Across Setting Generalization of Social Conventions in Aphasia: An Experimental Analysis of "Loose Training"

Cynthia K. Thompson and Marie E. Byrne
Pennsylvania State University, University Park, Pennsylvania

Generalization of treatment effects to stimulus conditions that differ from those in which training takes place is considered to be an essential portion of the treatment process. That is, therapeutic gains seen in the clinical environment must also be evident in the natural environment or "intervention cannot be considered effective" (Chapey, 1981, p. 364) and the value of treatment must be questioned (Guess, Keough, and Sailor, 1978; Warren, Rogers-Warren, Baer, and Guess, 1980).

Investigations of such generalization, referred to as stimulus generalization, are limited in the aphasia literature and the results of available studies are inconsistent. For example, studies by Goodkin, Diller and Shah (1973), Jennings and Lubinski (1981), and Schlanger and Schlanger (1970) have suggested improved responding in natural environments following aphasia treatment, whereas, Egolf and Chester (1977) reported limited stimulus generalization of the effects of treatment. Further, the generalization data for many of these studies were derived from questionnaires or from anecdotal reports provided by family members or rehabilitation personnel, a method of data collection which fails to control for many confounding variables. Thus, conclusive statements are difficult to make. In addition, research in this area has primarily consisted of case study reports or small group investigations in which experimental control was not demonstrated. Finally, of the stimulus generalization studies reviewed, none were focused on facilitation of generalized use of language when it was lacking.

One method for enhancing generalization across settings has been discussed by Stokes and Baer (1977), Baer (1980) and Costello (1983). This method, referred to as "loose training" is concerned with diminishing the difference between the treatment setting and the natural environment by allowing variation in language responses, discriminative stimuli, and feedback schedules employed in treatment. In "loose training" these parameters are varied or progressively altered to approximate those occurring in the natural environment. One study in the applied behavioral literature by Campbell and Stremmel-Campbell (1982) has investigated the effects of such treatment with language delayed children. However, the effect of "loose training" on language generalization in aphasia has not been examined.

PURPOSE

The purpose of this study was to investigate generalization of Broca's aphasic subjects' use of social conventions using a "loose training" strategy. Social conventions included greetings, self-disclosures and question responses. Generalization was assessed across settings and persons. Specific responses were trained in individual treatment sessions by the examiner, while generalized use of trained responses was assessed in novel social dyads consisting of the subject and an unfamiliar person or confederate. The "loose training" strategy included systematic and progressive alteration of the discriminative stimuli and schedule of feedback. This training represented a departure from "loose training" as described by Stokes and Baer (1977) in that specific language responses were trained.
METHOD

Subjects. Three Broca's aphasic subjects, one female and two males ranging in age from 47 to 62 years, participated in the study. Subjects were between 14 and 20 months post onset of a single left CVA and exhibited residual right hemiparesis. All subjects were native English speakers, had normal hearing acuity, and had completed high school prior to their infarction.

The diagnosis of Broca's aphasia was based on the Western Aphasia Battery (Kertesz, 1979). Performance on selected subtests is shown in Table 1. These data indicate that all subjects had relatively unimpaired auditory comprehension and could repeat simple words and phrases. However, prominent word retrieval difficulties and nonfluent verbal output were apparent.

Table 1. Subjects' performance on the Western Aphasia Battery.

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluency</td>
<td>1.0</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Comprehension</td>
<td>6.9</td>
<td>7.8</td>
<td>6.5</td>
</tr>
<tr>
<td>Repetition</td>
<td>3.2</td>
<td>3.9</td>
<td>3.0</td>
</tr>
<tr>
<td>Naming</td>
<td>2.9</td>
<td>3.4</td>
<td>2.6</td>
</tr>
<tr>
<td>Aphasia Quotient</td>
<td>28</td>
<td>34</td>
<td>26</td>
</tr>
</tbody>
</table>

A single-subject, multiple baseline design across subjects was used to assess the effects of treatment (Hersen and Barlow, 1976; Kazdin, 1982; McReynolds and Kearns, 1983). The dependent variable (percent correct production of social conventions in dyads) was measured during the baseline phase for all subjects. Treatment was then applied to Subject 1, while baseline conditions were continued for Subjects 2 and 3. Next, treatment was extended to Subject 2, while baseline data continued to be collected for Subject 3. Treatment was never introduced to Subject 3, who was included to control for the effects of repeated exposure to the probe condition. In this design, experimental control is demonstrated when behavioral change (relative to baseline) is noted when, and only when, treatment is applied.

Baseline. During the baseline phase, each subjects' ability to produce social conventions was assessed in novel social dyads. The social conventions of interest included eight responses (Greetings: Hello and Goodbye. Self-disclosures: I am fine, My name is ____, and I live in _____. Questions: How are you?, What is your name? and Where do you live?). To obtain these data, each subject engaged in a five-minute conversation with an unfamiliar person (confederate) in a comfortable, nontreatment room. Confederates included 58 undergraduate student volunteers who had had no experience with aphasic subjects. Prior to each five-minute probe, confederates were given written instructions to talk with someone for five minutes, during which time they were to greet the person and to attempt to obtain certain information while not divulging information about themselves unless asked (Appendix A). Each social dyad was then scored for the percent of the target social conventions produced correctly by the subject on a plus-minus basis. Greetings and self-disclosures were expected to occur within a 10-second response time following the initial greeting or request made by the confederate. Questions could occur any time

133
within the five-minute probe. To be considered correct, subjects were required to produce the complete expected response, as trained, except for self-disclosures, for which single-word responses which provided the appropriate information were accepted. For example, in response to "How are you?", "Fine" was an acceptable response. Articulatory errors which did not influence intelligibility were permitted, but self-corrections were not. All baseline sessions were videotaped through a one-way mirror.

Treatment. In the treatment condition, subjects were trained to produce targeted social conventions in a small, traditional therapy room. Four levels of training were implemented. At each level, the discriminative stimuli and feedback schedule were progressively altered to approximate those occurring in the natural environment (see Tables 2 and 3). Each subject received a minimum of eight treatment sessions on Levels 1 and 2 and five sessions on Levels 3 and 4. During each session, the eight social conventions were elicited five times each for a total of 40 trials per session. Criterion for movement through the treatment levels was set at 80% correct production of the target responses. At least every third treatment session was audiotaped.

Level 1. The first level of training was focused on imitation (Table 2). During each training trial, one of the eight social conventions was presented by the examiner for the subject to repeat. A 15-second response interval was allowed. The first verbal response to occur within that interval was scored. A response was considered to be correct when the complete response, as modelled, was produced. As in baseline, self-corrections were considered errors, but, articulation errors were permitted. Incorrect responses were modelled and a forward chaining procedure was used. Contingent verbal feedback such as "nice job" or "no, not quite" was provided for correct and incorrect responses on a fixed 1:1 ratio.

Level 2. During Level 2, the subject was required to produce the target social conventions in response to commands given by the examiner (Table 2). Each trial was begun by the examiner presenting a command such as "Tell me your name." A 15-second response time was given and responses were scored (+) or (-) using response criterion consistent with that used during Level 1. Feedback was given on a fixed 1:2 ratio and incorrect responses were modelled.

Level 3. During Level 3, the discriminative stimuli were further loosened for greetings and self-disclosures (Table 2). That is, greetings were elicited by the examiner saying Hello, Hi, Goodbye, or Bye, and self-disclosures were prompted by the examiner randomly asking one of three questions. For example, How are you?, How are you doing? or How are you feeling? was presented to elicit the response I am fine. Question responses continued to be elicited using commands as in Level 2. During each trial the examiner presented a selected stimulus and allowed 15 seconds for a response to occur which was then scored (+) or (-) according to established response criteria. Verbal feedback was reduced to a variable ratio 1:4 and incorrect responses were modelled.

Level 4. In the final level of treatment, the examiner and the subject engaged in five role playing situations during each session (Table 3). Each role play situation was initiated by the examiner saying Hi or Hello. The subject was then expected to respond with a greeting and to produce a question within a 30-second delay time. This delay served as a discriminative stimulus for question production. If a question was not produced, prompts and modelling procedures such as those used in Level 3 were implemented. The examiner then answered each question (providing natural feedback) and asked a new question. Each role play situation was continued until all targeted responses had been produced by the subject either spontaneously or following prompts.
Table 2. Expected responses, discriminative stimuli and Feedback Schedule used at Levels 1, 2 and 3 of training.

<table>
<thead>
<tr>
<th>Expected Responses</th>
<th>Discriminative Stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level 1</strong></td>
<td><strong>Level 2</strong></td>
</tr>
<tr>
<td>Hello</td>
<td>&quot;Hello.&quot;</td>
</tr>
<tr>
<td>How are you?</td>
<td>&quot;Ask me how I am.&quot;</td>
</tr>
<tr>
<td>I am fine.</td>
<td>&quot;Tell me how you are.&quot;</td>
</tr>
<tr>
<td>What is your name?</td>
<td>&quot;Ask me my name.&quot;</td>
</tr>
<tr>
<td>My name is ____</td>
<td>&quot;Tell me your name.&quot;</td>
</tr>
<tr>
<td>Where do you live?</td>
<td>&quot;Ask me where I live.&quot;</td>
</tr>
<tr>
<td>I live in ____</td>
<td>&quot;Tell me where you live.&quot;</td>
</tr>
<tr>
<td>Goodbye.</td>
<td>&quot;Tell me goodbye.&quot;</td>
</tr>
<tr>
<td>Feedback Schedule</td>
<td>FR1:1</td>
</tr>
<tr>
<td></td>
<td>VR1:4</td>
</tr>
</tbody>
</table>
Table 3. Treatment Level 4: Structured role play.

E. Initiates conversation: "hi" or "hello."

S. Replies "Hello" and asks question "How are you?" within 30 second response time. (incorrect responses were prompted)

E. Replies "I am fine" and asks "How are you?"

S. Replies "I am fine" and asks "What is your name?" within 30 second response time. (incorrect responses were prompted)

E. Replies "My name is ____." and asks "Who are you?"

S. Replies "My name is ____." and asks "Where do you live?" within 30 second response time. (incorrect responses were prompted)

E. Replies "I live in ____." and asks "Where do you live?"

S. Replies "I live in ____." "Goodbye." within 30 second response time. (incorrect responses were prompted)

Generalization Probes

Daily (following each treatment session), generalization probes were conducted to assess the subjects' ability to produce responses during social dyads. Procedures identical to those used during baseline were employed. These data served as the dependent variable throughout the study.

Follow-Up

To assess the maintenance of treatment effects, two follow-up probes were administered for each subject three months following the termination of treatment. Again, baseline procedures were employed.

Social Validity

In order to provide a guideline for measurement of treatment effectiveness, a social validity procedure was implemented. Specifically, a social comparison procedure as described by Kazdin (1982) was used to obtain data regarding normal non-brain-damaged individuals' use of social conventions in social dyads. Two normal adults, aged 48 and 56, and eight confederates served as participants. Using procedures identical to those used with the aphasic subjects in the novel social dyads, eight separate interactions were collected and analyzed. All sessions were videotaped. The percent of the social conventions of interest used by the normal adults in each dyad was calculated and the means and standard deviations were computed.

Reliability. Interobserver agreement was calculated on 18 of 56 training sessions (every third training session), 75 of 75 baseline, generalization and follow-up sessions, and 8 of 8 social comparison sessions. Point-to-point agreement between the examiner and an independent observer was obtained using the following formula:
\[
\% \text{ Agreement} = \frac{\text{Total \# Agreements}}{\text{Total \# Agreements} + \text{Disagreements}} \times 100
\]

Percent agreement for training sessions ranged from 96% to 100% with a mean of 99%, and interjudge agreement ranged from 88% to 100% with a mean of 98% on baseline, generalization and follow-up probes, and from 92% to 100% with a mean of 97% for the social comparison probes. These data suggest a high percentage of agreement across all subjects.

RESULTS AND DISCUSSION

Results for Subjects 1, 2 and 3 and the social comparison data are shown in Figure 1. The percent correct production of social conventions in dyads for each subject is shown during baseline, intervention, and follow-up phases of the study. Sessions in which levels of treatment were changed are marked by arrows. The mean for the social comparison group was 65% with a standard deviation of 17%. These data provided a range of responding for the normal group between 48% and 82% production of the targeted social conventions in social dyads. This range is represented by horizontal lines.

![Figure 1. Percent correct production of social conventions in dyads during Baseline, Intervention, and Follow-up phases of the study for Subjects 1, 2, and 3, and the mean and standard deviation of the social comparison group. Arrows indicate sessions in which levels of treatment were changed.]

Subject 1. Examination of the data for Subject 1 indicated that during the baseline phase, production of the social conventions of interest was at a low and stable level (Figure 1). Performance ranged from 13% to 25% correct, which was well below that of the social comparison group. Following baseline, Subject 1 received a total of 26 treatment sessions which resulted in a gradual increase in correct responding on generalization probes taken throughout
the four levels of treatment, with performance ranging from 13% to 75% accuracy. During Levels 1 and 2, training criterion (80% correct production of social conventions during treatment sessions) was reached within eight treatment sessions at each level, resulting in generalized responding within the normal range on 6 of 16 probes. During Levels 3 and 4, training criterion was reached within five sessions at each level, during which time responding on generalization probes was within the normal range on eight of ten probes. On follow-up probes, production of social conventions was maintained at 63% correct, a level which was superior to baseline and approximated the mean for the social comparison group.

Subject 2. Similar results were noted for Subject 2 (Figure 1). Baseline performance ranged from 0% to 13% correct. Treatment, which entailed a total of 30 sessions across the four levels, facilitated a gradual increase in percent correct production of social conventions during generalization probes, with performance ranging from 13% to 63% accuracy. During Level 1, training criterion was reached within eight sessions, resulting in little generalization to the dyad condition. Performance during this level of training never approximated that of the normal group. Thirteen treatment sessions were required during Level 2 to establish criterion level responding, which resulted in generalized performance within the normal range on 3 of 13 probe sessions. For Levels 3 and 4, training criterion was reached within five sessions at each level, during which time performance was within the normal range on eight of ten generalization probes. Responding at this level was maintained on follow-up probes.

Subject 3. Examination of the data for Subject 3, who was held in baseline throughout the study, indicated that performance was stable and at a low level over a nine-week period (Figure 1). Performance ranged from 0% to 25% correct production of social conventions during all dyads.

These data suggest that treatment was effective in facilitating generalized use of social conventions from individual treatment sessions to novel social dyads for the two Broca's aphasic subjects who received treatment and that treatment gains were subsequently maintained. That is, when treatment was applied to Subject 1, improved performance, beyond that which would have been expected based on projected baseline performance, was noted. This effect was replicated following a prolonged baseline phase for Subject 2 and experimental control was demonstrated. Further experimental control was demonstrated by the data derived from Subject 3, in that baserate performance was maintained throughout the study. While it cannot be inferred that Subject 3 would have responded to treatment in a manner consistent with Subjects 1 and 2, Subject 3's stable performance indicated that repeated exposure to the probe condition did not cause changes in behavior.

Social validation of the treatment effect was also shown when the aphasic subjects' performance during baseline and intervention phases was compared with that of the social comparison group. The data indicated that during baseline, the aphasic subjects performed below the range established by the normal group, however, when treatment was applied, performance gradually approximated the normal range. These data indicated that treatment was effective in facilitating behavioral change to a level consistent with that of normal non-brain-damaged individuals' behavior.

**Generalization of Response Types**

Analysis of the data with regard to response types revealed further findings of interest. The number of greetings, self-disclosures and question responses produced during social dyads for Subjects 1 and 2 are shown in Figures 2 and 3, respectively. Examination of Figure 2 showed that Subject 1 produced
Figure 2. Number of greetings, self-disclosures, and questions produced correctly in dyads during Baseline, Intervention, and Follow-up phases for Subject 1.

Figure 3. Number of greetings, self-disclosures, and questions produced correctly in dyads during Baseline, Intervention, and Follow-up phases for Subject 2.
greetings at a variable level during baseline. This performance pattern was continued on generalization probes taken throughout treatment. Thus, treatment can be said to have (perhaps) stabilized these responses. Self-disclosures, however, were produced at a low level during baseline and by the end of treatment all self-disclosures were produced accurately on generalization probes. Conversely, questions, which were never produced during baseline probes, also were not produced during 23 of 28 generalization and follow-up probes. These data indicated that question responses, which were produced at 80% accuracy during training, did not generalize to the dyad condition.

Similar trends were noted for Subject 2. Inspection of Figure 3 shows that greetings, self-disclosures and question responses were all produced at a low and stable rate during baseline and when treatment was applied generalized use of greetings and self-disclosures was evident, whereas question responses were maintained at baserate.

In summary, these data suggest that "loose training," or gradual and systematic change in the discriminative stimuli and feedback schedule used in treatment, enhances generative responding in some aphasic subjects and provides support for the use of this method for programming stimulus generalization. The findings from this study also support the notions underlying functional communication training (Aten, Caligiuri, and Holland, 1982; Davis and Wilcox, 1981). That is, when the training environment is structured to approximate the generalization environment or the natural environment, generalization is heightened. It is not possible to determine from these data, however, the extent to which each level of training was necessary or which had the greatest impact on generalization. The relative contribution of each component of the treatment awaits further investigation. Finally, it is necessary systematically to replicate the effects of "loose training" on the generative use of other linguistic structures by Broca's aphasic subjects in order to determine the lawfulness of these findings.

REFERENCES


140


**DISCUSSION**

Q: I'd like to compliment you on the social validation. I think it lends some credibility to statements you're making about the expected performance in aphasic patients. I thought that was a very nice way to describe the behavior of these patients in a functional way.

A: I think it's important to obtain social comparison data whenever possible to provide a basis for evaluating patient behavior.

Q: I have a question about Subject 3, who was your control. Since you're using a between subject design, there's some concern about the potential in the control subject vs. the two experimental subjects to perform if they were exposed to the treatment conditions. And I notice that the baseline on Subject 3 was a little flat. Do you think there was any difference in the variability of Subject 3's performance on the WAB? Since he was a flatter performer maybe he would have had less potential.

A: I don't think so. Quantitatively there was little difference in the Subjects' performance on the Western Aphasia Battery. The main reason we included Subject 3 was to assess what the effects of continued exposure to the probe would be. Someone might perhaps say that if a patient were engaged in that many interactions with new people, that he might begin to use those responses without ever having received training. That subject did not. It would have been nice to train the third subject to replicate the treatment effects seen in the first two subjects. However, Subject 3 was only included to add strength to the study--experimental control was
actually demonstrated in the first two subjects by prolonged baselines and sequential application of treatment.

Q: Were you surprised that the normal population only gave you a mean of 65% performance, and that they had variability in their performance as well?
A: I really was surprised. But you know, we really don't know what patients are supposed to do when they go out into the real world because we don't always know what non-brain-damaged people do. How often do people ask when they meet someone "Where do you live?", "What is your name?" We wanted to know if that is something that normal people do. I thought that the normals would perform at a higher level than they did and I also didn't expect to see as much variability. The standard deviation being 17%, some people never produced the responses and some people always produced them. But, still the mean responding by the normals was higher than that of the aphasic subjects.

Q: Would you speculate on the lack of generalization on the questions as opposed to the greetings and self-disclosures?
A: My impression of that finding is related to the fact that aphasic patients often do not initiate conversation. It has been shown in the literature—at least descriptively by Ulatowska and others—that aphasic subjects often experience great difficulty in initiating verbal exchange. That is, they most often respond when asked to respond. If someone were to ask an aphasic patient, for example, "What is your name?", the patient may respond by providing his/her name, but then might not reciprocate by asking a question. This is also a common observation made by spouses of aphasic patients. That is, spouses, in describing their husband or wife, often say "He (or she) doesn't talk to me." My impression is that this phenomenon may be related to a pragmatic language disorder relative to turn taking and sharing the burden of communication. Right now we're in the process of further analyzing these data, descriptively, to look at the difference between the normal subjects' and the aphasic subjects' use of initiation and also whether the initiations were verbal or nonverbal. So far we're seeing a large difference between aphasic patients and normals and very little difference between verbal and nonverbal initiations. It appears that Broca's aphasic patients, at this level of functioning, initiate very little, if at all, in any manner.

Q: Do you think your normals would have done better with the social conventions had they been a little closer in age?
A: It's possible. That's a good point. Confederates for both normal and aphasic subjects were all the same in terms of age. It's perhaps true that the normals and the aphasic subjects would have responded differently to different confederates. This study only shows what persons from both populations did under the specified conditions. An interesting thing was that the normal subjects, knowing that the confederates were students, asked questions such as "What classes are you taking?" or "How far along are you in your program?" The aphasic patients performed very differently.

Q: I have a question about Subject 3. He seemed to fail again and again when you were presenting the probes. Did he exhibit signs of frustration?
A: No more frustration than the other subjects.

Q: If you were to compare his performance with the performance of a hypothetical fourth subject who had maybe an initial baseline session and a final
probe, do you think that the hypothetical Subject number 4 would have performed identically with Subject number 3? Do you think that the frustration level had a negative impact on that Subject's performance?

A: I can't tell you that. I don't know. My subjective impression is that this fourth subject would have performed the same. But, again, I don't know. Do you have some ideas?

Q: No, I just wonder how frustrating the task was.

A: It wasn't extremely frustrating. The subjects were in the room for a five-minute period of time. Most of the time was spent by the confederate talking. My guess is it was no more frustrating than talking with anybody else.

Q: Was there any subject that showed a training effect for the questions?

A: Yes, we got a training effect on all responses, questions and self-disclosures and greetings, for both of the subjects. What we did not see was a generalization effect for the question responses.

Q: When Stokes and Baer reviewed the generalization literature, they came up with several techniques that have been used to train generalization. I'm just wondering if maybe one of those techniques would have been more effective in training generalization, because we know that aphasic subjects do have difficulty initiating. Wouldn't it be a good idea to go back and try to train generalization using a couple of different techniques?

A: Absolutely. I agree with you completely. In fact that's my intention for further research. We've looked already at programming common stimuli as a technique to facilitate generalization. Aphasic subjects are also resistant to question production across settings using this method. I don't know at this point what methods might serve to facilitate generative use of question responses. This remains an empirical question. I think we need to further investigate what particular strategies do facilitate generalization of specific response types.

Q: I have a question about Subject number 3. Pragmatically, it's been shown that aphasic patients typically do know when something's being requested of them. Did you get any kind of responding, such as body language? What kind of behavior were you getting from the subject?

A: That's an interesting question. In general we were getting the same kinds of behavior from Subject 3 that we were getting from the other subjects. Right now we're in the process of analyzing those data with regard to gestural responses—whether or not the subjects were attempting to provide or to request information at all. Communicative competence vs. verbal communicative behavior is essentially what we're looking at. I don't think that there was any difference between the subjects' behavior in any remarkable way. In terms of responding to questions, all of the subjects attempted to supply information by pointing, drawing in the air, etc., but were only successful after training had begun. But they did not seem to use these nonverbal behaviors to request information.

Q: Did you find a difference in the greeting category in terms of initiation vs. responding?

A: We didn't analyze that but that would be interesting. We scored a greeting response as correct if the subject initiated the greeting or if it occurred within 10 seconds after the confederate produced a greeting. It would be interesting to find out if the subjects did, in fact, initiate some responses and not others.
APPENDIX A

Instructions to Confederates

You are going to meet someone. I want you to go in the room where a person will be waiting and talk with him or her for five minutes. During your five minutes, I want you to do the following things:

1) greet the person,
2) inquire as to how the person is,
3) attempt to find out the person's name,
4) attempt to find out where the person lives,
5) say goodbye.

Do NOT GIVE YOUR NAME OR WHERE YOU LIVE UNLESS ASKED.

Use the remainder of time talking about anything that you wish. You might talk about families, weather, hobbies, etc. At the end of the five minute period, someone will tap on the door. Do not answer the door. This is your cue to wrap up the conversation and leave the room.

Thank you for your participation.