British Journal of Psychiatry (1989), **154**, 383–385

The Question of Therapists' Differential Effectiveness
A Sheffield Psychotherapy Project Addendum

DAVID A. SHAPIRO, JENNY FIRTH-COZENS and WILLIAM B. STILES

Further analysis of outcome data from the Sheffield Psychotherapy Project suggested that one of the principal therapists was responsible for most of the reported advantage of Prescriptive over Exploratory treatment.

The Sheffield Psychotherapy Project (Shapiro & Firth, 1987) compared the processes and outcomes of a cognitive–behavioural treatment called Prescriptive with those of an interpersonal–psychodynamic treatment called Exploratory. In a cross-over design, each patient was seen by the same therapist in eight weekly sessions of one treatment, followed by eight sessions of the other. This design was selected in order to maximise the study's sensitivity to differences between the two treatments, by holding constant individual differences among clients and therapists.

As previously reported (Shapiro & Firth, 1987), patients began treatment with clinical levels of symptomatology (mainly depression), made substantial gains during treatment, improved similarly regardless of the order in which they received the two treatments, and had maintained their gains at a three-month follow-up. On average, patients showed greater gains during the first eight-session period than during the second, and when this was statistically controlled showed a little more gain during Prescriptive than during Exploratory therapy.

Since the inception of our project, individual differences in therapists' effectiveness have become of substantive interest in their own right, rather than being dismissed as 'nuisance variables' to be eliminated from consideration by careful design (Lambert, 1989). For example, Luborsky *et al* (1985) found wide, statistically reliable differences in mean outcomes *within* groups of therapists practising each of three approaches: supportive–expressive psychotherapy, cognitive–behavioural therapy, and drug counselling.

Alongside such differences in outcome between therapists, evidence also exists of individual differences in therapists' contributions to the therapeutic process. In the Sheffield Psychotherapy Project, for example, the two principal therapists differed in the grammatical forms or literal meanings they used, irrespective of the type of treatment (Stiles *et al*, 1989).

It therefore becomes of interest to compare the effectiveness of different therapists in the two treatments in the Sheffield project. The 40 patients were seen by four therapists. Therapists 1 and 2 (this article's first and second authors) saw 18 and 16 patients respectively, while therapists 3 and 4 saw four and two patients respectively. This paper compares these therapists with respect to their patients' improvement in each of the two treatments, with particular attention to therapists 1 and 2, who each saw sufficient patients to enable attribution of any difference to the therapists rather than to the patients.

Method

The design and implementation of the Sheffield project are described in detail elsewhere (Shapiro & Firth, 1987). Each patient underwent an intake assessment including the Present State Examination (PSE; Wing *et al*, 1974), the Social Adjustment Scale (SAS; Weissman & Paykel, 1974), the Beck Depression Inventory (BDI; Beck *et al*, 1961), the Symptom Checklist-90 (SCH; Derogatis *et al*, 1973), and the O'Malley & Bachman (1979) Self-Esteem Scale (SE). Patients were assigned to therapists in accordance with scheduling and workload constraints, and as far as practicable we balanced the numbers of each sex seen by each therapist. Within each therapist's caseload, patients were randomly assigned to receive eight weekly sessions of either Exploratory or Prescriptive therapy. This randomisation was accomplished by the toss of a coin for each alternate patient of a given sex assigned to a given therapist, with the next patient of that sex seen by that therapist automatically assigned to the other treatment. After a second administration of the assessment battery, patients received the other treatment with the same therapist for a further eight weekly sessions, followed by third and fourth administrations of the assessment battery at termination and after a further three months. Therapists remained blind to each patient's assessment data until after the completion of treatment. As previously, we used a total symptom score from the PSE and a grand mean of all items on the SAS for statistical comparisons.

Patients were referred to a research clinic by general practitioners and psychiatrists. Inclusion criteria specified a total symptom score of at least 14 and an absence of

TABLE I
Residual gains during Prescriptive and Exploratory therapies[1]

Therapist and treatment	*PSE*[2]	*BDI*[3]	*SCL-90*[4]	*SAS*[5]	*SE*[6]
Therapist 1 (18 patients)					
Prescriptive	0.52	2.31	16.46	0.07	2.40
Exploratory	−2.62	−2.02	−12.51	−0.08	−3.64
t	1.70	2.27*	2.86*	2.68*	2.48*
Therapist 2 (16 patients)					
Prescriptive	0.43	0.71	7.56	0.03	−0.51
Exploratory	−0.12	−1.54	−16.84	0.03	−0.28
t	<1	1.15	2.50*	<1	<1
Therapists 3 & 4 (6 patients)					
Prescriptive	2.71	−0.16	3.25	0.09	−1.11
Exploratory	3.46	1.66	6.59	0.08	−0.28
t	<1	<1	<1	<1	<1

1. Higher scores indicate greater improvement. Some *n*s were slightly lower than the number of patients because of missing data on some measures.
2. Present State Examination total symptom score.
3. Beck Depression Inventory.
4. Symptom Checklist-90.
5. Social Adjustment Scale, overall mean.
6. O'Malley & Bachman Self-Esteem Scale total.
*$P<0.05$.

obsessional or psychotic symptoms on the PSE, a continuous history of psychological disorder not greater than two years, no significant recent change in psychotropic medication, current employment in a professional or managerial job, and a complaint that work was affected by psychological problems. There were 23 males and 17 females; their mean age was 40.7 years (range 27–62 years). Thirty patients were diagnosed depressed using the PSE–1D–Catego system of Wing *et al* (1974); the remainder had mainly anxiety disorders.

Therapists were clinical psychologists who had had prior training in both relationship-orientated and cognitive–behavioural methods and maintained confidence in the efficacy of both approaches. Therapists' adherence to treatment manuals and general treatment issues were addressed in weekly peer supervision.

The Prescriptive treatment, designed to represent cognitive and multimodal behavioural approaches, involved therapists selecting appropriate techniques from four areas: anxiety-control training, self-management procedures, cognitive restructuring, and a job-strain package. The Exploratory treatment, chosen to represent psychodynamic and interpersonal approaches, was based on Hobson's (1985) Conversational Model. Verbal response mode coding of these treatments confirmed a pattern of adherence to treatment manuals and large verbal process differences between treatments (Hardy & Shapiro, 1985; Stiles *et al*, 1988).

Results

As previously (Shapiro & Firth, 1987), we used residual gain scores (Cronbach & Furby, 1970; Mintz *et al*, 1979) to assess improvement across each period of treatment. For each measure within each period of treatment, we calculated the deviation of each patient's ending score from the regression line of the beginning score upon the ending score. In other words, we controlled statistically for differences in symptomatology between patients at the start of each eight-session period before comparing their symptoms at the end of that period. Table I shows the mean residual gains (i.e. mean improvement scores adjusted to take account of initial symptom levels) achieved in each treatment (regardless of whether it was administered first or second) for therapist 1, therapist 2, and therapists 3 and 4 combined. The significance of each treatment difference was evaluated by *t*-test for related samples (also shown in Table I).

We have already reported (Shapiro & Firth, 1987) the results for therapists combined. These suggested a slight advantage for Prescriptive therapy, which produced numerically larger gains on all measures, significant only on the SCL-90 (mean residual gains of 11.09 and −11.09 for Prescriptive and Exploratory therapy respectively, $P<0.05$) and marginally significant on the BDI and SE ($P<0.10$ in each case). However, examination of Table I shows that only therapist 1's clients showed clearly differential results, with significantly greater gains in Prescriptive on four of the five measures and a trend in the same direction on the fifth. Among the other therapists' results, only therapist 2's significant difference on the SCL-90 supported the inference that Prescriptive was superior: most of the numerical differences were negligible or reversed.

For all therapists the overall effectiveness (i.e. mean patient change across all 16 sessions) was approximately the same. There were no significant effects on any of the five measures (a) in *t*-tests comparing therapists 1 and 2 on residual gain scores from intake to termination and from intake to follow-up, or (b) in one-way analysis of variance comparing all therapists on these scores, except that residual gains on the PSE at follow-up by the patients of therapists 3 and 4

averaged slightly larger (5.99) than those by patients of therapist 1 (−2.42) or therapist 2 (0.85) ($F(2, 36) = 3.61$, $P < 0.05$).

Given that assignment to therapists was based on availability rather than being formally randomised, we checked the comparability of symptoms at intake of clients assigned to each therapist and to each treatment order. F ratios were less than 1 for all tests of main effects due to therapist and all tests of the interaction between therapist and treatment order for all five outcome variables, except for therapist differences on the BDI (F (2, 34) = 1.09, $P = 0.35$).

Discussion

The overall slight advantage of Prescriptive over Exploratory treatment appears mainly attributable to the differential effectiveness of therapist 1 in the former. Therapist 1 was not more effective than the other therapists overall, as his slightly greater effectiveness in Prescriptive was balanced by a slightly lesser effectiveness in Exploratory. (However, direct comparisons of therapists' effectiveness in each treatment were not significant, owing to small differences and small numbers of observations.) In addition, it is noteworthy that the one measure to show an advantage of Prescriptive therapy for both therapists 1 and 2 was the highly symptom-orientated SCL-90, which might be judged least favourable to Exploratory treatment.

Although patients were not assigned randomly to therapists, patients assigned to each therapist were very similar in symptomatology at the outset; the therapists rather than the patients were therefore the most likely source of the findings obtained.

Finding strong evidence of treatment differences for one therapist and weak or absent evidence for others underlines the importance of individual differences in therapeutic effectiveness, even in successfully manualised treatments. It lends added significance to detailed study of individual differences in therapists' in-session behaviour (Stiles *et al*, 1988) that may help to explain the differences in effectiveness. The simultaneous failure to find overall differences in therapist effectiveness across treatments suggests that, beyond general clinical competence, individual therapist characteristics may have a different impact depending upon which theoretical approach is employed.

This first Sheffield Psychotherapy Project was not designed as a study of therapist effects, but rather with a view to controlling these. With hindsight, the differential effectiveness revealed by the present analysis testifies to the failure of our attempt to design therapist effects out of the study. In response to the findings presented here, the factorial design of the Second Sheffield Psychotherapy Project (Shapiro *et al*, unpublished) addresses systematically the important question of therapist effects. Each of five therapists is randomly assigned 24 patients, 12 in Prescriptive and 12 in Exploratory therapy. The design and sample sizes will enable thorough comparisons between treatment and therapist effects on both process and outcome.

Meanwhile, the present findings are broadly consistent with clinical lore that each new therapist should try different approaches to find the one in which he or she is most effective.

Acknowledgements

We gratefully acknowledge clinical supervision by Robert F. Hobson and Ian Burnside, participation as therapists by Gillian Hardy and Glenys Parry, interviewing and data processing by Lisa Davies, Jan Jackson, Ann Macaskill, Brenda McWilliams, and Leslie Morrison, administration of the clinic by Mary Lou Hughes, training in the Present State Examination by Paul Bebbington and Traolach Brugha, and help with the Social Adjustment Scale by Eugene Paykel.

References

BECK, A. T., WARD, C. H., MENDELSON, M., *et al* (1961) An inventory for measuring depression. *Archives of General Psychiatry*, **4**, 561–571.

CRONBACH, L. J. & FURBY, L. (1970) How we should measure 'change' – or should we? *Psychological Bulletin*, **74**, 68–80.

DEROGATIS, L. R., LIPMAN, R. S. & COVI, M. D. (1973) SCL-90, an outpatient rating scale: preliminary report. *Psychopharmacology Bulletin*, **9**, 13–20.

HARDY, G. E. & SHAPIRO, D. A. (1985) Therapist verbal response modes in prescriptive vs. exploratory psychotherapy. *British Journal of Clinical Psychology*, **24**, 235–245.

HOBSON, R. F. (1985) *Forms of Feeling: the Heart of Psychotherapy*. London: Tavistock.

LAMBERT, M. J. (1989) The individual therapist's contribution to psychotherapy outcome. *Clinical Psychology Review*.

LUBORSKY, L., MCLELLAN, A. T., WOODY, G. E., *et al* (1985) Therapist success and its determinants. *Archives of General Psychiatry*, **42**, 602–611.

——, LUBORSKY, L. & CRISTOPH, P. (1979) Measuring the outcomes of psychotherapy: findings of the Penn Psychotherapy Project. *Journal of Consulting and Clinical Psychology*, **47**, 319–334.

O'MALLEY, P. M. & BACHMAN, J. G. (1979) Self-esteem and education: sex and cohort comparisons among high school seniors. *Journal of Personality and Social Psychology*, **37**, 1153–1159.

SHAPIRO, D. A. & FIRTH, J. (1987) Prescriptive vs. Exploratory psychotherapy: outcomes of the Sheffield Psychotherapy Project. *British Journal of Psychiatry*, **151**, 790–799.

STILES, W. B., SHAPIRO, D. A. & FIRTH-COZENS, J. (1988) Verbal response mode use in contrasting psychotherapies: a within-subjects comparison. *Journal of Consulting and Clinical Psychology*, **56**, 727–733.

——, —— & —— (1989) Therapist differences in the use of verbal response mode forms and intents. *Psychotherapy* (in press).

WEISSMAN, M. M. & PAYKEL, E. S. (1974) *The Depressed Woman*. Chicago: University of Chicago Press.

WING, J. K., COOPER, J. E. & SARTORIUS, N. (1974) *The Measurement and Classification of Psychiatric Symptoms*. Cambridge: Cambridge University Press.

*David A Shapiro, BA, MSc, PhD, *Senior Scientist, Medical Research Council/Economic and Social Research Council Social and Applied Psychology Unit, Department of Psychology, University of Sheffield, Sheffield S10 2TN*; Jenny Firth-Cozens, BA, MSc, *Lecturer, Department of Psychology, University of Leicester*; William B. Stiles, PhD, *Professor, Department of Psychology, Miami University*