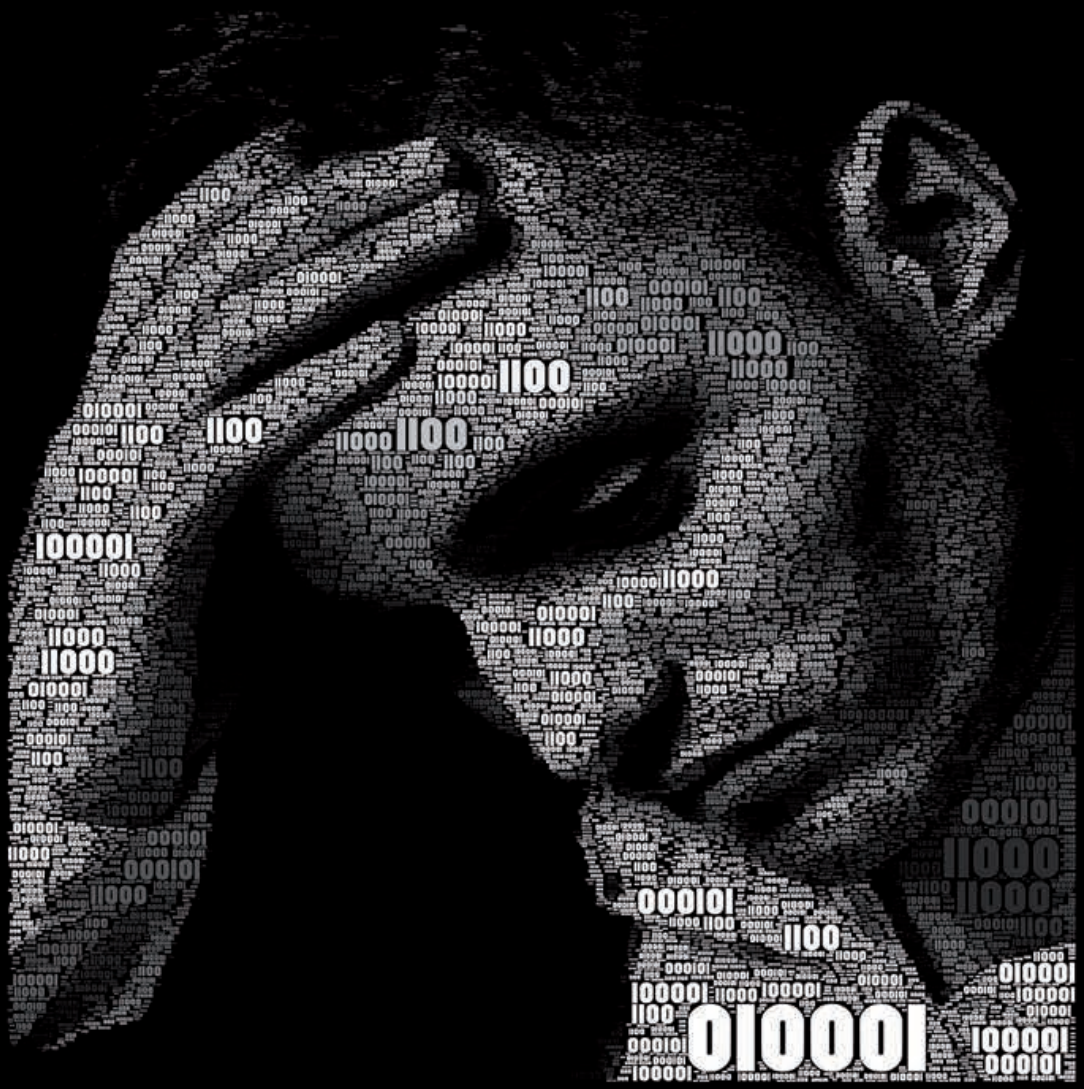
The Efficacy and Effectiveness of Online CBT

Jeroen Ruwaard



The Efficacy and Effectiveness of Online CBT

Jeroen Ruwaard

The studies in this thesis were supported by the Dutch Foundation of Mental Health (NFGV), ZonMw, the Netherlands Organisation for Health Research and Development, the Dutch Innovatiefonds Zorgverzekeraars and Interapy PLC.

Ruwaard, J. (2012). *The efficacy and effectiveness of online CBT*. Amsterdam: Department of Clinical Psychology, University of Amsterdam.
ISBN: 978-94-6191-588-7

Cover photo by Dundanim http://www.dreamstime.com/Dundanim_info.
Typeset in Latex, using the ILLC Dissertation Style package.
Printed by IpskampDrukkers BV, Enschede.

The Efficacy and Effectiveness of Online CBT

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom,
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit
op vrijdag 8 februari 2013, te 11:00 uur

door

Jan-Jeroen Ruwaard

geboren te Ede

Promotiecommissie:

Promotores	: Prof. dr. A. Lange
	: Prof. dr. P.M.G. Emmelkamp
Co-promotor	: Prof. dr. C.V. Dolan
Overige leden	: Prof. dr. G. Andersson
	: Prof. dr. W.J.M.J. Cuijpers
	: Prof. dr. A.W.M. Evers
	: Prof. dr. P. Prins
	: Prof. dr. A.H. Schene
	: Prof. dr. R.W.H.J. Wiers

Faculteit der Maatschappij- en Gedragswetenschappen

Contents

Chapter 1
General Introduction

In 1997, researchers at the University of Amsterdam developed one of the first psychotherapeutic applications of the World Wide Web. They implemented a standardized cognitive behavioural treatment (CBT) of posttraumatic stress symptoms in a website, and used this site to treat a small number of students with matching symptoms. To the surprise of the research team, the results of what they called *Interapy* were very encouraging. Despite the lack of face-to-face contact, symptoms of 19 of the 20 students had been reduced to normal levels after treatment (Lange, van de Ven, Schrieken, Bredeweg, & Emmelkamp, 2000). Subsequently, controlled trials confirmed the value of this pilot study, and showed that a large part of these improvements could be attributed to the intervention. In this dissertation, we explore the wider applicability of online CBT, in four randomized controlled trials (Chapter 2 to 5) and a practice study (Chapter 6). In the controlled studies ($N = 456$), we assess the efficacy of online CBT for work-related stress, mild to moderate depression, and symptoms of panic disorder and bulimia nervosa. In the fifth study, we examine the effectiveness of online CBT in routine clinical practice ($N = 1500$).

1.1 Online CBT

While effective treatments exist for most common mental health disorders (Roth & Fonagy, 2005), evidence suggests that only a minority of those affected are reached (World Health Organization, 2001; Lancet Global Mental Health Group, 2007). Patients do not seek help, are unable to access it, or receive treatment that does not correspond to current practice guidelines (Shafran et al., 2009). The barriers are

psychological, social, geographical, and systemic (Collins, Westra, Dozois, & Burns, 2004). Internet applications have reduced or removed such barriers in many areas of life. Psychotherapeutic internet interventions may support public mental health-care in a similar vein (F. Griffiths, Lindenmeyer, Powell, Lowe, & Thorogood, 2006; Emmelkamp, 2005).

In the past decade, researchers have developed psychotherapeutic internet interventions for a wide variety of mental health disorders (Barak, Hen, Boniel-Nissim, & Shapira, 2008). These interventions vary in terms of the intended purpose (prevention, early intervention, treatment, aftercare), the electronic means that are used to deliver the intervention (e-mail, e-chat, websites, video-conferencing, mobile applications), and in the level of professional support. In terms of the psychotherapeutic orientation, available interventions are surprisingly homogeneous: the vast majority of existing internet interventions are based on Cognitive Behavioural Therapy (CBT).

CBT is characterized by clear-cut therapeutic procedures, which are relatively easy to translate into an online format. In addition, there is general agreement on the effectiveness of CBT (Butler, Chapman, Forman, & Beck, 2006). Meta-analyses suggest that this effectiveness generalizes to internet interventions, although it should be noted that these meta-analyses also reveal significant heterogeneity in study outcomes (Cuijpers et al., 2009; Andersson & Cuijpers, 2009). There is growing consensus that professional guidance is a critical determinant of the efficacy of online CBT (Palmqvist, Carlbring, & Andersson, 2007; Spek et al., 2007).

1.1.1 Three types of online CBT

Self-help. Online self-help interventions do not include guidance from a mental health professional. Such interventions allow large-scale, low-cost dissemination of CBT. Online self-help appears to be effective for a variety of psychological symptoms, although its effects are statistically small and drop-out rates are very high (Kaltenthaler et al., 2006; Foroushani, Schneider, & Assareh, 2011; Spek et al., 2007; Andersson & Cuijpers, 2009). Even so, online self-help interventions are currently implemented in routine practice for preventive purposes, based on the assumption that interventions with small effects may have a noticeable impact if available to a large part of the population.

Guided self-help. Most internet interventions offer a mix of self-help and human guidance. In guided self-help, mental health professionals support clients in the use of the self-help material, through e-mail, e-chat, telephone support, or face-to-face sessions. This support is intended to be limited to approximately 60-90 minutes per treatment (e.g., Titov et al., 2010). The interventions are therefore alternatively labelled as 'low-intensity CBT' (Bennett-Levy et al., 2010). In comparison to online self-help, guided self-help is more expensive and less accessible. However, the professional support substantially increases the efficacy of, and adherence to, an intervention (Cuijpers et al., 2009; Andersson & Cuijpers, 2009; Spek et al., 2007).

Psychotherapy. Online psychotherapy is an extension of online guided self-help. Like guided self-help, online psychotherapy promotes self-efficacy through a combination of homework assignments and psycho-education. However, the role of the therapist is not limited to providing support or help if the client gets stuck. The online psychotherapist has the same role as the regular face-to-face CBT therapist (Barak, Klein, & Proudfoot, 2009): to explain each step of treatment, to give tailored feedback to the client feedback after each step, and to introduce the next step. Online treatment may be highly manualized, and thus provides therapists with specific guidelines and help in treating patients.

1.2 Interapy

The Interapy research program, which we introduced in the beginning of this chapter, represents one of the few systematic attempts to deliver full online psychotherapy, i.e., online CBT with extended therapist support that does not require any face-to-face contact between the client and a mental health professional.

1.2.1 General characteristics

In Interapy, clients and therapists work on treatment in convenient locations with internet access, including the home. The therapeutic procedure is delivered completely through a secure website. Clients are assessed in an elaborate online screening procedure, in which standardized and validated self-report questionnaires query back-

ground data, symptom severity and contra-indications for online treatment. During the treatment, clients and therapists interact through an asynchronous exchange of text messages via web-based e-mail. Their dialogue is governed by a highly detailed treatment manual, which defines a fixed sequence of homework assignments that implement common cognitive-behavioural therapeutic techniques. Therapist support consists of standardized, default feedback and instructions that are tailored by the therapists to the specific situation and specific alliance of the client. In this feedback, motivational techniques are applied to enhance the impact of the intervention, i.e., to ensure clients understand the purpose of the homework, set realistic goals, do their exercises as prescribed, and continue the treatment (Lange, 2006, chapter 3; van der Velden, Hoogduin, & Lange, 2010). These motivational techniques, which are seen as the 'common factors' of psychotherapy (Messer & Wampold, 2002), target clients' motivation for change, the therapeutic alliance (e.g., by expressing empathy and understanding), and self-esteem and self-efficacy (e.g., by complimenting the clients on their progress and accomplishments, instead of confronting them when their homework is insufficient).

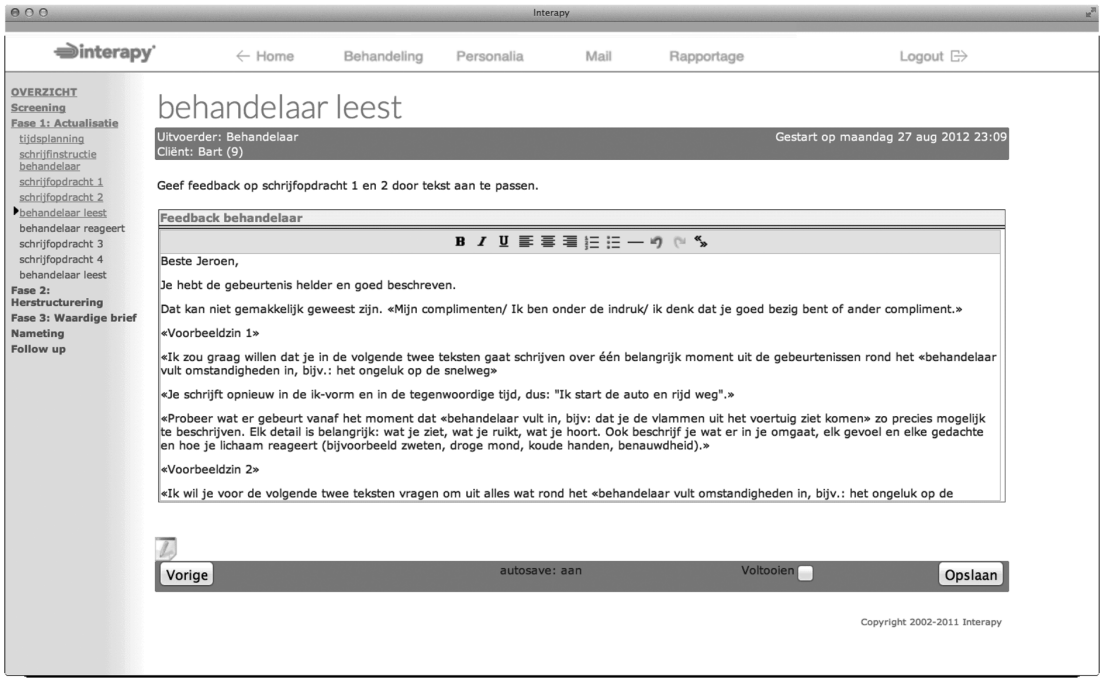


Figure 1.1: Screenshot of the Interapy treatment website.

1.2.2 Interapy for posttraumatic stress

The general principles of Interapy were developed from research into the treatment of posttraumatic stress (Lange, Van de Ven, Schrieken, & Schoutrop, 2002). Interapy addresses symptoms of posttraumatic stress through five weeks of online CBT, in which clients receive psycho-education, structured writing exercises, and regular instructions and feedback from a personal therapist. The treatment comprises three phases, which incorporate imaginatory exposure (phase I), cognitive reappraisal (phase II) and social sharing (phase III).

Interapy for posttraumatic stress was first tested in 1997, in the seminal pilot study by Alfred Lange and colleagues mentioned above (Lange, van de Ven, et al., 2000). Since results of this pilot were almost too good to be true, the researchers checked for errors in their data and their analyses. As these checks revealed no errors, the online intervention was judged to be surprisingly effective.

To determine the controlled effects of the online treatment, the Interapy research group conducted two randomized waiting-list controlled trials (Lange, van de Ven, Schrieken, & Emmelkamp, 2001; Lange, Rietdijk, et al., 2003). In both trials, online CBT was more effective in comparison to the waiting-list, with large between-group effect sizes relating to symptoms of posttraumatic stress and general psychopathology. Subsequent follow-ups showed that gains sustained up to 18 months after treatment (Lange et al., 2002; Lange, van de Ven, & Schrieken, 2003). These findings were replicated in two RCTs by an independent Swiss/German research group (Knaevelsrud & Maercker, 2007, 2009; Wagner, Knaevelsrud, & Maercker, 2006; Wagner & Maercker, 2007). Recent trials indicate that the treatment can be successfully used to reach out to adolescent victims of sexual abuse (Lange et al., 2011), and to victims of war and human rights violations in remote conflict areas such as Iraq (Wagner, Schulz, & Knaevelsrud, 2012).

1.3 This dissertation

1.3.1 Research aims

The research aims of this dissertation were two-fold:

1. To assess the efficacy of online cognitive behavioural treatment for common mental health disorders other than posttraumatic stress.

Given the positive effects of Interapy for posttraumatic stress, we had good reasons to think that online CBT would provide an effective treatment for other mental health disorders. Evidence indicated that computer-mediated therapy allowed for an effective therapeutic relationship, and that evidence-based therapeutic procedures could be delivered online. But there were, of course, no guarantees that Interapy would be equally efficacious in the management of disorders that are different from posttraumatic stress. Hence, we developed new protocols for the online treatment of four common mental health disorders on the basis of the general principles on which Interapy for posttraumatic stress was based, and checked the efficacy of these protocols in a series of randomized controlled trials.

2. To assess the effectiveness of online cognitive behavioural treatment in routine clinical practice.

Favourable results in clinical trials do not necessarily imply favourable results in routine clinical practice. RCTs provide information about the efficacy of a treatment, i.e., the capacity of a treatment to produce an effect under ideal conditions. But efficacy should not be confused with effectiveness, which refers to the capacity of the treatment to produce positive effects under routine practice conditions (Andersson, Carlbring, & Cuijpers, 2009; Nathan, Stuart, & Dolan, 2000; Seligman, 1995). Findings suggest that effects are generally better under controlled conditions, i.e., that efficacy $>$ effectiveness (Westbrook & Kirk, 2005). It was therefore important to check the degree to which the efficacy of online CBT, as determined in the controlled trials, would translate to treatment effectiveness in the routine clinical practice setting.

1.3.2 Outline

The dissertation comprises five studies. Chapters 2 to 5 describe four randomized controlled trials of online CBT for symptoms of work-related stress, depression, panic disorder, and bulimia nervosa. Chapter 6 presents the effects of online CBT in routine clinical practice.

Online CBT for work-related stress. Chapter 2 describes the controlled evaluation of online CBT for work-related stress. We developed this program in 2001. At that time, work-related stress had been identified by the European Union (EU) as the second most common work-related health problem, affecting an estimated 28% of EU workers (possibly a greater percentage in the Netherlands; Schaufeli & Kompier, 2001), accounting for a quarter of long-term absences from work, and amounting to a staggering 20 billion euros of annual costs to EU member states (Paoli & Merllie, 2001). A meta-analysis had identified individual-focused CBT as an effective intervention for work-related stress (van der Klink, Blonk, Schene, & van Dijk, 2001). Based on these findings, we developed a treatment protocol with which employees could be trained, over the internet, to apply new coping skills and stress reduction techniques in their work-environment.

Online CBT for depression. Chapter 3 presents the randomized trial of online CBT for depression, which we started in 2003. Colleagues in Australia and the UK had reported some success with online and computer-supported self-help in the treatment of depressive symptoms (Christensen, Griffiths, & Korten, 2002; Christensen, Griffiths, & Jorm, 2004; Proudfoot et al., 2003, 2004), but these positive results were countered by the negative results of a large RCT, in which unsupported online CBT was ineffective (Clarke et al., 2002). We hypothesized that results were inconclusive because existing programs did not include therapist support. Hence, we set out to develop a web-based treatment for depressive symptoms with scheduled therapist feedback, and assessed the efficacy of this treatment in an RCT.

Online CBT for panic disorder. Since 2002, the Interapy research group had been experimenting with online CBT for symptoms of panic disorder and agoraphobia. This

treatment was delivered through e-mail. Although the research suggested that this treatment was potentially effective (Jager, Emmelkamp, & Lange, 2004), many clients dropped out (Jager, Emmelkamp, & Lange, 2005). This raised doubts whether this research should be continued. However, since drop-out appeared to be related to technical difficulties only, and since reports of online treatments for panic disorder were generally positive (Carlbring, Westling, Ljungstrand, Ekselius, & Andersson, 2001; Carlbring, Ekselius, & Andersson, 2003; Carlbring et al., 2005; Klein & Richards, 2001; Klein, Richards, & Austin, 2006; Marks, Kenwright, McDonough, Whittaker, & Mataix-Cols, 2004), we decided to redesign the treatment as a web-based intervention, instead of depending on e-mail communication. Chapter 4 provides a description of this treatment and of the RCT that we conducted to assess its efficacy.

Online CBT for Bulimia Nervosa. Chapter 5 describes a fourth RCT, which we started in 2007. In this study, we assessed the efficacy of online CBT for Bulimia Nervosa. Since sufferers from Bulimia Nervosa tend to hide their eating habits out of guilt and shame (Fairburn & Cooper, 1982; Fairburn & Harrison, 2003), we supposed that online CBT could provide an acceptable low-threshold treatment alternative. At that time, a number of computer-supported and online (self-help) programs for eating disorders existed, but the effects of these programs were found to be modest at best (Marks, Cavanagh, & Gega, 2007, chapter 6). In contrast, a Swedish research group found strong effects of an online guided self-help treatment for bulimic symptoms (Ljotsson et al., 2007). Encouraged by these results, we developed an Interapy program and tested this program in an RCT. In contrast to the previous RCT's, we included three experimental comparison groups in this RCT. Participants were randomly allocated to online CBT, a waiting list, or to a pure (offline) self-help comparison group, in which participants received a self-help book. With the extra comparison group, we hoped to gain further insight in the role of therapist support.

Online CBT in routine clinical practice. In 2001, the University of Amsterdam and members of the Interapy research group founded the Interapy clinic. This was one of the first attempts to implement online mental healthcare in routine clinical practice. In 2008, the Interapy clinic had been offering online treatment for posttraumatic stress, work-related stress and depression for several years, and had just launched online

CBT for panic disorder. Outcome data, collected through routine outcome monitoring, started to provide a unique database. Although research in the field had been growing exponentially, most existing internet interventions had not been implemented in routine clinical practice. Some routine practice data had been published, but these concerned either online self-help (Christensen, Griffiths, Korten, Brittliffe, & Groves, 2004) or computer-supported therapy with face-to-face contact (e.g., Gega, Marks, & Mataix-Cols, 2004; Hayward, MacGregor, Peck, & Wilkes, 2007). While routine practice studies of therapist-assisted CBT were absent, the database of the Interapy clinic contained treatment outcome data of hundreds of clients. These data were waiting to tell how online CBT performed in routine clinical practice. Chapter 6 presents treatment outcome data of $N = 1500$ clients of the clinic.

Ruwaard, J., Lange, A., Bouwman, M., Broeksteeg, J., & Schrieken, B. (2007). E-Mailed Standardized Cognitive Behavioural Treatment of Work-Related Stress: A Randomized Controlled Trial. *Cognitive Behaviour Therapy, 36*(3), 179-192. doi: 10.1080/16506070701381863

Chapter 2

Online Cognitive Behavioural Treatment of Work-Related Stress

Objective To assess the effects of a 7-week standardized Cognitive Behavioural Treatment (CBT) of work-related stress conducted via e-mail. **Participants** 342 people applied for treatment in reaction to a newspaper article. Initial screening reduced the sample to a heterogeneous (sub)clinical group of 239 of participants. **Design** participants were randomly assigned to a waiting list condition (n = 62), or to immediate treatment (n = 177). A follow-up was conducted three years after the inception of the treatment. **Outcome Measures**: The Depression Anxiety Stress Scales (DASS-42) and the Emotional Exhaustion scale of the Maslach Burnout Inventory - General Survey (MBI-GS). **Results** Fifty participants (21%) dropped out. Both groups showed statistically significant improvements. Intention-to-treat ANCOVA's revealed that participants in the treatment condition improved significantly more than the participants in the waiting control condition ($.001 < P \leq .025$). In the treatment group, the effects were large to moderate [0.9 (stress) $\geq d \geq$ 0.5 (anxiety)]. The between-group effects ranged from d = 0.6 (stress) to d = 0.1 (anxiety). At follow-up, the effects were more pronounced, but this result requires replication in view of high attrition at follow-up. **Conclusion** The results warrant further research on Internet-driven standardized CBT for work-related stress. Such research should include the direct comparison of this treatment with face-to-face treatment, and should address the optimal level of therapist contact in Internet-driven treatment.

Introduction

Globalization, organizational restructuring, and the universal adoption of information technology have transformed the demands on employees in many organizations (Sparks, Faragher, & Cooper, 2001). The pressure to adapt to rapid technological and organizational demands has contributed to an increase in the prevalence of chronic stress among employees. Despite considerable efforts to counter this trend, the rates of work incapacity and sick leave due to work-related stress are high in the Netherlands compared with other European countries (Schaufeli & Kompier, 2001).

Stress is not necessarily harmful, i.e., a degree of stress is necessary to achieve desired levels of performance (van Doornen, 2001). However, stress may become a serious problem if people are exposed to high levels of stress over extended periods of time. Work-related stress is defined as excessive chronic stress caused by a mismatch between work-related demands and available coping skills (Health & Safety Executive, 2001). It is characterized by a variety of negative emotional and physiological reactions to aspects of the work itself, to the environment, and to the organization (Levi & Levi, 1995). Common psychological reactions to work-related stress are depression, anxiety, and emotional exhaustion (Cooper, Dewe, & O'Driscoll, 2001).

Cognitive Behavioural Therapy (CBT) and relaxation therapy are moderately effective in reducing work-related stress (van der Klink et al., 2001). In CBT, clients are instructed how to address the causes of their stress by changing their coping skills with regard to demands and pressures. In relaxation therapy, clients learn to lower their physiological arousal. In a meta-review, van der Klink et al. (2001) estimated the effects of CBT and relaxation therapy to be $d = .7$ and $d = .4$, respectively.

The implementation of these interventions, however, requires trained professionals, who are relatively few in number (World Health Organization, 2001). In addition, clients may be hesitant to seek help, due to limited mobility, travelling distance, or a lack of awareness concerning their condition, or the possibility of treatment. Finally, fear of stigmatization may discourage potential clients from seeking help. Given the obstacles to seeking face-to-face therapy, there is an increasing interest in methods to provide evidence-based treatment using the Internet (Barak, 1999; Emmelkamp, 2005; Newman, 2004). Online therapy may reduce the present burden on the available

mental health resources, and may provide a viable alternative for people, who face geographical, physical, psychological and/or financial barriers in seeking traditional, face-to-face care.

Critics of Internet-driven therapy have questioned its feasibility and have expressed ethical concerns (e.g. Bonger, 1988; Bauer, 2000). However, at present, the main question concerning online therapy is not whether, but rather how exactly it should be undertaken. Over the years, the accumulation of studies evaluating online therapy has resulted in a body of results. Reviews to date indicate that online therapy provides an effective alternative to existing methods for early intervention (e.g. K. M. Griffiths & Christensen, 2006; Wantland, Portillo, Holzemer, Slaughter, & McGhee, 2004). However, these reviews also show many challenging, open-ended issues that remain to be addressed. For example, the available evidence on the effectiveness of online therapy is limited to a relatively small number of disorders, and the effectiveness is found to vary considerably across studies (e.g. see Spek et al., 2007). Furthermore, dropout-rates tend to be high. Wantland et al. (2004) report an average dropout rate of 21%.

Studies vary greatly in the degree of therapist involvement. Some studies relied entirely on self-help material (e.g. Klein & Richards, 2001). Others explicitly involved a personal therapist, and included explicit structuring of the remote client-therapist interaction to facilitate the desired changes. The involvement of a therapist allows for personalization of the interventions and encourages the development of a therapeutic alliance (Knaevelsrud, Jager, & Maercker, 2004). In a meta-analysis of Internet-based interventions for anxiety and depression, Spek et al. (2007) found that the effects of interventions with therapist support were considerably larger than the effects of interventions without therapist support.

Zetterqvist, Maanmies, Ström, and Andersson (2003) tested the efficacy of e-mailed self-help for stress management with minimal therapist contact. The treatment involved relaxation training, development of problem-solving and time-management skills, cognitive restructuring, and behavioural exercise. Compared to a waiting list control group, participants in the programme displayed significantly greater improvement on the Perceived Stress Scale (PSS: S. Cohen, Kamarck, & Mermelstein, 1983) and the Hospital Anxiety and Depression Scale (HADS: Zigmond & Snaith, 1983).

The reported summary statistics suggest large to moderate (controlled) effect sizes of $d = .6$ (PSS), $d = .5$ (HADS Anxiety) and $d = .8$ (HADS Depression). Unfortunately, 26% of the participants dropped out and intention-to-treat analyses failed to reveal significant differences between the two groups.

Based on principles of online therapy developed earlier for the treatment of post-traumatic stress (Lange, Rietdijk, et al., 2003), we designed a standardized treatment of work-related stress, which can be conducted via e-mail. The treatment is based upon interventions that have proved their worth in clinical trials and clinical practice (de Jong & Emmelkamp, 2000; Lange, Richard, Gest, de Vries, & Lodder, 1998). The treatment involves psycho-education, awareness training, applied relaxation, cognitive restructuring, positive self-verbalization, social skills training, and time-management. The content of the treatment is similar to that of Zetterqvist et al. (2003). In contrast to the Zetterqvist et al. study, where client-therapist interaction was incidental and not standardized, the treatment manual includes explicit structuring of the client-therapist interaction. Specifically, it specifies the timing, nature, and frequency of the therapist instructions and feedback.

Below, we present the results of a randomized controlled trial that we conducted to assess the effects of our Internet-based standardized CBT on work-related stress. We compared the effects of treatment with those observed in a waiting list control group. We hypothesized that, in comparison to the waiting-list, the treatment would substantially reduce levels of stress. We also expected the treatment to reduce stress-related complaints, in particular depression, anxiety, and emotional exhaustion. In addition, we report the results of a three year follow-up study in which we assessed the long term effects among those who had completed treatment.

2.1 Treatment

The treatment comprised 7 phases, as described in detail below. Each phase required approximately 1 week to complete. The schedule allowed the participants to adjust the pace of completion within limits to their own situations. Treatment integrity was guaranteed by a manual that specified each step of the treatment in detail (e.g., the order and the nature of the assignments, the contents of the psycho-education, and

the timing of the therapist feedback). The manual also provided default text templates for feedback and instructions. The therapists tailored these texts to the specific needs of their clients. They were required to email their feedback after 1 working day. This provided the therapists with ample time to think through their feedback, or to discuss difficult cases with colleagues or a supervisor. The treatment manual specified 10 feedback-moments. Composing the feedback took about 30 minutes. Thus, a full treatment took about 5 hours of therapist time.

2.1.1 Treatment phases

Phase 1. Inducing awareness: monitoring and writing. This phase comprised one week of monitoring stress and two writing assignments to increase participants' awareness of somatic, cognitive, emotional, and behavioural signs of stress. Participants kept records of stressful events: they described the situation, associated feelings and thoughts, and rated the degree of stress on a 10-point scale. Subsequently, they conveyed their ratings graphically and reflected on the results. In addition, the participants spend two periods of 45 minutes writing about stressful events, in the present tense and in as great a detail as possible (Lepore & Smyth, 2002; Alford, Malouff, & Osland, 2005).

Phase 2. Relaxation. First, participants were taught muscle relaxation or mental relaxation, depending on their own preference. Subsequently, they practiced 6 times a day. In addition, they were encouraged to take short breaks during their work to reduce their overall pace, and to engage in physical exercise. Participants who reported sleeping complaints received instructions how to regain healthy sleeping habits: i.e., not sleeping during the day, not sleeping too long during the night, avoiding stressful activities (e.g., stressful conversations) before going to bed, and not trying to force themselves to sleep (Morin, 1993).

Phase 3. Worrying, rumination & challenging dysfunctional thoughts. Participants were taught to recognize events that trigger worrying and ruminating. Participants were instructed to limit ruminating to predetermined, fixed moments, rather than indulging in it at all hours of the day (McKay, Davis, & Fanning, 1997). When they

started worrying and ruminating spontaneously, they were required to make a short note, and to determine the next fixed moment for ruminating. Furthermore they were required to focus on their notes and to challenge dysfunctional (black/white) thinking, overgeneralization, catastrophic thinking, self-blaming, and neglect of positive aspects (Beck, 2005). They did this three times a week for 20 minutes.

Phase 4. Positive self-verbalization. Participants were taught to focus on their positive qualities by using the technique of positive self-verbalization (Lange et al., 1998; Lange, Richard, Kiestra, & van Oostendorp, 1997). They wrote an essay on their positive traits and summarized this on a small card. The participants were encouraged to read the summary on the card frequently, and read it aloud.

Phase 5. Positive assertiveness/social skills training and behavioural experiments. Participants received a short social skills training. They were encouraged to reflect on the reciprocity of the interactions with their fellow-workers. Subsequently, they received instructions how to experiment with new strategies to improve negative reciprocal behaviours. They were, for instance, taught to communicate wishes in a positive and constructive manner, to give compliments, and to express their interest in the well-being of others. If they questioned the validity of these strategies, the therapist helped to create behavioural experiments, in which the new approaches could be put to test in a simple and realistic manner.

Phase 6. Time management. Participants were made aware how they managed the demands of both their work and their private life. In the evening, participants reflected on the activities of the day, after which they planned the activities for the next day. The participants learned how to set priorities, to reject requests that they found too demanding, to plan time for relaxation and self-reward, and to avoid unrealistic scheduling. They were encouraged to take their time to decide consciously whether to start an activity immediately, postpone it, or to pass it on to someone else.

Phase 7. Future, reintegration and relapse prevention: the toolkit. Participants reflected on symptoms that might signal relapse. They were encouraged to formulate their personal 'relapse prevention toolkit', using the techniques they had mastered in

therapy. They were encouraged to print the toolkit, and place it in a visible place at home as a symbol of what they had mastered. Participants who were partially or fully on sick leave received suggestions for gradual reintegration.

2.2 Method

2.2.1 Design

The comparitive study comprised a waiting-list controlled pre-post trial. Participants were randomly assigned to two groups. One group started the 7-week treatment immediately (experimental group), while the other started it after 7 weeks (waiting list control group). Three years after the completion of the pretest, a follow-up study was conducted among the clients of the experimental group who had completed the therapy to determine the long-term effects of the intervention.

2.2.2 Participants

Recruitment. A national Dutch quality newspaper published an interview with one of the authors about Internet-driven therapy. The article announced the current study, and referred interested readers to a website. This site provided psycho-education on work-related stress, explained the purpose and design of the study, and contained an application form.

Screening. Respondents were screened by means of self-report questionnaires with respect to the following exclusion criteria: age $<$ 18 years, heightened risk of dissociation or psychosis, suicidal ideation, drug abuse, use of neuroleptic medication, and concurrent other treatment. In addition, respondents were required to download, print, and return a signed *Informed Consent form*. Excluded respondents were referred to other mental health institutions.

The participants completed a multiple-choice questionnaire to record age, gender, marital status, education, work contract, work years, working hours, duration of symptoms, sick leave status, and perceived causes of stress. In addition, alcohol dependency and drug abuse was registered. To assess the use of neuroleptics, participants answered open-ended questions concerning medicine brands and prescribed

doses. Risk of dissociation was assessed using the 5-item *Somatoform Dissociation Questionnaire* (SDQ-5; Nijenhuis, Spinhoven, van Dyck, van der Hart, & Vanderlinden, 1997). The internal consistency of this self-rate instrument is good (Cronbach's α = .80). A cut-off of 8 (used in this study) results in good discrimination between groups of patients and non-patients (sensitivity = .94; specificity = .96). Participants, who scored above the cut-off value of 13 on the *Screening Device for Psychotic Disorder* (SDPD; Lange, Schrieken, Blankers, van de Ven, & Slot, 2000 were excluded. The cut-off value of 13 was established in the Dutch norm group. The 7-item self-rate inventory (Cronbach's α = .82) is a good predictor of psychotic episodes. Agreement between self-report in a group of 33 schizophrenic patients and the ratings by their clinicians is high(r = .85) (Lange, Schrieken, et al., 2000). Suicidal ideation was measured by an inventory similar to that of Joiner et al. (2003). It comprises six multiple-choice questions, including "Do you currently have plans to end your life?" and "Are you currently feeling desperate?". Participants were excluded if they had a history of suicide attempt(s) within the past two years, or if there was immediate risk of suicide.

Randomization. One month after the publication of the newspaper article, participants were randomly assigned to the two experimental groups by means of the random number generator procedure of SPSS 10. Assuming a correlation of r = .5 between a pretest and posttest measurement, a significance level of .05 would require 49 participants per group to provide an 80% probability to detect a medium (d = .5) effect size (J. Cohen, 1988).

2.2.3 Procedure

Setting. Communications took place entirely through e-mail, without any face-to-face session. Participants underwent the treatment at home. Incidentally, telephone contact was necessary (e.g., to inquire about the reasons for dropout when e-mails remained unanswered).

Privacy. Several procedures secured the privacy of the participants. An e-mail server was set up exclusively for the study and located at a professional Internet host. The

server, running Linux OS, was protected by a firewall and remotely administered through an encrypted communication channel. The therapists worked from a central computer-room that was accessible only with proper authorization. Each therapist received a private password-protected account to login to the computers and to the e-mail server. Participants were informed of mail-client extensions to secure their e-mail, and received help in configuring these extensions on request.

Materials. Participants needed a personal computer connected to the Internet and an e-mail account. For screening and outcome measurement, validated questionnaires were delivered in the body text of e-mails and in documents that were attached to these e-mails. The attachments were formatted in a common cross-platform computer format (*Rich Text Format*).

Therapists. The therapists were 25 doctoral students in clinical psychology and one postgraduate student, aged between 22 to 42 years old (M = 26, SD= 4.7). They had followed advanced courses in CBT. Additional training taught the therapists how to tailor the feedback and instruction templates of the manual to the needs of their clients, how to increase motivation by adopting a stimulating empathetic attitude, how to avoid the pitfalls of electronic, text-based communication (e.g. Brennan & Ohaeri, 1999), and how to make use of the asynchronous nature of the communication to enhance the quality of the feedback (e.g., by discussing cases with one another or the supervisor). Once a week, a senior specialist in Internet-driven CBT supervised the therapists. The supervisor assigned participants to therapists on the basis of the availability of the therapists.

Posttest and follow-up. Immediately after treatment, participants received the posttest measurement (the outcome questionnaires and an evaluation questionnaire). Those who did not complete treatment received a telephone call to inquire about the causes of dropout. Dropouts were also asked to complete the posttest measurements. Three years after the pretest measurements, participants, who had completed treatment, were invited by e-mail to complete the outcome questionnaires.

2.2.4 Measures

The primary outcome measure was the Stress subscale from the *Depression Anxiety Stress Scales* (DASS-42: Lovibond & Lovibond, 1995; Dutch version: de Beurs, van Dyck, Marquenie, Lange, & Blonk, 2001). Secondary measures were the Depression and Anxiety subscales from the DASS and the Emotional Exhaustion subscale of the Dutch version of the *Maslach Burnout Inventory - General Survey* (MBI-GS: Maslach, Jackson, & Leiter, 1996; Dutch version by Schaufeli & van Dierendonck, 2000).

DASS. The DASS is a self-report instrument that assesses depressive symptoms, physical anxiety (fear), and mental stress (nervous tension). It contains 42 items, 14 per subscale, that relate to the experience of symptoms in the past week. The items are measured on a 4-point scale ranging between 0 (”did not apply to me”) to 3 (”applied to me very much, or most of the time”). Higher scores denote less favourable conditions.

All subscales of the Dutch adaptation are characterized by good internal consistencies (Cronbach's α between .94 and .97, present sample: .86 to .92). Test-retest reliabilities for the Depression, Anxiety and Stress scale are $r = .75$, $r = .89$ and $r = .79$, respectively (de Beurs et al., 2001) . Clinical cut-off scores of $c = 12$ and $c = 5$ for depression and anxiety are recommended by Nieuwenhuijsen, de Boer, Verbeek, Blonk, and van Dijk (2003). The corresponding percentile score of stress (14) observed in a large ($N = 1771$) non-clinical sample in the UK (Crawford & Henry, 2003) was used as the clinical cut-off in this study.

MBI-GS Emotional Exhaustion. The MBI-GS is a self-report questionnaire to assess burnout across professional occupations. The Emotional Exhaustion subscale assesses emotional fatigue, i.e., the feeling of being ”worn-out”. The scale contains five items that are measured on a 7-point scale scored from 0 to 6, where higher scores indicate higher levels of emotional exhaustion. The internal consistency of the subscale ranges between $.84 \leq \alpha \leq .90$ (present sample: $\alpha = .84$). The 8-month test-retest reliability is satisfactory ($.58 \leq r \leq .85$). Well-established clinical cut-off scores are unavailable for the Exhaustion scale. To indicate burnout, Brenninkmeijer and Van Yperen (2003) recommend a cut-off score of 2.67, which was used in the present study. With this

cut-off score, they found a false positive rate of 9.1% and a false negative rate of 13.8%.

2.2.5 Analysis

Intention-to-treat. The analyses were on an intention-to-treat basis and included all participants. Dropouts, who did not complete the posttest measurements, were assumed to have gained nothing. Their pretest scores also served as their posttest scores.

Statistical significance. One-way ANCOVA's (using the pretest scores as a covariate) were conducted to test the difference in means of the two groups at posttest ($\alpha = .05$). The statistical significance of the within-group pre-post gain scores were assessed using paired *t*-tests.

The assumptions of ANCOVA were met. The homogeneity of the regression coefficients in the two groups was confirmed by non-significant interactions between the covariates (pretest scores) and experimental condition. The distributions of the outcome variables were approximately normal, and the variance across the experimental groups was found to be homogeneous.

Effect size. To express the magnitude of the effects, gain scores on the outcome measures were standardized to Cohen's *d*'s (J. Cohen, 1988), representing the number of standard deviations separating the two means. Point estimates and 95% confidence intervals of *d* were determined both for the within- and the between-group effects following a procedure described in detail by Robey (2004). In this procedure, between effect sizes are calculated using the pooled standard deviation and confidence intervals are approximated from the central *t*-distribution.

Clinical relevance. We tested the higher probability of statistically reliable improvement and recovery after treatment compared to the control group with Fisher exact tests, and expressed the difference in this probability between the treatment and the control group in terms of odds ratios (Hillis & Woolson, 2002). We used the Reliable Change Index (RCI) to test the significance of individual improvement (Jacobson & Truax, 1991; $\alpha = .05$; critical RCI = 1.96; "no change" and "detoriation" were pooled

into a single "unimproved" category). Participants were considered recovered if they reliably improved from a pretest score above the cut-off to a posttest score below cut-off. Therefore, participants scoring below cut-off at pretest were excluded from the recovery analysis.

Follow-up. Participants who had completed the treatment were invited by e-mail to complete the follow-up measures on a specially constructed website. These participants include members of the treatment group and members of the control group, who followed the treatment at a later date. Pretest to follow-up data of those who participated in the follow-up were analyzed using repeated measures ANOVA's with time of measurement as the within factor. Differences between the means at the times of measurement were tested for significance using Bonferroni adjustments to keep the family-wise Type I error at α=.05.

2.3 Results

2.3.1 RCT

Participants. Of the 342 respondents who applied for treatment, 65 did not complete the screening and 38 met the exclusion criteria (See Figure 2.1 for details). The final sample (N = 239) comprised 143 females (60%) and 96 males (40%), aged between 22 and 60 (M = 44, SD = 8). Most of the participants were highly educated (84% completed tertiary education), and reported a wide variety of professional occupations. The majority (80%) were in full-time paid employment. Seven participants were unemployed, because of work-related stress. Ten participants suffered from stress in unpaid jobs (housewives, volunteers, and students). In 167 (70%) of the participants, the level of stress was above the clinical cut-off on the stress outcome measure (described in the measures section). On average, the duration of the reported symptoms was 30 months (SD = 34, Range = 1 - 180 months). A number of participants (38%) was on partial or full sick leave. All participants attributed their complaints to their job, although many (65%) reported additional (unspecified) personal issues.

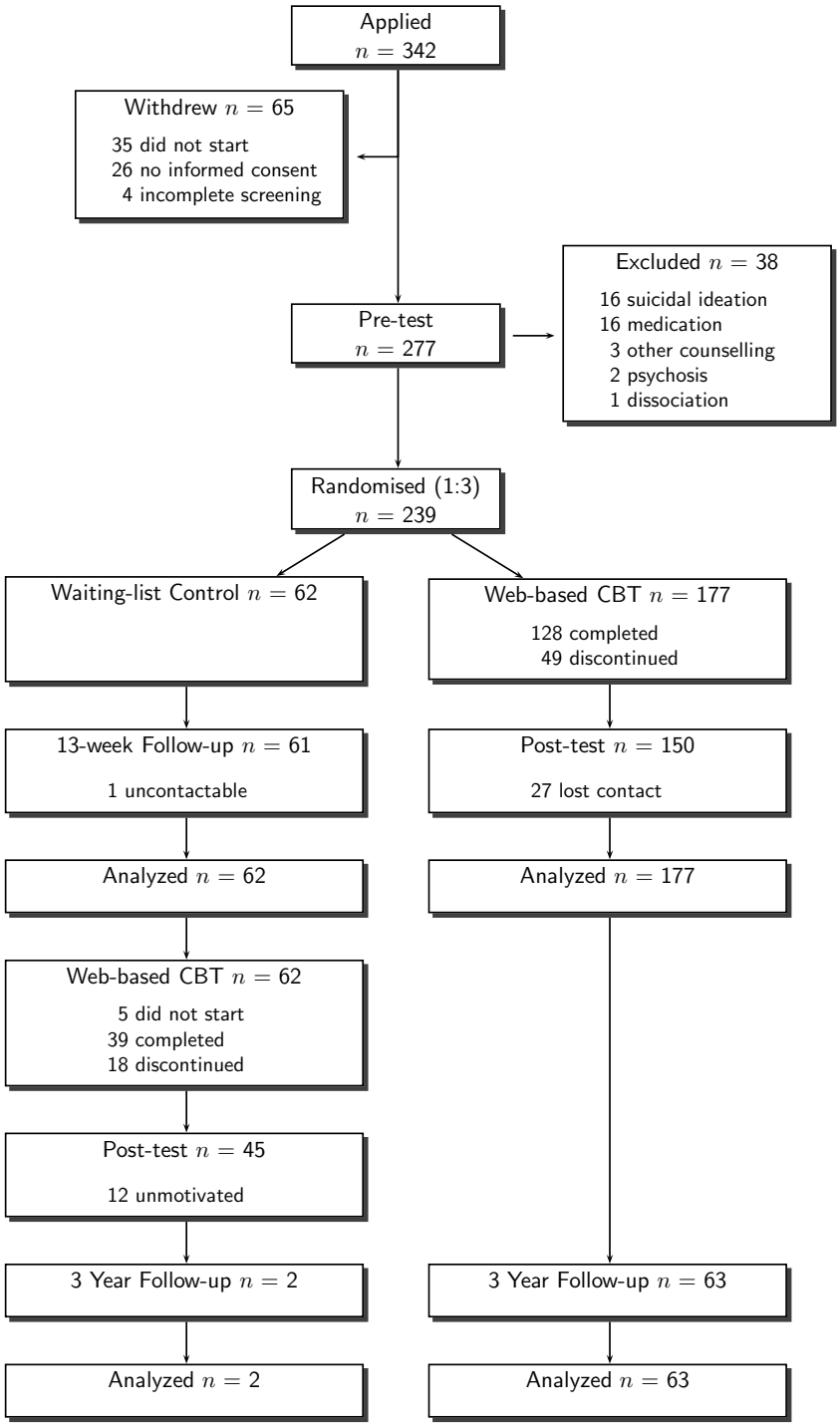


Figure 2.1: Flowchart of study participation.

Randomization. Because we had a fairly large sample of participants, we assigned three times as many participants to the immediate treatment condition, while retaining a large enough sample-size in the control condition to ensure sufficient statistical power to detect treatment effects. Thus, 177 participants (74%) were randomly assigned to immediate treatment, and 62 (26%) to the waiting list control condition.

Table 2.1: Characteristics of participants.

Characteristic[a]	Control $n = 62$		Treatment $n = 177$		Combined $N = 239$	
Demographic						
Gender : female	68%	42	57%	101	60%	143
Age (*M SD*)	42	9	44	8	44	8
Education : tertiary	83%	50	85%	144	84%	194
Marital status : married	27%	17	18%	32	21%	49
living together	42%	26	54%	96	51%	122
unmarried	21%	13	21%	38	21%	51
divorced/widow(er)	10%	6	6%	11	7%	17
Work						
Employment : fulltime	77%	48	81%	140	80%	188
Years in current profession (*M SD*)	9	9	11	9	10	9
Working hours per week (*M SD*)	35	13	36	13	36	13
Symptoms						
Duration of symptoms (months) (*M SD*)	26	29	32	36	30	34
Sickleave : no	56%	35	65%	114	63%	150
partial	23%	14	12%	21	15%	35
full	21%	13	23%	41	23%	54

[a]Values represent subsample percentage and size unless otherwise noted.

Table 2.1 provides an overview of the characteristics of the two groups. To check the randomization, χ^2 tests or *t* tests (where appropriate) were conducted with respect to the outcome measures, gender, age, marital status, education, work contract, work years, working hours, duration of symptoms, and sick leave status. As expected, no significant difference was found [$t(238) = .26$ - 1.75, $P = .08$ - $.80$; $\chi^2 = .002$ - 4.1, $P = .13$ - $.96$].

Dropout. Fifty participants (21%) dropped out from the study, 49 from the treatment group and one from the control group. Of these 50, 25 (50%) dropped out without providing a reason (see Figure 2.1). Twenty-two dropouts (44%) completed the posttest. None of the variables used to check the randomization were predictive of dropout.

Statistical significance. Table 2.2 shows the results of the pretest and posttest measurements on intention-to-treat basis. As hypothesized, the treatment group improved statistically significant with respect to stress [$t(176) = 11.76$, $P < .001$]

Table 2.2: RCT Results (Intention-to-Treat): Treatment ($n = 177$) vs. Waiting List Control ($n = 62$).

Measure[a]	**Pre**		**Post**[b,c]		**Effect size**		**ANCOVA**	
Group	*M*	*SD*	*M*	*SD*	*d*	$CI_{.95}$	$F_{1,236}$	*p*
DASS Stress								
Treatment	19.4	8.0	12.0	8.5	.6	±.3	23.9	$< .001$
Control	19.7	8.0	17.6	9.3				
DASS Depression								
Treatment	11.8	6.7	7.8	6.6	.4	±.3	10.0	.002
Control	12.1	7.3	10.7	7.3				
DASS Anxiety								
Treatment	8.2	6.2	5.2	5.2	.1	±.3	5.1	.025
Control	9.6	6.7	7.3	5.5				
MBI Em. Exh.								
Treatment	3.1	1.2	2.5	1.3	.3	±.2	8.8	.003
Control	3.4	1.2	3.2	1.2				

[a]DASS: Depression Anxiety Stress Scales; MBI: Maslach Burnout Inventory, General Survey. DASS scores are sum scores, MBI scores are averages.

[b]The means and standard deviations reflect the raw data, whereas the F-statistics reflect the results of the test of the difference in posttest means after pretest adjustments (ANCOVA).

[c]Pretest scores were carried forward for participants, who did not complete the posttest measurement ($n = 28$).

[d]ES: Effect size: Cohen's *d* point estimate and 95% confidence interval.

and the secondary outcome measures (all paired *t* tests $P < .0001$). However, the improvement of the waiting list control group also reached statistical significance with respect to stress ([$t(61) = 2.14$, $P < .036$), and all secondary outcome measures, except DASS depression ($t(61) = 1.46$, $P = .15$). Nonetheless, the ANCOVAs showed that the treatment group improved significantly more than the control group, on both stress and the secondary measures (See Table 2.2).

Effect size. The within-group effects in the treated group were large to moderate, ranging between $d = .9$ (stress) and $d = .5$ (anxiety, exhaustion). Because the control group also improved, the effect sizes of treatment as compared to no treatment were smaller: the between effects ranged between $d = .6$ (stress) and $d = .1$ (anxiety) (see Table 2.2).

Clinical relevance. Table 2.3 shows the improvement and recovery rates in the two experimental groups. In the treatment group, with respect to stress, 53% of the participants reliably improved, and half the clinical subgroup recovered from

Table 2.3: Clinical Relevance: Reliable Improvement and Recovery.

	Improved				Recovered					
Measure[a]	*n*	%	*OR*[b]	*p*[c]	c_n[d]	$c_{\%}$[d]	*n*	%	*OR*	*P*
DASS Stress										
Treatment	93	53%	2.9	$< .001$	124	70%	62	50%	5.1	$< .001$
Control	17	27%			43	69%	7	16%		
DASS Depression										
Treatment	64	36%	3.3	.001	78	44%	41	53%	3.5	.007
Control	9	15%			29	47%	7	24%		
DASS Anxiety										
Treatment	55	31%	1.1	.448	110	62%	38	35%	1.9	.097
Control	18	29%			41	66%	9	22%		
MBI Em. Exhaustion										
Treatment	62	35%	1.8	.047	106	60%	36	34%	2.6	.024
Control	14	23%			43	69%	7	16%		

[a] DASS: Depression Anxiety Stress Scales; Maslach Burnout Inventory, General Survey.

[b] Odds Ratio (OR): the ratio of the odds of recovery in the treatment group compared to the odds in the control group.

[c] The *P*-values represent results of Fisher Exact tests of the observed 2x2 tables.

[d] Column c_n and $c_{\%}$ represent the number and percentage of participants that scored in the clinical range at pretest.

clinical stress. The odds of improvement and recovery were significantly higher in the treatment condition than in the waiting list. With respect to the secondary measures, the effects varied. Concerning depression, the effects were similar to the effects on stress. The effects were smaller, but statistically significant, with regard to emotional exhaustion. With respect to anxiety however, the differences were not significant. Concerning anxiety, there was substantial improvement and recovery in the control group that was not compensated by much additional improvement and recovery in the treatment group.

2.3.2 Long-term follow-up

Participants. After about three years (the mean number of months to follow-up was 34, $SD = 2.1$), participants who had completed treatment ($N = 167$) were invited to complete the follow-up questionnaires. Of these 167, 73 could not be traced. Of the remaining 94 participants, 63 (67%) completed the follow-up questionnaires.

A few differences were found between participants who completed the follow-up and those who did not. The former were more highly educated ($\chi^2 = 4.9$; $P =$

Table 2.4: Results of the Three-Year Follow-up (n = 63).

	Pretest		Posttest		Follow-up		Pre-FU ES	
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*	$CI_{.95}$
DASS Stress	19.6	7.6	10.1	7.9	7.3	6.6	1.8	±.3
DASS Depression	11.6	6.5	5.6	5.8	3.3	4.1	1.5	±.3
DASS Anxiety	8.6	5.2	3.8	3.6	3.0	3.7	1.3	±.3
MBI Em. Exhaustion	3.4	1.4	2.5	1.3	1.7	1.0	1.6	±.3

[a]DASS: Depression Anxiety Stress Scales; MBI: Maslach Burnout Inventory, General Survey.

.034). In addition, their treatment length was shorter compared to those who did not participate in the follow-up ($t(117) = 3.3$, $P = .001$). At pretest and posttest, the two groups did not differ on stress, anxiety or depression ($t(164) = .27$ - 1.0, $P = .30$ - $.78$). However, the mean pretest score on emotional exhaustion of participants completing the follow-up ($M = 3.4$) was higher than that of participants who did not ($M = 3.0$; $t(164) = 2.47$, $P = .014$). At posttest this difference was not statistically significant [$t(157) = 1.19$, $P = .23$].

Long-term Effects. Table 2.4 shows the means and variances of the mean scores of the participants at pretest, posttest, and the 3-year follow-up. There was no relapse: the effects became more pronounced. The 3-year follow-up means show significantly less pathology than the posttest means. In terms of effect size, the gains from pretest to follow-up were large [1.3 (anxiety) $\leq$ d $\leq$ 1.8 (stress)].

With respect to stress, the main effect of time was significant, $F(2, 126) = 79.90$, $P < .0001$. The pre-post improvement ($P < .001$) was maintained and more pronounced at the 3-year follow-up (pretest to follow-up: $P < .001$; posttest to follow-up: $P = .020$). The same pattern emerged with respect to depression [$F(2, 126) = 60.73$, $P < .0001$; posttest to follow-up: $P = .002$] and emotional exhaustion [$F(2, 120) = 51.68$, $P < .0001$; posttest to follow-up: $P < .001$]. However, with respect to anxiety, the effects remained constant [$F(2, 126) = 49.52$, $P < .0001$; posttest to follow-up: $P = .34$].

Effects of other treatments. Between the posttest and the 3-year follow-up, 41% ($n = 26$) of the participants received additional treatment (psychotherapy, other forms of

counselling, or medication). Those who followed additional therapy tended to display more complaints at follow-up compared to those who did not undergo additional treatment, although the difference was statistically significant only with respect to emotional exhaustion [$t(60) = 3.1, P < .001$]). At pretest and posttest, the groups did not differ on the outcome measures.

Treatment satisfaction. After completing treatment, participants rated the overall value of the treatment on a scale of 1 to 10 with an average of 7.6 ($SD = 1.0$; Range: 5 - 10). The different treatment phases were consistently rated favourably. Increasing awareness through writing received the highest mean rating ($M = 8.1$; $SD = 1.4$), relaxation was rated lowest ($M = 7.1$; $SD = 1.8$). The participants rated various aspects of the relationship with their therapists on three-point multiple-choice response scales. Overall, the relationship was rated as pleasant (88%) and personal (75%), and was perceived to have grown during treatment (57%). Sixty-eight percent indicated that they had not missed face-to-face contact.

Mediating variables. Comparison of the outcome in completers of the treatment group and participants of the control group who completed the treatment at a later date revealed no significant differences. Therefore, the treatment completers in both groups were pooled (N = 167) to obtain more power for the detection of significant predictors of outcome in multiple regression. However, no significant mediating variables were found. After controlling for pre-treatment stress-levels, post-treatment stress levels were not predicted by gender [$F(1,163) = .52, P = .47$], age [$F(1,153) = .45, P = .45$], education [$F(1,156) = .72, P = .40$], working years [$F(1,157) = .004, P = .95$], working hours [$F(1,157) = .12, P = .73$], duration of symptoms [$F(1,110) = .10, P = .75$] or length of treatment [$F(1,115) = .001, P = .97$]. Similar results were obtained with the secondary measures.

2.4 Discussion

Compared to no treatment, e-mailed standardized CBT moderately reduced stress and induced small to moderate improvements on depression and emotional exhaustion. Fifty percent of the participants had recovered from clinical stress following treatment.

The effects on anxiety appeared to be small. However, after three-years the (uncontrolled) effects were large on all outcome measures. Participants were highly satisfied with their treatment and their therapists.

2.4.1 Evaluation of effect size

The substantial improvement of the participants in the waiting list control group considerably deflated the effects between treatment and control group. The improvement observed in the waiting-list control group is comparable to that observed in other studies (Lange, van de Ven, & Schrieken, 2003; Zetterqvist et al., 2003). Zetterqvist et al. (2003) provided several explanations for these improvements, including the beneficial effects of pretesting and the prospect of treatment. In the current study, the recruiting article and the website provided psycho-education that may have had beneficial effects. Possibly, the participants in the control condition benefited from the information they received, combined with the knowledge that they would start the active treatment in the near future. However, these explanations remain speculative. The use of different types of control groups may shed light on the source of the effects in the control group.

Notwithstanding the improvement in the control group, the observed between group effects are similar to those reported in the literature. The (controlled) effects on stress ($d = .6$) and depression ($d = .4$) compare well with the effects of face-to-face CBT ($d = .7$ and $d = .3$ respectively; van der Klink et al., 2001). In addition, the results corroborate results obtained by Zetterqvist et al. (2003) in an RCT of an Internet-driven stress-reduction program. In that study, the effect sizes were high, but intention-to-treat analyses failed to reveal statistically significant between effects. In contrast, all between effects in the present study were significant on an intention-to-treat basis. This was perhaps due to the larger sample size and the use of a more powerful statistical technique.

2.4.2 Maintenance

The uncontrolled three-year follow-up showed a high impact of the treatment with large effect sizes after three years, on stress and depression, and on emotional exhaustion and anxiety. This is consistent with our findings in the study on posttraumatic

stress (Lange, van de Ven, & Schrieken, 2003), in which effects also proved to be stable. We may conclude that participants, who complete this type of Internet-driven CBT, have a high chance to not relapse, and may even enjoy further improvements. These results are compelling, as the further improvements were not explained by any additional treatment, which some participants had sought in this three year period.

2.4.3 Limitations

Several aspects of this study limit the generalizability of the results. We will discuss the self-selection of the participants, the different time-frame between measurements in the experimental groups, and the dropout and attrition at follow-up.

Self-selection most clearly manifested itself in the high educational level of the participants. The high educational level is probably due to the fact that most participants were recruited through an article in a quality newspaper that is favoured by the well educated. Yet, high education seems to be rather typical of people who participate in Internet-based therapy (Andersson et al., 2006; Carlbring et al., 2001). Until the relation between education and Internet therapy is better understood, caution should be exercised in generalising the present results to less well-educated populations.

Participants in the treatment condition were posttested when they completed treatment. We expected them to complete treatment within the planned 7 weeks, but the average treatment took substantially longer (16 weeks). Consequently, the experimental group, on average, was posttested several weeks after the control condition. Therefore, the between-effects may have been confounded by (uncontrolled) effects of the mere passage of time. However, if present, this effect was probably small. Exploratory analyses did not reveal a significant effect of treatment duration on the outcome variables. Nevertheless, to avoid such interpretative issues in our current studies, we now multiply the protocol time by 1.5 to estimate the actual length of treatment.

During the study, a considerable number of participants dropped out. Although the observed dropout rate of 21% is common in online therapy (Wantland et al., 2004), it is disconcerting that most dropout occurred in the treatment group. No significant predictors of dropout were found, and many dropouts did not provide a reason for terminating their participation. However, therapists noticed that participants often

experienced problems in combining work and treatment demands. Perhaps some of the exercises required too much time for some participants or were otherwise too demanding. Thus, the treatment may have presented a practical burden in terms of time and effort that was not experienced by control group participants.

Finally, the observed long-term reduction of complaints is promising, but was uncontrolled, and suffered considerable attrition. The long-term follow-up needs replication.

2.4.4 E-mail versus website

In this study, the website was used only for basic psycho-education and recruitment. Communication between clients and therapists took place though e-mails. At present, the treatment is delivered completely through a structured, database-driven website. This has several advantages over e-mail exchanges. Privacy is ensured by transparent encryption techniques and login procedures. Further, therapists and clients now use a graphical interface to access treatment elements, which enhances access to the treatment dossier and process overview. Furthermore, the website provides real-time calculation of results and multimedia interfaces that support and facilitate the execution of the exercises. Preliminary (unpublished) findings suggest that the website-driven manual decreases the number of dropouts to a mere 15%, and that it enhances the improvement rates. It would be interesting to directly test the results of an e-mail versus website-driven treatment protocol for chronic stress.

2.4.5 Client-therapist relation and treatment integrity

Knaevelsrud and Maercker (2006) and Spek et al. (2007) observed that positive relationships between therapists and clients are quite feasible in Internet-driven therapy. This observation is supported by the positive evaluation of our participants of their therapists. Yet, this study does not permit conclusions regarding the effect of working alliance or the amount of therapist involvement in Internet-driven treatment. Given that some studies investigating Internet-based self-help with minimal therapist contact report reasonable results (e.g. Carlbring, Furmark, Steckzó, Ekselius, & Andersson, 2006), future research should focus explicitly on the role of the therapist:

How much therapist contact is optimal? What is the effect of working alliance in Internet-based treatment?

Internet-based treatment with standardized therapist contact may have distinct advantages over face-to-face treatment. In Internet-driven treatment it is easier to ensure treatment integrity than in manualized face-to-face treatment. In the present manual much attention is devoted to small details to ensure that the therapists use the manual in the intended way. The manual even specifies the motivating attitude that the therapists are required to adopt and convey to their clients. Furthermore, the weekly supervisions were intense, and exploited the full details of the exchanges between the therapists and their clients. The detailed manual provided the relatively young and inexperienced therapists with guidance and support in treating their clients. As a consequence, their results were on a par with the results of experienced therapists. Apart from the effects on clients, it is interesting to investigate the effects of Internet-driven manuals in the education and training of therapists.

Ruwaard, J., Schrieken, B., Schrijver, M., Broeksteeg, J., Dekker, J., Vermeulen, H., & Lange, A. (2009). Standardized Web-Based Cognitive Behavioural Therapy of Mild to Moderate Depression: A Randomized Controlled Trial with a Long-Term Follow-Up. *Cognitive Behaviour Therapy, 38*(4), 206-221. doi: 10.1080/16506070802408086

Chapter 3

Online Cognitive Behavioural Treatment of Mild to Moderate Depression

Background Depression is common but undertreated. Web-based self-help provides a widely accessible treatment alternative for mild to moderate depression. However, the lack of therapist guidance may limit its efficacy. **Objective** To assesses the efficacy of therapist-guided web-based Cognitive Behavioural Treatment (web-CBT) of mild to moderate depression. **Design** A randomised waiting-list controlled trial (RCT), with an 18-month follow-up. **Participants** The study included a community sample of 54 participants with chronic, moderate depression (immediate treatment: $n = 36$, waiting-list control: $n = 18$). **Measures** Primary outcome measures were the Beck Depression Inventory (BDI-IA) and the Depression scale of the Symptom Check List - Revised (SCL-90-R DEP). Secondary outcome measures were the Depression Anxiety Stress Scales (DASS-42) and the Well-being Questionnaire (W-BQ12). **Results** In the RCT, 5 participants (9%) dropped out. Intention-to-treat analyses of covariance (ANCOVA) revealed that participants in the treatment condition improved significantly more than the participants in the waiting-list control condition ($.011 < P < .015$). With regard to the primary measures, between-group effects were $d = 0.7$ (BDI-IA) and $d = 1.1$ (SCL-90-R DEP). Posttest SCL90-R DEP scores indicated recovery of 49% of the participants in the treatment group, compared with 6% in the control group (Odds Ratio = 14.5; $P < .004$). On average, the effects were stable up to 18 months ($n = 39$), although medication was a strong predictor of relapse. **Conclusion** The results demonstrate the efficacy of web-CBT of mild to moderate depression, and the importance of therapist guidance in psychological interventions.

Introduction

Although depression is common, costly, and a considerable source of personal suffering and diminished quality of life (Ustun, Ayuso-Mateos, Chatterji, Mathers, & Murray, 2004; Wang, Simon, & Kessler, 2003), it often goes untreated. Effective treatments are available; however, only a third of those affected gain access to these treatments (Collins et al., 2004). This suggests that evidence-based treatment for depression remains poorly accessible. It is, therefore, important to develop alternative avenues to disseminate such treatment.

Computerised and Internet-based cognitive-behaviour therapy (CBT) provides a viable and cost-effective treatment alternative for a range of mental health conditions, including depression (Emmelkamp, 2005; K. M. Griffiths & Christensen, 2006; Kaltenthaler et al., 2006; Riper et al., 2007; Wantland et al., 2004). Meta-reviews of Internet-based programs targeting depression revealed small to moderate mean effect sizes, ranging from $d = .3$ (Spek et al., 2007), and $d = .6$ (Barak et al., 2008), to $d = .7$ (Riper et al., 2007). However, these reviews also revealed significant heterogeneity in efficacy among studies. In a more qualitative review, Andersson (2006) concluded that differences between the studies preclude a single, pooled effect size.

Therapist support appears to be a significant moderator of the efficacy of Internet-based therapy. Palmqvist et al. (2007) found a strong correlation ($rho = .8$) between the amount of therapist guidance and treatment efficacy. In the meta-analysis of Spek et al. (2007), the mean effect size of programs with therapist input is large ($d = 1.0$), while the mean effect size of programs without therapist support is small ($d = .2$). However, there is a clear need for more research, because this review included only 12 trials. Of these, five trials included therapist support, and only one targeted depression (Andersson et al., 2005).

On the basis of principles of online therapy applied earlier for the treatment of posttraumatic stress (Lange, Rietdijk, et al., 2003) and work-related stress (Ruwaard, Lange, Bouwman, Broeksteeg, & Schrieken, 2007), we developed a standardised, therapist-guided, web-based treatment of depression. Like most computer-aided psychological interventions of depression, the treatment is highly structured and comprises a balanced sequence of cognitive-behavioural interventions. However, in

contrast to most programmes, this intervention includes scheduled therapist guidance. The treatment protocol specifies the timing, nature, and frequency of therapist feedback. In this respect, the treatment is similar to the programme studied by Andersson et al. (2005). However, there are clear strategic differences. Andersson et al. (2005) aimed to provide guided self-help. In contrast, we aimed to transform manualised face-to-face CBT into an Internet intervention. As such, the treatment includes more frequent therapist feedback. Further, our programme includes explicit techniques that serve to support the therapists in providing the treatment.

We present the treatment and the results of a randomised controlled trial in which we compared the effects of this treatment with those of a waiting-list control condition. In addition, we report the results of an 18-month follow-up study. In light of previous research (Andersson, Bergström, Holländare, Ekselius, & Carlbring, 2004; Spek et al., 2007), we examined several predictors of treatment outcome, including baseline depression intensity, gender, educational level, number of depressive episodes, and medication. We hypothesised that, in comparison to the waiting-list condition, the treatment would reduce depressive symptoms more, and that the improvements would be sustained over time.

3.1 Treatment

The treatment involves strategies from Cognitive Therapy (CT) and Behavioural Activation (Beck, Rush, Shaw, & Emery, 1979; Lewinsohn, Biglan, & Zeiss, 1976). In the cognitive interventions, clients learn to recognise maladaptive, negative thoughts and to substitute these for more realistic, constructive thoughts. In the behavioural interventions, clients are encouraged to engage in behaviours that either elicit positive reinforcement, avoid negative reinforcement from the environment and generally enhance feelings of self-respect (Cuijpers, van Straten, & Warmerdam, 2007; Hopko, Lejuez, Ruggiero, & Eifert, 2003).

Similar to manualised face-to-face CBT, the treatment comprises a balanced set of homework assignments and scheduled therapeutic sessions in which assignments are explained and adapted to the needs of the client. In web-based CBT, the homework assignments are based on web pages, which implement a personal interactive

workbook. At regular intervals, therapists use the contents of this workbook to post feedback and further instructions. The therapist and the client do not communicate in real-time (e.g., online chat, MSN messenger). Similar to e-mail, communication between the therapist and the client is asynchronous. This provides the therapists with the opportunity to reflect on their feedback or to discuss complex cases with colleagues or a supervisor.

Treatment integrity is enhanced by a computerised manual that specifies each step of treatment. This manual governs the order, nature, and contents of the assignments, and the timing of therapist feedback. Furthermore, the manual provides default feedback templates which the therapists personalise to their clients. The templates contain explicit instructions to the therapist (e.g. to start the feedback with a summary of the accomplishments of the client followed by a compliment). In addition, the templates provide feedback snippets for various scenarios. Ideally, clients may complete an assignment as intended and with satisfying results. However, when this is not the case, the templates provide the therapist with suggestions on how to deal with these situations.

A full treatment requires approximately 11 weeks to complete, although the schedule allows clients to adjust the pace to their own situation. The manual defines 8 treatment phases, with 86 client assignments and 21 feedback/instruction texts. Clients need about 2 to 4 hours per week to complete the assignments, and therapists need about 20 to 40 minutes to formulate their feedback. Hence, the treatment takes between 22 to 44 hours of client time, and about 7 to 14 hours of therapist time.

3.1.1 Treatment phases

Phase 1. Inducing awareness: writing. The first phase consists of two writing assignments (Lepore & Smyth, 2002). Clients make an inventory of how and when they experience depressive moods and try to determine the immediate causes of these moods. In the first essay, they are asked to describe in detail an event that elicits their depressive symptoms. In the second essay, they are asked to reflect on the first essay, by relating possible past and present events to their negative feelings and thoughts. After each essay, therapists provide advice, support, and specific commentary.

Phase 2. Inducing awareness: monitoring. Following Phase 1, mood monitoring is applied to enhance awareness. On 5 days, at three fixed time points, participants rate their mood on a 1-10 scale and provide a short description of the situation. These records provide clients and therapists with insight into the mood fluctuations during the day and high-impact events, thoughts, and emotions. In the feedback, provided after the second and fifth day, therapists emphasise the positive experiences reported by the clients (e.g.,"You are not always feeling bad"). In addition, the therapists explain that there are ways to counter the negative experiences (e.g.,"this may change").

Phase 3. Timemanagement. Building on the previous phase, clients are encouraged to structure their activities. For 1 week, they schedule their daily activities and evaluate their success in following this schedule. They are taught how to use their insight in possible daily mood fluctuations by scheduling positive activities at difficult moments. To this end, clients compile a list of pleasant activities. In addition, the therapists suggest several health-promoting activities. Clients receive detailed psycho-education and instruction concerning relaxation and breathing exercises (Öst, 1987), physical activities (Bosscher, 1991), and sleep hygiene strategies (Morin, 1993).

Phase 4: Cognitive restructuring: challenging negative thoughts. Phase 4 targets maladaptive, negative thinking. First, clients simply record negative thinking and its triggering events. Next, they learn to identify automatic, dysfunctional thinking in their notes (e.g., black-and-white thinking, overgeneralization, catastrophic thinking, self-blaming, and neglect of positive aspects). Finally, clients adopt a critical approach towards the maladaptive thoughts, by formulating alternative interpretations of the triggering events.

Phase 5: Cognitive restructuring through behavioural experiments. In Phase 5, clients reformulate their negative thoughts into so-called "If . . . Then . . . " hypotheses (e.g., "If I express my feelings, then they will make fun of me"). Next, the therapists help to create behavioural experiments in which these hypotheses are tested. Two such experiments are designed, executed and evaluated. The therapists ensure that the experiments are practicable and realistic.

Phase 6: Positive self-verbalization. Phase 6 targets clients' self-esteem. For 1 week, they focus on their positive qualities by using the technique of positive self-verbalization (Lange et al., 1998). First, they write an essay on their positive qualities. Next, they summarise on a small card their most powerful statements. The clients are encouraged to read aloud the contents of this card several times a day.

Phase 7. Social skills: Interacting with others. A depressive mood and a negative self-image result in diminished contact with others, which may, in turn, initiate a negative spiral through the lack of positive interaction. In Phase 7, clients break this spiral with two behavioural exercises. In the first, they compile a list of positive characteristics of an influential person (e.g., a wife or husband, a good friend), communicate this list to this person in a scheduled, face-to-face meeting and report the outcome of this meeting to the therapist. In the second experiment, clients compile a list of ways in which they can convey positive messages toward people in their surroundings, including family, friends, and colleagues. Again, these strategies are put into practice, and the therapist is informed of the results.

Phase 8. Relapse prevention: the 'toolkit'. In this last phase, clients reflect on symptoms that might signal relapse in the future. They compose a personal 'relapse prevention toolkit', a personal account of the techniques that helped them most during therapy. They are encouraged to print the toolkit, and keep it for future reference.

3.2 Method

3.2.1 Design

The efficacy study comprised a randomised waiting-list controlled pre-post trial. Participants were randomly assigned to two groups. One group started the 11-week treatment immediately (experimental), while the other started after 11 weeks (waiting-list control). Eighteen months after the start of the trial, a follow-up study was conducted among participants who completed therapy to assess treatment effects in the long term.

3.2.2 Participants

Enrollment. The recruitment procedure, which targeted people from the general population with mild to moderate depression, started with the publication of an article in a national Dutch newspaper announcing the study. The article referred readers to a website, which provided psycho-education regarding depression and its treatment, explained the purpose and design of the study, and contained an application form.

Screening. Respondents were screened by means of web-administered self-report questionnaires. The following exclusion criteria were applied: a score less than 10 (no depression) or more than 29 (severe depression) on the Beck Depression Inventory (BDI-IA: Beck, Ward, Mendelson, Mock, & Erbaugh, 1961; Beck et al., 1979), age under 18 years, heightened risk of dissociation or psychosis, suicidal ideation, drug and alcohol abuse, use of neuroleptic medication, unstable dosages of other psychiatric medication, concurrent psychotherapy, high anxiety levels, or a prevailing posttraumatic stress disorder or panic disorder. Finally, eligible respondents were required to download, print, and return a signed Informed Consent form. Respondents with severe depression (BDI-IA>29) were referred to a concurrent clinical trial of a face-to-face treatment for severe depression (Dekker et al., 2007). Other excluded respondents were referred to their GPs or received information concerning mental health centres in their vicinity.

Risk of dissociation was assessed using the five-item Somatoform Dissociation Questionnaire (SDQ-5; Nijenhuis1997), using a cut-off of 8. Respondents who scored above this cut-off also completed the longer, more specific Dissociation-Questionnaire (DIS-Q; Nijenhuis, Spinhoven, van Dyck, van der Hart, & Vanderlinden, 1998).

Risk of psychosis was attributed to respondents who scored above the cut-off value of 5 on the Hallucination scale of the Screening Device for Psychotic Disorder (SDPD; Lange, Schrieken, et al., 2000). Suicidal ideation was measured using an inventory similar to that of Joiner et al. (2003). Respondents with suicidal plans or a history of recent suicide attempts (within the past 3 years) were excluded. High anxiety levels were detected using the Anxiety and Stress subscales of the Depression Anxiety Stress Scales (described below in the Outcome Measures section). Prevailing posttraumatic stress or panic disorder was inferred in respondents who either scored higher than

36 on the revised Impact of Events Scale (IES-R; Weiss & Marmar, 1996; cut-off 36: Neal et al., 1994), or higher than 8 on the self-rated version of the Panic Disorder Severity Scale (PDSS-SR; Houck, Spiegel, Shear, & Rucci, 2002). Medication usage was assessed through open-ended questions concerning the use, including dosage, of prescription and non-prescription medicine.

Randomisation. One month after the publication of the newspaper article, participants were randomly assigned to the experimental groups by means of a random number generator. Participants were assigned to the treatment group and the control group in a 2:1 ratio. Unbalanced randomisation was considered a defensible ethical compromise between the interests of the participants and that of the study: the treatment under study was built upon evidence-based principles, but not readily available to the public, and the alternative for treatment was the waiting list (Avins, 1998; Edwards & Braunholtz, 2000).

3.2.3 Procedure

Setting. The trial included no face-to-face contact between the participants and any mental health professional. Participants and therapist were given an account to a private password-protected website. They used a common web-browser (e.g., Microsoft Internet Explorer or Firefox) to follow the complete therapeutic procedure, including the completion of the questionnaires and the therapeutic assignments.

Privacy. Several procedures were in place to secure the privacy of the participants. First, only the therapist and the participant were given access to individual treatments. Also, the website included a web-mail system, which allowed participants to contact their therapist outside the treatment regime. Thus, participants who shared an e-mail account with others (e.g. family members) did not have to use this shared account during treatment. Furthermore, the webserver was located at a professional Internet host, protected by a firewall, and remotely administered through an encrypted communication channel. All communication with the website was encrypted with the Hypertext Transfer Protocol over Secure Socket Layer (HTTPS).

Therapists. Eighteen therapists participated: 12 (67%) graduate-level clinical psychologists, and 6 (33%) therapists of the JellinekMentrum Mental Health Care Organisation Amsterdam. The therapists were supervised by a licensed clinical psychologist specialised in web-CBT. E-contact among the therapists and the supervisor was encouraged. In addition, weekly face-to-face supervision group sessions were held.

All therapists were trained in administering web-CBT. The therapists learned how to personalise the feedback, to avoid pitfalls of electronic, text-based communication (Brennan & Ohaeri, 1999), and make use of the asynchronous communication to enhance the quality of the feedback.

Posttest and Follow-up. Immediately after treatment, participants completed the posttest measurements on the website. These included the outcome questionnaires and an evaluation questionnaire. Eighteen months after treatment, those who completed treatment were invited by e-mail to complete the follow-up measures on the website. This group included members of the control group, who followed treatment after the 11-week waiting period.

3.2.4 Measures

The primary outcome measures were the Beck Depression Inventory (BDI-IA, Beck et al., 1979) and the Depression subscale of the Symptom Check List - Revised (SCL-90-R: Derogatis, 1977). Secondary measures were the Depression Anxiety Stress Scales (DASS-42: Lovibond & Lovibond, 1995 and the short-form Well-being Questionnaire (W-BQ12: Bradley, 2000). With the exception of the W-BQ12, higher scores indicate less favourable conditions. For the W-BQ12, higher scores indicate better levels of functioning.

BDI-IA. This self-rate instrument assesses (changes in) the intensity of depression. It comprises 21 items, scored by means of a 4-point Likert scale ranging from 0 to 3, yielding a total score between 0 and 63. It has good psychometrical properties, which have been confirmed in the Dutch population (T. Bouman, Luteijn, Albersnagel, & van der Ploeg, 1985; T. Bouman, 1994). It has an internal consistency coefficient of $\alpha = .86$, a 1-month test-retest reliability of $r = .82$, and it correlates $r = .72$ with

clinical ratings of depression (Beck, Steer, & Garbin, 1988). In this study, we used the commonly applied clinical cut-off of c = 10 (Beck et al., 1961; National Institute for Clinical Excellence, 2004a; Oliver & Simmons, 1984).

SCL-90-R Depression Scale. This scale comprises 16 items, which are scored on a 5-point Likert scale (0-4), indicating the rate of occurrence of depressive symptoms over the past week. The scale has good internal consistency (a = .90), and good convergent and discriminant validity (Arrindell & Ettema, 2003; Schmitz, Kruse, Heckrath, Alberti, & Tress, 1999). For this study, a cut-off score of 25 was used (Aben, Verhey, Lousberg, Lodder, & Honig, 2002; sensitivity: 88%, specificity: 61%). Of relevance to the present study is that computerised versions of the SCL-90-R are comparable to the paper-and-pencil version (Schmitz, Hartkamp, Brinschwitz, Michalek, & Tress, 2000; Vallejo, Jordan, Diaz, Comeche, & Ortega, 2007).

DASS-42. This scale measures negative affect by assessing the severity of symptoms of depression (DASS DEP), anxiety (DASS ANX), and mental stress (DASS STR). It comprises 42 items, 14 per subscale, that relate to the experience of symptoms in the past week. The items are rated on a 4-point Likert scale ranging between 0 (*did not apply to me*) to 3 (*applied to me very much, or most of the time*). All subscales of the Dutch adaptation are characterised by good internal consistencies (Cronbach's a between .94 and .97) and satisfactory 1-month test-retest reliabilities (Depression: r = .75, Anxiety: r = .89; Stress: r = .79; de Beurs et al., 2001; Nieuwenhuijsen et al., 2003).

W-BQ12. The 12-item W-BQ (Dutch version: Pouwer, Snoek, van der Ploeg, Ader, & Heine, 2000) is a measure of psychological well-being. It comprises three subscales measuring negative well-being, positive well-being and energy. Each subscale contains four items scored on a 4-point Likert scale, ranging from 0 (*never*) to 3 (*always*). The total score, used in this study, is derived by summing the subscale-scores, after inverting the Negative Well-being subscale score. This yields scores between 0 and 36. The W-BQ12 was administered at baseline and posttest, but not at follow-up, to reduce the response burden.

3.2.5 Analysis

Intention-to-treat. The analyses were conducted on an intention-to-treat basis. Dropouts, who did not complete the posttest measurements, were assumed to have gained nothing. Their pretest scores served as posttest scores.

Statistical significance. Two-tailed analyses of covariance (ANCOVAs) , using pretest scores as a covariate, were conducted to test the difference in means of the two groups at posttest. The assumptions of ANCOVA were tested and found to be satisfied. The homogeneity of the regression coefficients in the two groups was confirmed by non-significant interactions between the covariates (pretest scores) and experimental condition. Further, the distributions of the outcome variables were approximately normal (in the case of the DASS anxiety scores, normality was achieved by means of a square root transformation), and the variance across the groups was homogeneous. To balance Type-I and Type-II errors, we controlled for multiple testing by controlling the False Discovery Rate (FDR; Benjamini & Yekutieli, 2001). In this procedure, *P*-values are ordered ascendingly and then evaluated sequentially against ascending critical alpha levels, according to the formula a_i = (a/m) * i, where *i* denotes the rank of the ordered *P*-values, and *m* denotes the number of tests). Alternatively, ordered *P*-values, multiplied by *m * i*, can be evaluated against the nominal α significance level, which is the approach taken in this article.

Effect size. To express the magnitude of the effects, mean gain scores on the outcome measures were standardised to Cohen's *d* (J. Cohen, 1988), representing the number of standard deviations separating the two means. Point estimates and 95% confidence intervals of *d* were determined following a procedure described in detail by Robey (2004). In this procedure, effect sizes are calculated using the pooled standard deviation (of the pretest and posttest scores) and confidence intervals are approximated from the central *t*-distribution.

Clinical relevance. We tested the higher probability of statistically reliable individual recovery after treatment compared to the control group with two-sided Fisher's Exact Tests (α = .05), and expressed this difference as odds ratios (OR; Hillis & Woolson,

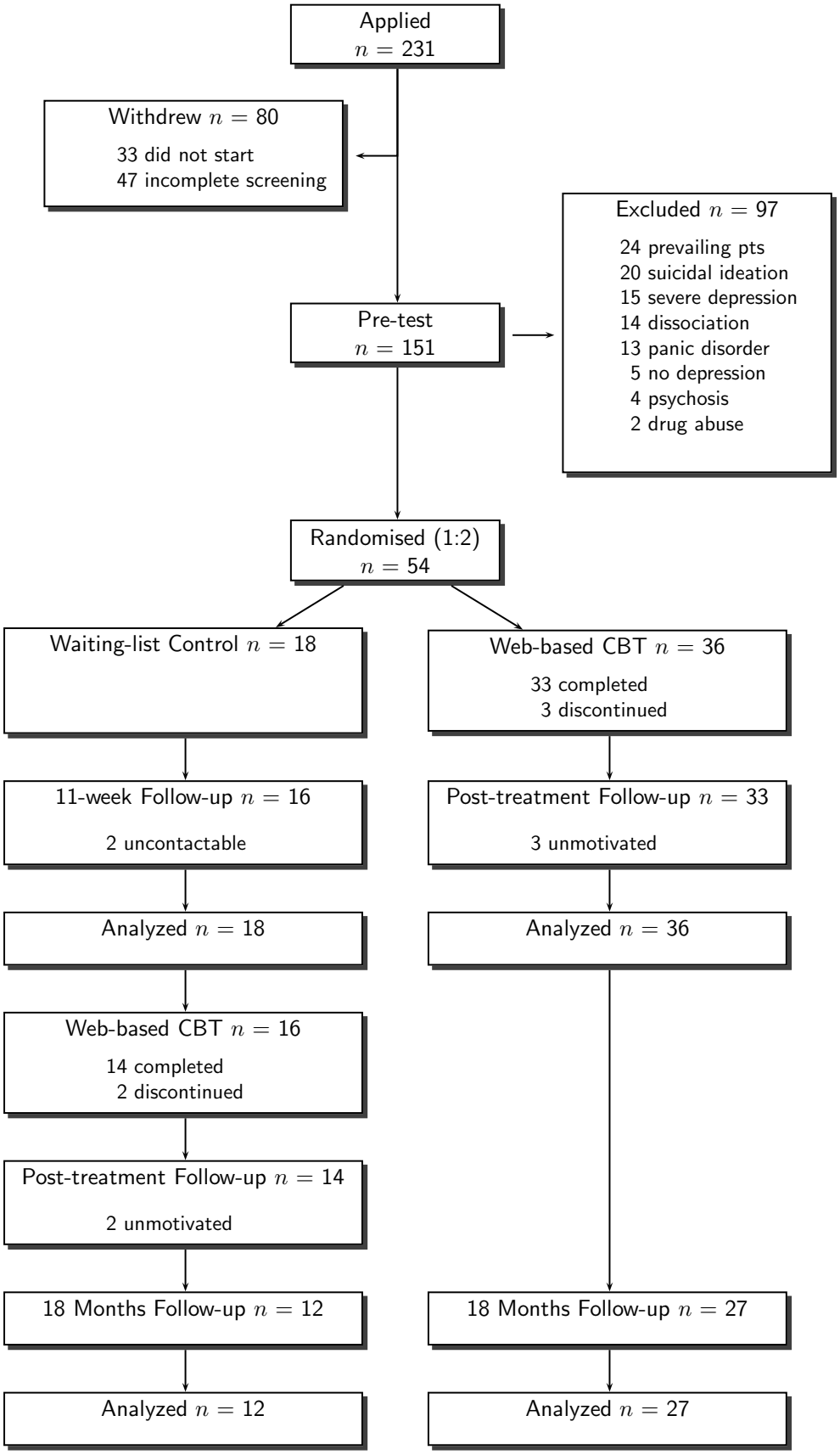


Figure 3.1: Flowchart of study participation.

2002). First, we used the Reliable Change Index (RCI) to test the significance of individual change (Jacobson & Truax, 1991; 'no change' and 'deterioration' were pooled into a single 'unimproved' category). Next, recovery was defined as reliable change from a pretest score above the (published) clinical cut-off to a posttest score below the cut-off. Therefore, participants scoring below cut-off at pretest were excluded from the recovery analysis.

Follow-up. Pre-treatment to follow-up data of those who participated in the follow-up were analyzed using repeated measures analysis of variance (ANOVA) with time of measurement as a 3-level within-subject factor (pretest/posttest/follow-up). Differences between the means at the times of measurement were tested for significance using simple contrasts and *p*-value adjustments to control the false discovery rate within each measure.

Mediating variables. To assess the strength of outcome predictors, we examined the correlation of several predictor variables with the change scores of the BDI-IA (pretest-posttest; pretest-follow-up). Additionally, the predictors were examined through multiple regression analyses.

3.3 Results

3.3.1 RCT

Enrollment and screening. In response to the newspaper article, 231 respondents applied for treatment, of whom 80 (34%) subsequently withdrew. Of the 151 participants who completed the screening, 97 (64%) were excluded, and 54 (36%) were included (cf. Figure 3.1). As shown in Table 3.1, excluded respondents scored less favourably on the BDI-IA [$t(148) = 5.8, P < .001$] and the other outcome measures, in accordance with the exclusion protocol. In addition, excluded respondents were younger [$t(149) = 3.3, P = .001$] and less educated [$\chi^2(1, \text{n} = 143) = 4.9, P = .03$].

Sample. The included sample ($N = 54$) comprised 37 women (69%) and 17 men (31%). On average, they were middle-aged ($M = 42$, $SD = 10$, range: 18 - 71),

unmarried (56%) and highly educated (69% completed tertiary education). They reported a median of 4 depressive episodes (interquartile range = 7.5, range: 0 - 50) over a median of 12 years (interquartile range = 14, range: 2 - 36 years). The median duration of the longest episode was 16 weeks (interquartile range = 45). Sixty percent had received previous treatment for depression, and 19% (n=10) applied for treatment while taking a stable dosage of an antidepressant. The participants were randomly assigned to immediate treatment (n = 36 participants; 67%) or to the waiting-list control condition (n = 18; 33%). With these groups sizes, the power of the study to detect a large (d = .8) between-group effect was approximately 75% for the strongest effect (evaluated at α = .05/6 = .008). To check the randomisation, t-tests and χ^2 tests were conducted with respect to the outcome measures, gender, age, marital status, education, duration of symptoms, and medication status. No significant differences were found (cf. Table 3.1).

Table 3.1: Characteristics of participants and excluded respondents.

	Control		Treatment		Excluded	
Characteristic	$n = 18$		$n = 36$		$N = 97$	
Demographic						
Gender : female	56%	10	75%	27	78%	76
Age (*M SD*)	42	9	42	10	35	12
Education : tertiary	67%	12	64%	23	45%	44
Marital status : married	22%	4	28%	10	31%	30
unmarried	61%	11	53%	19	56%	54
divorced/widow(er)	17%	3	19%	7	13%	13
Symptoms						
Duration of symptoms (*years; Median IQR* [a])	15	16	10	12	11	16
Multiple depressive episodes	72%	13	69%	25	78%	76
Number of depressive episodes (*Median IQR*)	4	7	5	6	3	3
Longest episode (weeks; *Median IQR*)	16	33	22	44	24	39
Treatment						
Received treatment before	67%	12	56%	20	61%	60
Pharmacotherapy	44%	8	30%	11	48%	47
Psychotherapy	56%	10	44%	76	43%	42
Antidepressant at pretest [b]	17%	3	20%	7		

[a] IQR: Interquartile range.

[b] All excluded respondents were excluded before baseline medication usage was recorded. Therefore, the cells in the columns of the excluded group are empty. Further, baseline medication of one participant of the treatment group was unknown at the time of analysis.

Dropout. Five participants (9%) dropped out of the study: three in the treatment group (8%) and two in the control group (11%; see Figure 3.1). Given the small number of drop-outs, predictors of drop-out were not subject to analysis.

Treatment duration. The treatment group took longer than the planned 11-weeks to complete treatment (a median of 16 weeks), and completed the posttest later than the control group (13 weeks). This resulted in interpretative difficulties, because the control group did not fully control for spontaneous improvement. There were no indications that longer treatments were more effective: correcting for pre-treatment scores, treatment duration did not correlate with posttest scores, -.10 $< r_{part} <$ .13. Nevertheless, to account for the discrepancy, the scores of the control group were extrapolated to 16 weeks. Because there were just two measurements, the only option for post-hoc statistical correction was to assume linear change. Given improvements in the control group, this correction deflated the treatment efficacy estimate and, therefore, resulted in a conservative evaluation.

Statistical significance. Table 3.2 shows the results in the intention-to-treat sample. Despite improvements in the control group, the improvement in the treatment group was significantly greater, both with regard to the primary outcome measures (adjusted $P <$ .012; Figure 2A), DASS Anxiety and W-BQ12 ($P <$ .015). With regard to DASS Stress and DASS Depression, the differential improvements between the two groups were not significant, although marginally so ($P <$ 0.07).

Effect size. Table 3.2 also shows the magnitude of the effects. Compared to no treatment, the primary depression measures revealed a large pooled effect size of treatment of d = .9. Large between-group effects were also found with respect to the secondary measures. With regard to DASS, the pooled effect was d = .8 and with regard to W-BQ it was d = 1.0 (c.f. Table 3.2).

Clinical relevance. Table 3.3 shows the recovery rates in the two experimental groups. The odds ratios indicated more reliable individual recovery in the treatment group compared to the waiting-list (OR $>$ 2.1). However, this difference was only significant with regard to SCL-90 DEP (OR = 14.5; adjusted $P <$ .0043).

Table 3.2: Web-CBT vs. waiting-list control (intention-to-treat analysis).

Measure[a]	**Pre**		**Post**[c]		**Effect Size**		**ANCOVA**[c]	
group[b]	*M*	*SD*	*M*	*SD*	*d*	$CI_{.95}$	$F_{1,51}$	p[d]
BDI-IA								
Web CBT	19.7	5.5	9.8	6.5	.7	±.6	7.6	.012
Control	21.3	5.3	15.6	7.6				
SCL-90-R Depr.								
Web CBT	39.9	8.6	26.7	8.1	1.1	±.7	10.9	.011
Control	38.1	7.5	34.4	10.5				
DASS Depression								
Web CBT	20.6	8.4	7.6	6.9	.5	±.6	3.4	.070
Control	20.8	8.1	11.8	9.8				
DASS Anxiety								
Web CBT	6.9	4.7	2.8	2.8	1.0	±.6	8.0	.013
Control	6.1	4.2	6.3	6.6				
DASS Stress								
Treatment	18.1	8.6	9.2	7.1	.8	±.7	4.1	.057
Control	15.5	7.9	13.3	8.8				
W-BQ12								
Web CBT	14.6	2.6	19.9	4.4	1.0	±.6	8.6	.015
Control	15.3	3.0	16.8	4.7				

[a]BDI-IA: Beck Depression Inventory; SCL-90-R DEP: Symptom Check List-Revised Depression scale; DASS: Depression Anxiety Stress Scales; W-BQ12: Well-Being Questionnaire. Higher scores indicate less favourable conditions, except for W-BQ12, where higher scores indicate higher levels of functioning.

[b]Web-CBT: $n = 36$; Waiting-list control: $n = 18$.

[c]Pretest scores of drop-outs ($n = 5$) were carried forward to the posttest.

[d]ANCOVA: Analysis of Covariance

[d]Listed *P*-values are adjusted for multiple testing.

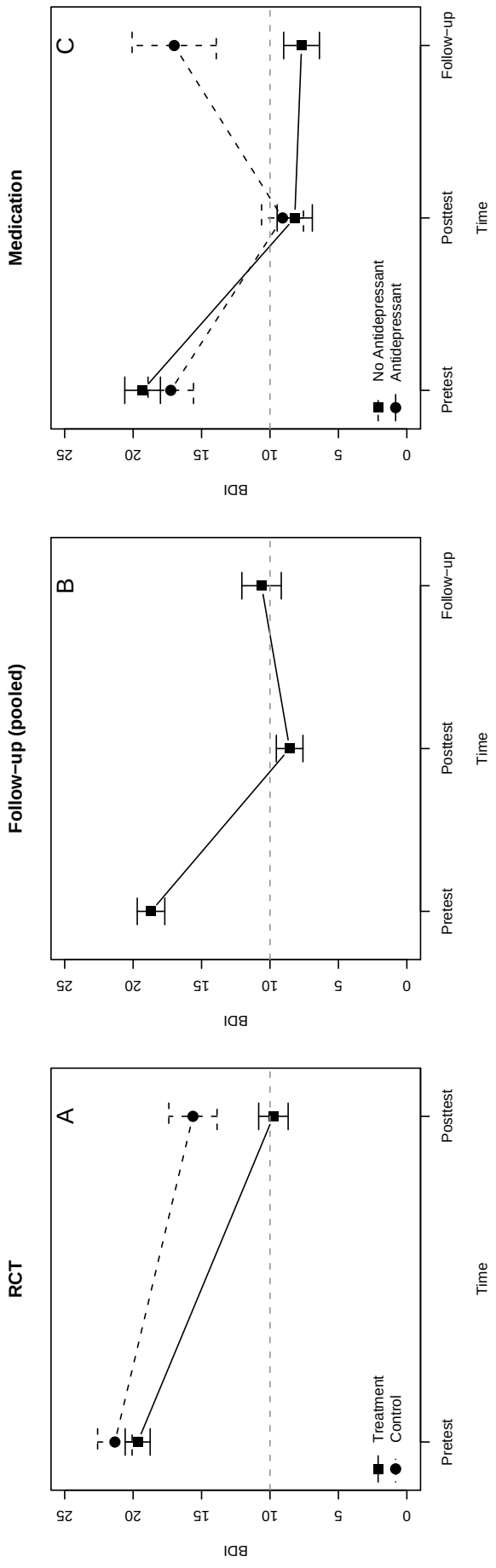


Figure 3.2: Effects of Web-CBT as measured by the Beck Depression Inventory (BDI-IA) in comparison to a waiting-list control group (A), from pretest to the 18-month follow-up (B), and as moderated by medication (C). The error bars represent ± 1 standard error. The dashed horizontal line represents the clinical cut-off (10).

Table 3.3: Clinical relevance: Reliable recovery (intention-to-treat Analysis).

	Pretest		Posttest[b]			
Measure[a]	*n*	%	*n*	%	*OR*[c]	*p*[d]
BDI-IA						
Treatment	36	100%	16	44%	2.1	.375
Control	18	100%	5	28%		
SCL-90-R DEP						
Treatment	35	97%	17	49%	14.5	.004
Control	17	94%	1	6%		

[a]BDI-IA: Beck Depression Inventory; SCL-90-R Dep: Symptom Check List Revised Depression Scale; Cut-off: BDI $<=$ 10; SCL-90-R DEP $<=$ 25.

[b]pretest scores of drop-outs were carried forward to the posttest.

[c]Odds Ratio (OR): the ratio of the odds of recovery in the treatment group compared to the odds in the control group.

[d]Listed *P*-values are adjusted for multiple testing.

Table 3.4: Pooled outcome of web-CBT at pretest, posttest, and after 18 months (n = 39).

	Pre		Post		FU		Pre-FU ES[b]	
Measure[a]	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*	$CI_{.95}$[b]
BDI-IA	18.5	6.6	8.6	6.1	10.6	9.0	1.0	$\pm$.4
SCL-90-R DEP	39.2	10.0	25.5	7.8	29.6	13.4	.8	$\pm$.4
DASS DEP	18.6	9.6	5.8	6.4	9.2	10.3	.9	$\pm$.5
DASS ANX	6.9	5.6	2.2	2.6	3.3	4.8	.7	$\pm$.4
DASS STR	17.1	9.2	8.2	6.8	9.8	8.1	.8	$\pm$.4

[a]BDI-IA: Beck Depression Inventory; SCL-90-R DEP: Symptom Check List - Revised Depression scale; DASS: Depression Anxiety Stress Scales.

[b]ES: Effect size. CI: confidence interval.

3.3.2 Pooled Outcome and Long-term Follow-up

Participants. After the 11-week waiting period, 16 of 18 participants of the control group embarked on the web-based treatment. To enhance the precision of the assessment of participant satisfaction and predictors of outcome, data from both groups were pooled ($N = 52$), using the second (pre-treatment) assessment of the control group as the pretest.

Dropout. In total, 47 participants (90%) completed treatment, and 5 participants (10%) dropped out. Visual inspection of the mean outcome scores suggested that drop-out was unrelated to baseline symptom severity. However, given the small number of drop-outs, predictors of drop-out were not analysed.

Follow-up attrition. Eighteen months after the start of the trial, the participants who had completed the treatment were invited for the long term follow-up. Thirty-nine participants (83%) responded. Baseline depression severity was unrelated to follow-up participation [$t(45) = 0.08$ - 1.5, $P = .93$ - $.13$].

Outcome. Table 3.4 shows the results among the trial completers ($n = 39$). The growth curves of the depression scores (illustrated by Figure 2B) indicated significant linear and quadratic time effects ($F[1,38] = 21.7$-47.6, $P < .001$). On average, participants improved significantly from pretest to posttest ($P < .001$; pooled $d = 1.4$) and maintained their improvements up to the 18-month follow-up with marginal changes from posttreatment (pretest to follow-up: $P < .001$; posttest to follow-up: p $< .13$). The same pattern emerged with respect to DASS Stress (posttest to follow-up: $P < .12$) and DASS Anxiety (posttest to follow-up: $P < .13$).

As indicated by posttest BDI-IA scores of 36 trial completers scoring above cut-off at pretest, 17 (47%) reliably recovered from depression. At follow-up, again 17 (47%) of the participants had recovered. Of the 17 who recovered at posttest, 13 (76%) maintained this improvement up to 18 months.

Participant satisfaction. Immediately after treatment, participants evaluated the overall value of the treatment on a scale ranging from 1 to 10, with an average of 7.7 ($SD = 1.2$). They rated their relationship with their therapist as pleasant (87%) and

Table 3.5: Predictors of reduction of depressive symptoms after web-based CBT. [a]

Measure	***r* post-test** [b]		***r* follow-up**	
Pretest depression severity	.68	***	.42	**
Gender	−.07		.11	
Age	.12		−.02	
Education	−.01		−.04	
Treatment duration	.06		.07	
Previous number of episodes	−.14		−.02	
Antidepressant (at pretest)	−.17		−.35	*
Additional treatment	.00		−.29	
Antidepressant	−.12		−.49	**
Psychotherapy	−.03		−.24	

[a] Outcome is defined as the difference score (Pretest-Posttest, n = 47; Pretest-Follow-up, n = 39) on the Beck Depression Inventory (BDI-IA).

[b] *: $P < .05$; **: $P < .01$; ***: $P < .001$.

personal (78%) and stated that the relationship had grown during treatment (78%). Most participants (89%) indicated that they had not missed face-to-face contact.

Predictors of outcome. As shown by Table 3.5, baseline BDI-IA scores correlated positively with posttreatment improvements. Further, the use of an antidepressant (either during or after treatment) correlated negatively with long-term improvements (see Figure 3.2). None of the other variables significantly predicted the improvements. Examination through multiple regression produced similar results.

Baseline medication status correlated negatively with long-term improvements. Although the effect was absent at posttest, taking an antidepressant during treatment ($n = 6$), was associated with a less favourable outcome at 18 months. At follow-up, the severity of depression in the antidepressant group was equal to baseline severity of depression.

Between the posttest and the 18-month follow-up, 16 (41%) of the 39 participants received additional treatment (psychotherapy only: $n = 7$; medication only: $n = 1$; both: $n = 8$). Although the overall association was not significant, participants receiving additional treatment showed a considerable higher level of symptoms at

follow-up (r_pb = -.29). Further examination showed that participants, who started taking medication after treatment, clearly displayed more severe symptoms at follow-up compared to those who did not [M = 16.8 vs. 8.0, $t(36)$ = 3.0, P = .005]. Again, at pretest and posttest, the two groups reported similar depression levels.

When medication usage both during and after treatment was considered, there was a significant medication effect (r_{pb} = -.52, $P < .001$). Figure 2C illustrates the effect. On average, those who took medication during or after the trial (n = 13) did not improve from pretest to follow-up (d = .04). In contrast, those who did not take any medication (n = 26) reported large improvements (d = 1.6).

3.4 Discussion

Compared with a waiting list control condition, standardised therapist-guided Web-based CBT induced large and clinically relevant improvements in depression, anxiety and well-being in a community sample of adults with chronic symptoms of mild to moderate depression. The follow-up indicated these effects to be persistent, although the use of medication was highly predictive of a negative long-term outcome. Despite the lack of face-to-face contact, participants were highly satisfied with the treatment and their therapist.

3.4.1 Primary results

After controlling for improvements in the control group, web-CBT had an effect size of d = .9. This compares well with the effects found in meta-analyses of face-to-face CT and CBT trials (d = .8: Gloaguen, Cottraux, Cucherat, & Blackburn, 1998; d = .9: Cuijpers et al., 2007) and therapist-guided web-CBT (Spek et al., 2007: d = 1.0). Furthermore, the low drop-out rate of 10% in the present trial was encouraging, given that poor adherence has been identified as one of the major challenges of online intervention (Eysenbach, 2005).

With regard to negative affect, results were less clear. Significant results were found with the anxiety subscale of the DASS. The effects on DASS depression and DASS stress just fell short of statistical significance. However, given the magnitude of observed effects ($.5 < d < 1.0$), we expect future trials to find positive results.

In accordance with face-to-face CBT, 47% of the participants in the web-CBT condition recovered from their depression immediately after treatment. However, when this rate was compared with the recovery rates in the control group, strong significant differences in recovery rates between the experimental groups were found with the SCL-90, but not with the BDI-IA. While the recovery rate on these measures were comparable in the treatment group (49% and 44%), the recovery rates in the control group diverged (SCL-90: 6% vs. BDI-IA: 28%). Both rates of spontaneous recovery appear extreme based on what is known about the natural course of untreated depression (Posternak & Miller, 2001). Future studies should aim to resolve this issue.

In accordance with previous follow-up studies of online therapist-guided CBT (Andersson et al., 2004; Lange, van de Ven, & Schrieken, 2003; Ruwaard et al., 2007; Spek et al., 2007; Wagner & Maercker, 2007) and face-to-face CT (Gloaguen et al., 1998), the effects were found to be stable in the long term. This is encouraging, given that the participants self-reported chronic and recurrent depressive symptoms before treatment. Apparently, this treatment made a lasting difference.

Importantly, the persistent improvements were not explained by additional treatment, which some participants had sought in the 18-month period. On the contrary, pharmacotherapy was highly predictive of relapse on the long term. We would have expected the use of an antidepressant to have resulted in slightly lower, rather than higher, follow-up scores (Pampallona, Bollini, Tibaldi, Kupelnick, & Munizza, 2004). Participants, who took an antidepressant during treatment, may have terminated their medication too soon, because of the short-term improvements as a result of the Internet treatment. Alternatively, these participants may have attributed the immediate improvements not to their own efforts but to the medication. Also, participants who started pharmacotherapy after treatment may have done so simply because they did not experience long-term benefit from Web-CBT. In any case, there is a clear need of updating the treatment manual to address combined therapy. Additionally, future trials should control for medication in a more rigorous manner.

The present treatment and the Internet-based guided self-help program of Andersson et al. (2005) appear equally effective, even though the present treatment requires more therapist time (2 hours vs. 7 to 14 hours). However, in Andersson et al. (2005), the attrition rate was higher (27% vs. 10% in the present treatment), and the effects

on anxiety were lower (d = .5 vs. d = 1.0). Furthermore, the differences in efficacy may be obscured by differences in control group improvements. Andersson and co-workers (2005) found a small effect of d = .2 in the control group, whereas the present study revealed a large effect in the control group (d = .8). Clearly, it would be interesting to compare both approaches directly.

3.4.2 Limitations

We did not make use of a structured diagnostic interview to establish a formal diagnosis of depression or dysthymia. Face-to-face interviewing was not an option, given that we aim to develop fully internet-based treatment. Instead, we used cut-off scores of self-report scales to determine the clinical status of the participants. Although the consequences of these cut-offs are known, observer-rated diagnoses may add to the validity of the results. Telephonic diagnostic interviews may provide a future solution.

A second limitation concerns the generalizability of the results. We excluded respondents with BDI-IA scores indicating severe depression ($>$ 29). Therefore, the results of the present study should not be generalised to individuals with severe depression. The generalizability of results might be further limited by our exclusion criteria. Note, however, that most excluded respondents were rejected because they suffered from posttraumatic stress or panic disorder, for which effective online alternative treatments exist, or because of dissociation and suicidal ideation, for which we presently consider online treatment unsuited given the lack of face-to-face contact.

Another limitation is that the period between pre-test and posttest differed somewhat between the groups studied in this trial. To control for this discrepancy, we used linear extrapolation to estimate the changes in the control group. As a result, the observed improvements in the control group were more pronounced. Thus, the extrapolation resulted in a conservative estimate of treatment efficacy. Nevertheless, replication studies should aim to avoid such corrections, for example by using the observed estimate of treatment duration (16 weeks), or by measuring both groups after a fixed interval, regardless of treatment progress. As a final limitation, we remind the reader that the long-term follow-up study was uncontrolled, and that it included only those participants who completed treatment. The positive outcome of the long-term follow-up needs to be corroborated, preferably by a comparative study.

3.4.3 Future directions

The present results certainly are encouraging and warrant further study of this programme. The treatment shares many characteristics with existing computer-aided CBT, but includes more therapist guidance and appears to be more effective. The results therefore underscore the suggestion of Palmqvist et al. (2007) and Spek et al. (2007) that the efficacy of internet-based CBT is related to the amount of therapist guidance. However, the cost-benefit of various approaches of internet-based CBT awaits further exploration. Future trials should aim for direct comparisons between programmes with varying amounts of therapist guidance, and should evaluate patient preference and treatment adherence as well as efficacy and costs.

Ruwaard, J., Broeksteeg, J., Schrieken, B., Emmelkamp, P., & Lange, A. (2010). Web-based therapist-assisted cognitive behavioral treatment of panic symptoms: A randomized controlled trial with a three-year follow-up. *Journal of Anxiety Disorders, 24*, 387-396. doi: 10.1016/j.janxdis.2010.01.010

Chapter 4

Online Cognitive Behavioural Treatment of Panic Symptoms

Background Internet-delivered treatment may reduce barriers to care in those unwilling or unable to access traditional forms of treatment. **Objective** To assesses the efficacy of web-based therapist-assisted cognitive behavioral treatment (web-CBT) of panic symptoms. **Design** A randomized waiting-list controlled trial with an uncontrolled three-year follow-up. **Participants** A community sample of 58 participants with chronic panic symptoms of varying severity (immediate treatment: $n = 27$, waiting-list control: $n = 31$). **Outcome Measures** The primary outcome measures were a one-week Panic Diary and the Panic Disorder Severity Scale Self-Report (PDSS-SR); secondary measures were the Agoraphobic Cognitions Questionnaire (ACQ), the Body Sensations Questionnaire (BSQ), the Mobility Inventory - Alone subscale (MI-AAL), and the Depression Anxiety Stress Scales (DASS-42). **Results** In the RCT, 54 participants (93%) completed posttest measurements. With regard to the primary outcome measures, intention-to-treat ANCOVA's revealed that participants in the treatment condition improved more than the participants in the waiting-list control condition ($P < .03$), with a pooled between-group effect size of $d = .7$. After three years ($n = 47$; 81% study compliance), effects were more pronounced. **Conclusion** The results demonstrate the efficacy of therapist-assisted Web-CBT in the treatment of panic symptoms.

Introduction

Panic Disorder With and Without Agoraphobia (PD/A) is a debilitating condition characterized by recurrent, unexpected panic attacks accompanied by persistent concern about future attacks and possible avoidance of situations and places in which attacks are expected to occur. With pharmacotherapy or psychotherapy, especially cognitive behaviour therapy (CBT), prospects of recovery are good (Roth & Fonagy, 2005). However, too few of those affected actually receive treatment (Collins et al., 2004; W. H. O. World Mental Health Survey Consortium, 2004). Given the chronic nature of untreated PD/A and its associated individual suffering and costs to society, it is important to find ways to increase access to treatment.

Internet-delivered treatment is in general more accessible than traditional forms of treatment. In addition, randomised controlled trials (RCTs) have demonstrated the feasibility of Web-based CBT in the treatment of mood and anxiety disorders, with moderate to large effect sizes found with programs targeting PD/A (Andersson, Cuijpers, Carlbring, & Lindefors, 2007; Barak et al., 2008; Reger & Gahm, 2009). At present, a variety of web-based treatments targeting PD/A are available, from fully self-administered therapy (Farvolden, Denisoff, Selby, Bagby, & Rudy, 2005), to self-help with minimal guidance (Marks et al., 2004) and guided self-help with limited therapist support via e-mail and/or telephone (Carlbring et al., 2006; Klein et al., 2006; Shandley et al., 2008).

In web-based CBT, therapist involvement is often minimized to increase availability and reduce costs. In a review of Palmqvist et al. (2007), the mean time spent per PD/A patient is 90 minutes (range 0-6 h). This is considerably less than the time spent in the 12-15 (45 min) sessions commonly provided in face-to-face CBT (Roth & Fonagy, 2005). Although most Web-based treatments are promising in terms of established effect sizes, the effect of therapist involvement on the outcome requires further study. For instance, dropout was huge in the Farvolden et al. (2005) study of a program without therapist involvement, with as few as 12 out of 1161 participants completing the program. As reduced contact has been shown to affect outcome negatively in face-to-face CBT (Sharp, Power, & Swanson, 2000), it is possible that this is also the case in web-based treatments. Meta-reviews of Internet programs

suggest that therapist involvement reduces dropout and that greater therapist input improves outcome (Palmqvist et al., 2007; Spek et al., 2007). At present, the most extensive web-based treatment of PD/A includes 6 h of therapist involvement. The effect of increased therapist involvement on outcome of web-based CBT remains to be determined.

On the basis of principles of online therapy, which we applied in the web-based treatment of posttraumatic stress (Lange, Rietdijk, et al., 2003), work-related stress (Ruwaard et al., 2007), and depression (Ruwaard et al., 2009), we developed a therapist-guided web-based treatment of panic symptoms. As in our other studies, the primary aim was to maximize adherence and outcome, rather than to reduce therapist time. Thus, our internet program involved more therapist time (about 5-9 h) than existing programs targeting PD/A.

Earlier, Jager et al. (2004) presented a case study of a preliminary version of this treatment, in which interventions were delivered by e-mail. This treatment was found to induce large reductions in panic symptoms. Here, we present the adapted web-based version of this treatment, and the results of a RCT in which we compared the effects of this treatment with those of a waiting-list. Prevalence rates of minor forms of panic disorder are known to be high and individuals with minor forms of panic disorder suffer a similar impact on their quality of life and functioning as those with full blown panic disorder (Batelaan, de Graaf, Van Balkom, Vollebergh, & Beekman, 2006; Kessler et al., 2006; Magruder & Calderone, 2000; W. H. O. World Mental Health Survey Consortium, 2004). Therefore, to increase the external validity of the study, we aimed to ascertain a community sample with varying intensities of panic symptoms. In addition, we assessed outcome three years after the start of treatment. We hypothesised that, in comparison to the waiting-list, the treatment would reduce panic symptoms more, and that the improvements would be sustained over time.

4.1 Treatment

The treatment involves common CBT strategies for panic disorder, such as psycho-education, awareness training, applied relaxation, cognitive restructuring and (intero-ceptive) exposure techniques. Similar to manualized face-to-face CBT, the treatment

comprises homework assignments and scheduled therapeutic sessions, in which assignments are explained and tailored to the needs of the client. In web-based CBT, the homework assignments are based on a web-based personal interactive workbook. At specific occasions indicated in the manual, therapists post feedback and further instructions on the basis of the contents of this workbook. Therapists take about 20-40 min to read a client's assignment, and to prepare feedback. The manual includes 14 of these feedback moments, so that a full treatment requires between 5 and 9 h of therapist time. Treatment integrity is guaranteed by a computerized manual that stipulates each step of treatment, including the order, the nature, and the contents of the assignments, and the timing of therapist feedback. Furthermore, the manual provides feedback templates, which the therapists adapt to the needs of their clients. These templates include suggested courses of action given various scenarios, such as problems in completing a given assignment.

The approximate duration of treatment is 11 weeks, in which clients work through seven treatment modules. The first module focuses on awareness of panic symptoms. Panic attacks are explained as the result of a cyclic process in which bodily reactions are misinterpreted and amplified (Clark, 1986). Next, clients are asked to do two writing assignments, in which they describe two past panic attacks in the light of this explanation. In the second module, clients learn to keep a Panic Diary, in which they describe, following each panic attack, the situation, their accompanying thoughts and bodily sensations, and in which they rate the intensity of fear on a 10-point scale. Clientsmaintain this diary throughout treatment. Next, in module 3, clients are taught a breathing retraining exercise and a progressive relaxation exercise with tension-release of the muscles. They are then encouraged to expose themselves to their feared situations. Here, the therapists add the paradoxical suggestion (Lange, 2006, chapter 10) that the occurrence of a panic attack would provide the opportunity to put the relaxation exercise into practice to counter the attack. Next, in the fourth module, clients are educated about different forms of exposure (in vitro/in vivo/ interoceptive) and instructed to seek out the situations in which they expect to suffer a panic attack. The clients discuss their planned exposure assignment and their expectations with the therapist. Following the exposure assignment, clients report the outcome and the therapist provide feedback. In module 5, clients learn about the role

of automatic negative thoughts and the principles of cognitive restructuring. They are asked to write two letters of advice to a hypothetical friend coping with similar fears and panic attacks. Their written advice can involve (1) questioning the likelihood of the dysfunctional thoughts, (2) offering new views concerning bodily sensations and panic attacks, and (3) helping in regaining a sense of control. The clients summarize their advice in a set of statements, and are instructed to read these statements out loud regularly, and to recall them when they fear a panic attack. Next, in module 6, transfer of change is implemented by reducing the frequency of therapist feedback. In a two-week period, clients do at least two exposure assignments, without any contact with the therapist. Finally, in module 7, clients reflect on symptoms that might signal relapse. They formulate a 'relapse prevention toolkit', i.e., a personal account of the techniques that proved most helpful during therapy. They are encouraged to print this toolkit on paper and to place it in a visible place at home as a symbolic aid to the future.

4.2 Method

4.2.1 Design

To assess the efficacy of the treatment, we ran a randomised waiting-list controlled pre-post trial, which was approved by the Ethics Committee of the Department of Psychology of the University of Amsterdam. The participants were randomly assigned to two groups. One group started the 11-week treatment immediately (experimental group), while the other started after 11 weeks (waiting-list control group). Three years after the start of the trial, all randomised participants were invited to complete follow-up measurements.

4.2.2 Participants

Enrollment. Radio broadcasts, newspaper articles, advertisements, and magazines interviews announced the study and referred to a website for additional information. This website provided background information on PD/A and on the purpose and design of the study. In addition, the site contained an application form. Power

analyses showed that approximately 70 participants (35 per group) were needed to detect a large, $d = .8$, between-group effect (with ANCOVA, an estimated pre-post correlation of .5 and Holm-Bonferroni corrections for comparisons on 10 outcome measures).

Screening. Respondents were screened through web-administered self-report questionnaires and a semi-structured 15-minute telephonic interview. Participants first completed the self-report questionnaires. Next, the results were used to prepare the clinical interview. In this interview, the presence of at least subsyndromal PD/A was established, according to the guidelines as listed in Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association, 2000), i.e., respondents were included only if they experienced at least one full or limited symptom panic attack in the previous month. Diagnoses were made by trained interviewers and checked by supervising researchers.

The following inclusion criteria applied: age 18 years or older; no heightened risk of dissociation, psychosis, or suicide; no history of serious mental illness; no other prevailing mental illnesses; absence of a medical condition that might explain panic symptoms; no drug abuse; no use of neuroleptic medication; no use of anxiolytics or antidepressants for less than 3 months or use of unstable doses (i.e., clients took medications as prescribed and in the targeted dosage); and no concurrent other treatment. Excluded respondents were referred to their general practitioner, to other mental health institutions or to other Web-CBT programs. Eligible respondents downloaded, signed and returned an Informed Consent form.

Risk of dissociation was assessed using the Somatoform Dissociation Questionnaire (SDQ-5; citeNPNijenhuis1997). Respondents who scored above the cut-off (8), completed the more specific Dissociation Questionnaire (DIS-Q) (Vanderlinden, van Dyck, Vandereycken, & Vertommen, 1991; cut-off: 3.0). The Screening Device for Psychotic Disorder (SDPD; Lange, Schrieken, et al., 2000) was used to assess the risk of psychotic episodes, with a cut-off value of 5 on the Hallucination scale. Hypochondriasis was assessed using the Whitely Index (Speckens, Spinhoven, Sloekers, Bolk, & van Hemert, 1996; cut-off: 15). The occurrence of a prevailing posttraumatic stress disorder was assessed using the Dutch version of the Impact of Events Scale - Revised (IES-R; Weiss & Marmar, 1996; cut-off 36: Neal et al., 1994). Obsessive compulsive

disorder, social phobia and specific phobia, and bipolar disorder were signalled using seven items from the Diagnostic Interview Schedule (DIS: Helzer & Robins, 1988).

Therapists. The therapists were 7 graduate students in clinical psychology, 2 post-graduate students, and 2 psychologists. They were supervised by two senior specialists in web-CBT. All therapists had followed advanced courses in CBT, and received additional training in administering web-CBT. They were taught how to use the feedback templates of the manual, to increase motivation by adopting a stimulating emphatic attitude, to avoid the pitfalls of electronic, text-based communication (e.g., Brennan & Ohaeri, 1999) and to profit from the asynchronous nature of the communication to enhance the quality of the feedback (i.e. by discussing cases with one another or with the supervisor). Participants were assigned to therapists by the supervisor based on the availability of the therapists.

4.2.3 Outcome measures

Primary outcome measures were the self-rate version of the Panic Disorder Severity Scale (PDSS-SR; Houck et al., 2002; Shear et al., 2001) and a one-week *Panic Diary*. Secondary measures were the *Avoidance when ALone* subscale of the *Mobility Inventory* (MI-AAL; Chambless, Caputo, Bright, & Gallager, 1984), the *Agoraphobic Cognitions Questionnaire* and the *Body Sensations Questionnaire* (ACQ/BSQ; Chambless et al., 1984) and the *Depression Anxiety Stress Scales* (DASS-42; Lovibond & Lovibond, 1995). For all measures, higher scores indicate greater symptom severity.

Panic diary. In the Panic Diary, participants monitored panic attacks occurring in a one-week period. Participants were instructed to report each distinct period that was characterized by a sudden onset of intense apprehension, fearfulness, or terror, possibly associated with feelings of impending doom (American Psychiatric Association, 2000). They were asked to note the occurrence of each attack and to rate the attack severity (on a 1-10 scale) and the specific symptoms experienced during the attack. To this end, participants checked each of 13 key symptoms of panic disorder occurring during the attack, e.g., “palpitations, pounding heart or accelerated heart rate”, “trembling or sweating”. Three outcome measures were derived from the

diaries: 1) one-week attack frequency, 2) average number of experienced symptoms and 3) average attack intensity. For the three year follow-up, the burden of keeping a one-week diary was expected to have a negative impact on the response rate. Therefore, in the follow-up, we asked participants to retrospectively estimate the frequency, the average number of symptoms, and the average intensity of panic attacks in the week prior to follow-up assessment.

PDSS-SR. This 7-item survey is the self-report version of the more commonly used clinician-rated Panic Disorder Severity Scale (Shear et al., 2001). It assesses the severity, in the past week, of the following seven key dimensions of panic disorder, i.e., frequency of panic attacks, distress during panic attacks, anticipatory anxiety, agoraphobic fear and avoidance, body sensation fear and avoidance, and work and social impairment. The responses to the items are expressed on a 5-point Likert scale ranging from 0 (*none*) to 4 (*extreme*), and summed to create an overall severity score (range 0-28). Psychometric studies revealed excellent internal consistency (Cronbach's α: > 9; Houck et al., 2002; Newman, Holmes, Zuellig, Kachin, & Behar, 2006), good two-week test-retest reliability (r = .84; Newman et al., 2006), and sensitivity to change after treatment (Houck et al., 2002). With a cut-off of 8, the clinician-rated PDSS is moderately effective in identifying patients with Panic Disorder (sensitivity: 83%; specificity: 64%; Shear et al, 2001). van der Meer and Burgerhout (2004) confirmed the usefulness of this cut-off in a psychometric study of the Dutch self-report version of the PDSS. In their study, a PDSS-SR cut-off of 8 discriminated between patients (n = 129) and non-patients (n = 131) with specificity of 100% and a sensitivity of 82%.

MI-AAL. This is a 27-item self-report questionnaire measuring agoraphobic avoidance in a variety of places and situations while alone. Participants rate the frequency of avoidance of these situations on a 5-point Likert scale (from 1, *never avoid*, to 5, *always avoid*). Item scores are averaged to obtain an overall measure of avoidance severity. In Dutch psychometric studies (de Beurs, 1993), the MI-AAL was shown to be reliable (α = .94), valid and sensitive to change with treatment. Respondents scoring 2.18 or up are more likely to belong to the clinical population (de Beurs, 1993).

ACQ/BSQ. Fear of fear was measured using the Body Sensations Questionnaire and the Agoraphobic Cognitions Questionnaire twin-scales (ACQ/BSQ: Chambless et al., 1984; Dutch version: T. K. Bouman, 1995, 1998; de Beurs, 1993). The BSQ is a measure of the intensity of anxiety/fear provoked by 17 physical sensations. The ACQ measures the frequency of 14 beliefs about the negative consequences of anxiety. For both measures, items are rated on a 5-point Likert score ranging from 0 to 4, and item scores are averaged to obtain the overall score. Both measures have good internal consistency ($\alpha >$ .89), and good test-retest reliability ($>$ .78). To reduce the response burden, the ACQ and BSQ were administered at baseline and posttest, but not at follow-up.

DASS. The DASS measures negative affect by assessing the severity of symptoms of depression (DASS DEP), anxiety (DASS ANX), and mental stress (DASS STR). It comprises 42 items, 14 per subscale, that relate to the experience of symptoms in the past week. The items are rated on a 4-point Likert scale ranging between 0 (*did not apply to me*) to 3 (*applied to me very much, or most of the time*). All subscales of the Dutch adaptation are characterized by good internal consistencies (Cronbach's α between .94 and .97), and satisfactory 1-month test-retest reliabilities (Depression: r = .75, Anxiety: r = .89; Stress: r = .79; de Beurs et al., 2001; Nieuwenhuijsen et al., 2003).

Impairment. During treatment, after each treatment phase, we asked participants to rate their past-week impairment caused by panic symptoms, to gain insight in the development of panic symptoms during the course of treatment. Participants rated the impact of panic symptoms on their daily functioning on a single 10-point item ranging from 1 (*no impairment*) to 10 (*severe impairment*).

4.2.4 Analyses

Intention-to-treat. The RCT analysis was conducted on an intention-to-treat basis and included all participants. Participants failing to complete posttest measurements were assumed to have gained nothing. Their pretest scores served as their posttest scores. No attempt was made to correct for missing data of the three-year follow-up,

because statistical imputation was considered inappropriate with such a long time-interval. However, in the analyses of the long-term outcome data we used mixed modeling (see below for details), which is an accepted method to account for missing data.

Statistical significance. Two-tailed ANCOVAs (using pretest scores as a covariate) were conducted to test the difference in means of the two groups at posttest, using Holm-Bonferroni adjustments (Holland & DiPonzio Copenhaver, 1988) to maintain overall Type-1 error α at .05. These analyses were run using the generalized linear model function (glm) of the statistical software program 'R' (R Development Core Team, 2008).

The assumptions of ANCOVA were examined and found to be satisfied. The distribution of most outcome variables was approximately normal, and the variance across the groups was homogeneous. With regard to DASS Depression, normality was achieved by means of a square root transformation. Further, the distribution of attack frequency was positively skewed, as was to be expected since these are count data. For this variable, a generalized linear model with a Quasi-Poisson distribution (see Maindonald & Braun, 2007) as the link function provided a more realistic ANCOVA regression model. The homogeneity of the regression coefficients in the two groups was confirmed by non-significant interactions between the covariates (pretest scores) and experimental condition. However, a significant group by covariate interaction was found with respect to the BSQ (the effects were more pronounced for higher baseline BSQ scores). As it would be improper to use the significance of the treatment factor as an indicator of effect in this case (Enqvist, 2005), we used the significance of the group by covariate interaction term as the outcome of interest.

Effect size. To express the magnitude of the effects, mean gain scores on the outcome measures were standardized to Cohen's *d* (J. Cohen, 1988), representing the number of standard deviations separating the two means. Point estimates and 95% confidence intervals of *d* were determined both for the within- and the between-group effects following a procedure described in detail by Robey (2004). Between effect sizes were calculated using the pooled standard deviation (of the pretest scores) and confidence intervals were approximated from the central *t*-distribution.

Clinical relevance. We tested the differential probability of a clinically relevant outcome on the primary outcome measures after treatment compared to the control group with two-sided Fisher's exact tests ($\alpha = .05$), and expressed this difference as odds ratios (OR; Hillis & Woolson, 2002). With regard to attack frequency, two clinically relevant outcomes were defined as follows: (a) a reduction in panic attack frequency of at least 50%, and (b) no full-blown panic attack at posttest, i.e. no panic attack with four or more of the key symptoms listed in DSM-IV (American Psychiatric Association, 2000). With regard to the PDSS-SR, a posttest score below the cut-off of 8 was considered clinically relevant. Participants scoring below cut-off at pretest had to remain below cut-off, and participants scoring above cut-off at pretest had to reliably improve to a score below cut-off. To account for measurement error, we used the Reliable Change Index (RCI; Jacobson & Truax, 1991) to test the significance of individual improvement. Participants had to improve at least 5 scale points before change was considered reliable.

Pooled Outcome and long-term follow-up. After the waiting period, the control group followed the web-based treatment too. To increase power of the long-term follow-up analyses, the data of the treatment and the control group were pooled, using the second (post-waiting period) assessment of the control group as the pretest. Pre-treatment to follow-up outcome data were analyzed through multilevel regression modeling (see Pinheiro & Bates, 2000) with time of measurement (at level 1) nested within participants (at level 2), and a random intercept and time-coefficient to account for individual differences between participants. Differences between the means at the times of measurement were tested for significance using simple contrasts. For these analyses, Holm-Bonferroni adjustments (Holland & DiPonzio Copenhaver, 1988) were used to maintain a familywise type I error of $\alpha = .05$ within each measure.

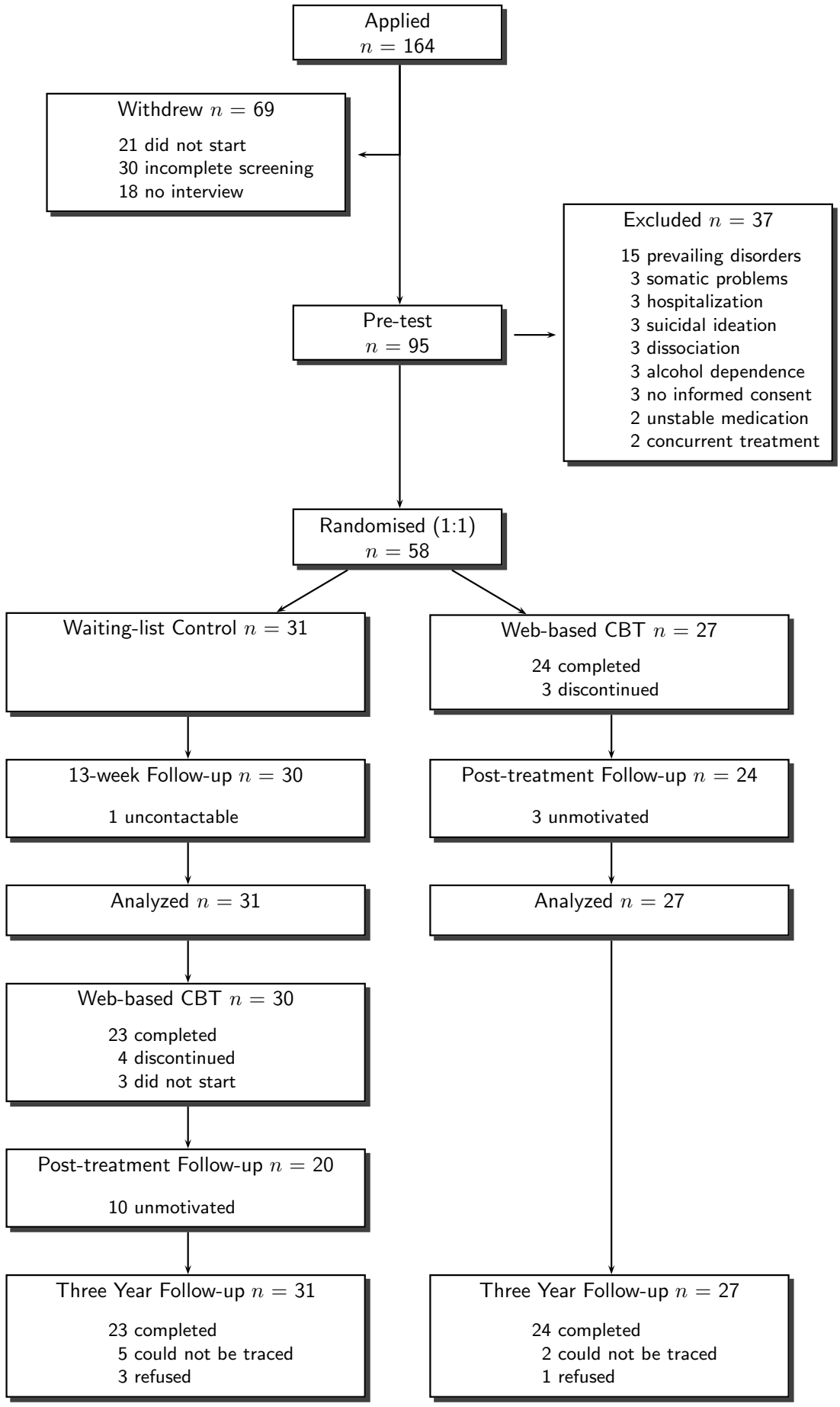


Figure 4.1: Participant flow.

4.3 Results

4.3.1 RCT

Enrollment and randomisation. Recruitment resulted in 164 respondents who applied for treatment. Of these, 69 (42%) did not complete the screening. Of the remaining 95 respondents, 37 (39%) met the exclusion criteria (cf. Figure 4.1), and 58 were randomly assigned either to immediate treatment ($n = 27$; 47%) or to the waiting-list ($n = 31$; 53%). With these groups sizes, the power of the study to detect a large ($d = .8$) between-group effect was approximately 71% for the strongest effect. To check the randomization, *t*-tests and χ^2 tests were conducted with respect to the outcome measures, gender, age, marital status, education, duration of symptoms, and medication status. As implied by Table 4.1, no significant differences were found. We concluded that the randomisation procedure had been successful.

Table 4.1: Characteristics of Participants.

	Control		**Treatment**	
Characteristic[a]	$n = 31$		$n = 27$	
Demographic				
Female	65%	20	75%	27
Age (*M, SD*)	39	10	38	11
Education : tertiary	48%	15	44%	12
With partner	55%	17	52%	14
Symptoms				
Years with symptoms (*M, SD*)	8	7	11	9
PDSS-SR > 8[b]	52%	16	56%	15
MI-AAL > 2.18[c]	58%	18	56%	15
Panic Diary: at least 1 full-blown attack[d]	65%	20	67%	18
Treatment				
Previous treatment for panic	33%	10	41%	11
Previous treatment for other disorder	65%	20	59%	16
Medication	29%	9	30%	8

[a] Values represent subsample percentage and size unless otherwise noted.

[b] PDSS-SR: Panic Disorder Severity Scale - Self-Report.

[c] MI-AAL: Mobile Inventory - Avoidance Alone subscale.

[d] A full-blown attack is an attack with four or more core symptoms of panic disorder as listed in the DSM-IV (American Psychiatric Association, 2000)

Baseline characteristics. As shown by Table 4.1, the participants, on average, were female (72%), middle-aged ($M = 38$, $SD = 10$, range: 20-69) and highly educated (47% completed tertiary education). Average duration of panic symptoms was 9 years ($SD = 8$, Range = 1 - 144 months). At pretest, symptoms levels were typical of clinical groups in the majority (52%-68%) of participants. About a third (36%, $n = 21$) of the participants had engaged in previous treatment for panic disorder, and 29% ($n = 17$) were on a stable dose of either an anxiolytic or an antidepressant.

Compliance. Trial attrition was low: 93% of the participants ($n = 54$) completed the posttest measures. Four participants dropped out: three in the treatment group and one in the control group (c.f. Figure 4.1). Given the small number of dropouts, predictors of dropout were not subject to analysis.

Statistical significance and effect size. Table 4.2 shows the results of the pretest and posttest measurements on intention-to-treat basis. With regard to the primary outcome measures, the treatment group improved significantly more than the control group ($P < .027$), with a pooled standardized mean difference in improvement between the two groups of $d = .7$. In addition, treated participants experienced significant larger reductions in fear-provoking bodily sensations (BSQ, $P < .025$) and general psychopathology (DASS) in comparison to untreated participants. However, between-group effects were moderate to small ($d < .3$), and non-significant ($P > .056$) with regard to avoidance (MI-AAL) and agoraphobic beliefs (ACQ).

Table 4.2: RCT results (intention-to-treat[a]): web-CBT (n = 27) vs. waiting-list control (n = 31).

Measure[b]	**Pretest**		**Posttest**		**Between ES[c]**		**ANCOVA**	
Group	*M*	*SD*	*M*	*SD*	*d*	$CI_{.95}$	$F_{1,55}$	p^{d}
One-week attack frequency								
Treatment	4.7	3.5	2.7	3.1	.6	±.4	13.3	.005
Control	3.9	2.6	3.6	3.4				
Symptom count (range: 1-13)								
Treatment	3.2	1.5	1.9	1.7	.8	±.5	8.4	.022
Control	3.0	1.8	3.0	1.9				
Attack intensity (range: 1-10)								
Treatment	4.5	2.2	2.3	2.1	1.1	±.5	16.7	.001
Control	4.1	2.0	4.2	2.3				
PDSS-SR (range: 0-28)								
Treatment	9.0	5.5	5.9	4.4	.4	±.4	7.3	.027
Control	9.4	5.4	8.4	4.6				
MI-AAL (range: 1-5)								
Treatment	2.6	1.0	2.2	1.0	.3	±.3	5.1	.056
Control	2.5	1.0	2.4	.9				
BSQ (range: 1-5)								
Treatment	2.3	.6	1.9	.5	.4	±.4	8.6	.010
Control	2.5	.7	2.4	.8				
ACQ(range: 1-5)								
Treatment	1.1	.6	.8	.5	.2	±.3	1.6	.210
Control	1.0	.6	.9	.6				
DASS Dep (range: 1-42)								
Treatment	9.5	8.0	3.9	4.1	.7	±.5	10.3	.015
Control	8.6	8.3	8.3	8.3				
DASS Anx (range: 1-42)								
Treatment	11.8	6.7	6.7	4.9	.5	±.4	8.8	.023
Control	14.2	8.2	12.9	9.2				
DASS Str (range: 1-42)								
Treatment	14.2	7.8	8.5	4.7	.8	±.5	9.3	.022
Control	12.6	8.1	12.9	9.3				

[a] All randomized participants were included in the analyses. Pretest scores of dropouts were carried forward to the posttest.

[b] PDSS-SR: Panic Disorder Severity Scale - Self-report; MI Alone: Mobile Inventory - Alone subscale; ACQ: Agoraphobic Cognitions Questionnaire; BSQ: Body Sensations Questionnaire; DASS: Depression Anxiety Stress Scales (Dep: Depression; Anx: Anxiety; Str: Stress). Higher scores indicate less favourable conditions.

[c] ES: effect size: Cohen's d point estimate and 95% confidence interval (d- $CI_{.95}$ to d + $CI_{.95}$).

[d] P-values were Holm-Bonferroni corrected for multiple testing.

Table 4.3: Web-CBT vs. waiting-list control: clinical relevance analysis (intention-to-treat)[a].

Measure[b]	Criterion	Criterion met %[b]		OR[c]	P[d]
		Treatment	Control		
Attack Frequency	-50%	52%	14%	6.3	.02
Full-Blown Attacks	0	70%	47%	2.7	.11
PDSS-SR	<8	70%	42%	3.2	.07

[a]Pretest scores of dropouts were carried forward to the posttest.

[b]PDSS-SR: Panic Disorder Severity Scale. Cut-off: < 8.; A full-blown attack is an attack with four or more core symptoms of panic disorder as listed in the DSM-IV (American Psychiatric Association, 2000)

[c]Odds ratio (OR): the ratio of the odds of a clinically relevant outcome in the treatment group and the odds in the control group.

[d]Listed *P*-values were Holm-Bonferroni adjusted for multiple testing.

Clinical relevance. Compared to the participants in the waiting list, treated participants were six times more likely to experience a 50% reduction in panic attacks ($P < .02$). Significant differences were not found with regard to the PDSS-SR and the number of people who experienced no full-blown panic attacks, despite a clear trend towards a more favourable outcome after treatment, as evidenced by the odds ratios in Table 4.3 (2.7 and 3.2).

4.3.2 Pooled outcome and long-term follow-up

Participants After the waiting period, 27 of the 31 participants in the control group embarked on the web-based treatment, of whom 23 (85%) completed treatment, and 20 (75%) started the posttest. Three years after the start of the trial, 47 of the 58 participants (81%) responded to the invitation to participate in the follow-up study. We ran a series of *t*-tests comparing those who participated in the follow-up and those who did not. These analyses failed to reveal a relation between follow-up participation and either pretest symptom severity or immediate (posttest) treatment response.

Outcome. Table 4.4 shows the results of the assessments in the pooled group at pretest, posttest, and after three years. Immediately after treatment, participants reported significant, moderate to large improvements ($.4 < d < .9$). Three years

Table 4.4: Web-CBT: pooled outcome at posttest ($N = 58$) and after three years ($n = 47$).[a]

	Pre		Post		FU		Pre-FU ES[c]	
Measure[b]	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*	$CI_{.95}$
Attack Frequency	4.2	3.4	2.4	2.6	.6	1.3	1.4	±.4
Symptom Count	3.1	1.7	2.0	1.7	1.5	2.7	.7	±.4
Attack intensity	4.3	2.2	2.6	2.3	1.4	2.6	1.2	±.4
PDSS-SR	8.6	5.0	5.5	4.0	4.0	4.7	.9	±.3
MI-AAL	2.5	1.0	2.1	.9	1.8	.8	.8	±.2
DASS Dep	8.9	8.2	4.3	5.4	4.1	6.2	.7	±.3
DASS Anx	12.3	8.2	6.3	5.6	4.8	5.8	1.1	±.4
DASS Str	13.5	8.6	8.4	6.6	6.2	5.4	1.0	±.4

[a]To account for attrition at follow-up (19%), pre-treatment to follow-up data were analyzed through multilevel regression modeling. Missing posttest data was imputed using pretest scores.

[b]PDSS-SR: Panic Disorder Severity Scale Self-Report; MI-AAL: Mobile Inventory - Alone subscale; DASS: Depression Anxiety Stress Scales (Dep: Depression; Anx: Anxiety; Str: Stress). Higher scores indicate less favourable conditions.

[c]ES: Effect size: Cohen's *d* point estimate and 95% confidence interval (d- $\text{CI}_{.95}$ to d + $\text{CI}_{.95}$).

after the start of the trial, participants reported significant further improvements with regard to the frequency and intensity of attacks ($z = 6.2$, $P < .001$; $z = 2.5$, $p = .01$) and avoidance (MI-AAL; $z = 3.1$, $P = .002$). Overall, compared to the immediate (posttest) effects, pretest to follow-up effect sizes were substantially higher ($.7 < d < 1.4$). A majority of the follow-up participants ($n = 32$; 68%) reported no panic attacks in the week preceding follow-up measurement. With regard to the PDSS-SR, 83% of the follow-up participants ($n = 39$) scored below the clinical cut-off (8).

Additional treatment. In the period between the posttest and follow-up, 12 participants (26% of those completing the follow-up) received further treatment (medication and/or psychotherapy) for their panic symptoms. With respect to the primary outcome measures, those who received additional treatment did not score differently from those without additional treatment.

Participant satisfaction. At posttest, participants rated the overall value of the treatment with a mean score of 8.6 on a 1 to 10 point scale ($SD = 1.3$; range: 4-

10), and their satisfaction with their therapists with a mean score of 9.0 ($SD = 1.2$). Eighty-one percent reported a large impact on their daily functioning, 81% thought the participant-therapist contact to be personal, and 71% indicated they had not missed face-to-face contact. At follow-up, these aspects were rated similarly. Participants were also asked to rate the degree to which they would recommend the treatment to others, on a 10-point scale ranging from 1 (*No*) to 10 (*Yes*). The average score on this item was $M = 8.9$ ($SD = 1.5$).

Progress during treatment. During treatment, after each treatment phase, participants rated the degree to which their panic symptoms impaired daily functioning. As shown by Figure 4.2, impairment declined over the course of treatment, with an effect size of $d = 1.2$ between the first and last measure. Noticeable improvement occurred only after the cognitive restructuring phase. Post-hoc, we tested mean impairment values before this phase with mean impairment values after this phase. This contrast was highly significant ($z = 8.2$, $P < .001$).

Predictors of outcome. Using post-treatment PDSS-SR scores and attack frequency as the outcome variables of interest, we tested several predictors of posttreatment symptomatology through two separate multiple regression analyses. Specifically, we tested the significance of pre-treatment symptom severity, agoraphobic avoidance (MI-AAL), gender, education level (low/high), pretest medication status (no/yes), and time of measurement (posttest/follow-up). Further, we explored interaction effects of

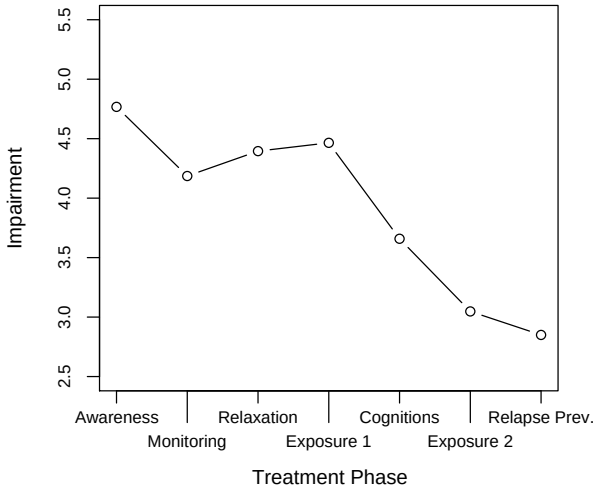


Figure 4.2: Change in impairment during treatment.

the time variable and the other variables, i.e. time by pre-treatment symptom severity, by gender, by education, MI-AAL, and by medication status.

With regard to the PDSS-SR, none of the interaction terms were significant. Hence, the model was simplified to include only the main effects. In this model, medication status [$t(50) = 2.0$; $P = .05$] and education [$t(50) = 0.3$; $P = .76$] had no effects. However, PDSS-SR pretest scores [$t(50) = 5.7$; $P < .001$] and gender [$t(50) = 3.0$; $P = .019$] predicted post-treatment symptoms significantly. Post treatment scores were higher for females, and for those displaying higher symptoms levels at pretest.

With regard to attack frequency, only the interaction of time and pre-treatment severity was significant [$t(50) = 3.8$; $P < .001$], signalling a significant larger decline in the number of panic attacks at follow-up compared to the posttest in those experiencing more attacks at pretest. In the reduced model, in which we retained only this interaction and the main effects, none of the variables predicted post-treatment scores significantly, including gender and pre-treatment agoraphobic avoidance (MI-AAL scores).

Given the regression results, we further explored the effect of pretest symptom severity. Based upon the PDSS-SR baseline score, we split the sample into two groups (based on the PDSS-SR cut-off of 8) and determined the effect size for each group on this measure. This revealed considerable higher effects in the group with higher pretest levels (posttest $d = 2.1$; follow-up $d = 2.3$).

4.4 Discussion

In comparison to a waiting-list, web-based therapist-assisted CBT induced moderate to large reductions in panic symptoms and general psychopathology in a heterogeneous community sample of clients suffering from chronic symptoms of panic disorder. Treated participants were six times more likely to experience a reduction in panic attacks of at least 50%. Client satisfaction was high. After three years, even though the participants never met their therapist, 80% returned for follow-up measurements. On the long term, improvements were found to be more pronounced.

4.4.1 RCT

One of the strengths of this study is that it included a community sample representative of the population that is encountered in everyday practice. Unlike most trials of internet-delivered CBT, we did not exclude respondents with subsyndromal PD/A. Subthreshold PD/A is common and constitutes a major proportion of the cases treated in routine care. Therefore, compared to most existing trials of Web-based CBT, the results of this trial are more likely to generalize to the applied setting.

The average between-effect size found in this study of d =.7 is considerably higher than the effects of internet-based programs without therapist support (d = .2, Spek et al., 2007). This lends further support to the hypothesis that therapist guidance is a critical determinant of the efficacy of internet-based psychotherapy (Palmqvist et al., 2007). The results also support the suggestion that scheduled therapist guidance increases adherence, as evidenced by the relatively low dropout rate of 13% found in this study, compared to the high dropout rates that are observed in online self-help programs (Eysenbach, 2005; Farvolden et al., 2005).

At first sight, the effect sizes appear somewhat smaller than those observed in trials with less therapist support (Carlbring et al., 2006, 2005, 2001; Klein et al., 2006). However, the larger sample variance in symptom severity in this trial precludes direct comparisons with these programs. Clearly, more research is needed concerning the relation between the amount of therapist involvement and outcome. Such research should involve the systematic manipulation of the amount of therapist input. However, given the large effect sizes that we observed among those with more severe symptom levels, we recommend that such studies include therapist involvement over and above the maximum of six hours of therapist time of existing web-based CBT of PD/A.

4.4.2 Three-year follow-up

Our follow-up study confirms and extends previous studies of web-based treatments demonstrating stable effects of online treatment (Carlbring et al., 2005, 2006; Klein et al., 2006; Shandley et al., 2008). In these trials, effects were maintained up to one year after treatment. Our three-year follow-up suggests that this also holds for the longer term. At follow-up, participants reported high client satisfaction, good

outcome characterized by large effect sizes, and further improvements compared to those observed at posttest. This, together with the fact that participants reported chronic symptoms at pretest, provides preliminary but encouraging evidence that online treatment is effective in inducing sustainable changes.

4.4.3 Limitations

One of the limitations of this study is that we did not use a formal clinical interview to assess treatment outcome. Instead, outcome was measured through self-report questionnaires. One of the key advantages of Internet-based therapy is the potential of reaching people whose access to treatment is poor. The fact that a clinical interview requires face-to-face contact with an expert could be counterproductive in that it may deter those who tend to avoid treatment because they are reticent about such contacts. Nevertheless, observer-rated diagnoses may add to the validity of the results. Telephonic diagnostic interviews may provide a future solution.

A second limitation of this trial is that the long-term effects cannot be attributed to the treatment alone, because the three-year follow-up study was uncontrolled. During the period between the posttest and the follow-up, participants were subject to many influences which we could not control. Specifically, we found that a quarter of the participants had followed additional treatment. However, there was no difference in long-term outcome between those who followed additional treatment and those who did not. We observed this before in our previous studies (Ruwaard et al., 2007, 2009). Although some clients need additional treatment, online therapy appears to be a sufficient intervention for the majority of clients. Another limitation of the long-term follow-up is that we compared retrospective panic attack estimates at follow-up with more accurate and conservative monitoring of panic attacks at pre- and posttest. de Beurs, Lange, and van Dyck (1992) found that retrospective estimates of panic attack frequency can be substantially higher than those obtained through monitoring. Thus, the observed long-term effects may represent an underestimate. Further comparative studies of the long-term benefits of this treatment are necessary to corroborate our findings.

Third, with the absence of a formal component analyses, it is not possible to identify key ingredients of the therapy. Through repeated measurements during treatment, we

found some support in favor of combined treatment - i.e., a combination of exposure techniques, cognitive interventions, and applied relaxation - because notable change occurred relatively late, after each technique had been applied. More incisive studies would be useful to assess the necessity of individual techniques. However, repeated measurement during treatment has wider use than a measure of efficacy alone. Our therapists indicated that they found the impairment scores very useful as an additional check on the progress of their clients.

4.4.4 Conclusion

Our study provides further evidence for the efficacy and feasibility of web-based therapist-assisted CBT. Web-based CBT is a promising online alternative to traditional treatments for panic symptoms, and we recommend exploring its effectiveness through field-testing in routine practice. From a cost-benefit point of view, future research should develop criteria to guide the identification of clients who require varying intensities of (web-based) treatment. Costs, however, should come second to outcome.

Ruwaard, J., Lange, A., Broeksteeg, B., Renteria-Agirre, A., Schrieken, B., Dolan, C.V., & Emmelkamp, P. (2012). Online cognitive behavioral treatment of bulimic symptoms: A randomized controlled trial. *Clinical Psychology and Psychotherapy*, epub ahead of print. doi: 10.1002/cpp.1767

Chapter 5

Online Cognitive Behavioural Treatment of Bulimic Symptoms

Background Manualised cognitive behavioral treatment (CBT) is underutilised in the treatment of bulimic symptoms. Internet-delivered treatment may reduce current barriers. **Objective** To assess the efficacy of a new online cognitive behavioral treatment of bulimic symptoms. **Method** Participants with bulimic symptoms ($N = 105$) were randomly allocated to online CBT, bibliotherapy or a waiting-list/delayed treatment condition. Data were gathered pre-treatment, post-treatment and at one-year follow-up. **Outcome Measures** Primary outcome measures were the Eating Disorder Examination-Questionnaire (EDE-Q) and the frequency of binge eating and purging episodes. Secondary outcome measure was the Body Attitude Test. **Results** Dropout from internet treatment was 26%. Intention-to-treat ANCOVAs of posttest data revealed that EDE-Q scores and the frequency of binging and purging reduced more in the online CBT group compared to bibliotherapy and the wait-list (pooled between-group effect size $d = .9$). At 1-year follow-up, improvements in the online CBT group had sustained. **Conclusion** This study identifies online CBT as a viable alternative in the treatment of bulimic symptoms.

Introduction

Bulimia Nervosa (BN) is an eating disorder characterized by cycles of binge eating followed by compensatory purging such as self-induced vomiting or laxative misuse. It typically affects young women, at an estimated life-time prevalence rate of 1% (Fairburn & Harrison, 2003). Subclinical BN, in which bulimic symptoms are present but below diagnostic threshold, has a higher estimated prevalence of 5% (Hoek & van Hoeken, 2003). Untreated, bulimic symptoms tend to persist or develop into different types of eating disorder. With post-treatment recovery rates of 30 to 50 percent, manualized cognitive behavioral therapy (CBT) is the current treatment of choice (Fairburn & Harrison, 2003; National Institute for Clinical Excellence, 2004b). Even so, evidence suggests that only a minority of patients with bulimic symptoms receive CBT in accordance with the guidelines (Crow, Mussell, Peterson, Knopke, & Mitchell, 1999; Hoek & van Hoeken, 2003; Mond, Hay, Rodgers, & Owen, 2007; Mond, Myers, Crosby, Hay, & Mitchell, 2010). The underutilization of manualized CBT has therefore been identified as a pressing issue in the treatment of bulimic symptoms (Wilson, 2005).

Internet-delivered treatment may reduce barriers to care in those unwilling or unable to access evidence-based treatment. Recent studies found encouraging effects of internet-delivered programs for the prevention and early intervention of eating disorders (Taylor et al., 2006; Jacobi et al., 2007; Paxton, McLean, Gollings, Faulkner, & Wertheim, 2007). With regard to the treatment of bulimic symptoms, results are mixed. The efficacy of Internet-delivered CBT for bulimic symptoms has been studied in five controlled trials. Large effects were observed by Ljotsson and colleagues in a waiting-list controlled trial of an online guided self-help treatment program (Ljotsson et al., 2007). In the four remaining trials, effects of online guided self-help were moderate-to-large in one (Sánchez-Ortiz et al., 2010), small-to-moderate in two (Fernandez-Aranda et al., 2009; Nevonen, Mark, Levin, Lindström, & Paulson-Karlsson, 2007), and small in one trial (Robinson & Serfaty, 2008). Treatment adherence was low across studies, with dropout rates varying from 35% to 82%. Most trials included face-to-face diagnostic interviews to assess inclusion criteria or outcome. This limits conclusions about the effects of online treatment per se, since it cannot be ruled out

that the effects were confounded by the additional face-to-face contact. It also limits the generalisability of the results, since the required face-to-face sessions may have deterred those who tend to avoid treatment because they are reticent about such contact.

Based on principles of online treatment, applied before in the online treatment of posttraumatic stress, work-related stress, depression and panic disorder (Lange, Rietdijk, et al., 2003; Ruwaard et al., 2007, 2009; Ruwaard, Broeksteeg, Schrieken, Emmelkamp, & Lange, 2010), we developed a therapist-guided internet-delivered CBT program for bulimic symptoms that does not require any face-to-face contact. In this article, we present the results of a trial in which we assessed the effects of this treatment and its relative efficacy in comparison to a waiting list and unsupported bibliotherapy. Bibliotherapy has been associated with small-to-moderate effects in the treatment of bulimic symptoms and other eating disorders (Hay, Bacaltchuk, Stefano, & Kashyap, 2009; Perkins, Murphy Rebecca, Schmidt Ulrike, & Williams, 2006). Since effects of online CBT were large and similar to face-to-face CBT in our previous trials (Lange, Rietdijk, et al., 2003; Ruwaard et al., 2007, 2009, 2010), we hypothesized that online CBT would be more effective than both bibliotherapy and the waiting-list in reducing bulimic symptoms.

5.1 Treatment

The online treatment is a twenty-week standardized program based on existing CBT manuals for BN (Lange, De Vries, Gest, & Van Oostendorp, 1994; Vanderlinden, Pieters, Probst, & Norré, 2004). It comprises awareness training, monitoring, planning and structuring of meals, exposure and response prevention, cognitive restructuring, behavioral experiments, mirror exposure, positive self-verbalization, and relapse prevention. Screening, treatment, and outcome assessments are conducted without face-to-face contact. Psycho-education, homework assignments, and therapist support are mostly standardized and are delivered through secure web pages that clients share with their personal therapist. Execution of the treatment protocol is governed by a computer program which automatically assigns tasks to clients and therapists. This program also sends automatic alerts and reminders when treatment progress deviates

too greatly from the intended pace.

The treatment is structured into five phases and defines 24 scheduled therapist feedback moments (phase 1-4: five feedback texts; phase 4: four feedback texts). Full treatment requires approximately 13 hours of therapist time.

Phase 1 serves to raise client awareness and defines the goals for treatment. Clients receive psycho-education on causes and maintaining factors of BN, and engage in self-monitoring and structured writing exercises to gain insight into their eating problems. Next, they are introduced to a generic set of self-control procedures designed to reduce the probability of binge eating and compensatory behavior. With therapist assistance, clients learn how to prepare a personal self-control plan by tailoring these generic procedures to their specific situation. In the early phase of treatment, the clients receive psycho-education with regard to the differences in dynamics between binge eating and compensatory behavior. While binge eating is considered to be triggered by a loss of impulse control, compensatory behavior is considered to be a deliberate, voluntary and 'rational' response to reduce weight increase after binge eating episodes (Lange et al., 1994). Accordingly, clients are requested to sign an agreement to quit compensatory behavior.

In phase 2, clients begin to apply their personalized self-control plan and post regular reports of successes and difficulties in executing the plan. Clients evaluate their plan based upon their experiences and make adjustments, if needed. Self-monitoring is intensified by planning meals and daily activities ahead of time, to gain better insight in the pattern of binge eating and compensatory behaviors, and to learn how to prevent these behaviors. In this phase, clients are also encouraged to disclose their eating problems to a trusted person.

In phase 3, cognitive restructuring techniques are introduced. Clients compose a list of short statements reflecting dysfunctional attitudes and opinions, and challenge these statements by formulating alternative interpretations. In addition, clients conduct several behavioral experiments to challenge key attitudes and opinions in real-life situations. They carefully prepare these experiments and anticipate on possible outcomes. Before the client actually conducts the experiment, the therapist ensures that the experiments are practicable and realistic, and have a good chance of success. As in phase 2, experiences are used to evaluate and adjust the self-control

plan. Phase 4 targets body image disturbance and self-esteem. The phase starts with an exposure exercise. Clients are asked to look at their body in the mirror, and to provide a detailed description of their body (Delinsky & Wilson, 2006). Next, they change perspective in a structured writing exercise. They write a supporting letter to an imaginary friend, who has similar concerns about body shape and weight. In a third exercise, clients write down a number of traits or features that they value in a person, apart from physical appearance. They rate both themselves and someone whom they admire on these traits, compare these ratings, and reflect on the results. Finally, they compose a list of positive self-statements from the material that they generated so far, which they read aloud a couple of times a day (Lange et al., 1997).

In the last phase, clients finalize their self-control plan and write a personal relapse prevention plan containing self-directed instructions on how to recognize signals that might predict a relapse, or how to deal with relapse, should it occur. Next, clients enter a 'test period' of two weeks in which they have no contact with the therapist. After this period, final adjustments to the relapse prevention plan are considered and treatment is ended.

5.2 Method

5.2.1 Design

The study was a randomised controlled trial with measurements at baseline, immediately after treatment, and one year after treatment (controlled trials registry: ISRCTN06477195). Through computerised permuted block randomisation (Beller, Gebski, & Keech, 2002), participants were allocated to online CBT, unsupported bibliotherapy or a waiting-list/delayed treatment control group. This randomisation method guaranteed an allocation ratio of 1:1:1.

5.2.2 Participants

Participants were recruited from the general Dutch community between August 2006 and September 2007. Since we aimed to generalize findings to a population with varying levels of bulimic symptoms, a formal diagnosis of BN was not an inclusion

requirement. To be included in the study, respondents had to report recurrent binge eating, inappropriate weight-control behaviour (either in the form of purging, physical exercise or dieting), and elevated concern with body shape and weight. Additional inclusion criteria were: age above 16 years, BMI $\geq$ 18, no heightened risk of dissociation or psychosis, no indications of automutilation, no history of suicide attempts within the past 3 years, and no suicidal ideation indicative of current plans, no drug and alcohol abuse, no use of neuroleptic medication or unstable dosages of other psychiatric medication, no concurrent psychotherapy, and no indication that another psychological disorder was prevailing.

5.2.3 Target sample size

Since this was the first controlled evaluation of the online treatment, power calculations could not be based on previous research findings. Therefore, we provisionally powered the study to detect a large between-group effect (a Cohen's *d* of .8). Approximately 105 participants (35 per group) were needed to detect this effect (with ANCOVA, an estimated pre-post correlation of .5, a 10% attrition factor and Bonferroni corrections for comparisons on 4 outcome measures). Therefore, recruitment stopped when 105 participants were included.

5.2.4 Procedure

Enrollment & screening Dutch magazine articles and information posters announced the study and referred to a website, which provided additional information and an application form. Applicants received an automated e-mail with instructions how to access a secure website, where they could start an online screening. To assess inclusion criteria, the online screening comprised - in addition to the measures of bulimic symptomatology described below - a biographic questionnaire, a substance abuse questionnaire, the Depression Anxiety Stress Scales (DASS; de Beurs et al., 2001; Lovibond & Lovibond, 1995), the Impact of Event Scale (IES; Horowitz, Wilner, & Alvarez, 1979), the Somatoform Dissociation Questionnaire (SDQ-5) (Nijenhuis et al., 1997), the Dissociation-Questionnaire (DIS-Q; Vanderlinden et al., 1991), the Screening Device for Psychotic Disorder (SDPD; Lange, Schrieken, et al., 2000), three items from the Self-Harm Inventory (SHI; Sansone, Wiederman, & Sansone, 1998),

and a measure of suicidal ideation based on the work of Joiner (Joiner et al., 2003). Medication usage was assessed through open-ended questions.

Trained diagnosticians used the questionnaire data to prepare and focus a 30-minute semi-structured interview. In this interview, results of the online screening were discussed with the patient, after which the diagnostician decided on the eligibility of the applicant. Excluded respondents were referred to their GPs or to mental health centres in their vicinity. Eligible respondents downloaded, signed and returned an *Informed Consent*. On receipt of this form, research assistants used the allocation sequence (which was concealed to the diagnosticians) to allocate participants to the experimental groups.

Experimental conditions Participants, allocated to online CBT, started treatment immediately. Participants in the bibliotherapy group received a hard-copy of *Overcoming bulimia and binge eating*, a Dutch self-help book for BN (*Overcoming bulimia and binge eating* by Vanderlinden (2002), which is based on the same cognitive-behavioral principles as applied in the online treatment. Apart from an initial introductory e-mail with instructions how to use this book, no support was provided to the participants. Participants in the waiting list control condition started online CBT after twenty weeks.

Therapists The treatments were provided by graduate students in clinical psychology or recently graduated psychologists ($N = 23$). All had followed advanced courses in CBT and received additional training in administering online CBT. They were supervised by two senior specialists in online CBT.

Posttest and follow-up After 20 weeks, participants received an e-mailed invitation for online post-test measurements. One year after the post-test, participants in the online CBT and bibliotherapy group were similarly invited to online follow-up measurements. Participants in the waiting list control group completed follow-up measurements at a later time (i.e., one year after they had finished treatment).

5.2.5 Outcome measures

Primary outcome measures were the frequency of binge eating and purging episodes and the Eating Disorder Examination-Questionnaire (EDE-Q; Fairburn & Beglin, 1994;

Nauta, Hospers, Kok, & Jansen, 2000). Secondary outcome measure was the Body Attitude Test (BAT; Probst, Vandereycken, Van Coppenolle, & Vanderlinden, 1995; Probst, van Coppenolle, & Vandereycken, 1998).

EDE-Q The EDE-Q is a 30-item self-report list assessing the severity of behavioural, emotional and cognitive symptoms of eating disorder over the past 28 days. Twenty-two items of the EDE-Q assess symptoms on a 7-point Likert-scale, scored from 0 to 6, with higher scores reflecting more severe symptoms. We used the mean score of these items as a global indicator of eating disorder severity (Cronbach's α = .9; 330-day test-retest reliability: r = .79; Mond, Hay, Rodgers, Owen, & Beumont, 2004a; Peterson et al., 2007). At a cut-off of 2.3, this indicator identifies patients with eating disorders with a sensitivity of 83% and a specificity of 96% (Mond, Hay, Rodgers, Owen, & Beumont, 2004b). EDE-Q contains eight additional items that address key behavioral aspects of different types of eating disorders. From these, we chose the *objective bulimic episodes* item to assess the frequency of binge eating, and the sum of the *self-induced vomiting* and *laxative misuse* items to assess the frequency of purging episodes. EDE-Q does not assess fasting and the extent to which exercise is compensatory. In addition, the reliability and validity of EDE-Q items assessing non-purging compensatory behaviors is not clear (Berg et al., 2011). Hence, change in non-purging compensatory behavior was not assessed in this study.

BAT The BAT is a 20-item self-report survey assessing body dissatisfaction. Items are scored on a 6-point Likert-scale, with higher scores indicating more dissatisfaction. We used the total score, ranged 20-120, as a global indicator of body dissatisfaction. At a cut-off of 36, this indicator has a sensitivity of 69% and a specificity of 80% in the detection of people with an eating disorder (Probst, Pieters, & Vanderlinden, 2008).

5.2.6 Analyses

Statistical significance and effect size. Between-group differences at posttest and follow-up were examined using two-tailed ANCOVA's, using pretest scores as a single covariate and Holm-Bonferroni adjustments to to maintain the family-wise significance level α at .05 (Holland & DiPonzio Copenhaver, 1988). Follow-up comparisons did not

include data of the waiting-list, because the participants in this group provided follow-up data at a later time. Binge eating and purging frequency data were highly skewed. Consequently, we reverted to rank-score ANCOVA (Conover & Iman, 1982; LaVange & Koch, 2006) to assess group differences with these data, and chose to summarize these variables as medians and interquartile ranges. Analyses included all participants, irrespective of treatment adherence and attrition. Missing values were handled through last-observation-carried-forward (LOCF) data imputation. Within-group gain scores and between-group differences were standardized to Cohen's *d*, using the pooled standard deviation of the pretest scores as the standardizer (J. Cohen, 1988). Confidence intervals around *d* were approximated from the central *t*-distribution (Robey, 2004).

Clinical significance. We tested the differential probability of a clinically relevant outcome after online treatment compared to both control groups with two-sided Fisher's exact tests, and expressed this difference as odds ratios (OR; Hillis & Woolson, 2002). For binge eating and purging, we defined abstinence (i.e., a frequency of 0) as the clinically relevant outcome. With regard to EDE-Q and BAT, we applied the principles of reliable clinically significant change (RCSC; Jacobson & Truax, 1991). We used the reliable change index to determine how many scale points a participant had to change to rule out measurement error (EDE-Q: 1.1 points; BAT: 18.5 points). RCSC was defined as reliable change from a pretest score above the clinical cut-off (EDE-Q: < 2.3; BAT: < 36) to a posttest score below the cut-off .

5.3 Results

5.3.1 Participants

Recruitment resulted in 273 respondents. Of these, 91 (33%) did not complete the screening. Of the remaining 182 respondents, 105 (58%) met inclusion criteria (cf. Figure 5.1). As shown in Table 5.1, baseline characteristics were equally distributed across the three experimental groups. Participants were female ($n = 104$; 99%) and about 31 years old. Self-reported illness history indicated chronic symptoms (the average duration of symptoms was $M = 11$ years, $SD = 9$), for which 62%

(n = 65) had received earlier treatment. Primary measures indicated moderate-to-severe bulimic symptoms (cf. Table 5.2). The vast majority (80%; n = 84) engaged in purging behavior to compensate for regular binge eating. Symptoms appeared to be more severe in the online CBT group. However, these differences were not significant (EDE-Q/BAT: largest ANOVA $F_{2,102}$=2.1,P = .12, Binge eating/Purging: largest Kruskal-Wallis $\chi^2_{df=2}$ = 4.3, P = .11). Comorbid symptoms were mild, as measured by participants' scores on the Depression Anxiety Stress Scales, which were slightly above clinical cut-offs of the subscales of this measure (cf. Table 5.1).

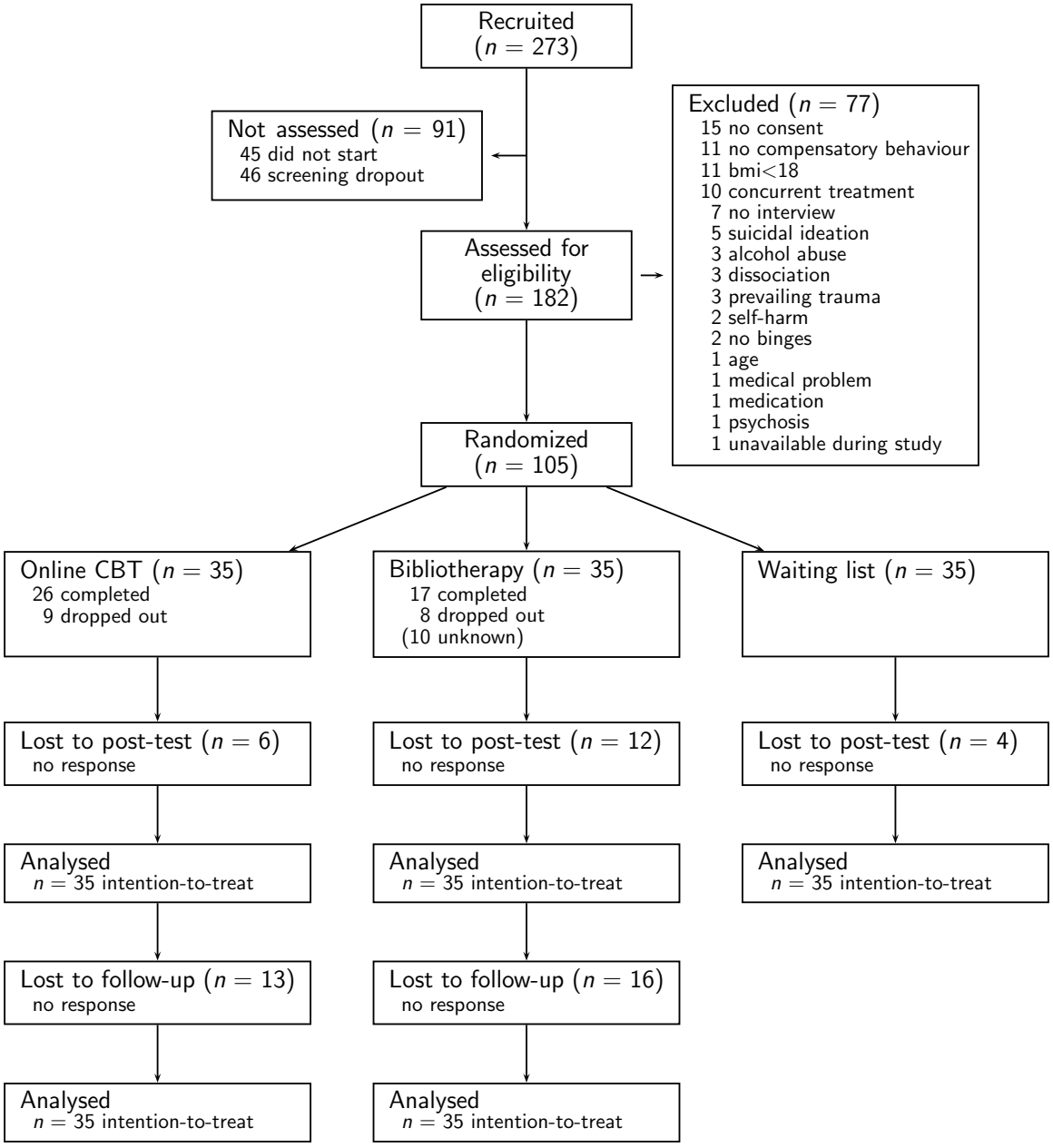


Figure 5.1: Flowchart of study participation.

Table 5.1: Baseline characteristics of the three experimental groups (N = 105).

Characteristic[a]	Wait list		Bibliotherapy		Online CBT	
Demographic						
Female	97%	34	100%	35	100%	35
Age (*M SD*)	30	10	31	9	32	11
Tertiary Education	83%	29	74%	26	78%	27
Living with partner	43%	15	49%	17	46%	16
BMI						
underweight	0%	0	0%	0	0%	0
normal	74%	26	78%	27	80%	28
overweight	20%	7	9%	3	6%	2
obesity	6%	2	14%	5	14%	5
Symptoms						
Years with symptoms (*M SD*)	10	8	11	9	13	12
Received treatment before	66%	23	54%	19	66%	23
Comorbidity[b]						
Depression (*M SD*)	15.4	9.3	12.3	10.2	14.3	10.0
Anxiety (*M SD*)	5.8	6.7	5.0	4.6	6.5	6.6
Stress (*M SD*)	14.3	9.0	13.7	10.0	16.3	9.5
Psychoactive Medication	14%	5	14%	5	14%	5

[a] Values represent subsample percentage and size unless otherwise noted.

[b] Assessed by the Depression Anxiety Stress Scales (DASS; Lovibond & Lovibond, 1995). Clinical cut-off scores: Depression: 12; Anxiety: 5; Stress: 14.

5.3.2 Adherence & attrition

In the online CBT group, 74% (n = 26) of the participants completed the full treatment. In the bibliotherapy group, 25 participants provided adherence data, of which 68% (n = 17) worked through all chapters of the book. At post-test, study attrition was 35% in the bibliotherapy group (n = 12), 17% (n = 6) in the online CBT group, and 11% (n = 4) in the waiting-list (a non-significant difference: Fisher's Exact test: P = .07). At follow-up, attrition was 37% (n = 13) in the online CBT group, and 46% (n = 16) in the bibliotherapy group (Fisher's Exact test: P = .31). In the period between the pretest and the posttest, additional psychological treatment was started by 31% (n = 11) of the bibliotherapy group, 11% (n = 4) of the waiting-list group and 6% (n = 2) of the online CBT group; a significant difference (Fisher's Exact Test: P = .004).

Table 5.2: Changes in bulimic symptomatology observed in an online CBT group, a bibliotherapy group and a waiting-list control condition, from pretest to 1-year follow-up (intention-to-treat sample).

Measure[a]	**Pretest**		**Posttest**				**1-YR Follow-up**			
	n = 105[b]		**n = 105**				**n = 70**[d]			
group	*M*	*SD*	*M*	*SD*	*d*	*t*[c]	*M*	*SD*	*d*	*t*
EDE-Q										
online CBT	4.2	1.0	2.7	1.5	1.5	6.9***	2.8	1.5	1.4	6.5***
bibliotherapy	3.8	1.0	3.2	1.2	.6	3.9**	2.6	1.3	1.2	5.4***
waiting-list	3.8	1.0	3.4	1.3	.4	2.4				
BAT										
online CBT	61.6	16.7	49.3	19.6	.7	5.6***	49.6	22.4	.7	4.5***
bibliotherapy	53.4	17.7	48.8	18.8	.3	2.4*	43.0	18.5	.6	3.7***
waiting-list	58.9	16.8	56.3	18.7	.1	1.2				
	Mdn[e]	*IQR*[e]	*Mdn*	*IQR*	Δ	Z[c]	*Mdn*	*IQR*	Δ	Z
Binge eating										
online CBT	15	16	2	7	13	4.2***	2	9	13	4.0***
bibliotherapy	10	15	8	10	2	2.4*	7	14	4	3.5***
waiting-list	8	12	5	7	3	1.9				
Purging										
online CBT	14	25	0	6	14	3.1**	1	6	13	3.0***
bibliotherapy	10	26	6	16	4	2.0*	1	12	9	2.9**
waiting-list	6	12	4	11	2	.2				

[a] EDE-Q: Eating Disorder Examination-Questionnaire; BAT: Body Attitude Test. Behavioral items of the EDE-Q self-report were used to assess one month frequencies of binge eating and purging (defined as the sum of the frequencies of self-induced vomiting and laxative misuse) .

[b] n = 35 per group

[c] t: paired t-test statistic; Z: Wilcoxon signed-rank test statistic. *: $p < .05$; **: $p < .01$; ***: $p < .001$

[d] Follow-up data of the waiting-list are not shown because follow-up measurements in this group were administered at a later time.

[e] Because the distributions of the frequencies of binge eating and purging episodes were highly skewed, the median (Mdn) and interquartile range (IQR) are used as summarizing statistics. Δ denotes the median difference (pretest - posttest, and pretest - follow-up).

5.3.3 Statistical significance and effect size

Within-group improvements. Summary statistics and tests of symptom reduction within the groups are shown in Table 5.2. Figure 5.2 provides a graphical summary. Bulimic symptoms reduced significantly in the online CBT group and the bibliotherapy group, but not in the waiting-list.

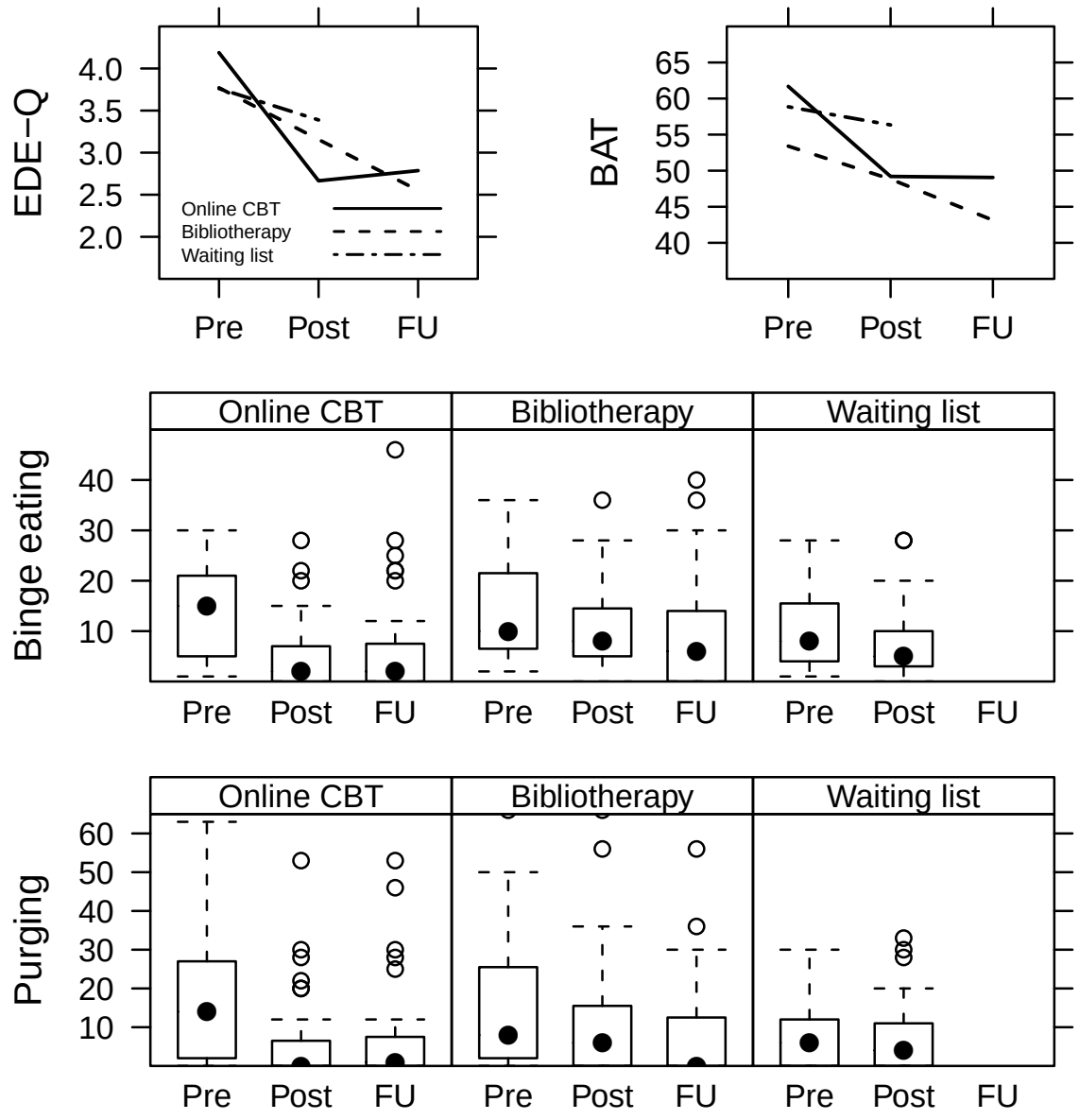


Figure 5.2: Change in bulimic symptoms, from pretest to 1-year follow-up, per experimental condition.

Between-group comparisons. Table 5.3 shows the results of the primary hypothesis tests of between-group differences. At posttest, binge eating frequency, purging frequency and global eating disorder symptom severity (EDE-Q) was significantly less in the online CBT group, compared to the bibliotherapy group and the waiting-list group ($P\ geq$ .04; pooled between-group effect size: d = .9 , range: .7 $< d <$ 1.2). Posttest differences in symptom severity between the bibliotherapy group and the waiting-list group were non-significant (d $<$ .2; P = .99). At follow-up, due to improvements in the bibliotherapy group, differences between the online CBT group and the bibliotherapy group were no longer significant (d $<$ .2; P = .99).

Table 5.3: Online CBT vs. bibliotherapy vs. a waiting-list control condition: pair-wise differences in reduction of bulimic symptomatology, at posttest and 1-year follow-up (intention-to-treat analysis).

			Posttest				**Follow-up**			
Measure[a]			**Ancova**		**Effect Size**[c]		**Ancova**		**Effect Size**[c]	
Contrast			$F_{1,67}$	p[b]	d	$CI_{.99}$	$F_{1,67}$	p[c]	d	$CI_{.99}$
EDE-Q										
online CBT	-	waiting-list	15.6	.003	1.2	± .5	[d]			
online CBT	-	bibliotherapy	8.7	.04	.9	± .5	.1	.99	.2	± .6
bibliotherapy	-	waiting-list	1.2	.99	.2	± .4				
Binge Eating										
online CBT	-	waiting-list	9.3	.04	.7	± .5				
online CBT	-	bibliotherapy	15.0	.003	1.0	± .5	1.3	.99	.3	± .4
bibliotherapy	-	waiting-list	.7	.99	-.2	± .5				
Purging										
online CBT	-	waiting-list	10.8	.02	.8	± .5				
online CBT	-	bibliotherapy	10.6	.02	.8	± .5	.5	.99	.2	± .5
bibliotherapy	-	waiting-list	.1	.99	.1	± .5				
BAT										
online CBT	-	waiting-list	10.1	.03	.6	± .4				
online CBT	-	bibliotherapy	5.0	.26	.5	± .3	.0	.99	.1	± .4
bibliotherapy	-	waiting-list	1.0	.99	.1	± .3				

[a] EDE-Q: Eating Disorder Examination-Questionnaire; BAT: Body Attitude Test. Behavioral items of the EDE-Q self-report were used to assess one month frequencies of binge eating and purging (defined as the sum of the frequencies of self-induced vomiting and laxative misuse)

[b] *P*-values were Holm-Bonferroni adjusted to correct for the total of 16 tests of significance.

[c] Cohen's *d* effect size; $CI_{.95}$: 95% confidence interval of *d*.

[d] Follow-up data only allowed for the comparison of online CBT and bibliotherapy.

LOCF-ITT may be overly liberal if attrition is unbalanced across experimental groups, or if posttest data is carried forward to impute missing follow-up data (Molenberghs et al., 2004). Therefore, we reran analyses using available data only (posttest: $n = 83$; follow-up: $n = 41$). In terms of statistical significance, results of the ITT analyses and the available data analyses were similar.

5.3.4 Abstinence and reliable clinically significant change

As shown in Table 5.4, participants in the online CBT group were more likely to be abstinent from purging in comparison to control group participants at post-test (OR: 7.2; $P = .02$), and more likely to achieve reliable clinically significant change with respect to EDE-Q (OR: 4.5; $P = .01$). At follow-up, differences between the online CBT group and the bibliotherapy group were not significant (OR: 1.0 to 1.4; $P = .99$).

5.3.5 Client satisfaction with online CBT

Participants following online CBT completed a post-test evaluation questionnaire ($n = 51$; including those participants who started treatment at a later time). On a 1 to 10 point scale, they rated the overall value of treatment with an average of 7.5 ($SD = 1.3$) and the contact with their therapists with an average of 8.6 ($SD = 1.9$). Seventy-six percent of the participants indicated they had not missed face-to-face contact, and 94% stated that they would recommend the treatment to others.

5.4 Discussion

As hypothesised, posttest reductions in bulimic symptoms were substantially greater with online CBT, in comparison to bibliotherapy and a waiting list. One year after treatment, differences between online CBT and bibliotherapy were no longer significant, due to improvements in the bibliotherapy group. Participants allocated to online CBT reported large, clinically relevant and stable reductions in bulimic symptoms. Notwithstanding the absence of face-to-face contact, treatment dropout was acceptable, and client satisfaction was high.

Results provide further support for the hypothesis that therapist support is a critical determinant of treatment efficacy and adherence in internet-delivered treatment

Table 5.4: Online CBT vs. control conditions (bibliotherapy and a waiting-list): rates of reliable clinically significant change (RCSC; EDE-Q/BAT) and abstinence (binge eating/purging), at posttest and 1-year follow-up (intention-to-treat analysis).

Measure[a]	**Pretest**	**Posttest**				**Follow-up**			
group	n	RCSC	OR	($CI_{.95}$)	p[c]	RCSC	OR	($CI_{.95}$)	p[c]
EDE-Q									
online CBT	33	42%	4.5	(1.5 - 13.9)	.01	42%	1.3	(.4 - 4.3)	.99
bibliotherapy	31	16%				32%			
waiting-list	34	12%							
BAT									
online CBT	32	22%	3.0	(.7 - 13.4)	.51	25%	1.0	(.3 - 4.0)	.99
bibliotherapy	29	14%				24%			
waiting-list	31	3%							
Binge eating									
online CBT	35	37%	3.5	(1.2 - 10.4)	.07	40%	1.3	(.4 - 3.8)	.99
bibliotherapy	35	14%				34%			
waiting-list	35	14%							
Purging									
online CBT	28	39%	7.2	(1.8 - 35.4)	.02	39%	1.4	(.4 - 5.2)	.99
bibliotherapy	26	8%				30%			
waiting-list	24	8%							

[a]EDE-Q: Eating Disorder Examination-Questionnaire; BAT: Body Attitude Test. Behavioral items of the EDE-Q self-report were used to assess one month frequencies of binge eating and purging (defined as the sum of the frequencies of self-induced vomiting and laxative misuse).

[b]n = number of participants with clinical symptom levels at pretest: EDE-Q: > 2.3; BAT: > 36; binge eating and purging: frequency > 0.

[c]*P*-values reflect the results of a Fisher's Exact test of the differences between the online CBT group and the control conditions (at posttest: the combined bibliotherapy group and the waiting-list group; at follow-up: the bibliotherapy group). P-values were Holm-Bonferroni adjusted to correct for multiple testing.

[d]Follow-up data of the waiting-list group are not shown because follow-up measurements in this group were administered at a later time.

(Andersson & Cuijpers, 2009; Cuijpers et al., 2009). The self-help book and the online treatment comprised similar therapeutic principles and strategies. Yet, at posttest, effects of online CBT group were significantly better than those of bibliotherapy, while the effects of bibliotherapy were not significantly different from those of the waiting-list. In addition, participants in the bibliotherapy group were more likely to drop out of the study or seek additional treatment. Formal cost-effectiveness studies are needed to further inform us on the relative benefits of guided and unguided interventions.

Our results imply that face-to-face sessions are an optional ingredient of online CBT. Past trials of internet-based CBT for bulimic symptoms included face-to-face sessions (Ljotsson et al., 2007; Sánchez-Ortiz et al., 2010; Fernandez-Aranda et al., 2009; Nevonen et al., 2007; Robinson & Serfaty, 2008), which limits conclusions about the effects of online treatment per se, and may restrict the results of these trials to those clients who are not deterred by face-to-face contact. These limitations do not apply to this trial, since the study did not involve any face-to-face contact. Our results suggest that face-to-face sessions may be an optional ingredient of online CBT. In the study of Sánchez-Ortiz and colleagues, face-to-face contact was replaced with telephone contact when participants preferred not to meet face to face (Sánchez-Ortiz et al., 2010). If existing online programs, which currently require face-to-face sessions, are adapted along similar lines, the uptake of these programs might increase.

In the online CBT group, improvements at post-test remained stable up to one-year after treatment. However, in the bibliotherapy group, participants strongly improved after the end of treatment, which resulted in non-significant differences between the two experimental conditions at follow-up. Since the bibliotherapy did not have effects at post-test, it is likely that bibliotherapy had an indirect delayed effect; the bibliotherapy might have stimulated participants to seek additional treatment at a later stage. The fact that 31% of the participants in the bibliotherapy group started additional treatment before post-test, supports this reasoning. Our results suggest that bibliotherapy increases the probability of recovery by promoting positive attitudes towards treatment in a considerable portion of patients.

5.4.1 Limitations

First, results of the long-term follow-up should be interpreted with caution. Study attrition at follow-up was high. Participants, who experienced a more favourable long-term outcome, may have been more inclined to respond to the follow-up invitations. Also, no data were available with regard to additional treatment between post-test and the long-term follow-up. This is a serious omission, since these data could have provided more support for our post-hoc hypothesis about the effect of additional treatment at follow-up.

Second, this study focused on purging as compensatory behavior. Quantitative assessment of change in non-purging compensatory behaviours was not considered, since these behaviours are not adequately tapped by the EDE-Q (Berg et al., 2011). It is possible that some participants replaced purging with non-purging compensatory behaviour, such as excessive physical exercise, dieting, and fasting. Future studies should address this possibility. Third, we relied on dimensional screening during the recruitment of the participants. Formal diagnoses would have required in-person interviews, which were incompatible with the goals of the study. In effect, some included participants probably did not fully meet the diagnostic criteria of bulimia nervosa. This might be considered a limitation. However, the dimensional approach increased the external validity of the trial, since patients often present with subclinical symptoms in clinical practice. In addition, our data indicate that we included a rather severe group of participants, who suffered from chronic symptoms, for which more than half of the participants had been treated before. Nonetheless, formal diagnoses could have complemented our results. We will add clinician-administered assessment to our future trials, since there are indications that in-person and telephone structured interviews yield comparable conclusions (Aziz & Kenford, 2004; Crippa et al., 2008). Fifth and finally, results would have been more compelling if treatment satisfaction could have been reported for bibliotherapy as well.

5.4.2 Conclusion

The present study identifies web-based, standardized, therapist-assisted CBT as a viable and promising treatment for bulimic symptoms. Internet-delivered treatment may provide an acceptable treatment alternative for bulimic patients, who are reticent about face-to-face contact.

Ruwaard, J., Lange, A., Schrieken, B., Dolan, C.V. & Emmelkamp, P. (2012). The effectiveness of online cognitive behavioral treatment in routine clinical practice. PLoS ONE, 7(7): e40089. doi: 10.1371/journal.pone.0040089

Chapter 6

Online Cognitive Behavioural Treatment in Routine Clinical Practice

Background Randomized controlled trails have identified online cognitive behavioral therapy as an efficacious intervention in the management of common mental health disorders. **Objective** To assess the effectiveness of online CBT in routine clinical practice. **Design** An uncontrolled before-after study, with measurements at baseline, posttest, 6-week follow-up, and 1-year follow-up. **Participants & Setting** 1500 adult patients (female: 67%; mean age: 40 years) with a GP referral for psychotherapy were treated at a Dutch online mental health clinic for symptoms of depression (n = 413), panic disorder (n = 139), posttraumatic stress (n = 478), or burnout (n = 470). **Interventions** Manualized, web-based, therapist-assisted CBT, of which the efficacy was previously demonstrated in a series of controlled trials. Standardized treatment duration varied from 5 weeks (online CBT for Posttraumatic stress) to 16 weeks (online CBT for Depression). **Main Outcome Measures** Validated self-report questionnaires of specific and general psychopathology, including the Beck Depression Inventory, the Impact of Event Scale, the Panic Disorder Severity Scale-Self Report, the Oldenburg Burnout Inventory, and the Depression Anxiety Stress Scales. **Results** Treatment adherence was 71% (n =1071). Study attrition was 21% at posttest, 33% at 6-week FU and 65% at 1-year FU. Mixed-model repeated measures regression identified large reductions in primary symptoms ($P < .001$; $d = 1.9 \pm 0.2$ to $d = 1.2 \pm 0.2$), which sustained up to one year after treatment. At posttest, rates of reliable improvement and recovery were 71% and 52% in the completer sample (full sample: 55%/40%). Patient satisfaction was high. **Conclusions** Results suggest that online therapist-assisted CBT may be as effective in routine practice as it is in clinical trials. Although pre-treatment withdrawal and long-term outcomes require further study, results warrant continued implementation of online CBT.

Introduction

In the past decade, there has been a rapid expansion in the research and development of internet-based psychotherapeutic interventions. As a result, we now know that online interventions are feasible and efficacious in the prevention and treatment of a wide variety of common mental health disorders (Cuijpers, van Straten, & Andersson, 2008; Andersson & Cuijpers, 2009; Cuijpers et al., 2009; Andrews, Cuijpers, Craske, McEvoy, & Titov, 2010). Although effect sizes vary with program characteristics (e.g., whether human support is included, or whether the aim is prevention or treatment), the benefits of various approaches are clear. Within the field, there is general agreement that online interventions are pivotal in improving the accessibility and uptake of evidence-based care (F. Griffiths et al., 2006).

While the benefits of online interventions have been firmly established in controlled research, the performance of these interventions in routine clinical practice is less clear (Andersson et al., 2009). Online interventions do not fit traditional healthcare systems, and raise legal, ethical, and professional issues that are only partially resolved by current guidelines (Childress, 2000). Consequently, the implementation of online treatment in routine clinical practice has progressed slowly, which has limited the options for effectiveness research. Some interventions have been evaluated in real-world contexts, with positive results (Christensen, Griffiths, Korten, et al., 2004; Ritterband et al., 2008; Riper et al., 2009; Marks et al., 2003; Kaldo, Larsen, & Andersson, 2004; Bergström et al., 2009, 2010; Postel, ter Huurne, de Haan, & de Jong, 2009). However, the current evidence base with regard to the effectiveness of online interventions in routine practice is small. Large-sample effectiveness studies are needed before wide-scale dissemination of online interventions can be recommended.

In this article, we present a study of the outcome of online therapist-assisted cognitive behavioral treatment (CBT) of 1500 patients, who were treated for symptoms of depression, panic disorder, posttraumatic stress, or burnout at a Dutch online mental health clinic. The efficacy of these treatments was previously demonstrated in seven randomized controlled trials, which included a total of 629 participants. (Lange et al., 2001; Lange, van de Ven, & Schrieken, 2003; Knaevelsrud & Maercker, 2007, 2009; Wagner et al., 2006; Wagner & Maercker, 2007; Ruwaard et al., 2007, 2009,

2010). The objective of the present study was to assess the external validity of these trials, by examining the effectiveness of these treatments in routine clinical practice. Given the outcome of the controlled trials, we expected the treatment to produce large, significant, and clinically significant reductions in the relevant symptoms of psychopathology.

6.1 Method

6.1.1 Study design & setting

This was an uncontrolled pre/post/follow-up study. Data were obtained from the electronic patients records of Interapy PLC, a Dutch online mental health clinic. These records provide data that are routinely collected before treatment, immediately after treatment, six weeks after treatment, and one year after treatment. In March 2009, we queried the electronic patient database of the clinic. Starting with the first record in the database (entry date: February, 2002), we retrieved consecutive records until we obtained data of $N = 1500$ patients, who had started treatment (entry date of last record: January, 2008).

6.1.2 Participants

Patients. Patients were Dutch adults, who were screened through a series of validated web-administered self-report questionnaires and a 30-minute semi-structured diagnostic telephone interview. The clinic did not accept applicants, who a) showed signs of heightened risk of dissociation, psychosis, suicidal ideation, alcohol or drug dependence, b) were recently hospitalized because of mental health problems, c) used neuroleptic medication, d) used unstable doses of other psychoactive medication, or e) suffered from a prevailing disorder for which the clinic could not provide treatment. As a final requirement, the clinic demanded that every patient was seen by a General Practitioner (GP) or another health professional. The screening procedure was open to all, at no costs. However, treatment did not start without a (confirmed) referral source. Referrers received electronic reports at intake, halfway during treatment, and at posttest. Since this was a routine practice service evaluation, study approval was

not obtained from an ethics committee. All patients approved the use of anonymized data through signed informed consent.

Therapists. Therapists were employed by or managed by the clinic. All had a university master's degree in clinical psychology, completed extensive training in CBT, and received additional training in delivering the specific treatment manuals. Most therapists were junior therapists, who were employed by the clinic immediately after their graduation. They were supervised by two licensed clinical psychologists. Psychiatric consultation was available when needed.

6.1.3 Interventions

All patients received web-based therapist-assisted CBT. Depending on presenting problems, patients were assigned to one of four manualized treatments for symptoms of depression, panic disorder, posttraumatic stress, or burnout. The treatment manuals were identical to those tested in previous controlled trials (Lange et al., 2001; Lange, van de Ven, & Schrieken, 2003; Knaevelsrud & Maercker, 2007, 2009; Wagner et al., 2006; Wagner & Maercker, 2007; Ruwaard et al., 2007, 2009, 2010). In these treatments, screening, treatment, and outcome measurement are conducted without any face-to-face contact. With exception of a diagnostic telephone interview, patients and therapists interact through a secure website, in the form of asynchronous text-messages (i.e., their dialogue resembles a structured e-mail conversation rather than a video-conference or an online chat-session). The manuals define fixed sequences of homework assignments that implement common CBT interventions, which are translated into a format suitable for delivery over the Internet. Therapist support consists of standardized, default feedback and instructions that are tailored by the therapist to the specific situation of the patient. In the feedback, motivational techniques are used to enhance the impact of the interventions, i.e., to ensure patients understand the purpose of the interventions, that they set realistic goals, that they do the exercises as prescribed, and that they continue treatment. Standardized duration and hours of therapist input of the treatments varies from 5 weeks and 9.5 hours *Web-CBT for Posttraumatic Stress* to 16 weeks and 19.5 hours *Web-CBT for Depression*.

Web-CBT for Depression is a 16-week treatment that includes symptom awareness training, structuring of daily activities, challenging of dysfunctional thinking patterns, positive self-verbalization, social skills training and relapse prevention (Ruwaard et al., 2009). *Web-CBT for Panic Symptoms* comprises 11 weeks of CBT, and includes symptom awareness training, applied relaxation, *in vitro* exposure, cognitive restructuring, *in vivo exposure* exposure and relapse prevention (Ruwaard et al., 2010). *Web-CBT for Posttraumatic Stress* takes 5 weeks, and consists of structured writing exercises that implement imaginary exposure, cognitive reappraisal and social sharing (Lange et al., 2001). *Web-CBT for Burnout* consists of 16-weeks of online CBT (Ruwaard et al., 2007). It comprises symptom awareness training, progressive relaxation, social skills training, positive self-verbalization (Lange et al., 1998), a rumination intervention, cognitive restructuring, time management training and relapse prevention.

6.1.4 Measures

Patients received automated e-mailed invitations to complete a set of validated self-report questionnaires on the website of the clinic, at pretest, posttest, 6-week follow-up and at 1-year follow-up.

Primary outcomes. Primary outcomes were pre- to post-treatment changes in the intensity of specific psychopathology (i.e., symptoms of depression, panic disorder, posttraumatic stress, or burnout), as measured through validated questionnaires that were different for each treatment manual. *Web-CBT for Depression* included the Beck Depression Inventory (BDI-IA; 22 items; range 0-65, Beck et al., 1988; T. Bouman, 1994), and the Depression subscale of the Depression Anxiety Stress Scales (14 items; range: 0-42; Lovibond & Lovibond, 1995; de Beurs et al., 2001). *Web-CBT for Panic Symptoms* included the Panic Disorder Severity Scale Self-report (PDSS-SR; 7 items; range: 0-45; Houck et al., 2002; Furukawa et al., 2009). *Web-CBT for Posttraumatic Stress* included the Impact of Event Scale (IES; 15 items; range 0-75; Horowitz et al., 1979; Brom & Kleber, 1985). *Web-CBT for Burnout* included the Oldenburg Burnout Inventory (OLBI; 16 items, range of mean global score: 1-4; Demerouti, Bakker, Nachreiner, & Schaufeli, 2001; Demerouti, Bakker, Vardakou, & Kantas, 2003), and the Stress subscale of the DASS (14 items; range 0-42). These primary symptom

measures were identical to those used in the clinical trials. Full descriptions of the questionnaires, which are often used in mental health outcome research, can be found in the reports of these trials (Lange et al., 2001; Lange, van de Ven, & Schrieken, 2003; Knaevelsrud & Maercker, 2007, 2009; Wagner et al., 2006; Wagner & Maercker, 2007; Ruwaard et al., 2007, 2009, 2010). The psychometric characteristics of these questionnaires are satisfactory (PDSS-SR, OLBI) to good (DASS, BDI, IES). In the present sample, Cronbach's alpha's ranged from α = .73 (OLBI) to α = .95 (DASS Depression).

Secondary outcomes. Secondary outcomes were pre-to-post-treatment changes in general psychopathology, and patient satisfaction. General psychopathology was assessed through the total score of the DASS (42 items; range: 0-126), which provides a global measure of negative emotional symptoms (Lovibond & Lovibond, 1995). In the present sample, the internal consistency of this measure was excellent (Cronbach's alpha: α = .97). Patient satisfaction was assessed through a brief evaluation questionnaire, which was administered at posttest. Patients rated the contact with their therapists (on a 1-10 scale), and stated whether they perceived the treatment as effective, whether they had missed the face to face contact, and whether they would recommend the treatment to others (*Yes/No/Don't know*).

6.1.5 Statistical analyses

Statistical significance and effect size We estimated treatment effects through multi-level mixed-model repeated measures regression (MMRM) (Verbeke & Molenberghs, 2000), using the statistical software package R (R Development Core Team, 2008). Time of measurement was coded to contrast mean baseline scores to a) mean scores at the short-term posttreatment assessments (i.e., the post-test and the 6-week follow-up), and b) mean scores at one-year follow-up. Separate analyses were conducted on data from each treatment and each outcome measure, using two-sided tests and Bonferroni corrections to ensure a family-wise significance level of α = .05. We fitted three-level regression models, with repeated measurements nested in patients at level 1, patients nested in therapists at level 2, and therapists at level 3 (Wampold & Serlin, 2000). Conditional intraclass correlations (ICC; Bickel, 2007) revealed that

2% of the variance in outcome was attributable to differences between therapists (mean ICC = .02; range: .001 - 0.13). To express effect sizes as Cohen's d *d* (J. Cohen, 1988), fixed effect regression estimates and associated 95% confidence intervals were divided by pretest standard deviations.

Reliable and clinically significant change Following principles set out by Jacobson and Truax (1991), pre- to post-treatment change scores of each patient were classified as follows: a) deterioration (change was negative and statistically reliable, i.e., it exceeded 1.96 times the standard error of the difference, b) no change (change was not statistically reliable), c) improvement (change was positive and statistically reliable), or d) recovery (change was positive, statistically reliable, and involved a change from a score above clinical cut-off to a score below this cut-off). Analyses were conducted on observed data of the full sample (i.e., $N = 1500$, assuming no change in the patients, who did not complete post-test measurements) as well as the completer sample (i.e., those patients, who completed the full treatment and post-treatment assessment).

6.2 Results

6.2.1 Sample selection

To obtain 1500 records of patients who started treatment, 3003 patient records were retrieved from the database. Figure 6.1 shows that 507 (17%) of the applicants did not start baseline assessment, 843 (28%) withdrew during assessment, 153 (5%) were referred to other mental health institutions, while 50% ($N = 1500$) started treatment. The available data did not allow for a systematic analysis of the reasons of voluntary pre-treatment withdrawal. Of the 1500 accepted patients, 413 (28%) started *Web-CBT for depression*, 139 (9%) started *Web-CBT for panic disorder*, 478 (32%) started *Web-CBT for posttraumatic stress*, and 470 (31%) started *Web-CBT for burnout*. Patients were treated by a total of 135 therapists (depression: $n = 74$; panic disorder: $n = 24$; posttraumatic stress: $n = 65$; burnout: $n = 51$).

6.2.2 Baseline characteristics

Table 6.1 shows baseline characteristics of treated patients ($N = 1500$). The sample comprised 1011 women (67%) and 489 men (33%), who were between 14 and 73 years old ($M = 40$; $SD = 11$; only two patients were younger than 16). Most ($n = 1052$; 71%) did not use psychiatric medication. The vast majority of patients scored above clinical cut-off on the primary outcome measures ($n = 1407$, 94%; range 73%-99% across treatments). Patients were referred by GP's (51%), other specialized mental health organizations (38%), or occupational health officers (11%).

Table 6.1: Baseline characteristics of patients, who followed online CBT for symptoms of burnout, depression, panic disorder, or posttraumatic stress.

	Depression		Panic		PTS		Burnout	
Characteristic[a]	$n = 413$		$n = 139$		$n = 478$		$n = 470$	
Female	280	68%	86	62%	369	77%	276	59%
Age (*M SD*)	40	11	37	11	38	12	41	9
Education								
low (secondary or less)	201	49%	73	54%	284	61%	198	43%
high (tertiary or more)	207	51%	63	46%	182	39%	261	57%
Computer skills (1-10; *M SD*)	7.8	1.9	8.1	1.8	7.6	1.9	6.9	2.5
Clinical symptoms[b]	399	97%	101	73%	458	96%	449	96%
Years with symptoms								
$<$ 1 year	109	26%	29	21%	153	32%	181	39%
1-3 years	142	35%	29	21%	159	34%	211	45%
$\geq$ 3 years	160	39%	79	58%	160	34%	75	16%
Medication								
no medication	237	59%	68	50%	362	76%	385	83%
antidepressant	115	28%	21	15%	53	11%	39	8%
anxiolyticum	23	6%	29	21%	36	8%	31	7%
combination	29	7%	19	14%	24	5%	8	2%

[a] Values represent subsample size and percentage unless otherwise noted. Counts do not add up to 1500 for every characteristic due to missing values (not exceeding 5% on any characteristic).

[b] As measured through the total score of the Oldenburg Burnout Inventory (burnout; cut-off: 2.18), the Beck Depression Inventory-IA (depression; cut-off: 9), the Panic Disorder Symptom Severity Self-Report (panic disorder; cut-off: 8) and the total score of the Impact of Event Scale (posttraumatic stress; cut-off: 24).

6.2.3 Adherence, duration and attrition

Treatment adherence, defined as the percentage of patients completing every step of the treatment program, was 71% (n = 1071). As shown in Figure 6.1, treatment adherence was highest in the burnout sample (77%; n = 364) and the posttraumatic stress sample (75%; n = 361), and lowest in the depression sample (62%) and the panic disorder sample (63%). Patients completed an average of 84% of the treatment protocol, ranging from 81% (*Web-CBT for depression*) to 87% (*Web-CBT for burnout*). Treatments took a median of 22 weeks with depression (IQR = 12), 19 weeks with panic disorder (IQR = 11), 8 weeks with posttraumatic stress (IQR = 6) and 20 weeks with burnout (IQR = 13). Study attrition, defined as the percentage of patients not completing post-treatment measurements, was 21% at post-test (n = 310/1500), 34% at 6-week follow-up (n = 512/1500), and 65% at one-year follow-up (662/1022).

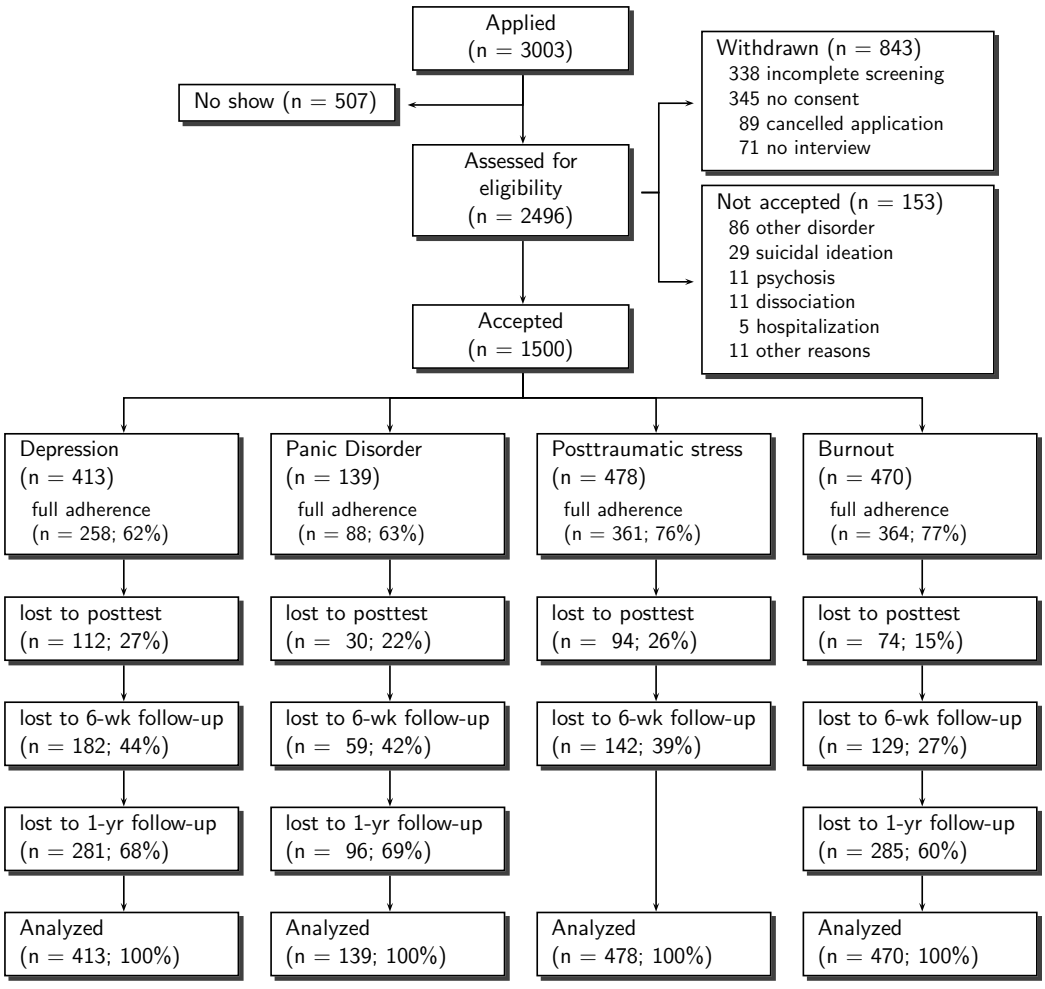


Figure 6.1: Flowchart of study participation.

Table 6.2: Means and standard deviations of measures of specific and general psychopathology, per measurement occasion.

	Measurement occasions											
Treatment	**Pretest**			**Posttest**			**6Wk. FU**			**1Yr. FU**		
measure[a]	*n* [b]	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Specific Psychopathology												
Depression												
BDI	405	24.2	8.0	301	8.4	7.7	249	8.7	7.7	131	9.0	7.6
DASS Depression		22.0	9.1		6.8	8.1		7.5	7.8		7.7	8.1
Panic												
PDSS-SR	136	11.5	5.8	109	4.5	4.2	80	4.8	4.2	44	3.1	4.3
Posttraumatic Stress												
IES Intrusion	477	24.6	7.3	384	13.0	9.3	336	11.7	9.6	[c]		
IES Avoidance		23.4	8.5		12.0	8.9		11.2	10.0			
Burnout												
OLBI Exhaustion	470	2.7	0.6	396	2.3	0.6	341	2.3	0.6	184	2.2	0.7
OLBI Disengagment		3.0	0.4		2.3	0.5		2.3	0.5		2.2	0.6
DASS Stress		19.7	8.8		7.8	7.1		7.9	7.1		7.3	7.9
General Psychopathology (DASS global score)												
	Pretest			**Posttest**			**6Wk. FU**			**1Yr. FU**		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Depression	413	53.6	21.1	301	20.0	19.6	249	21.1	19.5	131	20.4	18.7
Panic Disorder	136	47.3	24.6	108	20.2	22.5	80	20.0	20.6	44	15.2	17.2
Posttraumatic Stress	476	43.5	25.7	381	24.2	24.9	334	23.6	27.1			
Burnout	470	44.2	22.6	399	15.8	16.5	342	16.1	16.0	185	14.8	17.2

[a] BDI: Beck Depression Inventory, version IA (range: 0-63; cut-off: 10); DASS: Depression Anxiety Stress Scales (DASS-42; range: 0-126; cut-off: 30, DASS-Stress: cut-off: 14). PDSS-SR: Panic Disorder Severity Scale - Self report (range: 0-28; cut-off: 8); IES: Impact of Event Scale (Intrusion subscale: range: 0-35; Avoidance Subscale: range 0-40; sumscore cut-off: 24); OLBI: Oldenburg Burnout Inventory (OLBI: range 1-4; cut-off: 2.18).

[b] Pretest data were missing for 8 patients in the depression sample, 3 patients in the panic sample, and 1 patient in the PTS sample.

[c] The assessment protocol of the treatment manual for posttraumatic stress did not include a one-year follow-up.

Table 6.3: Regression analysis and effect sizes of changes in specific and general psychopathology.

Treatment	**Pre vs. Post, 6Wk. FU**				**Pre vs. 1Yr. FU**			
measure[a]	*b*	*SE*	*t*[c]	*d*[d]	*b*	*SE*	*t*	*d*
				Specific Psychopathology				
Depression								
BDI	-15.3	0.46	33.4	1.9±0.1	-14.2	0.82	17.2	1.8±0.2
DASS Depression	-14.8	0.52	28.4	1.6±0.1	-13.9	0.74	18.8	1.5±0.2
Panic								
PDSS-SR	-6.6	0.45	14.6	1.2±0.2	-7.5	0.68	11.1	1.3±0.3
Posttraumatic Stress								
IES Intrusion	-12.0	0.42	28.3	1.6±0.1	[e]			
IES Avoidance	-11.4	0.58	19.6	1.3±0.1				
Burnout								
OLBI Exhaustion	-0.4	0.03	12.7	0.7±0.1	-0.5	0.05	9.5	0.8±0.2
OLBI Disengagment	-0.7	0.03	23.6	1.6±0.1	-0.8	0.04	18.5	1.8±0.2
DASS Stress	-11.9	0.42	28.1	1.3±0.1	-12.1	0.62	19.4	1.4±0.1
				General Psychopathology (DASS global score)				
	Pre vs. Post, 6Wk. FU				**Pre vs. 1Yr. FU**			
	b	*SE*	*t*[c]	*d*[d]	*b*	*SE*	*t*	*d*
Depression	-32.9	1.23	26.8	1.6±0.1	-31.6	1.70	18.2	1.5±0.1
Panic Disorder	-26.1	2.20	11.9	1.3±0.2	-29.7	3.10	9.7	1.2±0.1
Posttraumatic Stress	-18.1	1.20	15.5	0.7±0.1				
Burnout	-28.2	1.05	27.4	1.2±0.1	-28.7	1.47	19.6	1.3±0.1

[a] BDI: Beck Depression Inventory, version IA; PDSS-SR: Panic Disorder Severity Scale - Self report; IES: Impact of Event Scale; OLBI: Oldenburg Burnout Inventory (OLBI); DASS: Depression Anxiety Stress Scales (DASS-42).

[b] b (SE): regression estimate and standard error of raw change score; Negative values of b represent symptom reductions.

[c] t: test statistic. N: depression: n = 413; panic disorder: n = 139; posttraumatic stress: n = 478; burnout: n = 470. Degrees of freedom of the t-test were conservatively set to the number of therapists in each sample (depression: n = 74; panic disorder: n = 24; posttraumatic stress: n = 65; burnout: n = 51). All regression parameters are significant at $P < .001$ after Bonferroni corrections for multiple testing.

[d] d: Cohen's d effect size (95% confidence interval).

[e] The assessment protocol of the treatment manual for posttraumatic stress did not include a one-year follow-up.

6.2.4 Statistical significance and effect size

Statistical significance. Table 6.2 shows means and standard deviations of patients' questionnaire scores at each measurement occasion. Table 6.3 shows results of the regression analyses of pre-to-post-treatment changes in symptom severity. Each measure of specific and general psychopathology indicated significant ($P < .001$) reductions in symptom severity, on the short term (at posttest and 6-week follow-up), as well as on the long term (at one-year follow-up).

Effect size. Point estimates and 95% confidence intervals of effect sizes are shown in Table 6.3. With regard to the primary outcome measures of specific symptom severity, short-term improvements represented a large pooled standardized effect size of $d = 1.4$ (range: $0.7 \leq d \leq 1.9$). One year after treatment, these effect sizes were found to be sustained. With regard to general psychopathology (as measured by the DASS total score), short-term improvements represented a large pooled effect size of $d = 1.2$ (range: $0.7 \leq d \leq 1.6$), and these effect sizes also sustained on the long term.

6.2.5 Reliable and clinically significant change

Results of the clinical significance analyses are shown in Table 6.4. On the short-term (at post-test and 6-weeks follow-up), reliable improvement was 72% among treatment completers ($n = 1046$, 71%), and 55% in the full intent-to-treat sample ($N = 1500$, conservatively assuming no change where scores were unavailable). Recovery (i.e., reliable improvement from a pretest score above cut-off to a post-treatment score below cut-off) was 51% in the completers analysis and 40% in the intent-to-treat analysis. Reliable deterioration was 2% or less. Available data of the 1-year follow-up ($n = 358$), which were almost exclusively provided by treatment completers (99%), revealed an average reliable improvement rate of 78% and a recovery rate of 59% (c.f. Table 6.5).

Table 6.4: Clinical significance of short-term changes in primary psychopathology. [a]

Treatment			Recovered		Improved		No change		Deteriorated	
measure[b]	*Sample*[c]	*n*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
Depression										
BDI	All	413	177	44%	68	17%	165	39%	3	1%
	Completer	258	158	61%	58	22%	41	16%	1	0%
Dass	All	413	189	46%	37	9%	183	44%	4	1%
Depression	Completer	258	170	66%	29	11%	55	21%	4	2%
Panic										
PDSS-SR	All	139	52	37%	14	10%	73	53%	0	0%
	Completer	87	42	48%	11	13%	34	39%	0	0%
Posttraumatic stress										
IES	All	478	190	40%	94	20%	193	40%	1	0%
	Completer	358	181	51%	93	26%	83	23%	1	0%
Burnout										
OLBI	All	470	134	29%	111	24%	220	47%	5	1%
	Completer	364	122	34%	104	29%	133	37%	5	1%
Dass Stress	All	470	205	44%	58	12%	204	43%	3	1%
	Completer	364	189	52%	53	15%	119	33%	3	1%

[a]Analyses of short-term changes were based on individual difference scores on the primary outcome measures, calculated as the pretest score minus the mean of the post-test and 6-week follow-up score.

[b]BDI: Beck Depression Inventory, version IA (cut-off: 10; reliable change: 7 scale points); DASS: Depression Anxiety Stress Scales (DASS Depression cut-off: 12; reliable change:5 scale points; DASS Stress cut-off: 14; reliable change: 7 scale points); IES: the Impact of Event Scale (cut-off: 24; reliable change: 12 scale points); OLBI: Oldenburg Burnout Inventory (cut-off: 2.18; reliable change: .37 scale points). PDSS-SR: Panic Disorder Severity Scale - Self report (cut-off: 8; reliable change: 5 scale points).

[c]All: all patients (assuming no change where data was missing); Completer: subsample of patients, who completed the full treatment and at least one post-treatment assessment.

Table 6.5: Clinical significance of changes in primary psychopathology, at one-year follow-up.

Treatment		Recovered		Improved		No change		Deteriorated	
measure[a]	*n*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
Depression									
BDI	131	73	56%	33	25%	25	19%	0	0%
Dass Depression		80	70%	19	17%	14	12%	2	2%
Panic									
PDSS-SR	43	28	65%	6	14%	9	21%	0	0%
Burnout									
OLBI	184	72	42%	52	30%	40	23%	8	5%
Dass Stress		95	52%	31	17%	54	29%	4	2%

[a]BDI: Beck Depression Inventory, version IA (cut-off: 10; reliable change: 7 scale points); DASS: Depression Anxiety Stress Scales (DASS Depression cut-off: 12; reliable change:5 scale points; DASS Stress cut-off: 14; reliable change: 7 scale points); OLBI: Oldenburg Burnout Inventory (cut-off: 2.18; reliable change: .37 scale points). PDSS-SR: Panic Disorder Severity Scale - Self report (cut-off: 8; reliable change: 5 scale points).

Table 6.6: Patient satisfaction.

	Depression		Panic		PTS		Burnout	
Aspect[a]	$n = 296$		$n = 108$		$n = 380$		$n = 323$	
Satisfaction with therapist (*M SD*)	8.4	1.5	8.7	1.6	8.6	1.5	8.6	1.2
Do you consider online therapy an effective method?								
Yes	240	81%	90	83%	298	78%	289	89%
No	30	10%	7	6%	29	8%	6	2%
Don't know	26	9%	11	10%	53	14%	28	9%
Did you miss face-to-face contact?								
Yes	35	33%	35	32%	110	29%	87	26%
No	163	55%	64	59%	231	61%	189	59%
Don't know	35	12%	9	8%	39	10%	47	47%
Would you recommend online treatment to others?								
Yes	260	88%	95	88%	339	89%	292	90%
No	18	6%	3	3%	18	5%	12	4%
Don't know	18	6%	10	9%	23	6%	19	6%

[a]Depression: $n = 296$; Panic Disorder: $n = 108$; Posttraumatic Stress: $n = 380$; Burnout: $n = 323$.

6.2.6 Patient satisfaction

Posttest evaluation data, which were available for $n = 1107$ patients, are summarized in Table 6.6. Patient satisfaction was high, with little variance between treatment samples. Patients gave high ratings to their therapists ($M = 8.5$ on a 1-10 scale; $SD = 1.5$). Although 30% ($n = 330$) of the patients indicated that they had missed face-to-face contact during therapy, 83% evaluated online therapy as effective, and 89% would recommend web-based treatment to others.

6.3 Discussion

6.3.1 Key findings

We assessed the effectiveness of online CBT in routine clinical practice among 1500 patients, who suffered from symptoms of depression, panic disorder, posttraumatic stress, or burnout. We found that effect sizes and recovery rates were comparable to, or somewhat better than, those observed in previous controlled trials (Lange et al.,

2001; Lange, van de Ven, & Schrieken, 2003; Knaevelsrud & Maercker, 2007, 2009; Wagner et al., 2006; Wagner & Maercker, 2007; Ruwaard et al., 2007, 2009, 2010), and comparable to those of face-to-face routine practice CBT (Westen & Morrison, 2001; Cahill et al., 2003; Hunsley & Lee, 2007; Minami et al., 2008). Post-treatment reductions of specific and general psychopathology were large (pooled $d = 1.4/1.2$; $P < .001$), about 50% of the patients recovered, and patient satisfaction was high. Our findings suggest that online CBT may be as effective in routine practice as it is in clinical trials.

While high dropout is a common problem in studies of online interventions, (Eysenbach, 2005; Farvolden et al., 2005; de Graaf, Huibers, Riper, Gerhards, & Arntz, 2009), our data show that acceptable adherence can be achieved with online treatment. We found that 79% of the patients completed post-treatment measurements, and that 71% completed every step of treatment. Although adherence rates were somewhat better in the controlled trials (83%), these findings still compare well to published adherence rates of Dutch routine practice mental healthcare, which tend to vary between 60% to 70% (GGZ Nederland, 2010).

As previously shown by Titov, Andrews, Kemp, and Robinson (2010), patients who seek online treatment have substantial disorders, and are not necessarily young or technologically sophisticated. Our data confirm and extend this finding. In terms of demographic characteristics and presenting problems, the patients of the online clinic are comparable to the patients that are generally seen in Dutch specialized mental healthcare. Our results indicate that online treatment provides an appropriate intervention for these patients.

6.3.2 Strengths & limitations

This study has several strengths. First, the treatments had been carried out before the study was planned. Treatment outcomes were routinely assessed as part of every-day practice. Accordingly, therapists and staff were not influenced by their participation in the evaluation study (the so-called Hawthorne effect; Leonard & Masatu, 2006). Second, the size of the sample is much larger than the sample sizes of previous effectiveness studies of therapist-assisted online CBT. Third, the sample included every patient who had started treatment in the studied timeframe. Thus, our results

could not be affected by a selection of well-responding patients. Fourth, patients were treated by a large number of relatively inexperienced therapists. Since we found little variance between the therapists in terms of treatment outcome, we concur with Wilson, who argued that "the capacity to train a diverse group of therapists to a criterion level of competence so that they can reliably administer a treatment protocol [...]] can be seen as a significant advance in the dissemination of effective treatment" (Wilson, 1998). Fifth, the therapeutic procedures and outcome measures in the study were identical to those used in the controlled trials. In both contexts, the same computerized treatment manuals were used. Hence, treatment integrity was guaranteed in both contexts. Sixth, there was no face-to-face contact at all. This considerably enhances the flexibility of the online treatment, since it provides the possibility to treat patients who live at distant locations. Paradoxically, this positive aspect could also be seen as a potential weakness, as will become clear in our subsequent discussion of the limitations of the study.

A first limitations of our study is that patient screening was conducted online and by telephone, without a structured clinical interview. The clinic made use of dimensional screening through validated self-report instruments for which norm tables have been established. This did not allow for formal DSM-IV diagnoses, but compared well with the DSM-IV categories by using the cut-off scores of the scales. Since there are indications that telephone interviews and face-to-face interviews yield comparable results (Cacciola, Alterman, Rutherford, McKay, & May, 1999), the clinic at present makes use of structured diagnostic interviews in telephone contacts during the screening, in addition to the dimensional screening. Second, we recall that a high percentage of the patients did not complete the long-term follow-up measures (65%). Although available data suggest that effects maintain up to one year after treatment, these results should be interpreted with caution. It would have been useful if we would have sent short questionnaires to non-responders, to assess whether there were any differences between responders and non-responders. Third, a considerable percentage of the applicants (40%) withdrew, during or even prior to the screening. While pre-treatment withdrawal rates are often not reported, research shows that high withdrawal rates are common, in both online and offline research (Issakidis & Andrews, 2004; Melville, Casey, & Kavanagh, 2010). It is a challenge to improve these

figures, but our data do not permit valid assessment of the reasons for and effects of pre-treatment withdrawal. The withdrawal might be caused by the ease with which one can apply for online therapy. This may result in impulsive applications (we found some indications that applicants, who started the screening without a referral, were less likely to start treatment). A second possibility is that patients withdraw from online treatment because they are unwilling to relinquish their anonymity (Lange & Ruwaard, 2010).

6.3.3 Clinical utility

The protocols that are used in this routine practice comprise a strong mix of cognitive behavioral procedures and techniques that enhance patient motivation and the patient-therapist relation. One particular advantage of the therapy format is that therapists do not have to respond immediately when patients post their homework. Due to the asynchronous communication, therapists have time to reflect on the best possible feedback and the best explanation of new homework assignments. If needed, they may even discuss a case with colleagues. In addition, the Interapy protocols provide help files, which the therapists may consult to find motivating phrases for their feedback. All these aspects make this a highly usable treatment for clinical practice.

The study shows how computer-mediated treatment can save time by freeing therapists of repetitive tasks such as the administration and scoring of outcome questionnaires. In a computerized environment, these tasks can be easily automated.

Online treatment is highly dependent on regulatory approval, professional codes of ethics and jurisdiction regulations, which vary considerably from country to country. But even when such barriers have been overcome, financial hurdles may still exist. When the Interapy clinic was founded, in 2001, costs of face-to-face treatment were fully reimbursed by Dutch public health insurance, while costs of online treatment were not. In effect, access to online treatment was limited due to a financial barrier. This changed in 2005, when health regulatory bodies recognized the online services of the clinic as reimbursable healthcare, under condition of a GP-referral for psychotherapy. Without this recognition, further implementation of online CBT would not have been feasible.

At present, cost-benefit comparisons between face-to-face treatment and online treatment are complicated by the limited availability of routine practice data and by unclear definitions of relevant health-economic variables. Nonetheless, scant research has shown that online treatment is cost-effective (Bergström et al., 2010; Tate, Finkelstein, Khavjou, & Gustafson, 2009), and it would be instructive to conduct a cost-effectiveness study for the present treatment as well. One important variable in this respect is the amount of therapist input that is required for a meaningful clinical response. Less input implies fewer costs, but may also limit the effects of the intervention (Palmqvist et al., 2007; Donker et al., 2009).

6.3.4 Conclusion

In sum, our study suggests strongly that RCT findings of online therapist-assisted CBT generalize well to routine clinical practice as carried out in the Interapy clinic. It would be interesting to examine whether results with this form of online CBT are similarly positive in other clinical contexts. While pre-treatment withdrawal and long-term outcome demand attention, our results suggest that online CBT provides a viable, effective, and acceptable treatment alternative for patients, who are unwilling or unable to seek traditional forms of mental healthcare. Future studies should confirm this suggestion through direct comparisons of online CBT and regular treatment options, preferably in the form of large-sample equivalence trials that are conducted in naturalistic settings. Meanwhile, based on our present results, we recommend further implementation of online CBT.

Chapter 7
General Discussion

Previous controlled trials established that Interapy, a particular form of online Cognitive Behaviour Therapy (CBT), yielded promising results in the treatment of post-traumatic stress. This raises the question whether similar positive results may be achieved with other mental health disorders, and whether online CBT performs well in routine clinical practice. In the preceding chapters, we presented several studies that were designed to examine the efficacy and effectiveness of online CBT. In this final chapter, we reflect on the aggregated evidence. What are our key findings? What are the strengths and limitations of the present studies? How do these findings inform clinical practice and future research initiatives? And finally, what general conclusions can be drawn with respect to the efficacy and effectiveness of online CBT?

7.1 Key Findings

7.1.1 The efficacy of online CBT

To explore the wider applicability of online CBT, we developed four new protocols for the online treatment of burnout, depression, panic disorder and bulimia nervosa, and we assessed the efficacy of each protocol in a randomised controlled trial (RCT; Ruwaard et al., 2007, 2009, 2010; Ruwaard, Lange, Broeksteeg, et al., 2012; Chapters 2-5). The characteristics of these trials are listed in Table 7.1. The RCT's included a total of $N = 456$ participants, who were randomly allocated to online CBT or a waiting-list/delayed treatment condition (in the RCT of the Bulimia Nervosa trial, bibliotherapy served as an additional experimental control). Outcome was assessed through well-validated self-report measures of symptom severity, which were administered at baseline (pre-test), immediately after treatment (post-test), and one to three years after treatment (long-term follow-up). Outcomes were analysed, on intention-to-treat basis, in terms of mean change in symptom severity over time, and in terms of clinically significant change (i.e., reliable recovery; Jacobson & Truax, 1991).

Dropout rates were encouragingly low: an (unweighted) average of 82% of the patients completed every step of treatment. Reductions in primary symptom severity were significantly larger with Interapy, in comparison to the experimental controls, as illustrated by the left forest plot in Figure 7.1(a). The standardised mean difference (d) in improvement between Interapy and the experimental controls ranged from $d = .7$ to $d = 1.0$, with a pooled effect size of $d = .8$ (95% CI: .6 to 1.0). These are large effects, and roughly equivalent to those of face-to-face CBT. As illustrated in Figure 7.1(b), our studies also revealed higher recovery rates with Interapy compared to those obtained with experimental controls. Across the trials, the average rate of reliable recovery was 47% in the online treatment groups, compared to 20% in the control groups, which equates to a significant odds ratio (OR) of 4.1 (95% CI: 2.4 to 7.0). One to three years after treatment, we found that treatment gains persisted.

7.1.2 The effectiveness of online CBT

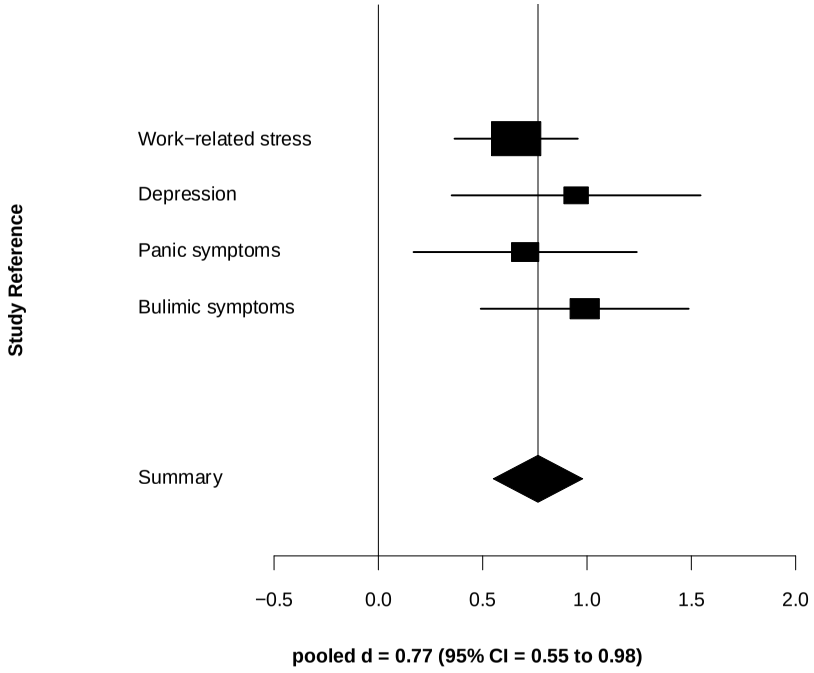
To assess the effectiveness of online cognitive behavioural treatment in routine clinical practice, we examined treatment outcome of $N = 1500$ patients of the Interapy clinic (Ruwaard, Lange, Schrieken, Dolan, & Emmelkamp, 2012; chapter 6). This was an uncontrolled, pre-test/post-test study, with two follow-ups. Data were collected from unselected, consecutive electronic patient records of the clinic, which contained scores of self-report questionnaires that had been routinely administered at pre-test, post-test, and at 6-weeks and 1-year follow-up. Patients were Dutch adults with a referral for psychotherapy, who started online CBT for depression ($n = 413$), post-traumatic stress ($n = 478$), panic disorder ($n = 139$), or burnout ($n = 470$). Outcome variables were treatment adherence, primary and secondary symptom severity, and recovery rates.

Due to routine outcome measurement, post-treatment data were available for 79% of the patients. Full treatment was completed by 71%, which is comparable to reported completion rates in Dutch mental healthcare (GGZ Nederland, 2010). Symptom reductions met selected benchmarks of naturalistic studies of face-to-face CBT, and were comparable to those observed in the RCT's, as shown in Figure 7.2. On the short-term (at post-test and six weeks follow-up), clients reported significant ($P < .001$) reductions in symptom severity, which represented a large pooled (uncontrolled) effect size of $d = 1.4$ (cf. Figure 7.3). Among treatment completers, 71% reliably improved and 52% experienced a clinically significant change (i.e., recovery). Follow-up measurements were difficult to interpret given an attrition rate of 67%. Nonetheless, available data suggest that improvements sustained up to one year after treatment. As in the clinical trials, routine practice clients expressed high satisfaction with the treatment. Patients gave high ratings to their therapists ($M = 8.5$ on a 1-10 scale; $SD = 1.5$). Although 30% of the patients indicated that they had missed face-to-face contact during therapy, 83% evaluated online therapy as effective, and 89% would recommend online treatment to others.

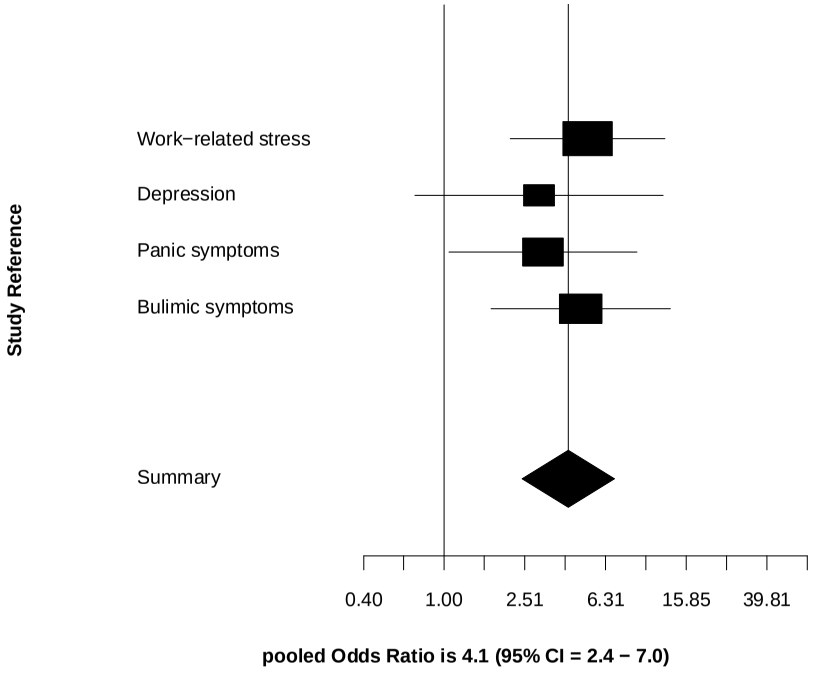
Table 7.1: Characteristics of the studies in this dissertation.

Study ID	Target problem	Participants	Type	Experimental conditions	N	Drop-out	Key Outcome Measures	Longest Follow-up
Ruwaard 2007	Work-related stress	Adult Community	RCT	Online CBT Waiting list	177 62	28%	Stress (DASS); Exhaustion (MBI); Depression (DASS); Anxiety (DASS)	3 Yrs.
Ruwaard 2009	Depression	Adult Community, BDI < 30	RCT	Online CBT Waiting list	36 18	8%	Depression (BDI, SCL-90); General psychopathology (DASS)	18 Mos.
Ruwaard 2010	Panic symptoms	Adult Community	RCT	Online CBT Waiting list	27 31	11%	Panic symptoms(PDSS-SR); Attack frequency & intensity (Panic diary)	3 Yrs.
Ruwaard 2012a	Bulimic symptoms	Adult Community	RCT	Online CBT Waiting list Bibliotherapy	35 35 35	26%	Binging and purging frequency & global eating disorder severity (EDE-Q); Body attitude (BAT)	1 Yr.
Ruwaard 2012b	Burnout Depression Panic symptoms Posttraumatic stress	Adult, Referred	Routine practice study	Online CBT	1500	29%	Specific psychopathology (OLBI/DASS Stress, BDI, PDSS-SR, IES); General psychopathology (DASS)	1 Yr.

BAT: Body Attitude Test; BDI: Beck Depression Inventory (Beck et al., 1979); DASS: Depression Anxiety Stress Scales (Lovibond & Lovibond, 1995); EDE-Q: Eating Disorder Examination-Questionnaire (Fairburn & Beglin, 1994); IES: Impact of Event Scale (Horowitz et al., 1979); MBI: Maslach Burnout Inventory (Maslach et al., 1996); OLBI: Oldenburgh Burnout Inventory (Demerouti et al., 2001); PDSS-SR: Panic Disorder Severity Scale, Self-rate version (Houck et al., 2002); Ruwaard 2012a: Ruwaard, Lange, Broeksteeg, et al., 2012; Ruwaard 2012b: Ruwaard, Lange, Schrieken, Dolan & Emmelkamp, 2012



(a) Cohen's d



(b) Odds Ratio

Figure 7.1: Forest plots of between-group effect sizes in the controlled trials of Interapy online CBT for work-related stress, depression, panic symptoms and bulimic symptoms, with regard to a) the standardised difference (Cohen's d) between the experimental groups in primary symptom severity improvements, and b) the relative odds (odds ratio) of reliable recovery. Cohen's d is considered small when $d = .2$, medium when $d = .5$ and large when $d = .8$. Likewise, the odds ratio is small when OR = 1.5, medium when OR = 3.5 and large when OR = 9.0

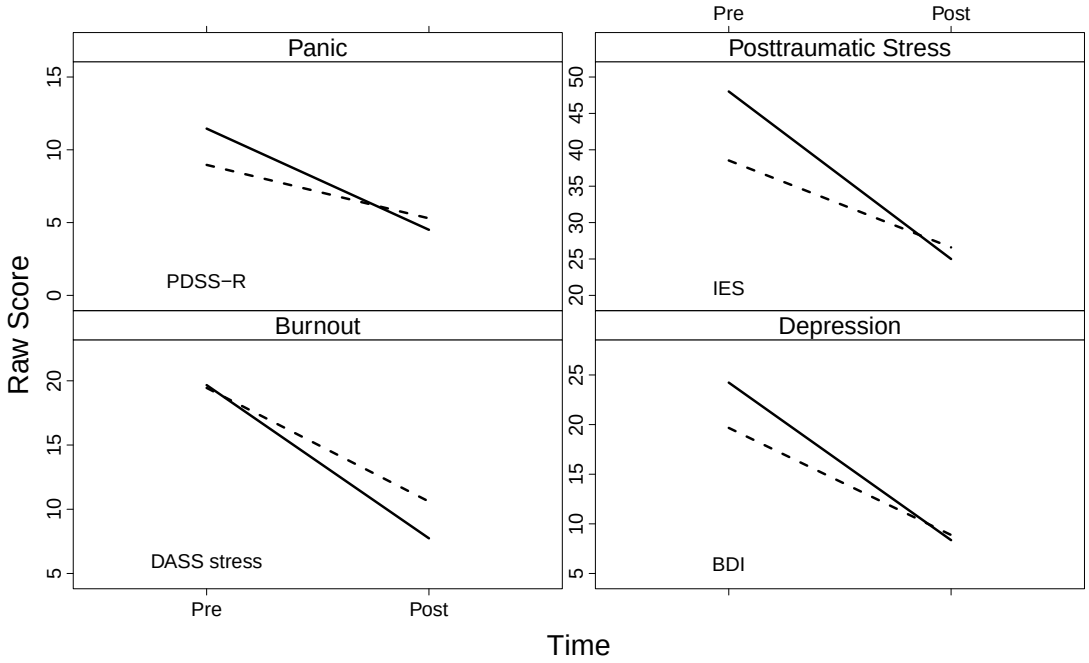


Figure 7.2: A comparison of RCT and routine practice data of pre- to post-treatment changes in primary symptom severity observed in clients, who followed online CBT. Solid lines represent mean scores of routine practice clients (Ruwaard, Lange, Schrieken, et al., 2012). Dashed lines represent mean scores of RCT participants, who completed online CBT (Ruwaard et al., 2007, 2009, 2010; Ruwaard, Lange, Broeksteeg, et al., 2012; Lange, Rietdijk, et al., 2003). Overall, treatment effects in routine practice were somewhat stronger than those in the controlled trials, because routine practice clients had more to gain from treatment. At pre-test, symptom severity levels of routine practice clients were higher in comparison to those of RCT participants.

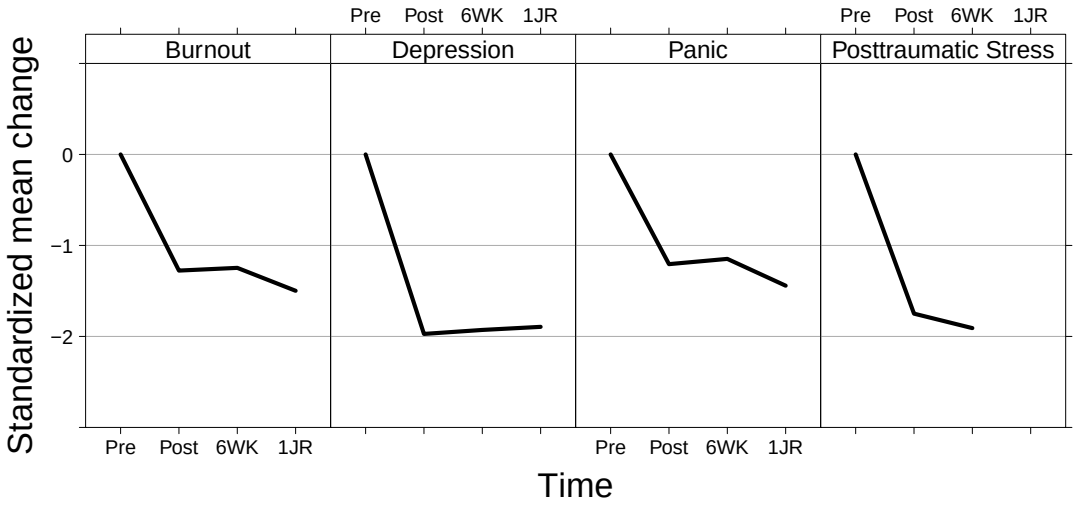


Figure 7.3: Standardised changes in primary symptom severity in the routine practice dataset, from pre-test to one year follow-up.

7.2 Strengths and limitations

7.2.1 Strengths

Therapist involvement. Our studies complement a field of research that is largely focused on the reduction of therapist involvement. The vast majority of studies of internet-delivered CBT have evaluated self-help or guided self-help formats, in which therapist support is either not provided or substantially reduced. The focus on self-management is understandable, given concerns that available resources are inadequate to meet the rising demands for mental healthcare. However, reduced therapist involvement may also limit the effects of an intervention (Sharp et al., 2000; Palmqvist et al., 2007; Cuijpers et al., 2009). We examined a form of online treatment that matches manualised face-to-face CBT, in which therapist involvement is considered crucial and therefore more extensive. By doing so, we addressed a neglected area in internet intervention research.

Large routine practice sample. Until recently, little was known about the performance of online therapist-assisted CBT in routine clinical practice (Andersson et al., 2009). By demonstrating that the results of controlled trials of online therapist-supported CBT generalise well to clinical practice, we substantially added to the existing evidence base. To our knowledge, our study resulted in one of largest routine clinical samples in therapist-assisted online CBT research. Encouragingly, similar findings with therapist-assisted online CBT are being reported from other online virtual clinics around the world (e.g., Hilvert-Bruce, Rossouw, Wong, Sunderland, & Andrews, 2012; Sunderland, Wong, Hilvert-Bruce, & Andrews, 2012; Bergström et al., 2010). Policymakers, regulatory bodies, and insurance companies may find an argument in these results to support further implementation of online CBT.

Acceptable adherence. Poor adherence has been identified as one of the major challenges of online interventions. Many participants do not complete full treatment. The problem is most pronounced for self-help interventions. For instance, dropout was huge in the Farvolden et al. (2005) study of a program without therapist involvement, with as few as 12 out of 1161 participants (1%) completing the program. But guided

self-help interventions are also subject to considerable dropout. In the typical trial of internet-based guided self-help, the majority of participants do not complete the full program. Eysenbach (2005) went as far to suggest that high attrition should be considered an inherent aspect of all internet-based interventions. Our data challenge this suggestion. Treatment adherence rate was 83% in the clinical trials, and 71% in routine practice. Both figures clearly demonstrate that acceptable adherence can be achieved in internet-based treatment, which is an encouraging finding. In the absence of a formal component analysis, no definite explanation for the better adherence can be given. The protocols that we studied comprise a strong mix of cognitive behavioural procedures and techniques that enhance client motivation and the therapeutic relationship (see, for example, Lange, 2006; van der Velden et al., 2010). Future studies should cast further light on the interaction of these three factors.

Strong effects. Effects were evaluated against conservative criteria of outcome. We corrected for multiple testing to ensure a family-wise error rate of $P < .05$, we corrected for missing values using the last-observation-carried-forward intention-to-treat technique (which is increasingly considered to be over-conservative), and we applied the robust criteria of clinically significant change. Yet, despite these conservative criteria, the effects of online CBT on primary symptoms and general psychopathology were significant and large across the trials.

Our data are in line with the hypothesis that the effects of human-supported interventions are larger than those of unsupported interventions. The pooled controlled effect size that we found in our trials is considerably higher than the effect size of unsupported internet interventions as found in previous meta-analyses (Cuijpers et al., 2009; Andersson & Cuijpers, 2009; Spek et al., 2007).

In comparison to other human-supported online interventions (e.g., Andersson et al., 2007; Andrews et al., 2010), the present interventions appear to have similar effects. For instance, between-group effects in our depression trial were comparable to those found in a study of internet-based guided self-help by Andersson et al. (2005). This suggests that the effectiveness of internet-based interventions may be less dependent of the frequency of therapist support, as was previously shown in a Australian trial (Klein et al., 2009). This result is puzzling, since one would expect that increased adherence would result in better effects. It may not be enough to

simply compare standardized outcomes of separate studies. Differences in research methodology and sample characteristics may obscure subtle but important differences in outcome. It would be interesting to directly compare the current interventions with variants with less frequent support (e.g., Donker et al., 2009). Such comparisons should include the careful examination of the relationship between adherence and effect size.

7.2.2 Limitations

No DSM-diagnoses. Throughout our studies, we relied on symptom severity ratings obtained through validated self-report questionnaires rather than DSM-IV diagnoses. Formal diagnoses would have required in-person diagnostic interviews, which were incompatible with the goals of our research (i.e., to assess the effects of online CBT with no face-to-face contact). In effect, some included participants probably did not fully meet the diagnostic criteria for the targeted disorders. This might be considered a limitation. However, our dimensional approach also increased the external validity of the trial, since patients often present with subclinical symptoms in clinical practice. In addition, our data indicate that we included a rather severe group of participants, who suffered from chronic symptoms. A large percentage of the participants had been treated before. Face-to-face diagnostics interviews limit conclusions about the effects of online treatment per se, since it cannot be ruled out that the effects are confounded by the additional face-to-face contact. They also limit the generalisability of the results, since the required face-to-face contacts may deter those who tend to avoid treatment, because they are reticent about such contacts. In addition, these diagnostic interviews create a geographical limitation, which precludes the possibility to treat clients who live at distant locations. Telephonic diagnostic interviews may provide a future solution. There are indications that telephone interviews and face-to-face interviews yield comparable diagnostic conclusions (Cacciola et al., 1999; Aziz & Kenford, 2004; Crippa et al., 2008).

Asymmetric post-test assessments. The internal validity of the first two RCT's suffered from asymmetric timing of symptom assessment between the experimental groups. We expected participants to complete treatment within the planned weeks,

but the average treatment took longer. Consequently, measurements in the experimental groups were taken some time after the measurements in the control condition. This resulted in interpretative difficulties, since the between-effects were possibly confounded by (uncontrolled) effects of the mere passage of time. In the study of online CBT for work-related stress (Chapter 2), post-hoc analyses failed to reveal a significant effect of treatment duration. This implied that the confounding effects were probably small. In the depression study (Chapter 3), corrections were necessary, since the participants in the control group reported strong reductions in depressive symptoms from pre- to post-test. To correct for this, we extrapolated the post-test scores of the control group, which resulted in more conservative estimates of the between-group effects. Although the post-hoc analyses and statistical corrections provided some resolution, it would of course had been better if measurements had been simultaneously taken in both groups. Fortunately, we were able to avoid asymmetric timing of assessments in the other RCT's.

Uncontrolled follow-ups. The present evidence with regard to the long-term outcome of online CBT is encouraging, but needs to be corroborated in future research. Attrition at follow-up was 19% to 37% in the RCT's, and 65% in the routine practice trial. Clients, who experienced a more favourable long-term outcome, may have been more inclined to respond to the follow-up invitations. An additional limitation of the follow-up is that it was uncontrolled, since participants in the control condition received treatment after post-test. During the period between the post-test and the follow-up, participants were subject to many influences which we could not control. Many clients reported chronic and recurrent depressive symptoms before treatment, and stable improvements up to three years after treatment. Apparently, treatment had made a lasting difference. However, these results should be interpreted with caution.

7.3 Implications for clinical practice

Suitability. Our findings suggest that online CBT is a good alternative to face-to-face CBT for adults presenting with mild to moderate-severe forms of depression, panic disorder, bulimia nervosa, post-traumatic stress, or burnout. Online CBT might be offered to clients who prefer more anonymous forms of treatment, to clients with

busy or irregular life-styles, to expats, to clients who are physically unable to attend regular face-to-face CBT, and to clients who are reluctant to follow regular mental healthcare. Clients should have basic reading and writing skills, and need to be reasonably proficient in using the computer and the internet. Education level and age are not important. Suicidal ideation, risk of psychosis and dissociation, previous hospitalisation, and alcohol- and drug abuse are risk factors that need to be assessed at intake. Concurrent medication usage is an option, when usage is stable. Residual symptoms at post-test should be checked and addressed.

Accessibility. Online CBT is a highly accessible treatment. In theory, it has no geographical limitation. This is perhaps best illustrated by a recent trial, in which traumatised victims of war violence in Iraq received online CBT from German therapists, who operated from Berlin (Wagner et al., 2012). In practice, however, its availability is highly dependent on local codes of conduct. Online interventions do not fit traditional healthcare systems, and raise legal, ethical, and professional issues that are only partially resolved by current guidelines. In addition, financial barriers may limit the implementation of online treatment. In 2001, the Interapy initiative was well-received. Clients, referrers, and policymakers welcomed the new online treatment option. However, while costs of face-to-face treatment were fully reimbursed by Dutch public health insurance, costs of online treatment were not. Fortunately, this changed in 2005, when health regulatory bodies officially recognised the online services of the clinic as reimbursable healthcare.

Efficiency. There are several ways in which online CBT saves time. First, online CBT saves time related to travelling and the scheduling of therapy sessions. Second, online CBT saves time by freeing therapists of repetitive tasks such as the administration and scoring of outcome questionnaires. In a computerised environment, these tasks can be easily automated. Third, it is important to recognise that manualised CBT is still practiced by only a minority of therapists in clinical practice (Shafran et al., 2009). Although total time invested in online CBT may be comparable to the time that is invested in manualised face-to-face CBT, it may save time by providing an attractive alternative to the less structured (and presumably less efficient) treatments that are generally provided in routine practice. Fourth, online CBT saves time because it is

associated with brief treatment duration. In the routine practice study, duration of treatment ranged from 8 to 22 weeks, which is considerably less than the average duration of treatment in Dutch specialised mental healthcare (1 year; Hilderink & Van 't Land, 2009).

Practicability. Online CBT provides an excellent tool to maintain treatment fidelity, since its treatment manual is computerised and highly standardised. It has been argued that treatment manuals enable therapists to administer treatments more effectively, reduce idiosyncrasy in therapeutic methods, increase the therapists' focus on specific treatment goals and techniques, and reduce differences in the quality of treatment delivery (Crits-Christoph et al., 1991). Our data underscore these suggestions. In our routine practice data, only 2% of the variance in outcome was attributable to differences between therapists. As illustrated in Figure 7.4, therapists generally achieved similar positive outcomes.

One particular advantage of online CBT is that therapists have time to reflect on the best possible feedback, due to the asynchronous nature of the therapeutic dialogue. Before sending their feedback, they can consult the help files that are provided by the treatment manual to find motivating phrases, discuss a case with colleagues, or call for supervision, if needed. As a result, the quality of feedback can be increased. In our trials, the treatments were provided by therapists who were relatively young and inexperienced. Yet, the outcomes of their treatments were on a par with the outcomes of treatments of experienced therapists. In our view, the extra time provided by the asynchronous therapeutic dialogue played a key part in this result.

7.4 Suggestions for future research

Replication studies. Although the efficacy of the new interventions was demonstrated, these findings should be confirmed in replication research. Each of these interventions was tested in just one RCT. It might be argued that the aggregated evidence of the five related studies strengthen the evidence of each individual intervention. However, according to the criteria of empirically validated treatment (e.g., Chambless, 1993), the new interventions should be considered experimental until at least one other group design study has been completed. Ideally, these replications will

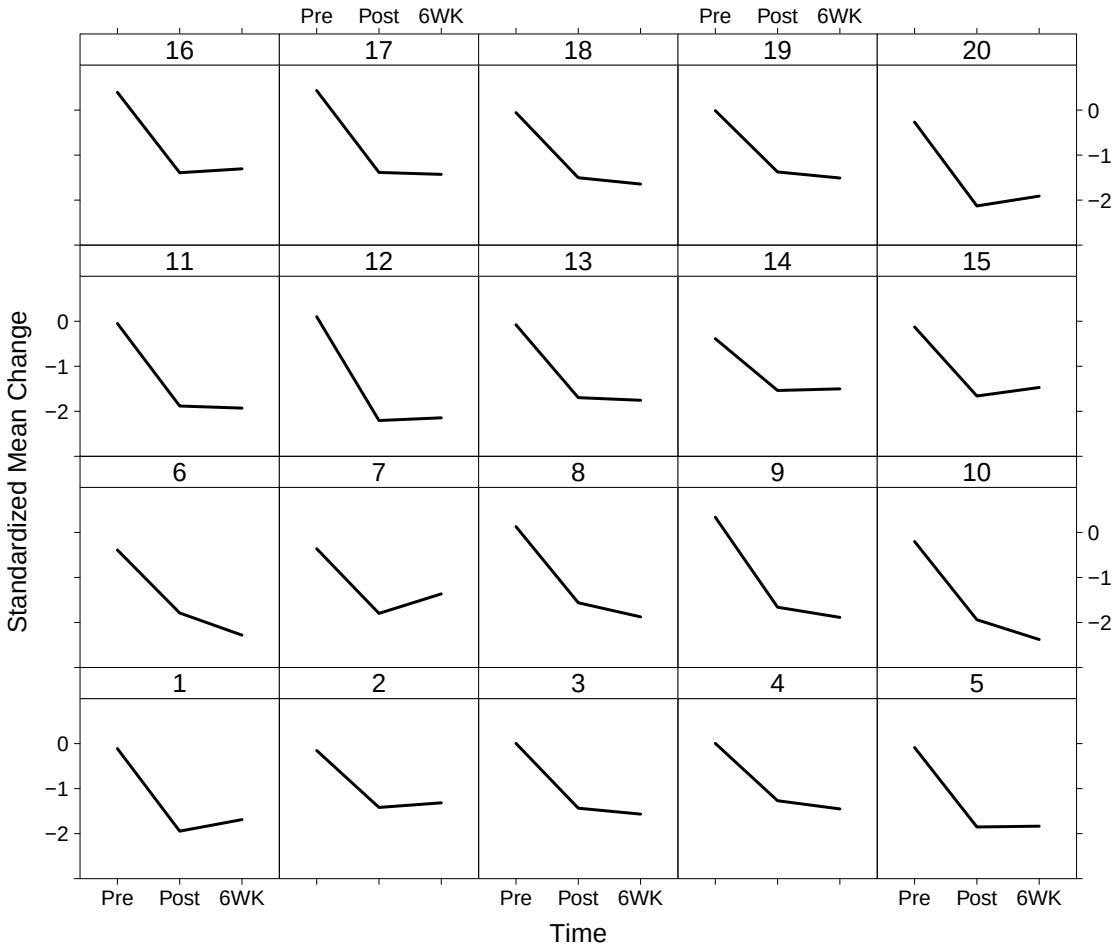


Figure 7.4: Mean pre-test to six week follow-up changes in primary symptom severity (standardised scores) as observed for clients of the twenty most active therapists in the routine practice dataset (number of clients per therapist ranges from 21 to 151).

be conducted by independent research groups. In our studies, we evaluated the treatments that we also developed. Although this is almost common practice in internet intervention research, independent replications would significantly strengthen the present evidence base.

Comparative trials. In the clinical trials, the treatments have been evaluated exclusively by comparison with waiting-list comparison groups. Any reference to the equivalence of online CBT and face-to-face therapy is based on comparisons of effect sizes that were collected in different contexts. There is a clear need for direct comparisons between online CBT and alternative treatment options, preferably in the form of large-sample equivalence trials that are conducted in naturalistic settings.

Cost-effectiveness. Modern healthcare is managed healthcare. The economic evaluation of an intervention is as important as the more traditional clinical evaluation. At present, cost-benefit comparisons between face-to-face treatment and online treatment are complicated by the limited availability of routine practice data and by unclear definitions of relevant health-economic variables. Nonetheless, scant research has shown that online treatment is cost-effective (e.g., Bergström et al., 2010; Tate et al., 2009), and it would be instructive to conduct a cost-effectiveness study for the present treatment as well.

Pre-treatment withdrawal. Pre-treatment withdrawal, voluntary drop-out that occurs during or even prior to the screening, demands attention. Throughout the studies, a considerable percentage of the applicants withdrew, during or even prior to the screening (40% in the routine practice study). While pre-treatment withdrawal rates are often not reported, research shows that high withdrawal rates are common, in both online and offline research (Issakidis & Andrews, 2004; Melville et al., 2010). The application of survival analysis methods, as suggested by Eysenbach (2005), might provide insight into the reasons for this withdrawal. Pre-treatment withdrawal might be caused by the ease with which one can apply for online therapy. This may result in impulsive applications (we found some indications that applicants, who started the screening without a referral, were less likely to start treatment). A second possibility is that patients withdraw from online treatment because they are unwilling to relinquish their anonymity (de Haas et al., 2009; Postel et al., 2009).

In the Netherlands, several organisations have started to provide free anonymous online CBT to clients who are unwilling to disclose their identity. Although this may be understandable given the alternative (no treatment), this raises questions about reimbursement, professional responsibility and parental consent (Lange & Ruwaard, 2010). Is it fair to provide treatment for free to some and not to others? Could providers of anonymous treatment be held responsible in case of crisis? Is it desirable that a person under 16 can start treatment when her parents are unaware of this? These questions require more discussion before wide-spread implementation of anonymous treatment should be considered.

Post-treatment booster sessions. When the effects of CBT are evaluated against conservative criteria of clinically significant change, as we did in this study, around 50% of the clients achieve high end-state functioning, i.e., recovery (Roth & Fonagy, 2005). Online CBT performed comparably in this respect. Although a large majority of clients experience a reliable change, a considerable percentage of client end treatment with residual symptoms that warrant further treatment. Currently, the options for continued online treatment are limited. The majority of the clients are referred back to their GP. For some clients with co-morbid symptoms, it makes sense to address symptoms by starting another treatment. In very rare cases, clients re-start the same treatment. However, the effects of these strategies are not clear. It may be fruitful to develop a generic booster session protocol in which clients with residual symptoms receive a series of follow-up sessions.

7.5 Conclusion

Our research strongly suggests that online therapist-assisted CBT is efficacious in the treatment of a variety of common adult mental health disorders. It is characterised by strong effects and clinically significant improvement, and by acceptable treatment adherence and high client satisfaction. Our findings further show that these effects are achieved both within the confines of the controlled trial and in the routine practice setting. Based on our results, we recommend further implementation of online CBT.

Summary

In 1997, researchers of the University of Amsterdam developed one of the first psychotherapeutic applications of the internet. They implemented a standardised cognitive behavioural treatment (CBT) of post-traumatic stress symptoms in a website, and used this site to treat a small number of students with matching symptoms. The results were surprisingly encouraging. Almost every participant reported a strong reduction in symptoms. Subsequently, controlled trials confirmed the value of this pilot study, and showed that a large part of the improvements could be attributed to the intervention. To implement the treatment in routine practice, the researchers founded an online clinic. With this clinic, the *Interapy* clinic, they introduced online psychotherapy in the Dutch mental healthcare system.

In this dissertation, we explore the wider applicability of online CBT, in four randomised controlled trials (Chapter 2 to 5) and a routine practice study (Chapter 6). The research included almost 2000 patients with a variety of psychological symptoms. In the controlled studies (N = 456), we assess the efficacy of online CBT for work-related stress, mild to moderate depression, and symptoms of panic disorder and bulimia nervosa. In the fifth study, we examine the effectiveness of online CBT in routine clinical practice (N = 1500).

Chapter 1, the introduction, starts with a brief description of the motivation behind our research. Epidemiological studies suggest that a considerable portion of the people with mental health problems do not receive adequate treatment. People do not seek help, or are not treated in accordance to current practice guidelines. Psychological, social, geographical, and organizational factors contribute to this gap between what we know and do in clinical psychotherapy. Internet applications have reduced or removed such barriers in many areas of life. Psychotherapeutic internet interventions may support public mental healthcare in a similar vein. The second part

of the introduction provides a brief overview of internet intervention research. Three different types of internet interventions are identified: online self-help (which does not include therapist support), online guided self-help (which includes limited therapist support, up to about 90 minutes per treatment, often in combination with face-to-face contact), and online psychotherapy (which includes more intensive therapist support, and no face-to-face contact). In this dissertation, we studied this last form of online treatment. Previous controlled trials established that online psychotherapy yields promising results in the treatment of post-traumatic stress. Would similar positive results be achieved with other mental health disorders? And would these results generalise to routine clinical practice?

Chapter 1 describes a randomised controlled trial of an online treatment for work-related stress. For this, a new treatment protocol was developed. Similar to the aforementioned online treatment of post-traumatic stress, the new treatment is fully online, manualised, and based on therapeutic techniques that are commonly applied in CBT. Clients and therapists communicate over the internet, without face to face contact. The nature and the sequence of the interventions is standardised, and therapists use default feedback texts which they adapt to the situation of their clients. The treatment comprises psycho-education, awareness training through self-monitoring and writing exercises, relaxation training, positive self-verbalisation, social skills training through behavioural experiments, time management and relapse prevention.

The trial was announced in national Dutch newspaper. Adult applicants were screened through online administration of validated questionnaires and a semi-structured telephone interview. Demographic data, symptom severity and contra-indications for online treatment were assessed. Applicants with suicidal ideation, alcohol- or substance abuse, or increased risk of dissociation of psychosis were excluded and referred to other mental health institutions. We assessed 342 applicants, of which 239 were accepted. Participants started online treatment immediately (n = 177), or after a waiting period of seven weeks (n = 62). Three years later, follow-up measurements were taken.

The online treatment had positive effects on stress, depression and emotional exhaustion. On relevant subscales of self-report questionnaires, the Depression Anxiety

Stress Scales (DASS) and the Maslach Burnout Inventory - General Survey (MBI), treated participants scored better in comparison to untreated participants, even after conservative corrections for dropout (26%) and multiple statistical testing ($P < 0.003$; $0.6 > d > 0.3$). The treatment increased the likelihood of recovery from stress symptoms from 16% to 50% (Odds Ratio = 5.1, $P < .0001$). At follow-up, three years after treatment, gains were found to be maintained. The study had some limitations. The average treatment took longer than estimated. Hence, post-test measurements in the two experimental groups were taken at slightly different times, which resulted in interpretative difficulties. In addition, the follow-up results could not be contrasted to a control group, because all participants had followed treatment when follow-up measurements were taken. Although these limitations demand further research, the results provided a first indication that online psychotherapy is efficacious for psychological problems other than post-traumatic stress.

Encouraged by the results of the first trial, we developed an online treatment for mild-to-moderate depression. **Chapter 3** describes the randomised controlled trial in which we assessed the efficacy of this intervention. The design of this study is similar to the first study. After online screening and a telephone interview, participants ($N = 54$) started treatment immediately $n = 36$), or after 13 weeks ($n = 18$). After 18 months, all participants were invited for online follow-up measurements. Changes in the severity of depressive symptoms and general psychopathology were assessed through the Beck Depression Inventory (BDI), the Depression Subscale of the Symptom Check List (SCL90-R), and the DASS.

Results were striking. Participants in the control condition reported strong reductions in depressive symptoms ($d = .9$) after the waiting period, despite the fact that they had not been offered treatment. However, since the observed improvements in the treatment group were very large ($d = 1.6$ - 1.9), the differences between both experimental groups remained large and significant (a *pooled* effect size of $d = .9$; $P < .03$). Treated participants were more likely to recover from their depressive symptoms (although it should be mentioned that this difference was significant on just one of the two outcome measures - the SCL90R depression subscale). Follow-up data showed that the improvements were stable on the long term. The large effect sizes and the positive evaluations of the patients carried a clear message. The results confirmed

that face-to-face contact is not a necessary requirement of efficacious treatment. Furthermore, the study provided further support to the hypothesis that the amount of therapist support is a critical determinant of the impact of an online intervention.

In the third study, which is presented in **chapter 4**, we focused on the online treatment of symptoms of panic disorder. Fifty-eight participants with clinical or subclinical panic symptoms followed a fourteen-week online treatment program, consisting of psycho-education, in vivo and in vitro exposure, relaxation training and cognitive restructuring. Half of the group started treatment immediately, and the other half started after fourteen weeks. In this study, we conducted a follow-up after three years. The severity of panic symptoms was assessed through a self-report measure, the Panic Disorder Severity Scale Self-Report (PDSS-SR), and a panic diary, in which participants recorded the intensity and frequency of their panic attacks. Bodily sensations, agoraphobic cognitions and avoidance and general psychopathology were assessed with the Bodily Sensations Questionnaire (the BSQ), the Agoraphobic Cognitions Inventory (ACQ), the Mobility Inventory (MI), and the DASS.

Drop-out was low: 89% of the participants completed the full treatment. Another strength of this study is that 81% of the participants could be traced for the three-year follow-up measurements. The treatment group improved significantly more in comparison to the control group ($P < .03$), with regard to every outcome measure, with exception of the agoraphobic measures. The significant effects were moderate to large ($.4 < d < 1.1$), and were more most pronounced in the panic diary data. Treated participants reported less frequent and less intensive panic attacks in comparison to untreated participants. The treatment increased the probability of a reduction of the number of panic attacks of at least 50% from 14% to 52%. The follow-up data suggested that the effect of treatment increased over time. Despite the positive results, the controlled effects of treatment were somewhat less than expected. This might be explained by our choice to include a heterogeneous group in the study. This explanation received some support from ad-hoc explorative analyses, which suggested that the treatment was more effective for participants, who presented with more severe symptoms.

Chapter 5 presents the design and results of the fourth and last RCT of the dissertation, in which we assessed the efficacy of online CBT in the management of

bulimic symptoms. This study had three experimental arms. One-hundred-and-five participants were randomly assigned to online treatment ($n = 35$), to a waiting list ($n = 35$), or to bibliotherapy ($n = 35$), in which participants received a self-help book, but no therapist guidance. The online treatment takes about twenty weeks, and comprises awareness training, exposure and response prevention, cognitive restructuring, behavioural experiments, and relapse prevention. The self-help book is based on a similar therapeutic approach (i.e., CBT), and comprises similar psychotherapeutic techniques. Primary outcome measures were the Eating Disorder Examination-Questionnaire (EDE-Q) and the frequency of binge eating and purging episodes. Disturbances in body satisfaction were measured through the Body Attitude Test (BAT).

Changes in the waiting list group were marginal. Participants, who followed online treatment, experienced significantly less bulimic symptoms than the participants in both control groups ($d = .9, P < .04$), while the changes in the bibliotherapy group did not differ significantly from the changes in the group assigned to the waiting list ($d = .2, P = .99$). The effect of online treatment on body satisfaction was less pronounced, and was significant only in comparison to the waiting list. The data suggested that online treatment increases the likelihood of recovery from bulimic symptoms from 20% to 40%. Interpretation of the results was complicated by the occurrence of a rather high drop-out rate in the bibliotherapy group (35%). In addition, we found that 31% of the participants in this group had sought additional treatment in the period between the pre-test and the post-test. This also might explain the results of the follow-up. In the online CBT group, improvements at post-test remained stable up to one-year after treatment. However, in the bibliotherapy group, participants strongly improved after the end of treatment, which resulted in non-significant differences between the two experimental groups at follow-up. Since the bibliotherapy did not have effects at post-test, it is likely that bibliotherapy had an indirect delayed effect. We concluded that online CBT is superior to no treatment and bibliotherapy on the short-term, and that bibliotherapy may increase the probability of recovery by promoting positive attitudes towards treatment in a considerable portion of patients.

Expectations with regard to the clinically utility of the results of controlled research are sometimes rather low. Since the conditions in routine practice are typically

complex, routine practice is only partially informed by controlled research, in which conditions are artificially simplified. It is therefore recommended to examine the performance of tested interventions in routine practice as well. This was the aim of a large naturalistic study of the effectiveness of online CBT, which is presented in **chapter 6**. This was an analysis of electronic patient records of $N = 1500$ clients of the Dutch Interapy clinic. Between 2002 and 2008, the clinic offered these patients online treatment for depression ($n = 413$), panic ($n = 139$), burnout ($n = 470$), or post-traumatic stress ($n = 478$). The study included consecutive patients, until the predetermined sample size was reached. The anonymised records provided demographic data, process data and patient satisfaction data, and pre/post/follow-up data of validated psychological self-report questionnaires of specific and general psychopathology, including the BDI, the PDSS-SR, the Impact of Event Scale (IES), the Oldenburg Burnout Inventory (OLBI), and the DASS. The online treatments, the procedures and effect measures that are used in the clinic were identical to those that were used in the controlled research. Therefore, because of the manualisation and the high degree of treatment fidelity, we expected that the effects of treatment in routine practice would be similar to those observed in the controlled studies.

Post-test measurements were available for 79% of the patients. The drop-out rate (the percentage of patients not completing the full treatment) was 29%, which is relatively low in comparison to other internet-based interventions and comparable to know drop-out rates in Dutch clinical practice. Clients evaluated the services of the clinic very positively. On the short-term (at post-test and six-weeks follow-up), clients reported significant improvements ($P < .001$) in specific and general psychopathology. The effects on the specific outcome measures ($0.7 < d < 1.9$) were somewhat better in comparison to those observed in the controlled studies. Among those, who completed treatment, 71% experienced a statistically reliable improvement, and 52% experienced a statistically reliable recovery. The results of the one-year follow-up should be interpreted with caution, since only 33% of the clients returned for follow-up measurements. However, the available data suggest that improvements are maintained up to one year after treatment.

In **chapter 7**, the general discussion, the studies are revisited from a birds-eye perspective, and re-evaluated within the context of the general aims of the research

program. Previous studies identified online CBT as an efficacious treatment alternative for post-traumatic stress. Would this also hold for psychological symptoms which have a less clear etiology, such as work-related stress, depression, panic attacks and bulimic symptoms? The controlled studies in this research suggest this to be the case. In addition, the routine practice study shows that these effects generalise to the applied setting. Online CBT appears to earn its place in the public health system. Treatment adherence is acceptable, treatment effects are comparable to regular face-to-face CBT, and clients highly appreciate this form of treatment.

Future research should focus on direct comparisons between online CBT and treatment alternatives, especially with regard to the cost-effectiveness and the optimal amount of human guidance. In addition, more attention should be given to clients who respond less to treatment. Some clients end treatment with residual symptoms, which calls for continued research efforts to enhance the effectiveness of treatment. For this, we need to increase our understanding of the effective components of online treatment. In addition, we may need more flexible forms of treatment. Finally, there is a clear need for independent replication research. In our studies, we evaluated the treatments that we also developed. Although this is common practice in internet intervention research, independent replications would significantly strengthen the evidence base of online CBT.

In the Netherlands, online CBT is reimbursed since 2005. Online CBT has been an accessible treatment alternative for patients with a variety of mental health symptoms for several years. At present, nonetheless, only a marginal fraction of Dutch mental healthcare is provided online. Referrers appear to be unable to see the wood for the trees. It is hoped that this dissertation will provide some clarification.

Samenvatting

In 1997 ontwikkelden onderzoekers van de Universiteit van Amsterdam één van de eerste psychotherapeutische toepassingen van het World Wide Web. Ze zetten een protocol voor de behandeling van posttraumatische stress om in een website en nodigden een aantal studenten met posttraumatische stressklachten uit om zich via deze site te laten behandelen. De resultaten waren verrassend gunstig. Vrijwel alle deelnemers rapporteerden na behandeling een sterke afname van hun klachten. Gecontroleerd vervolgonderzoek liet zien dat deze verbeteringen voor een groot gedeelte aan de behandeling mochten worden toegeschreven. Om de behandeling in de praktijk te brengen startten de onderzoekers vervolgens een online behandelinstituut. Met dit instituut, de *Interapy* kliniek, introduceerden zij online psychotherapie in de Nederlandse GGZ.

Dit proefschrift bouwt voort op het hierboven beschreven onderzoek. Het omvat vier gecontroleerde gerandomiseerde studies (hoofdstuk 2 tot en met 5) en een praktijkstudie (hoofdstuk 6) onder bijna 2000 mensen met verschillende psychische klachten. In de gecontroleerde studies ($N = 456$) onderzochten we de werkzaamheid van online psychotherapie bij werkgerelateerde stress, depressie, paniek en boulimie. In de praktijkstudie ($N = 1500$) onderzochten we de effectiviteit van online psychotherapie in de dagelijkse klinische praktijk.

Hoofdstuk 1, de inleiding, begint met de motivatie achter het onderzoek. Uit bevolkingsonderzoek blijkt dat een aanzienlijk deel van de mensen met psychische klachten daarvoor niet de meest adequate behandeling ontvangt. Mensen zoeken geen hulp, of worden niet conform de richtlijnen behandeld. De belemmerende factoren blijken psychologisch, sociaal, geografisch en organisatorisch van aard. Internettoepassingen hebben soortgelijke barrières in tal van maatschappelijke domeinen verminderd of zelfs opgeheven. Door de ontwikkeling van psychotherapeutische toepassingen

hoopt men hetzelfde te bereiken in de geestelijke gezondheidszorg. Het tweede deel van de inleiding gaat in op de verschillende vormen van online hulpverlening en wat er globaal bekend is over de resultaten hiervan voor de afname van psychische klachten. Er wordt een onderscheid gemaakt tussen pure zelfhulp (komt geen behandelaar aan te pas), begeleide zelfhulp (waarbij het contact met behandelaars wordt gelimiteerd tot ongeveer anderhalf uur, vaak in combinatie met face-to-face contacten) en online psychotherapie (waarbij terugdringen van het aantal begeleidingsuren niet het primaire uitgangspunt is en de dialoog tussen behandelaar en cliënt volledig via het internet verloopt). De studies in dit proefschrift richten zich op deze laatste vorm. Uit gecontroleerd onderzoek was gebleken dat online psychotherapie succesvol is in de behandeling van posttraumatische stress. Zou dat ook gelden voor andere psychische klachten? En zouden de resultaten ook generaliseren naar de klinische praktijk?

Hoofdstuk 2 beschrijft een gerandomiseerd gecontroleerd onderzoek naar de werkzaamheid van een online therapie voor werkgerelateerde stress. Daarvoor werd een nieuwe behandeling ontwikkeld. Net als de online behandeling van posttraumatische stress, is deze behandeling volledig online, strict geprotocolleerd en gebaseerd op technieken uit de cognitieve gedragstherapie (CGT). Cliënt en behandelaar hebben uitsluitend via het internet contact. De aard en de volgorde van de interventies ligt vast en voor het contact met de cliënten wordt gebruik gemaakt van standaardteksten die door de behandelaars worden gepersonaliseerd. De behandeling omvat psycho-educatie, bewustwording via klachtmonitoring en schrijfopdrachten, ontspanningsoefeningen, zelfbeeldverbetering via positieve zelfverbalisatie, gedragsexperimenten gericht op het verbeteren van sociale vaardigheden, timemanagement en terugvalpreventie.

Het onderzoek werd aangekondigd in een landelijk dagblad. Via online afgenomen gevalideerde vragenlijsten en een semi-gestructureerd telefonisch interview werden geïnteresseerden gescreend. Hiermee werden demografische gegevens, de ernst van de psychologische klachten en eventuele contra-indicaties voor behandeling in kaart gebracht. Bij risico op zelfmoord, een recente opname in een psychiatrische kliniek, alcohol- of middelenmisbruik of een verhoogde kans op dissociatie en psychose, werd doorverwezen naar de reguliere hulpverlening. De werving leidde tot 342 aanmeldingen, waarbij 239 maal tot toelating kon worden overgegaan. De deelnemers

startten direct met de online behandeling ($n = 177$) of pas na een wachtperiode van zeven weken ($n = 62$). Drie jaar later volgde een online follow-up.

De online behandeling bleek positieve effecten te hebben op stress, depressie en emotionele uitputting. Op relevante zelfrapportage vragenlijsten, de Depression Anxiety Stress Scales (DASS) en de Utrechtse Burnout Schaal (UBOS), scoorden behandelde deelnemers significant beter dan onbehandelde deelnemers, zelfs na strenge correctie voor dropout (26%) en meervoudig toetsen ($P < 0.003$; $0.6 > d > 0.3$). De behandeling verhoogde de kans op herstel van stressklachten van 16% naar 50% (Odds Ratio = 5.1, $P < .0001$). Tijdens de follow-up, drie jaar na behandeling, bleken de verbeteringen verder doorgezet. De studie kende enkele beperkingen. Zo bleek de behandeling langer te duren dan gepland, waardoor er een probleem ontstond in de timing van de metingen en de interpretatie van het verschil tussen de twee experimentele condities. Daarnaast konden de resultaten van de follow-up niet afgezet tegen een controlegroep, omdat alle deelnemers inmiddels waren behandeld. Hoewel vervolgonderzoek dus gewenst is, gaven de resultaten een eerste aanwijzing dat online psychotherapie ook werkzaam is bij andere klachten dan posttraumatische stress.

Gesteund door de resultaten van de eerste trial, werd een online behandeling ontwikkeld voor milde tot matige depressieklachten. **Hoofdstuk 3** beschrijft het gerandomiseerde gecontroleerde onderzoek naar de werkzaamheid van deze behandeling. De opzet van de studie is identiek aan de studie uit hoofdstuk twee. Na online screening en een telefonisch interview volgden de deelnemers ($N = 54$) de behandeling direct ($n = 36$) of pas na dertien weken ($n = 18$). Anderhalf jaar later werden alle deelnemers werden uitgenodigd voor een online follow-up. Veranderingen in depressieve klachten en algemene psychopathologie werden gemeten via de Beck Depression Inventory (BDI), de Depressie Subschaal van de Symptom Check List (SCL90-R) en de DASS.

De uitkomst was opvallend. Deelnemers in de controleconditie rapporteerden na de wachtperiode een grote afname van depressieve klachten ($d = 0.9$), ondanks het feit dat zij geen enkele behandeling hadden gevolgd. De gemeten vooruitgang in de behandelde groep was echter zeer groot ($d = 1.6$-1.9), waardoor alsnog een groot en significant verschil ontstond tussen de beide experimentele condities (een *gepoolde*

effectgrootte van $d = .9$; $P < .03$). Behandelde deelnemers herstelden ook relatief vaker van depressieve klachten dan de deelnemers in de wachtlijstconditie (al moet daarbij aangemerkt dat dit verschil significantie bereikte op slechts één van de twee primaire uitkomstmaten - de SCL-90R). De follow-up liet zien dat de verbeteringen stabiel waren gebleven. Deelnemers die tijdens of na de behandeling antidepressiva gebruikten, bleken echter wel teruggevallen. De grote effectsizes en de hoge mate van cliënttevredenheid die werden gevonden in deze studie gaven een duidelijk signaal af. De studie bevestigde het beeld dat *face-to-face* contact geen noodzakelijke voorwaarde is voor effectieve behandeling van depressieve klachten. Bovendien werd verdere ondersteuning gegeven aan de hypothese dat de mate van begeleiding een kritische determinant is van de impact van een online interventie.

In de derde studie, beschreven in **hoofdstuk 4**, richtten we ons op de online behandeling van paniekklachten. Achtenvijftig deelnemers met klinische of subklinische paniekklachten doorliepen een online programma van veertien weken, dat bestaat uit een combinatie van psycho-educatie, in vivo en in vitro exposure, ontspanningsoefeningen en cognitieve herstructurering. De helft van de deelnemers startte meteen en de andere helft na 14 weken. De follow-up werd dit keer na drie jaar uitgevoerd. De ernst van de paniekklachten werd in kaart gebracht via een zelfrapportagevragenlijst, de Panic Disorder Severity Scale Self-Report (PDSS-SR) en een paniekdagboek, waarin de deelnemers de intensiteit en frequentie van hun paniekaanvallen bijhielden. Lichamelijke sensaties, agoraphobische cognities en vermijding en algemene psychopathologie, werden gemeten via de Bodily Sensations Questionnaire (de BSQ), de Agoraphobic Cognitions Inventory (ACQ), de Mobility Inventory (MI) en de DASS.

De drop-out was laag: 89% van de deelnemers rondde de volledige behandeling af. Bijzonder was ook dat 81% van de deelnemers kon worden achterhaald voor de drie-jaar follow-up. De behandelgroep bleek op alle uikomstmaten significant meer verbeterd dan de controlegroep ($p < .03$), met uitzondering van de agorafobische maten (ACQ en MI). De significante effecten waren middelgroot tot groot ($.4 < d < 1.1$) en kwamen het sterkst tot uitdrukking in het paniekdagboek. Behandelde deelnemers rapporteerden duidelijk minder frequente en minder intensieve paniekaanvallen dan de onbehandelde deelnemers. De behandeling verhoogde de kans op een halvering van het aantal paniekaanvallen van 14% naar 52%. De follow-

up-metingen suggereerden dat het effect van de behandeling toeneemt op de lange termijn. Ondanks de positieve uitkomsten waren de gecontroleerde effecten van de behandeling toch wat minder hoog dan verwacht. Dit wordt mogelijk verklaard door de keuze om een heterogene groep te includeren voor dit onderzoek. Deze verklaring krijgt enige ondersteuning uit exploratieve analyses, die suggereren dat het effect van de behandeling groter is bij ernstigere aanvangsklachten.

Hoofdstuk 5 bevat een beschrijving van de opzet en resultaten van de vierde en laatste RCT van deze dissertatie, waarin de werkzaamheid van online CBT op boulimische klachten werd onderzocht. Dit onderzoek kende drie experimentele condities. Honderdenvijf deelnemers werden aselect toegewezen aan de online behandeling ($n = 35$), een wachtlijstconditie ($n = 35$) of een meer actieve pure zelfhulpconditie ($n = 35$) waarin een zelfhulpboek werd aangeboden zonder enige begeleiding van een behandelaar. De online behandeling, die ongeveer twintig weken duurt, omvat bewustwording van het klachtenpatroon via gedragsregistratie en schrijfopdrachten, exposure en responspreventie, cognitieve herstructurering, gedragsexperimenten, oefeningen voor de verbetering van lichaamsperceptie en zelfbeeld en het opstellen van een plan om terugval te voorkomen of te recouperen. Het zelfhulpboek is op dezelfde principes gebaseerd en omvat vergelijkbare therapeutische technieken. De primaire uitkomstmaten waren de frequentie van de eetbuiaanvallen en (purgerende) compensatoire maatregelen en de globale ernst van de eetbuistoornis. Deze variabelen werden via zelfrapportage vastgesteld middels de Eating Disorder Examination-Questionnaire (EDE-Q). Verstoringen in de lichaamsbeleving werden gemeten via de LichaamsAttitudeVragenlijst (LAV).

De veranderingen in de wachtlijstconditie waren marginaal. Deelnemers die de online behandeling volgden ervoeren significant minder boulimische klachten dan de deelnemers in beide controlecondities ($d = .9$, $P < .04$), terwijl de verbeteringen bij de deelnemers in de zelfhulpconditie niet noemenswaardig verschilden van de deelnemers in de wachtlijstconditie ($d = .2$, $P = .99$). Het effect van de online behandeling op verstoorde lichaamsperceptie was minder duidelijk en bereikte alleen significantie ten opzichte van de wachtlijstconditie. De studie gaf aanwijzingen dat de online behandelingen de kans op herstel van boulimische klachten verhoogt van ongeveer 20% naar 40%. De interpretatie van de resultaten werd bemoeilijkt door een

vrij forse uitval in de zelfhulpconditie (35%). Bovendien bleek 31% van de deelnemers in de zelfhulpconditie tijdens de periode tussen de voor- en nameting met een andere behandeling te zijn gestart. Mogelijk verklaart dit ook de resultaten van de follow-up, die één jaar na behandeling werd uitgevoerd. De effecten van de online behandeling bleken grotendeels stabiel. Maar het verschil tussen online behandeling en het zelfhulpboek was verdwenen, omdat de deelnemers in de zelfhulpcondititie duidelijk waren verbeterd in de periode tussen de nameting en de follow-up. Geconcludeerd werd dat online CBT op de korte termijn een duidelijke meerwaarde heeft boven geen behandeling en zelfhulp en dat de effecten van zelfhulp mogelijk voornamelijk moeten worden gezocht in het vergroten van de bereidheid bij de cliënt om verdere hulp te zoeken.

Verwachtingen over de klinische toepasbaarheid van de resultaten van gecontroleerd onderzoek zijn soms nogal somber. De klinische praktijk is weerbarstig en complex en zou daarom slechts matig worden geïnformeerd door gecontroleerde studies, waarin doorgaans juist geïdealiseerde en onnatuurlijk versimplificeerde situaties worden gecreëerd. Het verdient daarom aanbeveling om werkzame interventies ook te toetsen in de praktijksituatie. Dit was het doel van een grote naturalistische studie naar de effectiviteit van online CBT, die wordt gepresenteerd in **hoofdstuk 6**. Dit was een analyse van de electronische patiëntendossiers van $N = 1500$ cliënten van de Interapy kliniek. De cliënten volgden daar tussen 2002 en 2008 een online behandeling volgden voor depressie ($n = 413$), paniekklachten ($n = 139$), burnoutklachten ($n = 470$) of posttraumatische stress ($n = 478$).

De studie includeerde alle cliënten, op volgorde van aanmelding, totdat de van te voren vastgestelde steekproefgrootte was bereikt. De geanonimiseerde dossiers bevatten demografische gegevens, gegevens over het verloop van de behandeling en cliënttevredenheid en pre/post/follow-up data van gevalideerde psychologische zelfrapportagevragenlijsten voor specifieke en algemene psychopathologie, waaronder de BDI, de PDDS-SR, de Schokverwerkingslijst (SVL), de Oldenburg Burnout Inventory (OLBI) en de DASS. De online behandelingen van de Interapy kliniek waren identiek aan de behandelingen in de gecontroleerde onderzoeken. Ook de procedures en de effectmetingen kwamen grotendeels overeen. De verwachting was dan ook dat, door de hoge mate van geprotocolleerde standaardisatie, de effecten in de praktijk

ongeveer gelijk zouden zijn aan die van de gecontroleerde studies.

Nametingen bleken beschikbaar voor 79% van de cliënten. Het uitvalspercentage was met 29% vergelijkbaar met gepubliceerde uitvalspercentages in de Nederlandse GGz en relatief laag in vergelijking met andere vormen van online hulpverlening. Cliënten waren uitgesproken positief over de geboden zorg. Op de korte termijn (tijdens de nameting en follow-up na zes weken), rapporteerden cliënten significante verbeteringen ($P < .001$) in specifieke en algemene psychopathologie. De effecten op de specifieke uitkomstmaten ($0.7 < d < 1.9$) waren over de hele linie iets groter dan in de gecontroleerde studies. In de groep die de behandeling afmaakte, ervoer 71% een statistisch betrouwbare verbetering en 52% een klinisch significante verbetering. De resultaten van de follow-up na één jaar moeten voorzichtig geïnterpreteerd, omdat slechts 33% van de clienten hieraan deelnam. De beschikbare gegevens suggereren echter dat de verbeteringen standhouden op de lange termijn.

Hoofdstuk 7 heeft een beschouwend karakter. De studies worden vanuit vogelperspectief nogmaals besproken en opnieuw in verband gebracht met de oorspronkelijke doelstellingen van het onderzoeksprogramma. Online CBT bleek in eerder onderzoek een goed alternatief voor de behandeling van posttraumatische stress. Maar geldt dit ook voor psychologische klachten met een minder duidelijke etiologie, zoals werkgerelateerde stress, depressie, paniekaanvallen en boulimische klachten? De gecontroleerde studies in dit proefschrift geven sterke aanwijzingen dat dit inderdaad zo is. De veldstudie laat bovendien zien dat de effecten overeind blijven in de klinische praktijk. Online CBT lijkt haar plek in de geestelijke gezondheidszorg te verdienen. De therapietrouw is redelijk tot goed, de effecten doen niet onder voor reguliere CBT en cliënten blijken deze vorm van behandelen te waarderen.

Veel blijft ook onduidelijk, zoals dat hoort in de wetenschap. Vervolgonderzoek zou zich moeten richten op directe vergelijkingen tussen online CBT en andere behandelvormen, zodat een goed beeld ontstaat van de kosteneffectiviteit van online CBT en de minimaal noodzakelijke hoeveelheid begeleiding van de behandelaar. Online CBT blijkt daarnaast niet voor iedereen afdoende. Een deel van de cliënten eindigt de behandeling met restklachten. Er is dus alle reden om de behandelingen te verbeteren. Een beter begrip van de effectieve bestanddelen van de behandelprogramma's zou in dit kader behulpzaam zijn, net als de ontwikkeling van meer flexibele vormen

van online CBT. Het verdient ten slotte aanbeveling dat toekomstig onderzoek ook wordt uitgevoerd door andere onderzoeksgroepen. Zoals bij vrijwel al het onderzoek naar internet-behandelingen werden ook hier de behandelingen gevalideerd door de oorspronkelijke ontwikkelaars. Verdere onafhankelijke replicaties zouden een belangrijke toevoeging vormen op de overtuigingskracht van het bewijs.

Online CBT wordt sinds 2005 vergoed uit de basisverzekering. Online CBT is dus al enkele jaren een toegankelijk, efficiënt en reëel alternatief voor mensen met uiteenlopende psychische klachten. Toch vormt het aantal online behandelingen nog steeds een marginaal deel van het totaal aan behandelingen in de GGz. Verwijzers lijken door de bomen het bos niet meer te zien. Hopelijk verschaft dit proefschrift enige helderheid.

References

Aben, I., Verhey, F., Lousberg, R., Lodder, J., & Honig, A. (2002). Validity of the Beck Depression Inventory, Hospital Anxiety and Depression Scale, SCL-90, and Hamilton Depression Rating Scale as screening instruments for depression in stroke patients. *Psychosomatics*, *43*(5), 386-393. doi: 10.1176/appi.psy.43.5.386

Alford, W. K., Malouff, J. M., & Osland, K. S. (2005). Written emotional expression as a coping method in child protective services officers. *International Journal of Stress Management*, *12*, 177–187. doi: 10.1037/1072-5245.12.2.177

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders, DSM fourth edition - text revision*. Washington, DC: American Psychiatric Association.

Andersson, G. (2006). Internet-based cognitive-behavioral self-help for depression. *Expert Review of Neurotherapeutics*, *6*(11), 1637-42. doi: 10.1080/16506070310014466

Andersson, G., Bergström, J., Holländare, F., Carlbring, P., Kaldo, V., & Ekselius, L. (2005). Internet-based self-help for depression: Randomised controlled trial. *British Journal of Psychiatry*, *187*, 456-61. doi: 10.1192/bjp.187.5.456

Andersson, G., Bergström, J., Holländare, F., Ekselius, L., & Carlbring, P. (2004). Delivering CBT for depression via the internet. predicting outcome at 6-months follow-up. *Verhaltenstherapie*, *14*, 185-189. doi: 10.1159/000080914

Andersson, G., Carlbring, P., & Cuijpers, P. (2009). Internet interventions: Moving from efficacy to effectiveness. *Electronic Journal of Applied Psychology.*, *5*(2), 18-24. Retrieved from `http://ojs.lib.swin.edu.au/index.php/ejap`

Andersson, G., Carlbring, P., Holmström, A., Sparthan, E., Furmark, T., Nilsson-Ihrfelt, E., ... Ekselius, L. (2006). Internet-based self-help with therapist feedback and in vivo group exposure for social phobia: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *74*, 677-686. doi: 10.1037/0022-006X.74.4.677

Andersson, G., & Cuijpers, P. (2009). Internet-based and other computerized psychological treatments for adult depression: A meta-analysis. *Cognitive Behaviour Therapy*, *38*(4), 196 - 205. doi: 10.1080/16506070903318960

Andersson, G., Cuijpers, P., Carlbring, P., & Lindefors, N. (2007). Effects of internet-delivered cognitive behaviour therapy for anxiety and mood disorders. *Helix Review Series: Psychiatry*(2), 9-14.

Andrews, G., Cuijpers, P., Craske, M., McEvoy, P., & Titov, N. (2010). Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: A meta-analysis. *PLoS One*, *5*(10), e13196. doi: 10.1371/journal.pone.0013196

Arrindell, W., & Ettema, J. (2003). *SCL-90. Herziene handleiding bij een multidimensionele psychopathologie-indicator.* [SCL-90: Revised manual for a multidimensional indicator of psychopathology]. Lisse: Swets & Zeitlinger.

Avins, A. L. (1998). Can unequal be more fair? ethics, subject allocation, and randomised clinical trials. *Journal of Medical Ethics*, *24*(6), 401-8. doi: 10.1136/jme.24.6.401

Aziz, M. A., & Kenford, S. (2004). Comparability of telephone and face-to-face interviews in assessing patients with posttraumatic stress disorder. *Journal of Psychiatric Practice*, *10*(5), 307-13. doi: 00131746-200409000-00004

Barak, A. (1999). Psychological applications on the internet: A discipline on the threshold of a new millennium. *Applied and Preventive Psychology*, *8*, 231–245. doi: 10.1016/S0962-1849(05)80038-1

Barak, A., Hen, L., Boniel-Nissim, M., & Shapira, N. (2008). A comprehensive review and a meta-analysis of the effectiveness of internet-based psychotherapeutic interventions. *Journal of Technology in Human Services*, *26*. doi: 10.1080/15228830802094429

Barak, A., Klein, B., & Proudfoot, J. G. (2009). Defining internet-supported therapeutic interventions. *Annals of Behavioral Medicine*, *38*(1), 4-17.

Batelaan, N., de Graaf, R., Van Balkom, A. J., Vollebergh, W., & Beekman, A. (2006). De epidemiologie van paniek [The epidemiology of panic]. *Tijdschrift voor Psychiatrie*, *48*, 195–205.

Bauer, C. (2000). Limiting factors on the transformative powers of e-mail in patient-physician relationships: A critical analysis. *Health Communication*, *12*, 239–259. doi: 10.1207/S15327027HC1203_02

Beck, A. T. (2005). The current state of cognitive therapy: A 40-year retrospective. *Archives of General Psychiatry*, *62*, 953–959. doi: 10.1001/archpsyc.62.9.953

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8*(1), 77-100. doi: 10.1016/0272-7358(88)90050-5

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. doi: 10.1001/archpsyc.1961.01710120031004

Beller, E. M., Gebski, V., & Keech, A. C. (2002). Randomisation in clinical trials. *The Medical Journal of Australia*, *177*(10), 565-567.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, *29*, 1165–1188. doi: 10.1214/aos/1013699998

Bennett-Levy, J., Richards, D., Farrand, P., Christensen, H., Griffiths, K., Kavanagh, D., ... White, J. (2010). *Oxford guide to low intensity cbt interventions*. Oxford: Oxford University Press.

Berg, K. C., Stiles-Shields, E. C., Swanson, S. A., Peterson, C. B., Lebow, J., & Le Grange, D. (2011). Diagnostic concordance of the interview and questionnaire versions of the eating disorder examination. *International Journal of Eating Disorders*, epub ahead of print. doi: 10.10012/eat.20948

Bergström, J., Andersson, G., Karlsson, A., Andreewitch, S., Ruck, C., Carlbring, P., & Lindefors, N. (2009). An open study of the effectiveness of internet treatment for panic disorder delivered in a psychiatric setting. *Nordic Journal of Psychiatry*, *63*(1), 44-50. doi: 10.1080/08039480802191132

Bergström, J., Andersson, G., Ljotsson, B., Ruck, C., Andreewitch, S., Karlsson, A., ... Lindefors, N. (2010). Internet-versus group-administered cognitive behaviour therapy for panic disorder in a psychiatric setting: A randomised trial. *BMC Psychiatry*, *10*(1), 54. doi: 10.1186/1471-244X-10-54

Bickel, R. (2007). *Multilevel analysis for applied research: It's just regression!* New York: the Guilford Press.

Bonger, B. (1988). Clinician, microcomputers, and confidentiality. *Professional Psychology: Research and Practice*, *19*, 286–289. doi: 10.1037/0735-7028.19.3.286

Bosscher, R. J. (1991). *Running therapie bij depressie* [Running therapy for depression]. Amsterdam: Thesis publishers.

Bouman, T. (1994). Kort instrumenteel; De Beck Depression Inventory (BDI) [Measurement instruments; The Beck Depression Inventory (BDI)]. *Gedragstherapie*, *27*(1), 69-71.

Bouman, T., Luteijn, F., Albersnagel, F., & van der Ploeg, F. (1985). Enige ervaringen met de Beck Depression Inventory [Some experiences with the Beck Depression Inventory]. *Gedragstherapie*, *13*, 13-24.

Bouman, T. K. (1995). De Agoraphobic Cognitions Questionnaire (ACQ) [The Agoraphobic Cognitions Questionnaire (ACQ)]. *Gedragstherapie*, *27*, 69–72.

Bouman, T. K. (1998). De Body Sensations Questionnaire (BSQ) [The Body Sensations Questionnaire (BSQ)]. *Gedragstherapie*, *31*, 165–168.

Bradley, C. (2000). The 12-item Well-Being Questionnaire: Origins, current stage of development, and availability. *Diabetes Care*, *23*(6), 875. doi: 10.2337/diacare.23.6.875

Brennan, S. E., & Ohaeri, J. O. (1999). Why do electronic conversations seem less polite? The costs and benefits of hedging. In *Proceedings, International Joint Conference on Work Activities, Coordination, and Collaboration (WACC 1999)* (p. 227-235). San Fransisco, CA: ACM Press. doi: 10.1145/295665.295942

Brenninkmeijer, V., & Van Yperen, N. (2003). How to conduct research on burnout: Advantages and disadvantages of a unidimensional approach in burnout research. *Occupational and Environmental Medicine*, *60*, 16-20. doi: 10.1136/oem.60.suppl_1.i16

Brom, D., & Kleber, R. J. (1985). De SchokVerwerkingsLijst [The Impact of Event Scale]. *Nederlands Tijdschrift voor Psychologie*, *40*, 164-168.

Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: a review of meta-analyses. *Clinical Psychology Review*, *26*(1), 17-31. doi: 10.1016/j.cpr.2005.07.003

Cacciola, J. S., Alterman, A. I., Rutherford, M. J., McKay, J. R., & May, D. J. (1999). Comparability of telephone and in-person structured clinical interview for DSM-III-R (SCID) diagnoses. *Assessment*, *6*(3), 235-42. doi: 10.1177/107319119900600304

Cahill, J., Barkham, M., Hardy, G., Rees, A., Shapiro, D. A., Stiles, W. B., & Macaskill, N. (2003). Outcomes of patients completing and not completing cognitive therapy for depression. *British Journal of Clinical Psycholy*, *42*(Pt 2), 133-43. doi: 10.1348/014466503321903553

Carlbring, P., Ekselius, L., & Andersson, G. (2003). Treatment of panic disorder via the Internet: A randomized trial of CBT vs. applied relaxation. *Journal of Behavior Therapy and Experimental Psychiatry*, *34*, 129–140. doi: 10.1016/S0005-7916(03)00026-0

Carlbring, P., Furmark, T., Steckzó, J., Ekselius, L., & Andersson, G. (2006). An open study of internet-based bibiliotherapy with minimal contact via email for social phobia. *Clinical Psychologist*, *10*, 30–38. doi: 10.1080/13284200500378662

Carlbring, P., Nilsson-Ihrfelt, E., Waara, J., Kollenstam, C., Buhrman, M., Kaldo, V., ... Andersson, G. (2005). Treatment of panic disorder: Live therapy vs. self-help via the internet. *Behaviour Research and Therapy*, *43*, 1321–1333. doi: 10.1016/j.brat.2004.10.002

Carlbring, P., Westling, B., Ljungstrand, P., Ekselius, L., & Andersson, G. (2001). Treatment of panic disorder via the Internet: A randomized trial of a self-help program. *Behavior*

Therapy, *32*, 751–764. doi: 10.1016/S0005-7894(01)80019-8

Chambless, D. (1993). *Task force on promotion and dissemination of psychological procedures. a report adopted by the division 12 board, october 1993*. Washington: American Psychological Association.

Chambless, D., Caputo, G., Bright, P., & Gallager, R. (1984). Assessment of fear of fear in agoraphobics: The Body Sensations Questionnaire and the Agoraphobic Cognitions Questionnaire. *Journal of Consulting and Clinical Psychology*, *52*, 1090–1097. doi: 10.1037/0022-006X.52.6.1090

Childress, C. A. (2000). Ethical issues in providing online psychotherapeutic interventions. *Journal of Medical Internet Research*, *2*(1), E5.

Christensen, H., Griffiths, K. M., & Jorm, A. F. (2004). Delivering interventions for depression by using the internet: Randomised controlled trial. *British Medical Journal*, *328*(7434), 265. doi: 10.1136/bmj.37945.566632.EE

Christensen, H., Griffiths, K. M., & Korten, A. (2002). Web-based cognitive behavior therapy: analysis of site usage and changes in depression and anxiety scores. *Journal of Medical Internet Research*, *4*(1), e3.

Christensen, H., Griffiths, K. M., Korten, A. E., Brittliffe, K., & Groves, C. (2004). A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. *Journal of Medical Internet Research*, *6*(4), e46. doi: 10.2196/jmir.6.4.e46

Clark, D. M. (1986). A cognitive approach to panic. *Behavioural Research and Therapy*, *24*, 461–470. doi: 10.1016/0005-7967(86)90011-2

Clarke, G., Reid, E., Eubanks, D., O'Connor, E., DeBar, L. L., Kelleher, C., ... Nunley, S. (2002). Overcoming depression on the internet (odin): A randomized controlled trial of an internet depression skills intervention program. *Journal of Medical Internet Research*, *4*(3), E14. doi: 10.2196/jmir.4.3.e14

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

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*, 385–396. doi: 10.2307/2136404

Collins, K. A., Westra, H. A., Dozois, D. J., & Burns, D. D. (2004). Gaps in accessing treatment for anxiety and depression: Challenges for the delivery of care. *Clinical Psychology Review*, *24*(5), 583-616. doi: 10.1016/j.cpr.2004.06.001

Conover, W. J., & Iman, R. L. (1982). Analysis of covariance using the rank transformation. *Biometrics*, *38*(3), 715-724. doi: 10.2307/2530051

Cooper, C. L., Dewe, P. J., & O'Driscoll, M. P. (2001). *Organizational stress: A review and critique of theory, research, and applications*. California: Sage Publications, Thousand Oaks.

Crawford, J. R., & Henry, J. D. (2003). The Depression Anxiety Stress scales (DASS): Normative data and latent structure in a large non-clinical sample. *British Journal of Clinical Psychology*, *42*, 111–131. doi: 10.1348/014466503321903544

Crippa, J. A., de Lima Osorio, F., Del-Ben, C. M., Filho, A. S., da Silva Freitas, M. C., & Loureiro, S. R. (2008). Comparability between telephone and face-to-face structured clinical interview for dsm-iv in assessing social anxiety disorder. *Perspectives in Psychiatric Care*, *44*(4), 241-7. doi: 10.1111/j.1744-6163.2008.00183.x

Crits-Christoph, P., Baranackie, K., Kurcias, J., Beck, A., Carroll, K., Perry, K., ... Zitrin, C. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychother-*

apy Research, *1*(2), 81 - 91.

Crow, S., Mussell, M. P., Peterson, C., Knopke, A., & Mitchell, J. (1999). Prior treatment received by patients with bulimia nervosa. *International Journal of Eating Disorders*, *25*(1), 39-44. doi: 10.1002/(SICI)1098-108X(199901)25:1<39::AID-EAT5>3.0.CO;2-W

Cuijpers, P., Marks, I. M., van Straten, A., Cavanagh, K., Gega, L., & Andersson, G. (2009). Computer-aided psychotherapy for anxiety disorders: A meta-analytic review. *Cognitive Behaviour Therapy*, *38*(2), 66 - 82. doi: 10.1080/16506070802694776

Cuijpers, P., van Straten, A., & Andersson, G. (2008). Internet-administered cognitive behavior therapy for health problems: A systematic review. *Journal of Behavioral Medicine*, *31*(2), 169-77. doi: 10.1007/s10865-007-9144-1

Cuijpers, P., van Straten, A., & Warmerdam, L. (2007). Behavioral activation treatments of depression: A meta-analysis. *Clinical Psychology Review*, *27*(3), 318-326.

de Beurs, E. (1993). *The assessment and treatment of panic disorder with agoraphobia.* Amsterdam: Thesis Publishers.

de Beurs, E., Lange, A., & van Dyck, R. (1992). Self-monitoring of panic attacks and retrospective estimates of panic: Discordant findings. *Behaviour Research and Therapy*, *30*(4), 411-3. doi: 10.1016/0005-7967(92)90054-K

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 depression, anxiety and stress.] *Gedragstherapie*, *34*, 35–53.

de Graaf, L. E., Huibers, M. J., Riper, H., Gerhards, S. A., & Arntz, A. (2009). Use and acceptability of unsupported online computerized cognitive behavioral therapy for depression and associations with clinical outcome. *Journal of Affective Disorders*, *116*(3), 227-231. doi: 10.1016/j.jad.2008.12.009

de Haas, S., Hing, M., Schrieken, B., Tsagana, A., Wijnker, C., Michels, C., ... Lange, A. (2009). Behandeling via het internet voor jeugdige slachtoffers van seksueel geweld: Een pilotstudie. [Internet treatment for young victims of sexual abuse: A pilot study.] *Maandblad voor de Geestelijke Volksgezondheid*, *64*, 981-997.

de Jong, G. M., & Emmelkamp, P. M. (2000). Implementing a stress management training: Comparative trainer effectiveness. *Journal of Occupational Health Psychology*, *5*, 309–320. doi: 10.1037//1076-8998.5.2.309

Dekker, J. J., Koelen, J. A., Van, H. L., Schoevers, R. A., Peen, J., Hendriksen, M., ... De Jonghe, F. (2007). Speed of action: The relative efficacy of short psychodynamic supportive psychotherapy and pharmacotherapy in the first 8 weeks of a treatment algorithm for depression. *Journal of Affective Disorders*. doi: 10.1016/j.jad.2007.10.015

Delinsky, S. S., & Wilson, G. T. (2006). Mirror exposure for the treatment of body image disturbance. *International Journal of Eating Disorders*, *39*(2), 108-16.

Demerouti, E., Bakker, A., Vardakou, I., & Kantas, A. (2003). The convergent validity of two burnout instruments: A multitrait-multimethod analysis. *European Journal of Psychological Assessment*, *19*(12-23). doi: 10.1027//1015-5759.19.1.12

Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology*, *86*(3), 499-512. doi: 10.1037/0021-9010.86.3.499

Derogatis, L. (1977). *SCL-90R: Administration, scoring, and procedures manual for the R(evised) version.* Baltimore: John Hopkins University School of Medicine, Clinical Psychometrics Research Unit.

Donker, T., van Straten, A., Riper, H., Marks, I., Andersson, G., & Cuijpers, P. (2009). Implementation of internet-based preventive interventions for depression and anxiety: Role of support? The design of a randomized controlled trial. *Trials*, *10*(1), 59. doi: 10.1186/1745-6215-10-59

Edwards, S. J., & Braunholtz, D. A. (2000). Can unequal be more fair? A response to andrew avins. *Journal of Medical Ethics*, *26*(3), 179-82. doi: 10.1136/jme.26.3.179

Emmelkamp, P. M. (2005). Technological innovations in clinical assessment and psychotherapy. *Psychotherapy & Psychosomatics*, *74*, 336–343. doi: 10.1159/000087780

Enqvist, L. (2005). The mistreatment of covariate interaction terms in linear model analyses of behavioural and evolutionary ecology studies. *Animal Behaviour*, *70*, 967971. doi: 10.1016/j.anbehav.2005.01.016

Eysenbach, G. (2005). The law of attrition. *Journal of Medical Internet Research*, *7*(1), e11. doi: 10.2196/jmir.7.1.e11

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, *16*(4), 363-70.

Fairburn, C. G., & Cooper, P. J. (1982). Self-induced vomiting and bulimia nervosa: an undetected problem. *British Medical Journal (Clinical Research Ed.)*, *284*(6323), 1153-5. doi: 10.1136/bmj.284.6323.1153

Fairburn, C. G., & Harrison, P. J. (2003). Eating disorders. *The Lancet*, *361*(9355), 407-416. doi: 10.1016/S0140-6736(03)12378-1

Farvolden, P., Denisoff, E., Selby, P., Bagby, R. M., & Rudy, L. (2005). Usage and longitudinal effectiveness of a web-based self-help cognitive behavioral therapy program for panic disorder. *Journal of Medical Internet Research*, *7*(1), e7. doi: 10.2196/jmir.7.1.e7

Fernandez-Aranda, F., Nunez, A., Martinez, C., Krug, I., Cappozzo, M., Carrard, I., ... Lam, T. (2009). Internet-based cognitive-behavioral therapy for bulimia nervosa: A controlled study. *Cyberpsychology and Behavior*, *12*(1), 37-41. doi: 10.1089/cpb.2008.0123

Foroushani, P., Schneider, J., & Assareh, N. (2011). Meta-review of the effectiveness of computerised cbt in treating depression. *BMC Psychiatry*, *11*(1), 131. doi: 10.1186/1471-244X-11-131

Furukawa, T. A., Katherine Shear, M., Barlow, D. H., Gorman, J. M., Woods, S. W., Money, R., ... Leucht, S. (2009). Evidence-based guidelines for interpretation of the panic disorder severity scale. *Depression and Anxiety*, *26*(10), 922-9. doi: 10.1002/da.20532

Gega, L., Marks, I., & Mataix-Cols, D. (2004). Computer-aided cbt self-help for anxiety and depressive disorders: Experience of a london clinic and future directions. *Journal of Clinical Psychology*, *60*(2), 147-57. doi: 10.1002/jclp.10241

GGZ Nederland. (2010). *Zorg op waarde geschat: Update. Sectorrapport ggz 2010* [Valued care: An update.] Amersfoort, the Netherlands: GGZ Nederland (Dutch Association of Health and Addiction Care).

Gloaguen, V., Cottraux, J., Cucherat, M., & Blackburn, I. M. (1998). A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders*, *49*(1), 59-72. doi: 10.1016/S0165-0327(97)00199-7

Griffiths, F., Lindenmeyer, A., Powell, J., Lowe, P., & Thorogood, M. (2006). Why are health care interventions delivered over the internet? a systematic review of the published literature. *Journal of Medical Internet Research*, *8*(2), e10. doi: 10.2196/jmir.8.2.e10

Griffiths, K. M., & Christensen, H. (2006). Review of randomised controlled trials of internet interventions for mental disorders and related conditions. *Clinical Psychologist*, *10*(1), 16-29. doi: 10.1080/13284200500378696

Hay, P. P., Bacaltchuk, J., Stefano, S., & Kashyap, P. (2009). Psychological treatments for bulimia nervosa and binging. *Cochrane Database of Systematic Reviews*, *7*(4), CD000562. doi: 10.1002/14651858.CD000562.pub3

Hayward, L., MacGregor, A. D., Peck, D. F., & Wilkes, P. (2007). The feasibility and effectiveness of computer-guided cbt (fearfighter) in a rural area. *Behavioural and Cognitive Psychotherapy*, *35*(04), 409-419. doi: 10.1017/S1352465807003670

Health & Safety Executive. (2001). *Work-related stress: A short guide.* Sudbury: HSE Books.

Helzer, J. E., & Robins, L. N. (1988). The Diagnostic Interview Schedule: Its development, evolution, and use. *Social Psychiatry and Psychiatric Epidemiology*, *23*(1), 6-16. doi: 10.1007/BF01788437

Hilderink, I., & Van 't Land, H. (2009). *Ggz in tabellen 2008* [Statistics of specialized mental healthcare, 2008]. Utrecht: Trimbos-instituut.

Hillis, S. L., & Woolson, R. F. (2002). Analysis of categorized data: Use of the odds ratio as a measure of association. In M. T. Tsuang & M. Tohen (Eds.), *Textbook in psychiatric epidemiology* (2nd ed., pp. 35–63). Hoboken, New Jersey: J. Wiley & Sons Inc. doi: 10.1002/0471234311.ch2

Hilvert-Bruce, Z., Rossouw, P. J., Wong, N., Sunderland, M., & Andrews, G. (2012). Adherence as a determinant of effectiveness of internet cognitive behavioural therapy for anxiety and depressive disorders. *Behaviour Research and Therapy*, *50*(78), 463-468. doi: 10.1016/j.brat.2012.04.001

Hoek, H. W., & van Hoeken, D. (2003). Review of the prevalence and incidence of eating disorders. *International Journal of Eating Disorders*, *34*(4), 383-96. doi: 10.1002/eat.10222

Holland, B. S., & DiPonzio Copenhaver, M. (1988). Improved bonferroni-type multiple testing procedures. *Psychological Bulletin*, *104*(1), 145-149. doi: 10.1037//0033-2909.104.1.145

Hopko, D. R., Lejuez, C. W., Ruggiero, K. J., & Eifert, G. H. (2003). Contemporary behavioral activation treatments for depression: Procedures, principles, and progress. *Clinical Psychology Review*, *23*(5), 699-717. doi: 10.1016/S0272-7358(03)00070-9

Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of event scale: A measure of subjective stress. *Psychosomatic Medicine*, *41*(3), 209-18.

Houck, P. R., Spiegel, D. A., Shear, M. K., & Rucci, P. (2002). Reliability of the self-report version of the PDSS-SR. *Depression and Anxiety*, *15*, 183–185. doi: 10.1002/da.10049

Hunsley, J., & Lee, C. M. (2007). Research-informed benchmarks for psychological treatments: Efficacy studies, effectiveness studies, and beyond. *Professional Psychology: Research and Practice*, *38*(1), 21-33. doi: 10.1037/0735-7028.38.1.21

Issakidis, C., & Andrews, G. (2004). Pretreatment attrition and dropout in an outpatient clinic for anxiety disorders. *Acta Psychiatrica Scandinavica*, *109*(6), 426-433. doi: 10.1111/j.1600-0047.2004.00264.x

Jacobi, C., Morris, L., Beckers, C., Bronisch-Holtze, J., Winter, J., Winzelberg, A. J., & Taylor, C. B. (2007). Maintenance of internet-based prevention: A randomized controlled trial. *International Journal of Eating Disorders*, *40*(2), 114-9. doi: 10.1002/eat.20344

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19. doi: 10.1037//0022-006X.59.1.12

Jager, J., Emmelkamp, P., & Lange, A. (2004). Die behandlung von panikstorungen uber das internet: Eine fallbeschreibung [Treatment of panic disorder over the internet: A case description]. *Verhaltenstherapie*, *14*, 169–236. doi: 10.1159/000080916

Jager, J., Emmelkamp, P. M. G., & Lange, A. (2005). *Behandeling van paniekstoornis met en zonder agorafobie via internet. projectverslag, zonmw projectnummer: 1055.0008.* [Treatment of panic disorder with and without agoraphobia via the internet: Project report zonmw project no: 1055.0008)] (Tech. Rep.). Amsterdam: University of Amsterdam).

Joiner, T. E., Steer, R. A., Brown, G., Beck, A. T., Pettit, J. W., & Rudd, M. D. (2003). Worst-point suicidal plans: A dimension of suicidality predictive of past suicide attempts and eventual death by suicide. *Behaviour Research and Therapy*, *41*, 1469–1480. doi: 10.1016/S0005-7967(03)00070-6

Kaldo, V., Larsen, H. C., & Andersson, G. (2004). Internet-based cognitive-behavioral self-help treatment of tinnitus: Clinical effectivness and predictors of outcome. *American Journal of Audiology*, *13*, 185-192. doi: 10.1044/1059-0889(2004/023)

Kaltenthaler, E., Brazier, J., De Nigris, E., Tumur, I., Ferriter, M., Beverley, C., ... Sutcliffe, P. (2006). Computerised cognitive behaviour therapy for depression and anxiety update: A systematic review and economic evaluation. *Health Technology Assessment*, *10*(33), iii, xi-xiv, 1-168.

Kessler, R. C., Chiu, W. T., Jin, R., Ruscio, A. M., Shear, K., & Walters, E. E. (2006). The epidemiology of panic attacks, panic disorder, and agoraphobia in the national comorbidity survey replication. *Archives of General Psychiatry*, *63*(4), 415-24. doi: 10.1001/archpsyc.63.4.415

Klein, B., Austin, D., Pier, C., Kiropoulos, L., Shandley, K., Mitchell, J., ... Ciechomski, L. (2009). Internet-based treatment for panic disorder: Does frequency of therapist contact make a difference? *Cognitive Behaviour Therapy*, 1-14. doi: 10.1080/16506070802561132

Klein, B., & Richards, J. C. (2001). A brief internet-based treatment for panic disorder. *Behavioural and Cognitive Psychotherapy*, *29*, 113-117. doi: 10.1017/S1352465801001138

Klein, B., Richards, J. C., & Austin, D. W. (2006). Efficacy of internet therapy for panic disorder. *Journal of Behavioural Therapy & Experimental Psychiatry*, *37*(3), 213-38. doi: 10.1016/j.jbtep.2005.07.001

Knaevelsrud, C., Jager, J., & Maercker, A. (2004). Internet-psychotherapie: Wirksamkeit und besonderheiten der therapeutischen beziehung [Online-therapy: Effectiveness and peculiarities of the therapeutic relationship]. *Verhaltenstherapie*, *14*, 174-183. doi: 10.1159/000080913

Knaevelsrud, C., & Maercker, A. (2006). Does the quality of the working alliance predict treatment outcome in online psychotherapy for traumatized patients? *Journal of Medical Internet research*, *8*, e31. doi: 10.2196/jmir.8.4.e31

Knaevelsrud, C., & Maercker, A. (2007). Internet-based treatment for ptsd reduces distress and facilitates the development of a strong therapeutic alliance: A randomized controlled clinical trial. *BMC Psychiatry*, *7*, 13. doi: 10.1186/1471-244X-7-13

Knaevelsrud, C., & Maercker, A. (2009). Long-term effects of an internet-based treatment for posttraumatic stress. *Cognitive Behaviour Therapy*, *39*, 72–77. doi: 10.1080/16506070902999935

Lancet Global Mental Health Group. (2007). Scale up services for mental disorders: A call for action. *The Lancet*, *370*(9594), 1241-1252. doi: 10.1016/S0140-6736(07)61242-2

Lange, A. (2006). *Gedragsverandering in gezinnen. Cognitieve gedrags- en systeemtherapie.* [Behaviour change within the family] (8th ed.). Groningen: Martinus Nijhoff.

Lange, A., De Vries, M., Gest, A., & Van Oostendorp, E. (1994). A self-management program

for bulimia nervosa: The elements, the rationales, and a case study. *Eating Disorders: The Journal of Treatment & Prevention*, *2*(4), 329 - 340. doi: 10.1080/10640269408249131

Lange, A., Richard, R., Gest, A., de Vries, M., & Lodder, L. (1998). The effects of positive self-instruction; a controlled trial. *Cognitive Therapy and Research*, *22*, 225–236. doi: 10.1023/A:1018740725281

Lange, A., Richard, R., Kiestra, J., & van Oostendorp, E. (1997). Cognitive treatment through positive self-verbalization; a multiple case study. *Behavioural and Cognitive Psychotherapy*, *25*, 161–171. doi: 10.1017/S1352465800018361

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 standardized treatment of posttraumatic stress through the internet. *Journal of Consulting and Clinical Psychology*, *71*, 901–909. doi: 10.1037/0022-006X.71.5.901

Lange, A., & Ruwaard, J. (2010). Ethical dilemmas in online research and treatment of sexually abused adolescents. *Journal of Medical Internet Research*, *12*(5), e58. doi: 10.2196/jmir.1455

Lange, A., Ruwaard, J., Schrieken, B., de Haas, S., Broeksteeg, J., Renteria-Agirre, A., & Kuyper, A. (2011). Online behandeling van seksueel misbruikte adolescenten: Ethische dilemma's en resultaten van een geprotocolleerde trial [Online treatment of sexually abused adolescents: Ethical dilemmas en results of a controlled trial]. *Directieve Therapie*, *31*(3), 259-279.

Lange, A., Schrieken, B., Blankers, M., van de Ven, J.-P., & Slot, M. (2000). Constructie en validatie van de Gewaarwordingenlijst (GL); een hulpmiddel bij het signaleren van een verhoogde kans op psychosen [Construction of the Screening Device of Psychotic Disorder SDPD; an instrument for detecting a heightened likelyhood of psychosis]. *Directieve Therapie*, *20*, 162–173. doi: 10.1007/BF03060234

Lange, A., van de Ven, J.-P., & Schrieken, B. (2003). Interapy. Treatment of post-traumatic stress via the internet. *Cognitive Behaviour Therapy*, *32*, 110–124. doi: 10.1080/16506070310004935

Lange, A., van de Ven, J. P., Schrieken, B., & Emmelkamp, P. M. (2001). Interapy. Treatment of posttraumatic stress through the internet: A controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry*, *32*(2), 73-90. doi: 10.1016/S0005-7916(01)00023-4

Lange, A., Van de Ven, J. P., Schrieken, B., & Schoutrop, M. J. A. (2002). Interapy: A model for therapeutic writing through the internet. In S. J. Lepore & J. M. Smith (Eds.), *The writing cure: How expressive writing promotes health and emotional well being.* Washington, DC: American Psychiatric Association. doi: 10.1037/10451-011

Lange, A., van de Ven, J. P., Schrieken, B. A., Bredeweg, B., & Emmelkamp, P. M. (2000). Internet-mediated, protocol-driven treatment of psychological dysfunction. *Journal of Telemedecine and Telecare*, *6*(1), 15-21. doi: 10.1258/1357633001933880

LaVange, L. M., & Koch, G. G. (2006). Rank score tests. *Circulation*, *114*(23), 2528-33. doi: 10.1161/CIRCULATIONAHA.106.613638

Leonard, K., & Masatu, M. C. (2006). Outpatient process quality evaluation and the Hawthorne effect. *Social Science and Medicine*, *63*(9), 2330-2340. doi: 10.1016/j.socscimed.2006.06.003

Lepore, S. J., & Smyth, J. M. (2002). *The writing cure; how expressive writing promotes health and emotional well being*. Washington: APA books.

Levi, L., & Levi, I. (1995). *Guidance on work-related stress. Spice of life, or kiss of death?* Luxembourg: Office for Official Publications of the European Communities. Retrieved

October 20, 2006, from `http://ec.europa.eu/`.

Lewinsohn, P. M., Biglan, A., & Zeiss, A. M. (1976). Behavioral treatment of depression. In P. O. Davidson (Ed.), *The behavioral management of anxiety, depression and pain.* (p. 91-146). New York: Brunner/Mazel.

Ljotsson, B., Lundin, C., Mitsell, K., Carlbring, P., Ramklint, M., & Ghaderi, A. (2007). Remote treatment of bulimia nervosa and binge eating disorder: A randomized trial of internet-assisted cognitive behavioural therapy. *Behaviour Research and Therapy*, *45*(4), 649-61. doi: 10.1016/j.brat.2006.06.010

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales*. Sydney, Australia: Psychology Foundation.

Magruder, K. M., & Calderone, G. E. (2000). Public health consequences of different thresholds for the diagnosis of mental disorders. *Comprehensive Psychiatry*, *41*(2 Suppl 1), 14-8. doi: 10.1016/S0010-440X(00)80003-6

Maindonald, J. H., & Braun, J. (2007). *Data analysis and graphics using R*. Cambridge, UK: Cambridge University Press.

Marks, I. M., Cavanagh, K., & Gega, L. (2007). *Hands-on help: Computer-aided psychotherapy (maudsley monograph 49)*. Hove/New York: Psychology Press.

Marks, I. M., Kenwright, M., McDonough, M., Whittaker, M., & Mataix-Cols, D. (2004). Saving clinicians' time by delegating routine aspects of therapy to a computer: A randomized controlled trial in phobia/panic disorder. *Psychological Medicine*, *34*(1), 9-17. doi: 10.1017/S003329170300878X

Marks, I. M., Mataix-Cols, D., Kenwright, M., Cameron, R., Hirsch, S., & Gega, L. (2003). Pragmatic evaluation of computer-aided self-help for anxiety and depression. *British Journal of Psychiatry*, *183*, 57-65. doi: 10.1192/bjp.183.1.57

Maslach, C., Jackson, S. E., & Leiter, F. (1996). *Maslach Burnout Inventory: Manual (3rd ed.)*. Palo Alto: Consulting Psychologists Press.

McKay, M., Davis, M., & Fanning, P. (1997). *Thoughts & feelings: Taking control of your moods and life.* Oakland, CA: New Harbinger Publications.

Melville, K. M., Casey, L. M., & Kavanagh, D. J. (2010). Dropout from internet-based treatment for psychological disorders. *British Journal of Clinical Psychology*, *49*(4), 455-471. doi: 10.1348/014466509X472138

Messer, S. B., & Wampold, B. E. (2002). Let's face facts: Common factors are more potent than specific therapy ingredients. *Clinical Psychology: Science and Practice*, *9*(1), 2125. doi: 10.1093/clipsy.9.1.21

Minami, T., Wampold, B. E., Serlin, R. C., Hamilton, E. G., Brown, G. S. J., & Kircher, J. C. (2008). Benchmarking the effectiveness of psychotherapy treatment for adult depression in a managed care environment: A preliminary study. *Journal of Consulting and Clinical Psychology*, *76*(1), 116-124. doi: 10.1037/0022-006X.76.1.116

Molenberghs, G., Thijs, H., Jansen, I., Beunckens, C., Kenward, M. G., Mallinckrodt, C., & Carroll, R. J. (2004). Analyzing incomplete longitudinal clinical trial data. *Biostatistics*, *5*(3), 445-64. doi: 10.1093/biostatistics/5.3.445

Mond, J. M., Hay, P. J., Rodgers, B., & Owen, C. (2007). Health service utilization for eating disorders: Findings from a community-based study. *International Journal of Eating Disorders*, *40*(5), 399-408. doi: 10.1002/eat.20382

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. (2004a). Temporal stability of the eating disorder examination questionnaire. *International Journal of Eating Disorders*, *36*(2), 195-203. doi: 10.1002/eat.20017

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. V. (2004b). Validity of the eating disorder examination questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour Research and Therapy*, *42*(5), 551-567. doi: 10.1016/S0005-7967(03)00161-X

Mond, J. M., Myers, T. C., Crosby, R. D., Hay, P. J., & Mitchell, J. E. (2010). Bulimic eating disorders in primary care: Hidden morbidity still? *Journal of Clinical Psychology in Medical Settings*, *17*(1), 56-63. doi: 10.1007/s10880-009-9180-9

Morin, C. M. (1993). *Insomnia; psychological assessment and management.* New York: Guilford Presss.

Nathan, P. E., Stuart, S. P., & Dolan, S. L. (2000). Research on psychotherapy efficacy and effectiveness: Between Scylla and Charybdis? *Psychological Bulletin*, *126*(6), 964-981. doi: 10.1037/0033-2909.126.6.964

National Institute for Clinical Excellence. (2004a). *Depression: Management of depression in primary and secondary care. Clinical guideline 23.* London: NICE.

National Institute for Clinical Excellence. (2004b). *Eating disorders: Core interventions in the treatment and management of anorexia nervosa, bulimia nervosa and related eating disorders*. London: NICE.

Nauta, H., Hospers, H., Kok, G., & Jansen, A. (2000). A comparison between a cognitive and behavioral treatment for obese binge eaters and obese non-binge-eaters. *Behavior Therapy*, *31*(3), 441-461. doi: 10.1016/S0005-7894(00)80024-6

Neal, L. A., Busuttil, W., Rollins, J., Herepath, R., Strike, P., & Turnbull, G. (1994). Convergent validity of measures of post-traumatic stress disorder in a mixed military and civilian population. *Journal of Traumatic Stress*, *7*(3), 447-55. doi: 10.1007/BF02102789

Nevonen, L., Mark, M., Levin, B., Lindström, M., & Paulson-Karlsson, G. (2007). Evaluation of a new internet-based self-help guide for patients with bulimic symptoms in sweden. *Nordic Journal of Psychiatry*, *60*(6), 463-468. doi: 10.1080/08039480601021993

Newman, M. G. (2004). Technology in psychotherapy: An introduction. *Journal of Clinical Psychology*, *60*, 141-145. doi: 10.1002/jclp.10240

Newman, M. G., Holmes, M., Zuellig, A. R., Kachin, K. E., & Behar, E. (2006). The reliability and validity of the panic disorder self-report: A new diagnostic screening measure of panic disorder. *Psychol Assess*, *18*(1), 49-61. doi: 10.1037/1040-3590.18.1.49

Nieuwenhuijsen, K., de Boer, A. G., Verbeek, J. H., Blonk, R. W., & van Dijk, F. J. (2003). The Depression Anxiety Stress Scales (DASS): Detecting anxiety disorder and depression in employees absent from work because of mental health problems. *Occupational Environment*, *60*, 77–82. doi: 10.1136/oem.60.suppl_1.i77

Nijenhuis, E. R. S., Spinhoven, P., van Dyck, R., van der Hart, O., & Vanderlinden, J. (1997). The development of the somatoform dissociation questionnaire (SDQ-5) as a screening instrument for dissociative disorders. *Acta Psychiatrica Scandinavica*, *96*, 311-318. doi: 10.1111/j.1600-0447.1997.tb09922.x

Nijenhuis, E. R. S., Spinhoven, P., van Dyck, R., van der Hart, O., & Vanderlinden, J. (1998). Psychometric characteristics of the somatoform dissociation questionnaire: A replication study. *Psychotherapy & Psychosomatics*, *67*(1), 17-23. doi: 10.1159/000012254

Oliver, J. M., & Simmons, M. E. (1984). Depression as measured by the DSM-III and the Beck Depression Inventory in an unselected adult population. *Journal of Consulting and Clinical Psychology*, *52*(5), 892-898. doi: 10.1037/0022-006X.52.5.892

Öst, L. G. (1987). Applied relaxation: Description of a coping technique and review of controlled studies. *Behaviour Research and Therapy*, *25*(5), 397-409. doi:

10.1080/16506078809456264

Palmqvist, B., Carlbring, P., & Andersson, G. (2007). Internet-delivered treatments with or without therapist input: Does the therapist factor have implications for efficacy and cost? *Expert Review of Pharmacoeconomics & Outcomes Research*, *7*(3), 291-297. doi: 10.1586/14737167.7.3.291

Pampallona, S., Bollini, P., Tibaldi, G., Kupelnick, B., & Munizza, C. (2004). Combined pharmacotherapy and psychological treatment for depression: A systematic review. *Archives of General Psychiatry*, *61*(7), 714-719.

Paoli, P., & Merllie, D. (2001). *Third european survey on working conditions*. Luxembourg: European Foundation for the Improvement of Living and Working Conditions.

Paxton, S. J., McLean, S. A., Gollings, E. K., Faulkner, C., & Wertheim, E. H. (2007). Comparison of face-to-face and internet interventions for body image and eating problems in adult women: An RCT. *International Journal of Eating Disorders*, *40*(8), 692-704. doi: 10.1002/eat.20446

Perkins, S. J., Murphy Rebecca, R. M., Schmidt Ulrike, U. S., & Williams, C. (2006). Self-help and guided self-help for eating disorders. *Cochrane Database of Systematic Reviews*(3 Art. No.: CD004191). doi: 10.1002/14651858.CD004191.pub2

Peterson, C. B., Crosby, R. D., Wonderlich, S. A., Joiner, T., Crow, S. J., Mitchell, J. E., ... Grange, D. l. (2007). Psychometric properties of the Eating Disorder Examination-Questionnaire: Factor structure and internal consistency. *International Journal of Eating Disorders*, *40*(4), 386-389. doi: 10.1002/eat.20373

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-Plus*. New York: Springer.

Postel, M. G., ter Huurne, E., de Haan, H. A., & de Jong, C. A. (2009). Alcoholdebaas.nl: Drie jaar online hulpverlening voor probleemdrinkers. [Alcoholdebaas.nl: Three years of online help for problem drinkers.] *Maandblad Geestelijke volksgezondheid*, *64*(5).

Posternak, M. A., & Miller, I. (2001). Untreated short-term course of major depression: A meta-analysis of outcomes from studies using wait-list control groups. *Journal of Affective Disorders*, *66*(2-3), 139-146. doi: 10.1016/S0165-0327(00)00304-9

Pouwer, F., Snoek, F., van der Ploeg, H., Ader, H., & Heine, R. (2000). The Well-being Questionnaire: Evidence for a three factor structure with 12 items (W-BQ12). *Psychological Medicine*, *30*, 463-466. doi: 10.1017/S0033291700001719

Probst, M., Pieters, G., & Vanderlinden, J. (2008). Evaluation of body experience questionnaires in eating disorders in female patients (AN/BN) and nonclinical participants. *International Journal of Eating Disorders*, *41*(7), 657-665. doi: 10.1002/eat.20531

Probst, M., van Coppenolle, H., & Vandereycken, W. (1998). De LichaamsAttitudeVragenlijst: Validering en normering. [The Body Attitude Test: Validation and norm data.] In J. Simons (Ed.), *Actuele themata uit de psychomotorische therapie.* (p. 79-90). Leuven: Arco.

Probst, M., Vandereycken, W., Van Coppenolle, H., & Vanderlinden, J. (1995). The Body Attitude Test for patients with an eating disorder: Psychometric characteristics of a new questionnaire. *Eating Disorders*, *3*(2), 133 - 144. doi: 10.1080/10640269508249156

Proudfoot, J., Ryden, C., Everitt, B., Shapiro, D. A., Goldberg, D., Mann, A., ... Gray, J. A. (2004). Clinical efficacy of computerised cognitive-behavioural therapy for anxiety and depression in primary care: Randomised controlled trial. *British Journal of Psychiatry*, *185*, 46-54. doi: 10.1192/bjp.185.1.46

Proudfoot, J., Swain, S., Widmer, S., Watkins, E., Goldberg, D., Marks, I. M., ... Gray, J. A.

(2003). The development and beta-test of a computertherapy program for anxiety and depression: hurdles and lessons. *Computers in Human Behavior*, *19*, 277-289. doi: 10.1016/S0747-5632(02)00062-6

R Development Core Team. (2008). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `http://www.R-project.org`

Reger, M. A., & Gahm, G. A. (2009). A meta-analysis of the effects of internet- and computer-based cognitive-behavioral treatments for anxiety. *Journal of Clinical Psychology*, *65*(1), 53-75. doi: 10.1002/jclp.20536

Riper, H., Kramer, J., Conijn, B., Smit, F., Schippers, G., & Cuijpers, P. (2009). Translating effective web-based self-help for problem drinking into the real world. *Alcoholism, Clinical and Experimental Research*, *33*, 1401-1408. doi: 10.1111/j.1530-0277.2009.00970.x

Riper, H., Smit, F., van der Zanden, R., Conijn, B., Kramer, J., & Mutsaers, K. (2007). *E-mental health: High tech, high touch, high trust. Programmeringsstudie e-mental health in opdracht van het ministerie van VWS.* [High tech, high touch, high trust: Programming study e-mental health.] Utrecht: Trimbos-instituut.

Ritterband, L. M., Ardalan, K., Thorndike, F. P., Magee, J. C., Saylor, D. K., Cox, D. J., ... Borowitz, S. M. (2008). Real world use of an internet intervention for pediatric encopresis. *Journal of Medical Internet Research*, *10*(2), e16. doi: 10.2196/jmir.1081

Robey, R. R. (2004). Reporting point and interval estimates of effect-size for planned contrasts: Fixed within effect analyses of variance. *Journal of Fluency Disorders*, *29*, 307–341. doi: 10.1016/j.jfludis.2004.10.005

Robinson, P., & Serfaty, M. (2008). Getting better byte by byte: A pilot randomised controlled trial of email therapy for bulimia nervosa and binge eating disorder. *European Eating Disorders Review*, *16*(2), 84-93. doi: 10.1002/erv.818

Roth, A., & Fonagy, P. (2005). *What works for whom? a critical review of psychotherapy research.* (2nd ed.). New York: Guilford Press.

Ruwaard, J., Broeksteeg, J., Schrieken, B., Emmelkamp, P., & Lange, A. (2010). Web-based therapist-assisted cognitive behavioral treatment of panic symptoms: A randomized controlled trial with a three-year follow-up. *Journal of Anxiety Disorders*, *24*, 387–396. doi: 10.1016/j.janxdis.2010.01.010

Ruwaard, J., Lange, A., Bouwman, M., Broeksteeg, J., & Schrieken, B. (2007). E-mailed standardized cognitive behavioural treatment of work-related stress: A randomized controlled trial. *Cognitive Behaviour Therapy*, *36*(3), 179-192. doi: 10.1080/16506070701381863

Ruwaard, J., Lange, A., Broeksteeg, J., Renteria-Agirre, A., Schrieken, B., Dolan, C. V., & Emmelkamp, P. (2012). Online cognitivebehavioural treatment of bulimic symptoms: A randomized controlled trial. *Clinical Psychology & Psychotherapy*, epub ahead of print. doi: 10.1002/cpp.1767

Ruwaard, J., Lange, A., Schrieken, B., Dolan, C., & Emmelkamp, P. (2012). The effectiveness of online cbt in routine clinical practice. *PLoS ONE*, *7*(7), e40089. doi: doi:10.1371/journal.pone.0040089

Ruwaard, J., Schrieken, B., Schrijver, M., Broeksteeg, J., Dekker, J., Vermeulen, H., & Lange, A. (2009). Standardized web-based cognitive behavioural therapy of mild to moderate depression: A randomized controlled trial with a long-term follow-up. *Cognitive Behaviour Therapy*, *38*(4), 206-221. doi: 10.1080/16506070802408086

Sánchez-Ortiz, V. C., Munro, C., Stahl, D., House, J., Startup, H., Treasure, J., ... Schmidt, U.

(2010). A randomized controlled trial of internet-based cognitive-behavioural therapy for bulimia nervosa or related disorders in a student population. *Psychological Medicine*, 1-11. doi: 10.1017/S0033291710000711

Sansone, R. A., Wiederman, M. W., & Sansone, L. A. (1998). The Self-Harm Inventory (SHI): Development of a scale for identifying self-destructive behaviors and borderline personality disorder. *Journal of Clinical Psychology*, *54*(7), 973-83. doi: 10.1002/(SICI)1097-4679(199811)54:7<973::AID-JCLP11>3.0.CO;2-H

Schaufeli, W. B., & Kompier, M. A. J. (2001). Managing job stress in the Netherlands. *International Journal of Stress Management*, *8*, 15–34.

Schaufeli, W. B., & van Dierendonck, D. (2000). *Utrechtse Burnout Schaal: Handleiding.* [The Utrecht Burnout Scale: Manual.] Lisse: Swets & Zeitlinger.

Schmitz, N., Hartkamp, N., Brinschwitz, C., Michalek, S., & Tress, W. (2000). Comparison of the standard and the computerized versions of the Symptom Check List (SCL-90-R): A randomized trial. *Acta Psychiatrica Scandinavica*, *102*(2), 147-52. doi: 10.1034/j.1600-0447.2000.102002147.x

Schmitz, N., Kruse, J., Heckrath, C., Alberti, L., & Tress, W. (1999). Diagnosing mental disorders in primary care: The General Health Questionnaire (GHQ) and the Symptom Check List (SCL-90-R) as screening instruments. *Social Psychiatry and Psychiatric Epidemiology*, *34*(7), 360-6. doi: 10.1007/s001270050156

Seligman, M. E. (1995). The effectiveness of psychotherapy: The consumer reports study. *American Psychologist, December 1995*, *50*(12), 965-974. doi: 10.1037/0003-066X.50.12.965

Shafran, R., Clark, D. M., Fairburn, C. G., Arntz, A., Barlow, D. H., Ehlers, A., ... Wilson, G. T. (2009). Mind the gap: Improving the dissemination of CBT. *Behaviour Research and Therapy*, *47*(11), 902-909. doi: 10.1016/j.brat.2009.07.003

Shandley, K., Austin, D. W., Klein, B., Pier, C., Schattner, P., Pierce, D., & Wade, V. (2008). Therapist-assisted, internet-based treatment for panic disorder: Can general practitioners achieve comparable patient outcomes to psychologists? *Journal of Medical Internet Research*, *10*(2), e14. doi: 10.2196/jmir.1033

Sharp, D. M., Power, K. G., & Swanson, V. (2000). Reducing therapist contact in cognitive behaviour therapy for panic disorder and agoraphobia in primary care: Global measures of outcome in a randomised controlled trial. *British Journal of General Practice*, *50*(461), 963-8.

Shear, M. K., Rucci, P., Williams, J., Frank, E., Grochocinski, V., Bilt, J. V., ... Wang, T. (2001). Reliability and validity of the Panic Disorder Severity Scale: Replication and extension. *Journal of Psychiatric Research*, *35*, 293–296. doi: 10.1016/S0022-3956(01)00028-0

Sparks, K., Faragher, B., & Cooper, C. L. (2001). Well-being and occupational health in the 21st century workplace. *Journal of Occupational and Organizational Psychology*, *74*, 489–509. doi: 10.1348/096317901167497

Speckens, A. E., Spinhoven, P., Sloekers, P. P., Bolk, J. H., & van Hemert, A. M. (1996). A validation study of the Whitely Index, the Illness Attitude Scales, and the Somatosensory Amplification Scale in general medical and general practice patients. *Journal of Psychosomatic Research*, *40*(1), 95-104. doi: 10.1016/0022-3999(95)00561-7

Spek, V., Cuijpers, P., Nyklicek, I., Riper, H., Keyzer, J., & Pop, V. (2007). Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: A meta-analysis. *Psychological Medicine*, *37*, 319-328. doi: 10.1017/S0033291706008944

Sunderland, M., Wong, N., Hilvert-Bruce, Z., & Andrews, G. (2012). Investigating trajectories

of change in psychological distress amongst patients with depression and generalised anxiety disorder treated with internet cognitive behavioural therapy. *Behaviour Research and Therapy*, *50*(6), 374-380. doi: 10.1016/j.brat.2012.03.005,

Tate, D. F., Finkelstein, E. A., Khavjou, O., & Gustafson, A. (2009). Cost effectiveness of internet interventions: Review and recommendations. *Annals of Behavioral Medicine*, *38*(1), 40-45. doi: 10.1007/s12160-009-9131-6

Taylor, C. B., Bryson, S., Luce, K. H., Cunning, D., Doyle, A. C., Abascal, L. B., ... Wilfley, D. E. (2006). Prevention of eating disorders in at-risk college-age women. *Archives of General Psychiatry*, *63*(8), 881-8. doi: 10.1001/archpsyc.63.8.881

Titov, N., Andrews, G., Kemp, A., & Robinson, E. (2010). Characteristics of adults with anxiety or depression treated at an internet clinic: Comparison with a national survey and an outpatient clinic. *PLoS One*, *5*(5), e10885. doi: 10.1371/journal.pone.0010885

Ustun, T. B., Ayuso-Mateos, J. L., Chatterji, S., Mathers, C., & Murray, C. J. (2004). Global burden of depressive disorders in the year 2000. *British Journal of Psychiatry*, *184*, 386-92. doi: 10.1192/bjp.184.5.386

Vallejo, M. A., Jordan, C. M., Diaz, M. I., Comeche, M. I., & Ortega, J. (2007). Psychological assessment via the internet: A reliability and validity study of online (vs paper-and-pencil) versions of the General Health Questionnaire-28 (GHQ-28) and the Symptoms Check-List-90-Revised (SCL-90-R). *Journal of Medical Internet Research*, *9*(1), e2. doi: 10.2196/jmir.9.1.e2

van der Klink, J. J. L., Blonk, R. W. B., Schene, A. H., & van Dijk, F. J. H. (2001). The benefits of interventions for work-related stress. *American Journal of Public Health*, *91*, 270–276. doi: 10.2105/AJPH.91.2.270

Vanderlinden, J. (2002). *Boulimie en eetbuien overwinnen. Een gids voor patiënt, gezin en hulpverlener.* [Overcoming bulimia and binge eating. A guide for the patient, the family and the counselor.] Tielt: Lannoo.

Vanderlinden, J., Pieters, G., Probst, M., & Norré, J. (2004). *Behandelprotocol bij boulimia nervosa. Therapeutenboek.* [Treatment protocol for boulimia nervosa. Therapist manual.] Houten/Diegem: Bohn Stafleu van Loghum.

Vanderlinden, J., van Dyck, R., Vandereycken, W., & Vertommen, H. (1991). Dissociative experiences in the general population in the netherlands and belgium: A study with the Dissociative Questionnaire (DIS-Q). *Dissociation*, *14*(4), 180-184.

van der Meer, A., & Burgerhout, K. (2004). *Nederlandse versie van de panic disorder severity scale.* [Dutch version of the Panic Disorder Severity Scale.] Doctoraal scriptie, Universiteit van Amsterdam.

van der Velden, K., Hoogduin, K., & Lange, A. (2010). *Directieve therapie* [Directive therapy]. Amsterdam: Boom/SUN.

van Doornen, L. (2001). Psychofysiologie van burnout [psychophysiology of burnout]. In C. A. L. Hoogduin, W. B. Schaufeli, C. P. D. R. Schaap, & A. B. Bakker (Eds.), *Behandelingsstrategieën bij burnout* (pp. 40–50). Houten: Bohn Stafleu van Loghum.

Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. New York: Springer-Verlag.

W. H. O. World Mental Health Survey Consortium. (2004). Prevalence, severity, and unmet need for treatment of mental disorders in the world health organization world mental health surveys. *Journal of the American Medical Association*, *291*(21), 2581-2590. doi: 10.1001/jama.291.21.2581

Wagner, B., Knaevelsrud, C., & Maercker, A. (2006). Internet-based cognitive-behavioral

therapy for complicated grief: A randomized controlled trial. *Death Studies*, *30*(5), 429-53. doi: 10.1080/07481180600614385

Wagner, B., & Maercker, A. (2007). A 1.5-year follow-up of an internet-based intervention for complicated grief. *Journal of Traumatic Stress*, *20*(4), 625-629. doi: 10.1002/jts.20230

Wagner, B., Schulz, W., & Knaevelsrud, C. (2012). Efficacy of an internet-based intervention for posttraumatic stress disorder in iraq: A pilot study. *Psychiatry Research*, *195*(1), 85-88. doi: 10.1016/j.psychres.2011.07.026

Wampold, B. E., & Serlin, R. C. (2000). The consequence of ignoring a nested factor on measures of effect size in analysis of variance. *Psychological Methods*, *5*(4), 425-433. doi: 10.1037/1082-989X.5.4.425

Wang, P. S., Simon, G., & Kessler, R. C. (2003). The economic burden of depression and the cost-effectiveness of treatment. *International Journal of Methods in Psychiatric Research*, *12*(1), 22-33. doi: 10.1002/mpr.139

Wantland, D. J., Portillo, C. J., Holzemer, W. L., Slaughter, R., & McGhee, E. M. (2004). The effectiveness of web-based vs. non-web-based interventions: A meta-analysis of behavioural change outcomes. *Journal of Medical Internet Research*, *6*, e40. doi: 10.2196/jmir.6.4.e40

Weiss, D. S., & Marmar, C. R. (1996). The Impact of Event Scale Revised. In J. P. Wilson & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (p. 399-411). New York: Guilford. doi: 10.1007/978-0-387-70990-1_10

Westbrook, D., & Kirk, J. (2005). The clinical effectiveness of cognitive behaviour therapy: outcome for a large sample of adults treated in routine practice. *Behaviour Research and Therapy*, *43*(10), 1243-1261. doi: 10.1016/j.brat.2004.09.006

Westen, D., & Morrison, K. (2001). A multidimensional meta analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *69*(6), 875-899. doi: 10.1037/0022-006X.69.6.875

Wilson, G. T. (1998). Manual-based treatment and clinical practice. *Clinical Psychology and Scientific Practice*, *5*, 363-375. doi: 10.1111/j.1468-2850.1998.tb00156.x

Wilson, G. T. (2005). Psychological treatment of eating disorders. *Annual Review of Clinical Psychology*, *1*, 439-465. doi: 10.1146/annurev.clinpsy.1.102803.144250

World Health Organization. (2001). *Atlas: Mental health resources in the world 2001*. Geneva: WHO.

Zetterqvist, K., Maanmies, J., Ström, L., & Andersson, G. (2003). Randomized controlled trial of Internet-based stress management. *Cognitive Behaviour Therapy*, *32*, 151–160. doi: 10.1080/16506070302316

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavia*, *67*, 361–370. doi: 10.1111/j.1600-0447.1983.tb09716.x

Dankwoord

Alleen mijn naam staat op de voorkant van dit boek. En daar is heus wat voor te zeggen, gezien de uren die ik er in heb zitten. Maar aan dit proefschrift droegen velen bij. Ik vind het belangrijk hun namen hier te noemen. Don't try this at home. Niet zonder de hulp die ik heb gehad.

Mijn dank gaat allereerst uit naar mijn promotoren, prof. dr. Alfred Lange, prof. dr. Conor Dolan en prof. dr. Paul Emmelkamp. Freddy, ik sta bij je in het krijt. Dit alles was jouw idee. Het was een eer om door jou te worden opgeleid, vooral omdat ik één van je laatste promovendi zal zijn. Ik heb grote bewondering voor je vakmanschap, je scherpzinnigheid, je creativiteit en je geduld. Je was een goede meester. Ik maak een buiging en ga als vriend met je verder. Laten we snel weer eens skypen. Conor, jou dank ik voor je vertrouwen, je relativisme en je common sense. Alleen de beste methodologen brengen de statistiek terug tot aannamen die om begrijpelijke redenen wel of niet acceptabel zijn. Ik prijs me gelukkig dat jij tot die groep behoort. Cruciaal waren verder je correcties op mijn Engels, waarvoor ik je ook heel dankbaar ben. Paul, jij was meer op de achtergrond aanwezig, maar daarom niet minder belangrijk voor het welslagen van dit project. Je vermogen om snel en kort de meest wezenlijke zaken te benoemen is fenomenaal. Ik heb veel aan je feedback gehad.

Beste oud-collega's bij Interapy, aan jullie draag ik het proefschrift op. Jullie schreven de protocollen, screenden en behandelden, programmeerden mee aan het systeem, en regelden de financiering en de gang van zaken op kantoor. Dit proefschrift is voor jullie. Joyce Aalders, Tabe Aalders, Leire Abels-Reguera, Sofija Andjelkovic, Josje Arnoldi, Maartje Bakker, Mirte Bakker, Annelies van de Berg, Winneke Blees, Melissa Blommers, Arjan van Bentem, Stijn Bornewasser, Joris de Boer, Kyra de Boer, Lisette de Boer, Manon Bouwman, Janneke Broeksteeg, Karin Brunner, Wouter Diesveld, Elke Ederveen, Rob Faltin, Aagje Gest, Kiffin Gish, Geertje Hagendoorn, Ruud

Hermans, Marike Hoex, Pieter van Hoogstraten, Herry Hubert, Milena Hudcovicova, Lotte Isopuro, Janet Jager, Niels Kakes, Jacqueline Keers, Phebe de Klerk, Jasper Koolhaas, Hidde Kuiper, Jeroen Lijmer, Alda Losse, Bé van Lotringen, Marcelino Lopez, Helle Loonen Sinkbaek, Sumit Mehra, Claartje Michels, Bob Navis, Marije Nuijens, Suzanne Maris, Jamal Ouariachi, Anne Peerdeman, Joline Peters, Hinke Post, Denise Ras, Aitziber Renteria-Agirre, Allard Rietberg, Deirdre Rietdijk, Len van der Rijst, Carlijn Sanders, Rinke Scheijde, Bart Schrieken, Anne Geert Smit, Ignance Suy, Menno Schrijver, Claudi Seinhorst, Arunya Sivarajah, Jacobine van der Smagt, Mischa ter Smitten, Lisanne Snel, Konstantin Starojitski, Luuk Stegmann, Sam van Tienhoven, Nadia Tsaganá, Jean-Pierre van de Ven, Dominique Vijverberg, Joke Vroeginderweij, Herman Wassink, Ramon Weel, Marije Wegman, Carolien Wijnker, Sophie Zeyl en iedereen die ik hier dan toch nog vergeet: duizendmaal dank!

Ik dank ook de studenten voor hun hulp bij het onderzoek (voor zover ik ze hiervoor nog niet genoemd heb): Christel Adriaens, Marian Bond, Yonina Bartelse, Petra de Bruin, Barbera Cupido, Parco Faber, Marieke Holterman, Janine van Hinsberg, Willemein Huisman, Romy Koch, Pepijn Koopman, Annemiek Kolkert, Anne-Hilde Krips, Eva Lith, Jenean van Muylwijk, Bob Navis, Elise Peetoom, Nel van Poelgeest, Cornelie Renckens, Annemieke Schilder, Marjanneke Slinger, Sabine Stellingwerf, Nellie Stiny, Femke Truijens, Jos Verstraten, Marjolein Vos, Thea Wegner, Geeske Wicherds, Asja Zandbergen en Anna Zegel.

Mijn dank gaat uit naar de cliënten, de deelnemers aan de studies, voor het ter beschikking stellen van hun gegevens. Ik realiseer me dat het gaat om de mens achter het getal. Ik dank verder het Fonds Psychische Gezondheid, het Innovatiefonds Zorgverzekeraars, Interapy Nederland BV, Mentrum, Altrecht, GGZ Drenthe, Emergis, de Gelderse Roos, GGNet, Orbis en ZonMw voor de financiële en logistieke bijdragen aan het onderzoek.

Ik had dit proefschrift niet geschreven als Volkert Balk mij in 1995 niet had aangenomen als programmeur, als later Wilfred van Dijk, hoofd systeembeheer van de afdeling Psychologie van de UvA, niet oogluikend had toegestaan dat Bas Roosen en ik onverantwoord veel programmeertijd aan Interapy besteedden, en als Bert Bredeweg me niet had laten geloven dat promoveren een optie was. Dank!

Als werk veel tijd vraagt, dan hoop je daar vrienden te vinden. Met Bart Schrieken

en Janneke Broeksteeg is dat volkomen gelukt. Ik ben blij dat zij mijn paranimfen zijn. Paranimfen worden geacht om tijdens de verdediging bij te springen wanneer de promovendus flauwvalt of het niet meer weet. Zij zouden dat kunnen. Ik word geflankeerd door de twee beste internetbehandelaars die ik ken.

Dank aan mijn ouders, voor hun onvoorwaardelijke liefde en steun ("het is niet of, maar wanneer"), aan Roel en Mark, mijn broers, voor hun kameraadschap, aan Ellen en Maarten voor hun voorbeeld, en aan Marianne en Henk voor hun betrokkenheid (en hun lieve dochter). Dank ook aan Hajo, voor de woensdagmiddagen met de kinderen en de periodieke reset in donker Amsterdam.

En dank, een niet te beschrijven hoeveelheid dank, aan Vera, Koen en Paulijn. Onder de streep gaat er niets boven de liefde.

Zaandam Jeroen Ruwaard
oktober, 2012.

Curriculum Vitae

Jeroen Ruwaard was born June 2nd, 1970, in Ede, the Netherlands. After an education at the Gymnasium, he earned a Master of Science degree in Experimental Psychology at the University of Amsterdam, where he programmed the websites for the first clinical trials of the Interapy research group. In 2000, he joined Interapy PLC. His work for Interapy involved the technological and methodological aspects of e-therapy (i.e. system development, trial design and data analysis). He can be reached at info@jruwaard.nl.

Publications

Ruwaard, J., Lange, A., Broeksteeg, J., Renteria-Agirre, A., Schrieken, B., Dolan, C.V., & Emmelkamp, P. (2012). Online cognitive behavioral treatment of bulimic symptoms: A randomized controlled trial. *Clinical Psychology & Psychotherapy*, epub ahead of print. doi: 10.1002/cpp.1767

Ruwaard, J., Lange, A., Schrieken, B., Dolan, C.V., & Emmelkamp, P. (2012). The effectiveness of online cognitive behavioral treatment in routine clinical practice. *PLoS ONE*, 7(7): e40089. doi: 10.1371/journal.pone.0040089

Lange, A., Ruwaard, J., Schrieken, B., de Haas, S., Broeksteeg, J., Renteria Agirre, A., & Kuyper, A. (2011). Online behandeling van seksueel misbruikte adolescenten: Ethische dilemma's en resultaten van een geprotocolleerde trial [Online treatment of sexually abused adolescents: Ethical dilemma's and the results of a protocolled trial]. *Directieve Therapie, 31*(3), 259-279.

Ruwaard, J., Lange, A., Schrieken, B., & Emmelkamp, P. (2011). Efficacy and effectiveness of online cognitive behavioral treatment: A decade of Interapy research. In B.K. Wiederhold, S. Bouchard & G. Riva (Eds.), *Annual review of cybertherapy and telemedicine 2011. Advanced technologies in behavioral, social and neurosciences* (pp. 9-14). Amsterdam: The Interactive Media Institute and IOS Press. doi: 10.3233/978-1-60750-766-6-9

Lange, A., & Ruwaard, J. (2010). Ethical dilemmas in online research and treatment of sexually abused adolescents. *Journal of Medical Internet Research, 12*(5), e58. doi: 10.2196/jmir.1455

Ruwaard, J., & Lange, A. (2010). Boulimie online overwinnen: resultaten van een gecontroleerd onderzoek [Overcoming bulimia online: results of a controlled trial]. *Directieve Therapie, 30*(1), 17-34.

Ruwaard, J., Broeksteeg, J., Schrieken, B., Emmelkamp, P., & Lange, A. (2010). Web-based therapist-assisted cognitive behavioral treatment of panic symptoms: A randomized controlled trial with a three-year follow-up. *Journal of Anxiety Disorders, 24*, pp. 387-396. doi: 10.1016/j.janxdis.2010.01.010

Lange, A., Ruwaard, J., & Schrieken, B. (2009). Hulpverlening voor Psychische Problematiek via het World Wide Web [Help for psychological problems via the World Wide Web]. *De Psycholoog, 44*(12), 634-640.

Ruwaard, J., & Lange, A. (2009). [Book review: Hands-On Help: Computer-Aided Psychotherapy (Maudsley Monograph 49).], Cognitive Behaviour Therapy, 36,(2). doi: 10.1080/16506070802651792

Ruwaard, J., Schrieken, B., Schrijver, M., Broeksteeg, J., Dekker, J., Vermeulen, H., & Lange, A. (2009). Standardized web-based CBT of mild to moderate depression: A randomised controlled trial with a long-term follow-up. *Cognitive Behaviour Therapy, 38*(4), 206-221. doi: 10.1080/16506070802408086

Lange, A., Ruwaard, J., Schrieken, B., Broeksteeg, J., Van Tienhoven, S., Jager, J., & Emmelkamp, P. M. (2007). Geprotocolleerde cognitieve gedragstherapie van klinische en subklinische paniekstoornis, via het internet: de behandeling en de resultaten van een gecontroleerde gerandomiseerde trial. [Standardized cognitive behavioral treatment of clinical and subclinical panic disorder, through the internet: the treatment and results of a randomized controlled trial.] *Directieve Therapie, 27*(2), 1-13. doi: 10.1007/BF03056843

Ruwaard, J., & Lange, A. (2007). Recensie: Programmeringstudie E-Mental Health: High Tech, High Touch, High Trust. *Maandblad voor de Geestelijke Volksgezondheid, 62*(12), 1084-1086.

Ruwaard, J., Lange, A., Bouwman, M., Broeksteeg, J. & Schrieken, B. (2007). E-Mailed Standardized Cognitive Behavioural Treatment of Work-Related Stress: A Randomized Controlled Trial. *Cognitive Behaviour Therapy, 36*(3), 179-192. doi: 10.1080/16506070701381863

Lange, A., Schrieken, B., Scheijde, R., Broeksteeg, J., Ruwaard, J., Schrijver, M., S., M., van de Ven, J.-P., & Emmelkamp, P. M. G. (2005). Interapy: diagnostiek en geprotocolleerde behandeling van welomschreven stoornissen via internet. [Interapy: diagnosis and standardized treatment of well-defined disorders through the internet.] *Tijdschrift voor Pyschotherapie, 31*(5), 355-376. doi: 10.1007/BF03062162

Bierman, D. J., Roosen, B., & Ruwaard, J. (1998). On-line Exams, from Intranet to Internet. *Proceedings of the International BITE conference*, pp. 218-228.

Bredeweg, B., Koopman, P., Ruwaard, J., Lange, A., Schrieken, B., Van de Ven, J. P., & Roosen, B. (1998). Designing Counselling Systems for the WWW. Paper presented at *The Webnet World Conference of the WWW, Internet & Intranet*, Charlottesville, USA.]

In 1997, researchers at the University of Amsterdam developed one of the first psychotherapeutic applications of the World Wide Web. They implemented a standardized cognitive behavioural treatment (CBT) of post-traumatic stress symptoms in a website, and used this site to treat clients over the internet, without face-to-face contact. Over the years, the efficacy of this treatment was established in a series of controlled trials. In this dissertation, we explore the wider applicability of online CBT, in four randomized controlled trials and a practice study. In the controlled studies, we assess the efficacy of online CBT for work-related stress, mild to moderate depression, panic disorder and bulimia nervosa. In the fifth study, we examine the effectiveness of online CBT in routine clinical practice. The results suggest that online CBT provides a feasible alternative to existing treatment options for people who suffer from a variety of mental health disorders.

ISBN 978-94-6191-588-7