The impact of guidance on Internet-based mental health interventions — A systematic review

H. Baumeister [a,b,*], L. Reichler [a], M. Munzinger [a], J. Lin [a,b]

[a] Department of Rehabilitation Psychology and Psychotherapy, Institute of Psychology, University of Freiburg, Germany
[b] Medical Psychology and Medical Sociology, Faculty of Medicine, University of Freiburg, Germany

ARTICLE INFO

ABSTRACT

Introduction: The aim of this study was to systematically review the impact of guidance on the efficacy of Internet-based interventions.

Methods: Included were RCTs with a comparison of (1) guided vs. unguided interventions, (2) different doses of guidance, (3) different qualification levels of e-coaches, and (4) synchronous vs. asynchronous communication mode. Outcomes were symptom severity, completer rates and number of completed intervention modules. A systematic search of MEDLINE, CENTRAL and PsycINFO, PsycARTICLES and Psyndex (search date 4th June 2013) was conducted, as well as a hand search of trial-registers and the reference lists of included articles. Methodological quality was rated using the Cochrane Risk of Bias tool. Relevant study characteristics and outcome data were extracted. Random-effects analyses were conducted if appropriate.

Results: 5328 articles were retrieved of which 14 fulfilled inclusion criteria. Guided interventions were significantly superior to unguided interventions ((symptom severity: *standardized mean difference* (SMD) $= -.27$ [*95% CI*: $-.45$; $-.10$]), n = 8; completed modules: $SMD = .52$ [.37; .67], n = 7; completer rate: $OR = 2.76$ [1.68; 4.53], n = 6). The four trials that examined different levels of e-coach qualification showed no significant differences on either of the outcome measures. Only one trial each examined the remaining two research questions, without significant effects on either of the outcome measures.

Conclusions: Guidance is a beneficial feature of Internet-based interventions, although its effect is smaller than reported before when compared to unguided interventions. The qualification of the e-coaches seems of minor importance. However, methodological limitations need to be considered when interpreting these findings. Overall, the number of studies was small and mainly limited to depression and social phobia restricting the generalizability of the findings.

1. Introduction

Several reviews indicate that Internet-based interventions (IBIs) are efficacious in treating mental disorders (Richards and Richardson, 2012; Lin et al., 2013; Andrews et al., 2010), however, they also report substantial heterogeneity of examined treatment effects across included studies. A review by Richards and Richardson (2012) on the efficacy of Computer-based psychological treatments for depression, for example, highlighted a standardized mean difference (SMD) regarding depressive symptoms of $g = -.56$ in favor of Computer-based interventions compared to treatment as usual or waitlist, with single trial results ranging from -1.42 to 0.03. To dismantle this heterogeneity and to examine the efficacious components of IBIs, research focuses on the mechanisms

underlying the efficacy of IBIs as well as possible predictors of therapeutic success or failure (Andersson et al., 2009; Nordgreen et al., 2012; Richards and Richardson, 2012). One of the core factors discussed in this context is guidance as part of IBIs. There are automated interventions independent of human support (self-guided or unguided interventions, e.g. Christensen et al., 2006) and interventions with some kind of human support (guided interventions, e. g. Nobis et al., 2013). Literature so far suggest that users benefit more from IBIs when guidance is provided (Andersson and Titov, 2014; Richards and Richardson, 2012; Johansson and Andersson, 2012). Beyond the dichotomy of unguided versus guided interventions, the efficacy of IBIs might further vary depending on the quantity (dose–response relationship) and quality of guidance (e.g. qualification of e-coaches providing guidance and communication mode used for guidance). Subgroup analyses conducted in the aforementioned review on depression (Richards and Richardson, 2012), suggested a hierarchy with therapist-supported interventions being most efficacious ($g = .78$), followed by interventions supported by non-clinical staff ($g = .58$) and unguided interventions ($g = .36$).

* Corresponding author at: University of Freiburg, Institute of Psychology, Department of Rehabilitation Psychology and Psychotherapy, Engelbergerstr. 41, D-79085 Freiburg, Germany. Tel.: +49 761 203 3044.
E-mail address: baumeister@psychologie.uni-freiburg.de (H. Baumeister).

Moreover, studies with asynchronously provided support (e.g. email contact; $g = .70$) showed a larger pooled *SMD* than studies with synchronous support (e.g. chat; $g = .28$, Richards and Richardson, 2012).

While these findings are important to better understand the underlying mechanisms of IBIs, they need to be interpreted cautiously given their explorative character comparing results across trials. Confounding variables such as technological developments over time (unguided interventions were more frequently conducted in the early years of Internet intervention research; Richards and Richardson, 2012) might partly explain the aforementioned differences. Titov and colleagues' trials on Internet-based social phobia interventions, for example, indicated that the efficacy of unguided interventions can substantially be increased when adherence facilitating components such as automated prompts are incorporated (Titov et al., 2008, 2009a).

To improve the validity of findings on the impact of guidance, it therefore seems important to focus on trials that experimentally examined the effects of guidance in randomized controlled clinical trials with a direct comparison of the aforementioned variations of guidance (i.e. unguided vs. guided; interventions with different doses of guidance; qualification of e-coaches; asynchronous vs. synchronous). The present systematic review extends the current state of evidence regarding these subjects by investigating the following four research questions:

1. Is there a difference in treatment outcome between guided and unguided interventions?
2. Is there a difference in treatment outcome depending on the dose of guidance?
3. Is there a difference in treatment outcome depending on the qualification of the e-coaches?
4. Is there a difference in treatment outcome between guided interventions with synchronous or asynchronous communication?

2. Material and methods

2.1. Inclusion criteria

Randomized controlled trials were included if they fulfilled the following criteria: 1) adult participants ($\geq$18 years), 2) with a mental disorder according to relevant classification systems (e.g. DSM-V or ICD-10) including subthreshold disorders as well as dimensionally measured mental disturbances of the respective disorder, 3) published in English or German, 4) comparing variations of an IBI with regard to (a) guided vs. unguided interventions, (b) at least two guided interventions with different guidance intensities, (c) at least two guided interventions with different levels of qualification of the e-coaches, or (d) at least two guided interventions using synchronous vs. asynchronous communication modes for guidance. 5) Trials had to report (a) symptom severity at the time of the follow up or (b) adherence to the program as outcomes. Symptom severity was operationalized by using the sum-score of a validated rating scale or self-report questionnaire for assessing the symptoms in question. Adherence was operationalized following Donkin et al. (2011) as a) the mean number of modules completed and b) the percentage of persons that completed the whole treatment.

2.2. Literature search and selection of studies

A systematic database search and additional hand search was conducted (compare PRISMA flow chart Moher et al., 2009, Fig. 1). Search strategies were developed and applied for MEDLINE, PsychINFO, PsychARTICLES and Psyndex (via EBSCO) and CENTRAL (via Wiley Online Library) (search date 4th June 2013) (see Appendix 1). All search strategies linked keyword-based and text-based searches. Hand search was conducted by searching the literature references of the included studies found through database search. We sent emails to the contact authors of included studies requesting further information on possible eligible

studies. Additionally, the clinical trial registers ClinicalTrials.gov and the German Clinical Trials Register (drks-neu.uniklinik-freiburg.de) were searched for eligible trials. In a two-step process titles and abstracts were screened for eligibility by one assessor (LR) (screening phase, n = 5328). All studies not excluded in step one were examined in detail on an abstract and full text basis by two assessors (HB, LR; n = 195).

2.3. Data extraction

Two assessors (HB, LR) extracted the following data from the included studies: basic sample characteristics (sample size, sex, age), information on how studies dealt with missing values, mental disorder identification, duration of the treatment in weeks, number of intervention modules and outcome measures. For the relevant trial groups, we extracted sample size, mean values, standard deviations and frequency of the respective outcome measures. Details of the pre- and post-treatment severity outcome data can be found in Appendix 2. Missing values were determined based on the reported data where feasible or requested from the respective primary author of included trials.

In case of multiple assessment instruments used for the assessment of an outcome, the data selection followed a hierarchical selection process favoring rating scales over self-report questionnaires. In case of multiple assessment instruments of the same hierarchical level, we randomly chose one assessment instrument for the meta-analysis, except for trials that compared unguided and guided interventions for social phobia. Here, all three studies (Berger et al., 2011a; Titov et al., 2008, 2009a) measured social phobia symptom severity by means of both the Social Phobia Scale (SPS) (Mattick and Clarke, 1998) and the Social Interaction Anxiety Scale (SIAS) (Mattick and Clarke, 1998). This allowed us to conduct a sensitivity analysis examining the robustness of the results by comparing the pooled standardized mean difference of two assessment instruments used for the same outcome. The SPS was randomly selected for the main analysis, while the sensitivity analysis was based on the SIAS.

2.4. Assessment of methodological quality

The methodological quality of the included studies was assessed using the Cochrane Risk of Bias tool (Higgins and Altman, 2008). It includes the categories "random sequence generation", "allocation concealment", "blinding", "incomplete outcome data", "selective outcome reporting" and "other sources of bias". Blinding was subdivided into "blinding of participants and staff", "blinding of outcome": a) symptom severity, b) completed modules, and c) completer rate. The included studies were ranked on a three-step scale ("low", "unclear" and "high") regarding the risk of possible bias.

2.5. Data analysis

Meta-analyses were conducted using Review Manager 5.2 (The Cochrane Collaboration, 2012). Standardized mean differences (*SMD*) with 95% confidence intervals (*CI*) were computed for all continuous outcomes. For dichotomous variables, odds ratios (*OR*) with *95%-CI* were computed. Random effects meta-analyses were performed to compute overall estimates of treatment outcomes. The effect sizes of the primary studies are presented in forest plots. Heterogeneity was examined with the I^2 statistic (Higgins and Thompson, 2002; Higgins et al., 2003). In the event of considerable heterogeneity ($I^2 > 75\%$), study results were not aggregated in meta-analyses. Following Sterne et al. (2011), publication bias was not examined by using a funnel plot due to the small amount of included studies.

For the comparison of unguided vs. guided interventions, results were analyzed for the three subgroups of trials that examined participant samples with depression (respectively depressive symptoms), social phobia or other mental disorders. For further comparisons subgroup analyses were not feasible due to the low number of primary trials per comparison.