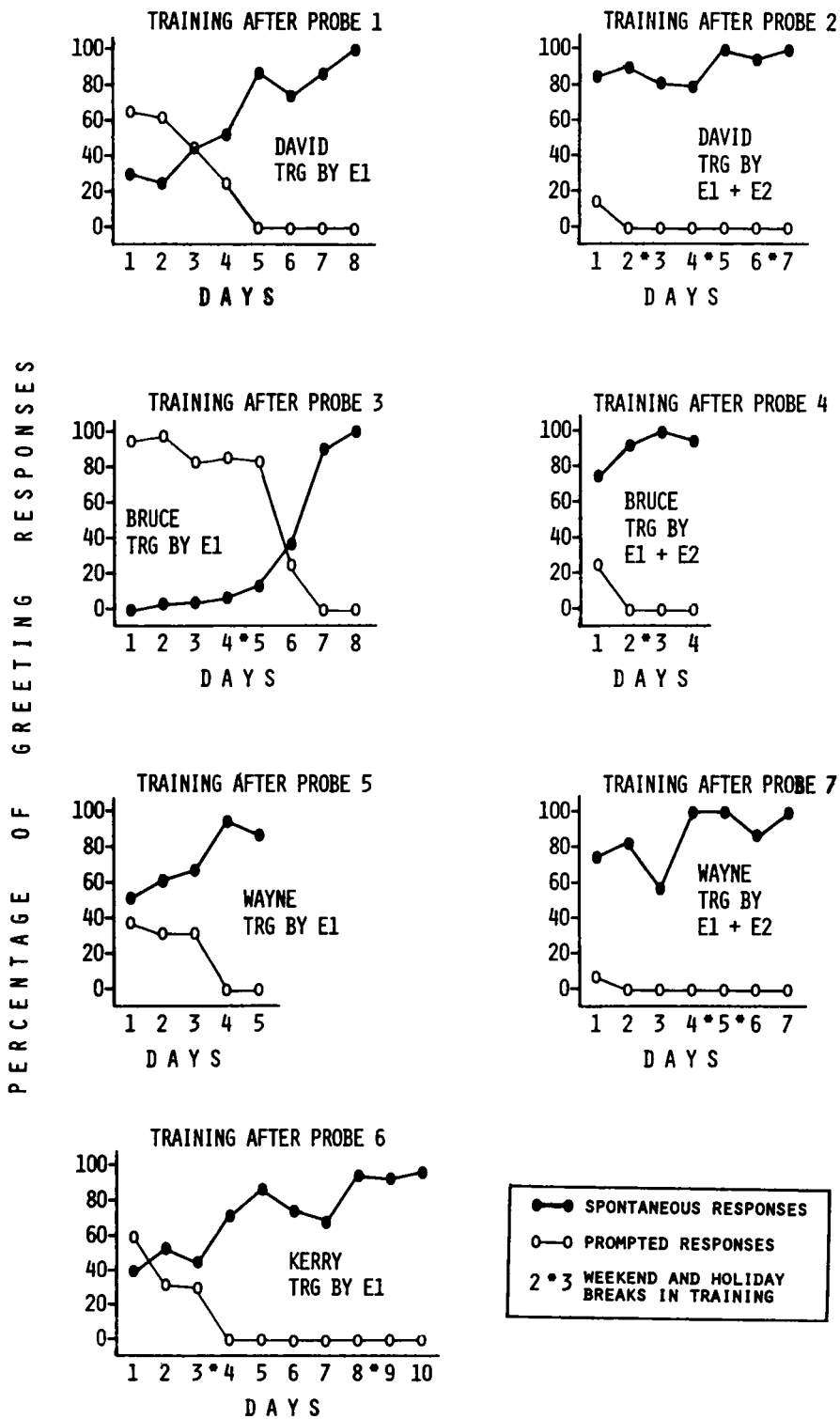
Fig. 3. Percentage of spontaneous greetings to trainers, and of prompts by trainers, during the one or two training phases required by each subject.

effectiveness of different probers in evoking waves was not consistent across any pair of subjects—there was no evidence of a generalized "greetableness" in probers.

Supplemental Results

Recorded only at an anecdotal level, generalization from the wave-greeting response to a related social greeting behavior was noted: David and Wayne both developed a distinctive "hello" response that usually accompanied their waves. Presumably, this response was in imitation of part of the adult social response, given contingent on waving by the subject.

In Bruce's case, it was found that the performance of the greeting response had overgeneralized, *i.e.*, greeting responses were given in inappropriate situations (*e.g.*, when his shoelace was being tied). After completion of the formal part of the study, discrimination training was successful in restricting Bruce's greetings to appropriate situations.

DISCUSSION

This study showed the experimental programming of the generalization of a greeting response. Reinforcement techniques of prompting and shaping developed the response in four retarded children, who, before the study, lacked any appropriate form of useful greeting behavior. The initial development of the generalized greeting response and the programming of its generalization to a large number of probers not involved in its training (when necessary) were examined in a multiple-baseline design, across subjects, showing the reliable functioning of both the training and the generalization-programming procedures. That is, no shift from near-zero baseline levels was seen in any subject before training; no subject failed to shift promptly after the training; three of four subjects required two trainers' input before extensive generalization, yet training by one or two trainers occurred at different times and after different numbers of baseline days for each subject. Some analysis of the generality of

the procedures' effects was made possible through the intersubject replication of this multiple-baseline design, and through the extended measurement of the behaviors over time.

With three of the four subjects, training and maintenance of the greeting response by two trainers was sufficient to promote the generalization of the response to more than 20 probers who had not participated in the training. One trainer/maintainer was found to be sufficient for similar generalization by Kerry; however, she did require a second phase of intensive training (by the same trainer) to correct response-topography shifts that became evident after a high initial level of generalization had been achieved.

Clearly, the widespread generalization evidenced after training by only two experimenters is a most practical and efficient outcome that has important implications for programming generalization across experimenters: after subjects experienced a diversity of training with a few experimenters, generalization was exhibited in conditions well beyond those of the experimental intervention.

Generalization to the probers was evident in the three settings in which most training took place: playroom, corridor, and courtyard. Although these were not a widely diverse range of environments, a number of anecdotal reports suggested that waves occurred in other settings as well, *e.g.*, the dining room and bathroom, as well as city shops, streets, parks, and other settings outside the dormitory confines, which the children occasionally visited under supervision.

The data showing differential generalization to different probers suggest a caution: when generalization is tested to only one prober (*e.g.*, Kale *et al.*, 1968) it is possible that the choice of the particular prober might well determine the conclusion drawn. The most marked example in the present study was that of David. During Probe 2 and the first three days of Probe 3, total generalization ranged from 9% to 38%. However, if it had been measured by any of a particular six probers, it would be concluded that no generalization had occurred at all. But if gen-

eralization were measured by two other specific probers over this seven-day period, it would have been judged to be about 50%. Similar distributions were found with Bruce and Wayne during their transitional probes. Thus, the large range of individual differences in the rates of responses received by the probers is a factor that should be considered in future research.

An "emergent" generalization was seen in each subject, *i.e.*, each subject showed a considerable upward trend in the generalization data, rather than displaying immediate high levels after training. In the case of David, Probe 3, the percentage of greeting responses increased from 6% to 89%; with Bruce, Probe 5 and the first eight or 10 days of Probe 6, from 74% to 100%; with Wayne, Probe 6, from 8% to 86%; and with Kerry, Probe 7, from 67% to 97%.

One possible explanation for such emergent generalization is the functional introduction, through training, of the child's social response into a natural shaping environment (Baer and Wolf, 1970). That is, after training of the greeting response, the subject contacted the systematic environmental social consequences contingent upon that response much more frequently, though not more consistently. This may have had the effect of steadily increasing the response rate. Unfortunately, the functional role of the probers' social consequences in increasing and/or maintaining the level of generalization was not examined in the present study because of time limitations. In future studies, the separate roles of the edible and social consequences might well be systematically examined.

Close examination of the generalization data also reveals a U-shaped function, most noticeable with David, and possible with Bruce and Wayne during the transitional probes. This function may be a combination of simple generalization effects, fading after training, but accomplishing the introduction of the behavior to a shaping environment.

The study endeavored to show the independence of time/trials factors from the generaliza-

tion-programming procedures, by allowing sufficient time for trends in the data to become clear before introducing training by additional experimenters. Even so, examination of the generalization data of David and Bruce, in light of the data of Wayne and Kerry, suggests the possibility that the introduction of the second trainer to David and Bruce was premature: their levels of generalization might have increased over an extended period of time without the introduction of E2. This possibility was partially examined with Wayne, where probes for generalization continued for 24 probe days before the second trainer was introduced. Systematic examination of such factors might be made by employing another multiple-baseline design similar to the present study. After a stable baseline rate has been established for all subjects, all could be trained by a first experimenter. Then, on a staggered basis, the second trainer could be introduced to one subject at a time. In such a design, the increasing lengths of time before introduction of the second trainer would control for a confounding of time and trials with the important procedure of introducing training and maintenance by the second trainer.

REFERENCES

Baer, D. M. and Sherman, J. A.   Behavior modification: clinical and educational applications. In H. W. Reese and L. P. Lipsitt (Eds.), *Experimental child psychology*. New York: Academic Press, 1970. Pp. 643-672.

Baer, D. M., Wolf, M. M., and Risley, T. R.   Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1968, 1, 91-97.

Corte, H. E., Wolf, M. M., and Locke, B. J.   A comparison of procedures for eliminating self-injurious behavior of retarded adolescents. *Journal of Applied Behavior Analysis*, 1971, 4, 201-213.

Garcia, E.   The training and generalization of a conversational speech form in nonverbal retardates. *Journal of Applied Behavior Analysis*, 1974, 7, 137-149.

Kale, R. J., Kaye, J. H., Whelan, P. A., and Hopkins, B. L.   The effects of reinforcement on the modification, maintenance, and generalization of social responses of mental patients. *Journal of Applied Behavior Analysis*, 1968, 1, 307-314.

Lovaas, O. I., Koegel, R., Simmons, J. Q., and Long, J. S. Some generalization and follow-up measures on autistic children in behavior therapy. *Journal of Applied Behavior Analysis*, 1973, **6**, 131-166.

Lovaas, O. I. and Simmons, J. Q. Manipulation of self-destruction in three retarded children. *Journal of Applied Behavior Analysis*, 1969, **2**, 143-157.

Redd, W. H. Generalization of adults' stimulus control of children's behavior. *Journal of Experimental Child Psychology*, 1970, **9**, 286-296.

Redd, W. H. and Birnbrauer, J. S. Adults as discriminative stimuli for different reinforcement contingencies with retarded children. *Journal of Experimental Child Psychology*, 1969, **1**, 440-447.

Patterson, G. R., McNeal, S., Hawkins, N., and Phelps, R. Reprogramming the social environment. *Journal of Child Psychology and Psychiatry*, 1967, **8**, 181-195.