come following stroke. In this study, aphasic patients were randomly assigned to speech therapists or untrained volunteers for 30 hours of individual treatment. The volunteers were provided with a detailed description of the patient’s communication problems, suggestions for the manner in which to interact with the patient, and instructed to "encourage the patient to communicate as well as possible". The speech therapists gave each patient such therapy as they thought appropriate for the 30 treatment hours. Response to treatment was measured with overall scores from the Functional Communication Profile (FCP) after 2, 4, 8, and 12 weeks of treatment and after cessation of treatment. Two pre-treatment baseline assessments using the FCP were taken one week apart. Results revealed that patients in both groups improved and that there were no significant differences between the groups at the pre-, within, and post-treatment intervals. Similar findings were obtained for both high and low level patients treated by speech therapists and volunteers.

Several important issues need to be considered before the results of this study can be used to support the employment of untrained volunteers in treatment of aphasic patients. Without the benefit of inter- and intra-examiner reliability data, the conclusion that aphasic patients will respond to treatment by untrained volunteers as well as to that provided by speech therapists, or vice versa, is intenable. In scoring the FCP, patients are rated as "normal", "good", "fair", "poor", and "none" on 45 communicative behaviours including broad performance categories, such as "reading letters", "understanding movies", "saying phrases" and "talking on the telephone". Determination of a rating requires familiarity with the patient and necessitates a subjective interpretation of the patient's abilities relative to his premorbid state. The examining clinicians in this study were not familiar with the patients they tested. David and her colleagues indicated that they were able to reduce inter-observer variation by providing suitable training, but do not tell us what this variation was. The fact that several therapists in different centres are involved in the administration of the FCP further heightens the need for reliability data.

A second major methodological problem that arises from the David study concerns the baseline pre-treatment measures. Baseline measures essentially imply steady state performance. Both groups of subjects reflected marked improvement from BL 1 to BL 2. Since the patients received no treatment during this time, this improvement must be attributed to other factors. Some of these might include (1) added spontaneous remission, (2) a placebo effect, or (3) increased patient familiarity with the test interview situation. Baseline measures are ordinarily employed in single case design research. In group studies baseline levels are essentially impossible to obtain inasmuch as some members of the group improve, some do not change and some may deteriorate during the baseline phase. Without the benefit of a stable baseline, treatment effects are difficult to ascertain.

Analysis of variance results indicated that both groups responded to treatment with significant changes in FCP scores. This improvement, however, was primarily restricted to baseline changes, and improvement from BL 2 to the first within-treatment measure. Beyond this point both groups improve minimally. This may result from the fact that 30 hours of treatment is minimal for an aphasic patient. The data provided in several large efficacy studies suggest that much more treatment is required to obtain positive changes.

Additional information is needed to determine the value of volunteers as surrogate clinicians. It is not stated whether a volunteer worked with one or several patients. If they worked with more than one patient, it is possible that some may have become more proficient therapists as the study progressed. It would be helpful to know if volunteers were screened before being assigned to patient care responsibilities. Information as to the educational levels and socioeconomic status of the volunteers would also be useful in determining their role as adjunct therapists. While the importance of significant others in the total rehabilitation of the aphasic person is vital, it is important to remember that volunteers in this study were directed by the speech therapists.

The authors state that the most essential finding of their study was that treatment had a positive effect on a group of late referrals as well as early referrals. This information is highly supportive of earlier investigations that have shown chronic aphasic patients to improve significantly following intensive speech and language therapy. The David et al study supports the contention that treatment, in any form, is beneficial to the aphasic patient. Information as to the effect of types of treatment, must come from further research. This project was approved by the Research and Development Committee of the Veterans Administration Medical Center, Portland, Oregon.

ROBERT C MARSHALL
LEE ANN C GOLPER
Department of Neurology,
Oregon Health Sciences University,
Veterans Administration Medical Center,
Portland, Oregon, 97201, USA

References

David replies:
We welcome Marshall and Golper’s comments on our paper. We did not include our own reliability data on the FCP in this paper because it is a well standardised test in its own right. However, for the reasons which Marshall and Golper cite we considered it important that the intercoder reliability should be checked. We undertook two studies: (1) Two therapists (RD and PE) together interviewed 15 aphasic
patients and then independently completed the profiles. Interobserver agreement was analysed using Pearson’s product moment correlation coefficient. There was significant agreement (p < 0.01) on all subsections and on the Overall scores (see table). (2) Fifteen therapists who were participating in the multicentre trial rated videotaped recordings of FCP interviews with four patients. The patients were selected to demonstrate a variety of types of aphasia and a range of severity. The testers were given an outline of the patients’ social and communicative background and photocopies of the reading material used and of their writing. The profiles were completed independently. Kendall’s Coefficient of Concordance was used to analyse agreement between the 15 observers and the results were converted to Spearman’s rank order coefficients (see table). All subsections and the Overall scores showed significant inter-rater agreement (p < 0.01). In addition, the standard error of the Overall scores obtained by the assessors indicated that approximately 95% fell within 4 points of the mean. The variation in subsection scores was slightly greater.

These findings support the high inter-rater agreement reported by the author of the test.¹ It was not possible for us to collect the assessors together again to examine test-retest reliability, but in view of the finding of high inter-rater reliability and the high test-retest correlation reported by Sarno¹ and Greenberg,² it was considered unlikely that our test-retest reliability would not also be high.

With respect to the change seen between the Baseline assessments, Fig 2 in the paper showed that this was mainly a feature of the High group, the Low group showing significant change only after the start of treatment. A future paper will contain more detailed analysis of the recovery curves of these groups and of other subgroups. However, as Marshall and Golper suggest, it is likely that more intensive treatment would produce more positive changes over a longer period in most patients. Unfortunately, the British speech therapy service is not usually able to provide many patients with more intensive treatment than was investigated in our study.

We are not able to answer Marshall and Golper’s request for more information on the patients. We did not collect information on their educational and socio-economic levels. They were not extensively screened before taking part, their own interest and apparent reliability being the main selection criteria. Most volunteers enjoyed the work and many saw several patients. They may indeed, as Marshall and Golper suggest, have become more proficient with increased experience but we would hope that the same can be said for speech therapists.

Table Results of three interscorer reliability trials of the FCP

<table>
<thead>
<tr>
<th>FCP subsections</th>
<th>Sarno 1965</th>
<th>1</th>
<th>2*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.95</td>
<td>0.97</td>
<td>0.93</td>
</tr>
<tr>
<td>Movement</td>
<td>0.95</td>
<td>0.87</td>
<td>0.67</td>
</tr>
<tr>
<td>Understanding</td>
<td>0.92</td>
<td>0.94</td>
<td>0.97</td>
</tr>
<tr>
<td>Reading</td>
<td>0.95</td>
<td>0.91</td>
<td>0.82</td>
</tr>
<tr>
<td>Other</td>
<td>0.87</td>
<td>0.90</td>
<td>0.84</td>
</tr>
</tbody>
</table>

1 Agreement between two observers rating 15 patients.
2 Agreement between 15 observers rating 4 patients.
All r values significant (p < 0.01).

The Functional Communication Profile (FCP)² can only complement but by no means replace a formal aphasia test. The FCP does assess language systematic abilities but only to a limited extent. It does not provide the speech therapist with the information necessary for planning a therapy programme aimed at the specific language deficit present in a given patient. Only a rigorous design of therapy can establish the difference between professional therapy by a speech therapist and unspecified stimulation by a volunteer. In view of the great variety of specific methods and techniques of aphasia therapy,³–¹¹ the multicentre approach is a serious disadvantage because uniformity of methods is not guaranteed. It could well be that some, if not the majority, of speech therapists applied rather nonspecific methods like auditory stimulation. In this case, they would have acted much like the volunteers. Volunteers, on the other hand, who see the patient in the speech therapy department and who get detailed information on FCP test results are likely to act like co-therapists.

In order to be effective, speech therapy must be given to aphasic patients at least three to four times a week over a period of six months and more (up to 12 months). To see a patient once a week is no better than no treatment.

No explanation is given for the finding that the recovery curves have a steep gradient only during the first two weeks of treatment although both types of treatment were continued over a period of 15 to 20 weeks. In the decisive two weeks the patients received at most four sessions of treatment. In other words, the recovery curves do not demonstrate that volunteer treatment is as good as professional treatment but rather that no continuous effect was obtained either way. This clearly shows the weaknesses of the type of professional speech therapy applied.

The great variation in performances (reported in table 4) would require the application of statistical methods for the

References


Sirs: The study by David et al¹² shows serious weaknesses in many aspects which invalidate the negative conclusions on the efficacy of speech therapy for aphasic patients. In both groups compared, there was a negative selection of subjects. Patient groups with a mean age of 65 to 70 years and a standard deviation of about 10 years can hardly be expected to show consider-