european psychotherapy

Rolf Sandell,[1] Jan Carlsson,[2] Johan Schubert,[2] Jeanette Broberg,[2]
Anna Lazar,[1,2] and Johan Blomberg[1]

Varieties of Therapeutic Experience and their Associations with Patient Outcome

Abstract

Therapist experience has been reported to be only weakly related to patient outcome in psychotherapy. The purpose of the present study was to explore the relations between patient outcome in treatment and therapists' level of experience. More than 400 patients in long-term psychotherapy provided the long-term outcome data, and these were linked to the level of experience of their therapists, numbering more than 200. Seven indicators of different varieties of therapeutic experience yielded mixed results, which was expected in consideration of the fact that therapeutic experience is not a unitary phenomenon. Accordingly, some indicators (such as number of years of post-license practice) were significantly and positively associated with patient outcome, some (such as duration of personal therapy) significantly and negatively related to outcome, and some (e.g., number of years in pre-license practice) only weakly or not at all. Alternative interpretations of the findings are discussed.

Introduction

Several research reviews on the value of experience of psychotherapists have concluded that the therapist's professional experience is at most weakly related to the treatment outcomes of his or her patients (Auerbach & Johnson, 1977; Berman & Norton, 1985; Beutler, Machado & Neufeldt, 1994; Hattie, Sharpley & Rogers, 1984; and for children and adolescents, Weisz, Weiss, Alicke & Klotz, 1987).

Stein and Lambert (1984, 1995) appear to have treated the issue in the most systematic and sophisticated manner, in two meta-analyses. In their first meta-analysis, they found the avera-

[1] Linköping University
[2] Stockholm County Council Institute of Psychotherapy

ge effect size equal to 0, that is, no difference between more and less experienced therapists. Although there were more studies with positive than negative results, about half of the studies did not show any difference whatsoever. An obvious but important conclusion of the authors was that experience effects were more evident where the variation in the therapists' experience was larger. This points to limitations in research on experience in relatively range-restricted groups of therapists: when therapists' levels of experience do not differ, outcome differences cannot depend on level of experience. Another important conclusion was that effects of experience were also more evident in more complicated forms of psychotherapy (in comparison with, for instance, simple forms of counseling or the routine application of specific behavioural techniques). This observation may be linked to a conclusion by Strupp and Hadley (1979), that low experience is a drawback especially in the therapists' work with long-standing and long-range problems.

In their second meta-analysis, Stein and Lambert (1995) reported somewhat more positive findings. They now focussed on training and concluded that "a variety of outcome sources are associated with modest effect sizes favoring more trained therapists" (p. 182). "Modest effect sizes" referred to d values between 0.2 and 0.3, but these are conventionally considered small effects (Cohen, 1988). The association with training was stronger for various measures of patient satisfaction and pre-post differences on psychological tests. An interesting observation, indirectly supporting one of the conclusions in Stein and Lambert's first review, was that less experienced therapists tended to have more dropouts among their patients. This was particularly obvious in longer and more complex forms of psychotherapy; the authors suggest so-called insight-oriented family therapy (in contrast to acute or crisis-oriented family counseling) as an example. Unless intention-to-treat is considered, by ignoring the drop-outs in estimating mean outcome, such dropout may make the less experienced therapists' cases appear more successful than they actually are.

These results are probably counter-intuitive to most of us and have not gone unchallenged. Several researchers (e.g., Rönnestad, Orlinsky, Parks, Davis et al., 1997, and Stein & Lambert, 1984, themselves) have noted that the null findings might have been due to the fact that the allegedly experienced therapists have indeed not been very experienced, if so there was in fact too little variation in therapist experience level to make much difference in patient outcome, a mechanism called restriction of range. Another reason for the unexpected findings may be that experience has been taken as a unitary phenomenon, which it probably is not. Stein and Lambert commented, in 1995 as well as 1984, that researchers usually had confounded training level, training quality, duration of clinical practice, case-load, and age, and suggested that difficulties in the definition of therapeutic experience had hampered research integration in the area. In general, however, therapeutic experience has been operationalized as number of years in practice or in terms of level of training.

It is likely that experience is somewhat more complicated than that. Sandell et al. (2001) applied a manual cluster analysis procedure to a set of experience-related variables and found four clusters of variables in a national random sample of Swedish therapists. Besides a central variety of experience, which they called seniority, three more specific clusters were suggested. Psychodynamic repute had psychoanalytic training and long personal therapy as its principal indicators (linked to seniority by way of long private practice). A third experience cluster was long and faithful service, indicated by age and number of pre-license years (and long experience of psychotherapy in psychiatric practice, inpatient or outpatient). A fourth variety was associated with a large accumulated caseload, particularly linked to cognitive-behavioural training and practice. The authors suggested that these factors might be related to success with patients in different ways, some positively, some not at all, and some even negatively.

The purpose of this study was to explore these relations between patient outcome and various parameters of therapist experience on the basis of data from the Stockholm Outcome of Psychotherapy and Psychoanalysis Project (STOPPP). The STOPPP consists of a number of data collections on patients in long-term psychotherapy and psychoanalysis and their therapists, as well as on various comparison samples.

Method

Design and procedure

The design was quasi-experimental, partly cross-sectional, partly longitudinal, based on a three-wave panel survey under "Caucus-race" conditions[1] (Carroll, 1865). This means that the panel members' treatment status in each panel wave was uncontrolled, some patients being in treatment, some waiting to start, and some having already terminated. In analyzing the observations in the panel, we unfolded, so-to-speak, the panel along a time scale, distributing the panel members on the time scale in accordance with their treatment status each wave, thus creating what we call an "unfolded panel design" or what Bell (1953) has referred to as an "accelerated longitudinal design." As we have noticed that our design is not easily understood, we shall

[1] During her adventures in Wonderland Alice had wept violently, creating a pool of tears, crowded with birds and animals that had fallen into it. The Dodo then suggested a Caucus-race to get them dry and declared that the best way to explain it was to do it: "First it marked out a race-course, in a sort of circle (the exact shape doesn't matter, it said), and then all the party were placed along the course, here and there. There was no 'one, two, three, and away,' but they began running when they liked, and left off when they liked, so that it was not easy to know when the race was over. However, when they had been running half an hour or so, and were quite dry again, the Dodo suddenly called out 'The race is over!' and they all crowded round it, panting, and asking 'But who has won?'" (Caroll, 1865/1982, p. 33). The Dodo's answer is familiar to every psychotherapy researcher.

describe it in terms of the following six steps (see also the account by Sandell et al., 2000, 2001).

(1) A sample of 756 persons was selected so as to ensure that it consisted of people who had terminated their treatments as well as people who were in the midst of it and people who had not yet started. Thus, we selected all persons who had been subsidized for the periods 1991-93 and 1992-94. These were 202, some having terminated and some being still in treatment. Additionally, we selected the first 554 persons on the waiting list for subsidization, because we anticipated that a number of these were not yet in treatment.

(2) A questionnaire was distributed to these 756 persons in 1994, and in 1995 and 1996 to all who had responded the previous year, each year with four reminders. After returns from 78%, 86% and 89%, respectively, for each year, this produced a panel of 446 persons, which was 59% of the initial sample of 756. An analysis of the attrition showed that patients with higher educational level and higher current level of functioning tended to respond significantly more often. However, the pattern of attrition did not differ between patients in different modalities of treatment.

(3) The "unfolding of the panel" is visualized in Figure 1. The 446 persons were divided in seven subgroups on the basis of their treatment status each panel wave. Thus, we had one group of 12 persons who had not commenced treatment in any of the three waves. This group was henceforth discarded altogether and is not shown in Figure 1. The remaining groups are ordered vertically in the left-most column of the figure. Thus, among the remaining 434 persons, we had one group of 20 who had not started treatment in the first panel wave 1994 or in the second 1995 but were in treatment in the third wave 1996. We had another group of 31 persons who had not started treatment in 1994 but had so in 1995 and were still in treatment 1996. Following the same reasoning we had four other groups in later phases of the treatment process, up to, and including, a group of 156 patients who had already finished their treatment when they were in the first panel wave. For natural reasons, we considered this last group as having reached a later stage of treatment than the rest of the groups, and similarly considered the first-mentioned group as being at an earlier stage of treatment than the others. Correspondingly, we could order the six subgroups along an ordinal or relative time scale, defined by the relations "before" and "after" or "earlier than" and "later than" such that each position on the scale is later than or after all positions to the left and earlier than or before all positions to the right.

(4) We then aligned the three panel waves in each of these subgroups, such that we assumed, for example, that the last wave before treatment was at the same time, relative to treatment, whether it was the first or second panel wave, and that the first wave in treatment was

at the same time, relative to treatment, whether it was the first, second, or third wave, and correspondingly through all waves in all groups. So, we were able to produce a sub-grouping of all (434 x 3 =) 1302 observations. This grouping is shown in the lowest row in Figure 1. Thus, for example, the 20 patients in group 1 in the third panel wave, the 31 in group 2 in their second panel wave, and the 151 in group 3 in their first panel wave were all in their first year of treatment and were therefore grouped together in the "early ongoing" treatment group, which thus consisted of 20 + 31 + 151 = 202 patients or observations. Thus, by pooling

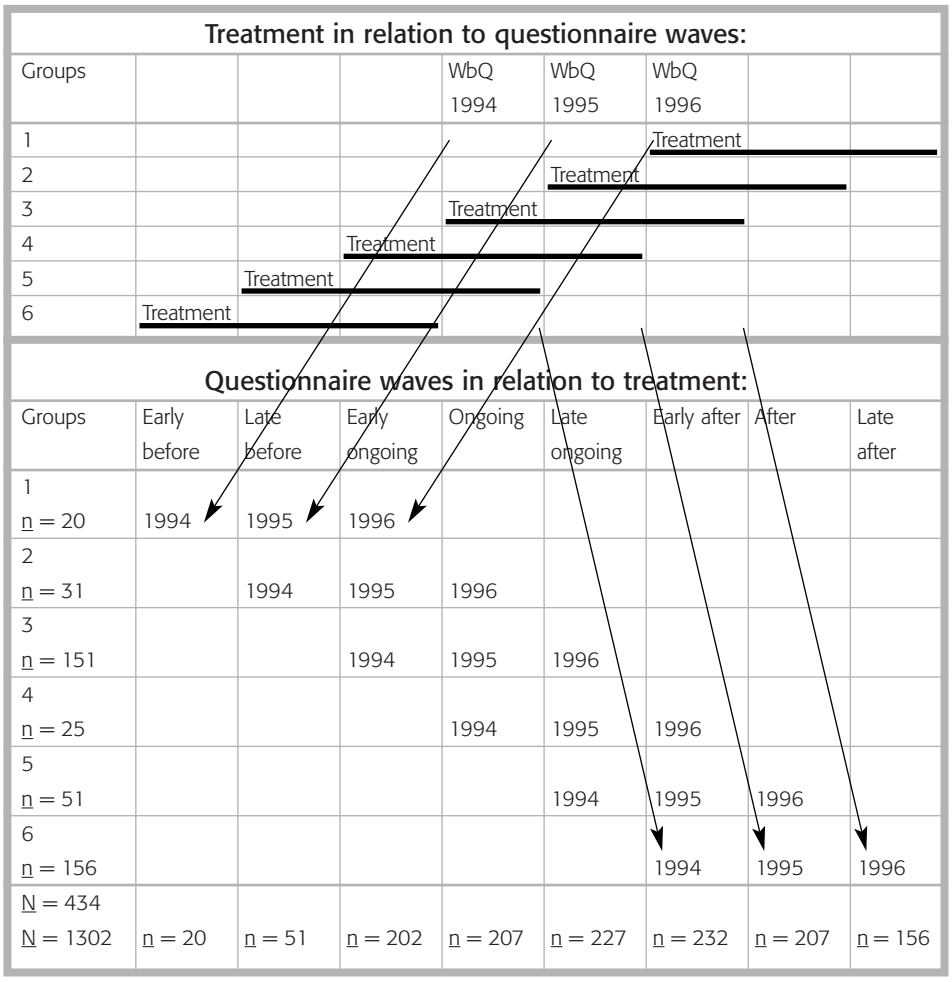

Figure 1. Unfolding the three-wave panel. Groups of cases at different stages of treatment, in relation to waves of administration of the Well-being Questionnaire (WbQ) (upper panel), and in analysis design (lower panel).

observations from different groups in different waves who were in the same relative phase of treatment, that is, relative to earlier and later phases, we had the observations distributed along eight steps on a relative time scale. The variation across time was partly within-subjects and partly between-subjects. Scattered missing data reduced the effective number of observations to between 1281 (for the SOCS) and 1287 (for the SCL-90 and the SAS), spread out across the eight-step time scale, from before (about two years) treatment to late (about three years) after treatment termination.
Having tested the correlations of our time scale with more than 30 different variables (patient and therapist characteristics), we have concluded that the distribution of the observations along the time scale is independent of obvious confounds.

(5) In order to compare different types of treatments, the observations in each step of the time scale were quasi-experimentally split in subgroups. Thus, of the 434 persons in the panel, 345 had long-term psychotherapy as their treatment or – in case they had been in more than one – their main treatment, in terms of number of sessions. They generated 1035 (3 waves x 345 persons) observations spread over different phases of psychotherapy. Correspondingly, we had 76 persons with psychoanalysis as their treatment (thus 228 observations) and 13 persons with various kinds of so-called low-dose treatments (39 observations in low-frequency individual therapy, group therapy, family therapy, etc.). When the treatment modalities, specifically, were being compared, we excluded the small low-dose treatment group.

(6) Correspondingly, in order to compare treatments with therapists with different levels of experience, the observations in each step of the time scale were divided in subgroups based on the therapists' level of experience according to different indicators. This required a linking of data from the patients with data from their respective therapists, so missing data from the therapists were naturally added to non-response from the patients, which naturally resulted in a smaller number of observations than the 1302 observations available from the patients alone. The results of the linking will be further detailed in the Treatment section, below.

Assessment Procedures

Patients' pretreatment status. Various diagnostic and assessment procedures were applied to the patients' referrals. These pre-treatment assessments will not be specified here (but see Blomberg, Lazar & Sandell, 2001).

Patient outcome measures. The Well-being Questionnaire (henceforth the WbQ), was designed to explore the patients' symptoms, social relations, and morale. The following standard self-rating scales were included:

- The Symptom Check List (SCL-90; Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974) contains 90 items representing various psychological and somatic signs of distress. The task is for the

patient to rate the extent to which he or she has been troubled with each during the last seven days. The scales are five steps, from 0 (not at all) to 4 (very much). The ratings are scored in various combinations. In this study we used the General Symptom Index (GSI), which is calculated as the mean rating across all 90 items. Reliability estimates in the three waves varied between .83 and .96.

- The Sense of Coherence Scale (SOCS; Antonovsky, 1987) is a 29-item self-rating instrument, designed to measure sense of coherence, the feeling of confidence that life is manageable, comprehensible, and meaningful. The items are questions or phrases about life experiences, and the rating scales are seven-step bipolar scales, the poles of which are contrasting, alternative responses to the item. Following Antonovsky's recommendations, a general score was calculated as the mean across all items. Reliability estimates in the three waves varied between .81 and .92.
- The Social Adjustment Scale (SAS; Weissman & Bothwell, 1976; Weissman, Prusoff, Thompson, Harding & Myers, 1978) contains 39 items, divided in six sections (Work, Friends and leisure time, Extended family, Partner, Children, Family [partner and children]). The original items were translated and revised to suit Swedish users in the 90's. The task is to rate to what extent, during the last two weeks, various kinds of contacts have been satisfying or unsatisfying. One of the important revisions was to introduce the same five-point rating scale for all items, ranging from every or almost every day (or every or almost every time) to not one day (or never/not once). A general score was computed as the mean across all items. Reliability estimates in the three waves varied between .74 and .80.

Besides these instruments, the WbQ contained the following sections, with standard items/questions on (a) demography and socio-economy, familial, vocational, and financial situation; (b) ongoing psychotherapy; (c) previous treatments, including psychotherapy, for psychological distress; (d) current health status and health care utilization past 12 months; (e) current and prior severity of psychological problems; (f) occupational activities (including studies) past 12 months.

Therapists

There were 294 therapists and analysts involved in treatment with the patients in this sample. These were all licensed by the National Board of Health and Social Welfare. Some were fully trained psychoanalyst, members of any of the two psychoanalytic societies in Sweden. In the fall 1995 a postal questionnaire, Therapeutic Identity (THID) was distributed to all. After four reminders, 209 (71%) had returned their questionnaires. Analyses of the attrition showed no systematic sources of dropout. This sample will henceforth be called the treatment provider sample (or provider sample, for short).

The THID has about 150 questions and/or items, divided in six sections, (a) demographics, academic and professional training (age; gender; graduate education [M.D., psychologist, social worker, etc.]; psychotherapeutic training for licensing purposes; auxiliary psychotherapeutic training more than one year; formal supervisory training; academic training or professional training outside psychology or psychotherapy; membership in professional associations), (b) professional experience (duration of psychotherapy practice before and after licensing; psychotherapy practice in different contexts [outpatient or inpatient psychiatry, private practice, etc.]; accumulated case-load in different categorizations [types of therapy, durations and frequencies, age and diagnostic groups of patients, etc.]; supervision taken and/or given past 12 months, (c) personal therapy or training analysis (rounds; kinds; frequencies; durations). The items were designed specifically for the project or adopted from the questionnaire used in the Society for Psychotherapy Research (SPR) network (Orlinsky, Ambühl et al., 1999). The most important information culled from these sections is summarized in Table 1. Section (d) is a set of six scales to rate one's allegiance to each of some major schools of psychotherapy. Sections (e) to (f) of the THID are three sets of items to chart the therapists therapeutical orientation (Blomberg & Sandell, in press). Items from sections (d) to (f) were not used in this study.

The THID had been standardized on a random sample of 325 licensed psychotherapists throughout Sweden, of which 227 had responded (70%) (Sandell et al., 2002). This sample will be called the national sample. There was significant non-response from therapists in the higher age categories who claimed that they had retired from work.

The treatments

In the referrals psychotherapy was defined as once- or twice-a-week treatment with a licensed psychotherapist, and psychoanalysis as three- to -five-times-a-week treatment with a fully trained psychoanalyst. Of the 434 cases with complete data from the patients, there were 337 with data from the treatment provider as well. Of these 274 were psychotherapy cases, 55 psychoanalysis cases, and 8 cases in low-dose therapies, spread across the nine-step time scale, yielding a total of 1011 observations of patient outcomes. (Due to missing data in the therapists' questionnaires the following analyses had to be based on smaller numbers of observations, different with different experience indicators.)

The treatments were not manualized or standardized with respect to duration, session frequency, technique etc. Without a protocol, further specification of the treatments has to be ex post facto, in terms of provider characteristics, on the basis of information in the THID. Thus, all therapists in the provider sample claimed to have a psychoanalytic or psychodynamic theoretical orientation. The majority (75%) were women and their mean age was 51.4 years. Mean number of years in the psychotherapeutic profession after licensing was 9.6 years and before licensing 10.7 years. Further details on the treatment providers are given in Table 1.

Patient Characteristics

The typical patient was a woman, single, divorced, or unmarried, with children. The majority (87%) had at least some university education and typically worked in the health-care-, education-, or social sector. The mean age was 36.4 (SD = 8.1).

When patients in psychotherapy and psychoanalysis were compared, there were relatively more men in the psychoanalysis group. Analysands were also somewhat older, had higher education and were more often married or divorced. There were no differences between the two groups with respect to DSM-diagnoses. However, in terms of previous experience of psychiatric treatments, more of the analysands had prior psychotherapeutic experiences, whereas psychotherapy patients had more often been hospitalized. Further details on the patient sample are given by Blomberg et al., (2001).

Norm Groups

To establish a standard for evaluating patient outcome in relation to "normality," the WbQ was also distributed in two non-clinical groups, (a) a random community sample of 400 persons between 20 and 69 years of age in Stockholm County; (b) a sample of 250 psychology students, demographically very similar to the clinical sample, according to pilot analyses of the referrals. The norm groups took the questionnaire only once, in May 1994. Without any reminders, the response rates in the two groups were 37% and 79%, respectively. The responders in the two groups had almost identical mean values on the self-rating scales, and they were therefore collapsed into one group.

Results

The therapist variables were selected on the basis of the cluster analysis of Sandell et al. (2002), which involved 15 different experience indicators. The cluster solution identified four clusters. For this study we have selected one pair of variables in each cluster that correlated more strongly with each other than with any other variables in the entire variable set. The correlations among these variables and their statistics in the national sample and the provider sample are given in Table 1. Whereas there was a natural division in subgroups on some of the experience variables (e.g., supervision training and psychoanalytic training have two natural categories, training and no training), subgroups on other experience variables were formed by splitting the distributions in three groups of cases as equal in size as possible, using the SPSS Categorize variables routine.

Table 1: **Distributions (%, unless otherwise specified) on Selected Background and Experience-Related Variables among Therapists in a Random National Swedish Sample (N = 227)**

Gender	
female	68
male	32
Age, in years	
– 44	14
45-49	23
50-54	39
55-60	16
61+	8
M (SD)	51.4 (6.5)
Basic academic training	
MD	11
psychologist	62
social worker	16
other	11
Psychotherapy training taken as a basis for licensing	
psychoanalytic	5
psychodynamic therapy, university training	19
psychodynamic therapy, extra-university training sites	41
child and adolescent psychotherapy	9
behavioural/cognitive-behavioural therapy	2
cognitive therapy	2
family therapy	12
group therapy	2
unspecified	8
Auxiliary psychotherapy-related training	
(›1 yr course(s) in the psychology or psychotherapy area)	55
Formal training as supervisor	43
Academic training or professional training beside the above	28
No. yrs in psychotherapy practice	
before licensing; *M (SD)*	10.7 (6.5)
after licensing; *M (SD)*	6.6 (4.2)
No. yrs doing psychotherapy in	
outpatient psychiatric practice; *M (SD)*	8.9 (7.8)
inpatient psychiatric practice; *M (SD)*	2.6 (5.1)
private practice; *M (SD)*	5.5 (6.5)

Accumulated no. patients in individual psychotherapy ("case-load")	
1-9	
10-24	7
25-49	15
50-99	21
100-199	25
200+	21
M (SD)	82.9 (125.8)
Been in supervision last 12 mos.	
regularly	54
occasionally, as needed	27
not at all	19
Has supervised colleagues last 12 mos.	
regularly	53
occasionally, on request	27
not at all	20
Number of personal therapies; *M (SD)*	2.3 (1.0)
Total duration (in years) of personal therapy; *M (SD)*	7.9 (3.6)
Total number of sessions of personal therapy (dose); *M (SD)*	566.7 (454.7)
Main kind of training therapy (in terms of duration)	
psychoanalysis	21
individual psychotherapy, generally psychodynamic	61
group therapy	11
behavioural/CBT/cognitive	1
unspecified	6

Note. **Boldface** indicates that the variable has been analyzed as a moderator variable in this study..

In the following analyses only the GSI, the mean ratings across all items on the SCL-90, was used, as this proved to be the most sensitive of the instruments used. We have analyzed how the patients developed from before treatment until after, in the following way. To obtain estimates of change rate (how much patients changed each stage, at an average, during and after treatment) for patients within each therapist subgroup, the GSI was related to the relative time scale in so-called linear regression analyses. The mean change rate in each therapist subgroup was then estimated by the linear trend (unstandardised b). The regression procedure also produced an estimate of mean baseline or initial level (before treatment), the so-called intercept (a). For the sake of convenience we have assumed that our time scale was equidistant (that each step on the scale was equally long) and that patients' change was linear (that they changed an equal amount each step).

Table 2: **SCL-90 GSI as a function of time in treatment (intercepts: a; slopes: b), in groups of cases with varying therapist experience**

	All treatments		Psychotherapy		Psychoanalysis	
Therapist experience	a	b	a	b	a	b
Post-license experience, years						
3-8	0.97	-0.016	0.99	-0.017	0.33	+0.066
9-10	1.07	-0.082***	1.11	-0.087***	0.88	-0.075(*)
11-35	1.08	-0.104***	1.07	-0.102***	1.15	-0.128***
Supervisor training						
yes	1.11	-0.085***	1.12	-0.081***	1.06	-0.109***
no	0.94	-0.039*	0.97	-0.046*	0.68	-0.045
Psychoanalytic training						
yes	0.99	-0.069***	0.98	-0.048	1.04	-0.104***
no	1.06	-0.066***	1.07	-0.070***	--	--
Personal therapy, no. sessions						
125-720	1.19	-0.105***	1.23	-0.115***	0.60	-0.058
725-1240	1.01	-0.061***	1.03	-0.063**	0.99	-0.094
1250-2500	0.90	-0.030	0.80	+0.007	1.08	-0.109***
Age, years						
41-50	1.08	-0.039	1.10	-0.038	1.04	-0.079
51-56	1.02	-0.073***	1.04	-0.074***	0.89	-0.067
57-80	1.05	-0.096***	1.05	-0.096***	1.14	-0.156***
Pre-license experience, years						
0-8	0.96	-0.042(*)	0.98	-0.039	0.99	-0.107**
9-11	1.19	-0.076***	1.20	-0.079***	1.21	-0.121
11.5-25	1.00	-0.087***	1.02	-0.090***	0.89	-0.071
Accumulated number of patients in individual therapy						
1-49	0.91	-0.035	0.89	-0.022	1.05	-0.100*
50-109	1.05	-0.074***	1.01	-0.057**	1.23	-0.151**
110-200	1.16	-0.092***	1.28	-0.121***	0.71	-0.053

Note. *** p ‹.001; ** p ‹.01; * p ‹.025; (*) p ‹.05. All intercepts were › 0 (p ‹ .01).

The regression parameters are all displayed in Table 2 and one example will be detailed in order to guide the reader. The first rows in the table body concern the first of the seniority indicators, number of years in psychotherapy practice after licensing. This variable was split in three groups, of roughly equal size (more than 300 observations in each group). When cases in different types of treatments were pooled, in the left-most column, the initial level of symptom

distress (a) was higher the longer the post-license experience, from 0.97 over 1.07 to 1.08. These differences were not significantly different from 0. When change was concerned (b), the cases with the least experienced therapists (with 3-8 years of post-license experience) changed at a low rate, -0.016. (It should be noted that the minus sign indicates a desirable decrease in symptom distress.) This was not significantly different from 0. As the therapists' experience increased, so did the change rate, over –0.082 to –0.104 in the most experienced group (more than 11 years). The two most experienced groups differed significantly from the least experienced group, t (∞) > 2.30, p <.05. How the patients' self-ratings changed across time in treatment, in the three groups, is shown in Figure 2.[2] Two reference lines are inserted

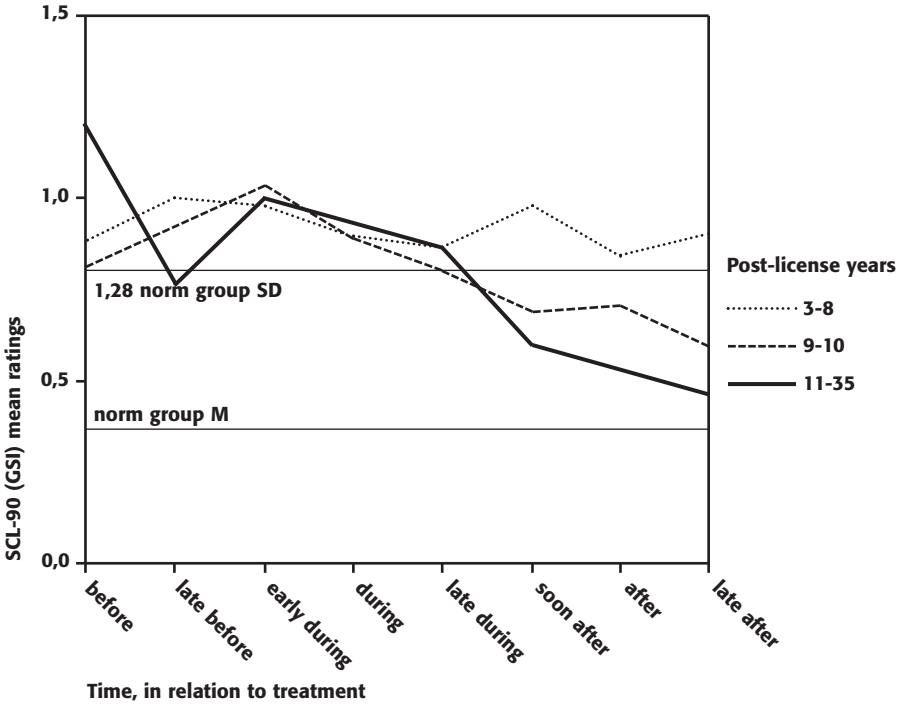

Figure 2. Mean trajectories across stages in treatment for patients with therapists with different amount of post-license experience. (Reference lines refer to normal mean M and caseness criterion [Derogatis & Lazarus, 1994].)

[2] The regression parameters in Table 2 are estimated on the exact size of each subgroup on the time scale, whereas the figure shows the means in each subgroup across the time scale, irrespective of the sizes of the subgroups. Especially, it should be noted that there were only 20 persons in total at the very first step of the time scale, which means, of course, that the subgroup means at this step are quite unstable. This is taken account of in the regression analyses (each step is weighed, so-to-speak, in relation to the number of observations on it), but not in the figures (all steps appear equally important).

in the figure, one indicating the mean in a normal population and one dividing the 10% highest-scoring persons in the normal population (1.28 SD above the mean). This has been suggested by Derogatis and Lazarus (1994) as best discriminating between persons who are or will become psychiatric patients and those who will not ("the caseness criterion").

When cases were divided in psychotherapy and psychoanalysis cases, the picture remained largely the same, as may be seen in the columns to the right in the table. In the psychoanalysis group the differences between the experience subgroups in change rate were more pronounced, and the patients of the least experienced analysts had indeed a trend towards more symptom distress (indicated by positive b). One may notice, however, that subdivision of the psychoanalysis cases created subgroups with large mean errors, because the subgroups became rather small. Therefore, slopes and slope differences of sizes that were significant in the larger psychotherapy subgroups failed to reach significance in the smaller psychoanalysis subgroups.

Turning, now, to the second seniority indicator, supervisor training, there was a slight but significant superiority for cases with therapists with such training, $t\ (\infty) = 2.02$, $p < .05$, which again was somewhat more pronounced among the psychoanalysis cases.

The second cluster of experience variables was labeled psychodynamic repute, because psychoanalytic training and being or having been in long personal therapy, particularly if a psychoanalysis, are high-prestige indicators among psychodynamic therapists in Sweden. As may be seen in Table 2, there was no difference between those with and those without psychoanalytic training in the undivided group, except that the patients of those without it started from a slightly, non-significantly, higher level. In the psychotherapy group the psychotherapists (those without psychoanalytic training) had patients with a significant decline in symptom distress, whereas the change among the cases of the therapists with psychoanalytic training did not differ from 0. A direct comparison between the two groups of cases did not reach significance, however.

When personal therapy is concerned, there is a very clear negative trend in the total group with increasing experience; from -.105 in the group with least personal therapy, over -.061, to -.030 in the group with most. The difference between the extreme groups is significant, $t\ (\infty) = 2.65$, $p < .01$. Obviously, however, there is a difference between the psychotherapy and the psychoanalysis cases here, because among the latter there is an opposite (though nonsignificant) trend.

The long-and-faithful-service cluster is indicated by therapist's age and number of years in psychotherapy practice before licensing. When therapists were divided in three equal age sub-

groups, the change parameter increased with increasing age. The difference between the cases of the youngest and the oldest therapists was not really significant, however, t (∞) = 1.94, p >.05 (two-tailed test).

Pre-license experience, in contrast to post-license experience, seemed to make little difference, as may be seen in Figure 3. The differences between the subgroups were not significant, whether in the total group or in any of the two treatments.

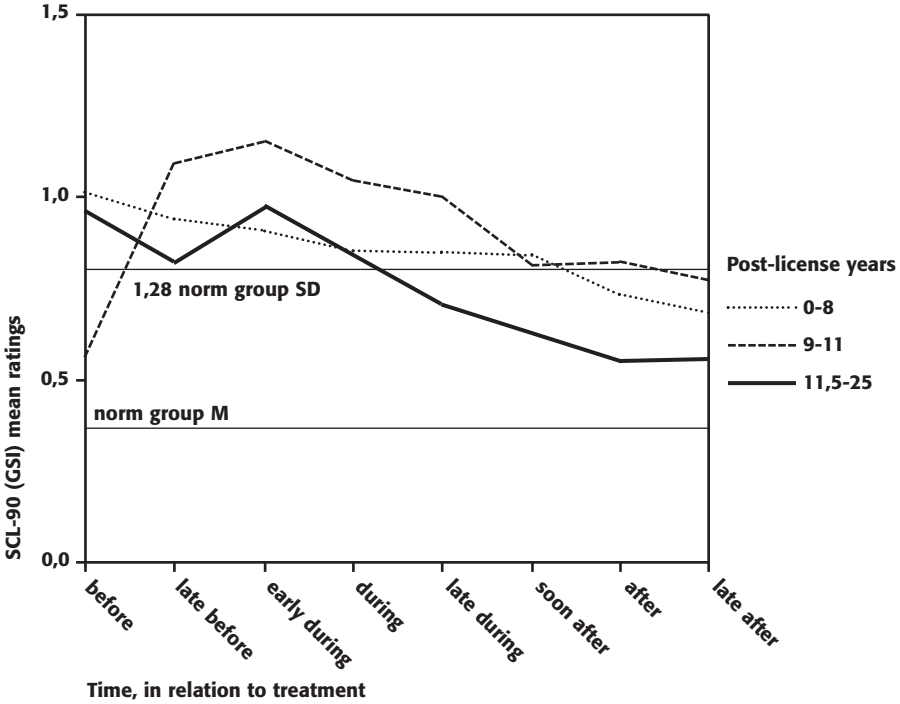

Figure 3. Mean trajectories across stages in treatment for patients with therapists with different amount of pre-license experience. (Reference lines refer to normal mean M and caseness criterion [Derogatis & Lazarus, 1994].)

As there were no therapists with CBT or CT training in the sample, the fourth cluster, case-load, had only case-load, the accumulated number of patients in individual therapy, as an indicator. The change among patients with the most experienced therapists did not differ significantly from those of the least experienced ones. There was an interaction with treatment, however, such that, in the psychotherapy group, the difference between the extreme groups was strongly significant, t (∞) = 3.39, p <.01, whereas, in the psychoanalysis group, there was a difference in the opposite direction, albeit not significant.

Discussion

Obviously, different experience indicators were associated in different ways with patient outcome. Although not shown, the findings were essentially replicated with the SOCS and the SAS, albeit somewhat less pronounced, so the associations are not specific to symptom distress. Whereas the seniority indicators (post-license years; supervisor training) were positively, and significantly, associated with patient outcome, the association between psychodynamic repute was negative, when personal therapy was concerned, and nil, when psychoanalytic training was concerned (among the psychotherapy cases). The variables in the long-and-faithful-service cluster seemed to matter little (age) or not at all (pre-license years), and the association between outcome and case-load was opposite among the psychotherapy and the psychoanalysis cases. No wonder, then, that generic experience level has not shown any strong relation to patient outcome.

Some of these findings are indeed quite interesting. For instance, pre-license and post-license experience seem to have different impact. The critical variable here is probably the difference before and after formal psychotherapy training. One may speculate that on the job experience before formal training lacks the theoretical underpinning that makes increasing on the job experience after formal training more and more useful to the patients. There may be reasons to question the value of "lay psychotherapy," that is, therapy offered by persons without formal training, even if they are in supervision.

Another intriguing finding concerns the variables in the psychodynamic repute cluster. Whereas the value of long training analysis seems fairly clear among the psychoanalysis cases, there is a direct and even clearer negative relation between length of personal therapy and outcome of psychotherapy. On this variable research has hitherto offered mixed findings (Macran & Shapiro, 1998). As has been reported elsewhere (Sandell, Carlsson, Schubert, Blomberg & Lazar, 2001), this negative value of long personal therapy is even more specifically linked to cases where the psychotherapist has been in particularly long personal psychoanalysis. Our interpretation of the findings on both personal therapy and psychoanalytic training is that the tricks of the psychoanalytic trade are irrelevant in psychotherapy (when the therapist has psychoanalytic training) or even detrimental (when the therapist has learned them by identification with his or her own analyst). If this is the case, the consequences for training in psychodynamic psychotherapy are far-reaching, when qualifications among applicants are rated, when teachers are recruited, and when techniques and approaches taught are concerned. It appears that the transfer value of psychoanalysis to psychotherapy has been overrated. Also, it would appear that persons contemplating a dynamic psychotherapy should carefully consider whether the therapist may be "too psychoanalytical."

However, when the importance and influence of different experience indicators are discussed, one had better consider carefully how the relations should be interpreted. Basically, there are three different possibilities. The interpretation that one might prefer is that the associations are direct treatment effects, for instance, that therapists become more skillful the longer they have been professionally active post-license. One alternative interpretation is that the associations are due to selection with increasing experience. Such selection means that the therapists who are more skillful tend to become more experienced. This may sound paradoxical but is not, really. Thus, for instance, more skillful therapists might continue to work as therapists, whereas less skillful ones might change to other specialties than psychotherapy. This attrition would be expected to accumulate with increasing years in the profession and with increasing age. The same mechanism might cause the more skillful therapists to train as supervisors. Considering the negative association with personal therapy, it might be the less skillful (and therefore less secure) therapists who seek to improve their performance by continuing personal therapy longer than more skillful ones. There is practically no way to control for this mechanism as long as one cannot randomize experience on therapists – and whereas that might be possible with training, in principle, it is certainly not with "number-of-years-type" of variables

A second alternative type of interpretation is patient selection, indeed of two kinds. The first is that more experienced therapists learn to select patients with better prognosis. Whereas there is no tendency for patients of more experienced therapists to commence their treatments on lower levels of symptom distress, there may be more subtle qualities with the patient which indicates how promising the case is, qualities that the more experienced therapist may be more sensitive to. The second kind of patient selection may be less reasonable but not totally unreasonable, that "the good patient" manages to spot the more experienced therapist. For instance, "good patients" may keep themselves better informed about therapists' qualifications or be more sensitive to skill and non-skill with the presumptive therapist during initial interviews and shun the less experienced ones on these bases. Patient selection is more easily controlled, simply by randomized assignment of therapists and patients to each other. The preconditions of this study precluded this control.

Whatever the mechanism, direct causality, selection with increasing experience, or patient selection, patients with therapists who have more of some (not all) kinds of experience show greater change – although this is not necessarily because of the therapists' experience. Some way, some aspects of therapist experience do matter. In this respect there are practical conclusions to be drawn from this study with respect to the selection of applicants for psychodynamic therapy training or therapist positions – and the selection of psychodynamic therapists by patients – and for training in psychodynamic therapy. Whether the findings would be replicated in the context of other types of therapy than the psychodynamic one remains to be researched.

References

- Antonovsky, A. (1987). *Unraveling the mystery of health: How people manage stress and stay well*. San Francisco: Jossey-Bass.
- Auerbach, A. H., & Johnson, M. (1977). Research on the therapist level of experience. In A. S. Gurman & A. M. Razin (Eds.), *Effective psychotherapy.* New York: Pergamon Press.
- Berman, J. S., & Norton, N. C. (1985). Does professional training make a therapist more effective? *Psychological Bulletin*, 98, 401-406.
- Bell, R. Q. (1953). Convergence: An accelerated longitudinal approach. *Child Development*, 24, 145-152.
- Beutler, L. E., Machado, P. P. P., & Neufeldt, S. A. (1994). Therapist variables. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 229-269). New York: Wiley.
- Blomberg, J., Lazar, A., & Sandell, R. (2001). Outcome of patients in long-term psychoanalytical treatments. First findings of the Stockholm Outcome of Psychotherapy and Psychoanalysis (STOPP) study. *Psychotherapy Research*, 11, 361-382.
- Carroll, L. (1865). *Alice's Adventures in Wonderland. In the Complete Illustrated Works of Lewis Carroll* (pp. 11-114). London: Chancellor Press, 1982.
- Cohen, J. (1988). *Statistical power analysis in the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Dawes, R. M. (1994). *House of cards. Psychology and psychotherapy built on myth*. New York: Free Press.
- Derogatis, L. R., & Lazarus, L. (1994). SCL-90-R, Brief Symptom Inventory, and matching clinical ratings scales. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 217-248). Hillsdale, N.J.: Lawrence Erlbaum.
- Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symtom Checklist (HSCL): A self-report symtom inventory. *Behavioral Science*, 19, 1-15.
- Jacobson, N. (1995). The overselling of therapy. *Family Therapy Networker*, 41-47.
- Macran, S. & Shapiro, D. A. (1998). The role of personal therapy for therapists: A review. *British Journal of Medical Psychology*, 71, 13-25.
- Orlinsky, D. E., Ambühl, H., Rönnestad, M. H., Davis, J., Gerin, P., Davis, M., Willutzki, U., Botermans, J.-F., Dazord, A., Cierpka, M., Aapro, N., Buchheim, P., Bae, S., Davidson, C., Friis-Jörgensen, E., Joo, E., Kalmykova, E., Meyerberg, J., Northcut, T., Parks, B., Scherb, E., Schröder, T., Shefler, G., Stiwne, D., Stuart, S., Tarragona, M., Vasco, A. B., & Wiseman, H. (1999). Development of psychotherapists: Concepts, questions, and methods of a collaborative international study. *Psychotherapy Research*, 9, 127-153.
- Rönnestad, M. H., Orlinsky, D. E., Parks, B. K., Davis, J. D., and the Society for Psychotherapy Research (SPR) Collaborative Research Network (1997). Supervisors of psychotherapy: Mapping experience level and supervisory confidence. *European Psychologist*, 2, 191-201.
- Sandell, R., Blomberg, J., Lazar, A., Carlsson, J., Broberg, J, & Schubert, J. (2000). Varieties of long-term outcome among patients in psychoanalysis and long-term psychotherapy. A review of findings in the Stockholm Outcome of Psychoanalysis and Psychotherapy Project (STOPPP). *International Journal of Psychoanalysis*, 81, 921-942.

- Sandell, R., Blomberg, J., Lazar, A., Carlsson, J., Broberg, J., & Schubert, J. (2001). Unterschiedliche Langzeitergebnisse von Psychoanalysen und Psychotherapien. Aus der Forschung des Stockholmer Projekts. *Psyche*, 55, 277-310.
- Sandell, R., Carlsson, J., Schubert, J., Blomberg, J., Lazar, A., & Broberg, J. (2001). *Therapists' therapies. The relation between therapists' personal/training therapies and patient outcome in long-term psychotherapy and psychoanalysis*. (Submitted)
- Stein, D. M., & Lambert, M. L. (1984). On the relationship between therapist experience and psychotherapy outcome. *Clinical Psychology Review*, 4, 127-142.
- Stein, D. M., & Lambert, M. L. (1995). Graduate training in psychotherapy: Are therapy outcomes enhanced? *Journal of Consulting and Clinical Psychology,* 63, 182-196.
- Strupp, H. H. (1996). The tripartite model and the Consumer Reports study. *American Psychologist*, 51, 1017-1024.
- Strupp, H. H., & Hadley, S. (1979). Specific versus nonspecific factors in psychotherapy: A controlled study of outcome. *Archives of General Psychiatry*, 36, 1125-1136.
- Weissman, M. M., & Bothwell, S. (1976). Assessment of social adjustment by patient self-report. *Archives of General Psychiatry*, 33, 1111-1115.
- Weissman, M. M., Prusoff, B. A., Thompson, W. D., Harding, P. S., & Myers, J. K. (1978). Social adjustment by self-report in a community sample and in psychiatric outpatients. *Journal of Nervous and Mental Disease*, 166, 317-326.
- Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology*, 55, 542-549.

Correspondence

concerning this article should be sent to:
Rolf Sandell
Department of Behavioural Sciences
Linköping University
S-581 83 Linköping, Sweden.
e-mail: rolsa@ibv.liu.se