Materials arising

References


David replies:

We welcome Marshall and Golpher'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 Golpher cite we considered it important that the inter-scorer reliability should be checked. We undertook two studies: (1) Two therapists (RD and PE) together interviewed 15 aphasic patients and calculated inter-scorer reliability. (2) Baseline measures were calculated using the FCP. Each set was scored by two different therapists. Inter-scorer reliability between therapists was 0.92. This high degree of reliability is essential in a study where the nature of the interaction is so crucial.
patients and then independently completed the profiles. Interrater agreement was analysed using Pearson’s product moment correlation coefficient. There was significant agreement (p < 0.01) on all subsections and on the Overall score (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 therapists were given an outline of the patients’ social and communicative background and photographs 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 Sarno2 and Greenberg,3 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 volunteers. We did not collect information on their educational and socioeconomic 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.

References


Sir: The study by David et al1 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-