ONE-SESSION GROUP THERAPY OF SPIDER PHOBIA: DIRECT VERSUS INDIRECT TREATMENTS

LARS-GÖRAN ÖST, INGELA FEREBEE, and TOMAS FURMARK

1Department of Psychology, Stockholm University, S-106 91 Stockholm, Sweden and 2Department of Psychology, University of Uppsala, Uppsala, Sweden

(Received 3 February 1997)

Summary—Forty-six patients with spider phobia, fulfilling the DSM-IV criteria for specific phobia, were assessed with behavioral, physiological and self-report measures. They were randomly assigned to three group treatment conditions: (1) direct treatment; (2) direct observation; and (3) indirect observation. All treatments were carried out in large groups of eight patients, and consisted of one 3 hr session of massed exposure and modelling. The results showed that on the behavioral test, measures and the specific self-report measures of spider phobia the direct treatment was significantly better than direct observation and indirect observation, which did not differ. On the physiological measures and the psychopathology self-report measures there were significant pre-post improvements, but no differences between the groups. The effects were maintained or furthered at the one year follow-up assessment. The proportion of clinically significantly improved patients were, at post-treatment, 75% in the direct treatment, 7% in the direct observation, and 31% in the indirect observation group. At follow-up, the corresponding figures were 75, 14, and 44%, respectively. The conclusion that can be drawn is that direct treatment is the treatment of choice. © 1997 Elsevier Science Ltd

INTRODUCTION

Specific phobias are the most common of the anxiety disorders with an estimated lifetime prevalence of 10–11% in the American population (DSM-IV; APA, 1994). Among the specific phobias it seems that spider phobia is the most common in the population (Bourdon et al., 1988). The therapy proven to be most successful for phobias is exposure (e.g. Chambless, 1990; Marks, 1987). Most exposure treatments have been therapist-directed where the therapist usually has weekly sessions with the patient until the problem is remedied. Recent research has shown that short intensive treatment during a single session produces just as good results as more spaced programs do and could be considered the treatment of choice for specific phobias (Hellström & Öst, 1995; Hellström, Fellenius & Öst, 1996; Öst, 1989a, 1996; Öst, Hellström & Kåver, 1992; Öst, Salkovskis & Hellström, 1991). This treatment has also been replicated in spider phobia (Arntz & Lavy, 1993), and there is even earlier research showing that brief treatment of animal phobias is effective (e.g. Bandura, Blanchard & Ritter, 1969).

One possibility to increase cost-effectiveness is to treat patients in groups. The literature contains a number of studies on specific phobias in which group treatments have been used. In acrophobia Ritter (1969a) used a small group of three patients and found that group contact desensitization was significantly better than non-contact desensitization and a control condition. Pendleton and Higgins (1983) also used groups of three or four patients and reported that negative practice and systematic desensitization were equally effective. In flying phobia Howard, Murphy and Clarke (1983) treated subjects in groups of two to three during 8 weekly 1 hr sessions and found that all active treatment conditions, i.e. systematic desensitization, flooding, implosion, and relaxation were more effective than no treatment but there was no difference between them. The specific phobia having the most group treatments is dental phobia. Wroblewski, Jacob and Rehm (1977) gave groups of two to five patients seven sessions during a 10-day period and found that symbolic modelling plus relaxation was more effective than symbolic modelling only or attention placebo when it comes to obtaining dental treatment after
therapy. Gatchel (1980) compared a group of eight patients treated with self-control desensitization with a group of five patients treated with education and discussion, and found the former to be significantly better than the latter and a control. Gauthier, Savard, Hallé and Dufour (1985) compared flooding and coping skills training, in a cross-over design, and administered the treatments in groups of three to four patients. They found no difference between the conditions. Jerremalm, Jansson and Öst (1986) treated dental phobics in groups of four during nine 90 min sessions, finding that applied relaxation and self-instructional training did equally well. Ning and Liddell (1991) compared massed (2 sessions per week for 2 weeks) with spaced (4 weekly sessions) anxiety management training, and found them to be equally effective. Nine patients were randomized to each condition but there is no information on the group size.

In a previous study on spider phobia from our laboratory (Öst, 1996) it was found that a large group of seven or eight patients yielded almost as good results as a small group of three or four patients. Despite a trend in favor of the small group on most measures this was only significant on the self-rating of anxiety during the behavioral test. In terms of clinically significant improvement the small group yielded 82% at post-treatment and 95% at follow-up, while the corresponding figures for the large group were 70 and 75%, respectively.

The conclusion that can be drawn from this brief review is that group treatment, mostly in small groups of three or four patients, has been used for dental phobia and to some extent for acrophobia, flying phobia, and spider phobia. However, the group format has not been tested for patients with blood-injury-injection phobia, or claustrophobia, which are very common in the general population.

One important component of the one-session treatment for animal phobias is the participant modelling that is fully integrated with exposure. Other forms of modelling have been investigated in previous research, primarily live modelling and symbolic modelling (see review by Rosenthal & Bandura, 1978). In live modelling the patient observes a model interacting with the animal in question, but does not actively interact herself/himself. In symbolic modelling the patient is shown a film or videotape of a model. These two types of modelling may utilize a patient as a model, but that is not a requirement, and any suitable person can act as a model.

Bandura et al. (1969) found that participant modelling (PM) was significantly better than symbolic modelling (SM) for snake phobics. In a second study on snake phobia (Bandura, Adams & Beyer, 1977) PM was significantly better than live modelling, and the same result was reported in a study of acrophobia by Ritter (1969b). In dental phobia, Shaw and Thoresen (1974) found that SM was equal to systematic desensitization and better than both placebo and waitlist control. Finally, Bernstein and Kleinknecht (1982) found equal effects of PM, SM and exposure in vivo for dental phobia. The conclusion that can be drawn from these studies is that for animal and acrophobia, PM is more effective than live modelling and symbolic modelling, but for dental phobia SM did as well as the comparison treatments.

The primary aim of the present study was to investigate the effects of two forms of modelling (live and symbolic) in comparison to the standard one-session treatment containing prolonged exposure and participant modelling, in group treatment of spider phobia. Live modelling in this study means that one patient is being treated while eight patients are watching the treatment from a distance of 3–5 feet. This treatment is called direct observation. Symbolic modelling, which in this study, is called indirect observation, means that the group of eight patients is watching the videotape of the patient being treated in the direct observation condition. In this way, the personal characteristics, anxiety reactions, etc. of the model, as well as the therapeutic assistance of the therapist are kept constant across the latter two conditions. Based on the studies by Bandura et al. (1969, 1977) cited above it was predicted that the direct treatment would be more effective than both indirect treatments, and that the direct observation (live modelling) would be more effective than indirect observation (symbolic modelling). This prediction holds for the behavioral test measures and the self-report measures of spider phobia, but not for the general questionnaire measures where only significant improvements across the groups were expected. It was also predicted that the effects would be maintained, or that further improvement would have taken place, at the 1 yr follow-up, which is the case in our previous one-session studies (Hellström & Öst, 1995; Hellström et al., 1996; Öst, 1996; Öst et al., 1991, 1992).
The secondary aim was to replicate earlier research by e.g. Bandura (1977) and Williams, Turner and Peer (1985) showing that perceived self-efficacy is a better predictor of approach behavior than anticipated or experienced anxiety.

METHOD

Subjects

The subjects for the study were recruited through advertisements in local newspapers or, were referred by their physicians in the Uppsala County. There were 46 patients, all women, and all of whom had been diagnosed with specific phobia of spiders according to DSM-IV (APA, 1994). In order to participate in the study, the following criteria had to be fulfilled: (1) be between the ages of 18 and 60 yr; (2) be afraid and exhibit avoidance of a number of situations where confrontation with spiders occurred, this being the primary problem for which the patient had sought help; (3) a minimum of 1 yr duration of the phobia; (4) be willing to participate in the study for a period of 3 weeks (and the 1 yr follow-up); (5) be incapable of inserting the hand in a plastic container with a spider (during the behavioral test); (6) have no other psychiatric problems requiring immediate treatment (7) have no psychotic or organic illnesses; (8) have no disease of the heart or lungs.

The patients were screened using a modified version of Anxiety Disorders Interview Schedule-IV (Brown, DiNardo & Barlow, 1994), yielding DSM-IV anxiety diagnoses. Of the 52 patients that were screened, 50 fulfilled the criteria and were included in the study. Two of the patients functioned as models and received individual treatment in the second condition. They were excluded from the analyses, as were two patients who did not show up for the treatment. The average age of the 46 patients was 29.5 yr (SD = 8.2; range 19–57), and the average age at which the phobia began was 6.6 yr (SD = 2.8; range 3–16). Twenty-nine (63%) of the patients were married or living together with a steady partner, 15 were single and two were divorced. There were 37 who worked or studied full-time, and 8 part-time, while one was unemployed. All of the patients were handicapped by their phobia for spiders in their daily lives or work.

Assessment

The patients were assessed prior to, following, and 1 yr after the treatment. All the screening-interviews and the behavioral tests were performed by two graduate students, and the therapist was in no way involved with the assessments.

Self-report measures. In order to measure the degree of spider phobia, the Spider Phobia Questionnaire (SPQ, Klorman, Weerts, Hastings, Melamed & Lang, 1974) and the Spider Questionnaire by Watts and Sharrock (1984) were used. The patients also self-assessed on 0-8 scales the degree of fear (SAFEAR) and avoidance (SAAV) in relation to each of a small, medium and large spider. There was also a 0–8 rating of overall degree of handicap (SAHAND) that the patient experienced from her spider phobia. The Fear Survey Schedule-III (FSS; Wolpe & Lang, 1964) was used to measure general phobic tendencies, the State–Trait Anxiety Inventory (STAI; Spielberger, Gorsuch & Lushene, 1970) was used to measure general anxiety, while the Beck Depression Inventory (BDI; Beck, 1967) and the Beck Anxiety Inventory (BAI; Beck, Epstein, Brown & Steer, 1988) were used to measure depression and anxiety, respectively.

Assessor rating. At the close of the screening interview, the therapist rated the severity of the patient’s spider phobia on a scale of 0–8; where 0 was free from symptoms, and 8 was extremely severe and disabling with all aspects of normal life affected by the phobia.

Behavioral test. A behavioral test was used to assess the patient’s avoidance of spiders and anxiety experienced. This test is described in detail in Öst et al. (1991). The behavioral test was scaled from 0 to 12; where 0 was refusal to enter the room, and 12 was holding the spider for at least 20 sec. Immediately after having received the instructions for the behavioral test the patient’s self-efficacy was measured by using a questionnaire describing the 13 steps of the test. The patient rated her confidence regarding being able to complete each step on a 0–100 scale. The level of self-efficacy was calculated as the proportion of tasks rated 20 or higher, and the strength of self-efficacy as the mean rating across the 13 steps. Patients were also asked to rate
(0–100) the anxiety they anticipated, and that they actually experienced at the point of interruption.

Procedure

Pre-treatment. The subjects who volunteered for the study were mailed the questionnaires described above. Each patient was then contacted by phone and an appointment was made for a screening interview. The interview took 60 min and its purpose was to establish whether or not the patient fulfilled the criteria necessary to be included in the study, and to conduct the behavioral test. After the interview, blood pressure (BP) and heart rate (HR) were measured with a portable digital blood pressure apparatus (model UA-751), followed by a 5 min rest period (adaptation), after which they were measured again (baseline). The patient then received instructions regarding the behavioral test where upon the assessor and the patient left the interview room and walked 20 m to conduct the behavioral test in another room. BP and HR were recorded once more at the point the patient interrupted the behavioral test. The interview, which was held approximately 1 week before treatment, ended with a brief description of the treatment the patient would undergo.

Post-treatment and follow-up. One week after treatment, the patient performed the behavioral test and answered the questionnaires again. The therapist also informed the patient of the voluntary maintenance program (Öst, 1989b) and the patient was instructed to continue her own self-exposure. One year later, the patient was called back to perform the behavioral test and fill in the questionnaires once again.

Design

After the pre-treatment assessment, the patients were randomly assigned to three different one-session treatment conditions: (1) direct treatment in groups of eight patients \( n = 16 \); (2) direct observation in groups of eight \( n = 16 \); and (3) indirect observation in groups of eight \( n = 16 \).

Treatments

General aspects. The therapist met the groups of patients for 30 min immediately before the treatment started. The purposes of this meeting were firstly, that the patients should share some of their experiences concerning spiders, and secondly, that the therapist should describe how the treatment was going to be carried out. When arriving at this meeting the patients did not know which condition they had been randomized to. They were told this at the end of the information session.

Direct treatment. Based on the description given by Öst (1989a), the one-session therapist-directed exposure includes a combination of gradual exposure and modelling, and is described in detail in Öst et al. (1991) and Öst (1997). In the present study, the treatment was carried out in groups of eight patients during a 3 hr session.

The patients were seated in a circle, and in front of each there was a small table with a plastic container on top. The therapist sat on an office chair on wheels to enable him to move around easily and help the patient most in need of his assistance at any time. For each patient, there was a set of four spiders of gradually increasing sizes.

The following pre-treatment instructions were given. The patients are told that the treatment is done as a team-work, and both the therapist and the patients have equal responsibility for achieving a good result. Moreover, the therapist will never do anything without first describing it to the patient, then demonstrating it, and finally getting the patient’s permission to do it. The therapist also informs the patients that even if the treatment will mean that they will be exposed to much more than they ever have encountered in natural situations this will not ‘break their personal record’ of anxiety in the phobic situation. Finally, the advantages of group treatment, especially the vicarious fear reduction, are emphasized.

The therapist usually demonstrates the different activities with one patient, while the other observes this. Then all the patients are instructed to perform the same activity and the therapist helps by using physical guidance to those who need assistance. Usually, all the patients need the therapist’s help, but to varying degrees. In some groups, it happens that one of the patients
spontaneously acts as a 'co-therapist' by helping the one sitting next to her when the therapist is occupied with one of the other patients. When the patients can handle the smallest spider with a clearly reduced anxiety it is time to proceed to the next, and so on until the patients have been exposed to all four spiders. The timing is geared by the progress of the average patient in the group, and it sometimes happens that one patient cannot proceed to the next spider, but has to continue working with the smaller one. The goal is that all patients should be able to have two spiders (2 and 3 cm in size) crawling on their hands.

Direct observation. At the end of the information session the patients in these groups were told that one of them was to be treated while the others observed this treatment at a distance of 4–6 feet. In order for the patients not to believe that the therapist had selected an especially 'easy patient' for treatment, the patient to receive the treatment was decided by lottery. The bowl of tickets contained eight blanks and one winning ticket, and the winner received the treatment. This patient got the ordinary individual one-session treatment, with the exception that the therapy room was full of observing fellow spider phobics. The session was video-taped.

Indirect observation. At the end of the information session, the patients in these groups were told that they were going to watch a videotape of a treatment session done in the direct observation condition the previous week. They were also informed that the tape depicted an ordinary individual session with the exception of the observers present in the room. The patients sat in the same therapy room used for the first two conditions and the therapist just turned the video on and left the patients for the duration of the tape. When this was concluded he returned to answer any questions that the patients might have.

Thus, all groups were treated in the same therapy room, and the two observation conditions had the same stimulus material, although one was live and the other on videotape.

Therapist

The therapist conducting the treatments was the first author of this paper. At the time of the study L.-G. Öst had 26 years of clinical experience with behavior therapy, and had treated eight small and four large groups before the start of this study.

Statistical methods

Non-parametric tests ($\chi^2$ and Fisher's exact probability test) were used for ordinal data. For each group of interval measures, repeated measurement ANOVAs using Bonferroni corrections were performed. All statistical tests were two-tailed, and performed using the statistical package SPSS for Windows release 6.1.3 supplied by SPSS Inc. (1995).

RESULTS

Attrition

Two of the patients randomized to the direct observation condition did not show up for treatment. At the 1 yr follow-up, six patients were unable to come for the assessment. Their data were estimated with the respective group means.

Pre-treatment differences

On three of the specific spider phobia measures there were significant pre-treatment differences. On SPQ, WS and SAFEAR the direct observation group had significantly higher means than the indirect observation group, and on WS, the direct treatment group was also significantly higher than the indirect observation group. For these measures analysis of covariance with the pre-treatment data as covariate were used.

Treatment time

The length of the two group treatment sessions in the direct treatment conditions was 205 and 208 min. respectively. The two individual treatment sessions in the direct observation condition were 120 and 135 min. respectively. These figures are all within the mean $\pm 1$ SD for group and individual treatment, respectively, for our previous three spider phobia studies.
Behavioral test

For the ANOVAs, a Bonferroni correction was used meaning an α-level of 0.05/3 = 0.016.

Steps in the behavioral test. Results on avoidance in the behavioral test are shown in Fig. 1 (left panel). The ANOVA yielded a significant group \( F(2,43) = 8.36, P < 0.001 \), and time factor \( F(2,86) = 94.15, P < 0.0001 \), while the interaction factor was non-significant. The subsequent one-way ANOVAs at post-treatment and follow-up yielded the following results. At post-treatment the DT group was significantly better than the two observation groups, and at follow-up DT was better than DO, while the IO group did not differ from either of these.

Self-rating of anxiety. The middle panel of Fig. 1 shows the anxiety level at the time of interruption of the behavioral test. An ANOVA indicated a significant time effect \( F(2,86) = 29.20, P < 0.0001 \), interaction effect \( F(4,86) = 3.46, P < 0.012 \), and a marginally significant group effect \( F(2,43) = 3.01, P < 0.06 \). Subsequent one-way ANOVAs yielded a marginal difference at post-treatment, but at follow-up, the DT group was significantly better than the DO group, while the IO group did not differ from either of these.

Assessor rating. The right panel of Fig. 1 shows the patients' degree of phobic severity rated by the assessor. An ANOVA indicated a significant group effect \( F(2,43) = 8.83, P < 0.005 \), time effect \( F(2,86) = 126.01, P < 0.00001 \), and an interaction effect \( F(4,86) = 9.07, P < 0.0001 \). Subsequent one-way ANOVAs showed that at post-treatment the DT group was significantly better than the two observation groups, and at follow-up DT and IO were both significantly better than DO.

Self-efficacy. The results on the self-efficacy measures are displayed in Fig. 2. On the level of self-efficacy (left panel) there was a significant group effect \( F(2,43) = 7.83, P < 0.005 \), time effect \( F(2,86) = 55.89, P < 0.0001 \), and interaction effect \( F(4,86) = 5.27, P < 0.001 \). The subsequent one-way ANOVAs showed that the DT group had a significantly higher level of self-efficacy than the DO and IO groups, both at post-treatment and at follow-up. On the strength of self-efficacy (right panel), there were also significant group \( F(2,43) = 8.17, P < 0.002 \), time \( F(2,86) = 41.93, P < 0.0001 \), and interaction effects \( F(4,86) = 4.52, P < 0.005 \). Subsequent analyses also showed that the DT condition had significantly higher strength of self-efficacy than the other two conditions, both at post-treatment and at follow-up.

Correlations

In order to assess the relationship between behavioral scores on the behavioral avoidance test on the one hand, and anxiety ratings and self-efficacy ratings on the other hand, product-
moment correlations and partial correlations were calculated. These are presented in Table 1. Using the product–moment correlations all three variables correlated significantly with the behavioral scores at post-treatment and follow-up assessment. However, when the effects of the other variables in the table were partialed out, both anticipatory and experienced anxiety lost their significant relationships to the behavioral scores, but the self-efficacy measure remained basically unchanged as a highly significant predictor of behavioral performance.

Physiological measures

The results on the physiological measures are presented in Table 2. For the ANOVAs, an \( \alpha \)-level of 0.05/3 = 0.016 was used (Bonferroni correction). These indicated significant time effects for SBP and DBP, but not for HR. There were no interaction or group effects.

Self-report measures

The results on the self-report measures of spider phobia are presented in Table 3 and the outcome on the psychopathology self-report measures are presented in Table 4. For the ANOVAs, an \( \alpha \)-level of 0.05/5 = 0.01 (Bonferroni correction) was used. These indicated significant time effects for all the individual measures, while the interaction effects only reached a marginal significance level (\( P < 0.05 \)). On SPQ, WS and SAFEAR there were significant group effects, which was also true in the ANOVAs covarying out the pre-treatment values. SAAV also yielded a significant group effect. The subsequent one-way ANOVAs showed that at post-treatment DT was better than DO on the SPQ, WS, and SAAV, while IO was better than DO on SPQ, and DT better than IO on SAAV. At follow-up, the results were quite consistent with DT and IO

---

Table 1. Correlations between behavioral scores, anxiety and self-efficacy measures at pre-, post- and follow-up assessment

<table>
<thead>
<tr>
<th>Variable</th>
<th>Product-moment correlations</th>
<th>Partial correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-</td>
<td>Post-</td>
</tr>
<tr>
<td>Anticipatory anxiety</td>
<td>0.14</td>
<td>-0.45 ( ^a )</td>
</tr>
<tr>
<td>Experienced anxiety</td>
<td>0.00</td>
<td>-0.39 ( ^a )</td>
</tr>
<tr>
<td>Self-efficacy: strength</td>
<td>0.67 ( ^c )</td>
<td>0.86 ( ^c )</td>
</tr>
</tbody>
</table>

\( ^a P < 0.01; ^b P < 0.001; ^c P < 0.0001. \)
Table 2. Means (SDs) on the physiological measures at pre-, post- and follow-up assessment

<table>
<thead>
<tr>
<th>Measure</th>
<th>Condition</th>
<th>Stage</th>
<th>Direct treatment</th>
<th>Direct observation</th>
<th>Indirect observation</th>
<th>ANOVA F-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBP</td>
<td></td>
<td>Pre</td>
<td>129.7 (15.1)</td>
<td>133.0 (12.0)</td>
<td>133.5 (12.1)</td>
<td>G:1.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>123.1 (13.0)</td>
<td>127.6 (11.9)</td>
<td>121.8 (18.7)</td>
<td>T:63.3d</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-up</td>
<td>123.7 (17.7)</td>
<td>132.6 (15.2)</td>
<td>121.6 (18.7)</td>
<td>I:1.02</td>
</tr>
<tr>
<td>DBP</td>
<td></td>
<td>Pre</td>
<td>92.5 (10.7)</td>
<td>94.4 (9.7)</td>
<td>92.8 (9.4)</td>
<td>G:0.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>83.6 (9.1)</td>
<td>85.4 (11.0)</td>
<td>84.1 (13.8)</td>
<td>T:10.82d</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-up</td>
<td>86.5 (14.9)</td>
<td>87.5 (14.4)</td>
<td>83.4 (11.8)</td>
<td>E:0.19</td>
</tr>
<tr>
<td>HR</td>
<td></td>
<td>Pre</td>
<td>84.5 (12.8)</td>
<td>82.6 (17.5)</td>
<td>86.9 (15.7)</td>
<td>G:0.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>78.1 (13.1)</td>
<td>84.3 (17.5)</td>
<td>77.2 (13.8)</td>
<td>T:1.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-up</td>
<td>79.1 (12.1)</td>
<td>80.7 (19.6)</td>
<td>86.2 (13.6)</td>
<td>I:1.11</td>
</tr>
</tbody>
</table>

G, group; T, time; I, interaction effect in the ANOVA. * P < 0.016; ** P < 0.005; *** P < 0.001; **** P < 0.0001.

...both being better than DO on SPQ, WS, SAFEAR, and SAAV, but not differing from each other.

...For the ANOVAs on the psychopathology measures, an α-level of 0.05/5 = 0.01 (Bonferroni correction) was used, and these yielded significant time effects on all measures, but no group or interaction effects. The significant time effects were followed by paired t-tests between pre-post, and pre-follow-up, for each individual group. These showed that the DT group yielded significant changes on all five measures, while the DO group changed significantly on all except BAI, and the IO group only changed on the STAI-S and the FSS.

Clinically significant improvement

Jacobson, Follette and Revenstorf (1984) criteria for ‘clinically significant improvement’ were used for this assessment. These criteria stipulate that the patient must, in addition to displaying a statistically reliable change, fall within the range of a normal group, or outside the range of the patient group at the post-treatment assessment. This was defined as $M \pm 2 SD$ in the direction of functionality. As no norm data exist for these variables, the latter alternative was used.

Behavioral test. The pre-treatment average was $4.82 + 2 (SD = 2.26) = 9.34$, i.e. 10, which means being able to touch the spider with one finger.

Self-rating of anxiety. The pre-treatment average was $83.0 - 2 (SD = 18.2) = 46.6$, i.e. 46. However, at pre-treatment assessment, a number of patients stopped the behavioral test at such a low step that they did not experience very much anxiety. When they approached the spider much more at post- and follow-up assessment their anxiety did not decrease from an already fairly low level. For these patients the Spider Phobia Questionnaire was used. The pre-treatment mean was $23.2 - 2 (SD = 3.7) = 15.8$, i.e. 16.

Assessor rating. The pre-treatment mean was $5.9 - 2 (SD = 0.88) = 3.83$, i.e. 4.

All three criteria were required in order to constitute clinical improvement. At post-treatment, 12 (75%) of the patients in the DT group, 1 (7%) in the DO group, and 5 (31%) in the IO group constituting clinical improvement.

Table 3. Means (SD) on the self-report spider phobia measures at pre-, post- and follow-up assessment

<table>
<thead>
<tr>
<th>Measure</th>
<th>Condition</th>
<th>Stage</th>
<th>Direct treatment</th>
<th>Direct observation</th>
<th>Indirect observation</th>
<th>ANOVA F-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPQ</td>
<td></td>
<td>Pre</td>
<td>23.3 (3.9)</td>
<td>25.4 (2.2)</td>
<td>21.1 (3.5)</td>
<td>G:8.75d</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>13.0 (3.7)</td>
<td>20.4 (5.9)</td>
<td>15.1 (6.2)</td>
<td>T:67.72d</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-up</td>
<td>11.1 (5.2)</td>
<td>19.3 (7.1)</td>
<td>12.6 (7.9)</td>
<td>I:2.83*</td>
</tr>
<tr>
<td>WS</td>
<td></td>
<td>Pre</td>
<td>27.8 (3.6)</td>
<td>29.1 (2.8)</td>
<td>22.9 (5.6)</td>
<td>G:6.10h</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>17.9 (6.9)</td>
<td>24.9 (6.7)</td>
<td>19.1 (6.7)</td>
<td>T:55.03g</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-up</td>
<td>15.9 (6.8)</td>
<td>22.9 (5.6)</td>
<td>15.1 (9.3)</td>
<td>I:3.38*</td>
</tr>
<tr>
<td>SAFEAR</td>
<td></td>
<td>Pre</td>
<td>17.3 (3.0)</td>
<td>19.7 (2.9)</td>
<td>16.8 (3.1)</td>
<td>G:7.09b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>7.8 (5.2)</td>
<td>11.9 (5.6)</td>
<td>11.3 (4.2)</td>
<td>T:70.07f</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-up</td>
<td>7.6 (4.0)</td>
<td>14.6 (5.1)</td>
<td>9.1 (5.4)</td>
<td>I:3.10*</td>
</tr>
<tr>
<td>SAAV</td>
<td></td>
<td>Pre</td>
<td>19.4 (3.5)</td>
<td>21.2 (3.1)</td>
<td>19.5 (3.2)</td>
<td>G:7.46g</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>7.5 (5.8)</td>
<td>12.8 (7.1)</td>
<td>13.4 (5.6)</td>
<td>T:66.00d</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-up</td>
<td>7.9 (4.7)</td>
<td>16.9 (6.1)</td>
<td>10.6 (5.3)</td>
<td>I:4.84*</td>
</tr>
<tr>
<td>SAHAND</td>
<td></td>
<td>Pre</td>
<td>6.0 (1.8)</td>
<td>5.4 (1.6)</td>
<td>4.8 (1.0)</td>
<td>G:0.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post</td>
<td>2.8 (1.7)</td>
<td>3.4 (1.9)</td>
<td>3.1 (1.4)</td>
<td>T:66.53d</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-up</td>
<td>2.1 (1.7)</td>
<td>2.9 (1.3)</td>
<td>2.4 (1.9)</td>
<td>I:2.71*</td>
</tr>
</tbody>
</table>

G, group; T, time; I, interaction effect in the ANOVA. * P < 0.05; ** P < 0.005; *** P < 0.001; **** P < 0.0001.
Table 4. Means (SD) on the psychopathology self-report measures at pre-, post- and follow-up

<table>
<thead>
<tr>
<th>Measure</th>
<th>Stage</th>
<th>Direct treatment</th>
<th>Direct observation</th>
<th>Indirect observation</th>
<th>ANOVA F-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSS-III</td>
<td>Pre</td>
<td>155.3 (35.2)</td>
<td>153.5 (33.4)</td>
<td>158.4 (30.3)</td>
<td>G:0.06</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>138.8 (35.5)</td>
<td>130.4 (42.0)</td>
<td>136.1 (26.5)</td>
<td>T:26.35*</td>
</tr>
<tr>
<td></td>
<td>F-up</td>
<td>136.3 (32.5)</td>
<td>134.9 (34.9)</td>
<td>132.1 (17.2)</td>
<td>I:0.52</td>
</tr>
<tr>
<td>STAI-T</td>
<td>Pre</td>
<td>41.4 (9.7)</td>
<td>43.1 (9.9)</td>
<td>37.6 (6.4)</td>
<td>G:0.90</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>38.4 (9.0)</td>
<td>34.6 (9.3)</td>
<td>35.3 (7.0)</td>
<td>T:21.90*</td>
</tr>
<tr>
<td></td>
<td>F-up</td>
<td>38.1 (9.2)</td>
<td>35.5 (8.1)</td>
<td>33.9 (6.4)</td>
<td>I:3.04*</td>
</tr>
<tr>
<td>STAI-S</td>
<td>Pre</td>
<td>41.3 (9.7)</td>
<td>41.2 (10.8)</td>
<td>40.6 (8.7)</td>
<td>G:0.10</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>33.9 (10.2)</td>
<td>34.9 (7.5)</td>
<td>35.0 (6.9)</td>
<td>T:21.78*</td>
</tr>
<tr>
<td></td>
<td>F-up</td>
<td>32.7 (7.9)</td>
<td>31.0 (7.3)</td>
<td>34.8 (8.4)</td>
<td>I:0.58</td>
</tr>
<tr>
<td>BAI</td>
<td>Pre</td>
<td>12.2 (5.5)</td>
<td>10.5 (6.3)</td>
<td>9.6 (5.5)</td>
<td>G:0.46</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>9.1 (5.6)</td>
<td>7.5 (4.1)</td>
<td>8.8 (3.7)</td>
<td>T:5.87*</td>
</tr>
<tr>
<td></td>
<td>F-up</td>
<td>7.8 (5.0)</td>
<td>7.7 (5.3)</td>
<td>7.5 (5.1)</td>
<td>I:0.50</td>
</tr>
<tr>
<td>BDI</td>
<td>Pre</td>
<td>10.2 (8.5)</td>
<td>9.0 (6.7)</td>
<td>6.9 (4.6)</td>
<td>G:0.63</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>5.1 (5.2)</td>
<td>2.9 (2.7)</td>
<td>5.1 (3.4)</td>
<td>T:17.66*</td>
</tr>
<tr>
<td></td>
<td>F-up</td>
<td>5.1 (5.2)</td>
<td>3.6 (5.2)</td>
<td>4.4 (4.0)</td>
<td>I:1.33</td>
</tr>
</tbody>
</table>

G = group, T = time, I = interaction effect in the ANOVA. * p < 0.05; b p < 0.005; * p < 0.001; d p < 0.0001.

The primary purpose of the present study was to compare three forms of group treatment for spider phobia: direct treatment, direct observation and indirect observation. It was predicted that on the behavioral test measures and the self-report measures of spider phobia, (1) the direct treatment (DT) would be more effective than both indirect treatments, and (2) the direct observation (DO) would be more effective than indirect observation (IO). The results show that the first prediction was corroborated to a large extent. On most measures, the DT did significantly better than the two other groups and on the remaining ones there was a trend in that direction. Furthermore, the conservative measure of clinically significant improvement clearly showed that the DT was superior: 75% fulfilled the criteria, compared to 7% for DO, and 31% for IO.

The second prediction was, however, not corroborated. Only one measure yielded a significant difference between these groups, the Spider Phobia Questionnaire, and on this IO was significantly better than DO. Most other measures showed a trend in the same direction.

The third prediction was that for the general psychopathology measures only significant improvements across the groups were expected, since there is no reason to believe that these treatments will differentially affect variables with means already within the normal range. This prediction was corroborated as the ANOVAs yielded significant time effects, but no group or interaction effects on these measures. However, a weak indication for a wider range of improvement for DT was obtained from the paired t-tests showing that the DT group changed significantly on all five measures, while DO did so on four, and IO on only two measures.

The fourth prediction was that the effects of treatment would be maintained, or that further improvement would have taken place at the 1 yr follow-up. This prediction was corroborated on all but the physiological measures, a finding which is in accordance with our previous one-session studies (Hellström & Öst, 1995; Hellström et al., 1996; Öst, 1996; Öst et al., 1991, 1992).

The second purpose was to replicate earlier studies showing that perceived self-efficacy was a better predictor of approach behavior in the behavioral test than anticipated or experienced anxiety. The results in this respect clearly showed that strength of self-efficacy was a better pre-
predictor of approach behavior than anticipatory or experienced anxiety, and it was the only variable remaining significant when the influence of the other variables was partialed out.

Before discussing these results we should compare the outcome of the direct treatment with that of the same condition in our previous study (Öst, 1996), i.e. the large group of seven or eight patients. On steps completed in the behavioral test that group changed from 5.2 to 10.1 and 10.5 at follow-up. The DT in this study changed from 5.7 to 10.1 and 10.1 at follow-up. On the self-rating of anxiety, the large group changed from 73.8 to 46.3 and 44.0, while the DT group changed from 81.3 to 51.3 and 40.3. On the assessor rating of phobic severity, the large group changed from 5.6 to 2.9 and 2.8, while the DT group changed from 5.6 to 2.3 and 1.6. Finally the proportion of clinically significant improvement was 70% at post and 75% at follow-up for the large group, while the DT group yielded 75% at both assessments. Thus, there is a striking similarity in the means of the same treatment condition across the two studies.

One methodological problem with the current study is the lack of an untreated control group. However, previous research on specific phobias reviewed by Öst (1997) indicated that out of 21 studies in which an active treatment was compared with no treatment or a wait list condition, the active treatment was significantly better in 90% of the studies. Furthermore, the few studies which have followed untreated phobics for a long period of time found a low proportion of spontaneous remission after 5 (Agras, Chapin & Oliveau, 1972) and 7 yr (Wittchen, 1991), respectively. Thus, it does not seem plausible that the 2–3 week interval between pre- and post-assessment could have led to a spontaneous remission that questions the effects of the direct treatment.

Another methodological problem is the difference in treatment time. The group treatment, by necessity, takes longer than the individual treatment, since the goal is to bring all eight patients to the same stage as is done individually. One possible way of controlling for this would be to treat two patients, one after the other, in the direct observation group. However, it is not at all possible to predict how long an individual treatment, or a group treatment for that matter, is going to take. Our experience shows that individual treatments vary from 45 min to 3 hr 15 min, and group treatments from 2 h 30 min to 3 h 30 min. Naturally, it would be possible to deviate completely from clinical practice and do the treatment for a specified time, say 3 hr, irrespective of whether it suits the patients or not.

The inferior results of the two modelling conditions are in agreement with earlier studies in snake phobia (Bandura et al., 1969; Bandura et al., 1977) and acrophobia (Ritter, 1969b). The differences between percentage of change for participant modelling and symbolic modelling in the Bandura et al. (1969) study and the present study are rather similar, and the same goes for the difference between participant modelling and live modelling in the Ritter (1969b) study and the present study.

In an unpublished study using the same design Götestam (1994) found no significant differences between the conditions. However, this may very well be due to a low power in that study, since only 8 subjects per cell were used.

What are the reasons for the worse outcome of the indirect treatments? The most probable explanation is that both exposure and modelling is more thorough in the DT. The patients in this condition have the opportunity to, and are strongly encouraged by the therapist to verbalize their catastrophic cognitions and test these out directly by exposure. They also have two types of models; the therapist functions as a ‘mastery model’ when demonstrating the various steps for the patients, and the fellow patients function as ‘coping models’ as they gradually overcome their fears. In the DO condition there is only an indirect exposure, since the patients observing the treatment never interact with the spiders directly. It is only the treated patient’s idiosyncratic catastrophic thoughts that are being verbalized and tested out. Naturally, these can coincide with those of the observers, but the DO patients can always argue that the spider could have behaved differently had they interacted with it. Furthermore, there is only one coping model (the treated patient) in this condition. In the IO condition the number of models are the same as in DO, but the exposure is quite different. The videotape is a much weaker form of exposure and the patients never had to fear that the spiders could crawl up on them.

Bandura (1977) suggested that it is the change in perceived self-efficacy that is the mechanism of change in various treatments for phobias. There are four sources of information that influ-
ence an individual's self-efficacy: enactive information, vicarious information, verbal information, and physiological information. A number of studies have shown that the most important factor is the enactive information, i.e., the individual's experience of her own achievement and performance in various critical situations. Of the three conditions in the present study, it is only the DT which gives the patients enactive information. This factor, together with the higher quality exposure in DT probably activate the fear-relevant memory structures in a better way, thus leading to a more effective emotional processing (Foa & Kozak, 1986).

At the post-treatment interview a number of the patients in the DO condition said that various steps in the treatment lead to a marked anxiety arousal that did not have time to abate before the next step was introduced. Thus a fairly high anxiety level was followed by another increase in anxiety, which made it difficult for the patients to focus on the treatment and learn from it. In the IO, on the other hand, the videotape never led to such a high level of anxiety in the first place, and the patients experienced a reduction before the videotape showed the next step.

How can the indirect treatments be improved upon? The DO condition could be slightly changed so that the anxiety levels of the observing patients are being considered before continuing with the next step. They could perhaps be asked to indicate their anxiety level (nonverbally) now and then, indicating to the therapist whether to prolong the current step, or continue to the next step. The indirect observation condition could be improved by using a number of different models, displaying various types of catastrophic thoughts and problems during the treatment. This would mean preparing a videotape that is edited to include various segments that are considered the most informative from a treatment point of view. Such a tape could also include commentary from the therapist that could be of further help for the patients in their own self-treatment after the video session. Further research regarding how to improve the cost-effectiveness of behavioral treatments in specific phobia should investigate the effects of a video modelling treatment as outlined above. Perhaps a treatment package including a versatile video and a specific manual for self-help treatment (Öst et al., 1991; Hellström & Öst, 1995) could be compared to therapist-administered treatment. Another interesting possibility would be to test even larger groups as e.g. a study by White, Keenan and Brooks (1992) on generalised anxiety disorder who used so called 'evening classes' of 20–24 patients. A related important question is how far the treatment has to go in order for the patient to function non-phobically in natural situations. In the current group treatment the goal was the same as for individual treatment, i.e. to have two spiders crawling on the patient's hands. However, we do not know if this is necessary for the patient to reach the goal in natural situations. If it is enough to handle just one, and a smaller spider, perhaps it would be possible to achieve that goal during a one-session treatment for a substantially larger group than in the current study.

The controlled studies of one-session treatment published to date involve spider phobia (Arntz & Lavy, 1993; Hellström & Öst, 1995; Öst, 1996; Öst et al., 1991), blood phobia (Hellström et al., 1996), and injection phobia (Öst et al., 1992). A recent study on flying phobia (Öst, Brandberg & Alm, 1997) showed that one session did as well as five sessions of exposure and cognitive restructuring at post-treatment. Finally, a recently completed study of claustrophobia showed equally good results. Thus, we can tentatively conclude that this is the treatment of choice for specific phobias. A possible continuation of this research on one-session treatment would be to investigate whether this method could be used for social phobia and agoraphobia, as well as the other specific phobias which it has not been tested for yet, e.g. choking phobia, dental phobia, thunder and lightning phobia, and vomiting phobia.

Acknowledgements—This research was partly supported by Grant 05452 from the Swedish Medical Research Council.

REFERENCES


Brown, DiNardo & Barlow. 1994. Author please provide details.


