



Negative effects of Internet-based cognitive behavior therapy

Monitoring and reporting deterioration and adverse and unwanted events

Alexander Rozental

© Alexander Rozental, Stockholm University 2016
Illustration on cover printed with permission © Michelle Eismann

ISBN 978-91-7649-597-1
ISBN 978-91-7649-598-8

Printed in Sweden by US-AB, Stockholm 2016
Distributor: Department of Psychology, Stockholm University

“Primum Non Nocere.”

*To my mom, Britt Rozental
To my dad, Maks Rozental
To my stepdad, Börge Österholm
Thanks for all your love and
support throughout the years.*

Abstract

Internet-based cognitive behavior therapy (ICBT) has the potential of providing many patients with an effective form of psychological treatment. However, despite helping to improve mental health and well-being, far from everyone seem to benefit. In some cases, negative effects may also emerge. The overall aim of the present thesis was to establish the occurrence and characteristics of such incidents in ICBT using a combination of quantitative and qualitative methods. **Study I** determined deterioration, non-response, and adverse and unwanted events in a sample of 133 patients undergoing ICBT for social anxiety disorder. The results indicated that up to 6.8% fared worse during the treatment period, depending on the self-report measure and time point, as determined using the Reliable Change Index (RCI), while the non-response rate was between 29.3 to 86.5% at post treatment assessment, and 12.9% experienced other negative effects. **Study II** investigated the responses to open-ended questions on adverse and unwanted events among 556 patients in four separate clinical trials of ICBT; social anxiety disorder, panic disorder, major depressive disorder, and procrastination. In total, 9.3% reported negative effects, with a qualitative content analysis revealing two categories and four subcategories; patient-related, i.e., gaining insight and experiencing new symptoms, and treatment-related, i.e., difficulties applying the treatment interventions and problems related to the treatment format. **Study III** explored the number of patients achieving reliable deterioration, as determined using the RCI on the individual raw scores of 2866 patients from 29 clinical trials of ICBT. The results showed that the deterioration rate was higher among patients in a control condition, 17.4%, in comparison to treatment, 5.8%. Predictors were related to decreased odds of deterioration for patients receiving treatment; clinical severity at pre treatment assessment, being in a relationship, having a university degree, and being older. As for the control condition, only clinical severity at pre treatment assessment was associated with decreased odds of deterioration. **Study IV** examined a newly developed self-report measure for monitoring and reporting adverse and unwanted events, the Negative Effects Questionnaire. The results suggested a six-factor solution with 32 items; symptoms, quality, dependency, stigma, hopelessness, and failure. One-third of the patients reported experiencing unpleasant memories, stress, and anxiety, with novel symptoms and a lack of quality in the treatment and therapeutic relationship having the greatest negative impact. The general finding of the present thesis is that negative

effects do occur in ICBT and that they are characterized by deterioration, non-response, and adverse and unwanted events, similar to psychological treatments delivered face-to-face. Researchers and clinicians in ICBT are recommended to monitor and report negative effects to prevent a negative treatment trend and further the understanding of what might contribute to their incidents. Future research should investigate the relationship between negative effects and treatment outcome, especially at follow-up, to examine if they are transient or enduring. Also, interviews could be conducted with those achieving reliable deterioration to explore if and how it is experienced by the patients and to see if it is attributed to the treatment interventions or other circumstances.

Keywords: Negative effects, Internet-based cognitive behavior therapy, deterioration, non-response, adverse and unwanted events, qualitative content analysis, individual patient data meta-analysis, exploratory factor analysis, Negative Effects Questionnaire.

Sammanfattning

Internetbaserad kognitiv beteendeterapi (IKBT) har goda förutsättningar att kunna bli en form av psykologisk behandling som på ett effektivt sätt hjälper patienter med att hantera sin psykiska ohälsa och förbättra sitt välmående. Trots detta är det dock långtifrån alla som tycks bli bättre. För en del kan det till och med resultera i negativa effekter. Det övergripande syftet med denna avhandling har således varit att undersöka förekomsten av sådana fall och hur dessa uttrycks, såväl med kvantitativa som kvalitativa metoder. **Studie I** fastställde andelen försämrade, oförändrade samt andra ogynnsamma eller oönskade händelser bland 133 personer som behandlades med IKBT för social ångest. Resultatet visade att uppemot 6,8 % försämrades under sin behandlingsperiod beroende på vilket självskattningsformulär respektive tidpunkt som studerades, beräknat enligt metoden Reliable Change Index (RCI). Likaså var 29,3 % till 86,5 % oförändrade vid eftermätningen samt att 12,9 % rapporterade andra former av negativa effekter. **Studie II** undersökte svaren på öppna frågor som gällde ogynnsamma eller oönskade händelser bland 556 patienter i fyra olika kliniska studier med IKBT; social ångest, paniksyndrom, egentlig depressionsepisod och prokrastinering. Totalt sett rapporterade 9,3 % att de hade erfårit negativa effekter, vilka analyserades med hjälp av kvalitativ innehållsanalys. Två övergripande kategorier och fyra subkategorier framkom; patientrelaterade, som ökad insikt respektive nya symptom, samt behandlingsrelaterade, som svårigheter att implementera behandlingsinterventionerna respektive problem med behandlingsformatet. **Studie III** utröntade andelen patienter som försämrades i enlighet med RCI, baserat på insamlad rådata från 2866 personer i 29 olika kliniska studier med IKBT. Resultatet visade att försämring var mer förekommande hos de som var i en kontrollgrupp, 17,4 %, jämfört med de som fick behandling, 5,8 %. Bland de som genomgick behandling existerade det även ett par prediktorer som innebar lägre odds för försämring; större svårigheter vid förmätningen, att befinna sig i en relation, att ha en universitetsutbildning respektive att vara äldre. För de som var i en kontrollgrupp var enbart större svårigheter vid förmätningen relaterat till lägre odds för försämring. **Studie IV** testade ett nykonstruerat självskattningsformulär; Negative Effects Questionnaire. Resultatet visade på en faktorlösning med sex faktorer och 32 påståenden; symptom, kvalitet, beroende, stigma, hopplöshet respektive misslyckande. En tredjedel av personerna svarade att de hade upplevt obehagliga minnen, stress och ångest, samtidigt som nya symptom och bristande kvalitet i både

behandlingen respektive den terapeutiska relationen hade haft störst negativ inverkan på dem. Den generella slutsatsen av denna avhandling är således att negativa effekter förekommer i IKBT och att de kännetecknas av försämring, ett oförändrat tillstånd samt andra ogynnsamma eller oönskade händelser, något som liknar tidigare forskning av psykologisk behandling som bedrivs ansikte-mot-ansikte. Forskare och behandlare i IKBT rekommenderas att övervaka och rapportera negativa effekter i syfte att förhindra en negativ utveckling i behandlingen samt för att öka kunskapen om vad som kan bidra till deras förekomst. Framtida forskning bör undersöka relationen mellan negativa effekter och behandlingsutfall utifrån längre tidsperspektiv för att se om dess påverkan är övergående eller ihållande. Vidare kan till exempel intervjuer utföras med de patienter som har försämrats för att ta reda på om och hur det uppfattas samt huruvida det har förorsakats av behandlingen eller andra omständigheter.

Nyckelord: Negativa effekter, Internetbaserad kognitiv beteendeterapi, försämring, oförändrat tillstånd, ogynnsamma eller oönskade händelser, kvalitativ innehållsanalys, individuell patient meta-analys, explorativ faktoranalys, Negative Effects Questionnaire.

List of publications

- I. Boettcher, J., Rozental, A., Andersson, G., & Carlbring, P. (2014). Side effects in Internet-based Interventions for social anxiety disorder. *Internet Interventions*, *1(1)*, 3-11.
- II. Rozental, A., Boettcher, J., Andersson, G., Schmidt, B., & Carlbring, P. (2015). Negative effects of Internet interventions: A qualitative content analysis of patients' experiences with treatments delivered online. *Cognitive Behaviour Therapy*, *44(3)*, 223-236.
- III. Rozental, A., Magnusson, K., Boettcher, J., Andersson, G., & Carlbring, P. (In press). For better or worse: An individual patient data meta-analysis of deterioration among participants receiving Internet-based cognitive behavior therapy. *Journal of Consulting and Clinical Psychology*.
- IV. Rozental, A., Kottorp, A., Boettcher, J., Andersson, G., & Carlbring, P. (2016). Negative effects of psychological treatments: An exploratory factor analysis of the Negative Effects Questionnaire for monitoring and reporting adverse and unwanted events. *PLoS ONE*, *11(6)*, e0157503.

The original papers are reproduced with permission from the publishers.

Papers I: Copyright © 2014 Elsevier (www.elsevier.com)

Paper II: Copyright © 2015 Taylor & Francis (www.tandfonline.com)

Paper III: Copyright © 2016 American Psychological Association (www.apa.org)

Paper IV: Copyright © 2016 PLoS (www.plos.org)

Abbreviations

CBM	Cognitive Bias Modification
CBT	Cognitive Behavior Therapy
CI	Confidence Interval
ETQ	Experiences of Therapy Questionnaire
GAD	Generalized Anxiety Disorder
HRS-D	Hamilton Rating Scale for Depression
ICBT	Internet-based Cognitive Behavior Therapy
INEP	Inventory for the Assessment of Negative Effects in Psychotherapy
IPDM	Individual Patient Data Meta-Analysis
LSAS-SR	Liebowitz Social Anxiety Scale – Self-Report
MADRS-S	Montgomery-Åsberg Depression Rating Scale – Self-Report
MDD	Major Depressive Disorder
NEQ	Negative Effects Questionnaire
OCD	Obsessive-Compulsive Disorder
OQ-45	Outcome Questionnaire-45
OR	Odds Ratio
PD/A	Panic Disorder with or without Agoraphobia
PTSD	Posttraumatic Stress Disorder
QOLI	Quality of Life Inventory
RCI	Reliable Change Index
SAD	Social Anxiety Disorder
SCID-I	Structured Clinical Interview for DSM-IV Axis I Disorders
UE-ATR	Unwanted Event to Adverse Treatment Reaction Checklist
VNIS	Vanderbilt Negative Indicators Scale

Contents

Introduction	12
Background.....	13
Negative effects	15
Definitions	15
Deterioration and non-response.....	19
Adverse and unwanted events	22
Internet-based cognitive behavior therapy	25
History	25
Efficacy and effectiveness.....	27
Negative effects of ICBT	28
Aims of the thesis	34
Method.....	35
Participants	35
Ethical considerations	36
Measures.....	37
Analyses	40
Results	42
Discussion.....	44
The deterioration effect revisited	45
Predicting deterioration	49
Adverse and unwanted events	52
Preventing negative effects	55
Limitations and future research directions	61
Conclusion	69
Acknowledgements	70
References	74

Introduction

Psychological treatments offer relief for those who are suffering from mental distress and whose well-being has been impaired. Evidence suggests that many patients who are troubled by different psychiatric disorders can benefit from treatment, thereby overcoming great difficulties and starting to live a more valued way of life. It is therefore encouraging that the access to the right type of care is becoming readily more available, with psychological treatments being disseminated through routine outpatient care and new means of delivery, such as, via the Internet. Although promising steps have been made in terms of determining their efficacy and effectiveness, far less attention has been given to their potential of also having negative effects for some patients. It seems like the possibility of causing harm in treatment was never really considered, with few researchers and clinicians acknowledging this issue throughout history. However, if psychological treatments are in fact powerful enough to reduce symptoms, there is also a likelihood of them inducing harm. As discussed by Foulkes (2010): “All clinicians need to acknowledge that any treatment that has the capacity to greatly help the patient can also in equal measure have the potential to cause harm.” (p. 189). Hence, without knowing what these negative effects are and how they could be prevented, there is a risk of inadvertently making patients experience events during treatment that are adverse and unwanted, or have a treatment outcome that is characterized by non-response and deterioration. Recently, investigations have indicated that up to one-tenth of all patients seem to fare worse during treatment, and a number of suggestions have been proposed with regard to other of incidents that could be detrimental. Still, what might be responsible for their occurrence is unknown, and how many are afflicted in different contexts remains unclear. The present thesis therefore intends to investigate an issue that has been largely unexplored within the research of psychological treatments, with the aim of providing an understanding of the occurrence and characteristics of negative effects. Hopefully, the results will influence how we monitor and report change in treatment so that it includes both its ups and downs, and making it possible to predict and prevent events from having a negative treatment trend.

Åkerö, July-August, 2016.

Background

One of the most renowned aphorisms among healthcare providers stems from the axiom “Primum, Non Nocere”, that is, “first, do no harm” or “above all, do no harm”. According to legend, it was originally dictated by Hippocrates and can be found in the Hippocratic oath – a set of principles that convey the moral and ethical stance of anyone who offer their help to heal others. A historical review by Smith (2005) suggests that the axiom is in fact supposed to read: “As to diseases, make a habit of two things – to help, or at least to do no harm” (p. 371), underlining the importance of balancing potential gains against possible costs, but also to be careful not to inflict pain or other types of negative effects. Although the exact wording of the axiom has been a topic of great debate, its influence on how medicine and research is performed cannot be underestimated. The Hippocratic oath is, for instance, pledged by many medical doctors around the world, while its basic premise has laid the foundation for patient safety and monitoring in clinical trials. Even among other healthcare providers, such as, clinicians administering psychological treatments, the notion of avoiding harm is a cornerstone in their daily practice. The American Psychological Association, for instance, includes the following formulation in its ethics code: “...take reasonable steps to avoid harming their clients/patients...” and “...to minimize harm where it is foreseeable and unavoidable.” (p. 1065; American Psychological Association, 2002). Likewise, clinicians are recommended to inform their patients of the risks involved so that they can make an educated decision about undergoing treatment: “...inform clients/patients of the developing nature of the treatment, the potential risks involved, alternative treatments that may be available, and the voluntary nature of their participation.” (p. 1072; American Psychological Association, 2002). Still, although being a fundamental concept in terms of stipulating a moral and ethical standpoint, harm has not received much attention outside the field of medicine when it comes to determining its occurrence and characteristics. In fact, a review of 132 published randomized controlled trials of psychological treatments found that only 28 (21%) contained information concerning harm (Jonsson, Alaie, Parling, & Arnberg, 2014). Vaughan, Goldstein, Alikakos, Cohen, and Serby (2014) also noted that this type of research is nine to twenty times less likely to mention anything that might reveal possible or actual cases where negative effects have occurred, compared to pharmacological investigations. However, despite seldom being reported, the idea that treatment could be

associated with events that may be detrimental for the patient is nothing new. Kächele and Schachter (2014) gave an overview of various negative effects in psychoanalytic psychotherapy, some being described in the literature as early as the turn of the last century. In addition, the first empirical evidence of their existence is believed to have been provided in the evaluation of the Cambridge-Somerville Youth Study (Powers & Witmer, 1951), which found that individuals in the treatment condition fared worse to a greater extent compared to those in the control condition, i.e., became more delinquent. Bergin (1966) also presented a review of several psychotherapy outcome studies showing that some patients deteriorate while undergoing treatment. Since then, these findings have been both updated and replicated in different contexts and samples, ranging from clinical trials to routine outpatient care, across various psychiatric disorders, and regardless of theoretical orientation. Furthermore, several books have also been published, investigating failures in treatment (Foa & Emmelkamp, 1983), providing definitions (Mays & Franks, 1985), sharing examples from clinical practice (Kottler & Carlson, 2003; Stuart, 1970), and inviting some of the most recognized clinicians and researchers at the time to give their perspective on the subject (Strupp, Hadley, & Gomes-Schwartz, 1977). Nonetheless, the current understanding of negative effects remains more or less the same, with little knowledge of what makes some patient deteriorate and not others, and why some seem to experience adverse and unwanted events. Moreover, the development and dissemination of psychological treatments delivered via the Internet are even less explored, making it unknown if these differ from those administered face-to-face with regard to such incidents. However, in a recent series of commentaries on negative effects, the importance of exploring its causes has been emphasized, spurring a renewed interest in its research (Barlow, 2010; Boisvert, 2010; Dimidjian & Hollon, 2010). Clinicians and researchers are now recommended to monitor and report negative effects of psychological treatments and not only their potential for alleviating mental distress and increasing well-being. Examining the risks of undergoing treatment together with its benefits should, in the end, result in a better understanding of its occurrence and characteristics and help healthcare providers truly adhere to the axiom of “first, do no harm”. As summarized by Scott and Young (2016): “...given the unrelenting popularity of therapies as a treatment for common mental disorders in primary and secondary care, it is important to develop a strategy for examining failed psychotherapy interventions. This is not an attempt to undermine therapies, but is likely to be beneficial. First, because it is likely to lead to improvements in techniques and practice, and second, because every branch of medicine learns from its mistakes; it is inconceivable that the same is not true for psychotherapies.” (p. 209).

Negative effects

Definitions

Negative effects refer to results or outcomes of a particular influence that are unwanted or undesirable and often having unappreciated or unfavorable implications for those afflicted. However, this view of negative effects is only one of many definitions that exist in relation to describing events that might be harmful in psychological treatments. Originally defined by Hadley and Strupp (1976), the concept of negative effects was raised as a critique against the use of the term *deterioration*: "...because it seems to describe more accurately the problem of patients getting worse as a function of the therapeutic influence, as opposed to other factors possibly extraneous to psychotherapy." (p. 1291). At the time, Bergin (1966) had provided evidence suggesting that a proportion of all patients fare worse during treatment, which was referred to as the "client-deterioration phenomenon" (p. 236), or, deterioration effect. The idea that some patients could fare worse as a result of psychological treatments was, however, criticized by Rachman (1971) and May (1971) on account of the difficulty involved in determining causality. Negative effects was thus proposed in response to this debate, but in contrast to deterioration this term includes every type of change in every area of functioning that may be negative for the patient, for example, stigma and interpersonal difficulties. This was in line with the theoretical concept of assessing treatment outcome in different domains, not only symptomatology, and to rely on more than one perspective in controlling for risks and benefits, i.e., patient, clinician, and significant others, the "tripartite model of mental health and therapeutic outcome" (p. 187; Strupp & Hadley, 1977). However, Mays and Franks (1985) argued that such a definition still does not resolve the problem of establishing a causal relationship between the treatment and harmful events, suggesting that other circumstances could be responsible for their incidence. Furthermore, according to the so-called tripartite standpoint, negative effects depend on what viewpoint is being applied, which implies that harmful events, as perceived by the patient, do not always correspond to the opinion of the clinician nor a relative. Therefore, defining events as negative would undoubtedly vary between assessors and could be unreliable as a way of monitoring and reporting their occurrence. Instead, the concept of *negative outcome* was introduced, referring to a decline between the onset and termination of treatment that is not limited to a negative change caused

by certain treatment interventions: “The term Negative Outcome is not restricted to those negative changes which are therapy-induced, and usage of the term does not therefore imply that the therapist is necessarily responsible for the negative change.” (p. 8). Mays and Franks (1985) contended that because of the difficulties associated with determining causality, it would be better to assess every type of decline that occur during treatment regardless of its origin. Albeit more in line with the term *side effects* commonly used in pharmacological research, i.e., including all events that are unanticipated (Curtin & Schulz, 2011), this also introduces problems with validity, as it becomes hard to establish what is driving these effects. Moreover, it again focuses on the successive worsening of a given condition, or, deterioration, potentially missing out other incidents that could be experienced as negative by the patient. A similar term often employed in medicine is *iatrogenic*, i.e., adverse reaction, but is seldom used in psychological treatments (Boisvert & Faust, 2002). Dimidjian and Hollon (2010) thus proposed another way of differentiating treatment outcome in order to facilitate an exploration of its risks and benefits, involving; *harmful treatment effect*, *no treatment effect*, and *beneficial treatment effect*. This is further divided by its progression; constant course of target problem, deteriorating course of target problem, and spontaneous remission of target problem, taking into account the natural processes of psychiatric disorders, that is, fluctuations in symptomatology that are expected in some conditions. Although some difficulties remain with regard to causality, this facilitates an inspection of what may be responsible for negative effects in treatments: “Disorders with a constant course do not change over time and provide a backdrop against which changes induced by treatment should be relatively easy to detect; harmful treatments leave the patients worse than when they started.” (p. 23; Dimidjian & Hollon, 2010). Interestingly, *non-response* is hereby perceived as harmful, which was not the case in earlier attempts of defining negative effects. However, it could be argued that the absence of any benefit might also be detrimental, should the delivery of treatment either have prevented the patient from receiving more adequate care or if it resulted in a prolongation of suffering. The occurrence of *new problems* is also recognized, something that might otherwise have been overlooked if only the condition for which the patient sought help for was scrutinized: “...such harm can be difficult to detect, particularly if attention is focused exclusively on the target problem or if the effects of treatment on the target problem are beneficial.” (pp. 24-25; Dimidjian & Hollon, 2010). Recently, Linden (2013) proposed a similar, yet even more extensive terminology of negative effects, including; *unwanted event*, *treatment-emergent reactions*, *adverse treatment reactions*, *malpractice reaction*, *treatment non-response*, *deterioration of illness*, *therapeutic risk*, and *contraindications*. Even though it is presented as checklist, the purpose of this classification is primarily to provide an aid for clinicians and researchers to become better at distinguishing various incidents, rather than being used a

specific instrument for their assessment. Analogous to contemporary efforts of defining negative effects, it also includes events that might be perceived as adverse and unwanted by the patient, apart from deterioration. “An example is emotional turmoil in the patient during treatment, independent of the fact that it is unavoidable or may even turn out to be positive in the long run.” (p. 288; Linden, 2013). This should allow an investigation of incidents that occur during treatment without prematurely regarding them as unrelated. It also recognizes the fact that some negative effects, albeit experienced as undesirable by the patient, could be beneficial in the long-run: “All UEs [Unwanted Events] should be taken as treatment related unless it has convincingly been shown that they are not. Furthermore, therapists must be aware that some treatment effects are ambiguous in quality, i.e., they are therapeutic and negative at the same time, but these should still be recognized as UE.” (p. 288; Linden, 2013). As such, this classification of negative effects is inclusive by nature, monitoring and reporting all events that occur. This could prove to be important later on, as it may reveal therapeutic risks, that is, unwanted events probably caused by treatment, which the patient needs to be informed about. In addition, negative effects that are highly plausible given certain patient characteristics could also result in identifying specific contraindications, thus preventing the administration of particular treatment interventions to those that are at risk of being harmed. Similar to Dimidjian and Hollon (2010), non-response is seen as detrimental, but the classification also includes events that are caused by incorrect or improperly applied treatment, i.e., malpractice reactions. Although an important topic, malpractice is perhaps more closely related to issues of abuse or misconduct, e.g., initiating a sexual relationship with a patient, than the negative effects that are related to a correctly performed treatment. Malpractice might also be more associated with the use of potentially harmful treatment interventions, for example, recovered memory techniques or rebirthing, than what is being evaluated in a large number of clinical trials and provided by clinicians with appropriate training. Beyerstein (2001) called these *fringe psychotherapies*, however, Lilienfeld (2007) argues that even psychological treatments with the highest possible evidence and quality of delivery could yield negative effects for some patients: “...one may suspect that some ESTs [Empirically Supported Therapies] could be PHTs [Potentially Harmful Therapies] if administered improperly. For example, many standard texts warn that an insufficient duration of exposure and response prevention – that is, termination of exposure before adequate habituation of anxiety has occurred – may result in a worsening of clients’ anxiety symptoms.” (p. 56). Hence, it is possible that the investigation of negative effects will eventually reveal that certain treatment interventions are harmful for some patients, despite being seen as safe and efficient based on current knowledge. Being able to separate harmful from harmless psychological treatments might therefore be a matter of evidence, suggesting that some treatment interventions currently being

applied could be deemed risky or even unethical in the future, as has been the case for conversion therapy and aversive conditioning (Haldeman, 1994).

The definitions of negative effects are, in other words, relatively varied, ranging from the more objective assessment of worsening between time points to the more subjective examination of events that are perceived or experienced as negative by clinicians, significant others, or the patients. Negative effects can also be attributed to the treatment interventions that are being implemented or other circumstances in the patients' lives, affecting the exploration of their causes and perhaps also what can be done to prevent them from occurring. The lack of consensus with regard to defining negative effects might be one of the reasons why so few studies have been carried out on this particular topic of psychological treatments (Linden & Schermuly-Haupt, 2014). However, a similar argument could be raised in relation to clinically significant change, i.e., improvement (Bauer, Lambert, & Nielsen, 2004), which has not hindered research from being conducted. As for the present thesis, the suggestions by Dimidjian and Hollon (2010) have been used as guidelines for monitoring and reporting negative effects, i.e., to use both quantitative and qualitative methods for exploring incidents that may prove to be harmful. Thus, both an increase in symptomatology, as assessed by statistical means, and the responses by the patients themselves have been examined to allow a greater breadth and depth in the ensuing investigations. This is also in line with the idea of using a nomothetic and an idiographic approach in inspecting treatment outcome, that is, finding a more general explanation for the causes of negative effects without disregarding the value of studying individual cases: "Greater emphasis on more individual idiographic approaches to studying the effects of psychological interventions would seem necessary if psychologists are to avoid harming their patients and if they are to better understand the causes of negative or iatrogenic effects from their treatment efforts." (p. 13; Barlow, 2010). In terms of the definitions applied, the present thesis utilizes the term negative effects to describe incidents of a potentially harmful nature more globally, whether or not these are in fact related to treatment. This is close to the proposition by Hadley and Strupp (1976), although it also includes events that can co-occur (Mays & Franks, 1985). In addition, deterioration is employed to depict the worsening of a given psychiatric disorder or increase in symptomatology, similar to what was proposed by Bergin (1966), while *adverse and unwanted events* is used to illustrate incidents that are unrelated to worsening, but may still be regarded as negative, for example, new symptoms, or, new problems, as referred by Dimidjian and Hollon (2010). Furthermore, non-response is implemented in those cases where no positive change is assumed to have occurred during treatment. Additional negative effects have been proposed, such as, drop out (Lilienfeld, 2007), and severe adverse events (Rozenal et al., 2014), e.g., suicide attempts, but are, however, not explored or discussed in great length in the present thesis.

Deterioration and non-response

The deterioration effect was put forward by Bergin (1966) during a point in the history of psychological treatments when their benefits were questioned by Eysenck (1952), arguing that the effects of such treatment interventions were nothing else than spontaneous remission. The main objective of many investigations at that time was therefore to provide evidence for their utility as a mean for alleviating mental distress. However, as noted by Bergin (1963) it also became clear that, for some patients, receiving treatment was associated with faring worse, i.e., an increase in symptomatology during the period which the patient received some form of care. This was based on two series of clinical trials, one involving patients with schizophrenia, known as the Wisconsin schizophrenia project (Truax, 1963), and the other comprised of psychoneurotic patients at the University of Chicago Counseling Center (Cartwright & Vogel, 1960). This led to the conclusion that: "...change does indeed occur in psychotherapy, but in two opposite directions." (p. 246; Bergin, 1963). In other words, not only were there indications of the benefits of psychological treatments, it was also suggested that it may have its risks. Bergin (1966) later refined this finding with four additional studies, and as more and more research was conducted with regard to treatment outcome it became unmistakable that deterioration did occur for a small yet significant proportion of all patients undergoing treatment. In a summary of the research that was conducted during that era, Lambert (2013) concluded that by the 1970's it was evident that 5-10% of the patients fared worse. By then, a vast amount of critiques and replies had been produced, in part because of the problems associated with establishing causality, but perhaps also due to the controversy of psychological treatments having negative effects. However, as the number of studies supporting this observation reached more than fifty, the deterioration effect was more or less recognized as something that could happen during treatment. From a methodological perspective, the procedure for assessing worsening was, on the other hand, not yet established. In most cases deterioration was equal to a negative change score between two points of measurement, e.g., pre and post treatment assessment, but to what degree that had do be achieved was not decided (Mohr et al., 1990). This is relevant particularly because of the fluctuation in scores that can be attributable to unreliability of the self-report measure, which means that the patient needs to exceed a certain limit if the change is to reflect a true change and not just measurement error (Martinovich, Saunders, & Howard, 1996). For instance, in terms of improvement, a patient has to attain a change score that goes beyond a statistically determined criterion, the Reliable Change Index (RCI), calculated based on the difference between two time points divided by the standard error of difference (Jacobson & Truax, 1991). Also, cutoffs for clinically significant change can be used to determine the number of patients that have recovered during treatment, i.e., moving from a dysfunctional to a

functional distribution, but for deterioration such a method is not deemed possible given the problem associated with finding a distribution distinct from what is already considered dysfunctional: “There is no obvious counterpart to our distributional cutoff for clinical significance in the assessment of deterioration rates” (p. 350; Jacobson, Follette, & Revenstorf, 1984). Since then, the RCI has been the most commonly used method for assessing worsening in psychological treatments, often with a negative change score above 1.96 standard deviations regarded as a reliable deterioration (L. Christensen & Mendoza, 1986). There are, however, some suggestions of other endpoints, such as, those proposed by Wise (2004); 1.96 deteriorated, 1.28 moderately deteriorated, and 0.84 mildly deteriorated, as it could be argued that even milder forms of worsening is detrimental for the patient.

As for more recent investigations of deterioration, there are a few studies that have provided evidence for its incidence. For instance, Ogles, Lambert, and Sawyer (1995) found that, based on the RCI for the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960), 2 (5%) deteriorated while undergoing Cognitive Behavior Therapy (CBT) and 1 (3%) in interpersonal psychotherapy ($N = 162$). Meanwhile, Scogin et al. (1996) analyzed five studies of CBT delivered as self-help treatments, showing that 2 (1%) fared worse on the HRSD when assessed by a clinician, compared to 16 (9%) for self-reports ($N = 165-188$), although it should be noted that this was affirmed by a one-point increase. However, the largest evaluation of deterioration to date has been made by Hansen, Lambert, and Forman (2002), using the RCI for the Outcome Questionnaire-45 (OQ-45; Lambert et al., 1996) on patients in routine outpatient care at six sites, most receiving CBT. The average rate was 496 (8.2%), ranging from 4 (3.2%) to 84 (14.1%) depending on location ($N = 6072$). Similarly, Mechler and Holmqvist (2016) analyzed the results of treatments in routine outpatient care, using the RCI for the Clinical Outcome in Routine Evaluation (Evans et al., 2000). Patients received either CBT or some form of psychodynamic psychotherapy, with 15 (1.8%) faring worse in primary care as well as 22 (6.9%) in a psychiatric setting ($N = 1157$). In sum, the average rate of deterioration seem to be 5-10%, although it should be noted that these numbers are mainly derived from clinical trials of CBT and with adult patients with Major Depressive Disorder (MDD) or some kind of anxiety disorder. Comparing theoretical orientations head-to-head in terms of negative effects has not been conducted and it is unclear if deterioration occurs more frequently in certain psychiatric disorders. There are, however, reasons to believe that these rates may be somewhat higher in some contexts. Warren, Nelson, Mondragon, Baldwin, and Burlingame (2010) found that, using the RCI for the Youth Outcome Questionnaire on children obtaining different psychological treatments, 226 (24.1%) deteriorated in community care as compared to 440 (14.3%) in managed care ($N = 4011$). Mohr (1995) also suggested that deterioration occurs to a greater extent among patients with Borderline Personality Disorder and Obsessive-Compulsive Disorder

(OCD), although this was described as highly tentative given the lack of empirical data. In addition, it has also been proposed that the incidence of worsening is higher in group psychotherapy, with a review by Roback (2000) demonstrating that about 16% can be classified as casualties, i.e., achieving enduring deterioration, although the criteria for such a definition are unclear and the samples have been relatively small.

In terms of non-response, the issue of treatment outcome and clinically significant change becomes important to address before determining the number of patients that do not respond to treatment as planned. For example, a review by Loerinc et al. (2015) on the response rate for CBT for various anxiety disorders revealed that 49.5% were classified as responders at post treatment assessment and 53.6% at follow-up. However, it was also found that these proportions depended on what standards were implemented, i.e., the RCI or a cutoff, intention-to-treat or complete cases, and independent blind assessors or self-report measures. Similar to examining deterioration, non-response also warrants methodological considerations, albeit this has received much less attention. In some cases, the absence of any variation, i.e., no change at all (± 0) between two points of measurement, is used for deciding whether a patient has not responded. Again, given measurement error this might not be particularly reliable, which is why the RCI is more frequently used, that is, a change score that does not exceed the limits of what is seen as a reliable change, e.g., ± 5 points. Applying this procedure, additional examples of investigating the response rate for CBT exist. Hofmann, Asnaani, Vonk, Sawyer, and Fang (2012) reviewed meta-analyses on CBT for a number of different psychiatric disorders in order to establish the average response rate, e.g., Panic Disorder with/without agoraphobia (PD/A) 77%, MDD 51-87%, Generalized Anxiety Disorder (GAD) 46%, and OCD 38-50%. Albeit quite promising results, there is still a considerable proportion of all patients who do not improve and might be non-responders. In terms of other studies that have looked at non-response, the evaluation by Hansen et al. (2002) assessed the incidence of no change, i.e., in accordance with the RCI, demonstrating that the total rate of non-response was 3448 (56.8%), ranging from 57 (45.6%) to 219 (60.7%) at the individual sites. Likewise, Mechler and Holmqvist (2016) found that 514 (61.2%) did not respond in primary care compared to 211 (66.6%) in a psychiatric setting, again relying on the RCI. In other words, it appears the numbers are fairly consistent across studies, with non-response occurring for approximately half of the patients, but it should be stressed that this is again based on adult patients primarily receiving CBT for MDD and anxiety disorders. However, the review by Warren et al. (2010) on children suggested that the rate might be lower in this context, 226 (31.6%) for community care and 440 (31.4%) for managed care, but this could be due to the higher degree of deterioration that has been seen in this population.

Adverse and unwanted events

In comparison to deterioration and non-response, the rate of adverse and unwanted events is far less explored, making it difficult to determine its true incidence in psychological treatments. Linden and Schermuly-Haupt (2014) have proposed a number of possible explanations for this; a bias towards the positive of treatment, the spectrum of negative effects is extremely broad, there is no consensus of what might constitute such incidents, problems with differentiating deterioration from other types of negative effects, and no generally accepted method for determining adverse and unwanted events. Given the different ways of defining negative effects, there is also no clear and coherent conceptualization of what to monitor and report, which makes it difficult to pursue a systematic investigation of events that are experienced as negative by the patient, but which is not characterized by worsening. However, there have been several attempts at examining their occurrence. Suh, Strupp, and O'Malley (1986) were early to assess negative effect through the development of the Vanderbilt Negative Indicators Scale (VNIS). This was a clinician-rating system that could be utilized to review video recordings of therapy sessions with the purpose of determining the occurrence of negative incidents in treatment, e.g., unrealistic expectations by the patient, deficiencies in the therapeutic commitment, inflexible use of treatment interventions, poor therapeutic relationship, and a poor match. Albeit clearly defined and theoretically sound, based on the ideas by Strupp and Hadley (1977), it never became widely used apart from two studies of psychodynamic psychotherapy (Binder & Strupp, 1997). However, it did reveal that inflexible use of treatment interventions was correlated with worse treatment outcome, possibly illustrating an association between the behavior of the clinician and negative effects. Since then, a number of more recent attempts have been made with regard to monitoring and reporting adverse and unwanted events. The Experience of Therapy Questionnaire (ETQ), a patient self-report measure, was developed and evaluated by Parker, Fletcher, Berk, and Paterson (2013) on patients having received or currently undergoing psychological treatments ($N = 707$), primarily general counseling or CBT. The results revealed several possible domains, such as, negative therapist, pre-occupying therapy, idealization of therapist, as well as passive therapist. It was also suggested that younger patients dropped out from treatment because the clinician was perceived as passive or did not provide a resolution to their current problems, and that many believed the treatment to be ineffective. However, the ETQ also includes items regarding the benefits of treatment and is thus not intended for specifically examining negative effects. Furthermore, it has not yet been administered with the aim of determining the rate of such incidents, making it unclear to what extent adverse and unwanted events exist. The Inventory for the Assessment of Negative Effects of Psychotherapy (INEP; Ladwig, Rief, & Nestoriuc,

2014), on the other hand, provides more detailed information regarding the investigation of negative effects, assessed among patients having received psychological treatments ($N = 195$), in this case, evenly distributed between theoretical orientations. The findings from the responses revealed a number of domains; interpersonal changes, intimate relationship, stigmatization, emotions, workplace, therapeutic malpractice, and family and friends. Also, with regard to the harmful events that emerged, several occurred at a quite high rate; hurtful statements by the therapist, 109 (55%), longer periods when things were not going well, 83 (42.6%), and, also, [feeling] less secure, more labile and less resilient, 70 (35.9%). This allows a closer inspection of what adverse and unwanted events seem to be more frequent in treatment, but it is unknown at what rate these occur for patients in general. In addition, the INEP also includes items that could be regarded as unrelated to negative effects caused by treatment, for instance, insurance problems, which may have more to do with policies, and malpractice issues, which should be more closely related to ethics and jurisdiction than the treatment interventions being implemented.

As for the occurrence of adverse and unwanted events in psychological treatments more generally, not much is known about their incidence. Buckley, Karasu, and Charles (1981) investigated negative effects among clinicians undergoing treatment as part of their professional development and found that 20 (21%) experienced events that were seen as harmful, e.g., acting out, withdrawal, and relationship distress. A number of clinical trials have also provided reports of adverse and unwanted events of some sort; increase in marital problems among patients with agoraphobia, 7 (33%) (Hand & Lamontagne, 1976), emergence of new symptoms among patients with OCD, 3 (14%) (Foa & Steketee, 1977), and increased tension among patients with anxiety symptoms, 4-7 (30-54%) (Heide & Borkovec, 1983). Typically, these negative effects have been observed in relation to CBT, although Mohr (1995) reviewed evidence of such events from a variety of theoretical orientations. In most cases, however, the samples have been quite small and the results are therefore hard to generalize, although the notion that some patients can become anxious as a result of progressive relaxation has been a useful finding for clinicians. More recently, Moritz et al. (2015) determined the occurrence and characteristics of negative effects among patients with OCD ($N = 87$), whom had received psychological treatments, either CBT or another theoretical orientation, i.e., non-CBT. As a way of determining adverse and unwanted events, adapted versions of the Unwanted Event to Adverse Treatment Reaction Checklist (UE-ATR; Linden, 2013) and the INEP were used. The results indicated that almost all of the patients, 80 (92.9%), had experienced some form of adverse treatment reaction, i.e., any type of unwanted event probably caused by treatment. In addition, 77 (88.8%) described malpractice reactions, i.e., incorrect or improperly applied treatment interventions, 47 (54.1%) endorsed treatment emergent reactions,

i.e., unwanted event caused by treatment, and 12 (14.1%) stated that they had witnessed unethical conduct by their clinician. A closer inspection of the results also showed that patients with a higher educational level reported more adverse and unwanted events, on average, albeit the mean difference was only 1.95 incidents. Further, most of the negative effects were related to deterioration, new symptoms, disappointment about the treatment outcome, familiarity with treatment content, and stigma. Interestingly, for patients receiving CBT, new symptoms were primarily related to the implementation of gradual exposure. However, the largest study with regard to the issue of adverse and unwanted events has been conducted by Crawford et al. (2016). Using a cross-sectional survey administered to the National Health Service in Great Britain, a vast number of patients in 220 sites were asked whether or not they had experienced any harmful events during treatment ($N = 76\ 950$). Of these, 184 (83.6%) sites were subsequently included, with 14 587 (19%) patients responding. About half received CBT, although other theoretical orientations were also common. In total, 763 (5.2%) reported events that had a lasting detrimental impact. An additional 1099 (7.7%) also reported that they were uncertain about the occurrence of negative effects. Furthermore, the likelihood of reporting adverse and unwanted events was explored in relation to different sociodemographic and treatment variables, revealing that the Odds Ratios (OR) were higher among those unsure about what type of psychological treatment they were receiving, OR 1.71, 95% Confidence Interval (CI) [1.42, 2.05]. Likewise, it was higher for longer treatments, >12 months, OR 1.58, 95% CI [1.07, 2.33], >26 sessions, OR 1.65, 95% CI [1.18, 2.29], and for patients with a sexual identity that was not heterosexual, OR 1.28-1.98, 95% CI [0.83-2.92], as well as for those with an ethnical background that was self-described as non-white, OR 1.65-2.86, 95% CI [1.00-4.69], although it should be noted that the confidence interval was not reliable for some variables. There are, of course, a number of limitations related to the investigation of negative effects by survey, particularly with regard to selection bias and the fact that the results may have more to do with the likelihood of reporting incidents than actually experiencing them. However, the findings are nonetheless important given the increased odds of such events if the patient is not provided with adequate information about what treatment is being given, if the treatment is too long, and if the patient belongs to a minority group. As a consequence, Parry, Crawford, and Duggan (2016) suggested that, although it is fairly uncommon to experience negative effects, patients should always be informed about the possible risks involved before undergoing psychological treatments: “At present, patients are often asked to provide consent to psychological treatment without any discussion of potential harms associated with these interventions. Although negative effects associated with psychological therapies are far less common than positive ones, the process of informed consent requires some consideration of both.” (p. 211).

Internet-based cognitive behavior therapy

History

Internet-based Cognitive Behavior Therapy (ICBT) refers to the extension of CBT by implementing modern information technology as its primary mean of delivery (Andersson, 2015). As such, the fundamental idea and use of treatment interventions are exactly the same; the only difference is its format and administration, particular in terms of the extent of a face-to-face contact. However, CBT is a broad construct that includes a wide range of theoretical orientations and underpinnings; stemming from basic and applied research, clinical practice, and different philosophical perspectives (Dobson, 2009). Hence, defining CBT depends on what standpoint is being used, but a more general description would involve the notion of psychiatric disorders being a result of cognitive features, i.e., information processing (Westbrook, Kennerley, & Kirk, 2011), and the influence of behaviors, i.e., excess or deficit of certain activities, contextual effects, and particular schedules of reinforcement and punishment (Farmer & Chapman, 2016). The objective during treatment is thus to investigate and determine how these mechanisms maintain a specific problem, either through a cognitive conceptualization (Tarrier & Johnson, 2015), which emphasizes the problem of maladaptive patterns of thinking, or a functional analysis (Ramnerö & Törneke, 2008), which stresses the consequences of behavior. Treatment interventions can subsequently be selected and applied depending on the nature of the ongoing difficulties and the individual case formulation, making it highly flexible in terms of adjusting the treatment to the specific patient. Cognitive techniques can be used, for example, to explore biases and assumptions about oneself, others, and the world, similar to the process involved in conducting research, i.e., formulating a hypothesis and testing its validity. Behavioral techniques are also useful for changing habitual performance of behaviors that prevents the patient from acting more flexibly or living a more valued way of life, e.g., scheduling activities that can help improve mood. The treatment is often restricted to 10-20 sessions, depending on the condition and complexity of the patient, with continuous assessments of progress being used in order to evaluate its benefits, or, lack thereof. Furthermore, the agenda is determined and agreed upon collaboratively, although the patient is required to study and practice between the sessions through the use of assignments, referred to as homework (Kazantzis, Deane, & Ronan, 2000). Therefore, most of the

within-session activity involves psychoeducation, feedback, clarification, and introducing particular treatment interventions, while the between-session activity revolves around implementation by the patient, e.g., conducting a survey among colleagues to investigate the belief that making a mistake at work would be disastrous. An essential part of CBT has thereby been the provision of texts and exercises, such as, chapters from a self-help book, which, in turn, has resulted in the dissemination of self-help treatments. Williams and Martinez (2008) label these as bibliotherapy, used either as a stand-alone alternative, i.e., low intensity treatments with no or only limited guidance from a clinician, or as an adjunct to a regular face-to-face contact, i.e., a high intensity treatment. In both cases, the results are encouraging, albeit with greater benefits when supported by a clinician (Lewis, Pearce, & Bisson, 2012), although it should be underlined that bibliotherapy has also received mixed outcomes (c.f., Febraro, Clum, Roodman, & Wright, 1999), and can often be confused with similar formats in evaluations of its efficacy, creating some misunderstanding as to what is being reviewed (Cuijpers, Donker, van Straten, Li, & Andersson, 2010). Nonetheless, because of its accessibility, the delivery of such treatments was later easily transferred to other media, first using support via the telephone (Carlbring et al., 2006), and, following the development of modern information technology (Carlbring & Andersson, 2006), to computers as well as over the Internet. Research on ICBT began twenty years ago in several countries in parallel and has currently been assessed in close to 200 clinical trials; for a range of psychiatric disorders and somatic illnesses and in different settings (Lindfors & Andersson, 2016). The access to this type of treatment has also become available in routine outpatient care, allowing patients to receive CBT via the Internet regardless of geographical location (Hedman et al., 2014). In principle, ICBT usually involves 8-12 weeks of treatment that is delivered in chapters, or, modules, with one module being distributed to the patient each week. Similar to a self-help book, texts and exercises help the patient learn more about a condition, practice certain skills, confront fearful thoughts or emotions, and expose oneself to situations that have previously been avoided. However, a notable distinction between the formats is the administration of modules in chunks, with predetermined deadlines, e.g., weekly assignments, which is believed to improve adherence (Nordin, Carlbring, Cuijpers, & Andersson, 2010). In addition, a screening process involving self-report measures and a structured clinical interview is usually applied together with a post treatment assessment, which also should lower attrition rates. Furthermore, many cases of ICBT include the provision of guidance from a clinician who gives feedback and assistance during the treatment period, similar to reviewing the progress and homework with a patient face-to-face, but also giving reminders to follow through with the modules and upload the results of the exercises (Andersson, Rozental, Rück, & Carlbring, 2015). A key difference is also the amount of time allocated to

each patient, which seldom extends ten minutes per week (Andersson, Carlbring, Berger, Almlöv, & Cuijpers, 2009). However, current evidence implies that it does not matter who is providing the support, e.g., clinicians, master's degree students, or technical personnel, at least with regard to treatment outcome (Almlöv et al., 2011). As for unguided ICBT, there are currently numerous promising examples of delivering treatment that is more or less self-guided. Overall, results and adherence are usually somewhat lower when compared to guidance (Baumeister, Reichler, Munzinger, & Lin, 2014), but recent investigations show that ICBT can be administered without support with similar benefits as guided ICBT (c.f., Titov et al., 2016).

Efficacy and effectiveness

A number of systematic reviews and meta-analyses have determined the efficacy of ICBT for various psychiatric disorders and somatic illnesses, with results being in the moderate to large range (Cohen's d) (Andersson, 2016). In terms of MDD, the between-group effect size at post treatment assessment is $d = 0.41$, but with greater benefits for guided ICBT, $d = 0.61$, than unguided ICBT, $d = 0.25$, when compared to wait-list control (Andersson & Cuijpers, 2009). Similar results were later obtained by Johansson and Andersson (2012), indicating a relationship between the amount of support being provided and treatment outcome; $d = 0.21$ for no guidance at all, $d = 0.44$ for contact solely during the screening procedure, and $d = 0.76$ for contact as well as guidance throughout the treatment period, comparing ICBT to wait-list control, discussion group, or care as usual. Hedman et al. (2014) have also published a review of patients undergoing ICBT in routine outpatient care, showing large improvements, within-group $d = 1.27$, and Andersson, Hesser, Hummerdal, Bergman-Nordgren, and Carlbring (2013) assessed its long-term effects, with benefits maintained at 3.5 year follow-up, although it should be pointed out that few clinical trials include such investigations. As for anxiety disorders, most findings with regard to efficacy and effectiveness stems from specific clinical trials (Andersson, 2015). For example, Andersson, Carlbring, Ljótsson, and Hedman (2013) reviewed a number of studies of ICBT for these conditions, revealing significant benefits, e.g., between-group $d = 1.00-1.10$ for PD/A, compared to wait-list control. In addition, a three year follow-up illustrated that treatment outcome was sustained over time, between-group $d = 0.70$, compared to wait-list control (Ruwaard, Broeksteeg, Schrieken, Emmelkamp, & Lange, 2010), and a study conducted in routine outpatient care showed comparable results, within-group $d = 1.20$ (Ruwaard, Lange, Schrieken, Dolan, & Emmelkamp, 2012). With regard to Social Anxiety Disorder (SAD), the findings suggest that large gains are made in treatment, between-group $d = 0.70-0.95$, when comparing against wait-list control

(Andersson, Carlbring, et al., 2013). A five-year follow-up provide evidence for maintained results (Hedman, Furmark, et al., 2011), and Hedman, Andersson, et al. (2011) have also found support for its implementation in routine outpatient care, within-group $d = 1.42$. Although not as explored, other anxiety disorders have been delivered and assessed via the Internet, such as, GAD, Posttraumatic Stress Disorder (PTSD), phobias, and OCD, with comparable benefits. Likewise, ICBT has been studied in relation to severe health anxiety, eating disorders, insomnia, pathological gambling, among others, as well as for children, youths, and among older populations, also with promising results (Lindfors & Andersson, 2016). Also, ICBT has been found to be comparable to CBT when delivered face-to-face for many psychiatric disorders, and has obtained promising results in relation to the management of many somatic illnesses, such as, irritable bowel syndrome, tinnitus, and sexual dysfunctions (Andersson, Cuijpers, Carlbring, Riper, & Hedman, 2014).

Negative effects of ICBT

Negative effects of psychological treatments has not been particularly well explored in either clinical trials or routine outpatient care, and even less so with regard to treatments delivered via the Internet. As such, the knowledge of their occurrence and characteristics has been lacking, with no reports of their incidence or suggestions on how to assess such events during treatment. Hence, in an attempt to improve the situation and provide recommendations on the monitoring and reporting of negative effects, researchers in the field of ICBT were invited to share their view and discuss theoretical concepts and practical ideas on the topic. The results of this process were then used to develop a consensus statement, including propositions of how to define and examine negative effects, as well as examples of their probable causes (Rozental et al., 2014). In essence, there was an agreement that the issue had not received enough attention, although the absence of any empirical data made it difficult to refer to negative effects as a problem in ICBT. However, given insufficient evidence, further investigation of such events was advised: “Researchers need to be mindful of the fact that some patients could experience adverse events and deteriorate during treatment, and systematically probe and report negative effects when performing clinical trials of Internet interventions.” (p. 14; Rozental et al., 2014). Thereby, the awareness among researchers and clinicians about the potential risks of ICBT should improve, in turn, establishing how frequently such events seem to occur. In addition, monitoring and reporting negative effects on a regular basis might also help to distinguish what is specific for this type of format and what is similar to psychological treatments provided face-to-face. On the one hand, it could be argued that ICBT is only a means of delivery and that its content should not

differ from CBT. Accordingly, the negative effects ought to be the same, which would further the understanding of their features in general. On the other hand, given the circumstances through which ICBT is distributed, i.e., no or few physical meetings, limited feedback, and being based entirely on self-help, there may be negative effects that are specific for such a treatment. However, similar to psychological treatments given face-to-face, the issue of causality remains the same. Rozental et al. (2014) therefore recommended the use of randomized controlled trials in order to determine if there could be a difference between treatment and control conditions, which may be useful for finding out what is caused by the treatment interventions and what is related to other circumstances. Furthermore, the importance of applying a mutual terminology was put forward as important to facilitate research on negative effects, for example, deterioration, non-response, and adverse and unwanted events. A number of ways for monitoring and reporting such events were also proposed, for instance, the implementation of the RCI or cutoffs for investigating deterioration, open-ended questions or self-reports measures for targeting the experiences of the patients, or a structured clinical interview as well as qualitative means for exploring different perspectives. Given the lack of research on the topic, specific variables warranting caution or further inquiry was not possible to provide, but Rozental et al. (2014) suggested several issues as reasonable to pursue with regard to exploring the causes of negative effects, for instance, disappointment, deficient treatment, therapist factors, therapeutic orientation, and patient characteristics.

The consensus statement is believed to have raised awareness of negative effects among researchers and clinicians in ICBT. Moreover, it has provided a conceptual framework through which the ensuing studies of the present thesis have been conducted. However, whether or not it has been influential in terms of generating more investigations and improving the knowledge of these incidents has been unclear. Thus, a review of all studies referencing the consensus statement since its issue in March 2014 was performed. As of the completion of the present thesis in August 2016, there were 62 publications citing the consensus statement according to the scientific database Scopus. Of these, 2 (3.2%) were a part of the present thesis and thus excluded for the purpose of this summation. An additional 5 (8.1%) were study protocols and did not include any empirical data. Also, 4 (6.5%) were systematic reviews or meta-analyses, 5 (8.1%) were theoretical reviews or non-empirical texts, and 8 (12.9%) were qualitative studies or surveys. Hence, the remaining 38 (61.3%) were deemed to be of such a nature that reporting harm would have been possible. Of these, 23 (60.5%) did include some type of information on negative effects. A summary of these results can be obtained in Table 1.

Table 1

Rates of deterioration, none-response, and adverse and unwanted events among clinical trials of ICBT and the means for reporting such incidents

	Deterioration		Non-response		Adverse and Unwanted Events	
	Rate ¹	Assessment	Rate	Assessment	Rate	Assessment
Bachem and Maercker (2016)	9 (8.7%)	Increase in symptoms	2 (1.9%)	No change in symptoms		
Kayrouz, Dear, Karin, Fogliati, and Titov (2016)	0 (0%)	Cutoff				
Hadjistavropoulos et al. (2016)	5-8 (1.1-1.7%)	RCI	170-241 (37.1-52.6%)	RCI		
Nieminen et al. (2016)			3 (7.0%)	No change in symptoms	20 (46.5%)	n.a.
Klein et al. (2016)	101-107 (19.8-21%)	Increase in symptoms				
Ebert et al. (2016)	9-11 (3.4-4.2%)	RCI				
Weise, Kleinstäuber, and Andersson (2016)	1-2 (1.6-3.2%)	RCI				
Nordgreen et al. (2016)	1-2 (0.5-1.2%)	RCI				
Dahlin et al. (2016)	0 (0%)	n.a.				
Saekow et al. (2015)	3 (21.4%)	Increase in symptoms				
Nordmo et al. (2015)	1 (2.7%)	RCI	10-13 (33.3-44.8%)	Cutoff		
Dear et al. (2015)	1-13 (2-17%)	Cutoff				

	Deterioration		Non-response		Adverse and Unwanted Events	
	Rate ¹	Assessment	Rate	Assessment	Rate	Assessment
Kaldo, Jernelov, et al. (2015)	0 (0%)	n.a.	35-57 (23.7-38.5%)	n.a.	11-16 (7.4-11%)	Open-ended question
Rozental, Forsell, Svensson, Andersson, and Carlbring (2015)	1-7 (0.7-4.7%)	RCI			13 (8.6%)	Open-ended question
Alfonsson, Olsson, and Hursti (2015)	0-3 (0-3%)	RCI				
Blom et al. (2015)	1-2 (2.1-4.2%)	Cutoff or increase in symptoms			9 (19%)	Open-ended question
Rheker, Andersson, and Weise (2015)	0-2 (0-1.8%)					
Hovland et al. (2015)	1 (25%)	Increase in symptoms			0 (0%)	Interview
Buntrock et al. (2015)	24-26 (4.6-5.9%)	RCI				
Lancee, Eisma, van Straten, and Kamphuis (2015)	10 (15.8%)				0 (0%)	Open-ended question
Hilgart et al. (2014)					1 (12.5%)	Self-report measure
Kivi et al. (2014)	8 (8.7%)	n.a.				
Enander et al. (2014)			² (5-10%)	CGI	11 (48%)	Open-ended question

ICBT = Internet-Based Cognitive Behavior Therapy; RCI = Reliable Change Index; n.a. = not available; CGI = Clinical Global Impressions

¹ Numbers are intention-to-treat and vary depending self-report measure, time point, and/or condition

² Only proportions were reported in the original study

Although this investigation was not performed following the principles of a systematic review, the results are considered encouraging, but reveal some of the difficulties associated with assessing negative effects, providing a backdrop for the present thesis. Albeit promising with regard to offering information about the risks of ICBT, there was a large variation in terms of how negative effects were in fact monitored. The RCI or a predetermined cutoff was often used to determine the occurrence of deterioration, in line with the recommendations by Rozental et al. (2014), but some relied on any increase in symptomatology or did not specify how such a worsening had been established, making it difficult to assess its reliability or compare rates across clinical trials. The same was found for non-response, which was less frequently reported. As for adverse and unwanted events, there were a few cases assessing such negative effects, mostly involving the distribution of open-ended questions. However, in comparison to deterioration, these were not as common to explore. Furthermore, in reviewing the investigations of negative effects, there were some difficulties related to separating treatment to control conditions in the presentation of the results, as well as the use of different procedures for handling missing data, i.e., intention-to-treat or complete case analyses, which could complicate the study of deterioration and non-response rates. Nonetheless, despite these limitations, the findings from this review might suggest that researchers and clinicians are becoming more aware of the need to monitor and report negative effects in ICBT.

Despite an increased awareness of negative effects in ICBT, the current understanding of their characteristics is unsatisfactory. Further research needs to be done in order to fully comprehend their features and what might be contributing to their incidents. In addition, assessing their frequency and comparing them to those occurring in psychological treatments delivered face-to-face could perhaps help to elucidate potential differences as well as similarities between the formats. This is deemed particularly important given that ICBT is starting to become more available in the routine outpatient care in many countries, providing a promising treatment alternative for many psychiatric disorders and somatic illnesses (Emmelkamp et al., 2014). Thus, examining possible detrimental effects parallel to its many benefits should be considered vital, especially with regard to the ethical issues surrounding the introduction of new ways of distributing psychological treatments. Moreover, in comparison to a face-to-face contact, studying negative effects in ICBT might also have numerous practical and theoretical advantages. Firstly, since data is entered by the patients themselves via computer-based self-report measures rather than paper-and-pencil, the risk of data loss and data distortion can be averted (Thorndike et al., 2009), potentially making information more accessible for analyses of deterioration and non-response. Novel methods for investigating negative effects can also easily be included in clinical trials as the distribution of self-report measures and open-ended

questions is administered online, facilitating the exploration of adverse and unwanted events. Secondly, because a large number of complete data sets already exist, performing comprehensive assessments of negative effects can be conducted quickly and efficiently, without first having to recruit patients and follow through with treatment. Thirdly, as the nature of ICBT involves a predetermined set of modules for a given psychiatric disorder, content and provision should be less susceptible to confounders than psychological treatments delivered face-to-face, avoiding some of the complications related to therapist drift and inadequate treatment integrity. This would especially be true for unguided ICBT, although even with guidance, the influence from a clinician could be presumed to be less than having a face-to-face contact. Assuming that this is correct, data gathered from ICBT, particularly on a week-by-week basis, could then be used to explore and pinpoint the impact of specific aspects related to content and the treatment interventions, e.g., whether there are fluctuations in deterioration, non-response, and adverse and unwanted events depending on what modules are introduced. This does not preclude the importance of carefully studying the role of the therapeutic relationship or features of the clinician in a face-to-face contact that may be detrimental for the patient. Instead, it should be regarded as a complement that could help overcome the complexities involved in understanding the multifaceted nature of negative effects of psychological treatments. In sum, even though the present thesis primarily intends to assess the occurrence and characteristics of incidents that are perceived as negative in ICBT, there are reasons to believe that the ensuing results and forthcoming investigations can be of great value in the research of negative effects of psychological treatments in general.

Aims of the thesis

The overall aim of the present thesis was to determine the occurrence and characteristics of negative effects in ICBT, using quantitative and qualitative methods and different samples of patients. The purpose was to further the understanding of how negative effects can be assessed, what their features are in ICBT, and if there are variables that predict their incidence. The thesis consists of four studies, each with a unique contribution to the field.

Study I

The intention with the first study in the thesis was to conduct the first ever investigation of negative effects in ICBT, using a sample of 133 patients in a clinical trial of ICBT that also included Cognitive Bias Modification (CBM) for SAD. The rates of deterioration and non-response were determined, open-ended questions were posed to explore the patients' experiences of adverse and unwanted events, and predictors of both were examined.

Study II

The purpose of the second study in the thesis was to perform an exploration of how negative experiences were perceived and to what extent they had a negative impact. In total, 556 patients in four separate clinical trials were asked to respond to open-ended questions on adverse and unwanted events, which were analyzed using a qualitative and inductive approach.

Study III

The objective of the third study in the thesis was to determine the occurrence of deterioration in ICBT at large, using the individual scores of 2866 patients in 29 clinical trials that were allocated to either receiving treatment or being in a control condition. Moreover, predictors of deterioration were assessed based on a number of variables that were chosen a priori.

Study IV

The goal of the fourth and final study in the thesis was to utilize the findings from study I-III, together with a literature review and a consensus statement among researchers in the field of ICBT (Rozenal et al., 2014), to develop and test a newly developed self-report measure for monitoring adverse and unwanted events, the Negative Effects Questionnaire (NEQ).

Method

Participants

Study I

Patients with SAD were recruited via the Internet and regional and national newspapers in Sweden. Individuals interested in participation applied through email and were asked to complete the screening procedure. In total, 282 patients filled out the self-report measures, of whom 179 patients were deemed eligible for inclusion and contacted in order to undergo a structured clinical interview. Inclusion and exclusion criteria for the study were; a) being at least 18 years old (of adult age according to Swedish regulations), b) having access to a computer with Internet c) meeting diagnostic criteria for SAD, d) no suicidal ideation e) an error rate of less than 25% in the first attention bias assessment, f) not receiving any other psychological treatment during the clinical trial, and g) no addition or modification of medication for psychiatric disorders, i.e., stable dosage at least three months prior to the treatment period. A total of 133 patients fulfilled all the criteria and were subsequently randomized to either ICBT with CBM ($N = 66$) or ICBT with control training ($N = 67$) (Boettcher, Hasselrot, Sund, Andersson, & Carlbring, 2013). A description of the recruitment and screening process can be obtained in the study protocol (Boettcher, Andersson, & Carlbring, 2013).

Study II

Patients in four separate clinical trials of ICBT were distributed open-ended questions regarding adverse and unwanted events at the end of treatment. Individuals were recruited via the Internet as well as regional and national newspapers in Sweden for the following psychiatric disorders or condition; SAD and/or PD/A ($N = 130$) delivered via a smartphone (Lindner, Ivanova, Ly, Andersson, & Carlbring, 2013), MDD ($N = 143$) distributed as guided ICBT or physical activity (Ström et al., 2013), SAD ($N = 133$) administered with or without CBM (Boettcher, Hasselrot, et al., 2013), and procrastination ($N = 150$) given as guided or unguided ICBT (Rozenal et al., 2015). In total, 556¹ patients were included. In terms of inclusion and exclusion criteria, screening procedure, and use of a structured clinical interview, these were

¹ Due to an error in Study II, the total number of participants was referred to as 558 when it should be 556.

the same as Study I, albeit with minor differences related to the focus of the given clinical trial, e.g., the type of self-report measures being distributed as well as the diagnostic cutoffs.

Study III

Patients from 29 clinical trials of ICBT were included, using the individual scores of $N = 2866$; 2118 (73.9%) in treatment and 748 (26.1%) allocated to some form of control condition. Three classes of psychiatric disorders or issues that may warrant clinical attention were included; anxiety disorders (PTSD², PD/A, and anxiety with/without depression), MDD (with/without dysthymia), as well as other (erectile dysfunction, relationship distress with spouse or intimate partner, and gambling disorder). An absolute majority were recruited via the Internet as well as regional and national newspapers in Sweden, with two exceptions; one clinical trial conducted in primary care and one in a university setting. Inclusion and exclusion criteria, screening procedure, and use of a structured clinical interview were the same as Study I-II, with small differences due to the focus of the specific clinical trial.

Study IV

Patients were recruited by two different means; a clinical trial of ICBT for SAD delivered via a smartphone ($N = 189$) (Miloff, Marklund, & Carlbring, 2015), and through the national media, i.e., an article on negative effects in the largest morning newspaper in Sweden as well as a feature on a national radio show on science on the Swedish public radio ($N = 464$). With regard to the former, inclusion and exclusion criteria, screening procedure, and use of a structured clinical interview were the same as Study I-III, but with slight differences related to the focus of the clinical trial. As for the latter, the only inclusion criterion was that the individual had undergone psychological treatment during the last two years (not necessarily ICBT), or, alternatively, was currently seeing a mental health professional. In sum, a total of 653 patients completed a newly developed self-report measure for monitoring adverse and unwanted events, the NEQ (Rozenal, Kottorp, Boettcher, Andersson, & Carlbring, 2016).

Ethical considerations

Ethical issues associated with the investigation of negative effects of ICBT were considered in relation to each of the studies included in the thesis. Firstly, the clinical trials were subjected to examination and had received ethical approval by the Regional Ethical Review Board in their respective

² The clinical trial of PTSD in Study III was performed using the *Diagnostics and Statistical Manual of Mental Disorders*, 4th Edition, where it was still classified as an anxiety disorder.

study location in Sweden. Thus, both the risk of deterioration and potentially harmful incidents had been reviewed prior to the commencement of any treatment intervention. This was relevant for Study I-IV, as they all were comprised of, or, to some extent, involved data where the patients had been randomized to either treatment or a control condition. Secondly, all of the patients were given an automatically derived identification code, e.g., 1234abcd, making it impossible to disclose the identity of an individual. Also, all clinical trials used a secure online interface requiring registration and electronic identification, i.e., Secure Sockets Layer certificate, as well as a two-step verification process, i.e., username and password and a randomly generated pin code, comparable to the systems used by many banks (Bennett, Bennett, & Griffiths, 2010). As for Study II, which consisted of a qualitative content analysis of the responses to open-ended questions on negative effects, data was examined and the results presented in a way that ensured no personal information would be exposed, i.e., removing sensitive or private details. Thirdly, written informed consent was collected in Study I-IV, either via post or submitted digitally, as a way of guaranteeing that the patients understood and agreed to the circumstances of participation. Lastly, with regard to the ethical issues associated with investigating deterioration, Study I and III consisted of data from completed clinical trials, making it impossible to, in hindsight, detect and intervene if a particular patient fared worse during the treatment period. However, as the overall aim of the thesis is to determine the occurrence and characteristics of negative effects, it is believed that the findings in the thesis could improve the monitoring and reporting of deterioration in ICBT, which may help to reverse negative treatment trends in the future.

Measures

Study I

Patients completed a screening procedure that involved sociodemographic information and five different self-report measures; the Liebowitz Social Anxiety Scale - Self-Report (LSAS-SR; Liebowitz, 1987), the Social Phobia Scale as well as the Social Interaction Anxiety Scale (Mattick & Clarke, 1998), the Montgomery-Åsberg Depression Rating Scale - Self-Report (MADRS-S; Svanborg & Åsberg, 1994), and the Quality of Life Inventory (QOLI; Frisch, Cornell, Villanueva, & Retzlaff, 1992). A structured clinical interview was also performed with those patients that were deemed eligible for inclusion, using the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I; First, Gibbon, Spitzer, & Williams, 1997). Furthermore, the self-report measures were distributed at mid assessment (2 weeks), post treatment assessment (11 weeks), and follow-up (4 months), and were used in the investigation of deterioration and non-response rates. The SCID-I was,

however, only utilized for screening purposes. Moreover, patients were asked at mid and post treatment assessment whether they had experienced adverse and unwanted events that might be associated with the treatment interventions (yes/no)³. If yes, the patients were requested to describe the incidence in their own words, and to consider its degree of negative impact on a Likert-scale that ranged 0-3 (0 indicating no effect at all, 3 representing a severe effect). An adapted version of the UE-ATR (Linden, 2013) was used to code the responses in order to determine if the adverse and unwanted events were in fact related to the treatment interventions, scored on a Likert-scale that ranged 0-5 (1 implying no relationship, 5 suggested it was related). Data was entered by the patients themselves on a secure online interface, minimizing the risk of data loss and data distortion (Thorndike et al., 2009), except for the SCID-I, which was conducted via telephone.

Study II

Patients filled out the sociodemographic information and self-report measures that were used in each of the clinical trials, similar to the screening procedure described for Study I. A structured clinical interview was also conducted using the SCID-I, except for one clinical trial of procrastination (Rozenal et al., 2015). None of these instruments were, however, used for the purpose of Study II, which relied on open-ended questions that were distributed at the post treatment assessment (8-11 weeks). All patients were instead asked about adverse and unwanted events occurring throughout the treatment period, i.e., “Have you, during the course of treatment experienced any unwanted events that you believe are related to treatment, or have you encountered any unwanted effects that could be attributable to treatment?”. The patient was also required to assess its negative impact on a Likert-scale that ranged 0-3, same as Study I, as well as to determine its effect when it happened, “...in relation to when they occurred?”, and now, “...in relation to how you feel today?”. Had the patient experienced more than one incidence, separate responses could be given for each of them. Data was entered by the patients themselves on a secure online interface, except for the SCID-I, which was conducted via telephone, similar to study I.

Study III

Patients finished the screening procedure for each of the clinical trials that were included in the ensuing Individual Patient Data Meta-Analysis (IPDM), which contained sociodemographic information and self-report measures similar to Study I-II. These were filled out also at post treatment assessment (4-12 weeks). A structured clinical interview was performed to determine eligibility for inclusion, most often being the SCID-I, albeit in four cases

³ Study I utilized the term side effects before it was established what definitions to use in the present thesis, but refers to the same incidents involved in adverse and unwanted events.

another diagnostic interview was used instead, e.g., the MINI-International Neuropsychiatric Interview (Sheehan et al., 1998). The primary outcome measure for each clinical trial was used for the investigation of deterioration rates, depending on what type of psychiatric disorder was being examined, e.g., the Beck Depression Inventory (Beck, Erbaugh, Ward, Mock, & Mendelsohn, 1961), the Penn State Worry Questionnaire (Meyer, Miller, Metzger, & Borkovec, 1990), and the Spider Phobia Questionnaire (Klorman, Weerts, Hastings, Melamed, & Lang, 1974). A complete list of all self-report measures can be found in Study III. Like Study I-II, data was entered by the patients themselves on a secure online interface, except for the structured clinical interviews, which were conducted via telephone.

Study IV

Patients recruited from the clinical trial completed a screening procedure similar to study I-III, which involved sociodemographic information and several self-report measures; LSAS-SR, the Generalized Anxiety Disorder – 7 Items (Spitzer, Kroenke, Williams, & Löwe, 2006), the Patient Health Questionnaire – 9 Items (Kroenke, Spitzer, & Williams, 2001), the QOLI, the Brunnsvikien Brief Quality of Life Inventory (Lindner et al., 2016), and the Mini-Social Phobia Inventory (Connor, Kobak, Churchill, Katzelnick, & Davidson, 2001). In addition, a structured clinical interview was performed to assess eligibility, i.e., the MINI. The self-report measures were filled out again at mid assessment (3 weeks) and post treatment assessment (6 weeks), and, as for the latter, together with the NEQ. Patients recruited through the media, however, only submitted sociodemographic information together with details regarding what type of psychological treatment they had or were currently receiving, the NEQ, and the Client Satisfaction Questionnaire - 8 Items (Larsen, Attkisson, Hargreaves, & Nguyen, 1979), the last being used in a psychotherapist's bachelor thesis. The NEQ, as distributed in Study IV, consisted of 60 items that were generated using the consensus statement for monitoring and reporting in ICBT (Rozental et al., 2014), the results from Study I-III, and a literature review. The NEQ contains three sections; one for endorsing specific items, "Did you experience this?", one pertaining their negative impact, "If yes – here is how negatively it affected me", scored on a Likert-scale that ranged 0-4 (0 indicating no effect at all, 4 representing an extreme effect), and one regarding what the negative effects is attributed to, "Probably caused by", which separated "The treatment I received" from "Other circumstances". Albeit, for the purpose of Study IV, only data on treatment-related negative effects were used. As for Study I-III, data was entered by the patients themselves on a secure online interface, except for the structured clinical interviews, which were conducted via telephone.

Analyses

Study I

Analyses consisted of assessing the number of patients that had deteriorated, defined as exceeding the RCI in a negative direction, as well as those not responding to treatment, i.e., having a change score between two points of measurements that is within the boundaries of the RCI in either direction, e.g., ± 5 points. In order to investigate deterioration and non-response, all of the self-report measures at pre, mid and post treatment assessment, as well as follow-up were used. Both complete case and intention-to-treat analyses were utilized, with last-observation-carried-forward being implemented to account for missing data. In terms of the reliability estimates for the RCI, Cronbach's α for each of the self-reported measures were derived from a clinical trial of ICBT for SAD by Hedman et al. (2010). Logistic regression was subsequently used to investigate possible predictors of faring worse, with a dummy coding for deterioration (yes/no) as the dependent variable. Similarly, logistic regression was conducted to assess potential predictors of the adverse and unwanted events that were reported by the patients, again using a dummy coded variable (yes/no). In terms of whether or not these were associated with treatment, two independent raters indicated if they might be attributed to the treatment interventions, with Cohen's Kappa being utilized as a measure of inter-rater agreement.

Study II

Analyses were comprised of descriptive statistics for the average degree of negative impact caused by the self-reported negative effects, as well as a qualitative content analysis of the responses made by the patients. In terms of the latter, an inductive, hypothesis generating approach was used in order to investigate patterns and concepts in the data, as suggested by Elo and Kyngäs (2008).

Study III

Analyses included determining the number of patients exceeding the RCI on the primary outcome measure in each clinical trial, as assessed by the change score between pre and post treatment assessment for each individual. A RCI of 0.84 was used as an indication of mild deterioration, as recommended by Wise (2004). Reliability estimates for each of the self-report measures were derived from test-retest reliabilities with a short time period between the two points of measurement, and, if possible, a normal population, as suggested by Edwards, Yarvis, Mueller, Zingale, and Wagman (1978). In addition, intention-to-treat was implemented for all analyses, with multiple imputation being used to account for missing data. Generalized linear mixed models were utilized to examine the change scores, and a binomial logistic model was performed to assess possible predictors of deterioration. Dummy coding

(yes/no) was used as the dependent variable in order to identify each patient exceeding the RCI in a negative direction.

Study IV

Analyses involved an exploratory factor analysis, utilizing principal axis factoring with an oblique rotation. The Kaiser-Meyer-Olkin measure of sampling adequacy, the Bartlett's Test of Sphericity, and the Determinant, were also used to check the suitability to perform the analyses. In addition, item-item correlations, multicollinearity, Eigenvalues and the scree test were implemented to investigate the number of items as well as factors to retain, together with a parallel analysis, i.e., comparing the factor solution with one derived from data that is produced at random, and a stability analysis, i.e., retesting the results on a randomly selected subsample derived from the data.

Results

Study I

Results indicated that deterioration, as assessed using the RCI, was achieved by 0.8-5.3%⁴ at the mid assessment, when using the primary outcome measures of SAD, as compared to 0.8-5.3% for the other outcome measures. These numbers were lower at the post treatment assessment, nil for SAD and 6.8% for the MADRS-S, as well as the follow-up, 0.8-2.3% for SAD, 1.5% for the QOLI, and 3.8% for the MADRS-S. As for non-response, the rates were 69.9-86.5% for SAD and 80.5-97% for the other outcome measures at mid assessment; 32.3-49.6% for SAD and 57.1-89.5% for the other outcome measures at post treatment assessment; as well as 29.3-50.4% for SAD and 66.9-86.5% for the other outcome measures at follow-up. With regard to self-reported adverse and unwanted events, 4.5% had experienced negative effects at mid assessment, with a moderate negative impact, $M = 1.25^5$ ($SD = 0.96$), and plausible relatedness to treatment, $M = 4.0^6$ ($SD = 1.07$). At post treatment assessment this number was 12.9%, with a similar negative impact $M = 1.59$ ($SD = 1.12$), and possible relatedness to treatment, $M = 4.38$ ($SD = 0.85$). Predictors of deterioration were also examined, albeit only indicating that clinical severity on the LSAS-SR at screening was related to increased odds of deterioration, OR 1.07, 95% CI [1.02, 1.13]. Predictors of adverse and unwanted events, as reported by the patients themselves, were, in turn, investigated, however, none revealing a potential relationship.

Study II

Results showed that 9.3% of the patients reported adverse and unwanted events during the treatment period, range 4.9%-12.6%. The average negative impact at the time of their occurrence was moderate, $M = 1.80^5$ ($SD = 1.00$), range 1.62-2.20, and smaller when the patients were asked to consider their long-term effects, $M = 1.07$ ($SD = 1.05$), range 0.81-1.32. The qualitative content analysis, in turn, revealed two categories and four subcategories; 1)

⁴ Rates in Study I span over different numbers depending on which self-report measure is being used.

⁵ Scored on a Likert-scale that ranged 0-3 (0 indicating no effect at all, 3 representing a severe effect).

⁶ Scored on a Likert-scale that ranged 0-5 (0 indicating no relationship with treatment, 5 being related).

patient-related negative effects, associated with a greater self-awareness and understanding about one's condition, a) insight, as well as experiencing new symptoms that were unexpected by the patients, e.g., anxiety, insomnia, and physical complaints, b) symptoms. In addition, 2) treatment-related negative effects also emerged, linked to a sense of failure, having a bad conscience, and facing difficulties when implementing the treatment interventions, a) implementation, as well as problems that were directly associated with the treatment format, e.g., activity level, time pressure, and feeling frustrated when meeting obstacles or not receiving adequate feedback, b) format.

Study III

Results revealed that among the 2866 patients that were included in the IPDM of 29 clinical trials of ICBT, 122 (5.8%) deteriorated in treatment, defined as exceeding the RCI on the primary outcome measure, compared to 130 (17.4%) in the control conditions. Moreover, predictors of deterioration were also explored, indicating that, relative to treatment, being in some form of control condition was associated with increased odds of deterioration, OR 3.10, 95% CI [2.21, 4.34]. Furthermore, clinical severity at pre treatment assessment was related to lower odds, OR 0.62, 95% CI [0.50, 0.77], and OR 0.51, 95% CI [0.51, 0.80], for treatment and control conditions, respectively. Moreover, being in a relationship, OR 0.58, 95% CI [0.35, 0.95], having a university degree, OR 0.54, 95% CI [0.33, 0.88], and being older, OR 0.78, 95% CI [0.62, 0.98], were all associated with lower odds of deterioration, albeit only for patients in treatment.

Study IV

Results suggested a six-factor solution, retaining 32 of the original 60 items; symptoms, quality, dependency, stigma, hopelessness, and failure, in total accounting for 57.64% of the variance. In terms of the most endorsed items, 38.4% of the patients stated "Unpleasant memories resurfaced" (Item 13), 37.7% reported "I felt like I was under more stress" (Item 2), and 37.2% described "I experienced more anxiety" (Item 3). Likewise, the items that had the highest self-rated negative impact were; "I felt that the quality of the treatment was poor" (Item 29), $M = 2.81^7$ ($SD = 1.10$), "I felt that the issue I was looking for help with got worse" (Item 12), $M = 2.68$ ($SD = 1.44$), and "Unpleasant memories resurfaced" (Item 13), $M = 2.62$ ($SD = 1.19$).

⁷ Scored on a Likert-scale that ranged 0-4 (0 indicating no effect at all, 4 representing an extreme effect).

Discussion

Negative effects belong to an essentially uncharted territory in the research of psychological treatments. Perhaps due to the ambition of developing and providing effective treatment interventions for mental distress, the possibility of inadvertently causing harm was rarely considered, and thus overlooked. Evaluating their efficacy in clinical trials and disseminating new ways of helping patients overcome psychiatric disorders were instead prioritized, while the backside of their implementation was seldom discussed and hardly ever investigated. Psychological treatments were, in the end, supposed to alleviate suffering and increase well-being, not induce negative effects, and, should at worst only result in non-response, but never ever cause harm. However, Powers and Witmer (1951) provided the first empirical evidence that in some cases treatment interventions meant to do good could result in unwanted implications, leading to outcomes that were the opposite of what was originally intended. Early on, it also became evident that psychological treatments that were delivered to improve the condition of many patients did not always produce the same results for everyone, suggesting that not all seem to benefit from their application, and that some might even fare worse. Bergin (1966) named this the “client-deterioration phenomenon” (p. 236), or, deterioration effect, sparking a debate on whether or not psychological treatments could in fact leave some patients feeling worse than they were prior to entering treatment. Following this initiative, research from a number of settings and including diverse samples have indicated that deterioration do seem to occur in about one-tenth of all patients, warranting strategies for detecting and reversing such a negative treatment trend (Jarrett, 2008). Still, what aspects might be contributing to its event remain largely unresolved, and it is not yet clear if this deterioration should be attributed to treatment or if there are other circumstances in the patients’ lives that are responsible. Similarly, natural variations inherent in many psychiatric disorder as well as measurement error could also be involved, making the investigation of its causes a complicated matter. Meanwhile, researchers and clinicians like Strupp et al. (1977), Stuart (1970), and Mays and Franks (1985), pointed out that other adverse and unwanted events might occur beside deterioration, such as, new symptoms, lowered self-esteem, and interpersonal difficulties, warranting broader and more diverse ways of looking at negative effects in order to truly understand their nature.

The present thesis could be seen as an extension of this ongoing debate, with the ambition of furthering the current knowledge of how frequently deterioration occurs and what may be responsible for its incidence. It also aims to provide a more thorough and systematic assessment of adverse and unwanted events, as experienced by patients themselves - aspects that are far less explored than deterioration in the research of psychological treatments. Moreover, the present thesis puts the study of negative effects in a somewhat different context, determining their occurrence and characteristics in relation to treatment interventions delivered via the Internet rather than face-to-face. As new means of disseminating psychological treatments are developing and becoming more and more available, inspecting the potential problems with their implementation is essential. Providing evidence for their benefits is only the first step in helping patients gain access to an effective type of care for psychiatric disorders. Just as important is the examination of for whom the treatment interventions does not work, as well as how often and why some patients seem to deteriorate or experience adverse and unwanted events during the course of their treatment. Hopefully, the results from the present thesis will aid researchers and clinicians in becoming more aware of the possibility that treatment interventions administered via the Internet and face-to-face can have negative effects for some patients, and that it is crucial to monitor and report their incidence in order to identify, help, and prevent those that are at risk of faring worse. Below, the findings from the present thesis are discussed in relation to prior research and theoretical concepts, while its limitations and the need for further study are reviewed.

The deterioration effect revisited

Deterioration can be defined as the worsening of a given psychiatric disorder or increase in symptomatology (Dimidjian & Hollon, 2010), and is probably the most straightforward and reliable way of determining negative effects in psychological treatments. Hence, it is easy to understand why it has been the focus of most reports on detrimental events; deterioration is less open to subjective notions of what constitute negative effects, statistical procedures or predetermined cutoffs offer clear guidelines on how to identify when a patient has fared worse, and reporting deterioration rates give researchers and clinicians numbers to compare and analyze when investigating negative effects between as well as within clinical trials. However, as reviewed in Rozentel et al. (2014), deterioration can be assessed differently depending on the particular time frame, self-report measure, and calculation of the RCI that is being used, particularly given the absence of other suitable means for distinguishing the negative analogy to clinically significant change (Jacobson, Follette, & Revenstorf, 1984). Both theoretical and statistical considerations thus have to be made prior to commencing the assessment of

deterioration, which, in turn, can affect the proportion of patients that exceed the endpoint for faring worse. With this in mind, the present thesis began the first ever exploration of deterioration in ICBT; in relation to a clinical trial for SAD that also included a component of CBM, Study I, as well as with regard to the individual scores of 2866 patients in 29 clinical trials, Study III.

In the first case, deterioration rates varied over time, with 0.8-5.3% at mid assessment for both the primary and other outcome measures, nil at post treatment assessment for the self-report measures of SAD and 6.8% for the MADRS-S, and 0.8-2.3% for the primary outcome measures at follow-up, as compared to 1.5% for the QOLI and 3.8% for the MADRS-S. In relation to psychological treatments delivered face-to-face, 5-10% of all patients are assumed to deteriorate, at least when it comes to adults with one of the more prevalent conditions, such as MDD or anxiety disorders (Hansen et al., 2002), indicating that the deterioration rates in Study I are comparable to what would be expected in those settings. Recently, clinical trials of ICBT have started reporting negative effects along with their positive outcomes, which give some credence to these results, for example, 1-3% for comorbid psychiatric disorders (Titov et al., 2016) as well as 1.6-3.2% for MDD (Andersson, Topooco, Havik, & Nordgreen, 2016). However, it should be noted that there also exist examples with far greater deterioration rates, e.g., up to 17% for chronic pain (Dear et al., 2015) and 15.8% for insomnia (Lancee et al., 2015), which implies that the numbers probably vary a lot depending on the treatment interventions provided, what condition is treated, and the characteristics of an individual. In terms of the variations in patients faring worse that can be observed over time in Study I, these patterns need to be scrutinized in additional clinical trials before any definitive conclusions can be drawn of their significance. Although, it is reasonable to assume that certain treatment interventions implemented at certain occasions may affect patients differently, causing fluctuations in symptomatology that would be detected on the self-report measures. For instance, this should be expected from the gradual exposure to social situations or stimuli, which is a central element of CBT. As discussed by Castonguay, Boswell, Constantino, Goldfried, and Hill (2010), some parts of treatment will always be associated with more symptoms and an increase in distress, such as, when confronting fearful thoughts or emotions in exposure with response prevention in OCD, even though it is likely to be a transient phenomenon and beneficial in the long run. However, if the deterioration is maintained, or, possibly even arise to post treatment assessment or follow-up, this change may reflect a negative effect that is detrimental to the patient. In Study I, a small proportion was in fact identified as deteriorated at both of these time points, suggesting that some might have fared worse and stayed worse in terms of social anxiety, depression, and quality of life. The reason as to why different self-report measures would yield different deterioration rates in the same clinical trial remains unclear. Although it could be the case that deterioration as well as

improvement occur in several and separate domains, not only in relation to a given psychiatric disorder (Strupp & Hadley, 1977), such as, having less anxiety after treatment but still experiencing lowered quality of life due to unmet treatment expectations. As for non-response, the incidence was quite high at mid assessment, 69.9-86.5% for SAD, and 80.5-97% for the other outcome measures. However, this was not stable, with numbers dropping to 32.3-49.6% and 57.1-89.5% at the post treatment assessment, as well as 29.3-50.4% and 66.9-86.5% at follow-up, respectively. This finding is in line with what is expected in CBT delivered face-to-face, which suggests that about half the patients do not respond (Hofmann et al., 2012). Interestingly, non-response was much higher and remained so for the other outcome measures compared to the self-report measures of SAD. Although this might be because the treatment was primarily intended for SAD, wherefore it did not have a similar influence on other domains during the assessment period.

In the second case, deterioration rates were explored among a much larger set of patients undergoing ICBT, using the individual scores from a vast quantity of clinical trials. Given that deterioration occurs on proportionally few occasions and with large variation between samples, establishing more reliable estimates of the number of patients that deteriorate is necessary, which is why an IPDM was implemented in Study III. The results revealed that 122 (5.8%) deteriorated in treatment, as compared to 130 (17.4%) in the control conditions, the latter mainly consisting of wait-list control. Again, these numbers closely resemble the deterioration rate of 5-10% that is often referred to in psychological treatments delivered face-to-face, mirroring the proposition originally made by Bergin (1966) almost half a century ago. Likewise, the findings are similar to those by Hatfield, McCullough, Frantz, and Krieger (2010) and Mechler and Holmqvist (2016) for patients in routine outpatient care, providing a comparison between the two means of delivering treatment interventions for psychiatric disorders. More important, however, is the separation of patients receiving treatment and being assigned a control condition through randomization, as few attempts have previously been made to investigate which of these are linked to greater deterioration rates. Dimidjian and Hollon (2010) describe this as fundamental for understanding negative effects of psychological treatments, as it allows a closer inspection of what might be responsible for patients faring worse. In this perspective, the results were somewhat unexpected, as a much larger proportion of the patients deteriorated while waiting to be allocated to treatment than actually receiving it. This runs contrary to many previous findings of patients who improve during wait-list control, with a meta-analysis by Posternak and Miller (2001) on MDD suggesting that up to one-fifth of all patients could be experiencing a spontaneous remission without receiving any treatment. However, it might be the case that examining the average change between different time points misses the fact that, while some of the patients improve, certain patients actually fare worse, as the variability of the individual scores

may go unnoticed. This was already mentioned by Bergin (1963) in relation to investigating the efficacy of psychological treatments, suggesting that, apart from those who improve and do not respond, there exist a number of patients who deteriorate. Albeit less explored than improvement, there have been a few studies examining the deterioration rates among patients waiting for treatment. For instance, Young (2006) noticed that 3.5% fared worse on during an average waiting period of one month prior to entering treatment. Jacobson, Follette, Revenstorf, et al. (1984) also found that 17.3% of the couples awaiting help for marriage distress deteriorated, in both cases in face-to-face settings. Hence, it could be the case that deterioration occurs not only for patients that are undergoing treatment, but also for a significant proportion being allocated to wait-list control. If true, this runs contrary to the popular conception of this being a harmless comparator in clinical trials. It also raises ethical concerns with regard to the use of wait-list control in clinical trials. According to the Declaration of Helsinki, which is used as a guideline in conducting medical and psychosocial research (World Medical Association, 2013): “Measures to minimise the risks must be implemented” and “Medical research involving human subjects may only be conducted if the importance of the objective outweighs the risks and burdens to the research subjects” (p. 2192). On the one hand, given the evidence for a higher deterioration rate among patients waiting for treatment, these principles could make the usage of such a comparator unethical. On the other hand, there are a number of methodological issues in favor of wait-list control, e.g., statistical power, logistical advantages, and to establish the efficacy of a particular treatment compared to not receiving anything at all. Furthermore, it is not clear if other comparators are less harmful, e.g., treatment as usual, counseling, or, a non-active alternative, such as, discussion forum. Also, it is plausible that some patients would be experiencing deterioration regardless of waiting for treatment or not, and as they are in the process of eventually receiving help, wait-list control could be considered an ethical alternative. Nonetheless, given the finding that worsening occur to a large extent during the waiting period, it should be imperative to also monitor the well-being of patients waiting for treatment in order to prevent deterioration and possibly even dropout. It might also become increasingly important in the future to be more restrictive in terms of choosing a comparator, that is, selecting wait-list control only in those cases where it is absolutely necessary, rather than using it as a standard approach. In addition, greater emphasis of both the benefits and risks of treatment should be made during the informed consent process so that patients signing up for a clinical trial have the information to make an educated decision on whether to participate. However, informing patients about harm and monitoring deterioration during the waiting period may prove vital not only in a research setting, but should perhaps also be a part of standard practice among patients in line for treatment in routine care.

Predicting deterioration

Determining the deterioration rates in psychological treatments provides a reasonable estimate of the proportion of patients that are likely to fare worse. However, apart from informing researchers and clinicians of the occurrence of negative effects, these numbers are unlikely to help them foresee what patients are more prone to deteriorate. In order to understand if there are any specific characteristics associated with deterioration, an investigation of predictors is required, involving a closer look at what may put someone at risk of worsening during treatment. Similar to exploring the contingencies for a positive response (Delgado, Moreea, & Lutz, 2016), there could also exist aspects related to an undesirable outcome. Barlow (2010) discussed this with regard to psychological treatments delivered face-to-face, advising the use of both idiographic and nomothetic approaches to get a better picture of the advantages and disadvantages of treatment. Rozentel et al. (2014) also presented several plausible variables that could be responsible for negative effects in treatment interventions via the Internet, e.g., deficiencies in the implementation of treatment, feelings of disappointment and demoralization, and problems with the therapeutic relationship, recommending that more research should be done to study their importance. Still, despite an increased awareness of the fact that some patients can in fact fare worse, few attempts have been made to examine what might be contributing to its occurrence. Hence, the present thesis conducted the first ever assessment of predictors of deterioration in ICBT, as performed in relation to Study I and III.

In the first instance, a set of variables deemed related to deterioration were investigated; clinical severity at pre treatment assessment according to the LSAS-SR, civil status, prior psychological treatment, age, and gender, indicating that none were linked to faring worse at either mid assessment or follow-up. However, at post treatment assessment, patients exhibiting higher symptomatology of SAD had increased odds of deteriorating, OR 1.07, 95% CI [1.02, 1.13], suggesting that those exhibiting more symptoms prior to commencing treatment are more likely to deteriorate. Albeit not related to deterioration per se, H. Christensen, Griffiths, and Farrer (2009) did find similar results with regard to dropout in ICBT, which might suggest that clinical severity predicts worsening, and, in turn, the risk of dropping out and ending treatment prematurely. This seems reasonable given that patients with more severe psychiatric disorders may perhaps require more intensive care than is possible to obtain via the Internet, at least in terms of identifying those that do not respond and adapting their treatment. Because ICBT is often delivered according to a predetermined time schedule as well as fixed order of content, particularly in clinical trials, flexibility and responsiveness is limited, sometimes creating problems with adherence and implementation (Bendelin et al., 2011), and, potentially, making it difficult to effectively prevent deterioration in time. Furthermore, it is not unlikely that higher

symptomatology is associated with greater comorbidity, as is often the case when it comes to psychiatric disorders (Kessler et al., 2005). This could be another explanation as to why some do not benefit from and even deteriorate during treatment, again, potentially being more challenging to detect and manage via the Internet, as compared to face-to-face.

In the second instance, predictors of deterioration were explored among a much larger and more heterogeneous sample, including patients with several different psychiatric disorders or conditions that require clinical attention. This was conducted because of the relatively small number of patients that deteriorate during a single clinical trial, which most likely leaves the study of predictors underpowered to detect any meaningful differences. Edwards et al. (1978) pointed out that “Very large samples of patients would have to be used to develop a large enough group for reliable determination of predictors of deterioration” (p. 286), making an IPDM highly suitable for examining what might be associated with faring worse. A set of variables selected a priori were thereby examined; clinical severity at pre treatment assessment based on the primary outcome measure in each clinical trial, civil status, prior psychological treatment, prior or ongoing psychotropic medication, sick leave, educational level, and age, with different results depending on whether the patients were receiving treatment or being in a control condition. First, waiting for treatment, which was most often utilized as a comparator, was related to increased odds of deterioration, OR 3.10, 95% CI [2.21, 4.34], in line with the proportionally larger deterioration rate for these patients, and, again, stressing the importance of monitoring their well-being. Second, having more symptoms prior to treatment was, somewhat surprisingly, linked to lower odds, OR 0.62, 95% CI [0.50, 0.77], and OR 0.51, 95% CI [0.51, 0.80], respectively. This runs contrary to the findings in Study I, which instead suggested that clinical severity at pre treatment assessment was associated with increased odds. However, one possible explanation is that regression to the mean might exert an influence, with patients having scores that are considerably above the mean for the rest of the sample tend to demonstrate a much greater symptom reduction (Hsu, 1995). In other words, it could be the case that higher symptomatology is not necessarily related to faring worse in ICBT, with Bower et al. (2013) obtaining similar results in relation to low intensity treatment interventions for MDD, indicating that more severely depressed patients seem to benefit as much from treatment as those that are less depressed. Of course, this does not preclude individual cases from deteriorating, but the relationship may not be clear enough to predict what patients are at risk. Alternatively, there could be a distinction between psychiatric disorders, with some being more related to deterioration than others if the clinical severity is greater at pre treatment assessment. However, a post hoc analysis of the three classes that were used in Study III; anxiety disorders, MDD, and other, did not reveal a difference, implying that this is not supported, although it could still be assumed for conditions that

were not included, e.g., bipolar disorder, schizophrenia, and substance abuse. Third, for patients in treatment, being in a relationship, OR 0.58, 95% CI [0.35, 0.95], having a university degree, OR 0.54, 95% CI [0.33, 0.88], and being older, OR 0.78, 95% CI [0.62, 0.98], were all associated with lower odds of deterioration. Karyotaki et al. (2015) did in fact find a comparable relationship with dropout in ICBT, suggesting that patients of male gender, lower educational level, and younger age were more likely to end their treatment prematurely. It is plausible that these variables have an association with both deterioration and dropping out, which might be explained by their influence on resilience and social support (Davydov, Stewart, Ritchie, & Chaudieu, 2010). Of particular interest is, however, the connection between education and deterioration, demonstrating that patients without a university degree may be more inclined to fare worse. Given that ICBT can be very demanding in terms of the amount of time and effort that need to be put into reading the texts and completing the exercises, it is possible that patients with a lower educational level struggle with understanding the content and implementing the treatment interventions, thus making it hard to keep up with the pace of treatment. This particular issue has been found to have an impact on adherence in ICBT (Waller & Gilbody, 2009), and a review of the more frequently used English self-help manuals for MDD indicate that they often require a reading age level of 12.6-15.4 years, while a large proportion of the patients using them are actually below 11 (Martinez, Whitfield, Dafters, & Williams, 2007). It is therefore reasonable to assume that many of the clinical trials included in Study I and III also involved texts that are quite complicated to comprehend. In a face-to-face setting this may not result in any problems as questions and queries can be resolved immediately during the session. However, in ICBT, where the feedback from a clinician often is limited and the circumstances do not always permit any further clarification, this may be more difficult. Research gives some support to this assumption, suggesting that many patients would have preferred additional guidance and support from a clinician while undergoing treatment via the Internet (Halmetoja, Malmquist, Carlbring, & Andersson, 2014). This implies that more assistance might be necessary for those that experience difficulties when it comes to grasping the content or performing the exercises, such as, conducting regular checkups together with the patient in order to overcome potential barriers for implementation and sort out any misunderstandings (Bengtsson, Nordin, & Carlbring, 2015). Increasing patient involvement and providing more support by a clinician has in fact been shown to improve the results of treatment, for example, in one case of bibliotherapy for insomnia (Kaldo, Ramnerö, & Jernelöv, 2015), giving some credence to this notion.

Adverse and unwanted events

Deterioration has never been the only way of defining negative effects of psychological treatments. In fact, researchers and clinicians were early to discuss different types of adverse and unwanted events that also could occur during treatment, but whose incidence do not always result in the worsening of a given psychiatric disorder or increase in symptomatology. Hadley and Strupp (1976) presented a number of issues that might be seen as negative by the patient, clinician, or society, based on a survey on the topic, ranging from interpersonal difficulties and new symptoms, to stigma as well as lowered self-esteem. Lately, non-response, dropout, and dependency, in other words, becoming overly reliant on the clinician or treatment, have also been put forward as detrimental (Linden, 2013), suggesting that negative effects can take many different forms depending on how they are conceptualized and what perspective is being used. As such, adverse and unwanted events are more complicated to study, which is probably why less has been done on the topic in comparison to deterioration. However, self-reports by the patients, either through open-ended questions, interviews, or quantitative means, could be used as a way of determining their occurrence and characteristics, as has previously been performed in relation to psychological treatments delivered face-to-face (Ladwig et al., 2014; Parker et al., 2013; Sachs, 1983). Thus, the present thesis conducted the first ever investigation of adverse and unwanted events in ICBT, using a variety of different methods in order to explore and understand the nature and implications of these negative effects; an unstructured format that allowed the patients to describe their experiences in their own words in Study I and II, and a more systematic approach based on a novel self-report measure that was developed with the specific aim of monitoring adverse and unwanted events, as reviewed in Study IV.

In the first case, open-ended questions were posed in both studies, that is, probing for situations or circumstances that were perceived as negative by the patients themselves. This indicated that 4.5-12.9% experienced events that were regarded as adverse and unwanted in Study I, depending on the time point being assessed, compared to 9.3% in Study II, range 4.9-12.6%, varying between the four clinical trials that were a part of the investigation. In each instance, patients were asked to rate how negatively they perceived their incidence, suggesting a moderate influence, $M = 1.59^8$ ($SD = 1.12$) and $M = 1.07-1.80$ ($SD = 1.00-1.05$), respectively, also shifting in relation to the time point. Hence, adverse and unwanted events were indeed experienced during treatment, as described by the patients, but these seemed to have only a modest impact. Study I also used an adapted version of the UE-ATR as a way of classifying the negative effects that had been reported; emergence of

⁸ Scored on a Likert-scale that ranged 0-3 (0 indicating no effect at all, 3 representing a severe effect).

new symptoms, deterioration of symptoms, negative well-being, lack of clear treatment results, non-compliance, as well as changes in work situation, stigmatization, and other. These were, in turn, assessed by two independent raters and considered to be quite closely related to treatment, $M = 4.0-4.38^9$ ($SD = 1.07-0.85$), with slightly higher ratings at post treatment assessment, and suggesting a possible concurrence between patients and researchers with regard to which of these are attributable to treatment. In terms of Study II, the adverse and unwanted events were explored using qualitative content analysis in order to detect recurrent patterns in the responses. This revealed two distinct categories that separated the negative effects into patient-related and treatment-related, i.e., those being associated with gaining insight and experiencing new symptoms, and those in connection to difficulties with implementation and the format through which the treatment interventions were delivered. Albeit demonstrating certain disparities, if analyzed together, the results from Study I and II do indicate some overlap. In both instances, worsening and concerns seemingly unrelated to the initial condition of the patient appeared, most notably, stress, anxiety, and insomnia, implying that these might represent quite frequently occurring negative effects. In addition, negative well-being also seem related to what was defined as symptoms in Study II, while insight may constitute a distinct issue, perhaps characterized by thoughts or emotions associated with more existential matters and grief. This is rarely mentioned in the research, but warrants further investigation as it is likely for some patients to experience distress when confronting lifelong problems and realizing they could have been helped long ago. Furthermore, the lack of clear treatment results and non-compliance, as found in Study I, could be related to difficulties performing the treatment interventions or adjusting to the format - aspects that can probably leave the patient feeling inadequate and significantly affect motivation. As for the other negative effects that were reported in Study I, these adverse and unwanted events might have a closer association with such issues as occupational problems, shame, and interpersonal difficulties, rather than the treatment interventions per se. Albeit not included as a separate or distinct category or subcategory in Study II, examples of such incidents affecting the patient negatively were indeed found in a number of the individual responses, implying that other circumstances may sometimes exert an influence, possibly making it harder to differentiate what is driving negative effects in psychological treatments, as discussed by Mays and Franks (1980).

In the second case, a novel self-report measure intended to facilitate the monitoring of adverse and unwanted events was developed and evaluated. This extended the findings from Study I and II by using the results together with a consensus statement among researchers in the field of ICBT (Rozenal

⁹ Scored on a Likert-scale that ranged 0-5 (0 indicating no relationship with treatment, 5 being related).

et al., 2014) and a literature review in the generation of items. In comparison to the two previous studies, which applied qualitative methods, Study IV allowed a more systematic and quantitative approach of exploring negative effects that might prove useful in order to determine their relationship with treatment outcome. Of the original 60 items, 32 were retained, suggesting a six-factor solution that accounted for 57.64% of the variance; symptoms, quality, dependency, stigma, hopelessness, and failure. Of these, the first explained as much as 36.58%, a substantial proportion when compared to all other factors, totaling only 21.06%. Again, this suggests that deterioration and new symptoms may represent commonly experienced negative effects, perhaps contributing with the largest part of their negative impact. However, additional adverse and unwanted events did occur, and with regard to the other factors that emerged, the quality of the treatment could be particularly important to consider, despite explaining just 8.71%. This is sensible since it is essential for psychological treatments to instill hope and provide a clear and coherent rationale that can help instigate change. Hence, if the content or delivery is perceived as deficient or substandard, for instance, by introducing too many errors or inconsistencies or demonstrating technical breakdowns, this could potentially impose confusion and frustration that might result in hopelessness and worsening. Moreover, as proposed in relation to Study III, difficulties following through with treatment may also be responsible, as can be expected if there is a large discrepancy between educational level and the requirements to execute the treatment interventions. This underlines anew the importance of providing treatments that are accessible to all patients regardless of their socioeconomic status, cultural and ethnic background, or physical and cognitive disabilities. In addition, it might be the case that the actual provision of treatment can violate the expectations of some patients, that is, opposing a preconceived idea of how it should be supplied and what is needed in order for them to improve. This should be investigated further, as there may be a mismatch between what some patients want and what a clinician or treatment provider has to offer, especially in terms of the format. There is some support for this notion, with higher treatment credibility being associated with a better treatment outcome (El Alaoui, Ljótsson, et al., 2015), possibly reflecting how the patient sees the treatment interventions. Moreover, given the importance of goal setting in psychological treatments (Ramnerö & Törneke, 2015), this could also represent an area that warrants additional investigation, for instance, to what degree the goals of treatment are clear to the patient. Furthermore, concerns or ruptures in the therapeutic relationship is another possible explanation for the quality of the treatment being regarded as poor by some patients. With regard to the importance of a trusting and supportive working alliance for treatment outcome (Martin, Garske, & Davis, 2000), it is reasonable to assume that it also can have a negative impact if deficient, as has been found in face-to-face settings (Saxon & Barkham, 2012), and, in some cases, also for ICBT (Nordgren,

Carlbring, Linna, & Andersson, 2013). In sum, the reasons for insufficient quality in psychological treatments might have different explanations and should be explored more in the future. This is deemed particularly important because it would be fairly easy to influence and improve, as opposed to both deterioration and new symptoms, thereby making it possible to prevent some negative effects from occurring. As for the other factors that were retained, they contributed with notably less variance, range 2.11-4.13%. However, they were still regarded worthy of retention based on both statistical and theoretical grounds. It could, for example, be the case that they occur less regularly in treatment, yet still have negative implications for those afflicted, as would be expected of hopelessness and failure (Mohr, 1995). It might also vary depending on the treatment, psychiatric disorder, and context in focus, such as, dependency potentially being more prevalent in some types of less structured or long-term psychological treatments. Certain characteristics of the patient may also mediate what negative effects appear. It is plausible that those with interpersonal difficulties experience less adverse and unwanted events in treatment interventions delivered via the Internet because of the absence of a face-to-face contact, although, this is something that requires further investigation. Interestingly, all of the items that were specifically related to ICBT were removed, as these did not correlate sufficiently enough with the rest of the self-report measure, e.g., “I wasn’t satisfied by the user interface in which the treatment was being delivered” (Item 58). However, this could be due to the fact that many patients in Study IV were recruited from the media and may not have had any experience of undergoing ICBT. In addition, the items that were removed might have been associated with the ease of use, enjoyment, and layout of treatment, rather than negative effects per se, making other self-report measures more useful in order to determine problems related to the graphical or technical issues that may arise in ICBT, e.g., the Internet Evaluation and Utility Questionnaire (Ritterband et al., 2008) and the System Usability Scale (Brooke, 1996).

Preventing negative effects

The findings from the present thesis have two major implications for the implementation of ICBT in both clinical trials and routine outpatient care; first, deterioration occurs to such a degree that its monitoring and reporting should be considered mandatory, particularly for those waiting for treatment, given the proportionally higher number faring worse among these patients. In addition, there seem to exist various variables that are associated with increased or decreased odds of worsening, which might be of importance when considering for whom this type of treatment is appropriate and if there are ways of preventing deterioration from occurring. Second, adverse and unwanted events are also to be expected, ranging from the experience of new

symptoms and lack of quality, to dependency, stigma, hopelessness, and failure, possibly warranting other means for detecting their incidence during treatment and the development of strategies for their prevention. Based on these results, a number of suggestions on how to improve the delivery of psychological treatments via the Internet are provided, primarily from the perspective of finding methods for averting negative effects and reversing a negative treatment trend. The recommendations that follow are based on the identification and management of patients at risk of deteriorating, as seen in Study I and III, and the probability of experiencing other events that might be perceived as adverse or unwanted, as found in Study I, II, and IV.

In the first instance, a review of the research on continuous symptom monitoring is at place, as there already exist evidence on how to detect and manage deterioration in psychological treatments by using statistical and actuarial methods (Boswell, Kraus, Miller, & Lambert, 2015). In principle, the idea behind such an approach is that the condition of a patient should be monitored continuously throughout treatment via a self-report measure that is being distributed weekly. The scores are then compared against a vast amount of data, thereby allowing a more reliable assessment of whether or not they deviate from an expected trajectory. Hence, should the particular patient diverge from what is projected, that is, based on a certain degree of symptomatology at pre treatment assessment and the anticipated slope at a specific time point, the researcher or clinician is notified that the patient is not on track (Lambert, 2007). In this way, non-response and deterioration can be identified in time, making it possible to determine the reason for such an event and reverse a negative treatment trend. Furthermore, this could also be accompanied by recommendations of what treatment interventions are appropriate to apply, e.g., using motivational interviewing with the patient in order to overcome potential barriers for adherence. At present, this approach has been examined in nine naturalistic studies including over 4000 patients, proving that it can have a considerable impact on the number of patients that are either not responding or faring worse. The findings indicate that those who were followed using a weekly self-report measure deteriorated at a much lower rate than patients who received no between-session monitoring, i.e., treatment-as-usual, 15.2% compared to 23.2% (Lambert et al., 2002). Providing problem-solving to the clinician, i.e., advising what treatment interventions that could improve the condition of the patient, decreased these numbers even further, 8.5% in comparison to 13.6% for those only being monitored continuously, and 19.1% in treatment-as-usual (Whipple et al., 2003). Similar tendencies were found for non-response, albeit less striking than for worsening; in the first case, 50% compared to 58.9%, in the latter, 42.4%, 53.4%, and 55.7%, respectively. Interestingly, this approach does not appear to affect improvement rates significantly, instead it seems like the greatest benefits of continuous symptom monitoring are made by patients who are at risk of non-response or faring worse. Likewise, the number of

sessions required in order to achieve remission are also unaffected, at least for those on track, but increased by 1.5-4.7 for patients that are identified as non-responders or deteriorated, after which most are able to recover (Lambert, 2013). Hence, introducing a similar system to psychological treatments in general would primarily be of interest for detecting and helping those that are not improving as predicted, rather than for boosting effects or cut costs. From an ethical standpoint, however, it would be indefensible not to implement a procedure that has already been shown to minimize risks for the patient. From a financial perspective it should also be a greater return on investment to have more patients improve in the long-run despite an increase in expenses in the short-term. Moreover, such a system could also overcome some of the problems related to recent findings indicating that clinicians are often poor at recognizing if patients have worsened (Hatfield et al., 2010), that few have received information on negative effects during their basic clinical training (Bystedt, Rozental, Andersson, Boettcher, & Carlbring, 2014), and that the concept of negative effects is perceived as vague and hard to grasp (Jonsson, Johanson, Nilsson, & Lindblad, 2015). The use of continuous symptom monitoring would of course have to rely on comparable statistical and actuarial methods as the ones developed by Lambert et al. (2001), which were originally based on the administration of the OQ-45. However, given enough data, trajectories for improvement, non-response, and deterioration should be possible to distinguish for other self-report measures as well (Rozental et al., 2014), for example, the LSAS-SR in terms of SAD. Even without such information, establishing endpoints for reliable deterioration might also be possible by only using the RCI, or, alternatively, some other criterion for worsening, e.g., predetermined cutoffs for mild, moderate, and severe depression (Seggar, Lambert, & Hansen, 2002). Thus, implementing ways of continuously monitoring the condition of the patient is already available and should not be especially complex or expensive to introduce in either clinical trials or routine outpatient care, particularly with regard to psychological treatments delivered via the Internet. Given the use of computer-based self-report measures already distributed directly online instead of paper-and-pencil, the administration of such a system would be automatic, highly responsive, and give immediate feedback to the clinician or researcher on the progress of a particular patient. Depending on the course of treatment, it could then proceed as before, be changed according to what may be suitable, or, in the more severe cases, ended, so that the patient is referred to an appropriate alternative, for example, a face-to-face contact. This concept fits well with a stepped-cared model in which those who are on track can complete the treatment interventions at a predetermined pace and be dismissed upon remission, while patients who do not respond are given successively more intensive options (Nordgreen et al., 2016). Moreover, apart from identifying those that are not improving and providing a better procedure for dealing with non-response and deterioration, it might also

introduce greater flexibility in terms of duration, guidance, as well as what treatment interventions are needed in order to recover. For instance, it may be the case that patients with a certain psychiatric disorder would benefit from the addition of content or exercises that were not originally included, e.g., managing insomnia in MDD, practicing self-assertiveness in SAD, or transdiagnostic issues, e.g., overcoming perfectionism and procrastination. Flexible, or, tailored ICBT, has been examined in relation to depression and its comorbidities, indicating greater effectiveness for patients with higher symptomatology at pre treatment assessment, as compared to a similar but untailored treatment (Johansson et al., 2012). Berger, Boettcher, and Caspar (2014) also demonstrated that individualized ICBT was more effective in treating comorbidities among patients with anxiety disorders when compared to standard procedures. Likewise, Nordgren et al. (2014) found support for tailored ICBT for patients with anxiety disorders in a primary care setting, with benefits maintained at the one-year follow-up and proving to be quite cost-effective, albeit without a comparison to an untailored alternative. However, whether or not this type of more adaptable treatment interventions would yield lower rates of either non-response or deterioration than less flexible options remains to be seen and needs to be investigated further.

In the second instance, few investigations have been made with regard to the monitoring and prevention of adverse and unwanted events, making a review of prior research complicated. Although, a few notable exceptions exist in terms of exploring what might be experienced as negative by the patient during treatment and developing reliable ways for its assessment. Suh et al. (1986) were probably among the first to address this issue, proposing that the VNIS could be used as a method for identifying deficiencies and problems in psychological treatments. Since then, the ETQ, the UE-ATR, and the INEP have been proposed, each with different items, definitions, and perspectives of what may constitute negative effects and how these should be distinguished. Regardless of theoretical or practical standpoint, however, neither the clinician-rated instrument nor the self-report measures have been used to a great extent, and seldom together with other outcome measures, making the relationship between adverse and unwanted events and treatment outcome rather unclear. Hence, it is still unknown if these negative effects can have an impact that is detrimental, and, in turn, whether these have any long-term consequences. Nevertheless, the evidence indicates that some patients experience certain incidences during treatment as negative, with an estimation from the present thesis showing that the rates lie between 4.5-12.9%, i.e., when using open-ended questions to probe for their occurrence. The responses by patients having received help for mental distress in the United Kingdom give some credence to these numbers, with reports of 5.2% stating that they had lasting negative effects from undergoing treatment (Crawford et al., 2016). Similarly, El Alaoui, Hedman, Ljótsson, and Lindfors (2015) found that 7.8% experienced adverse and unwanted events

in ICBT for SAD, with 3.2% reporting that it also had a very acute effect¹⁰. Distributing some form of self-report measure thus seems like an appropriate method for monitoring these events, for instance, using the NEQ. This would facilitate the assessment of negative effects that are not necessarily related to non-response and deterioration, making it possible to determine if these are imperative to address. For instance, insufficient quality could be a sign of problems understanding or implementing certain treatment interventions, which would then be discussed together with the patient in order to adapt the level of difficulty, clarify a subject matter, or propose alternative exercises that are easier to complete. This could prove to be particularly important with regard to patients who refrain themselves from asking questions and giving feedback, perhaps due to interpersonal reasons, or who might hold expectations of treatment that are not matched by its delivery. In addition, probing for such incidents may be especially relevant in ICBT, where it is not always feasible to determine how things are apprehended by the patient. As for new symptoms, investigating their occurrence could prove important in understanding how they are associated with the execution of certain treatment interventions, such as, increased anxiety during gradual exposure, and should perhaps be brought up for discussion during treatment to make sure it does not affect adherence negatively. However, distributing such a self-report measure warrants methodological considerations, particularly in terms of when it should be filled out by the patient. Adverse and unwanted events are often explored at post treatment assessment or retrospectively, thereby making it impossible to, in hindsight, detect and manage negative effects as they occur. Thus, in order to be effective as a mean for preventing negative effects from having an impact, it has to be administered regularly during treatment, perhaps after a few weeks or sessions, at mid assessment, and prior to its termination. This would also enable a closer inspection of how adverse and unwanted events at different time points are related to treatment outcome, as it has been hypothesized that negative experiences linked to the implementation of certain treatment interventions should not be enduring for most patients (Castonguay et al., 2010). Furthermore, the use of such a self-report measure at several time points other than post treatment assessment may potentially overcome some of the problems associated with recall bias, primacy and recency effects, and social desirability. For instance, incidents that occurred at the beginning of treatment or quite close to its completion might be easier for the patient to recognize, regardless of their long-term consequences, while others are easily forgotten (Schwarz, 1999). Similarly, if negative effects are only evaluated in retrospect, the reporting of their occurrence will likely be contaminated by the benefits of treatment, or,

¹⁰ Refers to the same open-ended questions used in Study I and Study II in the present thesis, where “very acute effect” corresponds to 3 on a Likert-scale that ranged 0-3 (0 indicating no effect at all, 3 representing a severe effect).

lack thereof, as experienced by the patient, which could either increase or decrease the ability to provide valid responses (Krosnick, 1999). In addition, assessing negative incidences more consistently averts some of the problems related to missing values at post treatment assessment, that is, whether or not such an event is due to experiences that were adverse or unwanted, as it could be used for exploring their connection to dropout. Considering the frequent use of computer-based self-report measures in ICBT, probing for negative effects in this way should be fairly easy to achieve and constitutes an important first step in preventing their occurrence in treatment.

In both instances, the prevention of negative effects would probably also profit from providing clear statements declaring both the advantages and risks of undergoing treatment. In comparison to many medical treatments, such as, surgical procedures and the use of psychotropic medication, it is common practice to monitor adverse and unwanted events and to notify about both the ups and downs of their implementation (Wysowski & Swartz, 2005). Thereby the patient can make an educated decision of its use that also takes the probability of experiencing negative effects into account. However, this has typically not been the case in terms of psychological treatments (Walker, Logan, Clark, & Leukefeld, 2005), perhaps because the idea of inadvertently causing harm is still largely unknown. In fact, a survey among clinicians made in the United States indicates that only a small number have considered this possibility, and that few discuss the limits and dangers of undergoing treatment while gathering informed consent from the patient (Sarkozy, 2010). Interestingly, most of the respondents in this study believed that addressing negative effects early on is important and should not interfere with the therapeutic relationship or the potential of benefitting from the treatment interventions. This is in line with what has previously been shown in the research (Handelsman, 1990), indicating that there is no evidence of affecting treatment expectations by discussing risks together with the patient. Furthermore, the respondents often brought up ethical issues as a reason for providing this information to the patient, with both codes of conduct and professional guidelines being referenced. Still, mentioning risks were not particularly common, perhaps because a clear and distinct outline on how to collect informed consent is lacking, which makes it difficult for a clinician to know what should be included in its procedure. Sarkozy (2010) also argues that another explanation might be the fact that most clinicians frequently underestimate the possibility that their patients may deteriorate or experience adverse and unwanted events. This was shown early on by Dawes, Faust, and Meehl (1989), indicating that statistical and actuarial methods are far better than clinical judgment for determining improvement and deterioration. In other words, relying only on impression and experience and believing that a particular patient in treatment is not at risk seem to impede the ability to accurately assess if negative effects have occurred. Thus, acknowledging the fact that clinicians, on average, are relatively poor at identifying patients that

have deteriorated, and offering information about the risks of undergoing psychological treatments, should prevent some of the adverse and unwanted events that can arise. This could, for example, be comprised of better and more standardized ways of managing informed consent in clinical trials and routine outpatient care, such as, providing a disclosure of both improvement and deterioration rates in a written format. In addition, some of the more frequently occurring adverse and unwanted events can be described so that patients are fully aware of the possibility of experiencing similar incidents. Thereby, negative effects, at least those that are expected as a consequence of undergoing treatment, can be explained and perhaps even circumvented. This seems particularly relevant with regard to ICBT, where the treatment interventions are mostly delivered in writing. Including paragraphs of how patients might come to perceive certain parts may help them understand why they react in a particular manner, for instance, illustrating that it is common to experience thoughts or emotions of existential character at the beginning of treatment.

Limitations and future research directions

The findings in the present thesis are based on research that utilizes both quantitative and qualitative methods. As such, the results and the ensuing recommendations are derived from several different sources. It varies from responses by patients to open-ended questions and outcome measures to assessments by independent raters. It also uses a combination of analyses, ranging from statistically determining non-response, deterioration and its possible predictors, to distinguishing underlying factors in a novel self-report measure as well as using qualitative content analysis to exploring patients' experiences of negative incidents. One of its greatest strengths is therefore related to the breadth and depth with regard to how negative effects have been assessed, perhaps resulting in an understanding of their occurrence and characteristics that is more comprehensive than only relying on numbers, or, alternatively, depending solely on the validity of one patient. In comparison to other attempts that have examined the issue, this might overcome some of the problems associated with the different definitions and investigations of negative effects that exist. Furthermore, it integrates several perspectives that do not always overlap, i.e., what is perceived as negative by the patient may not correspond with the opinion of the clinician. This was discussed by Strupp and Hadley (1977), proposing that improvement and deterioration need to be considered multifaceted, warranting a broad range of procedures in determining the benefits and risks of treatment. Similarly, Stuart (1970) discussed some of the problems with trusting only one type of measurement when evaluating treatment outcome, advising clinicians and researchers to incorporate many different viewpoints. Hence, the versatility of methods that

have been applied could perhaps shed some light on the complex nature of negative effects by going beyond the use of a single source of information. Another strength of the present thesis is the successive accumulation of knowledge that has helped shape and refine how non-response, deterioration, and adverse and unwanted events have been explored from study to study. Starting out with a consensus statement among researchers in the field of ICBT (Rozental et al., 2014), the different features and ways of determining negative effects were conceptualized prior to any empirical study. A first attempt was then conducted in relation to a specific clinical trial, after which more diverse and advanced efforts were performed in larger and much more heterogeneous samples in order to get a better picture of their characteristics. Lastly, the findings from each of these investigations were used to develop a novel self-report measure of adverse and unwanted events that might be useful in the future when trying to understand how negative effects are being perceived by the patients themselves. In this way, all studies are connected to each other, chronologically and methodologically, hopefully creating a line of argument that is seen as clear and coherent, and with results that will be regarded as valid and reliable.

Given the somewhat exploratory nature of examining negative effects of psychological treatments, there are, however, a number of limitations that need to be considered when reviewing the results and implications of the present thesis. First, the clinical significance of non-response, deterioration, and adverse and unwanted events still remains unclear. Undoubtedly, it is not uncommon for patients to remain unchanged, fare worse, and experience negative incidents during treatment, but whether or not these instances have a detrimental impact is unknown. Identifying non-response and deterioration by employing statistical procedures or predetermined cutoffs, e.g., the RCI, does not translate into the preservation or exacerbation of a given condition, nor does it have to be something that is even perceived as harmful by the patients themselves. This is particularly true for Study III, where a more lenient endpoint of 0.84 standard deviations was employed instead of the conventional 1.96. According to Wise (2004), this makes the assessment less reliable by decreasing the level of confidence from 95% to 80%, but it also enhances the prospect of identifying even less severe cases of deterioration. In theory, patients experiencing milder forms of worsening would thereby be possible to observe, which might be important from an ethical perspective. However, it is uncertain if this actually translates into a raw score change on the original self-report measure that is perceived as negative by the patient and if this is meaningful to report. For instance, it might be the case that the fluctuations lie within the natural variations that are expected for a given psychiatric disorder, that it only reflects a transient phenomenon related to the implementation of specific treatment interventions, or that outcome measures are not sensitive enough to accurately distinguish non-response and deterioration. One way of investigating this issue would be to conduct

in-depth interviews with patients identified as unchanged or deteriorated, which would allow an inspection of how these instances are apprehended. Another way would be to use additional time points for assessing worsening so that it would be possible to determine its long-term consequences, e.g., examining if deterioration between pre and post treatment is also maintained at follow-up. Likewise, exploring on which occasion patients deteriorate might also provide valuable insights of what is responsible for its incidence. Comparable to the concept of sudden gains in psychological treatments, i.e., a large decrease in symptomatology following a specific session, there may also exist examples of sudden worsening. Aderka, Nickerson, Bøe, and Hofmann (2012) found support for sudden gains in a meta-analysis of MDD and anxiety disorders, especially in CBT. However, the idea that there may also be episodes of swift deterioration has not been examined, offering an important and interesting future research direction in the investigation of negative effects. For instance, if self-report measures are distributed weekly, as in continuous symptom monitoring, tracking even minor fluctuations in a negative direction is possible, which could pinpoint treatment interventions that seem to precede worsening. Given enough data, this might help predict what should be used or refrained from with certain patients, thus preventing negative effects from occurring. In addition, with this type of monitoring, qualitative methods may also be used to assess the causes for deterioration, as perceived by the patient or clinician, similar to the implementation of ecological momentary assessment (Moskowitz & Young, 2006). Thereby it would be possible to investigate if there is anything related to the clinician's performance that is regarded as negative, or, alternatively, if the clinician apprehends any event that could be responsible for the patient faring worse.

Second, the distribution of reliable assessments is essential for identifying worsening among patients in an adequate manner, warranting a cautious selection of outcome measures. This could perhaps also be accompanied by the use of additional self-report measures in order to explore deterioration along more than one dimension, e.g., mood, self-esteem, and quality of life. For instance, it might be the case that a patient improves with regard to symptomatology of PD/A, but remains unchanged in terms of well-being. This was in fact explored in Study I, revealing some divergence between self-report measures, however, not so in Study III due to the large number of different types of assessments. Further research is thus needed to explore how deterioration rates are influenced by the outcome measures that are being used. Furthermore, determining the risks of psychological treatments is highly affected by the constraints of the self-report measure, i.e., the range of scores that are possible to attain (Wang, Zhang, McArdle, & Salthouse, 2008). This is of particular relevance for the assessment of deterioration given the fact that a patient cannot deteriorate indefinitely because of ceiling and floor effects. Hence, there is a risk that worsening might not be feasible to distinguish if the individual lies close to the maximum or minimum of a

self-report measure at the beginning of treatment. In clinical trials this might be less of a problem as the inclusion and exclusion criteria should preclude the most severe cases from entering, however, the potential for ceiling and floor effects during treatment still exist. As for Study III, this was examined more closely, revealing that only 55 (1.9%) patients were close to reaching an extreme, with a majority later improving rather than becoming worsened, suggesting that restriction of range might not be as challenging as it has been conceived in prior research. Nonetheless, further study need to be conducted on this issue in order to explore if this is true for other contexts.

Third, the definitions being used in the present thesis to describe various types of negative effects of psychological treatments stem from a consensus statement among researchers in the field of ICBT (Rozenal et al., 2014), together with a comprehensive literature review. Hence, the terms should correspond to what many researchers and clinicians would consider being detrimental or harmful. However, given the lack of a mutual terminology, different views of what constitute negative effects exist. As for deterioration the situation seems to be more or less clear, with the RCI often being used to define worsening, but in terms of non-response and adverse and unwanted events there are different opinions of their validity. Dimidjian and Hollon (2010) suggest that no treatment effect, i.e., non-response, could potentially be negative to the patient if the delivery of a given treatment has restricted the access to a more effective alternative: “If an inert treatment is inaccurately assumed to be beneficial, it still may be costly in terms of time, expense, and other resources.” (p. 24). According to this perspective, non-response is not perceived as negative per se and may be related to an ongoing condition. Furthermore, lack of results may not have to be detrimental if treatment has deterred deterioration, for instance, in more chronic cases where there is little room for improvement. However, if the treatment were erroneously seen as beneficial, while other actions would have been more forthcoming, this might have prolonged the suffering of the patient and should therefore be regarded as harmful, e.g., receiving counseling instead of CBT for PD/A (Baldwin et al., 2014). Linden (2013) made a similar argument by defining non-response as “Lack of improvement in spite of treatment. It is a UE [Unwanted Event]; it can be or cannot be an ATR [Adverse Treatment Reaction] or an MPR [Malpractice Reaction].” (p. 288). In other words, non-response is never anticipated or particularly welcomed, but the relationship with treatment remains unclear until a causal link between them or the clinician can be established. The same goes, however, for deterioration, and despite the fact that the present thesis has helped increased the current understanding of worsening, the problem with causality still remains, as first described by Rachman (1971). Similar to determining the passing or enduring nature of negative effects by in-depth interviews, one way of overcoming this issue is to invite patients to share their views of what contributed to the occurrence of non-response and deterioration to better understand their mechanisms.

With regard to adverse and unwanted events, the situation is even more complicated given the lack of an agreement as to what negative effects exist and how they should be monitored. There are currently a number of different self-report measures to assess such incidents, each with its own set of items and theoretical constructs, together with checklists for distinguishing harm more broadly. Hence, the definition of adverse and unwanted events will probably differ between investigations, which suggests that the findings of the present thesis may not correspond to the terms used in other studies. Great effort has, however, been made to outline the concepts so that it will resonate well with the literature and be as coherent as possible. This does not prevent some overlap or that definitions are being used differently, but it should ease interpretation and make it clear as to how they were defined.

Forth, exploring such a complex issue as negative effects of psychological treatments will ultimately impose restrictions in generalizability depending on what types of analyses have been performed, the samples being included, and the treatment interventions that have been delivered. Even if the present thesis can be regarded as comprehensive with regard to its methodology and the patients that were recruited, there are also limitations as to how well the results can be transferred to other populations and contexts. In particular, ICBT may or may not resemble CBT or other instances where psychological treatments are being administered face-to-face. Given that the treatment is mostly distributed as predefined modules with restricted correspondence from a clinician, the negative effects that were explored might have more to do with content and delivery than, for instance, the therapeutic relationship. Although this could be an advantage in terms of preventing some of the problems associated with therapist drift and inadequate treatment integrity, there is also a risk of not being able to generalize the findings to seeing a clinician in routine outpatient care. As discussed by Rozental et al. (2014), ICBT might inhibit adverse and unwanted events related to the relationship between certain patients and their clinicians, but for others the association could just as well be the opposite. For the most part, there seem to be more similarities than differences between the two formats, with the findings in the present thesis resembling the research of negative effects in general. However, without a head-to-head comparison and further investigation the issue remains unresolved, warranting some caution as to how the results and their following conclusions can be interpreted. Also, since a majority of the studies on negative effects have been made in relation to CBT, including the present thesis, it is somewhat unclear if there exist any variation between theoretical orientations. For instance, Berk and Parker (2009) discussed some of the issues that might be related to specific psychological treatments, e.g., greater risk of lowered self-esteem due to difficulties implementing treatment interventions in cases where patient responsibility is emphasized, as in CBT. However, the empirical evidence for such notions is lacking, making a more systematic approach to exploring negative effects across

theoretical orientations important, such as, comparing rates of non-response and deterioration, and distributing the NEQ in different clinical trials. Lastly, albeit the present thesis involved various psychiatric disorders, conditions, and issues that may warrant clinical attention, the nature of the clinical trials from which data was aggregated affects the generalizability of the results. Inclusion and exclusion criteria always restrict some patients from entering a given treatment, particularly those with greater symptom severity as well as certain comorbidities. From an ethical perspective this seems reasonable given that the delivery of treatment interventions via the Internet makes it harder to assess suicidal ideation and provide more immediate care, such as, allocation to inpatient care. Nonetheless, this also leads to limitations in terms of how well the results can be transferred to routine outpatient care and patients with more severe or multifaceted difficulties. This is especially relevant for the investigation of negative effects, as it could be argued that those with greater problems experience negative incidents to a larger extent. Similarly, the present thesis mainly consisted of patients with MDD and anxiety disorders, which limits the possibility of generalizing the findings to other populations, such as, patients with personality disorders, whom have been hypothesized to be at risk of harmful events, according to Mohr (1995). Thus, the results from the present thesis might not be representative of the negative effects that occur in other situations, making the study of such events in additional populations as well as naturalistic settings essential in the future. The same is true for patients with sociodemographics that differ from those in the clinical trials that were included in the present thesis. Across Study I-IV, a majority was in their late thirties and in a relationship, as well as having an employment and a university degree, which may not be the case for everyone in need of psychological treatment. However, this is comparable to treatment-seeking individuals in general (Vessey & Howard, 1993), implying that those with better resources and social support could simply be more inclined to request help for their ongoing difficulties. On the other hand, ICBT might attract patients with certain characteristics, as put forward by Arnberg, Linton, Hultcrantz, Heintz, and Jonsson (2014) in a systematic review: “The high level of educational attainment and employment rates among the participants raise concerns about whether the effects found in most RCTs [Randomized Controlled Trials] can be generalized to those who today are underserved.” (p. 11). Even compared to the general population in the age of 20-64 in Sweden, where 16% have some form of university education (< 3 years) and 23% have at least a bachelor’s degree (Statistics Sweden, 2014), the patients in the present thesis had, on average, a higher educational level¹¹. Hence, concluding that the findings can be transferred to populations other than the one included in the studies may not be a valid assumption, suggesting that more research needs to be done in

¹¹ Study III indicated that 64.3% had studied at a university or had a university degree.

order to investigate whether the occurrence and type of negative effects vary depending on sociodemographics. For instance, it would be interesting to explore if education has an effect on treatment outcome, and, in turn, if it has an impact on the rates of non-response and deterioration.

Fifth, although the present thesis utilizes both quantitative and qualitative methods for investigating negative effects, the studies still rely heavily on self-reports made by the patients. On the one hand this could be seen as a major strength, given that the perspective of the individual more accurately reflects what is actually perceived as negative during treatment. On the other hand it is unclear if a clinician or significant others would agree to all of the incidents that are reported or that the patient has in fact worsened. Also, some events may be seen as negative by the surrounding but not the patient, as in the case of becoming more self-assertive. Thus, it could be important to examine if and how the experience of the patient, clinician, and significant others differ from each other with regard to negative effects. For example, exploring in what way they recognize that change has occurred in terms of non-response, deterioration, and adverse and unwanted events, perhaps by using parallel assessments, could possibly help researchers and clinicians become even better at detecting and reversing a negative treatment trend. Additionally, developing and testing a novel self-report measure imposes several restrictions that need to be explored further in order to fully establish its validity and reliability. Given that the NEQ was assessed via exploratory factor analysis suggests that the results are preliminary and need to be replicated using additional samples and in different contexts. In particular, both the factors and specific items being retained are based on decisions made by a researcher and the data that was available, implying that some information that could be potentially relevant in certain settings may be lost or left out. For instance, highly unusual yet very negative incidents might also exist, but remain unknown or unexplored because they are not included in the self-report measure. This needs to be considered when distributing the NEQ to different patients, especially those having psychiatric disorders that were not a part of the exploratory factor analysis, e.g., bipolar disorder, schizophrenia, and substance abuse, as well as other treatment settings, i.e., inpatient care, child and adolescent psychiatry, and geriatrics. In addition, the relationship between reporting adverse and unwanted events and actual treatment outcome needs to be established, as it is currently unknown if the negative effects being endorsed are in fact detrimental. The same goes for the specific factors that were found and if they should be seen as separate or interrelated concepts. Since the first factor pertaining symptoms explained such a large proportion of the total variance, it is plausible that this might be most important with regard to determining adverse and unwanted events, while the others are subordinate. It is also possible that, while assessing the same underlying construct, i.e., negative effects of psychological treatments, some of the factors might be regarded as mediators, e.g., quality and stigma.

Coming research should investigate the factor solution that was obtained in Study IV to see if it is possible to generalize the results and if any of the factors need to be revised. Furthermore, because of the many complications associated with distributing a self-report measure, particularly on adverse and unwanted events, for instance, recall bias, primacy and recency effects, and social desirability, its administration has to be carefully considered in the future. This could be characterized by monitoring negative effects throughout treatment and not only in retrospect, as well as investigating if it is possible to develop a shorter yet equivalent version of the NEQ. Similarly, in order to avoid difficulties due to reactivity (Haynes & Horn, 1982), i.e., affecting the responses solely by using assessments or the wordings of the instructions or items, other ways of exploring the patients' experiences could become important to consider. For example, cognitive interviews, where an individual responds to the self-report measure verbally instead of in writing, would allow the researcher to pose follow-up questions and elaborate on certain themes. This was not implemented to any greater extent as part of the present thesis, which may have affected the findings as well as conclusions from Study I, II, and IV, but should be performed in the future.

Conclusion

The present thesis provides evidence for the occurrence of non-response, deterioration, and adverse and unwanted events during ICBT, some of which might have a detrimental impact on treatment outcome and the well-being of the patient. Therefore, researchers and clinicians are advised to monitor their incidence during treatment in order to prevent harm and to reverse a negative treatment trend. A number of recommendations on how to become better at detecting negative effects are proposed, including the use of continuous symptom monitoring and the distribution of a novel self-report measure for investigating adverse and unwanted events. As discussed by Lambert (2010), relying on clinical judgment alone or disregarding that negative effects exist is not only ill-advised and incorrect, it is also highly unethical. It is thus time to put pride and orthodoxy aside and recognize that it is very likely to have one or more patients not responding, deteriorating, or experiencing events that are adverse and unwanted in treatment. That way negative effects can be expected, predicted, and averted so that the patient is never at risk and the pledge of “first, do no harm” is truly fulfilled.

Acknowledgements

Per Carlbring, my main supervisor. Thanks for believing in me when I sent you an email almost four years ago with the idea of conducting the first ever randomized controlled trial on the treatment of procrastination. Thanks for encouraging me to apply for doctoral studies and thanks for letting me be a part of such an interesting and important research project as negative effects of psychological treatments. You have the gift of making every student and doctoral candidate feel welcomed, listened to, and brave enough to take a step into the unknown.

Gerhard Andersson, my co-supervisor. Thanks for being an inspiration and role model for so many students and doctoral candidates, including myself. Thanks for your feedback on my master's thesis, which read; "Excellent. Hope you'll go into research", making feel confident enough to one day pursue a scientific career. You always lift the spirit of people around you and believe that everyone is destined to achieve something great.

Johanna Boettcher, my co-supervisor. Thanks for sharing your expertise on the subject of negative effects of psychological treatments and for letting me continue do research on something you helped initiate. Thanks for opening the door to your home and for having me as your guest; your hospitality, warmth, and kindness has been of great support in my doctoral studies.

David Forsström, my close friend and colleague. Thanks for sticking with me during all these years; from being a novice to becoming a doctor and through setbacks and success. Thanks for making me laugh and take those well-deserved breaks, for stealing my candy, for all the beers and burgers, and for discussing all aspects of doing research. This is just the beginning.

All of the colleagues at the Division of Clinical Psychology at Stockholm University; Jonas Ramnerö, Peter Lilliengren, Jan Bergström, Lotta Reuterskiöld, Anna-Clara Hellstadius, Margareta Simonsson-Sarnecki, Ewa Mörtberg, Lars Klintwall, Stephan Hau, Markus Jansson-Fröjmark, Gunilla Berglund, Therese Anderbro, Tobias Lundgren, and Andrzej Werbart. Thanks for a great time together and for many years to come.

All of the doctoral candidates at the Division of Clinical Psychology at Stockholm University; Philip Lindner, Ekaterina Ivanova, Alexander Miloff, Rikard Sunnhed, Mårten Tyrberg, Jakob Jonsson, Lichen Ma, Camilla von Below, Lillian Döllinger, and Jonas Rafi. Thanks for many stimulating collaborations and interesting discussions together at the department.

Anders Kottorp, my co-author. Thanks for being the best teacher I have had and for being so supportive of working together.

Kristoffer Magnusson, my co-author. Thanks for your contribution, patience, and for explaining complicated things in a way that even I can understand.

Brad Schmidt, my co-author. Thanks for providing valuable input and for making me express what I want to say in a clear and coherent manner.

The consensus group, my co-authors; Gerhard Andersson, Johanna Boettcher, David Ebert, Pim Cuijpers, Christine Knaevelsrud, Brjánn Ljótsson, Viktor Kaldo, Nickolai Titov, and Per Carlbring. Thanks for joining me on what would be the starting point of a new area of research.

Jonas Ramnerö and Ulf Jonsson, my reviewers. Thanks for scrutinizing my doctoral thesis and giving me suggestions on how to improve it even further.

George Vlaescu, my technical support. Thanks for making every study run smoothly and for always being so kind and helpful.

Anna Malmquist and Karin Zetterqvist Nelson, my supervisors on my master's thesis. Thanks for teaching me the basics of conducting research and for introducing me to the wonders of qualitative methods.

Daniel Bruchfeld, my close friend and mentor. Thanks for taking me under your wings when I was just a "gymnast" and started working as a clinician and teacher. Thanks not only for helping me become more appreciative of different theories and perspectives on mental distress, but also for believing in me and making me laugh about all of the events we have been through.

Klaas Wijma, my clinical supervisor during my master's degree. Thanks for sharing your wisdom on the psychology of learning, functional analysis, and cognitive behavior therapy. Thanks for introducing the term procrastination, you truly helped shape my future career as both a clinician and scientist.

Helene Semb, my current clinical supervisor. Thanks for being supportive, encouraging and always on point in relation to the problems I encounter with

my patients. You have an incredible gift of getting to the core of the problem and making me feel a lot more confident when I leave one of our sessions.

All of the students I have supervised; Samuel Bystedt, Erik Forsell, Andreas Svensson, Niklas Laninge, Lina Mårtensson, Malin Skoglund, Linnéa Trosell, Alexander Örtenholm, Andreas Landström, Molly Zetterberg, Angelica Norström, Peter Alhashwa, Maria Palmer, Daniel Attar, Annakarin Söderberg, Julia Kiiskinen, Josefine Almquist Tangen, and Meri Mirsoyan. Thanks for having me as your supervisor and for teaching me a lot when it comes to conducting research.

All of my friends and fellow board members of the Swedish Association for Behaviour Therapy; Linda Jüris, Diana Fabian, Henrik Engelmark, Lena Fagerström, Amira Hentati, Erik Hjalmarsson, Monica Buhrman, Lise Bergman Nordgren, Li Wolf, Daniel Björkander, Sarah Vigerland, Sara Elfstrand, and many more over the years. Thanks for two exciting terms as part of the board and for all of the great work we have done together.

Andreas Schill, Andreas Iwarsson, and all of my colleagues at the clinic PBM. Thanks for having me as a clinician during my doctoral studies. Thanks for letting me be a part of your team and being so supportive of my ambition of combining both research and clinical work.

All of my friends and colleagues at Psykologifabriken; John Airaksinen, Hoa Ly, Linda Backman, Oskar Henrikson, Jenny Rickardson, Niklas Laninge, Katarina Blom, and Kajsa Asplund. Thanks for letting me be a part of such a creative and inspirational environment.

Lina Wennersten, my co-author. Thanks for all your work on our book and for teaching me the essentials in writing and getting my message through.

Ingrid Ericson, my publisher. Thanks for seeing the potential in our book and for being both a friend and one of my greatest fans.

Hana Jamali, my close friend. Thanks for being one of the most important people in my life and supporting me during the good times and the bad. Thanks for encouraging me to do things I never thought possible, for being such a good listener, and for making me think about what matters the most.

Rebecca Jungermann and Mikael Jungermann, my close friends. Thanks for being so supportive of my work and for all of our interesting discussions about psychology, people, and everyday things. Thanks for letting me stay at your place and for all of the lovely breakfasts we have had together.

Julie Meldahl and Sophie Karlsson, my close friends. Thanks for all of the support throughout the years and the great times we have had together.

All of the classmates during my master's degree. Thanks for making our years together unforgettable.

Gunvor Larsson Abbad and all of the teachers and staff at the Department of Behavioural Sciences and Learning at Linköping University. Thanks for providing a great foundation for so many future clinical psychologists.

Johanna Johansson, my close friend. Thanks for being one of my oldest and dearest companions. Thanks for always lifting my spirit, saying all of the things that I need to hear, and showing your support when I need it the most.

Adam Hjorthén, my close friend. Thanks for being an important part of my life and for sharing your ideas about research, the academia, politics, and all of the little things, such as, the difference between a cake and a pastry.

David Ekenbäck and Petter Säfwenbergl, my close friends. Thanks for reminding me that there is more to life than work and for making me laugh about all of my weird traits and features.

Andreas Nilson, Andreas Davidsson, and Philip Kling, my close friends. Thanks for making me open up, have fun, and enjoy life.

All my friends; Therese Ottavainen Dommang, Johannes Wallén, Linda Rothschild Wallén, Jesper Vinge, Janina Vinge, Ilona Eismann, Anna Gustavsson, and many more for being so warm-hearted and considerate.

All my family; Britt Rozental, Maks Rozental, Börge Österholm, Victor Rozental, Kristoffer Rozental, Feben Birhane, the Österholm-Croomes, the Schenholms, the Hagströms, the Pousettes, the Hallströms, the Ahlbäcks, the Lindfors, the Mellas, the Lindvalls, as well as everyone else around me who have given me so much support in my life.

Thelma, my dog. Thanks for all of our walks together during the completion of this thesis, giving me time to focus on something else for a while.

André Rozental, my husband. Thanks for allowing me to pursue my dreams and encouraging me to do the things I love the most. Thanks for letting me be the person that I am and always appreciating what I have to say. You are the most important person in my life, without you this thesis would not have been possible.

References

- Aderka, I. M., Nickerson, A., Bøe, H. J., & Hofmann, S. G. (2012). Sudden gains during psychological treatments of anxiety and depression: A meta-analysis. *Journal of Consulting and Clinical Psychology, 80*(1), 93-101.
- Alfonsson, S., Olsson, E., & Hursti, T. (2015). The effects of therapist support and treatment presentation on the clinical outcomes of an Internet based applied relaxation program. *Internet Interventions, 2*(3), 289-296.
- Almlöv, J., Carlbring, P., Kallqvist, K., Paxling, B., Cuijpers, P., & Andersson, G. (2011). Therapist effects in guided Internet-delivered CBT for anxiety disorders. *Behavioural and Cognitive Psychotherapy, 39*(3), 311-322.
- American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*(12), 1060-1073.
- Andersson, G. (2015). *The Internet and CBT: A clinical guide*. Boca Raton, FL: CRC Press.
- Andersson, G. (2016). Internet-delivered psychological treatments. *Annual Review of Clinical Psychology, 12*, 157-179.
- Andersson, G., Carlbring, P., Berger, T., Almlöv, J., & Cuijpers, P. (2009). What makes Internet therapy work? *Cognitive Behaviour Therapy, 38*(1), 55-60.
- Andersson, G., Carlbring, P., Ljótsson, B., & Hedman, E. (2013). Guided Internet-based CBT for common mental disorders. *Journal of Contemporary Psychotherapy, 43*(4), 223-233.
- 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.
- Andersson, G., Cuijpers, P., Carlbring, P., Riper, H., & Hedman, E. (2014). Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: A systematic review and meta-analysis. *World Psychiatry, 13*(3), 288-295.
- Andersson, G., Hesser, H., Hummerdal, D., Bergman-Nordgren, L., & Carlbring, P. (2013). A 3.5-year follow-up of Internet-delivered cognitive behavior therapy for major depression. *Journal of Mental Health, 22*(2), 155-164.
- Andersson, G., Rozental, A., Rück, C., & Carlbring, P. (2015). Guided Internet-delivered CBT: Can it really be as good as seeing a therapist? *The Behavior Therapist, 35*, 123-126.

- Andersson, G., Topooco, N., Havik, O., & Nordgreen, T. (2016). Internet-supported versus face-to-face cognitive behavior therapy for depression. *Expert Review of Neurotherapeutics*, *16*(1), 55-60.
- Arnberg, F. K., Linton, S. J., Hultcrantz, M., Heintz, E., & Jonsson, U. (2014). Internet-Delivered Psychological Treatments for Mood and Anxiety Disorders: A Systematic Review of Their Efficacy, Safety, and Cost-Effectiveness. *Plos One*, *9*(5), e98118.
- Bachem, R., & Maercker, A. (2016). Self-help interventions for adjustment disorder problems: A randomized waiting-list controlled study in a sample of burglary victims. *Cognitive Behaviour Therapy*, *45*(5), 397-413.
- Baldwin, D. S., Anderson, I. M., Nutt, D. J., Allgulander, C., Bandelow, B., den Boer, J. A., . . . Lidbetter, N. (2014). Evidence-based pharmacological treatment of anxiety disorders, post-traumatic stress disorder and obsessive-compulsive disorder: A revision of the 2005 guidelines from the British Association for Psychopharmacology. *Journal of Psychopharmacology*, *28*(5), 403-439.
- Barlow, D. H. (2010). Negative effects from psychological treatments: A perspective. *American Psychologist*, *65*(1), 13-20.
- Bauer, S., Lambert, M. J., & Nielsen, S. L. (2004). Clinical significance methods: A comparison of statistical techniques. *Journal of Personality Assessment*, *82*(1), 60-70.
- Baumeister, H., Reichler, L., Munzinger, M., & Lin, J. (2014). The impact of guidance on Internet-based mental health interventions - A systematic review. *Internet Interventions*, *1*(4), 205-215.
- Beck, A. T., Erbaugh, J., Ward, C. H., Mock, J., & Mendelsohn, M. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*(6), 561-571.
- Bendelin, N., Hesser, H., Dahl, J., Carlbring, P., Nelson, K., & Andersson, G. (2011). Experiences of guided Internet-based cognitive-behavioural treatment for depression: A qualitative study. *BMC Psychiatry*, *11*(1), 107.
- Bengtsson, J., Nordin, S., & Carlbring, P. (2015). Therapists' experiences of conducting cognitive behavioural therapy online vis-à-vis face-to-face. *Cognitive Behaviour Therapy*, *44*(6), 470-479.
- Bennett, K., Bennett, A. J., & Griffiths, K. M. (2010). Security considerations for e-mental health interventions. *Journal of Medical Internet Research*, *12*(5), e61.
- Berger, T., Boettcher, J., & Caspar, F. (2014). Internet-based guided self-help for several anxiety disorders: A randomized controlled trial comparing a tailored with a standardized disorder-specific approach. *Psychotherapy*, *51*(2), 207-219.
- Bergin, A. E. (1963). The empirical emphasis in psychotherapy: A symposium - The effects of psychotherapy: Negative results revisited. *Journal of Counseling Psychology*, *10*(3), 244-250.

- Bergin, A. E. (1966). Some implications of psychotherapy research for therapeutic practice. *Journal of Abnormal Psychology, 71*(4), 235-246.
- Berk, M., & Parker, G. (2009). The elephant on the couch: Side-effects of psychotherapy. *Australian and New Zealand Journal of Psychiatry, 43*(9), 787-794.
- Beyerstein, B. L. (2001). Fringe psychotherapies: The public at risk. *The Scientific Review of Alternative Medicine, 5*(2), 70-79.
- Binder, J. L., & Strupp, H. H. (1997). "Negative process": A recurrently discovered and underestimated facet of therapeutic process and outcome in the individual psychotherapy of adults. *Clinical Psychology: Science and Practice, 4*(2), 121-139.
- Blom, K., Tarkian Tillgren, H., Wiklund, T., Danlycke, E., Forssén, M., Söderström, A., . . . Kaldo, V. (2015). Internet- vs. group-delivered cognitive behavior therapy for insomnia: A randomized controlled non-inferiority trial. *Behaviour Research and Therapy, 70*, 47-55.
- Boettcher, J., Andersson, G., & Carlbring, P. (2013). Combining attention training with cognitive-behavior therapy in Internet-based self-help for social anxiety: Study protocol for a randomized controlled trial. *Trials, 14*, 68.
- Boettcher, J., Hasselrot, J., Sund, E., Andersson, G., & Carlbring, P. (2013). Combining attention training with Internet-based cognitive-behavioural self-help for social anxiety: A randomised controlled trial. *Cognitive Behaviour Therapy, 43*(1), 34-48.
- Boisvert, C. M. (2010). Negative treatment effects: Is it time for a black box warning? *American Psychologist, 65*(7), 680-681.
- Boisvert, C. M., & Faust, D. (2002). Iatrogenic symptoms in psychotherapy - A theoretical exploration of the potential impact of labels, language, and belief systems. *American Journal of Psychotherapy, 56*(2), 244-259.
- Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research, 25*(1), 6-19.
- Bower, P., Kontopantelis, E., Sutton, A., Kendrick, T., Richards, D. A., Gilbody, S., . . . Liu, E. T. (2013). Influence of initial severity of depression on effectiveness of low intensity interventions: Meta-analysis of individual patient data. *BMJ, 346*, f540.
- Brooke, J. (1996). SUS - A quick and dirty usability scale. *Usability Evaluation in Industry, 189*(194), 4-7.
- Buckley, P., Karasu, T. B., & Charles, E. (1981). Psychotherapists view their personal therapy. *Psychotherapy: Theory, Research & Practice, 18*(3), 299-305.
- Buntrock, C., Ebert, D., Lehr, D., Riper, H., Smit, F., Cuijpers, P., & Berking, M. (2015). Effectiveness of a web-based cognitive behavioural intervention for subthreshold depression: Pragmatic

- randomised controlled trial. *Psychotherapy and Psychosomatics*, 84(6), 348-358.
- Bystedt, S., Rozental, A., Andersson, G., Boettcher, J., & Carlbring, P. (2014). Clinicians' perspectives on negative effects of psychological treatments. *Cognitive Behaviour Therapy*, 43(4), 319-331.
- Carlbring, P., & Andersson, G. (2006). Internet and psychological treatment. How well can they be combined? *Computers in Human Behavior*, 22(3), 545-553.
- Carlbring, P., Bohman, S., Brunt, S., Buhrman, M., Westling, B. E., Ekselius, L., & Andersson, G. (2006). Remote treatment of panic disorder: A randomized trial of Internet-based cognitive behavior therapy supplemented with telephone calls. *American Journal of Psychiatry*, 163(12), 2119-2125.
- Cartwright, R. D., & Vogel, J. L. (1960). A comparison of changes in psychoneurotic patients during matched periods of therapy and no therapy. *Journal of Consulting Psychology*, 24(2), 121-127.
- Castonguay, L. G., Boswell, J. F., Constantino, M. J., Goldfried, M. R., & Hill, C. E. (2010). Training implications of harmful effects of psychological treatments. *American Psychologist*, 65(1), 34-49.
- Christensen, H., Griffiths, K. M., & Farrer, L. (2009). Adherence in Internet interventions for anxiety and depression: Systematic review. *Journal of Medical Internet Research*, 11(2), e13.
- Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject - An alteration of the RC index. *Behavior Therapy*, 17(3), 305-308.
- Connor, K. M., Kobak, K. A., Churchill, L. E., Katzelnick, D., & Davidson, J. R. (2001). Mini-SPIN: A brief screening assessment for generalized social anxiety disorder. *Depression and Anxiety*, 14(2), 137-140.
- Crawford, M. J., Thana, L., Farquharson, L., Palmer, L., Hancock, E., Bassett, P., . . . Parry, G. D. (2016). Patient experience of negative effects of psychological treatment: Results of a national survey. *British Journal of Psychiatry*, 208(3), 260-265.
- Cuijpers, P., Donker, T., van Straten, A., Li, J., & Andersson, G. (2010). Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. *Psychological Medicine*, 40(12), 1943-1957.
- Curtin, F., & Schulz, P. (2011). Assessing the benefit: Risk ratio of a drug - Randomized and naturalistic evidence. *Dialogues in Clinical Neuroscience*, 13(2), 183-190.
- Dahlin, M., Andersson, G., Magnusson, K., Johansson, T., Sjögren, J., Håkansson, A., . . . Carlbring, P. (2016). Internet-delivered acceptance-based behaviour therapy for generalized anxiety disorder: A randomized controlled trial. *Behaviour Research and Therapy*, 77, 86-95.

- Davydov, D. M., Stewart, R., Ritchie, K., & Chaudieu, I. (2010). Resilience and mental health. *Clinical Psychology Review, 30*(5), 479-495.
- Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243*(4899), 1668-1674.
- Dear, B. F., Gandy, M., Karin, E., Staples, L. G., Johnston, L., Fogliati, V. J., . . . Perry, K. N. (2015). The Pain Course: A randomised controlled trial examining an internet-delivered pain management program when provided with different levels of clinician support. *Pain, 156*(10), 1920-1935.
- Delgadillo, J., Moreea, O., & Lutz, W. (2016). Different people respond differently to therapy: A demonstration using patient profiling and risk stratification. *Behaviour Research and Therapy, 79*, 15-22.
- Dimidjian, S., & Hollon, S. D. (2010). How would we know if psychotherapy were harmful? *American Psychologist, 65*(1), 21-33.
- Dobson, K. S. (2009). *Handbook of cognitive-behavioral therapies*. New York, NY: Guilford Press.
- Ebert, D. D., Heber, E., Berking, M., Riper, H., Cuijpers, P., Funk, B., & Lehr, D. (2016). Self-guided internet-based and mobile-based stress management for employees: Results of a randomised controlled trial. *Occupational and Environmental Medicine, 73*(5), 315-323.
- Edwards, D. W., Yarvis, R. M., Mueller, D. P., Zingale, H. C., & Wagman, W. J. (1978). Test taking and the stability of adjustment scales: Can we assess patient deterioration? *Evaluation Quarterly, 2*(2), 275-291.
- El Alaoui, S., Hedman, E., Ljótsson, B., & Lindefors, N. (2015). Long-term effectiveness and outcome predictors of therapist-guided internet-based cognitive-behavioural therapy for social anxiety disorder in routine psychiatric care. *BMJ Open, 5*(6), e007902.
- El Alaoui, S., Ljótsson, B., Hedman, E., Kaldo, V., Andersson, E., Rück, C., . . . Lindefors, N. (2015). Predictors of symptomatic change and adherence in internet-based cognitive behaviour therapy for social anxiety disorder in routine psychiatric care. *Plos One, 10*(4), e0124258.
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107-115.
- Emmelkamp, P. M., David, D., Beckers, T., Muris, P., Cuijpers, P., Lutz, W., . . . Vervliet, B. (2014). Advancing psychotherapy and evidence-based psychological interventions. *International Journal of Methods in Psychiatric Research, 23*(1), 58-91.
- Enander, J., Ivanov, V. Z., Andersson, E., Mataix-Cols, D., Ljotsson, B., & Ruck, C. (2014). Therapist-guided, Internet-based cognitive-behavioural therapy for body dysmorphic disorder (BDD-NET): A feasibility study. *BMJ Open, 4*(9), e005923.
- Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical outcomes in routine evaluation. *Journal of Mental Health, 9*(3), 247-255.

- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology, 16*(5), 319-324.
- Farmer, R. F., & Chapman, A. L. (2016). *Behavioral interventions in cognitive behavior therapy*. Washington, DC: American Psychological Association.
- Febbraro, G. A., Clum, G. A., Roodman, A. A., & Wright, J. H. (1999). The limits of bibliotherapy: A study of the differential effectiveness of self-administered interventions in individuals with panic attacks. *Behavior Therapy, 30*(2), 209-222.
- First, M. B., Gibbon, M., Spitzer, R. L., & Williams, J. B. (1997). *Structured clinical interview for DSM-IV Axis I Disorders (SCID-I)*. Washington, DC: American Psychiatric Press.
- Foa, E. B., & Emmelkamp, P. M. G. (1983). *Failures in behavior therapy*. New York, NY: John Wiley & Sons.
- Foa, E. B., & Steketee, G. (1977). Emergent fears during treatment of three obsessive compulsives: Symptom substitution or deconditioning? *Journal of Behavior Therapy and Experimental Psychiatry, 8*(4), 353-358.
- Foulkes, P. (2010). The therapist as a vital factor in side-effects of psychotherapy. *The Australian and New Zealand Journal of Psychiatry, 44*(2), 189.
- Frisch, M. B., Cornell, J., Villanueva, M., & Retzlaff, P. J. (1992). Clinical validation of the Quality of Life Inventory. A measure of life satisfaction for use in treatment planning and outcome assessment. *Psychological Assessment, 4*(1), 92-101.
- Hadjistavropoulos, H. D., Nugent, M. M., Alberts, N. M., Staples, L., Dear, B. F., & Titov, N. (2016). Transdiagnostic Internet-delivered cognitive behaviour therapy in Canada: An open trial comparing results of a specialized online clinic and nonspecialized community clinics. *Journal of Anxiety Disorders, 42*, 19-29.
- Hadley, S. W., & Strupp, H. H. (1976). Contemporary views of negative effects in psychotherapy - Integrated account. *Archives of General Psychiatry, 33*(11), 1291-1302.
- Haldeman, D. C. (1994). The practice and ethics of sexual orientation conversion therapy. *Journal of Consulting and Clinical Psychology, 62*(2), 221-227.
- Halmetoja, C. O., Malmquist, A., Carlbring, P., & Andersson, G. (2014). Experiences of internet-delivered cognitive behavior therapy for social anxiety disorder four years later: A qualitative study. *Internet Interventions, 1*(3), 158-163.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry, 23*(1), 56-62.
- Hand, I., & Lamontagne, Y. (1976). The exacerbation of interpersonal problems after rapid phobia-removal. *Psychotherapy: Theory, Research & Practice, 13*(4), 405-411.

- Handelsman, M. M. (1990). Do written consent forms influence clients' first impressions of therapists? *Professional Psychology: Research and Practice, 21*(6), 451-454.
- Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice, 9*(3), 329-343.
- Hatfield, D., McCullough, L., Frantz, S. H. B., & Krieger, K. (2010). Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. *Clinical Psychology & Psychotherapy, 17*(1), 25-32.
- Haynes, S. N., & Horn, W. F. (1982). Reactivity in behavioral observation: A methodological and conceptual critique. *Behavioral Assessment, 4*, 369-385.
- Hedman, E., Andersson, G., Ljótsson, B., Andersson, E., Rück, C., Mörtberg, E., & Lindefors, N. (2011). Internet-based cognitive behavior therapy vs. cognitive behavioral group therapy for social anxiety disorder: A randomized controlled non-inferiority trial. *Plos One, 6*(3), e18001.
- Hedman, E., Furmark, T., Carlbring, P., Ljotsson, B., Ruck, C., Lindefors, N., & Andersson, G. (2011). A 5-Year follow-up of internet-based cognitive behavior therapy for social anxiety disorder. *Journal of Medical Internet Research, 13*(2), e39.
- Hedman, E., Ljótsson, B., Kaldo, V., Hesser, H., El Alaoui, S., Kraepelien, M., . . . Andersson, G. (2014). Effectiveness of Internet-based cognitive behaviour therapy for depression in routine psychiatric care. *Journal of Affective Disorders, 155*, 49-58.
- Hedman, E., Ljótsson, B., Rück, C., Furmark, T., Carlbring, P., Lindefors, N., & Andersson, G. (2010). Internet administration of self-report measures commonly used in research on social anxiety disorder: A psychometric evaluation. *Computers in Human Behavior, 26*(4), 736-740.
- Heide, F. J., & Borkovec, T. (1983). Relaxation-induced anxiety: Paradoxical anxiety enhancement due to relaxation training. *Journal of Consulting and Clinical Psychology, 51*(2), 171-182.
- Hilgart, M., Ritterband, L., Baxter, K., Alfano, A., Ratliff, C., Kinzie, M., . . . Garber, S. (2014). Development and perceived utility and impact of a skin care Internet intervention. *Internet Interventions, 1*(3), 149-157.
- Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research, 36*(5), 427-440.
- Hovland, A., Johansen, H., Sjøbø, T., Vøllestad, J., Nordhus, I. H., Pallesen, S., . . . Nordgreen, T. (2015). A Feasibility study on combining Internet-based cognitive behaviour therapy with physical exercise as treatment for panic disorder - Treatment protocol and preliminary results. *Cognitive Behaviour Therapy, 44*(4), 275-287.

- Hsu, L. M. (1995). Regression toward the mean associated with measurement error and the identification of improvement and deterioration in psychotherapy. *Journal of Consulting and Clinical Psychology, 63*(1), 141-144.
- Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15*(4), 336-352.
- Jacobson, N. S., Follette, W. C., Revenstorf, D., Baucom, D. H., Hahlweg, K., & Margolin, G. (1984). Variability in outcome and clinical significance of behavioral marital therapy: A reanalysis of outcome data. *Journal of Consulting and Clinical Psychology, 52*(4), 497-504.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12-19.
- Jarrett, C. (2008). When therapy causes harm. *Psychologist, 21*(1), 10-12.
- Johansson, R., & Andersson, G. (2012). Internet-based psychological treatments for depression. *Expert Review of Neurotherapeutics, 12*(7), 861-870.
- Johansson, R., Sjöberg, E., Sjögren, M., Johnsson, E., Carlbring, P., Andersson, T., . . . Andersson, G. (2012). Tailored vs. standardized internet-based cognitive behavior therapy for depression and comorbid symptoms: A randomized controlled trial. *Plos One, 7*(5), e36905.
- Jonsson, U., Alaie, I., Parling, T., & Arnberg, F. K. (2014). Reporting of harms in randomized controlled trials of psychological interventions for mental and behavioral disorders: A review of current practice. *Contemporary Clinical Trials, 38*(1), 1-8.
- Jonsson, U., Johanson, J., Nilsson, E., & Lindblad, F. (2015). Adverse effects of psychological therapy: An exploratory study of practitioners' experiences from child and adolescent psychiatry. *Clinical Child Psychology and Psychiatry, 21*(3), 432-446.
- Kaldo, V., Jernelov, S., Blom, K., Ljotsson, B., Brodin, M., Jorgensen, M., . . . Lindefors, N. (2015). Guided internet cognitive behavioral therapy for insomnia compared to a control treatment - A randomized trial. *Behaviour Research and Therapy, 71*, 90-100.
- Kaldo, V., Ramnerö, J., & Jernelöv, S. (2015). Involving clients in treatment methods: A neglected interaction in the therapeutic relationship. *Journal of Consulting and Clinical Psychology, 83*(6), 1136-1141.
- Karyotaki, E., Kleiboer, A., Smit, F., Turner, D. T., Pastor, A. M., Andersson, G., . . . Cuijpers, P. (2015). Predictors of treatment dropout in self-guided web-based interventions for depression: An 'individual patient data' meta-analysis. *Psychological Medicine, 45*(13), 2717-2726.
- Kayrouz, R., Dear, B. F., Karin, E., Fogliati, V. J., & Titov, N. (2016). A pilot study of a clinician-guided internet-delivered cognitive behavioural therapy for anxiety and depression among Arabs in

- Australia, presented in both English and Arabic languages. *Internet Interventions*, 5, 5-11.
- Kazantzis, N., Deane, F. P., & Ronan, K. R. (2000). Homework assignments in cognitive and behavioral therapy: A meta-analysis. *Clinical Psychology: Science and Practice*, 7(2), 189-202.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6), 593-602.
- Kivi, M., Eriksson, M. C. M., Hange, D., Petersson, E.-L., Vernmark, K., Johansson, B., & Björkelund, C. (2014). Internet-based therapy for mild to moderate depression in Swedish primary care: Short term results from the PRIM-NET randomized controlled trial. *Cognitive Behaviour Therapy*, 43(4), 289-298.
- Klein, J. P., Berger, T., Schröder, J., Spith, C., Meyer, B., Caspar, F., . . . Hohagen, F. (2016). Effects of a psychological internet intervention in the treatment of mild to moderate depressive symptoms: Results of the evident study, a randomized controlled trial. *Psychotherapy and Psychosomatics*, 85(4), 218-228.
- Klorman, R., Weerts, T. C., Hastings, J. E., Melamed, B. G., & Lang, P. J. (1974). Psychometric description of some specific-fear questionnaires. *Behavior Therapy*, 5(3), 401-409.
- Kottler, J. A., & Carlson, J. (2003). *Bad therapy: Master therapists share their worst failures*. New York, NY: Taylor & Francis Books, Inc.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606-613.
- Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, 50(1), 537-567.
- Kächele, H., & Schachter, J. (2014). On side effects, destructive processes, and negative outcomes in psychoanalytic therapies: Why is it difficult for psychoanalysts to acknowledge and address treatment failures? *Contemporary Psychoanalysis*, 50(1-2), 233-258.
- Ladwig, I., Rief, W., & Nestoriuc, Y. (2014). What are the risks and side effects to psychotherapy? - Development of an inventory for the assessment of negative effects of psychotherapy (INEP). *Verhaltenstherapie*, 24(4), 252-263.
- Lambert, M. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research*, 17(1), 1-14.
- Lambert, M. (2010). *Prevention of treatment failure - The use of measuring, monitoring, and feedback in clinical practice*. Washington, DC: American Psychological Association.
- Lambert, M. (2013). Outcome in psychotherapy: The past and important advances. *Psychotherapy*, 50(1), 42-51.
- Lambert, M., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., & Yanchar, S. C. (1996). The reliability and

- validity of the Outcome Questionnaire. *Clinical Psychology & Psychotherapy*, 3(4), 249-258.
- Lambert, M., Whipple, J. L., Smart, D. W., Vermeersch, D. A., Nielsen, S. L., & Hawkins, E. J. (2001). The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research*, 11(1), 49-68.
- Lambert, M., Whipple, J. L., Vermeersch, D. A., Smart, D. W., Hawkins, E. J., Nielsen, S. L., & Goates, M. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology & Psychotherapy*, 9(2), 91-103.
- Lancee, J., Eisma, M. C., van Straten, A., & Kamphuis, J. H. (2015). Sleep-related safety behaviors and dysfunctional beliefs mediate the efficacy of online CBT for insomnia: A randomized controlled trial. *Cognitive Behaviour Therapy*, 44(5), 406-422.
- Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: development of a general scale. *Evaluation and Program Planning*, 2(3), 197-207.
- Lewis, C., Pearce, J., & Bisson, J. I. (2012). Efficacy, cost-effectiveness and acceptability of self-help interventions for anxiety disorders: systematic review. *The British Journal of Psychiatry*, 200(1), 15-21.
- Liebowitz, M. R. (1987). Social phobia. *Modern Problems of Pharmacopsychiatry*, 22, 141-173.
- Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science*, 2(1), 53-70.
- Lindfors, N., & Andersson, G. (2016). *Guided Internet-based treatments in psychiatry*. Cham, Switzerland: Springer International Publishing.
- Linden, M. (2013). How to define, find and classify side effects in psychotherapy: From unwanted events to adverse treatment reactions. *Clinical Psychology & Psychotherapy*, 20(4), 286-296.
- Linden, M., & Schermuly-Haupt, M.-L. (2014). Definition, assessment and rate of psychotherapy side effects. *World Psychiatry*, 13(3), 306-309.
- Lindner, P., Frykheden, O., Forsström, D., Andersson, E., Ljótsson, B., Hedman, E., . . . Carlbring, P. (2016). The Brunnsviken Brief Quality of life scale (BBQ): Development and psychometric evaluation. *Cognitive Behaviour Therapy*, 45(3), 182-195.
- Lindner, P., Ivanova, E., Ly, K., Andersson, G., & Carlbring, P. (2013). Guided and unguided CBT for social anxiety disorder and/or panic disorder via the Internet and a smartphone application: Study protocol for a randomised controlled trial. *Trials*, 14(1), 437.
- Loerinc, A. G., Meuret, A. E., Twohig, M. P., Rosenfield, D., Bluett, E. J., & Craske, M. G. (2015). Response rates for CBT for anxiety disorders: need for standardized criteria. *Clinical Psychology Review*, 42, 72-82.
- Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-

- analytic review. *Journal of Consulting and Clinical Psychology*, 68(3), 438-450.
- Martinez, R., Whitfield, G., Dafters, R., & Williams, C. (2007). Can people read self-help manuals for depression? A challenge for the stepped care model and book prescription schemes. *Behavioural and Cognitive Psychotherapy*, 36(1), 89-97.
- Martinovich, Z., Saunders, S., & Howard, K. I. (1996). Some comments on "assessing clinical significance". *Psychotherapy Research*, 6(2), 124-132.
- Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, 36(4), 455-470.
- May, P. (1971). For better or for worse? Psychotherapy and variance change: A critical review of the literature. *The Journal of Nervous and Mental Disease*, 152(3), 184-192.
- Mays, D. T., & Franks, C. M. (1980). Getting worse: Psychotherapy or no treatment - The jury should still be out. *Professional Psychology*, 11(1), 78-92.
- Mays, D. T., & Franks, C. M. (1985). *Negative Outcome in psychotherapy and what to do about it*. New York, NY: Springer Publishing Company, Inc.
- Mechler, J., & Holmqvist, R. (2016). Deteriorated and unchanged patients in psychological treatment in Swedish primary care and psychiatry. *Nordic Journal of Psychiatry*, 70(1), 16-23.
- Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, 28(6), 487-495.
- Miloff, A., Marklund, A., & Carlbring, P. (2015). The Challenger app for social anxiety disorder: New advances in mobile psychological treatment. *Internet Interventions*, 2(4), 382-391.
- Mohr, D. C. (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology: Science and Practice*, 2(1), 1-27.
- Mohr, D. C., Beutler, L. E., Engle, D., Shohamsalomon, V., Bergan, J., Kaszniak, A. W., & Yost, E. B. (1990). Identification of patients at risk for nonresponse and negative outcome in psychotherapy. *Journal of Consulting and Clinical Psychology*, 58(5), 622-628.
- Moritz, S., Fieker, M., Hottenrott, B., Seeralan, T., Cludius, B., Kolbeck, K., . . . Nestoriuc, Y. (2015). No pain, no gain? Adverse effects of psychotherapy in obsessive-compulsive disorder and its relationship to treatment gains. *Journal of Obsessive-Compulsive and Related Disorders*, 5, 61-66.
- Moskowitz, D. S., & Young, S. N. (2006). Ecological momentary assessment: What it is and why it is a method of the future in clinical psychopharmacology. *Journal of Psychiatry & Neuroscience: JPN*, 31(1), 13-20.
- Nieminen, K., Berg, I., Frankenstein, K., Viita, L., Larsson, K., Persson, U., . . . Wijma, K. (2016). Internet-provided cognitive behaviour therapy

- of posttraumatic stress symptoms following childbirth - A randomized controlled trial. *Cognitive Behaviour Therapy*, 45(4), 287-306.
- Nordgreen, T., Haug, T., Öst, L. G., Andersson, G., Carlbring, P., Kvale, G., . . . Havik, O. E. (2016). Stepped care versus direct face-to-face cognitive behavior therapy for social anxiety disorder and panic disorder: A randomized effectiveness trial. *Behavior Therapy*, 47(2), 166-183.
- Nordgren, L. B., Carlbring, P., Linna, E., & Andersson, G. (2013). Role of the working alliance on treatment outcome in tailored internet-based cognitive behavioural therapy for anxiety disorders: randomized controlled pilot trial. *JMIR Research Protocols*, 2(1), e4.
- Nordgren, L. B., Hedman, E., Etienne, J., Bodin, J., Kadowaki, Å., Eriksson, S., . . . Carlbring, P. (2014). Effectiveness and cost-effectiveness of individually tailored Internet-delivered cognitive behavior therapy for anxiety disorders in a primary care population: A randomized controlled trial. *Behaviour Research and Therapy*, 59, 1-11.
- Nordin, S., Carlbring, P., Cuijpers, P., & Andersson, G. (2010). Expanding the limits of bibliotherapy for panic disorder: Randomized trial of self-help without support but with a clear deadline. *Behavior Therapy*, 41(3), 267-276.
- Nordmo, M., Sinding, A. I., Carlbring, P., Andersson, G., Havik, O. E., & Nordgreen, T. (2015). Internet-delivered cognitive behavioural therapy with and without an initial face-to-face psychoeducation session for social anxiety disorder: A pilot randomized controlled trial. *Internet Interventions*, 2(4), 429-436.
- Ogles, B. M., Lambert, M. J., & Sawyer, J. D. (1995). Clinical significance of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Journal of Consulting and Clinical Psychology*, 63(2), 321-326.
- Parker, G., Fletcher, K., Berk, M., & Paterson, A. (2013). Development of a measure quantifying adverse psychotherapeutic ingredients: The Experiences of Therapy Questionnaire (ETQ). *Psychiatry Research*, 206(2-3), 293-301.
- Parry, G. D., Crawford, M. J., & Duggan, C. (2016). Iatrogenic harm from psychological therapies - Time to move on. *The British Journal of Psychiatry*, 208(3), 210-212.
- 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.
- Powers, E., & Witmer, H. L. (1951). *An experiment in the prevention of delinquency : The Cambridge Somerville youth study*. New York, Ny: Columbia University Press.
- Rachman, S. (1971). *The effects of psychotherapy*. Oxford, UK: Pergamon.
- Ramnerö, J., & Törneke, N. (2008). *The ABCs of human behavior*. Oakland, CA: New Harbinger Publications, Inc.

- Ramnerö, J., & Törneke, N. (2015). On having a goal: Goals as representations or behavior. *The Psychological Record*, 65(1), 89-99.
- Rheker, J., Andersson, G., & Weise, C. (2015). The role of "on demand" therapist guidance vs. no support in the treatment of tinnitus via the internet: A randomized controlled trial. *Internet Interventions*, 2(2), 189-199.
- Ritterband, L., Ardan, K., Thorndike, F., Magee, J., Saylor, D., Cox, D., . . . Borowitz, S. (2008). Real world use of an Internet intervention for pediatric encopresis. *Journal of Medical Internet Research*, 10(2), e16.
- Roback, H. B. (2000). Adverse outcomes in group psychotherapy: Risk factors, prevention, and research directions. *The Journal of Psychotherapy Practice and Research*, 9(3), 113-122.
- Rozental, A., Andersson, G., Boettcher, J., Ebert, D. D., Cuijpers, P., Knaevelsrud, C., . . . Carlbring, P. (2014). Consensus statement on defining and measuring negative effects of Internet interventions. *Internet Interventions*, 1(1), 12-19.
- Rozental, A., Forsell, E., Svensson, A., Andersson, G., & Carlbring, P. (2015). Internet-based cognitive-behavior therapy for procrastination: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 83(4), 808-824.
- Rozental, A., Kottorp, A., Boettcher, J., Andersson, G., & Carlbring, P. (2016). Negative effects of psychological treatments: An exploratory factor analysis of an instrument measuring adverse and unwanted events. *Plos One*, 11(6), e0157503.
- 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(4), 387-396.
- 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.
- Sachs, J. S. (1983). Negative factors in brief psychotherapy - An empirical assessment. *Journal of Consulting and Clinical Psychology*, 51(4), 557-564.
- Saekow, J., Jones, M., Gibbs, E., Jacobi, C., Fitzsimmons-Craft, E. E., Wilfley, D., & Barr Taylor, C. (2015). StudentBodies-eating disorders: A randomized controlled trial of a coached online intervention for subclinical eating disorders. *Internet Interventions*, 2(4), 419-428.
- Sarkozy, N. (2010). *Addressing negative effects of psychotherapy during the informed consent process: The licensed psychologists' perspectives*. (Unpublished doctoral dissertation), Philadelphia College of Osteopathic Medicine, Philadelphia, PA.

- Saxon, D., & Barkham, M. (2012). Patterns of therapist variability: Therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology, 80*(4), 535-546.
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54*(2), 93-105.
- Scogin, F., Floyd, M., Jamison, C., Ackerson, J., Landreville, P., & Bissonnette, L. (1996). Negative outcomes: What is the evidence on self-administered treatments? *Journal of Consulting and Clinical Psychology, 64*(5), 1086-1089.
- Scott, J., & Young, A. H. (2016). Psychotherapies should be assessed for both benefit and harm. *The British Journal of Psychiatry, 208*(3), 208-209.
- Seggar, L. B., Lambert, M. J., & Hansen, N. B. (2002). Assessing clinical significance: Application to the Beck Depression Inventory. *Behavior Therapy, 33*(2), 253-269.
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., . . . Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 59*(20), 22-33.
- Smith, C. M. (2005). Origin and uses of primum non nocere - Above all, do no harm! *The Journal of Clinical Pharmacology, 45*(4), 371-377.
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder. *Archives of Internal Medicine, 166*(10), 1092-1097.
- Statistics Sweden. (2014). *Educational attainment and employment of the population 2014* (Vol. UF 84 SM 1601). Stockholm: Statistics Sweden.
- Strupp, H. H., & Hadley, S. W. (1977). A tripartite model of mental health and therapeutic outcomes. With special reference to negative effects in psychotherapy. *The American Psychologist, 32*(3), 187-196.
- Strupp, H. H., Hadley, S. W., & Gomes-Schwartz, B. (1977). *When things get worse: The problem of negative effects in psychotherapy*. Washington, DC: Jason Aronson, Inc.
- Ström, M., Uckelstam, C. J., Andersson, G., Hassmen, P., Umefjord, G., & Carlbring, P. (2013). Internet-delivered therapist-guided physical activity for mild to moderate depression: A randomized controlled trial. *PeerJ, 1*, e178.
- Stuart, R. B. (1970). *Trick or treatment: How and when psychotherapy fails*. Champaign, IL: Research Press.
- Suh, C. S., Strupp, H. H., & O'Malley, S. S. (1986). The Vanderbilt process measures: The Psychotherapy Process Scale (VPPS) and the Negative Indicators Scale (VNIS). In L. S. Greenberg (Ed.), *The psychotherapeutic process: A research handbook* (pp. 285-323). New York, NY: Guilford Press.

- Svanborg, P., & Åsberg, M. (1994). A new self-rating scale for depression and anxiety states based on the Comprehensive Psychopathological Rating Scale. *Acta Psychiatrica Scandinavica*, 89(1), 21-28.
- Tarrier, N., & Johnson, J. (2015). *Case formulation in cognitive behaviour therapy: The treatment of challenging and complex cases*. Abingdon-on-Thames, UK: Routledge.
- Thorndike, F. P., Carlbring, P., Smyth, F. L., Magee, J. C., Gonder-Frederick, L., Öst, L.-G., & Ritterband, L. M. (2009). Web-based measurement: Effect of completing single or multiple items per webpage. *Computers in Human Behavior*, 25(2), 393-401.
- Titov, N., Fogliati, V. J., Staples, L. G., Gandy, M., Johnston, L., Wootton, B., . . . Dear, B. F. (2016). Treating anxiety and depression in older adults: randomised controlled trial comparing guided v. self-guided internet-delivered cognitive-behavioural therapy. *British Journal of Psychiatry Open*, 2(1), 50-58.
- Truax, C. B. (1963). Effective ingredients in psychotherapy. *Journal of Consulting Psychology*, 10(3), 256-263.
- Vaughan, B., Goldstein, M. H., Alikakos, M., Cohen, L. J., & Serby, M. J. (2014). Frequency of reporting of adverse events in randomized controlled trials of psychotherapy vs. psychopharmacotherapy. *Comprehensive Psychiatry*, 55(4), 849-855.
- Vessey, J. T., & Howard, K. I. (1993). Who seeks psychotherapy? *Psychotherapy: Theory, Research, Practice, Training*, 30(4), 546-553.
- Walker, R., Logan, T., Clark, J. J., & Leukefeld, C. (2005). Informed consent to undergo treatment for substance abuse: A recommended approach. *Journal of Substance Abuse Treatment*, 29(4), 241-251.
- Waller, R., & Gilbody, S. (2009). Barriers to the uptake of computerized cognitive behavioural therapy: A systematic review of the quantitative and qualitative evidence. *Psychological Medicine*, 39(5), 705-712.
- Wang, L., Zhang, Z., McArdle, J. J., & Salthouse, T. A. (2008). Investigating ceiling effects in longitudinal data analysis. *Multivariate Behavioral Research*, 43(3), 476-496.
- Warren, J. S., Nelson, P. L., Mondragon, S. A., Baldwin, S. A., & Burlingame, G. M. (2010). Youth psychotherapy change trajectories and outcomes in usual care: Community mental health versus managed care settings. *Journal of Consulting and Clinical Psychology*, 78(2), 144-155.
- Weise, C., Kleinstäuber, M., & Andersson, G. (2016). Internet-delivered cognitive-behavior therapy for tinnitus: A randomized controlled trial. *Psychosomatic Medicine*, 78(4), 501-510.
- Westbrook, D., Kennerley, H., & Kirk, J. (2011). *An introduction to cognitive behaviour therapy: Skills and applications*. Thousands Oaks, CA: Sage publications.
- Whipple, J. L., Lambert, M. J., Vermeersch, D. A., Smart, D. W., Nielsen, S. L., & Hawkins, E. J. (2003). Improving the effects of

- psychotherapy: The use of early identification of treatment failure and problem-solving strategies in routine practice. *Journal of Counseling Psychology*, 50(1), 59-68.
- Williams, C., & Martinez, R. (2008). Increasing access to CBT: stepped care and CBT self-help models in practice. *Behavioural and Cognitive Psychotherapy*, 36(6), 675-683.
- Wise, E. A. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendations for future directions. *Journal of Personality Assessment*, 82(1), 50-59.
- World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *Journal of the American Medical Association*, 48(3), 2191-2194.
- Wysowski, D. K., & Swartz, L. (2005). Adverse drug event surveillance and drug withdrawals in the United States, 1969-2002 - The importance of reporting suspected reactions. *Archives of Internal Medicine*, 165(12), 1363-1369.
- Young, C. (2006). What happens when people wait for therapy? Assessing the clinical significance of the changes observed over the waiting period for clients referred to a primary care psychology service. *Primary Care Mental Health*, 4(2), 113-119.