BACHELOR THESIS CLINICAL PSYCHOLOGY

DEPARTMENT OF PSYCHOLOGY, SOCIAL SCIENCES FACULTY

UNIVERSITY OF AMSTERDAM

Effectiveness of website-based therapy versus e-mail-based and face-to-face therapy in the treatment of Depression and Anxiety Disorders

Markus van Alphen (0452017)

KP/ BKT12-41/1 June/2006 – 12 ECT

Supervisor**: Prof. Dr. A. Lange**

Abstract

What exactly the effective ingredients for change are in an intervention is still a subject of research. One of these ingredients is the client-therapist contact. A question is, whether an intervention is any less effective when the face-to-face element of the contact is filtered out. To answer this question, the literature on effect sizes for various website-based and e-mail-based interventions for the treatment of populations with clinical levels of depression and anxiety is investigated. However, no studies were found in which the material was delivered by e-mail. The general conclusions drawn from website-based studies are 1) that website-based interventions seem equally efficacious as face-to-face interventions; 2) a therapeutic relationship is created in a website-based intervention, albeit a different relationship than is created in traditional face-to-face settings; 3) The amount and nature of the therapist contact is of very direct influence on the effect size; 4) the medium used for contact by the therapist is of lesser importance than the message; and 5) structural elements are important.

Effectiveness of website-based therapy versus e-mail-based and face-to-face therapy in the treatment of Depression and Anxiety Disorders

Table of Contents

1 Introduction

Pressure from society in the form of government, government policy, governmental agencies dealing with the issue of health (and the financing thereof) and health insurance companies to deliver cost-effective as well as efficacious treatment of mental health problems is ever mounting. Professionals no longer have the liberty to use their preferred approach but are required, wherever possible to use protocol-driven, scientifically proven-to-be-effective treatments. In the treatment of depression and anxiety it is cognitive behavioural therapy that to date carries most support for its effectiveness (for example: Lambert, 2004).

However, the question as to what exactly the effective ingredients for change are is still a subject of research. Common factor research seems to indicate that the classical Rogerian variables of warmth, unconditional positive regard and congruency are major contributors (Trijsburg, Colijn & Holmes, 2005). The question is, nonetheless, whether client-therapist contact is any less warm or congruent when the face-to-face element of the contact is filtered out. Preliminary studies seem to indicate that interventions without face-to-face contact are equally effective as those with face-to-face contact (Proudfoot, 2004), in any case when considering interventions that may be delivered according to a fixed protocol.

The implementation of face-to-face versus non-face-to-face intervention comparison is not as simple as the above might suggest. Several factors can be identified up-front as important. Screening, as first example. What are the inclusion and exclusion criteria for a client to be considered for either form of therapy? When these are not similar, the comparison is irrelevant. If initial screening and perhaps also final assessment is performed face-to-face, does this also have a therapeutic effect that is different than when screening and assessments are performed on-line?

When considering the type of non-face-to-face intervention, several different approaches can be identified. Is the intervention delivered on a stand-alone computer, by e-mail, over the telephone, via a chat-window, online via a web site or any combination of the above, for example? How much client-therapist contact do the various implementations provide? Is there a timed delivery of the various sessions with fixed milestones and how are users of these services motivated not to drop out?

As to the legal and ethical issues (privacy, for example), we acknowledge their presence but do not intend to discuss these aspects within the ambit of this study.

In order to come to the research question for this paper, we first need to consider how to quantitatively compare face-to-face interventions with non-face-to-face interventions. The choice here is the effect size: The number of standard deviation units improvement in the post-treatment scores versus the pre-treatment scores (normally a reduction in the level of symptoms), usually referred to as *Cohen's d*. By using effect size one is able to make a comparison across studies irrespective of the dependent measure used, assuming of course that the instrument with which the dependent measure is reported provides valid measurements. Generally, effect sizes between 0.2 and 0.5 may be interpreted as medium, 0.5 to 0.8 as medium to large and effect sizes larger than 0.8 may be regarded as large (Cohen, 1988). In calculating the effect size, the manner of dealing with dropout is commonly dealt with in one of two ways: case-wise exclusion or using the intention-to-treat paradigm. The former method is simple: Data from dropout is simply omitted in the analysis. This can paint an overly optimistic picture of the results, as only those who actually completed the intervention (and are perhaps the most likely to benefit from the treatment) are included in the calculation of effect size. The intention-to-treat paradigm attempts to

correct this by asking dropouts to complete post-treatment assessment anyway, irrespective how much of the intervention they had completed. Furthermore, those individuals who do not return their post-treatment assessment, whether they had dropped out or had completed the intervention, their latest (usually pre-treatment) scores are carried forward to post-treatment. That is, they are assumed not to have changed any further since the last assessment. This results in a more conservative calculation of effect size.

The second issue is pragmatic: In order to limit the size of the paper, only interventions for depression and anxiety are considered. For one, because for both these categories there are several well-documented studies reporting effect sizes for face-to-face therapies, especially using Cognitive Behavioural Therapy (CBT). Secondly, most non-face-to-face interventions to some degree use the principles of CBT, making for easier comparison. It is after all the effect of the face-to-face element of the therapist contact that we are trying to vary across conditions, as it were. Obviously, random controlled trials (RCT) where only the quantity, nature or quality of therapist contact is varied would provide the best data.

A third issue concerns relevance: From the point of view of interventions, the reduction in depression or anxiety levels for individuals with initially only marginally heightened scores with respect to the general population, does not really say much about the efficacy of the intervention. Although such persons might benefit from CBT, as the initial levels are marginally heightened, the possible improvement is limited, perhaps resulting in a floor effect. Moreover, these are also not the people who would ordinarily seek the services of a therapist. A useful comparison can only be made when the populations studied are a reflection of the population requiring help, that is, the clinical population or those with clinically relevant levels. Therefore we limit

ourselves further to studies containing populations with clinical levels of depression or anxiety, or those from which data pertaining to this clinical (sub)population can be extracted.

A last issue concerns the means of delivery: Since around 2000, it is especially the Internet and Internet-delivered therapies that have aroused interest. Not only has the Internet become a mature technology, but it has also become one of the most widespread media forms available to an ever-increasing population. The World Wide Web and e-mail have drastically changed the world and the way in which we communicate. A natural extension of this trend is to make the therapeutic process virtual too. In a sense the division of studies into those that are website-based or e-mail-based is a little arbitrary, as many of the interventions reported use a combination of both. In order to be able to differentiate, where the actual material is provided on a web site the intervention is regarded as website-based. Obviously, the amount and type of e-mail contact remains an important variable. However, we can already note at this stage that no studies were found in which the material was delivered by e-mail. This does not mean that this kind of intervention does not exist, but only that no one has evaluated such an intervention. The question is, of course, would an intervention delivered by e-mail have any advantages over and above an intervention via a website? Privacy issues (which we were not going to discuss in this paper) perhaps? About the efficacy of e-mail based interventions we can be brief: We simply do not know whether or not they are efficacious.

To sum up, we will investigate the literature on effect sizes found for various website-based interventions for the treatment of populations with clinical levels of depression and anxiety.

We commence by reporting the dearth of scientific studies published, first on the treatment of depression and then on the treatment of anxiety. Subsequently we discuss these results and attempt to discover factors that affect the efficaciousness of the interventions. Finally we conclude this paper with recommendations that may be put to the test in future trials of website-based and e-mail-based interventions.

2. *Studies on effectiveness: Treatment of Depression.*

Christensen, Griffiths, and Korten (2002) reported reductions in depression and anxiety scores, measured with the Goldberg scales (cited by authors: Goldberg, Bridges, Duncan-Jones & Grayson, 1988), for some 1500 registrants to the MoodGYM (more about MoodGYM in the following paragraphs) site. Unfortunately, these results do not provide clear-cut effect sizes that are interpretable within the framework of this paper. Fortunately, the MoodGYM program was also evaluated in a randomised controlled trial, which we will now report on without further ado.

Christensen, Griffiths, and Jorm (2004) randomly assigned 525 individuals with increased depressive symptoms to one of three groups: A website-based CBT program, a website with information or a control (attention placebo) condition (three-way pre-post design). Subjects were acquired via a mailing to a random sample of 27000 persons aged 18 – 52 years living in Canberra, Australia. Of the 6122 persons who responded, 722 had a score of 22 or higher on the Kessler psychological distress scale (cited by authors: Andrews & Slade, 2001), had access to the internet, were not getting professional help and were willing to participate. Only 525 persons returned

consent forms and pre-intervention questionnaires and were randomised to the various intervention groups.

The three conditions entailed the following:

- The website-based CBT program MoodGYM (see http://moodgym.anu.edu.au/, note, however, that the MoodGYM program has subsequently been upgraded and is now called MoodGYM Mark II). This intervention consisted of five interactive modules with weekly assignments, with revision of the five modules in the sixth session. Each session was made available sequentially at one-week intervals, giving a total intervention time of six weeks. As to the nature of the CBT-program, the authors do not specify of what the intervention consisted and remain vague by saying that it was a "cognitive behavioural therapy" intervention. From a previous publication (see above; Christensen, Griffiths, & Korten, 2002) and assuming the program had not significantly changed, the five modules of MoodGYM dealt with 1) the principles of CBT, 2) dysfunctional thoughts and overcoming them, 3) behavioural methods to overcome dysfunctional thinking, assertiveness and self-esteem training, 4) stress, relaxation, pleasant events and parenting styles and 5) problem solving and dealing with relationship break-up. Lay interviewers called each participant weekly by telephone to direct the use of the website, but there was no therapist contact whatsoever. Of the 182 individuals assigned to this group, 46 (25%) dropped out.
- The information group accessed a web site providing depression literacy and evidence-based information on the treatment of depression. This group was also contacted weekly by telephone by lay interviewers. As for the

CBT group, there was no therapist contact whatsoever. Of the 165 individuals assigned to this group, 25 (15%) dropped out.

- Individuals in the control group were also called weekly, but whereas in the CBT and Information groups the contact was strictly about navigating the web site, these conversations were directed at discussing lifestyle and environmental factors that could have an influence on depression. The idea behind this form of "attention placebo" control group is to factor out the possibility that the attention given to the individuals in one of the active treatments caused the improvement in scores at post-treatment (the so-called 'Hawthorne effect'; Arnold, Cooper & Robertson, 1998). Of the 178 individuals assigned to this group, 19 (11%) dropped out. From the report it is not clear whether subjects in the control group were offered access to the CBT intervention (or any other form of treatment) after completion of the program.

Of the 525 individuals randomised, 79% completed the intervention and 83% returned post-intervention questionnaires. The dropout rate was significantly higher for the CBT condition than for the other two conditions. Unfortunately the authors do not report attrition analysis across conditions other than in terms of number of cases.

The dependent measure for this study was the Center for Epidemiologic Studies depression score (CES-D; cited by authors: Radloff, 1977). A score of 16 or more may be regarded as indicating moderate clinical depression, 28 or more severe clinical depression. Of the 525 individuals randomised, 369 individuals had a CES-D score of 16 or over. The effect sizes relative to baseline for these individuals, based on the intention-to-treat paradigm, were 0.9 for the CBT condition, 0.75 for the Information condition and 0.25 for the Attention placebo condition, showing a

preference for the CBT variant. The authors concede that these results are post-treatment. The true acid test would require a follow-up after a year, say, to ensure that improvements are maintained.

The question might rightly be raised to what extent the control group could be regarded as an attention placebo. Individuals were telephoned and lifestyle and environmental factors that could have an influence on depression were discussed. This seems to approximate a form of depression literacy improvement or can even be seen as an intervention in its own right. Probably, especially as the intervention and therefore the waiting period was of such short duration, a wait-list control with a weekly phone call to inform that participation in the intervention would begin shortly would have been a better implementation. The medium effect size (0.65 post-treatment for CBT) is slightly less than the range from 0.7 to 1.2 that the authors report as typical effect sizes for brief cognitive therapy. Some others points should be borne in mind. Firstly, although the authors report fairly high completion rates, on average only half of the assignments were actually completed by subjects in the CBT condition. This raises the question as to how far subjects in the CBT condition could be regarded as having undergone CBT. Perhaps, as so little of the material was actually worked through, the CBT condition might be more comparable to the envisaged information condition – a surmise that finds support in the comparability of outcome sizes (only 0.15 standard deviation units difference) for these two conditions. Secondly, subjects in the information condition only accessed the site 4.5 times on average, but this does not reflect how much time they actually spent reading the material presented. The improvement in psychological depression literacy, for example, for the information group is only 0.4 standard deviation units more than for the CBT group (and 0.7 more than the control group), which seems to imply that site material was read to some

extent, but to what extent is a matter of speculation. On a positive note, it is pleasing to see that subjects were drawn from the general public, not students nor a population already in the clinical circuit. This improves the possibility to generalise results to the population that is best served by Internet based therapy (disregarding cost issues): Those who for whatever reasons cannot or will not visit a therapist for ordinary face-to-face therapy.

Clarke et al. (2005) randomly assigned 255 individuals to one of three conditions: Treatment as usual (TAU) control group, a self-help CBT website with postcard reminders or a self-help CBT website with telephonic reminders in a three-way pre-post design. As no significant differences were found between the results for the two types of reminders (and this being the only difference between these two intervention groups), the results for these two conditions are combined here for the purposes of our report.

Subjects were drawn via a health maintenance organization (HMO) in the USA. Of the approximately 440000 members, 6030 had a chart diagnosis of depression and had received depression medication or psychotherapy in the 30 days prior to selection (What exactly a chart diagnosis means is not entirely clear). A second age and gender-matched group of 6021 “non-depressed” individuals was additionally selected. These 12051 individuals were approached by mail and 291 responded. Of these, 255 participants, of which 55 (22%) were from the “non-depressed” group, were randomised into the three conditions: 100 into TAU, 75 CBT with postcard and 80 CBT with telephone reminders.

The intervention groups were directed to a website-based program (see www.feelbetter.org) for training in cognitive restructuring without any therapist

contact. The program did not contain any behavioural therapy or behavioural activation techniques. The training comprised seven chapters of material adapted from CBT manuals, augmented with interactive examples and tutorials. There was no time limit set for completion of a particular chapter. The intervention group was reminded by non-clinical staff two, eight and 13 weeks after enrolment, either by telephone, or by postcard. The nature of the reminder was to direct the participant to the web site by giving them the URL (web address), assisting them in retrieving forgotten passwords and telling them about a feature of the program so that they would be enticed to visit the web site, thereby to continue with the program.

On-line assessments were performed at baseline and at 5, 10 and 16 weeks after enrolment, for which they were rewarded with a gift certificate of US$ 5, 10, 15 and 20 respectively. However, post-baseline assessments were still only filled out by 64%, 68% and 66% respectively. The authors report that they dealt with missing data by using restricted maximum likelihood estimation, a method that uses the slope of the data per subject to estimate missing data, rather than carrying forward last observations. This method is less conservative than intention-to-treat and more conservative than case-wise exclusion. As it is not ascertainable how many subjects actually completed the whole intervention, the attrition rate can only be estimated at 18% (if we assume that someone filling out at least one post-baseline assessment has not dropped out).

The dependent measure, the CES-D, has two cut-off scores of 16 or greater (moderately depressed) and 28 or greater (severely depressed) (cited by authors: Radloff, 1977). The authors report an effect size of 0.5 (post-intervention versus baseline versus TAU control group) for those with clinical depression scores, for which they use the cut-off score of CES-D >= 20 (n = 191).

In a similar, previous study without reminders (Clarke et al., 2002), no effect was found, according to the authors mainly due to the fact that Internet visits by participants were limited.

Whereas on face value the comparison of website-based CBT with treatment as usual (TAU) -medication, psychotherapy, etc- sounds interesting, careful consideration makes for some reserve. Two important points to bear in mind about this study are a) that the control group, as well as the intervention group, TAU was the order of the day and b) most of the individuals (78%) were recruited from patients who had received depression medication or psychotherapy in the previous 30 days and had a "chart" diagnosis of depression. This means that one cannot really speak of treatment gains, as the intervention for the most is an adjunct to current depression treatment. Although the authors had attempted to recruit a sample from the "non-depressed" population, unfortunately they do not report separately for these two groups other than their comparability at baseline. Although a medium effect size of 0.5 seems small, seen in the light of an adjunct to treatment as usual, which includes face-to-face CBT among other forms of treatment, this effect is encouraging in itself.

Taken together with their previous study without reminders (Clarke et al., 2002), in which no effect was found, the conclusion may be drawn that participants somehow need to be motivated not to drop out and that reminders have some effect in this regard. The major weaknesses in both studies are, however, that the intervention is open-ended (that is, no timed delivery nor required completion of a module before continuing) and it is not ascertainable how much of the material has effectively been worked through by each participant. Furthermore, the face-to-face element is not the only thing that has been factored out, but the entire therapist.

Andersson et al. (2005) randomised 117 individuals with moderate depression to a website-based CBT treatment with minimal therapist contact or a waitlist control condition using a two-way pre-post design. Participants were recruited via newspaper articles in Sweden. Of the 343 individuals who responded, filled in the on-line assessments and sent informed consent forms, 117 met the inclusion criteria, of which the most important were: A probability of at least 0.55 for the diagnosis of a major depression according to the Composite International Diagnostic Interview-Short Form (CIDI-SF; cited by authors: World Health Organization, 1999); a score in the range 15 to 30 on the Montgomery Åsberg Depression Rating Scale (MADRS-S; cited by authors: Svanborg & Åsberg, 1994), indicating mild to moderate depression, 18 years or older, no co-morbid psychiatric disorders (including a bipolar disorder) or prior CBT treatment. Of the 57 randomised into the treatment group, 21(37%) dropped out and of the 60 randomised into the waitlist control group, 11 (18%) dropped out.

The treatment condition consisted of five modules (in Swedish) delivered via a website, based on the principles of cognitive therapy and behavioural activation: 1) introduction, 2) behavioural activation, 3) cognitive restructuring, 4) sleep and physical health and 5) relapse prevention and future goals. Each module was concluded with questions on the content of the module, the answers to which were e-mailed to the therapist, who would give feedback within 24 hours and simultaneously provide access to the next module if sufficient progress had been made. The advised duration was 8 weeks. Therapist time spent per participant, including assessments, was approximately two hours.

Participants in both the treatment as well as the waitlist control groups were expected to participate in a discussion group. Waitlist control participants commenced the intervention after the treatment group had completed theirs.

The dependent measure was the Beck Depression Inventory (BDI; cited by authors: Beck, Ward, Mendelson, Mock, & Erbuagh, 1961), filled out on-line. Dropout data was excluded in the calculation of effect size and a Cohen's *d* of 0.9 (a large effect) was reported post-treatment versus pre-treatment versus control group, results being maintained at six-month follow-up. As secondary measure the MADRS-S was also used and rendered an effect size of 0.8.

Though the authors call the intervention "self-help", it is actually a minimal contact therapy. What makes this study good is the fact that completion of the material is recorded and results are indicative for those individuals who actually completed the CBT material. This means that the attrition rate is somewhat higher than those reported for the other studies, yet also results in a higher effect size (with the exception of the study by Lange et al., 2005, see below). The addition of a discussion group did not, however, have any significant effect in the waitlist control group, even though the authors had expected one. The authors do not report how they ensured that the discussion group was actually (sufficiently) made use of, but do report that the amount of activity in the control group was significantly more than in the treatment group. This does mean that there is the possibility of an underestimation of the effect size (because control subjects were more active in the discussion group), but as the post-waitlist measures had not significantly changed from pre-waitlist, this effect should be negligible.

Lange et al. (2005) randomly assigned 57 individuals with heightened depression levels to a website-based CBT condition or a psycho-education control condition in a two-way pre-post design.

Subjects were sought in the Netherlands via an article in a national newspaper, advertisements and the interapy web site (see www.interapy.nl). Of the 231 respondents, 57 completed the screening, sent informed consent forms and met the criteria for inclusion in the study. To be included, individuals were screened with online self-report questionnaires and had to meet the following major criteria: 18 years or older of age, no alcohol or drugs misuse, not currently under psychological or psychiatric treatment, no antidepressant or other psycho-active medication and a depression score between 10 and 28 on the Dutch translation of the Beck Depression Inventory (BDI; cited by authors: Bouman, 1994). This depression range indicates mild to severe depression. Furthermore heightened suicide risk or co-morbid dissociative, psychotic, panic or posttraumatic stress disorders were excluded from the study.

The CBT program consisted of an eight module therapy delivered via the Interapy website over a period of 11 weeks. The modules contained 1) awareness of depression and a writing exercise, 2) awareness and the registration of depression levels, 3) activation, relaxation and sleep, 4) cognitive restructuring – the challenging of negative thoughts, 5) cognitive restructuring – testing negative thoughts via experiments, 6) positive self-talk, 7) interaction with others and 8) relapse prevention. Each module contained some form of registration or writing exercise, which would lead to interaction with the therapist and agreement on homework exercises. A client was required to do more than two hours of homework per week. Each client was assigned to one of 25 therapists, all of whom had (the equivalent of) a Master's degree in clinical psychology and had received training in the implementation of the Interapy CBT protocol. For both the psycho-education group and the CBT group, before randomisation, all participants had received psycho-education via the Interapy

website. In the psycho-education condition, a seven-week waiting period was implemented, after which post waiting period assessments were made and the participants could also commence with the CBT therapy.

Of the 40 randomised to the CBT condition, eight (20%) dropped out and of the 17 randomised to the psycho-education condition, three (27%) dropped out.

As dependent measures, the scores of the following three instruments (all as cited by authors) were used: The BDI mentioned above, the depression subscale of the Dutch version of the SCL-90 (Arrindell & Ettema, 2003) and the depression scale of the Dutch version of the Depression, Anxiety and Stress Scale (DASS; De Beurs, Van Dyck, Merquenie, Lange & Blonk, 2001). Data from dropout was excluded in the analysis, rendering an average effect size (post-pre-control) of 1.1 standard deviation units – a large effect.

This study gives support to the notion that the therapist is not what needs to be factored out and, on the other hand, that face-to-face contact is not essential. Intensive therapist contact in their program ensured large effect sizes of 1.8 standard deviation units reduction in levels of depression (1.1 when the effect of the psycho-education is deducted), similar or even better than effects found in face-to-face CBT interventions. It is a pity that the size of the sample is somewhat small and that the follow-up period is relatively short (at six weeks post-treatment).

3. Studies on effectiveness: Treatment of Anxiety.

Anxiety disorders come in many forms, such as specific phobias, obsessive-compulsive disorders, generalised anxiety disorder, posttraumatic stress disorder, panic disorder (with or without agoraphobia), etc. (Barlow, 2002) Although we have not limited ourselves to any specific anxiety disorder, it seems that certain forms of

anxiety disorders –mostly panic and posttraumatic stress- lend themselves better to a protocol-driven intervention that may be delivered without face-to-face contact.

3.1

Lange, Van de Ven, Schrieken and Emmelkamp (2001) randomly assigned 30 students who had experienced a traumatic event at least 3 months ago to a website-based intervention with psycho-education or a psycho-education waitlist control condition (two-way pre-post design). Participants were recruited from a pool of 500 students in Amsterdam, Netherlands, in return for course credit points. Of the 41 potential participants, only 30 fulfilled the inclusion criteria and were prepared to undergo the intervention. Potential participants were excluded if they showed extremely high depression scores (>58 for women and >53 for men) on the depression subscale of the SCL-90 (cited by authors: Arrindell & Ettema, 2003), had an inclination to psychological dissociation, were at risk for psychoses, were substance abusers, underwent current traumatic experiences or were currently being treated. Screening was performed on-line.

The treatment condition comprised 10 writing sessions of 45 minutes over a period of 5 weeks divided into three phases: 1) self-confrontation, 2) cognitive reappraisal and 3) sharing and farewell ritual. In the middle of each phase, the therapist (one of seven students in Clinical Psychology at the University of Amsterdam who had followed advanced courses in behavioural and cognitive psychotherapy and were under regular, expert supervision) would provide feedback on the writings of the participant and instructions how as to proceed.

Both the treatment as well as the control participants received psycho-education before commencing with the screening. Participants randomised to the

control condition were offered the active treatment at the conclusion of the treatment period for the intervention condition.

Dependent measures (all as cited by authors) were the Impact of Events Scale (IES; Kleber & Brom, 1986), the anxiety, depression, somatisation and sleeping problems subscales of the SCL-90 (Arrindell & Ettema, 1986) and the Profile of Mood States (POMS; Wald & Mellenbergh, 1990).

Of the 30 students randomised, two dropped out of the experimental condition (13%) and three from the control condition (20%). Effect sizes (post-pre-control) were: 1.1 for intrusion on the IES, 0.7 for avoidance on the IES, 1.0 for depressed mood (SCL-90), 1.1 for somatisation (SCL-90), 0.9 for depressed mood (POMS), 0.9 for fatigue (POMS), 0.9 for tension (POMS), 0.6 for vigour (POMS) and 0.5 for anger (POMS).

Although this study is based on a student population, it is regarded as relevant as the students concerned had actually undergone a traumatic event and were suffering from the consequences thereof. There is no certainty as to how many of the subjects would have filled the criteria of the DSM-IV, as the screening was performed self-report and on-line. The further limitations are, of course, the small sample size and no follow-up assessment. It is not entirely clear how much therapist time was spent with each participant and several measures for effect size make for difficult interpretation. From the point of view of anxiety, it is strange that a large effect size of 0.7 is found for the improvement in avoidance scores and that the effect size for the anxiety subscale of the SCL-90 is not mentioned (I calculate 0.9 – a large effect also).

Carlbring, Westling, Ljungstrand, Ekselius and Andersson (2001) assigned 41 individuals fulfilling the criteria for a DSM-IV panic disorder (PD) diagnosis to a

website-based intervention or a waitlist control condition (two-way pre-post design). Participants were recruited through newspaper articles and a link from the Swedish National Association for people suffering from PD. The first 500 respondents filled out a computerised interview the Composite International Diagnostic Interview-shortened form (CIDI-sf; cited by authors: World Health Organization, 1999) and if a PD diagnosis seemed likely, the PD section of the Anxiety Disorders Interview Schedule (ADIS-IV; cited by authors: Di Nardo, Brown & Barlow, 1994) was e-mailed to them. The main inclusion criteria were the fulfilment of DSM-IV criteria for PD, duration of the disorder for at least one year, age between 18 and 60 and no co-morbid disorders. After randomisation, four individuals dropped out from the treatment condition (20%) and one from the waitlist control condition (5%).

The intervention consisted of a two-week baseline period, during which participants kept a diary with anxiety ratings and details of panic attacks. Thereafter six modules based on self-help manuals were provided on the Internet site. Each module ended with five to eight questions that the participant was required to answer via e-mail. Feedback was given within 24 hours and on the basis of these answers the password to the next module would be provided. A time limit of 14 days was set for each module, participants being reminded by e-mail. The modules contained 1) psycho education, 2) breathing retraining, 3) thought processes in relation to anxiety, 4) interoceptive exposure, 5) exposure in vivo and 6) relapse prevention. The total amount of therapist time, including assessment, was approximately 90 minutes per participant.

The waitlist control group was time-matched so that whenever someone from the intervention group was ready to fill out post-treatment assessments, the

corresponding control group person would do the same. After filling out the assessment, the control group participant would commence with the active treatment.

The authors used several dependent measures (all as cited by authors): Body Sensations Questionnaire (BSQ; Chambless, Caputo, Bright, & Gallagher, 1984), Agoraphobic Cognitions Questionnaire (ACQ; Chambless et al., 1984), Mobility Inventory for Agoraphobia (MI; Chambless, Caputo, Jasin, Gracely, & Williams, 1985), Beck Anxiety Inventory (BAI; Beck, Epstein, Brown, & Steer, 1988), Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbuagh, 1961), Quality of Life Inventory (QOLI; Frisch, Cornell, Villanueva, & Retzlaff, 1992) and the panic diary. On all of these measures clinically significant improvements with respect to waitlist control were found with the exception of the Mobility Inventory and the Quality of Life Inventory. For clinical significance the authors use the criteria: no occurrence of a full-blown panic attack and no limited symptom attacks in the two weeks after treatment.

For an estimate of effect size based on the intention-to-treat paradigm, we calculate an approximate 0.9 standard deviation units improvement (a large effect) average in the treatment condition versus the control condition. The authors did not report the effect sizes per measure.

This study seems to have been set up very thoroughly and the authors seem to have taken a lot of effort to make their results measurable and explainable. What is missing is, of course, a follow-up after a reasonable amount of time (say one year). The relatively small size of their sample is also regrettable. Whilst a criticism might be that effect sizes were not reported, the fact that several measures are included make for easier comparison with other studies.

Richards and Alvarenga (2002) evaluated an internet-based intervention for nine people meeting the DSM-IV criteria for a panic disorder (PD) in a one-way pre-post design. Participants had been recruited via newspaper articles in Victoria, Australia. Of the 30 people who responded, only 14 individuals met the DSM-IV criteria for PD and only nine completed the intervention, giving an attrition rate of 36%.

The initial assessment was performed using the Prime-MD structured clinical interview (cited by authors: Spitzer, Williams, Kroenke, Linzer, & DeGruy, 1994), augmented with the Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV; cited by authors: Di Nardo, Brown, & Barlow, 1994) when there was any doubt about the PD diagnosis. The diagnosis was also confirmed with the Panic Disorder Severity Scale (PDSS; cited by authors: Shear et al., 1997). All assessments were performed face-to-face.

The intervention consisted of a five module program to be completed in 5 – 8 weeks: 1) Information about the nature of panic, 2) causes of panic and its effects, 3) cognitive, physiological and behavioural components of panic and their interaction, 4) negative and self-defeating cognitions and how to change them and 5) strategies for coping with panic attacks. Neither interoceptive exposure, nor in vivo exposure was included in the intervention. Each section of the material was concluded with a quiz, with immediate feedback on the correctness of the given answers. Participants were contacted on a weekly basis to monitor progress and answer any questions. It is not clear whether this contact included any therapeutic assistance, or if only progress and administrative matters were discussed. Post-intervention assessment was performed face-to-face three months after completion of the intervention. Total therapist time, including assessments, was about five hours per participant.

The dependent measures (all as cited by authors) used were the Panic Disorder Severity Scale (PDSS; Shear et al., 1997), the Anxiety Sensitivity Index (ASI; Riess, Peterson, Gursky, & McNally, 1986), the Body Vigilance Scale (BVS; Schmidt, Lerew, & Trakowski, 1997) and the Body Sensations Interpretation Questionnaire (BSIQ; Clark et al., 1997). Data from dropout was excluded. An effect size of 0.3 was found for the reduction of panic disorder severity measured with the PDSS, though only due to reduction in panic frequency and distress during panic attacks. The other measures showed no significant effects.

The major criticisms on this study are that the size of the group is very limited and the lack of control group makes for limited interpretation of the results. Furthermore, can the improvement in PDSS scores be attributed to the website-based intervention, or was the face-to-face therapist contact in the assessment procedure (both before and after the intervention) the therapeutic effective ingredient? It is well known that the diagnostic process itself also has a therapeutic effect (see for example Finn & Tonsager, 1992) and with such a small effect size it would be impossible to determine the effective ingredient. The authors concede that the exclusion of interoceptive exposure exercises probably explain the limited effect of the intervention. Again, only a RCT including and excluding exposure exercises (and why not include in vivo exposure as well) could give empirical support to this supposition. It is pleasing to note that follow-up assessments took place three months after completion, giving some time for improvements to maintain themselves, although a longer follow-up would have been preferable.

Lange et al. (2003), following their previous study (Lange et al., 2001), randomly assigned 184 participants suffering from posttraumatic stress to an active

treatment condition or a psycho-education waitlist control condition (two-way pre-post design). As their study was largely similar to their previous study, we will limit ourselves here to the differences. In this study participants were not sought from the student population, but from the general population via newspaper articles, etc. Of the 916 individuals who performed the screening and pre-test, 479 did not meet the inclusion criteria (same as previous study, except that incest victims were also excluded), 253 did not return their informed consent forms and the remaining 184 were randomised into the treatment (n=122) or control (n=62) condition.

The POMS (cited by authors: Wald & Mellenbergh, 1990) was dropped as dependent measure. Drop-out from the study was 30 in the control group (48%) and 44 from the intervention group (36%). The effect sizes reported for this study (post-pre-control) were: 1.3 for intrusion and 1.4 for avoidance on the IES, 1.0 for depression, 0.8 for anxiety, 0.7 for somatisation and 0.6 for sleeping problems on the SCL-90. These effect sizes are generally larger than those found in their previous study.

This shows that the results obtained in the general population were even better than those in the student population. Of course, this could be as the initiatory agent in this second study is the participant itself, that is, the participant actively sought help. This could explain better results. On the other hand, the high dropout rate is a cause for concern. The authors mention that dropouts were significantly more likely to be men, were younger and less likely to live with a partner.

Carlbring, Ekselius and Andersson (2003) randomly assigned 22 individuals with a DSM-IV panic disorder diagnosis into one of two treatment groups: CBT or Applied Relaxation (AR) in a two-way pre-post design. Both interventions were given

via the Internet, with minimal therapist contact via e-mail (30 minutes per client). In a previous study (mentioned above; Carlbring et al., 2001), they had already ascertained that waitlist controls showed insignificant improvements over time and chose for this study to compare AR and CBT interventions delivered via the Internet, as AR and CBT delivered face-to-face had given similar results in other studies (cited by authors: Öst, & Westling, 1995). Their second objective was to reduce the amount of therapist contact in comparison to their previous study, from 90 minutes to 30 minutes per participant.

Participants were recruited from a waitlist of individuals interested in taking part in an internet-based treatment for PD. Inclusion criteria were the same as in their previous study, except that the DSM-IV criteria for PD were assessed face-to-face with the Structured Clinical Interview for DSM-IV (SCID-IV; cited by authors: First, Gibbon, Spitzer, & Williams, 1997). Of the 53 individuals interviewed, 22 fulfilled the inclusion criteria and were randomised to one of the intervention groups. Three persons dropped out of the CBT condition (27%) and two from the AR condition (18%).

The CBT program for PD was the same as their previous study and the AR program comprised the following nine modules: 1) psycho education, 2) rationale, 3) progressive muscle relaxation – long version, 4) progressive muscle relaxation – short version, 5) conditioned relaxation, 6) differential relaxation, 7) quick relaxation, 8) applied relaxation and 9) relapse prevention. Furthermore, a CD with three relaxation instructions was provided and those participants with a cellular phone were sent SMS reminders to relax twice every workday. For both interventions, each module concluded with five to eight questions that were e-mailed to the therapist. Standardized e-mail feedback was given within seven days together with the password

to the next module if the answers were deemed satisfactory by the therapist. By using standardized responses, therapist time was reduced to 30 minutes per participant. There was no time limit set for participants to complete any particular module, yet an email message was sent after four months to inquire about progress. Post-treatment assessments were performed seven months after commencement of the program.

The same dependent measures as in their previous study were used. Overall effect sizes based on the intention-to-treat paradigm were small (CBT 0.4; AR 0.7) in comparison to what was expected, but did not differ significantly for the two groups. The authors think that the open-endedness (i.e. lack of time limit) of the modules and low completion rates (on average only 56% of the CBT material was completed) negatively affected results.

There are a few additional criticisms to be made over and above those made on the earlier study (Carlbring et al., 2001). That so little of the CBT material was completed raises the question whether the program has actually been worked through to sufficient extent so as to be called a CBT intervention. This is the same criticism levelled at the depression intervention of Christensen, Griffiths, and Jorm (2004). Another confounding introduced by these authors is the use of SMS-messages to a subgroup in the AR condition. It simply is not clear what the effect is.

By comparing these two studies we can conclude (a pity that the authors do not do this) that the reduction in therapist time could have negatively impacted the effect size. Unfortunately there are alternative explanations: The dropping of the time limit for completion of each module, or the longer time before responding to e-mailed answers to questions (seven days in stead of one) could also have caused the reduction in efficacy. It is a pity that the authors did not vary only *one* variable, for example the *amount* of therapist contact rather than the amount and quality of therapist contact and

the dropping of the time limit for completion. On the safe side, we could conclude that all three these variables made some impact but we are unable to say anything more with any degree of certainty.

Kenwright, Marks, Gega and Mataix-Cols (2004) assigned 10 individuals with an ICD-10 diagnosis of phobia or panic disorder to a website-based intervention in a one-way pre-post design. Participants were recruited from 266 persons referred to a clinic in West London. Via a telephonic interview, using an interview checklist of the ICD-10 criteria for phobia or PD (cited by authors: World Health Organisation, 1992), 56 met the suitability criteria. Only 10 had access to the internet at home, the remaining 46 either refused treatment (13) or were treated at the clinic (33, of which 16 dropped out).

The intervention (the FearFighter program) is a 9 module, practise-based (exposure) self-help intervention based on CBT principles (see www.FearFighter.com), originally developed for use on a stand-alone computer and subsequently also made available via the Internet. The modules consist of: 1) Welcome and introduction, 2) principles of CBT, 3) identifying triggers, 4) the value of recruiting a CBT co-therapist and hints on how to find one, 5) setting good goals and testing them, 6) coping strategies, 7) practising personal coping strategies during both imagined and live CBT homework, 8) progress review and 9) troubleshooting.

The therapist contacted participants at home, by telephone in seven 10-minute sessions to give participants support.

Dependent measures (all as cited by authors) used were the Fear Questionnaire (FQ; Marks & Mathews, 1979) and the Work and Social Adjustment scale (WSA;

Mundt, Marks, Shear, & Greist, 2002). The average effect size one month after end of treatment versus pre-treatment was 1.0 – a large effect. A subsequent RCT is reported by Schneider, Mataix-Cols, Marks and Bachofen (2005) below.

This, being a pilot study, has the obvious flaws of not being a RCT and an extremely small sample size. Other problems in interpreting the results are that the types of anxiety presented by the participants were quite diverse: Agoraphobia with panic, social phobia, insect phobia, claustrophobia, some with co-morbid depression and generalised anxiety disorder. Nonetheless, an average effect size of 1.0 one-month post treatment is encouraging, especially when considering that the participants did not visit the clinic, nor had seen their therapist face-to-face and had received around two hours of therapist contact throughout the intervention.

Schneider et al (2005) randomised 68 participants to a website-based intervention or a minimal CBT condition in a two-way pre-post design. This study was an extension of the FearFighter program described above (Kenwright et al., 2004). Participants were recruited from referrals to a self-help clinic in the U.K. Of the 94 individuals who applied for the trial, 68 met the inclusion criteria, the major criteria being ICD-10 agoraphobia (with or without panic), social phobia or specific phobia.

The two conditions comprised:

- FearFighter via Internet: This was the same as in the previous study (Kenwright et al., 2004). Of the 45 participants randomised to this condition, 12 (27%) dropped out.
- Managing Anxiety via Internet: This comprised a similar program to that of FearFighter, but excluded any exposure instructions. The program was divided into modules: 1) Introduction, 2) balancing health

worries with positive thoughts and breathing, 3) health goals and progressive muscle relaxation, 4) coping strategies and passive progressive muscle relaxation, 5) dealing with daily hassles and 6) improving social contacts. Of the 23 participants randomised to this condition, 8 (35%) dropped out.

In both conditions, participants were telephoned six times (approximately 18 minutes per call) to review progress, provide support and extra treatment advice. Of this time, 80% was spent on giving support.

The dependent measures (cited by authors) were Main Problem and Goals (Marks, 1986) the global phobia item of the FQ (Marks & Mathews, 1979) and the WSA (Mundt et al., 2002). The participants were asked to self-assess several times during the treatment and were also assessed at screening, post-treatment and at follow-up by an experienced psychologist not involved in the intervention, who was blind to the treatment condition. Using these blind assessments, effect sizes (pre-post and pre-follow-up) were: 1.1 and 1.2 for the FearFighter condition and 1.2 and 0.9 for the Managing Anxiety condition. These effect sizes are different to those reported by the authors, as they calculated effect size using the standard deviation at the pre-treatment measurement whereas it is customary to use the pooled standard deviation of the pre-treatment and post-treatment (or follow-up) measurements.

The authors had expected the FearFighter condition to render better improvements than the Managing Anxiety condition, as the latter contained no exposure instructions. The difference in effect size at follow-up was 0.3 (and was –0.1 at post-treatment). Although therapists were instructed not to give exposure instructions to participants in the Managing Anxiety condition, the amount of therapist contact was the same for both conditions, maybe explaining the similarity in effect

size. The major criticisms on this study are the limited sample size and the short follow-up period. Furthermore, the authors claim to use intention-to-treat analysis and then state that pre-treatment scores are *not* carried forward, implying exclusion of data for which no post-treatment assessment is available. As regards our study, the RCT component (that is, including or excluding exposure instructions) is not really relevant, as both conditions were delivered over the Internet and both had similar forms and quantity of therapist contact.

Klein, Richards and Austin (in press) randomised 55 people with PD to a website-based CBT with e-mail contact condition, a CBT manual-driven condition or an information only control condition in a three-way pre-post design.

Participants were recruited from 130 Australians who had made contact with a panic website. The main inclusion criteria were: A DSM-IV PD primary diagnosis and no other therapy or self-help during the intervention. The PD diagnosis was assessed telephonically using the Panic Attack Questionnaire (PAQ; cited by authors: Norton, Dorward, & Cox, 1986) and if on the basis of the PAQ a PD diagnosis seemed likely, the ADIS-IV (cited by authors: Brown, Di Nardo, & Barlow, 1994) was used.

Participants were sequentially assigned to one of three conditions:

- The on-line panic program consisted six modules delivered during a six-week period: 1) Introduction, 2-5) learning modules and 6) relapse prevention. The modules were based on common treatment methods used in the CBT treatment of PD such as controlled breathing, cognitive restructuring, interoceptive exposure and situational exposure. Therapists (trained and practising in the CBT treatment of PD) contacted clients by e-mail, guiding and supporting them through

the on-line program. Therapists were required to answer e-mails within 24 hours and spent an average of 330 minutes per client in providing this contact. Of the 19 participants, one (5%) dropped out from this condition.

- The manualised CBT workbook condition provided the same ingredients as the on-line version, except that the material was presented in a manual (the MAP-3; cited by authors: Barlow & Craske, 2000). Therapist contact was once weekly by telephone and an average of 245 minutes was spent per client. Of the 18 participants in this condition, 3 (17%) dropped out.
- The information control group: participants were informed that a therapist would become available in six weeks time. Each participant was called weekly by telephone and provided minimal support by the therapist (empathic listening), which entailed an average of 65 minutes contact time. Of the 18 participants in this condition, 5 (28%) dropped out. After the waiting period they were re-assessed and assigned to one of the two active conditions above.

The following dependent measures were used (all as cited by authors): The PDSS (Shear et al., 1997), which was administered telephonically, the Anxiety Sensitivity Profile (ASP; Taylor & Cox, 1998), the DASS (Lovibond & Lovibond, 1995), the ACQ (Chambless et al., 1984) the BVS (Schmidt, Lerew, & Trakowski, 1997) and some other measures to determine satisfaction with the intervention, all performed on-line.

The intention-to-treat paradigm was followed giving effect sizes (pre-post-information control) of: 3.1 on the PDSS, 0.6 on the BVS, 0.9 on the ACQ, 1.3 on the

ASP, 0.7, 0.6 and 0.8 on the depression, anxiety and stress subscales of the DASS for the online PD intervention. This yields an average effect size 1.1, which improves to 1,7 at three-month follow-up. The online PD intervention performed on average 0.2 better at post-treatment and 0,5 better at follow-up than the manualised condition.

The obvious shortcoming in this study is the small sample size.

Carlbring et al. (2005) randomly assigned 49 individuals with a DSM-IV panic disorder diagnosis to a live face-to-face CBT group (LIVE) or an internet-based CBT group (IT) in a two-way pre-post design. Participants were recruited from a waitlist of individuals interested in taking part in an internet-based treatment for PD. Inclusion criteria were similar to those in their previous study (Carlbring, Ekselius, & Andersson, 2003), in brief fulfilling the criteria for a DSM-IV PD, age between 18-60, no co-morbid psychiatric disorders and limited depressive symptoms. A major difference in the screening procedure was that a computerised screening was used rather than a face-to-face assessment. This screening consisted of the self-rated version of the Montgomery Åsberg Depression Rating Scale (MADRS-SR; cited by authors: Svanborg & Åsberg, 1994), and 53 questions derived from the PD sections of the Composite International Diagnostic Interview (CIDI 2.1; cited by authors: World Health Organisation, 1997), the ADIS-IV (cited by authors: Di Nardo, Brown, & Barlow, 1994) and the PD and social phobia sections of the SCID-IV (cited by authors: First et al., 1997). If this computerised interview showed that an individual fulfilled the inclusion criteria, a face-to-face SCID-IV interview was performed to confirm the diagnosis.

Of the 427 individuals who completed the computerised interview, 64 fulfilled the inclusion criteria and were subsequently called-up for the SCID interview. Of the

59 that showed up, 49 fulfilled the inclusion criteria and were randomised to one of the intervention groups. Three persons dropped out of the IT condition (12%) and three from the LIVE condition (12%).

Both groups followed a similar program during a ten-week period largely similar to that in their previous study (Carlbring, Ekselius, & Andersson, 2003), with the exception that in the LIVE version, participants received 10 weekly individual sessions of 45-60 minutes in which the same material was covered as in the IT condition. The modules consisted of: 1-2) psycho-education and socialisation, 3) breathing retraining and hyperventilation test, 4-5) cognitive restructuring, 6-7) interoceptive exposure, 8-9) in vivo exposure and 10) relapse prevention. However, the IT condition was no longer as open-ended: Each module included 3-8 essay questions the answers to which were e-mailed to the researchers, so that they could determine whether participants had understood the material and performed their homework exercises. Feedback was given within 36 hours and if sufficient progress had been made, the password for the next module was e-mailed. An additional interactive multiple-choice exercise, for which 100% correct answers were required (with immediate feedback), was included at the end of each module. Participants were also required to post at least one message in an on-line discussion group. Therapist time per participant, including administration, was on average 150 minutes.

The authors used similar dependent measures (all as cited by authors) as for their previous study (Carlbring, Ekselius, & Andersson, 2003): The ACQ and BSQ (Chambless et al., 1984), MI (Chambless et al., 1985), BAI (Beck et al., 1988), BDI (Beck et al., 1961), QOLI (Frisch et al., 1992). Although screening was performed on-line, the dependent measures were filled in pencil-and-paper, one month post-treatment and follow-up at 12 months.

No significant differences in effect sizes were found between the two groups, implying that an internet-based treatment is equally efficacious (average effect size 0.8 post-treatment and 0.8 at 12-month follow-up) as a face-to-face CBT intervention (average 1.0 post-treatment and 0.9 at follow-up). The effect sizes for the internet-delivered treatment per measure (post-treatment and at 12-month follow-up) were: BSQ 1.5 and 1.4; ACQ 1.2 and 1.3; MI (alone) 0.6 and 0.7; MI (accompanied) 0.7 and 0.6; BAI 0.9 and 0.9; BDI 0.8 and 0.8; MADRS 0.9 and 1.0 and QOLI –0.4 and –0.3. Clients generally found the pace of 10 weekly sessions a little too high.

This subsequent study corrects some of the caveats in their previous studies. Firstly, the conditions face-to-face and non-face-to-face are far more comparable. Of course, the LIVE condition implied far more therapist contact (450-600 minutes versus 150 minutes for the IT condition), yet the difference in effect size is non-significant. This again shows that the face-to-face element is not essential for efficacy. This, together with ensuring that modules are worked through and with regular e-mail contact with the therapist shows an increase of 0.4 standard deviation units in effect size.

Carlbring et al. (in press) randomly assigned 60 individuals with a DSM-IV panic disorder diagnosis to an internet-based CBT group or a waitlist control group in a two-way pre-post design. Participants were recruited from a waitlist of individuals interested in taking part in an internet-based treatment for PD. Inclusion criteria were similar to those in their previous study (Carlbring et al., 2005), in brief fulfilling the criteria for a DSM-IV PD, age between 18-60, no co-morbid psychiatric disorders and limited depressive symptoms. Screening procedure was the same as for the previous

study, except that the SCID interview was undertaken telephonically rather than face-to-face.

Of the 358 individuals who completed the computerised interview, 104 fulfilled the inclusion criteria and were subsequently called-up for the SCID interview. Of those interviewed 60 fulfilled the inclusion criteria and were randomised to one of the two conditions. One person dropped out of the active treatment condition (3%) and three persons did not return their post-treatment assessments. Persons on the waitlist were offered the active treatment at the end of the 10-week active treatment period.

The treatment condition was the same as that in their previous study (Carlbring et al., 2005). Feedback on emailed homework was given within 24 hours (instead of 36) and therapist time per participant, including administration, was an average of 120 minutes. In addition to what was matter of course for the previous study, therapists called participants by telephone once a week to provide feedback and answer questions. This resulted in an additional 120 minutes per participant in therapist contact.

The authors used the same dependent measures as for their previous study (Carlbring et al., 2005), except that they were now administered via the Internet. Follow-up was after nine months instead of 12 months. Effect sizes were calculated according to the intention-to-treat paradigm: BSQ 2.0, ACQ 1.5, MI 0.7 (alone) and 0.5 (accompanied), BAI 1.4, BDI 0.6, MADRS 0.8 and QOLI 0.5 at post-treatment. This renders an average effect size over all measures of 1.1, which dropped to 1.0 at 9-month follow-up. This is better than the 0.8 found in their previous study, implying that the telephone contacted accounted for an additional 0.3 in effect size.

Compliance was also improved –this probably explaining the increase in effect size- as 80% of the group completed all modules in the 10-week period and an average

of 8.9 modules (out of 10) were completed, compared with the 7.4 modules in the previous study without telephone calls.

The study has been well set up, although the number of participants remains relatively small and the follow-up at nine months could rather have been standardised at one year to make comparison with their previous study. The only other possible confounding is that assessments were now performed via the Internet and not pencil-and-paper.

Andersson et al. (in press) randomly assigned 64 individuals fulfilling the DSM-IV criteria for social phobia to a multi-modal CBT program or a waitlist control group in a two-way pre-post design. Participants were recruited via newspaper articles in Sweden and links on the World Wide Web. The inclusion criteria mainly consisted of the fulfilling of the DSM-IV criteria for social phobia, be afraid of giving a public speech, not be overly depressive and having social phobia as primary diagnosis. Subjects were first screened on-line using the Social Phobia Screening Questionnaire (SPSQ; cited by authors: Furmarket al., 1999) and the MADRS-S (cited by authors: Svanborg & Åsberg, 1994). Suitable participants were then contacted by telephone and a SCID interview (cited by authors: First et al., 1997) conducted to confirm the diagnosis. Of the 237 individuals who applied, 163 fulfilled the screening criteria and the first 100 contacted by telephone underwent the SCID interview. Of these, 64 fulfilled all inclusion criteria and were randomised to the active condition or to a waitlist control group. Those on the waitlist were offered the treatment discussed below (Carlbring, Furmark, Steczkó, Ekselius & Andersson, 2006) after the waiting period had been completed. Two individuals dropped out from the active condition (6%).

The treatment consisted of nine modules based on CBT principles, delivered on-line: 1) Social phobia, its symptoms, etiological factors and CBT, 2) social phobia model, relationship between thoughts, feelings, behaviour and cognitive symptoms, automatic thoughts, 3) thinking errors and cognitive distortions, registration and challenging of automatic thoughts, 4) behavioural experiments, 5) principles of exposure and reality testing, 6) self-focus, shifting of focus, attention training and safety behaviours, 7) exposure assignments, 8) communication and assertiveness and 9) perfectionism, procrastination, self-confidence and relapse prevention. Each module concluded with three to eight essay questions to be emailed to the therapist to determine if the participant had understood the material and performed the homework exercises. Feedback was given within 24 hours and, if sufficient progress had been made, the password for the next module was e-mailed. An additional interactive multiple-choice exercise, for which 95% correct answers were required (with immediate feedback), was included at the end of each module. Participants were also required to post at least one message in an on-line discussion group. 38% did not complete all modules in the 9-week period and on average 7.5 of the 9 modules were completed by the participants. Therapist time per participant for the on-line portion of the intervention, including administration, was 180 minutes on average. In addition to the on-line program, participants followed two three-hour group (6 – 8 per group) exposure sessions at Uppsala University. Attendance at the first session was 100% and 59% for the second session.

Therapists were three clinical licensed psychologists, two with a Master's or PhD and two students completing their last semester of their Master's degree.

The following dependent measures were administered pencil-and-paper (all as cited by authors): Liebowitz Social Anxiety Scale self-report (LSAS-SR; Baker,

Heinrichs, Kim, & Hoffman, 2002; Liebowitz, 1987), the Social Phobia Scale (SPS) and Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998), the SPSQ (Furmark et al., 1999), the Personal Report on Confidence as a Speaker (PRCS; Paul, 1966), the BAI (Beck et al., 1988), the MADRS-S (Svanborg & Åsberg, 1994) and the QOLI (Frisch et al., 1992). The intention-to-treat paradigm was followed. The effect sizes at post-treatment and at one-year follow-up were 0.9 and 1.3 on the LSAS-SR, 1.0 and 1.1 on the SPS, 1.2 and 1.1 on the SIAS, 1.2 and 1.5 on the SPSQ, 0.6 and 0.7 on the PRCS, 0.8 and 0.6 on the BAI, 0.8 and 0.7 on the MADRS and 0.6 and 0.4 on the QOLI, giving an average effect size of 0.9 at post-treatment and 0.9 at one-year follow-up.

One limitation of this study is that the number of participants is relatively small. The fact that follow-up was done one year after treatment shows that improvements were maintained. An effect size of 0.9 –a large effect- is promising. The question of amount of material completed remains a sticky one, but that on average 7.5 of the 9 modules were completed looks promising. As the authors also concede, it is difficult to untangle in this study whether the effect was the result of the Internet-based program or the exposure sessions – the following study shows that the effect is largely the result of the Internet-based program.

Carlbring et al. (2006) assigned 26 individuals fulfilling the DSM-IV criteria for social phobia to a website-based CBT program. Participants were recruited from the waitlist at conclusion of the active treatment for the study reported above (Andersson et al., in press). The inclusion criteria were the same. Of the 30 individuals on the waitlist, four individuals were excluded and of the 26 who commenced the treatment, 10 dropped out (38%). The treatment was the same, except that no group

exposure sessions were held and therapists were students completing their last semester of their Master's degree. The same dependent measures were used, except that the PRCS was dropped. The effect sizes using the intention-to-treat paradigm at post-treatment and at six-month follow-up were 0.9 and 1.4 on the LSAS-SR, 0.9 and 1.4 on the SPS, 1.1 and 1.5 on the SIAS, 1.4 and 2.0 on the SPSQ, 0.5 and 1.0 on the BAI, 0.8 and 1.0 on the MADRS and 0.7 and 0.7 on the QOLI, giving an average effect size of 0.9 at post-treatment and 1.3 at six-month follow-up. These effect sizes are not only large, but better than those reported for the intervention with group exposure sessions.

One limitation of this study is that it does not have a control group, although, the active treatment group being a waitlist control group from a previous study, the authors did test that in the waitlist condition before this active treatment started, no significant improvements had been made. The number of participants is relatively small and the follow-up at six months on the short side.

4. Discussion

The studies above each have a story to tell in their own right. However, our concern in this paper is to determine efficaciousness of a certain kind of therapeutic process –that is, one delivered via the Internet- and attempt to discover factors that affect the efficacy.

A matter that makes for difficult comparison is that even though we use Cohen's *d* as measure, the measures on which the calculation of this Cohen's *d* is based is not arbitrary. Each study uses its own preferred measures and it would go beyond the ambit of this paper to discuss whether or not the improvement of depression scores measured with the BDI should be included in the calculation of effect size in an intervention for the treatment of PD, for example. We choose therefore to compare on the basis of average effect size over all the measures used by the authors in their respective studies. Unless specifically stated otherwise, the effect size reported is the nett pre-post-control effect size.

A second choice that needs to be made concerns follow-up. The problem with follow-up is that few studies have used the same follow-up period. This raises two concerns: 1) Some of the improvement only becomes manifest as the client puts the newly learned behaviours into practice, that is, effects may not yet be fully present at post-treatment and 2) Some of the material learnt will not be put into practice and gradually fall into the background. So what is an ideal follow-up period? I believe that a period of 12 months is ideal, as this will allow both factors mentioned above to work themselves out and stabilise. Of all the studies reported on here, only two have a follow-up period of 12 months, the remainder ranging from no follow-up at all (that is, only post-treatment) to nine months. The average follow-up period was 16 weeks

(SD=16). For the sake of comparability, we are necessitated to use the post-treatment versus pre-treatment calculation of effect size.

As regards the differences in dealing with dropout, most studies have used the intention-to-treat paradigm. The question as to the actual difference the used paradigm makes in the reported effect size cannot really be answered. In the comparison below we therefore acknowledge the existence of this confounding, yet ignore its effect: The effect size as reported is used.

A final general comment is that on sample size, which ranges from 10 to 182 individuals in the active condition. If dropout is deducted this range reduces to from 8 to 155, whereas a sample size of at least 20, rather 30 individuals would be preferable. However, using a cut-off of 20, only nine of the 16 studies would remain. We have included all the studies in our comparison, irrespective of sample size.

In our subsequent discussion on factors that may affect website-based interventions, it may be useful to refer to Table 1 and Table 2 below, in which we have attempted to summarise some of the details extracted from the articles reported on:

Authors	*Number of participants*	*Attrition (%)*	*Effect size at post-treatment*	*Screening*	*Intervention*	*Contact with therapist*	*Duration*	*Follow-up after*	*Other*
Clark et al. (2002)	299 RCT	26%	n.s.	CES-D >= 20 (N=236)	1. CBT 2. Treatment as usual	None		4, 8, 16, 32 weeks	Website-based.
Christensen, Griffiths & Korten (2002)				Goldberg Depression Scale	CBT	None			Website-based.
Christensen, Griffiths, & Jorm (2004)	525 RCT	21%	0.65 (CBT) 0.5 (Info)	CES-D >= 16 (N=369)	1. CBT 2. Information 3. Attention placebo	None	6 weeks	intervention	Website-based. Intention-to-treat
Clark et al. (2005)	255 RCT	18%	0.5 (vs treatment as usual)	CES-D >= 20 (N=191)	1. CBT 2. Treatment as usual	None		5, 10, 16 weeks	Website-based with reminders. Restricted maximum likelihood estimation
Andersson et al. (2005)	117 RCT	27%	0.9	MADRS-S 15-30	1. CBT 2. Waitlist	Minimal. approx 2 hours	8 weeks	intervention, 6 months	Website-based. Case-wise exclusion
Lange, et al (2005)	57 RCT	19%	1.1	BDI 10-29	1. CBT 2. Psycho-education	Intensive	11 weeks	6 weeks	Website-based. Case-wise exclusion

Table 1: Comparison of effect size and some other possibly relevant factors for internet-based depression studies.

Authors	*Number of participants*	*Attrition (%)*	*Effect size*	*Screening*	*Intervention*	*Contact with therapist*	*Duration*	*Follow-up after*	*Other*
Lange et al (2001)	30 RCT	17%	1.1	PTSD	1) protocol-driven 2) waitlist control		5 weeks	6 weeks	Website-based. Case-wise exclusion.
Carlbring et al (2001)	41 RCT	12%	0.9	DSM-IV criteria for panic disorder	1) CBT via internet 2) waitlist control	Minimal via e-mail. 90 minutes	7 – 12 weeks	Post-treatment	Intention-to-treat
Richards & Alvarenga (2002)	9	36%	0.3	DSM-IV criteria for panic disorder	Information	Regular telephonic contact. 5 hours	5 – 8 weeks	3 months	Website-based CBT-based. Case-wise exclusion.
Lange et al (2003)	184 RCT	40%	1.0	PTSD	1) protocol-driven 2) waitlist control		5 weeks	6 weeks	Website-based. Case-wise exclusion.
Carlbring, Ekselius & Andersson (2003)	22 RCT	23%	0.4 (CBT) 0.7 (AR)	DSM-IV criteria for panic disorder	1) CBT via internet 2) Applied relaxation	Minimal via e-mail 30 minutes	Variable – post-treatment assessment after 7 months	Approx 3 months (assuming 4 month duration)	Website-based. Intention-to-treat.
Kenwright et al. (2004)	10	0	1.0	ICD-10 phobia or panic disorder	FearFighter via internet	Minimal via telephone. 113 minutes	12 weeks	1 month	Website-based. Pilot study

Table 2: Comparison of effect size and some other possibly relevant factors for internet-based anxiety studies.

Schneider et al. (2005)	68 RCT	29%	1.1 (1.2 at follow-up)	ICD-10 phobia or panic disorder	1) FearFighter via internet 2) Minimal CBT	Minimal via telephone. 115 minutes excluding assessments	10 weeks	1 month	Website-based
Klein, Richards & Austin (in press)	55 RCT	16%	1.1 (1.7 at follow-up)	DSM-IV criteria for PD	1) CBT via internet 2) CBT manual 3) Information	Via e-mail. 330 minutes	6 weeks	3 months	Website-based. Intention-to-treat.
Carlbring et al (2005)	49 RCT	12%	0.8 (internet) 1.0 (face-to-face)	DSM-IV criteria for panic disorder	1) CBT via internet 2) CBT face-to-face	Minimal via e-mail 150 minutes	10 weeks	12 months	Website-based. Intention-to-treat.
Andersson et al. (in press)	64 RCT	6%	0.9 (0.9 at follow-up)	DSM-IV criteria for social phobia	1) CBT + group exposure sessions 2) waitlist control	Minimal via e-mail. 180 minutes + exposure sessions (2 x 3 hours)	9 weeks	12 months	Website-based. Intention-to-treat.
Carlbring et al. (2006)	26	38%	0.9 (1.3 at follow-up)	DSM-IV criteria for social phobia	CBT via internet	Minimal via e-mail. 180 minutes	9 weeks	6 months	Website-based. Intention-to-treat.
Carlbring et al. (in press)	60 RCT	2%	1.1 (1.0 at follow-up)	DSM-IV criteria for panic disorder	1) CBT via internet 2) Waitlist control	Minimal via e-mail & telephone. 235 minutes	10 weeks	9 months	Website-based. Intention-to-treat.

Table 2 (continued): Comparison of effect size and some other possibly relevant factors for internet-based anxiety studies.

The efficaciousness of website-based interventions receives support from the studies reported above. We will view this statement in the light of the following issues:

- face-to-face versus non-face-to-face
- amount of therapist contact
- nature of therapist contact
- mechanisms to prevent dropout or stimulate completion
- screening issues

4.1 Face-to-face versus non-face-to-face

Glancing over Table 1 and Table 2 we see a range of effect sizes from zero to 1.1. However, not all of these studies can help us understand the effect of the face-to-face element. The studies in which we can untangle this factor are those by Lange et al. (2001, 2003 and 2005), the studies by Carlbring et al. (2005 and 2006) and the study by Andersson et al. (in press).

In the studies by Lange et al. (2001, 2003 and 2005) the face-to-face element is entirely absent, yet the amount of therapist contact is similar in quantity (even a little less) than in a traditional face-to-face delivery of a CBT intervention. The nature of the therapist contact in their studies is personal and therein also similar to a face-to-face intervention. The only element that is effectively varied is the face-to-face element. We see effect sizes around 1.0 relative to a psycho-education control group – a large effect.

Another potent indicator is the study by Carlbring et al. (2005), where an Internet-delivered intervention is compared to a live, face-to-face CBT intervention. Although especially the amount and somewhat the nature of the therapist contact is

different, a non-significant difference in effect size of 0.2 gives strong support for the notion that the face-to-face element is not essential.

Perhaps the most potent indicator is to be found in comparing the studies on the treatment of social phobia by Andersson et al. (in press) and Carlbring et al. (2006). In the original RCT by Andersson et al. (in press), a minimal e-mail contact, website-based CBT intervention, supplemented with two three-hour group exposure sessions was compared to a waitlist control, rendering an effect size of 0.9. In the subsequent study by Carlbring et al. (2006) the waitlist control group of the previous study was given the exact same intervention, but without the group exposure sessions, thereby removing 6 hours of face-to-face contact between clients and therapist. The effect size was essentially the same. One could argue that the exposure sessions themselves were not effective, but this does not alter the fact that clients *did have* face-to-face contact with their therapist in the one condition and *did not have* face-to-face contact in the other.

The general conclusion that can be drawn from the above is that the face-to-face element is not essential, that is, interventions are equally efficacious whether delivered face-to-face as when delivered non-face-to-face.

4.2 Amount of therapist contact

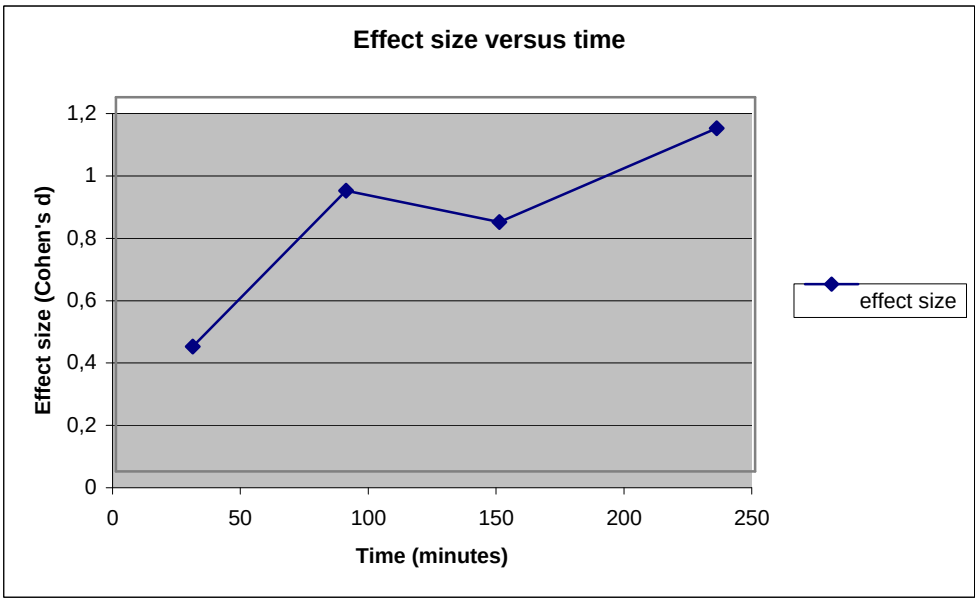
The amount of therapist contact is, as expected, an important variable.

If, for example, we glance at all the depression studies we note that those with no therapist contact render effect sizes from zero to 0.7, the one study with two hours therapist contact an effect size of 0.9 and the study with intensive therapist contact (say 11 hours, if therapist time was one hour per week for the 11-week

duration) renders an effect size of 1.1. Although the law of diminishing returns seems to be in operation, the amount of therapist time still is a factor.

The studies in which the amount of therapist contact is consciously varied are those by Carlbring et al (2001) and Carlbring, Ekselius and Andersson (2003) in the treatment of PD. In the former study therapist contact was 90 minutes per client via e-mail and in the latter only 30 minutes. Although one can rightly argue that the nature of the contact was also varied, this was probably the result of the reduction in contact time and not vice versa. The drop in effect size is dramatic: From 0.9 to 0.4. If we add the other PD studies by Carlbring et al. (2005) –150 minutes and an effect size of 0.8- and by Carlbring et al. (in press) –235 minutes and an effect size of 1.1- we see the trend confirmed, as shown in Figure 1 below.

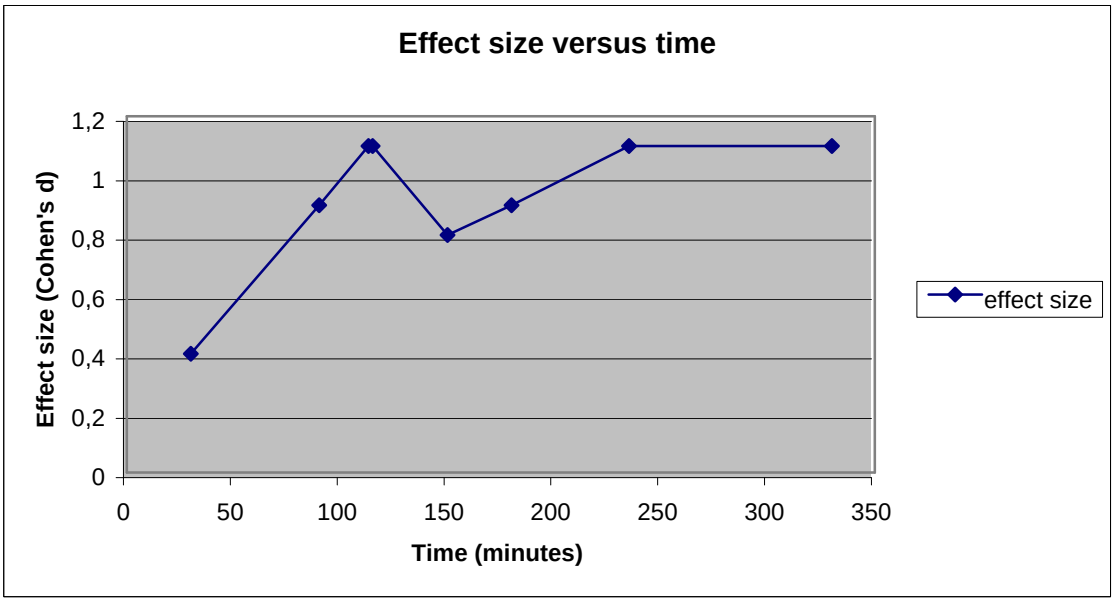
Figure 1. Effect sizes as a function of therapist time in the PD studies by Carlbring et al. and Andersson et al.



If we add the data of the other minimal contact anxiety studies a similar trend is apparent: As the amount of therapist contact time increases, the effect size increases

reaching a plateau around 1.1, which seems to be about the best effect size for minimal contact interventions via the internet (see Figure 2).

Figure 2. Effect size versus therapist time for minimal contact anxiety interventions



4.3 Nature of therapist contact

The nature of the therapist contact is a little stickier to interpret. Does the nature vary according to the medium, for example? Is the *nature* of the contact different by telephone than by e-mail? Probably it is: When writing an e-mail the therapist can choose his or her words more carefully, consult with colleagues, etc., whereas by telephone this possibility simply does not exist. If we compare the *types* of therapist contact in the studies, we have those with no contact whatsoever, those with e-mail contact, those with telephone contact and those with web-site contact and one with a combination of e-mail and telephone.

Those interventions without any contact can best be compared with bibliotherapy, except that the material is presented via the Internet instead of in a book. If we look at the four depression studies with no therapist contact, we see an average effect size of 0.6, which compares well with the envisioned “information-

only" control group of the study by Christensen, Griffiths, and Jorm (2004), with an effect size of 0.5.

Telephonic contact was used in the studies by Richards and Alvarenga (2002), Kenwright et al. (2004), Schneider et al. (2005) and in combination with e-mail contact in the study by Carlbring et al. (in press). With the exception of the study by Richards and Alvarenga (2002) with an effect size of only 0.3, these studies have comparable effect sizes of around 1.1.

E-mail contact is the most-used type of contact, rendering effect sizes between 0.4 and 1.1. The variation seems to be best explained by the *amount* of therapist time spent by e-mail (see 4.2 above).

Website-based contact is used by Lange et al (2001, 2003 and 2005). In the studies for treatment of PTSD (Lange et al., 2001 and Lange et al., 2003) the effect size ranges from 0.9 to 1.0 and for depression (Lange et al., 2005) 1.1, all when compared to psycho-education control. If we add the psycho-education effect in the depression study (psycho-education was given in both the active condition and the control condition) we have an effect size of 1.8 –which is a very large effect- arguably as good or better than a face-to-face delivered CBT intervention.

If we rank the types of therapist contact according to depth of personal contact, we would probably do so in the order: no contact (average effect size 0.6), e-mail contact (average effect size 0.9), mixed email/telephone (average effect size 1.1), telephone (average effect size 0.8) and website-based (average effect size 1.0). Although the average effect sizes do not exactly follow this suggested trend, one should take into account that these averages include pilot studies and studies with very small sample sizes.

4.4 Mechanisms to prevent dropout or stimulate completion

What makes an intervention efficacious is of course very difficult to ascertain if the material that is supposed to make up the body of the intervention is not worked through. The effect size suffers accordingly if less of the material is worked through. This is illustrated well in the studies by Christensen, Griffiths and Jorm (2004) where only half of the assignments were actually completed by those in the active (CBT) condition (concomitant effect size 0.7) and by Carlbring, Ekselius and Andersson (2003), where a similar amount of material was completed (concomitant effect size 0.4). The opposite is also true: In the study of Carlbring et al. (2005) 7.4 of the 10 modules were completed on average (effect size 0.8) whereas in their later study (Carlbring et al., in press) 8.9 of the 10 modules were completed on average (effect size 1.1).

What then works and what does not work? A timed delivery seems to have a positive effect. In the studies by Carlbring et al. we see this illustrated: No time limit or timed delivery in the study by Carlbring, Ekselius and Andersson (2003) and an effect size of 0.4; compared with a ten-week time limit and no access to the next module until the previous one has been completed and evaluated by the therapist in the study by Carlbring et al. (2005), effect size 0.8 and Carlbring et al. (in press), effect size 1.1.

Reminders seem to have a positive effect too. This can already be seen in the studies by Clark et al. (2002) –no reminders and a non-significant effect size- and Clark et al. (2005) – with reminders and an effect size of 0.5. But what is the nature of a reminder? A weekly telephone call by the therapist to augment the web-delivered material is not only of therapeutic value, but also has a reminder function. Regular e-mail contact fulfils the same function. Perhaps the active ingredient in reminders is

ensuring regular attention is given to the client. This seems to be supported in the study by Clarke et al. (2005) where no significant difference was found between reminding clients by telephone or reminding them by postcard.

Rewarding clients for filling in post-baseline assessments as Clark et al. (2005) did, still only resulted in less than 70% of their clients actually returning them.

The conclusion seems to be that ensuring regular *attention* from the therapist or non-clinical staff has a positive effect, as does ensuring structured, milestone-based (that is, timed delivery with completion check) delivery of the material.

4.5 Screening issues

Screening is another sticky issue. Firstly, how does one know that those randomised into the various conditions are a good reflection of those needing help? If we look at the number of respondents and the number eventually randomised, it seems as if many more respond than are actually randomised: On average only 34% of respondents are randomised (range 4% to 88%).

Secondly, what is the difference in therapeutic effect if screening is done on-line, by e-mail, pencil-and-paper or face-to-face? No answer is directly at hand.

5. Conclusion.

Website-based interventions seem equally efficacious as face-to-face interventions. This does not mean that ‘everything goes’, as there are specific ingredients that make an intervention efficacious. Of course, the material and how it is presented does have an effect. But just as in common factor research, part of the effect is due to the therapeutic relationship. On a website-based intervention this therapeutic relationship *is* created, albeit a different relationship that is created in traditional face-to-face

settings. Some of the active ingredients have been extracted in this paper. The amount and nature of the therapist contact is of very direct influence on the effect size. On the amount a certain plateau effect can be seen – future research might spend time in varying this ingredient. As to whether the contact is by telephone, e-mail or on-line we can be brief: the medium is of lesser importance than the message, or in other words, as long as the client feels attended to. The structure is important. Well-defined modules concluded by essay questions or homework that needs to be handed in and evaluated by the therapist work positively. Open-endedness and vagueness do not motivate.

Search strategy

Literature for this study was sought on the internet, amongst others using Psychinfo, Pubmed, Picarta, Academic Search Primer, Education Resources Information Center and Science Direct and Google.

The preliminary search terms used were

- "internet" and "therapy" and "depression"
- "internet" and "therapy" and "anxiety"
- "email" and "therapy" and "depression"
- "email" and "therapy" and "anxiety".

The word "internet" has also been replaced with "website-based", "internet-based", "web-driven", "online", "on-line" and "e-therapy".

The word "therapy" has also been replaced with "intervention" and "interventions"

The word "email" has also been replaced with "E-mail"

The word "anxiety" has also been replaced with "phobia", "phobic", "PTSD", "posttraumatic stress disorder", "panic disorder"

Furthermore, literature lists of found documents were scanned for possible other interesting articles.

Literature

Andersson, G., Bergström, J., Holländare, F., Carlbring, P., Kaldo-Sandström, V., & Ekselius, L. (2005). Internet-based self-help for depression: a randomised controlled trial. *British Journal of Psychiatry, 187,* 456 - 461.

Andersson, G., Carlbring, P., Holström, A., Sparthan, L., Furmark, T., Nilsson-Ihrfelt, E., Buhrman, M., & Ekselius, L. (in press). Internet-based self-help with therapist feedback and in-vivo group exposure for social phobia: a randomised controlled trial. *Journal of Consulting and Clinical Psychology.*

Andrews, G., & Slade, T. (2001). Interpreting scores on the Kessler psychological distress scale. *Aust N Z J Public Health, 25,* 494 - 497.

Arnold, J., Cooper, C. L., & Robertson, I. T. (1998). *Work Psychology: Understanding human behaviour in the workplace.* 3rd edition. Harlow: Pearson Education.

Arrindell, W. A., & Ettema, J. M. (2003). *SCL-90: Handleiding bij een multidimensionele psychopathalogie-indicator [Manual for a multi-dimensional indicator for psychopathology]*. Lisse: Swets & Zeitlinger.

Baker, S. L., Heinrichs, N., Kim, H.J., & Hofmann, S. G. (2002). The Liebowitz social anxiety scale as a self-report instrument: A preliminary psychometric analysis. *Behaviour Research and Therapy, 40,* 701 - 715.

Barlow, D. H. (2002). *Anxiety and its disorders*. 2nd edition. New York/London: Guildford.

Barlow, D.H., & Craske, M.G. (2000). *Mastery of your anxiety and panic: MAP-3.* New York: Graywind Publications.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *Journal of Consulting and Clinical Psychology, 56,* 893 - 897.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561 - 571.

Bouman, T. (1994). Kort Instrumenteel; de Beck Depression Inventory (BDI) [The Beck Depression Inventory]. *Gedragstherapie, 27*(1), 69 - 71.

Carlbring, P., Bohman, S., Brunt, S., Buhrman, M., Westling, B. E., Ekselius, L., & Andersson, G. (in press). Remote treatment of panic disorder: A randomised trial of Internet-based cognitive behavioral therapy supplemented with telephone calls. *American Journal of Psychiatry.*

Carlbring, P., Ekselius, L., & Andersson, G. (2003). Treatment of panic disorder via the internet: A randomised trial of CBT vs. applied relaxation. *Journal of Behavior Therapy and Experimental Psychiatry, 34,* 129 - 140.

Carlbring, P., Furmark, T., Steczkó, J., Ekselius, L., & Andersson, G. (2006). An open study of internet-based bibliotherapy with minimal therapist contact via email for social phobia. *Clinical Psychologist, 10,* 30 - 38.

Carlbring, P., Nilsson-Ihrfelt, E., Waara, J., Kollenstam, C., Buhrman, M., Kaldo, V., Söderberg, M., Ekselius, L., & Andersson, G. (2005). Treatment of panic disorder: live therapy vs. self-help via the internet. *Behaviour Research and Therapy, 43,* 1321 - 1333.

Carlbring, P., Westling, B. E., Ljungstrand, P., Ekselius, L., & Andersson, G. (2001). Treatment of panic disorder via the internet: A randomised trial of a self-help program. *Behavior Therapy, 32,* 751 - 764.

Chambless, D. L., Caputo, G. C., Bright, P., & Gallagher, R. (1984). Assessment of fear in agoraphobics: The Body Sensations Questionnaire and the Agoraphobic Cognitions Questionnaire. *Journal of Consulting and Clinical Psychology, 52,* 1090 - 1097.

Chambless, D. L., Caputo, G. C., Jasin, S. E., Gracely, E. J., & Williams, C. (1985). The Mobility Inventory for Agoraphobia. *Behaviour research and therapy, 23,* 35 - 44.

Christensen, H., Griffiths, K. M., & Jorm, A. F. (2004). Delivering interventions for depression by using the internet: randomised controlled trial. *BMJ,* doi:10.1136/bmj.37945.566632.EE.

Christensen, H., Griffiths, K. M., & Korten, A. (2002). Website-based cognitive behavior therapy: Analysis of site usage and changes in depression and anxiety scores. *J Med Internet Res 2002;4*(1):e3. <URL: http://www.jmir.org/2002/1/e3/>.

Clark, D. M., Salkovskis, P. M., Breitholtz, E., Westling, B. E., Ost, L. -G., Koehler, K. A., Jeavons, A., & Gelder, M. (1997). Misinterpretation of body sensations in panic disorder. *Journal of Consulting and Clinical Psychology, 65,* 203 - 213.

Clarke, G., Eubanks, D., Reid, E., Kelleher, C., O'Connor, E., DeBar, L. L., Lynch, F., Nunley, S. & Gullion, C. (2005). Overcoming depression on the internet (ODIN) (2): A randomised trial of a self-help depression skills program with reminders. *J Med Internet Res 2005;7*(2):e16. <URL: http://www.jmir.org/2005/2/e16/>.

Clarke, G., Reid, E., Eubanks, D., O'Connor, E., DeBar, L. L., Kelleher, C., Lynch, F., & Nunley, S. (2002). Overcoming depression on the internet (ODIN): A randomized controlled trial of an internet depression skills intervention program. *J Med Internet Res 2002;4*(3):e14 <URL: http://www.jmir.org/2002/3/e14/>.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates.

De Beurs, E., Van Dyck, R., Marquenie, L. A., Lange, A., & Blonk, R. W. B. (2001). De DASS; een vragenlijst voor het meten van depressie, angst en stress [The DASS; a questionnaire for the measurement of depression, anxiety and stress]. *Gedragstherapie, 34*(1), 35 - 53.

Di Nardo, P., Brown, T. A., & Barlow, D. H. (1994). *Anxiety Disorder Interview Schedule for DSM-IV.* San Antonio, TX: The Psychological Corporation.

Finn, S., & Tonsager, M. (1992). Therapeutic effects of providing MMPI-2 test feedback to college students awaiting therapy. *Psychological Assessment , 4 ,* 278–287.

Frisch, M. B., Cornell, J., Villanueva, M., & Retzlaff, P. J. (1992). Clinical validation of the Quality of Life Inventory. A measure of life satisfaction for use in treatment planning and outcome assessment. *Psychological Assessment, 4,* 92 - 101.

Furmark, T., Tillfors, M., Everrz, P. O., Marteinsdottir, I., Gefvert, O., & Fredrikson, M. (1999). Social phobia in the general population: Prevalence and sociodemographic profile. *Social Psychiatry and Psychiatric Epidemiology, 34,* 416 - 424.

Goldberg, D., Bridges, K., Duncan-Jones, P., & Grayson, D. (1988). Detecting anxiety and depression in general medical settings. *BMJ 1988 Oct 8*;297(6653): 897 - 899.

Kenwright, M., Marks, I. M., Gega, L., & Mataix-Cols, D. (2004). Computer-aided self-help for phobia/panic via internet at home: a pilot study. *British Journal of Psychiatry, 184,* 448 - 449.

Kleber, R. J., & Brom, D. (1986). *Traumatische ervaringen, gevolgen en verwerking [Traumatic experiences, consequences and processing]*. Lisse: Swets & Zeitlinger.

Klein, B., Richards, J. C., & Austin, D. W. (in press). Efficacy of internet therapy for panic disorder. *Journal of Behavior Therapy and Experimental Psychiatry*.

Lambert, M. J. (2004). *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed.). New York: Wiley.

Lange, A., Rietdijk, D., Hudcovicova, M., Van de Ven, J. -P., Schrieken, B., & Emmelkamp, P. M. G. (2003). Interapy: A controlled randomized trial of the standarized treatment of posttraumatic stress through the internet. *Journal of Consulting and Clinical Psychology, 71,* 901 - 909.

Lange, A., Van de Ven, J. -P., Schrieken, B., & Emmelkamp, P. M. G. (2001). Interapy. Treatment of posttraumatic stress through the internet: A controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry, 32*, 73 - 90.

Lange, A., Vermeulen, H., Renckens, C. H., Schrijver, M., van de Ven, J. -P., Schrieken, B., & Dekker, J. (2005). De geprotocolleerde Interapy-behandeling van depressie via het internet; resultaten van een gerandomiseerde trial [The protocol-driven Interapy treatment of depression via the internet; results of a randomised trial]. *Directieve therapie, 25*, 27 - 50.

Liebowitz, M. R. (1987). Social phobia. *Modern Problems of Pharmacopsychiatry, 22*, 141 - 173.

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for depression, anxiety and stress scales.* Sydney, Australia: The Psychological Foundation of Australia.

Marks, I. M. (1986). *Behavioural Psychotherapy: Maudsley Pocket Book of Clinical Management*. Bristol: John Wright.

Marks, I. M., & Mathews, A. M. (1979). Brief standard self-rating for phobic patients. *Behav Res Ther, 23*, 563 - 569.

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy, 36*, 455 – 470.

Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. M. (2002). The work and social adjustment scale: a simple measure of impairment in functioning. *British Journal of Psychiatry, 180*, 461 - 464.

Norcross, J. C. (Ed.). (2002). *Psychotherapy relationships that work. Therapist contributions and responsiveness to patients.* Oxford/New York: Oxford University Press.

Norton, G. R., Dorward, J., & Cox, B. J. (1986). Factors associated with panic in non-clinical subjects. *Behavior Therapy, 17*, 239 - 252.

Paul, G. (1966). *Insight versus desensitisation in psychotherapy: An experiment in anxiety reduction.* Palo Alto, CA: Stanford University Press.

Proudfoot, J. G. (2004). Computer-based treatment for anxiety and depression: is it feasible? Is it effective? *Neuroscience and Biobehavioral Reviews, 28,* 353 - 363.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Appl Psychological Measurement, 1*, 383 - 401.

Richards, J. C., & Alvarenga, M. E. (2002). Extension and replication of an internet-based treatment program for panic disorder. *Cognitive Behaviour Therapy, 31*, 41 - 47.

Riess, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behavior Research and Therapy, 29*, 161 - 166.

Schmidt, N. B., Lerew, D. R., & Trakowski, K. H. (1997). Body vigilance in panic disorder: evaluating attention to bodily perturbations. *Journal of Consulting and Clinical Psychology, 65*, 214 - 220.

Schneider, A. J., Mataix-Cols, D., Marks, I. M., & Bachofen, M. (2005). Internet-guided self-help with or without exposure therapy for phobic and panic disorders: a randomised controlled trial. *Psychotherapy and Psychosomatics, 74*, 154 - 164.

Shear, M. K., Brown, T. A., Barlow, D. H., Money, R., Sholomskas, D. E., Woods, S. W., Gorman, J. M., & Papp, L. A. (1997). Multicenter collaborative panic disorder severity scale. *American Journal of Psychiatry, 154*(11), 1571 - 1575.

Spitzer, R. L., Williams, J. B. W., Kroenke, K., Linzer, M., & DeGruy, F. V., III. (1994). Utility of a new procedure for diagnosing mental disorders in primary care: the Prime-MD 1000 study. *Journal of the American Medical Association, 272*, 1749.

Svanborg, P., & Åsberg, M. (1994). A new self-rating scale for depression and anxiety states based on the comprehensive psychopathological rating scale. *Acta Psychiatrica Scandinavica, 89,* 21 – 28.

Trijsburg, R. W., Colijn, S., & Holmes, J. (2005). Psychotherapy integration. In G. O. Gabbard, J. S. Beck, & J. Holmes (Eds.), *Oxford textbook of psychotherapy* (ch. 10; pp. 95 - 107). Oxford: Oxford University Press.

Wald, F .D., & Mellenbergh, G. J. (1990). De verkorte versie van de Nederlandse vertaling van de profile of mood states [The short version of the Dutch translation of the Profile of Mood States]. *Nederlands Tijdschrift voor de Psychologie, 45,* 86 - 90.

World Health Organization. (1999). *The World Health Organization Composite International Diagnostic Interview Short-Form (CIDI-sf) version 1.1*. Retrieved 2 May 2006 from http://www.who.int/msa/cidi/cidisf.htm.

World Health Organisation (1992). *International statistical classification of diseases and related health problems (ICD-10).* Geneva: WHO.