Social phobia is an anxiety disorder that remained practically unidentified until the mid-1980s (Heimberg, 1989a; Liebowitz, Gorman, Fyer and Klein, 1985), despite attaining prevalence rates of between 3% and 13% (Kessler, McGonagle, Zhao et al., 1994).

This phobic disorder is usually complicated by work absenteeism, drug and/or anxiolytics abuse, alcoholism and depression (Barlow, DiNardo, Vermilyea and Blanchard, 1986; Bowen, Cipywnyk, D’Arcy and Keegan, 1984; Chambless, Cherney, Caputo and Rheinstein, 1987; Higgins and Marlatt, 1975; Kushner, Sher and Beitman, 1990; Schneier, Martin, Liebowitz et al., 1989). In some cases these problems are the expression of an undiagnosed social phobia, so that the prevalence of this clinical condition may be greater than estimated (Stravynski, Lamontagne and Lavallee, 1986).

The central characteristic of social phobia is excessive and persistent fear of social situations in which the patient is exposed to the observation or scrutiny of others (American Psychiatric Association 1994). These social fears can attain diverse degrees of generalisation. The DSM-IV distinguishes a generalised subtype of social phobia that is applicable to those people who fear the majority of social situations. However, Heimberg, Holt, Schneier et al. (1993) distinguish two additional subtypes: circumscribed subtype, applicable to those who fear only one or two discrete situations, and non-generalised subtype, applicable to those that, demonstrating adaptive functioning in some social areas, feel anxiety in a minimal number of interactive situations. Social phobia presents a high level of comorbidity with other disorders of axis I (Brewerton, Lydiard, Ballenger and Herzog, 1993; Bulik, Beidel, Duchmann and Weltzin, 1991; Disalver, Qamar and Del Medico, 1992; Schneier, Martin et al., 1989; Schwalberg, Barlow, Alger and Howard, 1992; van Amerigen, Mancini, Styan and Donison, 1991) and axis II (Herbert, Hope and Bellack, 1992; Holt, Heimberg and Hope, 1992; Turner et al., 1992).

The most studied cognitive-behavioural treatments are social skills training, exposure techniques, cognitive restructuring techniques and the combination of exposure and cognitive restructuring. Relaxation techniques, systematic desensitisation and training in coping with anxiety have been studied to a lesser extent. Exposure techniques are considered as the preferred treatment for...
phobic disorders (Echeburúa, 1990; Echeburúa and Salaberría, 1991; Marks, 1991), or at least as a fundamental component of any effective therapeutic package (Echeburúa, 1993; Heimberg, 1989b; Heimberg and Juster, 1995), though their application to social phobia is quite recent (Echeburúa and Salaberría, 1991). Also, there is a wide consensus on the fact that cognitive mediation is present in social phobia more than in the rest of anxiety disorders (cf. Beck and Emery, 1985; Butler, 1989; Heimberg and Barlow, 1988), though the application of cognitive restructuring techniques has not provided conclusive results (Echeburúa, 1993; Heimberg, 1989b; Heimberg and Juster, 1995; Salaberría and Echeburúa, 1995). Feske and Chambless (1995) affirmed, after a meta-analysis of 21 studies, that exposure techniques are equally effective applied alone or in combination with cognitive restructuring techniques. Taylor (1996), whose meta-analysis included 42 studies, found that all the treatments considered in his work — placebo pills, exposure, cognitive restructuring (without exposure exercises), integrated combination of cognitive restructuring and exposure, social skills training — gave effect sizes superior to those of the waiting list condition, but that only the combination of cognitive restructuring and exposure produced better results than those of the group treated with placebo pills.

The aim of the present meta-analysis is to estimate the effectiveness of exposure techniques, cognitive restructuring techniques and social skills training, and combinations of them, in social phobia patients. Furthermore, we shall identify the variables that moderate therapeutic effectiveness in this phobic disorder. The main differences between the present study and the two previously-published meta-analyses are that our work:

a) Includes all studies that examine exposure techniques, social skills training, cognitive restructuring techniques and their combinations in a design with at least pretest-posttest data (this implies the inclusion of, for example, groups that receive only cognitive restructuring without exposure, or that receive cognitive restructuring after exposure).

b) Excludes groups treated with placebo pills, exclusively or in combination with the cognitive-behavioural treatments considered, given the confusion caused, for our purposes, on comparing a group with only psychological treatment with others that add the expectations of receiving pharmacological treatment.

c) Includes studies carried out with Spanish populations.

d) Does not treat work with waiting list groups as independent studies, as is the case in Taylor (1996).

e) Includes calculation of effect sizes differentiated for between-groups and within-group designs.

f) Includes all types of measure directly related to the descriptions of clinical conditions in the nosological systems employed (e.g., scales applied by interviewers, ‘Fear of Negative Evaluation Scale’ or ‘Social Avoidance and Distress Scale’).

g) Excludes measures not directly related to the descriptions of clinical conditions in the nosological systems (e.g., irrational beliefs or depressive state), since studies differ in the assessment of type and number of additional constructs analysed. It is our view that, in principle, the inclusion of constructs not considered in the definition of social phobia may generate an additional heterogeneity that would lead to confusion as regards the magnitude of the differences between treatments.

h) Employs data-analysis techniques based on weighted least squared. Currently, classical statistical techniques, which do not weight studies according to their precision, are not recommended in meta-analyses (Cooper and Hedges, 1994).

METHOD

Distinction between “research report” and “study”

In the meta-analyses carried out we distinguish between “research report” and “study”. By research report we understand the framework employed to report the results of one or more comparisons of groups in relation to the effectiveness of psychological treatments for social phobia. By study we understand the comparison between a group receiving psychological treatment and a control group. In those cases where a treated group was not compared with a non-treated group, but an assessment was made of the therapeutic gain comparing a posttest measure with a pretest measure for the treated group, we also considered it as an independent study.

Inclusion/exclusion criteria for research reports

With the aim of homogenising the empirical sample and applying the meta-analytic technique in an appropriate way, we established the following conceptual and methodological criteria for including or excluding a research report:

a) The research report should examine the effectiveness of exposure techniques, cognitive restructuring techniques, social skills training or a combination of these treatments.
b) Subjects treated should receive a diagnosis of social phobia.

c) The concept of social phobia employed in the research report should coincide explicitly with the definitions included in the ICD nosological system (World Health Organization, 1978, 1979, 1992) or that of the DSM (American Psychiatric Association, 1980, 1987, 1994).

d) The date of publication (or date of carrying out, if the study is unpublished) should fall within the period 1980-1997 inclusive. The beginning of this period is determined by the date of publication of the third edition of the Diagnostic and Statistical Manual of Mental Disorders.

e) The research should have a group design, with at least 5 experimental subjects.

f) The research report should provide sufficient data to allow estimation of effect sizes.

g) Research reports that do not present pretest data were excluded.

h) Also excluded were reports that employed single-case design methodology, since it is impossible to combine in a quantitative manner the results of group designs with those of single-case designs.

Literature search

In order to minimise selection biases, we used various processes in the search for research reports:

a) Computerised search. Using the databases PsycLIT, MEDLINE and PSICODOC. The descriptors employed were: “social phobia”, “social anxiety”, “treatment” and “therapy”.


c) Review of bibliographical references of the research reports already located, as a source of previous primary studies.

d) Request for papers from experts, published and unpublished.

The literature search carried out allowed us to locate 25 research reports that fulfilled the selection criteria, providing a total of 39 studies.

Coding of studies

In order to define operationally the variables to be coded we prepared a manual with specific guidelines and a coding protocol. The quality of the coding process was studied through the selection of a random sample of 33% of the total reports found, which was coded by two independent coders.

The variables whose possible moderating effect was examined were classified, following Lipsey (1994), in three large sections (see Table 1):

a) **Substantive variables**: these are intrinsic to the scope of the research being meta-analysed. They include: treatment, subject and context variables. For example, type of psychological treatment applied.

b) **Methodological variables**: deriving from aspects related to the research design and methodology. For example, type of design used.

c) **Extrinsic variables**: related to neither the methodology nor the scope of the research. Although such variables should not affect the results, they may occasionally be relevant. For example, date of the study.

**Calculation of effect size**

The index of effect size (ES) used is the standardised mean difference $d$ (Hedges and Olkin, 1985), considering the following definitions according to the nature of the design used in the assessed study: (a) For within group designs (pretest-posttest), $d$ is defined as the difference between the mean of the pretest and the mean of the posttest divided by the overall within-group standard deviation.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Moderator variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Treatment variables</strong>:</td>
<td>1) Psychological treatment employed, 2) Gender of professionals, 3) Experience of professionals, 4) Home tasks of the treatment, 5) Follow-up programme of the treatment, 6) Utilisation of agents external to the therapeutic group, 7) Therapeutic contract, 8) Training modality, 9) Number of therapists, 10) Number of treatment sessions, 11) Duration of treatment, 12) Intensity of treatment, 13) Absenteeism.</td>
</tr>
<tr>
<td><strong>Subject variables</strong>:</td>
<td>1) Age, 2) Gender, 3) Socio-economic level, 4) Educational level, 5) Diagnostic system employed, 6) Social phobia subtype, 7) Mean duration of the social phobia, 8) Comorbidity, 9) Presence of previous treatments, 10) Type of previous treatments (psychological, pharmacological, or both).</td>
</tr>
<tr>
<td><strong>Context variables</strong>:</td>
<td>1) Continent, 2) Country.</td>
</tr>
<tr>
<td><strong>Methodological variables</strong>:</td>
<td>1) Subject recruitment mode, 2) Measures to preserve integrity of the treatment, 3) Type of design, 4) Random assignment to groups, 5) Type of control group, 6) Total sample size, 7) Sample size of treatment group, 8) Sample size of control group, 9) Experimental mortality, 10) Design quality, 11) Months of follow-up, 12) Number of dependent variables.</td>
</tr>
<tr>
<td><strong>Extrinsic variables</strong>:</td>
<td>1) Technical qualifications of first researcher, 2) Form of presentation of the study, 3) Range of publication, 4) Date of presentation.</td>
</tr>
</tbody>
</table>
deviation (or, failing that, the average of the standard deviations of the pretest and posttest); (b) For between-groups designs (both experimental and quasi-experimental), with pretest and posttest measures, $d$ is defined as $d = d^e - d^c$, with $d^e$ and $d^c$ being the standardised mean differences between the pretest and posttest of the experimental and control groups, respectively. A $d$ value above zero indicates a beneficial effect for the subjects of the treated group, while a $d$ value below zero indicates a detrimental effect.

With the aim of optimising the comparison between meta-analysed treatments, we maximised the homogeneity of the ESs. We selected only the dependent variables directly related to the definition of social phobia, that is, fear, avoidance and/or uneasiness in social situations, which implies deterioration in social, work and/or academic adaptation, and excluded the rest (trait anxiety, depressive mood, self-esteem, locus of control, irrational beliefs, etc.).

For each study we calculated a maximum of two ESs, one for the posttest assessment and another for the assessment of the longer follow-up. For each point in time we averaged the $d$ values derived from the dependent variables that coincided with the established criteria.

**Statistical analysis techniques**

We calculated a confidence interval around the mean ES in order to estimate the population effect size and whether that effect was significantly different from zero (null effectiveness). We also applied a $\chi^2$ test of homogeneity of all effect sizes around the mean ES (Hedges, 1994; Hedges and Olkin, 1985).

Given that the homogeneity test may be less powerful when applied to a small group of studies, as in our case (cf. Sánchez-Meca and Marín-Martínez, 1997), we decided to check the influence of variables that may theoretically be moderating the results, even though the homogeneity test was not found to be statistically significant. Following the meta-analytic approach of Hedges and Olkin (1985), for the qualitative variables we carried out analyses of variance weighted by the inverse of the variance of each ES. For the quantitative variables we applied simple regression analyses weighted by the inverse of the variance of each ES.

**RESULTS**

**Reliability of the coding**

Inter-coder reliability for the moderator variables attained the values established by Orwin (1994) for guaranteeing coding reliability (a value of at least 0.80 for agreement rate and for Pearson’s correlation coefficient, and a value of at least 0.70 for Cohen’s $Kappa$ and the intra-class correlation). The reliability of the calculations of effect size (intra-class correlation and Pearson’s correlation) was greater than the value of 0.98 for the posttest and follow-up, indicating an excellent consensus in the selection of dependent variables to be included in the calculation of the effect size.

**Study of publication bias**

All of the studies included in the present meta-analysis are contained in research reports published in specialist journals, so that a possible threat to the validity of the results obtained derives from a potential publication bias, that is, that the publishers may be uneven in their treatment of the studies presented for publication, as a function of the statistical significance reported.

Following Orwin (1983), we calculated the “index of tolerance of null results”. According to this index, there would have to be more than 180 unpublished studies (and not considered by the meta-analyst) filed away with the publishers of the journals for the results of our meta-analysis to be invalidated. We can therefore conclude that it is highly improbable that our results are affected by the publication bias.

**Meta-analysis in the posttest and follow-up**

Five studies provide data only from the posttest and 34 from the posttest and follow-up. The median of the time interval between posttest and follow-up was three months.

The mean effect size (weighted by the inverse of the variance) was 0.769 in the posttest and 0.953 in the follow-up, the parametric values being far from the null value. These results allow us to affirm, in global terms, that the cognitive-behavioural treatments reviewed are clearly effective for social phobia. Moreover, considering the orientative classification proposed by Cohen (1988), the mean value obtained in our meta-analysis approaches a high magnitude in the posttest ($d = 0.80$), and a higher one still in the follow-up.

The homogeneity test was not found to be statistically significant [posttest: $Q_T(38) = 19.163$, $p > .05$; follow-up: $Q_T(33) = 25.318$, $p > .05$], so that we can assume homogeneity of the different studies among themselves, despite being derived from studies that differ in a large number of characteristics, both substantive and methodological; the parametric values were representative of those found in the meta-analysed studies.
A more intuitive interpretation of the effect sizes found can be obtained through their transformation into a correlation coefficient ($r = \frac{d}{\sqrt{d^2 + 4}}$) in order to construct the binomial presentation of the effect size, BESD (binomial effect size display) proposed by Rosenthal (1991). A correlation coefficient of 0.36, for example, is equivalent to an improvement rate of 68% in the treatment groups, as against a rate of only 32% in the control groups, representing a differential rate of 36% between the two groups.

According to the model of meta-analysis applied in our study, the fact of obtaining non-significant results of the homogeneity tests should have put a halt to the analysis, precluding the search for possible moderator variables of the effect sizes found in the empirical studies. Put another way, these results should have led us to the conclusion that the studies meta-analysed here present homogeneous effectiveness in the psychological treatment of social phobia and, consequently, the mean effect sizes obtained, and their confidence intervals, represent in a valid way the set of studies in the meta-analysis. In fact, these results concur, in general terms, with those obtained in the meta-analyses by Feske and Chambless (1995) and by Taylor (1996).

Nevertheless, this conclusion may be seen as simplistic, since the non-significant result obtained in the homogeneity test can be interpreted in different ways (cf., e.g., Hall and Rosenthal, 1991). In fact, the homogeneity test may be non-significant due to a lack of statistical power, especially with a fairly small number of studies, as in our research, $K=39$ (cf., e.g., Harwell, 1997; Sánchez-Meca and Marín-Martínez, 1997).

It is interesting to note that the differences found between the values of effectiveness for the different studies fail to reach statistical significance as a function of: (a) the cognitive and/or behavioural technique employed; (b) the diagnostic system used for diagnosing the patient’s anxiety disorder; (c) presence of the generalised subtype of social phobia; (d) the mean duration of the clinical condition; (e) the presence of previous treatments; and (f) the type of design.

Table 2 shows the mean ESs obtained ($d_+$) for the treatments considered in our meta-analysis. Also included are the limits of the confidence interval at 95% ($Li$; $Ls$), the number of studies from which the ES is derived ($K$) and the sample size accumulated for each set of studies ($N$).

### Differences between posttest and follow-up

In order to examine the changes in effectiveness that occurred between the posttest and the follow-up, we selected the 34 studies that provided data in posttest and follow-up. In general, effect sizes tended to decrease with respect to the posttest in the follow-up assessments, a finding usually attributed to a fading of the therapeutic effect of the treatment studied. However, in our research we found, as did Taylor (1996), that effect sizes were greater in the follow-up (as against the posttest), and that this difference was marginally significant [$T(33) = -1.946, p = .06$].

Nevertheless, a possible threat to the internal validity of the relationship between increase in therapeutic effectiveness and the point at which assessment was made may be constituted by selective mortality. In order to analyse the relationship between experimental mortality and magnitude of therapeutic effect, we defined two variables: (a) Increase in posttest-follow-up experimental mortality ($\Delta M_{\text{post-flwup}}$), as the difference between pretest-follow-up mortality and posttest-follow-up mortality for each study; and (b) increase in effect size posttest-follow-up ($\Delta T_{E\text{post-flwup}}$), as the difference between $ES$ in the follow-up and $ES$ in the posttest for each study.

Through the construction of a simple regression model,
we found that $D_{M_{post-flwup}} \text{ was related to } \Delta T_{E_{post-flwup}}$ $[F(1, 22) = 11.537, p = .003]$, explaining 31.4% of its variance. However, this regression model appeared to be affected by an outlier. The result in question belongs to a study included in the report by Mersch et al. (1995), and derives from the follow-up carried out with 3 of the 7 patients that completed the posttest assessment, giving an experimental mortality of 62.5%. After elimination of this outlier the relationship no longer showed statistical significance $[F(1, 21) = 0.997, p = .329]$.

**DISCUSSION**

Considering in a global way the effectiveness of social skills training, exposure techniques and cognitive restructuring techniques, we can conclude that they are fairly effective. Smith, Glass and Miller (1980), despite the methodological and content differences between their study and our own, provide an alternative practical illustration of the meaning of effect size obtained. They point out that nine months of work to teach primary school children to read translates into an effect size of 0.67, showing that the effectiveness of the treatments analysed is clearly substantial for social phobia patients.

The studies that examine the effectiveness of multi-component therapeutic packages — made up of exposure techniques plus social skills training or cognitive restructuring techniques — are not superior to the studies that employ only exposure techniques, even though the number of subjects treated with exposure, alone or in combination with others, is greater than that of subjects treated with cognitive restructuring or social skills training.

This fact is interpreted by some authors as representing a lack of support for the therapeutic principles underpinning social skills training and/or cognitive restructuring techniques in the treatment of social phobia (Feske and Chambless, 1995; Hope, Heimberg and Bruch, 1995; Mattick and Peters, 1988; Mersch, 1995; Scholing and Emmelkamp, 1993; Stravynski, Marks and Yule, 1982). In this regard, the absence of differences between treatments would derive from the fact that the different techniques possess common therapeutic elements that are effective in the treatment of phobic disorders. It is probable that exposure to feared social stimuli has been an element shared by the treatments applied to patients in the different studies. In fact, social skills training includes exposure *in vivo* to the phobic stimuli.

However, in the studies that examine cognitive restructuring techniques and that explicitly controlled the exclusion of elements of exposure, we found ESs similar to those found for exposure techniques applied exclusively (ES for posttest and follow-up, respectively: 0.61 and 1.28, in Emmelkamp, Mersch, Vissia and van der Helm, 1985; 0.98 and 1.77, in Mattick, Peters and Clarke, 1989; 0.59 and 1.05, in Mersch Emmelkamp, Bögels and van der Sleen, 1989). This could lead us to the conclusion that the effectiveness of cognitive restructuring techniques is not related to the inclusion of an explicit component of exposure to the phobic stimuli.

Given that we are considering a complex phobic condition and a quantitative improvement criterion, an alternative hypothesis may reside in the existence of multiple mechanisms of action through which both exposure techniques and cognitive restructuring techniques may be effective, each type of technique acting on different aspects of the social phobia (e.g., exposure techniques on avoidance behaviour and high vegetative activation, and cognitive restructuring techniques on dysfunctional cognitive content). In this case, though, the combination of the two techniques should prove superior to each technique used alone (since the patient is treated in multiple areas), and this is not the case. Nevertheless, we might ask ourselves whether response to treatment depends on the phobic profile of the patient. According to this hypothesis, patients with a predominantly cognitive response would benefit more from cognitive restructuring procedures, whilst patients with a predominantly physiological and motor response would benefit more from exposure to the phobic stimuli. The meta-analysis carried out does not permit the testing of this hypothesis, though the results of primary studies designed ad hoc do not support the hypothesis of customised treatment, at least for systematic desensitisation, social skills training, applied relaxation, Ellis’s rational emotive therapy and training in self-instructions (Jerremalm, Jansson and Öst, 1986; Mersch, Emmelkamp, Bögels and van der Sleen, 1989; Trower et al., 1978).

For Marks (1991), the equivalence of cognitive treatments and exposure techniques is due to the fact that cognitive restructuring techniques implicitly include a component of exposure to the feared social stimuli (cognitive exposure). Nevertheless, we believe it a little contrived to consider that the Socratic dialogue (discussion of beliefs such as “I should be perfect and not make mistakes”, or “It would be terrible if no-one loved me”) without explicit prescription of exposure (or self-exposure) tasks could constitute an element of exposure.

In our study we found no statistically significant differences between the studies according to administration format of the treatment (individual vs. group), although
the tendencies in the ESs are similar to those obtained by Moreno (1999), who found that the most effective treatments had been administered in group format. If we were to find a real difference in favour of group treatments, we could minimise the threat that the effectiveness of the treatments studied resides in what Frank (1988) called “common elements of psychotherapy”. This conception encompasses, for example, the therapist-patient relationship, the explanation to the patient of a conceptual scheme for understanding his/her problems (a “myth”, in this author’s words, since its validity is presupposed), or the provision of a therapeutic ritual to the patient (that is, a series of rules, techniques and exercises which, used correctly lead, a priori, to “cure”). Some of these elements are more typical of clinical contexts than contexts of research into therapeutic results.

Given the structure of the groups treated with these techniques, confirmation of the tendency in favour of treatment application in group format would provide empirical support for the principle of exposure to phobic stimuli as an explanation of the improvement achieved, regardless of the specific treatment imparted. It is pertinent to point out that the fact of imparting a treatment through a series of group sessions of two hours’ duration could convert virtually any treatment (even that in which the explicit prescription of exposure tasks is avoided) into a treatment of exposure to ansiogenic stimuli with sufficient duration to produce habituation, according to the parameters generally recommended for producing therapeutic exposure sessions (Marks, 1991). In fact, cognitive treatments without explicitly implemented exposure techniques were carried out in group format with series of 6 to 8 sessions of 120 to 150 minutes each. It would remain to weight the therapeutic effect of the inherent exposure to that situation (exposure not programmed or defined as such to the patient, in one group of 4 to 7 persons with supposedly equivalent problems, and over a period of 20 hours divided into 8 weekly sessions).

The observed increase in effectiveness of treatments between posttest and follow-up is an unusual phenomenon in the scientific literature, even if it may be due, as Taylor (1996) points out, to the application of additional treatment for those patients that require it. A possible alternative explanation, ruling out the phenomenon of selective mortality, is that, even though researchers may administer additional treatment to patients that suffer a relapse, it is also sometimes the case that patients who manage to confront the feared social stimuli (the reader will recall that exposure techniques have mainly been applied with tasks to carry out at home) may continue implementing self-exposure techniques by themselves in their own social environment, just as prescribed by the therapist during the treatment phase. Self-exposure may be as effective as other variants of exposure techniques (Marks, 1991, p. 156), and may constitute the basis of the therapeutic gain that tends to be found in some studies between posttest and follow-up. Nevertheless, in patients that received cognitive restructuring techniques without (explicit) exposure tasks, an additional improvement is likewise observed in the follow-up, and which is difficult to explain from the hypothesis of action mechanisms with a basis in exposure.

REFERENCES


