Fachbereich Erziehungswissenschaft und Psychologie der Freien Universität Berlin

Mechanisms of Change in Internet-Based Interventions for Depression

Mechanismen der Veränderung in der Onlinebehandlung der Depression

Dissertation
zur Erlangung des akademischen Grades

Doktor der Philosophie (Dr. phil.)

vorgelegt von
Pavle Zagorscak, M.Sc.
Berlin, Januar 2019

Erstgutachterin: Prof. Dr. Christine Knaevelsrud

Zweitgutachterin: Prof. Dr. Babette Renneberg

Tag der Disputation: 29. April 2019

“You're wondering who I am:
Machine or mannequin?

[…]

I'm not a robot without emotions
I'm not what you see
I've come to help you
With your problems,
So we can be free.”

Styx, ‘Mr. Roboto’, 1983

Danksagungen

Während nach einem afrikanischen Sprichwort ein ganzes Dorf von Nöten ist, um ein Kind großzuziehen, scheinen es im Fall meines akademischen Erwachsenwerdens gleich zwei Arbeitsbereiche gewesen zu sein: Die „Klinisch-Psychologische Intervention“ unter Christine Knaevelsrud sowie die „Klinische Psychologie und Psychotherapie“ unter Babette Renneberg. Jede Mitarbeiterin und jeder Mitarbeiter dieser Teams hat in den vergangenen sechs Jahren auf die eine oder andere Art meinen Werdegang beeinflusst und dafür werde ich immer dankbar sein. Im Speziellen gilt mein Dank allen, die mit mir gemeinsam dem Großprojekt „TK-DepressionsCoach“ zum Erfolg verhalfen – als Forschende, Studierende, Beratende, Beratene oder Mitarbeitende von Vilua und der Techniker Krankenkasse.

Ganz besonders danke ich meiner Doktormutter Christine für die vielen Freiheiten und Entwicklungsmöglichkeiten, die ich als Doktorand unter ihrer Betreuung genoss und das unverwüstliche Vertrauen in mich und meine Fähigkeiten – gerade in den Augenblicken, in denen ich ihnen selbst nicht genug vertraute. Babette möchte ich dafür danken, dass sie immer eine offene Tür, ein offenes Ohr und einen guten Rat für mich und meine Anliegen hatte.

Daniel und Rebekka danke ich dafür, dass sie die ersten Schritte im „TK-DepressionsCoach“ mit mir gegangen sind. Christina, Johannes, Jana, Lars, Raphael, Sarah und den zwei Sebastians für Freundschaft, Kollegialität, die moralische und inhaltliche Unterstützung meiner Arbeit sowie das ein oder andere gehobene Glas nach Feierabend. Mike danke ich für den Humor im Rahmen seiner Englisch-Korrekturen, der mir die finale Abfassung der Arbeit erheiterte.

Nicht zuletzt möchte ich Stephan Heinzel, Nina Knoll und Sarah Schumacher für die Bereitschaft danken, neben meinen Gutachterinnen meine Promotionskomission zu besetzen und damit den letzten Schritt meiner Dissertation zu begleiten.

Es war nicht immer leicht, die vielen Aufgaben im „TK-DepressionsCoach“ zu bewältigen, weswegen ich ganz besonders den Menschen in meinem privaten Umfeld zu Dank verpflichtet bin: Meine „Kleeblätter“ in Hessen und Berlin und meine Familie waren immer für mich da, um die guten Tage zu feiern und die schlechten Tage zu betrauern. Danke für eure Liebe.

Schließlich danke ich Manuel für die letzten Jahre, für die Genauigkeit, die methodische Expertise, das Ertragen meiner Launen und am aller meisten für die Freundschaft. Es ist möglich, dass ich das alles ohne dich durchgestanden hätte. Aber es wäre nicht so schön gewesen.

Abstract

Unipolar depression ranks first on the World Health Organization's list of diseases responsible for global health burden. Different approaches of pharmacotherapy and face-to-face psychotherapy have been proven efficacious in reducing symptoms of depression and are available for individuals affected by the disorder. Yet, a majority of those individuals do not receive evidence-based treatment. Barriers such as difficulties attending appointments at inconvenient locations and at inconvenient times might be responsible for this shortfall in mental health care. Internet-based Interventions (IBI) show potential to overcome such barriers by offering psychoeducation, treatment tools, and (in some cases) guidance by clinicians independent of time and location. While several randomized controlled trials provide convincing evidence for the utility of IBI in reducing symptoms of depression, little is known about the mechanisms that account for in-treatment symptom change. Possible contributors to change within IBI for depression may be features of the intervention (e.g., therapeutic support by clinicians), variables related to treatment processes (e.g., treatment uptake, therapeutic alliance, outcome expectations) or variables specific to individuals (e.g., sociodemographic or socioeconomical characteristics).

This thesis presents the findings from three studies that seek to broaden our understanding of in-treatment symptom change in IBI. All three studies have been conducted with data collected from 1089 mildly to moderately depressed individuals drawn from the client-base of a public health insurance company. They were randomized to receive weekly feedback that was either fully-standardized or individualized by a counselor within an otherwise identical cognitive-behavioral IBI for depression.

STUDY 1 investigated differences between the treatment conditions concerning changes on clinical (depression, anxiety, perseverative thinking) and psychosocial variables (emotional self-efficacy, quality of life, and perceived social support). The between-condition effects were estimated after the intervention as well as 3, 6, and 12 months after participants finished the

program. Results revealed large within-group effects for depressive symptom reduction across conditions. However, between-group differences were nonsignificant for all outcomes at all measurement occasions.

STUDY 2 compared the contributions of outcome expectations, therapeutic alliance (agreement on tasks and goals; bond), extra-therapeutic stressors, and the uptake of specific treatment components to weekly symptom change in both treatment arms. Results showed that reductions of extra-therapeutic stressors during the intervention and high agreements on tasks and goals of treatment were associated with depressive symptom reductions in both conditions. While the level of extra-therapeutic stress at baseline was only predictive of in-treatment symptom change in the fully-standardized condition, bond ratings were associated with symptom deterioration during the final week of treatment in the individualized condition only.

STUDY 3 investigated whether distinct groups of participants experienced discernable symptom courses during the two treatment variants. In addition, we examined whether participants' psychosocial, socioeconomic or clinical characteristics were associated with membership in these groups. The results suggested that patterns of change and the associated groups did not differ across conditions. In both treatment variants most individuals (62.5%) were classified as "immediate improvers" with substantial improvement, commencing even before the start of treatment. Another class (37.4%) of individuals was labeled "delayed improvers" for their symptoms improved less overall and did not change up until week three of treatment. Individuals with higher perceived social support had higher odds of being classified as "immediate improvers". In contrast, individuals fulfilling the criteria for a current MDD in a structured clinical interview (SCID-I) and individuals with high outcome expectations had higher odds of being classified as "delayed improvers".

In summary, the results stressed the similarities between standardized and individualized feedback in IBI for depression concerning efficacy and patterns of change. At the same time the findings highlighted that individuals' socio-demographic and clinical features

influence both the mechanisms of change and symptom courses in IBI for depression, and they had differential effects depending on whether the feedback was standardized or individualized.

Abstract in deutscher Sprache

Unipolare Depression belegt weltweit den ersten Platz in der von der Weltgesundheitsorganisation herausgegebenen Liste der Erkrankungen, die mit der höchsten Krankheitslast einhergehen. Aktuell stehen Betroffenen verschiedene wirksame Behandlungsansätze der Pharmakotherapie und der Psychotherapie zur Verfügung. Nichtsdestotrotz nimmt nur ein Bruchteil dieser Menschen eine evidenz-basierte Therapie in Anspruch. Gründe für diese Unterversorgung liegen zum Teil darin, dass diese Angebote für manche Menschen örtlich schwierig zu erreichen sind oder die zu vereinbarenden Termine nicht mit ihrem sonstigen Lebensalltag zusammenpassen. Internet-basierte Interventionen (IBI) weisen das Potential auf, diese Barrieren zu überwinden, da sie zeitlich und örtlich flexibel Psychoedukation, therapeutische Werkzeuge und (in manchen Fällen) Begleitung durch Kliniker*innen bieten. Während zahlreiche randomisiert-kontrollierte Studien die Wirksamkeit von IBI in der Reduktion depressiver Symptome belegen, ist wenig über die Veränderungsmechanismen bekannt.

Dabei könnten sowohl Aspekte der Intervention selbst (z.B. die therapeutische Begleitung durch Kliniker*innen) als auch Indikatoren des Behandlungsprozesses (z.B. Ausmaß der Inanspruchnahme der Behandlung, Therapieallianz, Erfolgserwartungen) und individuumsspezifische Variablen (z.B. soziodemografische oder sozioökonomische Eigenschaften) einen Beitrag zur Erklärung der Symptomveränderung während der Behandlung leisten.

Die vorliegende Doktorarbeit stellt drei Studien vor, die Erkenntnisse zu diesem Thema beitragen. Alle drei Studien basieren auf einer Stichprobe von 1089 leicht- bis mittelgradig depressive belasteten Versicherten einer gesetzlichen Krankenkasse. Die Teilnehmenden wurden randomisiert einer von zwei Bedingungen zugeteilt: Im Rahmen einer ansonsten identischen kognitiv-behavioralen IBI für Depression, erhielten sie wöchentlich entweder

automatisches, voll-standardisiertes Feedback oder Feedback, welches durch ein*e Berater*in individualisiert wurde.

In STUDIE 1 wurden Unterschiede zwischen den beiden Experimentalgruppen in der Veränderung von klinischen (Depression, Angst, Grübeln) und psychosozialen Variablen (emotionale Selbstwirksamkeit, Lebensqualität, wahrgenommene soziale Unterstützung) untersucht. Die Zwischengruppeneffekte wurden nach der Intervention sowie 3, 6 und 12 Monate nach der Teilnahme evaluiert. Während große Innergruppeneffekte zeigten, dass die depressive Symptomatik in beiden Bedingungen signifikant zurückging, waren die Zwischengruppen-Effekte auf keinem der selbstberichteten Outcome-Maße und zu keinem der untersuchten Zeitpunkte signifikant.

In STUDIE 2 wurden Erfolgserwartungen, therapeutische Allianz (agreement on tasks and goals; bond), extra-therapeutische Stressoren sowie die Inanspruchnahme spezifischer Behandlungselemente hinsichtlich ihres Beitrags zur Erklärung von wöchentlichen Veränderungen depressiver Symptome in beiden Bedingungen untersucht. Die Ergebnisse zeigten, dass eine Verringerung extra-therapeutischer Stressoren und hohe task- und goal-Werte in beiden Bedingungen mit einer Verringerung der Symptombelastung einhergingen. Während die Ausgangswerte des extra-therapeutischen Stresses nur in der voll-standardisierten Bedingung mit Symptomverschlechterung zusammenhingen, waren hohe bond-ratings nur in der individualisierten Variante mit einer Symptomverschlechterung in der letzten Behandlungswoche assoziiert.

In STUDIE 3 wurde in beiden Bedingungen untersucht, ob es verschiedene Teilnehmendengruppen gibt, die distinkte Muster der Symptomveränderung aufweisen. Zusätzlich wurde geprüft ob psychosoziale, klinische oder sozioökonomische Eigenschaften der Teilnehmenden eine Zugehörigkeit zu diesen Gruppen vorhersagen. Die Ergebnisse zeigten, dass sich die Veränderungsmuster und die zugehörigen Gruppen nicht zwischen den Bedingungen unterschieden. In beiden Varianten wiesen die meisten Teilnehmenden (62.5%)

ein Symptommuster über die Zeit auf, welches als „unmittelbare Verbesserung“ charakterisiert werden konnte. Diese Gruppe zeichnete sich durch substantielle Symptomreduktion über die Zeit aus, welche bereits vor Behandlungsbeginn ihren Anfang nahm. Die Symptomentwicklung der zweiten Gruppe (37.4%) kann als „verzögerte Verbesserung“ beschrieben werden. Hier stagnierten die Symptome bis zur dritten Behandlungswoche und es war insgesamt eine geringere Veränderung über die Dauer der Behandlung zu verzeichnen. Hohe wahrgenommene soziale Unterstützung sagte die Mitgliedschaft in der Gruppe mit „unmittelbarer Verbesserung“ vorher, während Teilnehmende mit einer aktuellen depressiven Episode (nach strukturiertem klinischen Interview, SKID-I) sowie Individuen mit hohen Erfolgserwartungen mit höherer Wahrscheinlichkeit der Gruppe mit „verzögerter Verbesserung“ angehörten.

Zusammenfassend unterstreichen die Ergebnisse die Ähnlichkeiten zwischen IBI mit individualisiertem und standardisiertem Feedback sowohl hinsichtlich der Effektivität als auch hinsichtlich resultierender Muster der Symptomveränderung. Zugleich weisen die Befunde darauf hin, dass die soziodemografischen und klinischen Charakteristika der Teilnehmenden – je nach Experimentalbedingung unterschiedlichen – Einfluss auf den Veränderungsprozess haben.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 1
THEORETICAL BACKGROUND

1.1 Introduction

The global spread of internet technology has fundamentally changed the quantity and quality of available information as well as the communication between individuals. A recent representative study showed that about 90% of the entire German population use the internet (Koch & Frees, 2017). The average German adult spends up to 4.4 hours online per day; that time is spent interacting with others, viewing, sharing and creating content, seeking entertainment or purchasing goods and services (Ernst & Young GmbH, 2017). As internet technologies have impacted virtually every area of life, they have also changed individual approaches to communicating about mental health problems and obtaining information about them (Birnbaum, Rizvi, Correll, Kane, & Confino, 2017; Naslund, Grande, Aschbrenner, & Elwyn, 2014).

For example, an analysis of data from *Google Trends* (2018) shows that depression is the category of mental disorder[1] with the highest search frequency in Germany. As depicted in FIGURE 1.1, the popularity of the search term "depression" gradually increased during the last 10 years, with an outlier peek in November 2009 coinciding with the suicide of popular soccer player Robert Enke. Importantly, individuals do not only use the internet to find information on this mental health topic but a growing number of individuals participate in treatments for their depressive symptoms that are offered partly or exclusively through the internet (Drozd et al., 2016). This trend is further mirrored by the number of studies per year on these so called "Internet-based interventions" (abbreviated: IBI) for depression, which has steadily increased for more than ten years (Drozd et al., 2016).

[1] Even though the term "depression" commonly refers to Major Depressive Disorder, it is in fact an umbrella term, encompassing multiple distinct disorder categories. Please see CHAPTER 1.2 for definitions and distinctions.

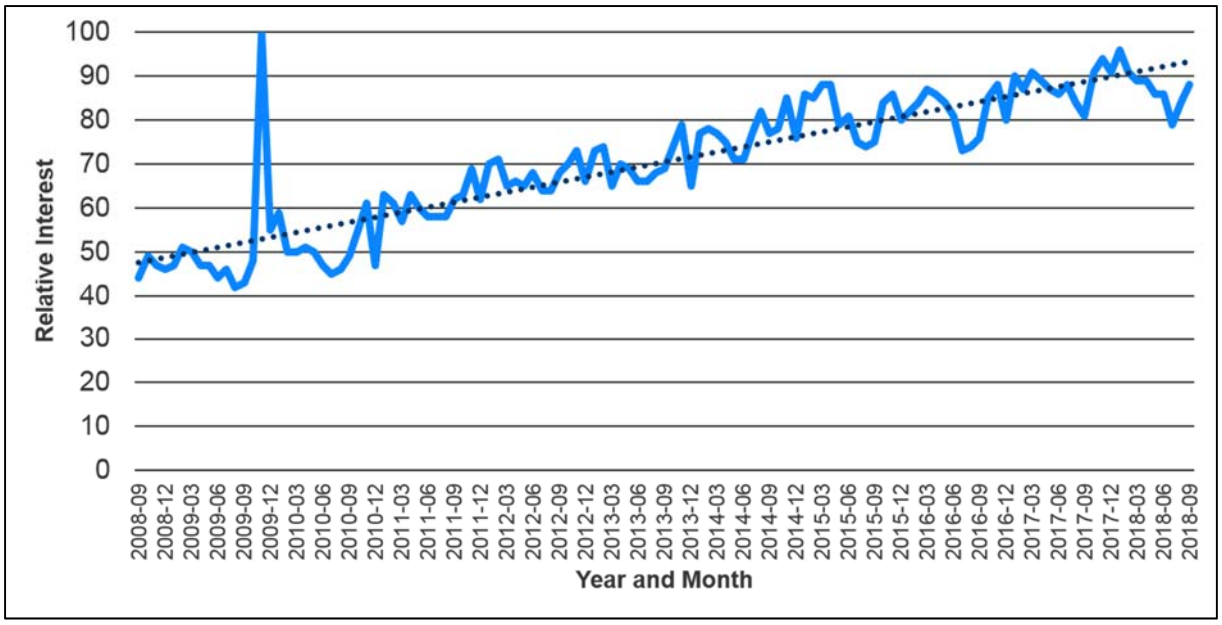


FIGURE 1.1 Monthly relative interest in the search term "depression" in Germany from September 2008 to September 2018 with linear trend line (dotted). Data is generated by calculating the share the search term "depression" has in all Google searches performed in Germany in each month. Data is indexed to 100 (highest value in November 2009 is set as 100, other values are calculated relative to that). Source: Publicly available *Google Trends* data set, own analysis.

While several randomized controlled trials found convincing evidence for the utility of IBI in reducing symptoms of depression (e.g., Karyotaki et al., 2018, 2017), little is known about the mechanisms that account for in-treatment symptom change in IBI with different levels of therapeutic support. The thesis at hand seeks to add knowledge on this important topic by presenting the results of three studies on interventional and individual predictors of change in IBI for depression.

The following chapters contextualize the development of these studies by defining depressive disorders and quantifying their prevalence (CHAPTERS 1.2 – 1.3). Additionally, face-to-face treatments for depression and their limitations are briefly described (CHAPTER 1.4). Next, IBI will be defined and introduced from a theoretical and empirical perspective (CHAPTER 1.5). Finally, I will discuss shortcomings of existing trials on IBI for depression (CHAPTER 1.6) and provide a short overview of the three studies presented in this thesis (CHAPTER 1.7).

1.2 Definitions of Depression

According to the current editions of the *International Classification of Diseases for Mortality and Morbidity Statistics* (ICD-11; WHO, 2018) and the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; APA, 2013) depression in adults is characterized by depressed mood and/or anhedonia or loss of interest. These core symptoms are frequently accompanied by other symptoms such as fatigue, difficulties concentrating, indecisiveness, lessened feeling of self-worth, feelings of guilt, thoughts about death or suicide, changes in appetite, sleep patterns or motor activity (APA, American Psychiatric Association, 2013). Depending on the combination, duration, frequency, intensity and assumed etiology of these symptoms, DSM-5 subsumes different disorder categories under the umbrella term "depression"[2] or "unipolar depression" (as opposed to bipolar disorders). These diagnostic categories are summarized in TABLE 1.1.

[2] These categories deviate from the ICD-11, which uses the umbrella term "mood disorders" (F30-F39) and includes manic episodes and bipolar disorders as well as mixed episodes of anxiety and depression. Further, Premenstrual Dysphoric Disorder is sorted into another category (Diseases of the female genital system/ Premenstrual Disturbances). DSM-5 categorization was preferred due to the clearer focus on (unipolar) depressed mood.

TABLE 1.1 (CONT. ON FOLLOWING PAGE)
Different diagnostic categories of depressive disorder in adults according to DSM-5

Common features	Changes in mood accompanied by somatic changes • **with** clinically significant distress or impairment in social, occupational or other important areas of functioning • **without** hypomanic or manic episodes (lifetime) • **not better explained as** any psychotic disorder or schizoaffective disorder			
Assumptions about etiology	*mainly not attributable* to periodical hormonal changes, substance use or medical conditions		mainly attributable to *periodic hormonal changes*	mainly attributable to *substance/medication use or* another *medical condition*
Diagnoses	**Major Depressive Disorder (MDD; recurrent or single episodes)**	**Persistent Depressive Disorder (Dysthymia)**	**Premenstrual Dysphoric Disorder**	**Substance/Medication-Induced Depressive Disorder** *or* **Depressive Disorder Due to Another Medical Condition**
Associated Symptoms				
I. Depressed Mood	Yes	Yes	Yes	Yes
II. Anhedonia or Diminished Interest in Activities	Yes	No	Yes	Yes
III. Increase/Decrease in Appetite	Yes or: significant weight change	Yes	Yes or: specific food craving	No
IV. Insomnia or Hypersomnia	Yes	Yes	Yes	No
V. Psychomotor Agitation or Retardation	Yes	No	No	No
VI. Decrease in Energy or Fatigue	Yes i.e., loss of energy	Yes, i.e., low energy	Yes i.e., lethargy	No
VII. Reduced Self-Esteem or Feelings of guilt	Yes i.e., feeling worthless or guilty	Yes i.e., low self-esteem, excl. guilt.	No	No

Associated Symptoms	**Major Depressive Disorder (MDD; recurrent or single episodes)**	**Persistent Depressive Disorder (Dysthymia)**	**Premenstrual Dysphoric Disorder**	**Substance/Medication-Induced Depressive Disorder** *or* **Depressive Disorder Due to Another Medical Condition**
VIII. Reduced Ability to concentrate, think or indicisiveness	Yes	Yes	Yes i.e., subjective difficulty in concentrating	No
IX. Thoughts of death, suicidal ideation, plans or behavior	Yes	No	No	No
X. Additional/Other Symptoms	No	1) Hopelessness	1) Affective lability, 2) Irritability/Anger, 3) Anxiety/Tension 4) feeling overwhelmed 5) physical symptoms (e.g., breast tenderness, pain)	No
Time criterion for symptom occurrence	present for *nearly every day*, nearly entire day for at *least two weeks*	present more than *half of the days* for more than *two years* with no symptom-free periods >2 months	present in *most final weeks before menses*; improving within few days after menses, become minimal/absent in the week postmenses	present *during or soon after* use of, withdrawal from or intoxication with *substance/medication* *or* symptom is the *direct pathophysiological consequence* of another *medical condition*
Required Number of Symptoms	5, including at least one core symptom (I. or II.)	3, including symptom I. Symptom severity may be milder than in MDD	5, specific combinations required	Not specified

*Note***.** Criteria derived from DSM-5 (APA, 2014). "Other Specified Depressive Disorder" and "Unspecified Depressive Disorder" are excluded from this summary due to the lack of specific symptoms, time frames or etiological assumptions.

1.3 The Relevance of Depression for Global Mental Health

The chapter at hand offers an overview of the prevalence of Major Depressive Disorder (MDD) and Persistent Depressive Disorder (Dysthymia)[3], highlighting their individual and societal relevance. Further, findings on epidemiology and comorbidity of unipolar depression will be briefly summarized in the current chapter as they are relevant for discussions about treatment and treatment uptake provided in CHAPTER 1.4.

Due to the focus of this doctoral thesis, affective disorders that are attributable to periodical hormonal changes, substance use or other medical factors will not be reviewed further. Comprehensive summaries of research on these disorders can be found elsewhere (e.g., Cosci, Fava, & Sonino, 2015; Hantsoo & Epperson, 2015; Tolliver & Anton, 2015).

1.3.1 Prevalence

With lifetime prevalence rates of 11-21% Major Depressive Disorder is the most common mental disorder internationally[4] (Kessler et al., 2005; Kessler, Petukhova, Sampson, Zaslavsky, & Wittchen, 2012; Kessler & Bromet, 2013). While lifetime prevalence rates for dysthymia are reported to be much lower (1-3%), the induced amount of suffering and loss of productivity is considerable, given the persistence of the disorder (Blanco et al., 2010; Ferrari et al., 2013). According to the Global Burden of Disease study (WHO, 2017), at any given time (point prevalence), about 2.4% of the German adult population is suffering from unipolar depression (MDD: 1.8%; dysthymia: 0.7%). Similar prevalence rates are reported worldwide, with slightly higher rates in high-income countries and the "Americas" (Ferrari et al., 2013; Thornicroft et al., 2017; WHO, 2017). Given these prevalence rates, unipolar depression was

[3] The DSM-IV categories of "chronic major depressive disorder" and "dysthymia/dysthymic disorder" were merged into the category "Persistent Depressive Disorder" in DSM-5. Due to the relative novelty of DSM-5, the majority of studies cited as well as the three studies presented in this dissertation follow DSM-IV nomenclature.
[4] This holds true on the level of the specific diagnosis. As a group, anxiety disorders are more common than mood disorders (Kessler et al., 2005; 2012).

responsible for over 54 million disability-adjusted life years (DALYs)[5] making it the leading cause of disability worldwide (WHO, 2017).

1.3.2 Epidemiology

The occurrence of unipolar depression is not distributed equally across the entire population. Significantly increased odds of being diagnosed with MDD or Dysthymia are reported for women, individuals over the age of 60, single/divorced or widowed individuals, individuals living under economically disadvantaged circumstances and individuals with lower levels of formal education (Arias-de la Torre, Vilagut, Martín, Molina, & Alonso, 2018; Salk, Hyde, & Abramson, 2017; Vandeleur et al., 2017). Individuals belonging to a social, ethnic or sexual minority[6] also show a heightened risk for developing MDD or Dysthymia (e.g., Lewis et al., 2017; Lucassen, Stasiak, Samra, Frampton, & Merry, 2017; Schouler-Ocak, Aichberger, Penka, Kluge, & Heinz, 2015).

1.3.3 Comorbidity

The importance and impact of unipolar depression is further underlined by high comorbidity rates. Most individuals who fulfill the criteria for MDD or dysthymia also fulfill the criteria for at least one other mental disorder: In a nationally representative sample in Germany, Jacobi et al. (2014) found that only 32.6% (95% CI: 26.9 - 38.9) of individuals fulfilling the diagnostic criteria for dysthymia or MDD during the last 12 months had *no* comorbid mental disorder diagnosis during the same time period. In the presence of unipolar depression, 21.5% (95% CI: 16.5 – 27.4) presented with one additional diagnosis, 16% (95% CI: 11.4 - 22.0) with two, and 29.9% (95% CI: 23.5 - 37.2) with three or more. Essentially, depressed individuals with comorbid mental disorders show significantly increased impairments and morbidity when compared to individuals who have only one mental disorder

[5] DALYs are the sum of Years of Life Lost (YLL) (i.e., average life expectancy reduced by the age of premature death/suicide associated with a disorder/disease) and Years Lived with Disability (i.e., the prevalence of a disorder multiplied by a factor representative of the short- and long-term loss of health associated with that disorder).
[6] i.e., individuals identifying as non-cisgendered or non-heterosexual, people of color, individuals who are defined as migrants, individuals who are physically handicapped.

(Laursen, Musliner, Benros, Vestergaard, & Munk-Olsen, 2016). Essentially, depressed individuals with comorbid mental disorders show significantly increased impairment and morbidity when compared to individuals who have only report one mental disorder (Laursen et al., 2016).

Notably, not all mental disorders co-occur with unipolar depression to the same degree. A WHO study across ten countries (de Jonge et al., 2018) estimated correlations between lifetime diagnoses of mental disorders. Regarding diagnoses of unipolar depression (Dysthymia and MDD) the authors found the highest correlations with other internalizing disorders such as generalized anxiety disorder (r =.61 -.65), PTSD (r =.46 -.56), social anxiety disorder (r =.49 -.50) and agoraphobia (r =.44 -.46). However, correlations with externalizing disorder diagnoses were considerable as well (e.g., Attention Deficit Hyperactivity Disorder: r =.39 -.42; Oppositional Defiant and Conduct Disorder r = 0.34 - 0.36; Substance Use Disorder: r =.27 -.29). A meta-analysis by Friborg (2014) on patients receiving inpatient or outpatient services further demonstrated high comorbidities between MDD or dysthymia and personality disorders. On average, about 40% of individuals with MDD also fulfilled the criteria for any personality disorder. The rates were higher for individuals diagnosed with dysthymia (60%). The authors summarized that for the group of patients with unipolar depression, personality disorder comorbidity "was lowest in cluster A, higher in cluster B, but highest in cluster C"[7] (Friborg et al., 2014, p. 7).

In summary, this chapter demonstrates that depression is frequent and the amount of the associated individual and societal detriments is further heightened by high comorbidity-rates. At the same time, prevalence of unipolar depression is especially pronounced in certain subsets of the population. Overall, this suggests a high demand for evidence-based treatments

[7] The clusters of personality disorders according to DSM-5 (APA, 2013) are Cluster A (odd or eccentric) encompassing paranoid, schizoid, and schizotypical personality disorder; Cluster B (dramatic, emotional or erratic) encompassing antisocial, borderline, histrionic, and narcissistic personality disorder; Cluster C (anxious or fearful) encompassing avoidant, dependent, and obsessive-compulsive personality disorder.

such as face-to-face psychotherapy and pharmacotherapy (i.e., “conventional treatment approaches”). The following chapter will briefly introduce studies on the efficacy of these two approaches and will especially focus on barriers to their uptake.

1.4 Barriers to the Uptake of Conventional Evidence-Based Treatments for Depression

Results of multiple meta-analyses including hundreds of RCTs support the notion that both pharmacotherapy and different forms of face-to-face psychotherapy (e.g., cognitive-behavioral therapy, psychodynamic therapy) are efficacious treatments for unipolar depression (e.g., Barth et al., 2016; Cipriani et al., 2018; Kamenov, Twomey, Cabello, Prina, & Ayuso-Mateos, 2017). Contrasted against passive control groups, both treatment approaches yield medium to large effect sizes[8] on the primary outcome of depressive symptom reduction (Barth et al., 2016; Cipriani et al., 2018). Accordingly, these approaches are considered to be the current "best practice" and have been included in national guidelines on treating depression (e.g., DGPPN, 2017).

Pharmacotherapy and face-to-face psychotherapy are both available (to a varying degree) in middle- and high-income countries. However, a recent meta-analysis, covering studies from 21 countries, concluded that only 16.5% of all individuals that fulfilled the criteria for MDD during the last 12 months received minimally adequate, albeit not necessarily evidence-based treatment[9] (Thornicroft et al., 2017). Similar rates[10] are reported in a German nationally representative study for individuals suffering from any mental disorder (Jacobi et al., 2014). Notably, minority groups as well as socially or economically disadvantaged populations show proportionally lower rates of treatment uptake, despite their heightened rates of depressive symptoms (e.g., Miranda, Soffer, Polanco-Roman, Wheeler, & Moore, 2015; Steele et al., 2016; Thornicroft et al., 2017). This raises questions on the reasons for this treatment gap and possible ways to circumvent it.

[8] Throughout this dissertation effect sizes will be evaluated based on the conventions proposed by Cohen (1992)
[9] Thornicroft et al. (2017) defined minimally adequate treatment as at least one month of medication, plus at least four visits to any type of medical doctor *or* at least eight visits with any professional including religious or spiritual advisor, social worker or counsellor, which seems a low threshold that does not necessarily rise to the level of being considered "evidence-based". Stricter criteria will probably result in substantially lower rates.
[10] Jacobi et al. (2014) report that 11% of individuals with one disorder diagnosis during the last year had "any contact to the health care system" in the same period, whereas the rates increase with the number of comorbid diagnoses (up to 40% of individuals with four or more comorbid disorder diagnoses had any contact with the health care system).

According to Thornicroft et al. (2017) there are three stages towards receiving adequate treatment for depression and each of these stages presents individuals with specific barriers (e.g., Harvey & Gumport, 2015). The stages and barriers are collectively illustrated in FIGURE 1.2.

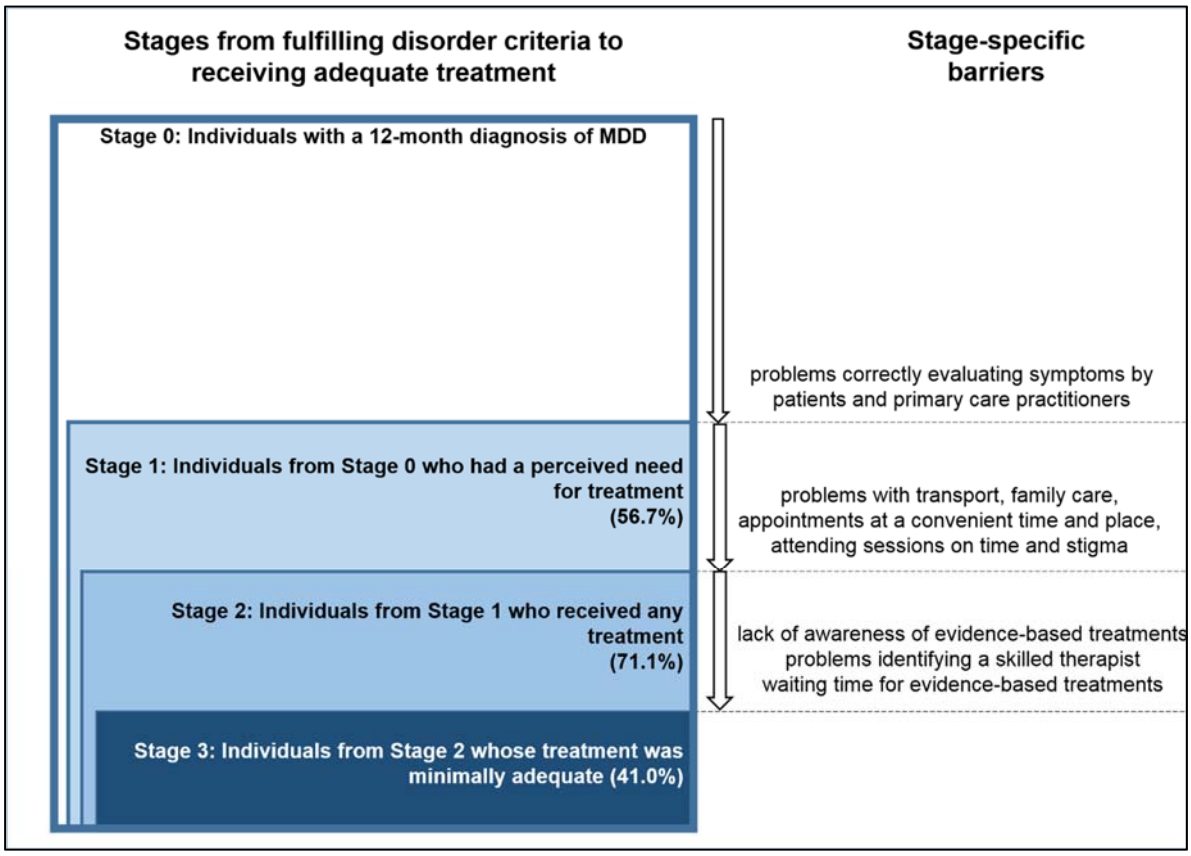


FIGURE 1.2 Stages from fulfilling disorder criteria to receiving adequate treatment and associated barriers. Rates/shares for each stage obtained from Thornicroft et al. (2017); barriers are summarized on the basis of the review by Harvey and Gumport (2015). Sizes of boxes are approximations of the actual percentages.

Stage 1: Perceiving a need for treatment. Individuals may have difficulties in correctly evaluating their own symptoms and choosing appropriate reactions due to a lack of mental health literacy (Bonabi et al., 2016). The problem is exacerbated by detection problems in primary care where about half of the individuals fulfilling the criteria for MDD fail to be diagnosed by primary care practitioners (Beesdo-Baum et al., 2018). It is estimated that only 56% of individuals who fulfill the criteria for MDD perceive a need for (depression) treatment at all (Thornicroft et al., 2017).

Stage 2. Receiving any kind of treatment. Even if depressed individuals feel that they need support with their symptoms, lack of internal and external resources inhibits an ultimate decision to (regularly) attend treatment. Overall, about 71% of individuals that perceive a treatment need also obtain *some* kind of treatment (Thornicroft et al., 2017). In a narrative review on barriers to treatment uptake, Harvey and Gumport (2015) summarized that the central problems on the patient level pertain to logistic issues such as "transport, childcare, appointments at a convenient time and place, identifying a skilled therapist, attending sessions on time and overcoming stigma" (Harvey & Gumport, 2015, p. 41). A detailed look at the types of barriers at this stage underlines that they are likely associated with socioeconomic and psychosocial disadvantages (e.g., dependence on public transport, not being able reconcile regular appointments with a precarious employment or with familial obligations). Indeed, empirical studies show that membership in socially marginalized groups, which might need treatment the most, reduces the likelihood of receiving treatment for mental disorders (Conner et al., 2010; Gulliver, Griffiths, & Christensen, 2010).

Stage 3. Receiving evidence-based treatment. More than a third of patients who start the uptake of psychotherapy discontinue their treatment within or before the first five preparatory sessions (Jacobi, Uhmann, & Hoyer, 2011), partly due to the barriers listed in Stage 2. Additionally, about 10 to 30% of individuals with depression seek complementary or alternative medicine (CAM) instead of evidence-based interventions (Solomon & Adams, 2015). Reasons include a lack of knowledge on what might constitute "evidence-based" treatments and the importance of treatment adherence as well as a reduced feeling of stigma derived from the uptake of CAM as opposed to psychotherapy (Hansen & Kristoffersen, 2016). Another important reason lies on a structural level and pertains to limited access to adequate treatment, especially in rural areas. Data from Germany indicates that psychological psychotherapists or specialized medical staff are scarce and capacity restrictions result in multiple months of waiting times. For example, the average patient in Brandenburg has to wait

for 29.4 weeks (German average: 19.9) for the start of guideline-based psychotherapy (BPtK, 2018). Taken together, about 59% of patients receiving “any kind of treatment” either do not receive a minimally adequate dosage of evidence-based treatments or they instead use CAM.

1.4.2 Circumventing barriers

Some of the stage-specific barriers can be targeted by awareness campaigns and measures that increase mental health literacy or diagnostic competencies in the general population (e.g., Chang, 2008), in specific target groups (e.g., Kutcher, Bagnell, & Wei, 2015) or in primary care practitioners (e.g., Gilbody, Whitty, Grimshaw, & Thomas, 2003). These measures pertain to stages 1 and 3 where knowledge about depression and evidence-based treatments are key issues. However, barriers on stage 2 (receiving any kind of treatment) are ingrained in the nature of face-to-face treatments for they are related to the requirement of co-presence of patient and therapist at a certain place (i.e., the treatment facility) and at a certain time (i.e., appointments).

Thus, these barriers might be effectively reduced by a different approach to treatment such as internet-based interventions.

1.5 Internet-based Interventions

1.5.1 Defining Internet-based Interventions

The umbrella term "Internet-based Interventions" (IBI) subsumes the use of internet-ready devices (e.g. personal computers, smartphones, tablets) for accompaniment and provision of (psycho)therapeutic measures. In that, the medium can be used to 1) provide information to patients, 2) offer training, tools, and tasks for patients, *and/or* 3) communicate between patients and therapists[11] (Zagorscak & Knaevelsrud, 2019).

Berger, Stolz and Schulz (2013) further clarify that depending on the function of the medium different approaches of IBI exist (see FIGURE 1.3 for a summary):

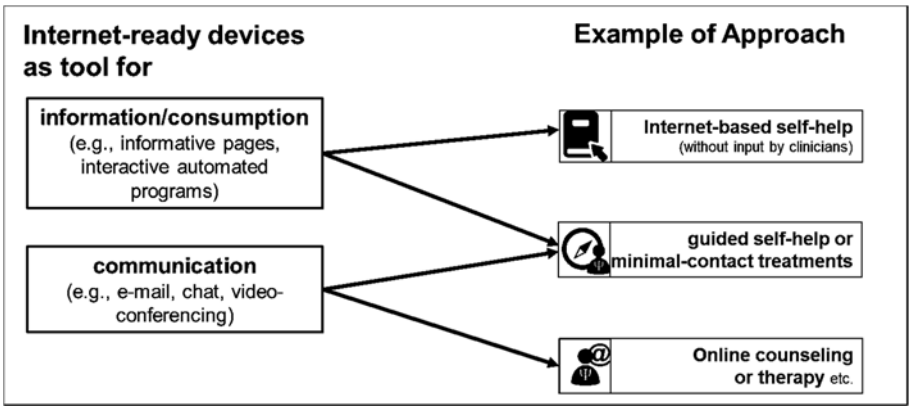


FIGURE 1.3. Functions of internet-ready devices and resulting approaches of internet-based interventions. Adapted from Berger, Stolz & Schulz (2013).

Focus on information. If internet-ready devices are used solely as a tool to provide or consume structured and comprehensive information (e.g., about a psychological problem or disorder, psychoeducation) and/or to offer automated (treatment) tools (e.g. for tracking the symptoms or protocoling thoughts) to patients, the approach is best summarized as "internet-based self-help". In that case, no "guidance" or contact to other humans (e.g., therapists) is

[11] Please note that for reasons of readability only the terms "patient" and "therapist" are used. Depending on the setting, symptom severity and qualification of those providing guidance, wording such as "client" and "counselor" or "provider" and "user/recipient" might be more appropriate.

offered. Thus, these approaches are also referred to as “unguided” IBI (e.g., Wijnen, Lokman, Leone, Evers, & Smit, 2018).

Focus on communication. If internet-ready devices are used solely as a communication tool between patients and therapists (e.g., using text, audio- or video-messaging tools) and no standardized or automated online-information or tools are offered, the approach is best summarized as “online counseling” or “online (psycho)therapy” due to the structural similarities with the respective face-to-face approaches.

Hybrid forms. If internet-ready devices are used to provide (standardized or automated) information and/or treatment tools and at the same time feedback or guidance is provided online, this approach is referred to as “guided self-help” or “guided IBI”. Usually, these approaches offer fully-standardized or automatically tailored (e.g., differing information for men and women) psychoeducation and tasks that can be completed online. Afterwards, a therapist assesses the work product of the patient and gives feedback or encouragement. This feedback is usually offered in the form of written messages but might also consist of accompanying phone calls, chats or video-conferencing (Richards & Richardson, 2012). Some of these approaches offer only very little guidance (e.g., a few semi-standardized lines of written encouragement weekly) and are thus called “minimal-contact treatments” (e.g., Lappalainen, Langrial, Oinas-Kukkonen, Tolvanen, & Lappalainen, 2015).

Blended treatment. Each of these three basic categories of IBI can be combined with face-to-face therapy. These combinations are called “blended treatment”. It may be considered as a separate category of IBI. However, blended treatments can usually be decomposed into an IBI facet and a face-to-face facet. Concordantly, most trials on blended treatments have investigated rather separate than complementary forms of IBI and face-to-face psychotherapy (Wentzel, van der Vaart, Bohlmeijer, & van Gemert-Pijnen, 2016).

1.5.2 Efficacy of IBI for depression

Dozens of randomized controlled trials have investigated whether IBI are more effective than passive control groups (usually waitlists). These studies have been summarized in four recent meta-analyses (Karyotaki et al., 2018, 2017; Königbauer, Letsch, Doebler, Ebert, & Baumeister, 2017; Sztein, Koransky, Fegan, & Himelhoch, 2018) and a number of older reviews and meta-analyses (e.g., Andersson & Cuijpers, 2009; Bhattacharya, Kelley, & Bhattacharjee, 2012; Johansson & Andersson, 2012; Richards & Richardson, 2012; So et al., 2013). In short, all meta-analyses concluded on the efficacy of IBI for depression. The reported standardized controlled effect sizes obtained from the four most recent meta-analyses are high for guided interventions and small for unguided interventions (see TABLE 1.2). Moreover, there have been first randomized controlled trials comparing IBI with similarly structured face-to-face treatments (i.e., same duration, same topics addressed, similar number of patient-therapist contacts), which have also been synthesized in a recently updated meta-analysis (Carlbring, Andersson, Cuijpers, Riper, & Hedman-Lagerlöf, 2018). The authors did not find any meaningful differences in symptom reductions between IBI and face-to-face treatments. While these results are encouraging for the utility of IBI, it is important to stress their preliminary nature. The meta-analysis included 20 trials on various somatic and psychiatric conditions. Thus, the number of included trials on IBI for depression ($k = 4$) is still too small to conclude evaluation on the equivalence of both approaches for this specific target population.

TABLE 1.2
Results of recent meta-analyses on the efficacy of internet-based interventions for depression.

Meta-Analysis	**Experimental IBI Conditions (and target populations)**	**Comparator Condition**	**Included Studies (*k*) and Participants (*n*)**	**Controlled Effect Size (Symptom Reduction) *or* Comparative Odds of Remission Pre-Post**	**Controlled Effect Sizes Follow-Up**
Königbauer et al., 2017	guided and unguided (depressed individuals; clinical interview)	WLC	*k* = 19 *n* = 1650	SMD = 0.90 (95% CI: 0.73 - 1.07)	---
Karyotaki et al., 2017	Unguided (depressed individuals; clinical interview or self-report)	WLC, CAU, AC	*k* = 13 *n* = 3876	SMD = 0.27	---
Karyotaki et al., 2018	Guided (depressed individuals; clinical interview or self-report)	WLC, CAU, AC	*k* = 24 *n* = 4889	OR = 2.41 (95% CI: 2.07–2.79)	---
Sztein et al., 2018	CBT only; guided and unguided (mildly to moderately depressed individuals; self-report)	WLC	*k* = 14 *n* = 1631	SMD = 0.74 (95% CI: 0.62 - 0.86)	3-6 months follow up: SMD = 0.83 (95% CI: 0.69 - 0.99)

Note. WLC = waitlist control, CAU = care as usual, AC = attention placebo control, CBT = cognitive-behavioral therapy, SMD = standardized mean difference (difference in mean outcome between groups divided by the standard deviation of outcome among participants), OR = odds ratio.

1.5.3 Benefits and challenges

Apart from the promising results on their efficacy, IBI offer several benefits that might help circumvent the barriers to treatment uptake listed in SECTION 1.4.2. These benefits are briefly summarized in this section and are discussed in more detail elsewhere (e.g., Andersson & Titov, 2014; Zagorscak & Knaevelsrud, 2019).

First, IBI can be offered independently of the location of a user (i.e., geographical flexibility) and with an increased flexibility regarding the duration, frequency and time they are used (i.e., temporal flexibility)[12]

Second, the availability of internet-ready devices in everyday situations facilitates the convenient use of therapeutic tools and information where they are needed (e.g., Beattie, Shaw, Kaur, & Kessler, 2009). Additionally, in most forms of IBI all information that is exchanged between patient and therapist is recorded, safely stored and retrievable on the patient's demand.

Third, the perceived risk of being stigmatized through the uptake of IBI is lower than for face-to-face psychotherapy (Bathje, Kim, Rau, Bassiouny, & Kim, 2014; Griffiths, Lindenmeyer, Powell, Lowe, & Thorogood, 2006; Wong, Bonn, Tam, & Wong, 2018).

Fourth, as treatment capacity is a limited and costly resource, IBI bear the potential to offer more efficient treatment to patients: In stark contrast to the time investment of face-to-face therapy, which is usually 50 minutes *per session*, a meta-analysis reported guidance time in IBI for depression and anxiety to be only between 53 and 150 minutes *per entire intervention* (Pihlaja et al., 2018).

Fifth, a meta-analysis summarized that IBI have a beneficial effect on empowerment and self-efficacy (Samoocha, Bruinvels, Elbers, Anema, & van der Beek, 2010) and patients perceived empowerment as a unique benefit of IBI in qualitative studies (Wallin, Mattsson, & Olsson, 2016).

In summary, the list of assets of IBI highlights the potential to circumvent barriers to the uptake of treatment for mental disorders. At the same time, the use of digital communication bears a number of challenges that patients and therapists are faced with.

[12] While the amount of geographical flexibility is high in all IBI, temporal flexibility is the highest in internet-based self-help and lower in guided self-help, where a patient might have to wait for her therapist's feedback in order to proceed. Online-counseling or therapy that seeks to duplicate the face-to-face setting suffers from the same temporal restrictions as a consequence (e.g., temporal co-presence might be required for video-conferencing).

Zagorscak and Knaevelsrud (2019) argued that in order for IBI to be successfully applied, there are challenges on three levels. These challenges are described briefly here, and discussed in more detail in the original publication.

Intervention level. A lack of physical co-presence of patient and therapist and (depending on the form of IBI, see SECTION 1.5.1) the absence of nonverbal communication cues increases the probability of misunderstandings and decreases the probability of noticing emotional nuances (Beattie et al., 2009). Further, the mode of communication might impede diagnosis in general and in particular the detection of a (suicidal) crisis (e.g., from written messages alone). This critique has been contested in studies demonstrating that diagnostic questionnaires and interviews lead to comparable conclusions regardless of whether they are offered online, by telephone or face-to-face (e.g., Fine et al., 2013; Hajebi et al., 2012; Hines, Douglas, & Mahmood, 2010; Vallejo, Jordán, Díaz, Comeche, & Ortega, 2007). However, the precondition is that reliable and valid diagnostic instruments are *indeed* used regularly and repeatedly. Consequently, implementing frequent screenings for disorder symptoms and suicidality and predetermined procedures for patients in crisis or experiencing symptom deterioration are important criteria for assessing the quality of existing IBI (Klein et al., 2018).

Technical level. To participate in IBI, both the patient and the therapist (if involved) need to have a stable internet connection and internet-ready devices that are suitable to the purpose at hand (e.g., recording and playing audio, installing certain software). In addition, the data and the communication of patients (with therapists) need to be secured. Typically, this entails encrypted messages stored on servers devised for that purpose, password-protected access or two-factor authentication (i.e., transaction authentication numbers for one-time use in addition to a stable password) (Zagorscak & Knaevelsrud, 2019).

Patient level. Most IBI rely heavily on written communication. Thus, individuals with physical or psychological impediments (e.g., blindness, mental disabilities, reduced attention

span or motivation, illiteracy) might experience difficulties when using IBI (Zagorscak & Knaevelsrud, 2019). Further, by definition, IBI make use of mobile or stationary computers. Thus, they are not suitable to individuals that do not feel comfortable or competent using such devices. Finally, most research on IBI has been done on non-suicidal individuals with mild to moderate symptom severity and with no or unreported comorbidities (Andersson & Titov, 2014). Part of the reason for this lack of studies is the linear nature and disorder-specificity of a majority of the available IBI. Nevertheless, first trials suggest that IBI are efficacious and suitable for individuals with more severe depressive symptoms, suicidal ideation (e.g., Meyer et al., 2015; van Spijker et al., 2018) and comorbid disorders (Johansson, Sjoberg, et al., 2012). However, the number of these trials is small and prohibits final conclusions on the appropriateness of IBI for these target groups.

In summary, there are challenges on each of the three levels that can be met by organizational or technical adaptations (e.g., buying a new device, changing inclusion or exclusion criteria, implementing regular screenings for suicidal symptoms). However, some of the challenges evidently indicate a lack of research. The problem, whether IBI is suitable for patient populations with severe symptoms, suicidal ideation and/or comorbidities is only one illustrative example. It symbolizes a general research question that can be broadly summarized as:

"*What kind of IBI works in what way for what kind of individual?"*

Specific aspects of this general question are at the core of this dissertation. They are derived from existing controversies and gaps in the literature, which the following section details.

1.6 Current gaps in research on IBI for depression

This chapter will give an overview of previously conducted studies and scientific discussions that led to the development of the three studies this dissertation is based on. For that purpose, it will extend on some general shortcomings in the literature on IBI for depression. It will further focus on the theoretical background of each of the three original studies. In particular, it will detail controversies surrounding the importance of human "guidance" in IBI (STUDY 1), discuss the current knowledge and research gaps concerning factors contributing to the success of IBI (STUDY 2), and give an overview of what is known about depressive symptom course developments (i.e., "patterns of change") during IBI (STUDY 3).

1.6.1 General shortcomings

Sampling bias. While IBI have promising potential to circumvent barriers to treatment (SECTION 1.5.3) and thus reach previously undertreated populations, a common critique is that they fail to deliver on this promise. In a review Arnberg, Linton, Hultcrantz, Heintz and Jonsson (2014) summarized, that 53-61% of participants in IBI already had a history of other psychological treatments, reported high employment rates and a high degree of education. Most participants were women. Similar sample statistics were reported in recent German-based trials on IBI for depression (e.g., Späth et al., 2017). Arnberg et al. (2014) concluded that such findings "raise concerns about whether the effects found in most RCTs [on IBI for depression] can be generalized to those who today are underserved" (p. 11). This sampling bias in trials on IBI seems to be rooted partly in the way recruitment of participants is carried out. So far, the largest meta-analysis on the efficacy of IBI for depression conducted by Karyotaki et al. (2018) included 24 trials and 15 of those recruited *self-selected* community-based or occupational samples (e.g. through advertisements in newspapers). At least partly clinically recruited samples (e.g., from waitlist for other treatments, through primary care or from inpatient or outpatient treatment facilities) are included in the remaining nine trials (Hallgren et al., 2015;

Johansson, Ekbladh, et al., 2012; Johansson, Sjoberg, et al., 2012; Kenter, Cuijpers, Beekman, & van Straten, 2016; Kivi et al., 2014; Klein et al., 2016; Newby et al., 2013; Nobis et al., 2015; Sheeber et al., 2012). TABLE 1.3 provides an overview of the studies with clinical recruitment strategies.

TABLE 1.3
Randomized-controlled trials on IBI for depression with (partly) clinical recruitment

Trial	Recruitment source	Specifics of recruitment
Hallgren et al. (2015)	Clinical	Invitation after scoring >9 on the PHQ-9 in a participating primary care facility
Johansson, Ekbladh et al. (2012)	Community and Clinical	1) Newspaper advertisement 2) Individuals in waitlist for another IBI
Johansson, Sjoberg et al. (2012)	Community and Clinical	1) Newspaper advertisement 2) Individuals in waitlist for another IBI
Kenter et al. (2016)	Clinical	Individuals in waitlist for face-to-face treatment
Kivi et al. (2014)	Clinical	Recruitment by primary care practitioners inviting suitable patients to participate
Klein et al. (2016)	Community and Clinical	Self-selected; advertisement (flyers) in inpatient and outpatient medical and psychological clinics. Advertisement through other channels (online forums, newspaper and radio, health insurance)
Newby et al. (2013)	Community and Clinical	1) Online advertisement 2) Individuals in waitlist for another IBI
Nobis et al. (2015)[a]	Clinical	Online and offline advertisement; individuals diagnosed with diabetes were informed by health insurance representatives about the trial.
Sheeber et al. (2012)[b]	Clinical	Mothers of children participating in a school-based intervention were invited to participate if they reported elevated symptoms of depression

Note. Trial selection based on the meta-analysis of Karyotaki et al. (2018). [a]Trial included only individuals diagnosed with diabetes. [b]Trial included only rural mothers of children who participated in the "Head Start Classroom" program.

A closer look at these nine trials reveals another interesting pattern regarding the representativeness of the samples: In most cases they consisted of *self-selected* individuals as well. These individuals were either already recruited and waitlisted for another IBI (or in one case face-to-face psychotherapy) or they were made aware of the trial through different forms of advertisement. Only in four trials were the individuals recruited by being directly approached by a primary care practitioner, an insurance representative or a study employee (*external selection*). Unfortunately, findings from these four studies do not generalize well to the broader

field of IBI for depression. This is due to the fact that two focused on very specific populations (depressed individuals diagnosed with diabetes, Nobis et al., 2015; rural mothers, Sheeber et al., 2012), one investigated an unguided self-help intervention (Hallgren et al., 2015) and the fourth included face-to-face meetings and phone calls with study therapists and might thus be considered to be a "blended treatment" (Kivi et al., 2014).

Insufficient Sample Size (Statistical Power). Another aspect that is evident from recent meta-analyses on IBI for depression (Karyotaki et al., 2018, 2017; Königbauer et al., 2017; Sztein et al., 2018) is the small sample size in the included studies (see also TABLE 1.4). Depending on inclusion and exclusion criteria, the meta-analyses were based on total sample sizes of $n = 1631$ to $n = 4998$ stemming from a total of $k = 14$ to $k = 24$ individual trials. Only seven of the 24 trials included in the largest meta-analyses featured treatment conditions with sample sizes of $n > 100$. Such small sample sizes are only appropriate when two experimental conditions are compared on a limited set of variables and *large* between-group effects are expected. However, these conditions are not met when active treatment conditions are compared (see TABLE 1.4 for example of underpowered trials comparing active conditions). Further, studying mechanisms of change in IBI requires statistical models that simultaneously include multiple predictors of symptom development and consider their interrelations, which requires large sample sizes as well (Kazdin, 2007).

In summary, there is a need for sufficiently powered studies on IBI that include more diverse means of recruitment (i.e., self-selection and external selection of participants) in order to increase the validity and generalizability of findings. The relevance of these general problems will be further highlighted in the following sections, for underpowered trials with biased samples have also been used – in part - to generate knowledge about the benefits of guidance in IBI (see SECTION 1.6.2), the contributions of different treatment-specific and nonspecific factors to the success of IBI (SECTION 1.6.3), and clinical and socio-demographic predictors of depressive symptom courses in IBI (SECTION 1.6.4).

1.6.2 Shortcomings of Research on the Importance of Guidance in IBI for Depression

Since the first large meta-analysis on IBI for depression (and anxiety) (Spek et al., 2007) concluded that whether guidance is provided "differentiates between large and small effect sizes" (p. 327), the scientific consensus obtained through reviews and meta-analyses seemed to favor guided over unguided IBI with regard to efficacy on primary outcomes (Baumeister, Reichler, Munzinger, & Lin, 2014; Johansson & Andersson, 2012; Richards & Richardson, 2012).

Strikingly, what was defined as guidance varied fundamentally across trials (Baumeister et al., 2014; Berger, 2017). The aforementioned reviews and meta-analyses on this topic in IBI for depression thus applied a broad definition, and included trials in the "guided" category that featured any form of direct contact between a provider (e.g., therapist) and recipient (e.g., patient) of IBI (Johansson & Andersson, 2012; Richards & Richardson, 2012). This coarse definition had to be used in order to account for the existing inter-trial heterogeneity in intensity, frequency, and purpose of guidance. Further differences exist in qualifications of providers or the mode/channel of communication. For example, Richards and Richardson (2012) summarized that twelve studies in their meta-analysis included no support at all, ten studies included a clinician who provided feedback on homework, and twenty studies featured administrative support in "logistical or administrative ways and used receptionists, nurses, lay people, research coordinators, administrative staff, or technicians." (Richards & Richardson, 2012, p. 336). If guidance was provided, it mostly took place through written messages, but also by phone, through chats or through face-to-face meetings in-between IBI sessions.

Other inconsistencies arise in the definition of the opposite category as well: Some studies included "contact-on-demand" in their "unguided" conditions, meaning that participants could contact the research team at any time if problems (e.g., technical issues) arose (Boss et al., 2015). In other studies, individuals in the unguided condition had contact to the research

team before the start of the intervention for diagnostic purposes (Johansson & Andersson, 2012), which was proven to positively affect symptoms of depression in a trial on IBI for social anxiety (Boettcher, Berger, & Renneberg, 2012).

Critical evaluation of meta-analyses. The variability in the definition of *guidance* is problematic, for the coarse integration of qualitatively and quantitatively different support conditions might conceal potentially existing differences between them. Concordantly, meta-analyses that group findings "from studies that measured different things, manipulated different variables, and tested different subject populations" (Sharpe, 1997, p. 882) have been critiqued in general for decades for the potential lack of validity of their results (e.g., Sharpe, 1997; Szajewska, 2018). This means, with respect to guidance, that imprecise definitions in conjunction with other confounding factors like selective sampling and different measurement practices lead to effect estimates that represent an uninterpretable blend of all these aspects. A closer look at the studies included in the two most influential reviews and meta-analyses reveals that they too might suffer from this so called "apples and oranges" problem (Sharpe, 1997) for the variable "guidance" was confounded with other trial- and setting-specific conditions that are likely to influence outcome.

In the first review and meta-analysis, Richards and Richardson (2012) compared trials based on support that was offered (i.e., unguided trials were compared with trials with administrational guidance and therapeutic guidance). Notably, only 31% of the included *unguided* trials recruited participants from clinical settings (primary care, secondary care, hospitals, outpatient clinics), whereas the remaining trials used community-based convenience samples. In contrast, with 65% the rate of trials that recruited their participants from clinical settings was significantly higher in the *guided* categories.

In the second review, Johansson and Andersson (2012) created four separate categories of trials and compared their controlled effect sizes: 1) Trials with no contact, 2) Trials with contact before the intervention (e.g., diagnostic interviews), 3) Trials with contact during the

intervention (e.g., regular feedback) and 4) trials with contact before and during the intervention. Again, their grouping displays problematic patterns. First, half of the eight trials in the "no contact" category targeted adolescents, while all trials in the other categories targeted adults (including one trial targeted at senior citizens). Second, 75% of the controlled effect sizes in the "no contact" category were based on comparisons with *"care as usual"*. The same is true for 50% of the trials in category 2 (contact before the intervention), none of the trials in category 3 (contact during the intervention), and 21% of the trials in category 4 (contact during and before the intervention). The controlled effect sizes of all other trials were based on comparisons with *untreated waitlists*.

These trial- and setting-specific conditions are problematic for a number of reasons. Previous meta-analyses demonstrate that, IBI trials with adolescents yield lower effect sizes than those with adults (Pennant et al., 2015). The same is true for studies recruiting from community populations as opposed to more severely affected clinical populations (Bower et al., 2013) and for comparisons against "care as usual" as opposed to waitlist control groups (Cuijpers et al., 2013). FIGURE 1.4 shows how the results of Johansson and Andersson (2012) would change if these problematic studies (i.e., demographic bias or comparator bias) were removed from the data. More specifically, the figure shows, that the differences between the four categories cease to exist or are drastically reduced when only comparable trials are included.

While this is just an illustration that has its own shortcomings (e.g., too few studies per category), it demonstrates potential validity problems of these reviews and meta-analyses that might have shaped the scientific discourse on the topic of guidance in IBI. Finally, it is important to note that one recent meta-analysis on IBI for depression without these biases exists (Königbauer et al., 2017). In this meta-analysis only trials targeting individuals with a diagnosis of MDD obtained from a clinical interview were included. The only comparator condition was

waitlist and 17 of the 19 included trials had the same recruitment strategy/source. The authors of this meta-analysis did not find "guidance" to be a significant moderator of outcome.

Direct experimental comparisons of varying levels of guidance. The validity of the core finding that „guided" interventions are superior to „unguided" interventions is further taken into question by the lack of direct experimental comparisons that would support the majority of meta-analytic conclusions. So far, only five studies have directly contrasted the same or similar IBI for depression and manipulated the quantity or quality of guidance that was provided (Andersson, Sarkohi, Karlsson, Bjarehed, & Hesser, 2013; Berger, Hammerli, Gubser, Andersson, & Caspar, 2011; Kelders, Bohlmeijer, Pots, & van Gemert-Pijnen, 2015; Titov et al., 2010; Vernmark et al., 2010). TABLE 1.4 provides details on these five studies. All comparisons yielded no significant differences between the different guidance conditions. Importantly, four out of the five studies stated having possibly conducted underpowered trials or criticized their own sample size for being too small. Indeed, a sensitivity analysis performed on all these studies reveals that most trials were only sufficiently powered to reliably detect medium to large between-group effects (see TABLE 1.4). This finding further underlines the necessity for studies with larger sample sizes that provide sufficient power to detect effect size differences between guided and unguided conditions.

TABLE 1.4
Overview of studies that directly compared different qualities or quantities of guidance in IBI for depression.

Study	Conditions	Sample Size	Result	Statements about Sample Size in Discussion Section	Result of Sensitivity Analysis[b]
Titov et al. (2010)	1) iCBT with weekly guidance by clinician 2) iCBT with weekly assistance by technician 3) WLC	*N* = 141	Depressive symptom reduction pre-post: 1 = 2 > 3	"The relatively small sample size is one limitation of this study." (p. 8)	*d* = .5
Vernmark et al. (2010)	1) E-Mail-Therapy 2) Guided self-help 3) WLC	*N* = 88	Depressive symptom reduction pre-post: 1 = 2 2 = 3 1 > 3	"The study was underpowered to detect differences between the two active treatments." (p. 375)	*d* = 73
Berger et al. (2011)	1) Guided iCBT 2) Unguided iCBT 3) WLC	*N* = 76	Depressive symptom reduction pre-post: 1 = 2 > 3	"The study was underpowered to detect small differences between the guided and unguided condition." (p. 263)	*d* = .84
Andersson et al. (2013)	1) E-Mail-Therapy 2) Guided self-help	*N* = 47	Reduction in negative thinking pre-post: 1 = 2	"Some insignificant differences between the two treatments […] may be due to our small sample size." (p. 32)	*d* = .83
Kelders et al. (2015)	1) Human support[a] 2) Automated support	*N* = 239	Depressive symptom reduction pre-post: 1 = 2	---	*d* = .36

Note. iCBT = internet-based cognitive behavioral therapy, WLC = waitlist control group. [a]This study featured eight conditions the participants were randomized to. However, all eight conditions were grouped as either human or automated support and compared accordingly. [b]Result of own sensitivity analysis indicates the smallest detectable between-group effect size with α=.05 and β =.2.

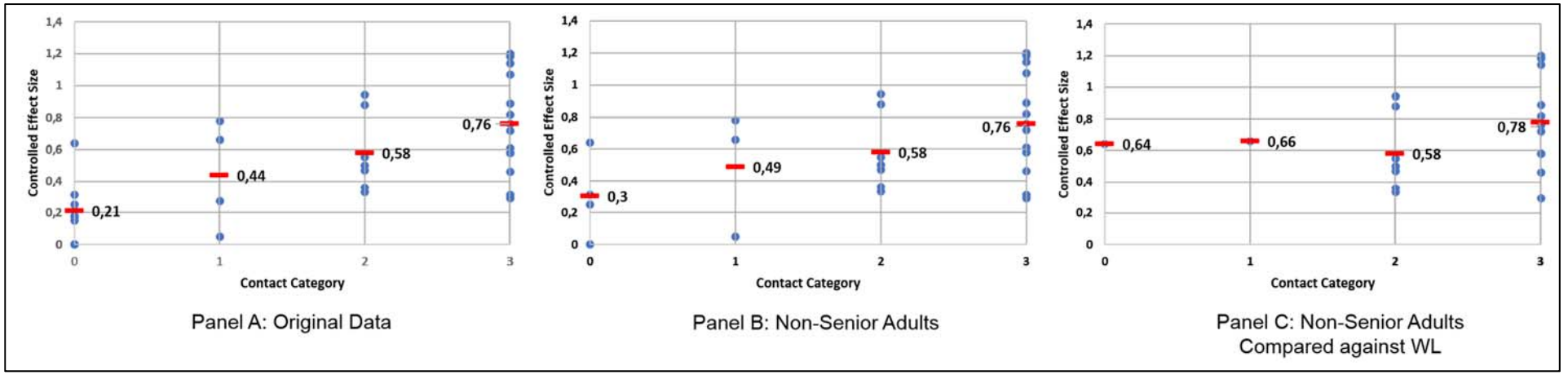


FIGURE 1.4. Depiction of controlled effect sizes from the review by Johansson and Andersson (2012) and changes if demographic bias and comparator bias is removed. Each dot represents one of the trials included into the original study. Red lines represent means of controlled effect sizes per category. WL = waitlist. Contact categories: 0 = none, 1 = before the intervention, 2 = during the intervention, 3 = before and during the intervention. Panel A shows the results of the original review. Trials on adolescents and seniors are removed in Panel B, additionally in Panel C trials are removed if they used "care as usual" as comparator condition. Data source: Johansson and Andersson (2012); own re-analysis of data after study removal with statistical methods used in the original publication (grand mean total of all studies per category).

In summary, while most reviews and meta-analyses support the notion that guided IBI for depression is superior to unguided IBI regarding efficacy in depressive symptom reduction, there are some methodological shortcomings of the conducted studies, such as sampling and comparator biases. In addition, experimental comparisons of different quantities and qualities of guidance or feedback do not support their conclusions, but there are very few such trials and the majority of studies are too heavily underpowered to detect even moderate group differences between two active treatments.

1.6.3 Research on Factors Contributing to the Success of IBI

Research on factors contributing to the success of IBI is still in its infancy and only a small number of limited studies exists. Thus, this section will introduce this topic from the historical perspective of research on contributing factors in face-to-face psychotherapy.

Theoretical perspectives. Identifying the elements of treatment that are beneficial, harmful or irrelevant to outcome is crucial in order to understand and improve psychotherapy. As one of the first researchers Rosenzweig (1936) formulated his theoretical assumptions about what contributes to the success of psychotherapy. Even before reliable evidence existed, he proposed that different therapeutic approaches are only *apparently* diverse and that all bona fide psychotherapies are equally effective. He suspected that the therapist's personality, attitude and general behavior are the most important contributors to treatment success and that given these factors "it is of comparatively little consequence what particular method that therapist uses" (Rosenzweig, 1936, p. 415).

Concerning the supposed lack of differences between different schools of therapy, he summarized his verdict with a quote stemming from the Dodo bird in the novel *Alice's Adventures in Wonderland*: "*Everybody* has won, and *all* must have prizes" (Carroll, 1922, p. 33). The influence of Rosenzweig's ideas extends into the present and both proponents and skeptics of this so-called "Dodo-bird-verdict" exist up until today (e.g., Asarnow & Ougrin,

2017; César González-Blanch & Laura Carral-Fernández, 2017). The proponents argue that the *only* relevant contributors to the success of treatment are *factors common to all psychotherapies*. Examples of such factors include the therapeutic relationship and related aspects such as the therapists' warmth or empathy as well as patients' hopes or expectations towards the succes of treatment (Wampold, 2015). The skeptics maintain, that there are additional *factors specific to any given treatment* that contribute as well. These factors are centered around the idea of therapeutic schools and are involved with the respective theoretical orientation and the applied methods or therapeutic techniques that are derived from it (Thomas, 2006).

The four-factor-model. It is apparent that the dichotomy of broad categories of "common" and "specific" factors seems limited in its informative value for research and practice. Consequently, more fine-grained models were suggested. One of the most influential model propositions is the four-factor-model by Lambert (1992) that was further developed by Miller, Duncan and Hubble (1996). The adapted model by Miller et al. (1996) differentiated between three aspects that were previously subsumed under "common factors":

1) relationship factors (originally labeled "common factors" by Lambert, 1992)
2) expectations (towards treatment (outcome))
3) patient factors or extra-therapeutic factors

In addition, the "specific factor" component by Lambert (1992) was labeled more precisely as

4) model/ techniques.

On the basis of a non-empirical literature review, Lambert (1992) further estimated the amount of contribution of each of these four categories to therapeutic outcome. These numbers together with examples of constructs representative of the four factors are illustrated in FIGURE 1.5.

As Stenzel and Berking (2012) summarized, numerous other theories on contributing factors exist such as the Generic Model of Psychotherapy (Orlinsky & Howard, 1987) or Grawe's contributing factors model (e.g., Smith & Grawe, 2003) to name just two examples.

However, none of these models were as widely and internationally dissiminated and received as the four-factor-model. The main reasons lie in its simplicity and generalizability: The other models can be constructed as derivatives, as all of their suggested categories fit in one of the four proposed by Lambert (1992) and Miller et al. (1996).

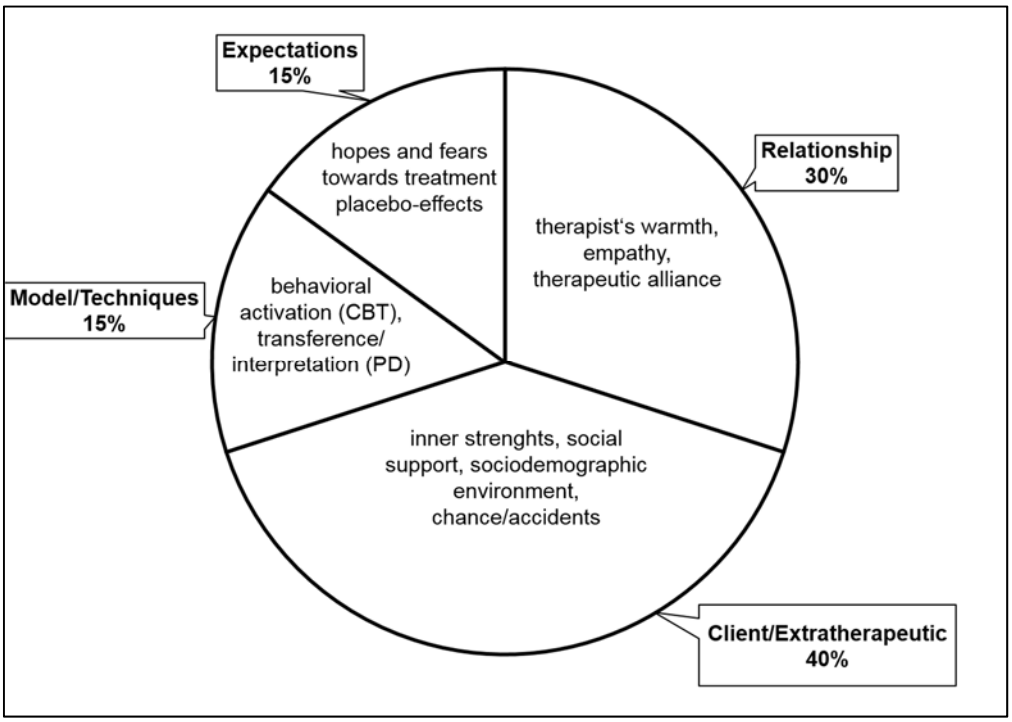


FIGURE 1.5. Four-factor-model by Lambert (1992) and Miller et al. (1996) with estimated percentage of contribution to psychotherapeutic outcome by Lambert (1992). Note, that the examples per factor are merely illustrative and not representative. CBT = cognitive-behavioral therapy, PD = psychodynamic therapy.

Empirical perspectives. The main reason that the debate about contributing factors is still ongoing is that it has not been empirically settled (for a summary see Cuijpers, Reijnders, & Huibers, 2019). On the one hand, there are large comparative meta-analyses that demonstrate the equivalence of different therapeutic approaches supporting the notion of the dodo-bird-verdict (Baardseth et al., 2013; Wampold et al., 1997). On the other hand, there are meta-analyses that report a small but significant superiority of cognitive-behavioral therapy over other therapeutic schools in general (Marcus, O'Connell, Norris, & Sawaqdeh, 2014) or for the treatment of anxiety disorders and depression in particular (Driessen et al., 2010; Tolin, 2010).

Since both sides critiqued the other's study selection and methodological approaches, the matter remains unresolved (e.g., Beutler, 2002; Mulder, Murray, & Rucklidge, 2017; Tolin, 2014; Wampold et al., 2017).

An interesting indirect empirical test of contributing factors in the treatment of depression was performed in a meta-analysis by Cuijpers and colleagues (2012). The authors investigated the efficacy of non-directive supportive treatments (NDST) for adult depression and compared them to specific psychotherapeutic treatments (e.g., cognitive-behavioral or psychodynamic), pharmacotherapy (antidepressants) and passive control groups. The authors defined NDST as "any unstructured therapy without specific psychological techniques other than those common to all approaches (e.g., helping people talk about their experiences and emotions and offering empathy), that is not aimed at solutions or acquiring new skills […]" (Cuijpers et al., 2012, p. 282).

The assumptions of the authors were that *only* common factors are active in NDST (i.e., relationship factors, expectations, extra-therapeutic factors), whereas *additional* specific contributing factors are at work in the comparator treatments that are aligned with a particular therapeutic school. Further, comparisons with waitlist control groups were used to discern the contributions of extra-therapeutic factors, which are assumed to be the only relevant contributors in this condition (Cuijpers et al., 2012).

The study found that common factors seem to be the most important and account for half of the improvement during treatment. The other half of the contribution is shared by extra-therapeutic factors (33.3%) and to a lesser degree treatment-specific factors (17.1%). As expectations and placebo effects may be at work in waitlists as well and NDST might feature some strategies that are similar to those in specific treatments, Cuijpers et al.'s (2012) assumptions seem like rather coarse approximations of reality. Nevertheless, the results bear a striking resemblance to the theoretically proposed numbers within the four-factor-model (see FIGURE 1.5) and thus support its validity and informative value.

Studies on contributing factors in IBI for depression. Given that the debate on contributing factors in face-to-face psychotherapy is decades older than the one concerning IBI, it is not surprising that no comprehensive explanatory model of treatment success in IBI exists. However, in predicting outcome, some studies on IBI have focused on expectations (e.g., Boettcher, Renneberg, & Berger, 2013), the therapeutic alliance (Flückiger, Del Re, Wampold, & Horvath, 2018) or on comparisons between different specific treatments (e.g., Ljótsson et al., 2014). These studies are discussed in more detail in CHAPTER 3.

In short, while there are no studies on extra-therapeutic predictors of symptom change in IBI, the studies on expectations and specific factors yield inconsistent and partly contradictory results regarding their importance. The only factor that was researched in a larger number of IBI studies is therapeutic alliance. A recent meta-analysis showed it to be significantly associated with outcome in IBI (Flückiger et al., 2018). However, it is unclear which of the subfacets of the construct (i.e., bond between patient and therapist, agreement on tasks or goals) are responsible for this finding. There is some preliminary evidence for the notion that the emotional bond between patient and therapist is less important in IBI than in face-to-face treatments, while an agreement on tasks and goals might be more important in this setting (Berger, 2017). Further the causal and temporal interrelations between therapeutic alliance and symptom change are unclear (i.e., whether symptom change predicts subsequent alliance ratings, alliance ratings predict subsequent symptom change or if it is a bi-directional association).

In summary, there is insufficient and partly contradictory evidence on the relative importance of the factors suggested by Lambert (1992) and Miller et al. (1996) in IBI. Importantly, if studies considered contributing factors, they focussed on one of these predictors, neglecting to controll for others. No study has investigated variables indicative of all four factors of the model simultaneously.

1.6.4 Research on Depressive Symptom Courses during IBI

The diagnosis of an MDD might suggests a certain uniformity in the group of patients carrying this diagnosis (see CHAPTER 1.2), however, even the diagnostic code formalized in the DSM-5 differs "based on whether this is a single or recurrent episode, current severity, presence of psychotic features and remission status" (APA, 2013, p. 162). This indicates that individuals diagnosed with MDD might display a considerable heterogeneity with regard to their current, past and future symptoms.

Empirical perspectives. Reviewing studies on long-term trajectories of depressive symptoms in the general population, Musliner, Munk-Olsen, Eaton and Zandi (2016) found that most researchers reported either three or four distinct classes of depressed patients. Patients in these classes differed regarding symptom chronicity and symptom severity. Clinical and socio-demographic variables such as lower income, lower education, stressful life events and female gender were linked to memberships in higher-severity groups and this membership was in turn associated with unfavorable clinical long-term outcomes (Musliner et al., 2016). Apart from naturalistic symptom courses, previous studies suggest considerable heterogeneity in symptom trajectories of depressed patients during face-to-face treatment as well (e.g., Aderka, Nickerson, Boe, & Hofmann, 2012; Cuijpers, van Lier, van Straten, & Donker, 2005; Lutz, Stulz, & Köck, 2009; Melchior et al., 2016; Schlagert & Hiller, 2017; Stulz & Lutz, 2007; Uher et al., 2010). Most of these studies identified distinct subgroups of patients that 1) benefit early in treatment with a high overall improvement, 2) benefit later in treatment and/or improve less, 3) do not benefit at all. Both, findings from naturalistic settings and from treatment-studies highlight the importance of this research area: In a first step, research on this topic helps to identify individuals who are at risk for unfavorable developments. In a second step, interventions that are tailored to these individuals' needs can be developed (Khan, Faucett, Lichtenberg, Kirsch, & Brown, 2012; Manen et al., 2015).

Studies on symptom courses during internet-based treatment. To date, only four studies have investigated depressive symptom courses during and after IBI (Batterham et al., 2017, 2018; Lutz et al., 2017; Sunderland, Wong, Hilvert-Bruce, & Andrews, 2012). To some degree, their results mirror the findings from face-to-face studies in that the IBI trials found at least two distinguishable classes of participants. The studies classified most individuals as improvers (76 – 95%), while a smaller class of individuals showed delayed or weaker response or even deterioration throughout treatment. In addition, consistent with findings from the face-to-face setting, Lutz et al. (2017) demonstrated that those who experience early improvement also achieve the best outcome overall. Still, a closer look at characteristics of these four studies reveals significant shortcomings in the generalizability of the results due to the intervention under research, the target population, the investigated outcome or the time frame where change was modeled (see TABLE 1.5 for an overview of specific study characteristics):

First, two out of the four studies did not investigate IBI for depression in general, but rather applied narrow interventions targeted at suicidal ideation (Batterham et al., 2018) or insomnia (Batterham et al., 2017), respectively.

Second, all of the studies focused on unguided IBI, albeit Lutz et al. (2017) offered additional guidance to more severely affected individuals (written messages). Unfortunately, the reported results did not differentiate between those who received guidance and those who did not.

Third, while one of the studies did not focus on fine-grained changes during treatment but rather on long-term outcomes of IBI (Batterham et al., 2018), another did not consider the treatment in its entirety and instead focused on early changes before and during the first four weeks of treatment (Lutz et al., 2017).

Fourth, the study of Sunderland et al. (2012) did not report symptom courses of depressive symptoms. The authors used a measure of non-specific psychological distress that

is best understood as a blend of symptoms of anxiety and depression (Kessler Psychological Distress Sclae, K10; Kessler et al., 2002).

Fifth, regarding socio-demographic or clinical characteristics associated with the trajectory classes reported in these studies, there are only few consistent findings. Overall, individuals that tend to have the worst baseline symptom constellations (e.g., higher symptom severity, lower physical health, socio-demographic disadvantages) tend to have the highest probability of membership in the least favorable trajectory class.

In summary, the few studies that exist on trajectories of depressive symptom change during unguided or fully-standardized IBI have a number of shortcomings limiting their informative value. No studies exist on depressive symptom courses in IBI with individualized feedback (guidance) by therapists.

TABLE 1.5
Summary of studies on depressive symptom courses in Internet-based Interventions (IBI)

Study	Study intervention	Target group	Target outcome	Investigated Time Frame (Measurement Occasions)	Sample Size	Resulting trajectory classes (with share of sample)	Patient characteristics associated with membership in least favorable class[b]
Batterham et al., (2017)	Unguided IBI for insomnia	Individuals with clinical insomnia and subclinical depression	Depressive symptoms (PHQ-9)	(1) Baseline (2) 4x biweekly during treatment (3) Post (4) FU-6months (5) FU-12 months (6) FU-18 months	*n* = 1149	(1) improving (95%) (2) stable/ deteriorating (5%)	More severe baseline depression; younger age, limited comfort with the internet
Batterham et al., (2018)	Unguided IBI for suicidal thinking	Individuals who currently experience suicidal thoughts	Depressive symptoms (CES-D)	(1) Baseline (2) Post (3) FU-6months (4) FU-12 months	*n* = 418	(1) high severity decreasing moderately over time (81%) (2) moderate severity with small decrease over time (19%)	Not in a relationship; not fulltime employed; high baseline insomnia and burdensomeness; low baseline belongingness
Lutz et al., (2017)	Unguided IBI for depression; additional guidance for severely depressed patients	Individuals with mild to moderate depressive symptoms	Depressive symptoms (PHQ-9)	(1) Pre-treatment screening (2) Pre-treatment registration (3) week 2 (4) week 4	*n* = 409	(1) early response after screening (45%) (2) early response after registration (39%) (3) early deterioration (16%)	Lower baseline physical health
Sunderland et al., (2012)[a]	Unguided IBI for depression	Individuals with symptoms of depression	non-specific psychological distress (K10).	(1) Baseline (2) 5 x biweekly during treatment (3) Post	*n* = 302	(1) responders (76%) (2) low responders (24%)	Higher baseline psychological distress

Note. PHQ-9, patient health questionnaire (9-item depression subscale); CES-D, Center for Epidemiological Studies Depression-Questionnaire; K10, Kessler-10 psychological distress scale. [a]study reports on two separate interventions, only results of depression trial are reported in Table. [b]class (3) in Lutz et al. (2017) and class (2) in all other trials.

1.7 The three studies at the core of this dissertation

As CHAPTER 1.6 demonstrated, there are gaps in the current literature on IBI for depression concerning the relevance of guidance, the four factors proposed by Lambert (1992) and patient-characteristics associated with discernable patterns of symptom change during treatment. Overall, existing knowledge about IBI for depression is mostly derived from trials with small samples of self-selected participants.

The following three chapters will present three studies that contribute new knowledge on these questions, while overcoming the methodological shortcomings.

- **Study 1** compares the effects of an unguided (standardized) and a guided (individualized) variant of the same IBI for depression on clinical outcomes. This study is presented in CHAPTER 2.
- **Study 2** investigates and compares the contributions of variables derived from the four-factor-model on weekly symptom change during standardized and individualized IBI for depression. This study is presented in CHAPTER 3.
- **Study 3** investigates whether distinct groups of patients experience qualitatively and quantitatively discernable symptom courses during standardized and individualized IBI for depression and whether patients' psychosocial or clinical characteristics are associated with membership in these groups. This study is presented in CHAPTER 4.

Overall, this thesis seeks to inform the development of IBI for depression by identifying features of interventions, patients, and the therapeutic process that are relevant for the success of treatment. Knowledge on these factors is essential to improve interventions and their tailoring to the patients' needs.

All three studies have been conducted with data collected from 1089 mildly to moderately depressed individuals drawn from the client-base of a public health insurance company, including both self-selected and externally selected individuals. They were

randomized to receive weekly feedback that was either fully-standardized and automated or individualized by a clinician within an otherwise identical cognitive-behavioral IBI for depression encompassing seven modules ("TK-DepressionsCoach"). All participants completed tasks on expressive writing, positive behavioral activation, cognitive restructuring and relapse prevention. The intervention is described more comprehensively in each of the following chapters and detailed in an unpublished intervention manual (Zagorscak, Sommer, Haug, & Knaevelsrud, 2014; available upon request).

CHAPTER 2

BENEFITS OF INDIVIDUALIZED FEEDBACK IN INTERNET-BASED INTERVENTIONS FOR DEPRESSION: A RANDOMIZED CONTROLLED TRIAL

Zagorscak, P., Heinrich, M., Sommer, D., Wagner, B. & Knaevelsrud, C. (2018). Benefits of Individualized Feedback in Internet-Based Interventions for Depression: A Randomized Controlled Trial. *Psychotherapy and Psychosomatics*, *87* (1), 32–45. https://doi.org/10.1159/000481515

Supplementary materials

The following supplementary materials related to STUDY 1 are available in the APPENDIX:

- *Online Table 1*. Change predicted by antidepressant medication, additional e-mail contacts and pre-interventional symptom load.
- *Online Table 2.* Estimated within-group changes for both primary outcome measures under various NMAR conditions.
- *Online Table 3*. Rates of reliable change, remission and recovery under various NMAR conditions.
- *Online Table 4*. Means and standard deviations for each outcome at each measurement occasion.

CHAPTER 3

FACTORS CONTRIBUTING TO SYMPTOM CHANGE IN STANDARDIZED AND INDIVIDUALIZED INTERNET-BASED INTERVENTIONS FOR DEPRESSION: A RANDOMIZED-CONTROLLED TRIAL.

A revised version of the following paper was published in the journal *Psychotherapy*:
Zagorscak, P., Heinrich, M., Schulze, J., Böttcher, J. & Knaevelsrud C. (in press). Factors contributing to symptom change in standardized and individualized internet-based interventions for depression: A randomized-controlled trial. *Psychotherapy*.

Factors contributing to symptom change in standardized and individualized internet-based interventions for depression: A randomized-controlled trial.

Short title: Contributing factors in IBI for depression

Pavle Zagorscak, Manuel Heinrich, Julian Schulze,
Johanna Böttcher, and Christine Knaevelsrud

Freie Universität Berlin

Author Note

Pavle Zagorscak, Clinical-Psychological Intervention, Freie Universität Berlin; Manuel Heinrich, Clinical-Psychological Intervention, Freie Universität Berlin; Julian Schulze, Psychological Assessment, Differential and Personality Psychology, Freie Universität Berlin; Johanna Böttcher, Clinical Psychology and Psychotherapy, Freie Universität Berlin; Christine Knaevelsrud, Clinical-Psychological Intervention, Freie Universität Berlin.

This trial was funded by the German public health insurance company "Techniker Krankenkasse" and prospectively registered (New Zealand Clinical Trials Registry, URL: https://www.anzctr.org.au (ID: ACTRN12614000312640)). The funding body was not involved in the study design, collection, analysis, and interpretation of the data; in the writing of the report; or in the decision to submit the article for publication. The authors do not report any conflicts of interest related to this publication.

Correspondence concerning this article should be addressed to Pavle Zagorscak, Department of Clinical-Psychological Intervention, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany. E-mail: pav.zag@fu-berlin.de

Abstract

Research suggests four categories of outcome predictors in face-to-face therapy (i.e., treatment expectations, extra-therapeutic factors, relationship factors, factors specific to a treatment approach/technique). However, it is unclear whether these factors are relevant in standardized and individualized internet-based interventions (IBI). To investigate this question, 1089 mildly to moderately depressed adults undergoing six weeks of cognitive-behavioral IBI for depression were randomized to either receive weekly individualized feedback (IF) or fully-standardized feedback (SF). The following variables corresponding to the four categories were tested regarding associations with depressive symptom change during multiple treatment periods within a multi-group structural equation model: 1) expectations, 2) extra-therapeutic stressors and stress change during treatment, 3) mid-treatment working alliance (task, bond) and 4) uptake of treatment-specific components (logins, specific tool use). Results confirm the importance of extra-therapeutic stressors and working alliance: an increase in stressors was associated with detrimental depressive symptom developments in both arms. Initial stress was related to deteriorations in SF, but not in IF, suggesting a stress buffer effect of individual guidance. Mid-treatment task ratings were related to prior symptom changes in both conditions, bond was only predictive for symptom changes in IF. Indicators of specific treatment component uptake and expectations showed no associations with symptom change.

Keywords: mechanisms of change, internet-based treatment, depression, working alliance, stress

Factors contributing to symptom change in standardized and individualized internet-based interventions for depression: A randomized-controlled trial.

Major Depressive Disorder is one of the most significant mental health challenges and account for 8.6% of years lived with disease worldwide (WHO, 2008). Consequently, psychotherapeutic interventions for depression have been researched intensively. Internet-based interventions (IBI) increasingly augmented the available treatment options for depressed individuals during the last twenty years. IBIs typically include mental health information and exercises. In some programs, patients receive professional guidance, which is mostly delivered in the form of semi-standardized e-mail feedback but may also take place through phone calls, chats or additional face-to-face meetings (Richards & Richardson, 2012). While meta-analyses confirm the efficacy of IBI in the treatment of depression (Karyotaki et al., 2017; Königbauer, Letsch, Doebler, Ebert, & Baumeister, 2017), factors contributing to treatment outcomes are not well understood. Nevertheless, most IBIs are based on treatment techniques (e.g., CBT) that are derived from literature on face-to-face therapy, where comprehensive theoretical models of contributing factors have already been proposed.

Contributing Factors in Face-to-Face Therapy

Reviewing the literature, Lambert (1992) and Miller, Duncan & Hubble (1996) suggested four categories of factors contributing to the outcome of face-to-face therapy ("four-factor-model"): 1) *relationship factors* (e.g., therapeutic bond and similar aspects such as agreement on tasks and goals, therapist's empathy or encouragement); 2) *client's expectations towards treatment*; 3) *specific factors or techniques* (i.e., therapeutic methods specific to one treatment approach), and 4) *extra-therapeutic factors* (e.g., client characteristics or changes in the client's life events and stressors outside of therapy).

While recent reviews confirmed these categories (e.g., Wampold, 2015), there is an ongoing debate on the relative importance of contributing factors. For example, some meta-analyses emphasize the role of factors specific to a given treatment approach (e.g., Marcus, O'Connell, Norris, & Sawaqdeh, 2014), while others deemphasize it (e.g., Baardseth et al., 2013). Only one meta-analysis attempted to empirically quantify the contributions of factors common to all treatment approaches and specific factors (Cuijpers et al., 2012). The authors constrasted the effects of non-directive supportive treatments ("common factors only") with specific treatments (e.g. cognitive-behavioral or psychodynamic treatments) and passive control groups on symptoms of depression. Results suggested that extra-therapeutic factors accounted for 33.3%, common factors for 49.6% and specific factors for 17.1% of overall improvement (Cuijpers et al., 2012). The effects were aggregated across multiple studies that did not address multiple contributing factors themselves, which is a limitation of the study design. Studying multiple predictors of change and their interrelations simultaneously requires large sample sizes and repeated assessments, requirements more easily realized in IBI. The current study aims at examining the relationship of different predictors of change in a large IBI study with depressive patients to close this research gap. Following the four-factor-model (Lambert, 1992; Miller et al., 1996), we assessed indicators of common as well as specific mechanisms of change, of external factors, and of client expectations. Overall, it is unclear, to what extent the four-factor-model (Lambert, 1992; Miller et al., 1996) transfers to the online context. Nonetheless, some previous studies have tested single predictors of outcome in IBI that correspond to the suggested categories.

Working Alliance as Common Factor in Internet-Based Interventions

There are only few studies that investigated the contributions of common factors to outcome in IBI. On the one hand, research on face-to-face treatments indicates that therapists'

general behaviors and personal characteristics do have a small measurable influence on outcome (Wampold, 2015), but these associations seem to be very low or non-significant in IBI (Almlöv, Carlbring, Berger, Cuijpers, & Andersson, 2009; Holländare et al., 2016). Features unique to IBI may explain these findings: Even in guided interventions, communication is usually highly standardized and supervisors revise feedback messages. As a result, all therapists behave fairly consistent across all patients, and therapists' characteristics do not seem to impact treatment substantially (Baldwin et al., 2011). On the other hand, there is variance in how patients perceive their therapists and treatments. Accordingly, a recent meta-analysis reported significant alliance-outcome relationships in IBI (Flückiger, Del Re, Wampold, & Horvath, 2018). However, results were inconsistent across sub constructs of working alliance. In a review on this topic, Berger (2017) summarized that a number of studies did not find significant alliance-outcome associations. In cases where the associations were significant "it was rather the client's agreement with the tasks and goals of the treatment than the bond between therapist and patient that predicted outcome" (p. 516). It is important to note that significant findings in previous studies might have emerged due to a lack of control of other variables, such as symptom severity before alliance assessments. Lately, trials on face-to-face interventions demonstrated that alliance is dependent on early symptom change (Zilcha-Mano, Dinger, McCarthy, & Barber, 2014) and the predictive validity of alliance ratings was drastically reduced in some trials when prior symptom change was accounted for (e.g., Constantino et al., 2017). Recent papers promote more detailed analyses of therapeutic alliance and their interactions with outcome over multiple time points during and after treatment (e.g, Falkenstrom, Ekeblad, & Holmqvist, 2016). These types of analyses have not been implemented in studies on IBI yet.

Client's Expectations towards Internet-based Interventions

The majority of studies exploring the role of clients' treatment expectations stems from IBIs targeting social phobia and results confirm that positive expectations towards outcome of treatment predict favorable symptom courses (Boettcher, Renneberg, & Berger, 2013; El Alaoui et al., 2015; Hedman, Ljotsson, & Lindefors, 2012). Interestingly, Nordgreen et al. (2012) only found significant associations with changes in social anxiety symptoms in the *unguided* but not in the *guided* arm of their trial. In contrast, expectations were a strong predictor of outcome in another study, where *guided* IBI for depression was provided (El Alaoui et al., 2016). To date, there are too few studies to conclude on the importance of expectations as predictors of symptom change, especially considering possible differences between guided and unguided treatments.

Treatment-Specific Factors in Internet-based Interventions

There is conflicting evidence from treatment component studies on the importance of specific therapeutic techniques. For example, research on internet-based CBT with and without behavioral strategies (such as exposure or behavioral activation) found them to provide small incremental benefits (Ljótsson et al., 2014) or no benefits at all (Christensen, Griffiths, Mackinnon, & Brittliffe, 2006; Schneider, Mataix-Cols, Marks, & Bachofen, 2005).

While there is value in assessing the contributions of specific factors through comparisons of different treatments in randomized controlled trials, Kazdin (2005) critiqued these approaches due to their coarseness. Following this line of reasoning, Sieverink and colleagues (2017) argued for opening up the so-called „black-box" of internet-based interventions by tracking the uptake of specific components of treatments that represent certain theoretical concepts. The authors suggest that the use of detailed log-data (e.g., login-duration, use of specific treatment tools, and completion of certain treatment modules) might shed light

on “the usage (the dose) that is needed to reach certain effects (the response)” (Sieverink et al., 2017, p. 2). First attempts at investigating the dose-response relationships in IBI for depressive symptoms found promising associations between the uptake of specific treatment components and overall outcome (Donkin et al., 2013; Whitton et al., 2015). However, these studies did not consider symptom changes immediately after the uptake of treatment components but instead focused on overall pre-post improvements, which impedes conclusions about the real source of symptom change.

Extra-therapeutic Factors in Internet-based Interventions

While the client’s characteristics and life circumstances are an essential contributing factor to therapeutic success in face-to-face treatments (Cuijpers et al., 2012), the two most recent meta-analyses on IBI for depression did not identify any socio-demographic variables as significant predictors of improvements (Karyotaki et al., 2017; Königbauer et al., 2017). Importantly, other extra-therapeutic factors, such as changes in life-circumstances (e.g., in family life, working situation, financial situation) have not been included as predictors of outcome in IBI. Accordingly, only static baseline characteristics of patients have been researched, while the influence of changes in life circumstances during therapy on outcome has not been included in trials on IBI yet.

The Current Study

In summary, while common and specific factors have been considered as predictors of outcome in IBIs to some degree, the results of existing studies are inconclusive. Furthermore, only single predictors of outcome were investigated so far. The current study aims at shedding more light on mechanisms of change in IBI for depression. Due to the lack of established explanatory models for outcomes in IBI, the aim is to identify relevant variables in an exploratory fashion which is nevertheless guided by the theoretical propositions of the four-

factor-model. Different common factors (e.g., working alliance, expectations towards treatment), specific factors (e.g., uptake of specific treatment tools) and extra-therapeutic-factors (e.g., distress due to external life circumstances) are assessed regarding their predictive value for depressive symptom change over multiple time periods during and after treatment in a large sample of clinically depressed adults. Following the previously outlined recommendations and inconclusive results (Falkenstrom et al., 2016), this study aims to evaluate working alliance as both a predictor and a criterion. To our knowledge, this is the first trial that examines multiple predictors of change and their interaction within a comprehensive model in IBI. The role of these factors is assessed in an individualized form (IF: semi-standardized weekly written feedback from a counselor; contact-on-demand) and a standardized form (SF: fully-standardized written weekly feedback; contact-on-demand) of the same IBI for depression within a two-arm randomized controlled trial with parallel group assignment. Consequently, the current study will investigate and compare mechanisms of change in IBI for depression with varying levels of written guidance (individualized vs. standardized feedback).

Method

Data was collected within a randomized controlled trial exploring the benefits of individualized, semi-standardized feedback compared to standardized feedback in otherwise identical internet-based cognitive-behavioral-therapy (CBT) for individuals with mild to moderate depression. Further details on this trial can be found in a previous publication, reporting on the comparative efficacy of individualized vs. standardized feedback in the form of a "black-box" evaluation [omitted for blinded review]. Given the established efficacy of internet-based CBT, only two active treatment arms were compared due to the ethical

implications of withholding an efficacious treatment from patients in need (Lemmens, Müller, Arntz, & Huibers, 2016).

Participants

Individuals were recruited from the client base of a German health insurance company. Nationwide recruitment started in March 2014 and ceased as planned after one year. Only participants meeting the criteria for mild to moderate depression (scores between 14 and 28) according to the Beck Depression Inventory II (BDI-II; Hautzinger, Keller, & Kühner, 2006) and not at risk for suicide (score ≤ 1 on BDI-II item 9) were included. Additionally, participants with current mania/hypomania or psychotic symptoms (lifetime) were excluded after a structured clinical interview by phone (SCID-I). A detailed flow-chart is provided in Figure 1. Overall, $N = 1089$ individuals (IF: $n = 555$; SF: $n = 534$), with a mean age of 45.7 (SD = 11.3) years were included. The majority was female (65.6%). No significant differences between the experimental groups were found at pre-treatment assessment (all p-values > .05, see Table 1).

Treatment

The intervention encompassed seven consecutive modules (M1-M7). The completion of one module took one week on average, resulting in a treatment period of six to eight weeks. Most treatment topics were offered for two modules each (i.e., M1 and M2: Exploring Thoughts, Feelings, and Behaviors; M3 and M4: Behavioral Activation; M5 and M6: Cognitive Restructuring and Interpretational Bias Training), while a single module focused on relapse prevention (M7). Figure 2 illustrates primary goals and treatment methods. In the IF condition, each participant was assigned to a personal counselor who provided semi-standardized, written feedback on the platform after each module. In the SF condition, participants automatically received fully-standardized feedback. Treatment content was the same for both groups. Independent of feedback, participants of both conditions could contact the research team (SF-

condition) or the individually-assigned counselor (IF-condition) upon demand (e.g., in case of technical problems). A previous publication revealed that both treatment arms yielded large pre-post effects on depression, as well as improvements in all other outcomes. Between-group differences were statistically nonsignificant in pre-post comparisons as well as 3, 6 and 12 months later [omitted for blinded review].

Treatment Allocation

After providing informed consent, eligible participants were randomized to either the IF- or the SF-condition. Treatment allocation was performed automatically by a computer-based random number generator supported by the host website.

Ethical Approval

The Research Ethics Committee of [omitted for blinded review] approved the protocol before recruitment of participants commenced. The trial was preregistered (New Zealand Clinical Trials Registry, URL: https://www.anzctr.org.au (ID: [omitted for blinded review])).

Measures

Figure 2 provides an overview of the measurement occasions. The present study only reports outcomes that are relevant for investigating predictors of change. Other outcomes that were assessed within the randomized controlled trial can be obtained from the public trial registry.

Patient Health Questionnaire-9. The severity of depression was assessed using the DSM based PHQ-9 (Löwe, Spitzer, Zipfel, & Herzog, 2003). Participants rated the frequency of depression-related behaviors/feelings during the past two weeks on a 4-point rating scale [*not at all* (0) to *nearly every day* (3)].

Expectations. Expectations towards treatment were assessed before the intervention with five established seven-point semantic differentials by Ajzen (1991) that were originally

developed to assess Theory of Planned Behavior constructs. Previous clinical studies that used similar (Lin, Updegraff, & Pakpour, 2016; Mausbach et al., 2013) or identical scales (Mendez, Rodrigues, Cornélio, Gallani, & Godin, 2010) demonstrated good psychometric properties of the instrument in this context. The original wording was adapted to address expectations towards future participation in IBI. Patients rated, whether participation in the intervention during the next six weeks would be "*beneficial"* to "*harmful,"* "*pleasant"* to "*unpleasant,"* "*good"* to "*bad,"* "*meaningful"* to "*meaningless," "favorable"* to *"unfavorable"* for them.

Working Alliance. Working alliance was assessed mid-treatment (at the beginning of M5) using the Working Alliance Inventory-Revised (Hatcher & Gillaspy, 2006) as adapted by Berger, Boettcher, and Caspar (2014) for online self-help programs. In contrast to the original version, items of the goal and task component assess the individuals' perception of the program rather than the therapist's efforts (e.g., "I know what I can expect as a result of using the online program"). The third component assesses the perceived bond between participants and their counselor (IF condition) or the research team responsible for the intervention (SF condition). As highlighted by Falkenstrom, Hatcher, Skjulsvik, Larsson, and Holmqvist (2015), the goal and task component of the WAI have shown high factor inter-correlations across different studies (Falkenström, Hatcher, & Holmqvist, 2015; Falkenström, Hatcher, Skjulsvik, et al., 2015; Munder, Wilmers, Leonhart, Linster, & Barth, 2010) and are not differentiable. Consequently, items assessing task and goal were summarized within one factor. The bond component represented the second factor.

Patient Health Questionnaire - Stress Module. The amount of psychosocial stress was assessed using the *Patient Health Questionnaire-Stress Module* (PHQ-S; Löwe et al., 2003). Participants rated the degree to which they felt stress corresponding to ten different psychosocial events (i.e., difficulties with their partner or family or financial worries) on a

three-point rating scale [*not impaired* (0) to *strongly impaired* (2)]. The mean score was used as a proxy for the participants' overall burden. The PHQ-S was assessed prior to the intervention and at post-assessment.

Uptake of specific treatment components. The platform hosting the intervention automatically tracked several quantitative indicators of the participants' intensity of treatment use. The *number of logins* and the time spent on the platform (*login-duration*) was counted for each treatment module. Participants could access the treatment module they currently work with as well as the content of completed modules at any time. There is no upper bound regarding the number of logins and minutes participants were allowed to spend working on treatment tasks. As a proxy of specific treatment component use, the *number of written letters* in expressive writing tasks (M1, M2 and M7), *the number of entries in the activity planner* (M3 and M4) as well as the *number of completed sessions of the interpretational bias training* (M5 and M6) and the *number of written thought protocols* (M6) were automatically tracked. To address the question whether specific treatment components account for symptom change, indicators were averaged over modules addressing the same topic with the same techniques (i.e., M1+M2; M3+M4; M5+M6; M7). Averaged indicators for login time, login duration and specific tool use were included as predictors of symptom change in further analyses.

Statistical Analysis

Structural Equation Models. Mean changes, as well as predictors of change, were investigated using multi-group (MG) latent difference (LD) models (Steyer, Eid, & Schwenkmezger, 1997). The MG approach allows for differences in means, variances, and co-variances between the IF and SF condition and was therefore favored.

Measurement Invariance. Strong measurement invariance is needed to compare means (Millsap, 2011). Measurement invariance was tested in the following sequence: A model

assuming configural invariance (free loadings and intercepts across groups and time) was followed by assuming weak factorial invariance (loadings constrained equal across time and group) and strong factorial invariance (equal loadings and intercepts across groups and time).

Mean change in depression. As a precondition for investigating predictors of mean change in depression, the amount of depressive symptom change during each treatment period was estimated, tested for significance and compared across groups using Wald tests. Due to the focus on changes during time periods, where specific treatment components were offered (see Figure 2), depressive symptom changes were calculated between M1 and M3, M3 and M5, M5 and M7 as well as M7 and post-assessment.

Predictors of change. To identify significant predictors of symptom change, depression change scores for each treatment period were regressed on expectations, baseline stress, changes in stress between pre- and post-assessment and depressive symptom change during the previous treatment period. Furthermore, bond and task ratings, obtained at the beginning of M5, were used as predictors of symptom change during subsequent treatment periods (M5-M7, M7-Post). Additionally, depressive symptom change during each treatment period was regressed on usage behavior (e.g., login duration, specific tool use) during that period and the previous one ("lagged path"). For example, usage behavior between M1 and M3 was used to predict depressive symptom change between M1 and M3 as well as between M3 and M5.

In order to account for the possible role of working alliance as both a predictor and a criterion, expectations, change in depressive symptoms during treatment periods prior to M5, as well as initial stress and depressive symptom-load were tested as predictors of working alliance components at M5.

Estimation and Model Fit

Taking the complexity of the model into account, parcels were formed for depressive burden as well as for the task/goal and bond component of the WAI. Models were estimated using the MLR estimator as implemented in M*plus* 8. Model fit was evaluated taking the RMSEA (< .05 good fit; .05 < RMSEA ≤ .08 acceptable fit), the CFI (.97 ≤ CFI ≤ 1.00 good fit; .95 ≤ CFI < .97 acceptable fit), and the SRMR (0 ≤ SRMR ≤ .05, good fit; .05 < SRMR ≤ .10 acceptable fit) as well as the chi-square test of exact model fit into account (Schermelleh-Engel, Moosbrugger, & Müller, 2003). Due to the large sample size, more weight was put on RMSEA, CFI and SRMR. In addition, a decrease in CFI smaller than .010 (RMSEA < .015 and SRMR < .030) was considered as acceptable when testing for measurement invariance (Chen, 2007).

Missing Data

Overall, 234 (21.5%) participants did not complete the intervention (e.g., did not start working with every treatment module), with higher rates in the SF condition, 25.8% vs. 17.3%, $\chi^2(1, N = 1089) = 11.780$, $p = .001$, $\varphi = -.10$. Missing data patterns for the PHQ-9 ratings are monotone up to M7 with increasing rates of missingness from 2.2% (IF) and 2.8% (SF) at M1 to 17.3% (IF) and 25.8% (SF) at M7. All randomized participants were invited to complete the post-assessment. As a consequence, the rate of missingness at post-assessment deviated from the strictly monotone pattern: 21.4% (IF) and 24.2% (SF). Except for group comparisons in usage behaviors (pairwise deletion), missing data was handled using Multiple Imputation (MI) via chained equations using MICE in *R* 3.4.3 (100 data sets) under the assumption that data are missing at random (van Buuren & Groothuis-Oudshoorn, 2011). An inclusive imputation strategy was used (Collins, Schafer, & Kam, 2001), that is, the imputation model included all variables used in the later analysis. MI was conducted separately for each group to maintain

consistency between group assignment, change and association with predictors. All results are pooled across all imputed data sets.

Results

Primary Data Analysis

Primary analyses revealed that some participants displayed irregular patterns of uptake. For example, 97.5% of the participants did not log-in more than five times at M1; however, some participants logged in more than 30 times (maybe due to technical problems). Consequently, the upper 5% of the number of letters and login-durations and the upper 2.5% of planner-entries and login-frequencies were replaced with less extreme values while maintaining the ordering of all individuals within the sample (Winsorizing) (Wilcox, 2003).

Measurement Invariance

Under the assumption of strong measurement invariance, a baseline model (only containing depressive symptom load over time) showed a good approximate fit to the data: average $\chi^2(136) = 269.725$, average RMSEA = .042, average CFI = .984, average SRMR = .037. The same is true for the final model (including all predictor and criterion variables): average $\chi^2(918) = 1151.300$, average RMSEA = .026, average CFI = .977, average SRMR = .036. Results of model comparisons further support the assumption of strong measurement invariance in both cases (model comparisons are detailed in Appendix B/Table B.1).

Change in Depressive Symptoms over Time and across Conditions

Within-group effect size estimates revealed statistically significant improvements in depression severity from M1 to M3, M3 to M5 and from M5 to M7 in both treatment conditions (all p-values < .001; range of effect estimates: $d_w = -0.461$ to $d_w = -0.686$). However, symptoms increased significantly from M7 to post-assessment in the SF-condition, $M_{LD} = 0.094$, $p < .001$, $d_w = 0.410$, but remained stable in the IF-condition, $M_{LD} = 0.025$, $p = .117$. Depressive symptom

change during each treatment period was tested for between-group differences by constraining all corresponding means to be equal across conditions and testing whether such a constraint leads to a worsened model fit (Wald test).

This indicated a significant between-group difference in change overall, W = 12.591, $df = 4$, $p = .014$. Post-hoc pairwise comparisons revealed that the increase in symptom load from M7 to post-assessment was stronger in the SF condition ($\Delta LD_{SF\text{-}IF} = 0.069$, SE = 0.023, $p = .003$, $d_W = -0.297$). All other between-group differences were non-significant. Details on within-group and between-group differences in mean change are available in Table B.2 (Appendix B).

Predictors of Depressive Symptom Change

The following sections summarize predictors of change in depressive symptoms. Significant standardized and unstandardized regression weights are summarized in Table 2 and are illustrated in Figure 3. Regression weights and significance tests of all other paths are made available in Table B.3 in (Appendix B). Only unstandardized regression weights are reported in-text due to being better comparable across groups. Please note, that numerical increases in change scores over time can be summarized as less favorable symptom courses (i.e., stronger deterioration or weaker improvement), whereas decreases in change scores indicate more favorable symptom courses (i.e., stronger improvement or weaker deterioration; please refer to Appendix A for additional information on interpreting change scores).

Prior depressive symptom change. In both treatment conditions, changes in depressive symptoms (Δ) during one treatment period were negatively associated with symptom changes during the subsequent treatment period ($\Delta_{M1\text{-}M3} \rightarrow \Delta_{M3\text{-}M5}$, IF: $b = -0.177$, $p = .040$; SF: $b = -0.257$, $p = .038$; $\Delta_{M3\text{-}M5} \rightarrow \Delta_{M5\text{-}M7}$, IF: $b = -0.264$, $p = .005$; SF: $b = -0.294$, $p < .001$; $\Delta_{M5\text{-}M7} \rightarrow \Delta_{M7\text{-}Post}$, IF: $b = -0.423$, $p < .001$; SF: $b = -0.242$, $p = .006$). On average, more favorable symptom

courses during one treatment phase were associated with less favorable symptom courses during the subsequent treatment phase.

Expectations towards treatment. Expectations were not associated with symptom change during any treatment period.

Specific factor "uptake of treatment components". None of the indicators of treatment uptake showed significant associations with symptom change during the respective treatment module or the following treatment module (lagged path).

Extra-therapeutic factor "external psychosocial stressors". In the IF condition, pre-post changes in PHQ-Stress ratings were positively associated with depressive symptom change during all modules (M1-M3: $b = 0.156$, $p = .006$; M3-M5: $b = 0.168$, $p = .006$; M5-M7: $b = 0.196$, $p < .001$; M7-Post: $b = 0.133$, $p = .010$). In the SF-condition, the same pattern emerged, but the association between PHQ-stress change and depressive symptom change during cognitive restructuring (M5-M7) was non-significant (M1-M3: $b = 0.124$, $p = .022$; M3-M5: $b = 0.184$, $p = .013$; M5-M7: $b = 0.082$, $p = .132$; M7-Post: $b = 0.244$, $p < .001$). On average, individuals that felt less stressed by external psychosocial events between pre- and post-assessment, also experienced more favorable depressive symptom courses during all (IF) or most (SF) treatment periods under investigation.

Common factor "working alliance". In the IF condition, task ratings assessed at M5 showed negative associations ($b = -0.089$, $p < .001$) with depressive symptom change between M7 and post-assessment, whereas bond ratings assessed at M5 showed positive associations with symptom changes during that later period ($b = 0.051$, $p = .014$). On average, individuals in the IF-condition that reported higher task-ratings mid-treatment (M5), reported more favorable symptom courses during the final period of the intervention. In contrast, individuals in the IF-condition that reported a stronger bond to their counselor mid-treatment reported less

favorable symptom courses during the final period of the intervention. Neither task nor bond ratings showed significant associations with later depression change in the SF-condition.

Predictors of Mid-Treatment Working Alliance Ratings

Expectations. In both treatment-conditions, expectations towards treatment at pre-assessment were positively associated with mid-treatment task-ratings (IF: $b = 0.278$, $p < .001$; SF: $b = 0.224$, $p < .001$). On average, individuals that reported more optimistic expectations before treatment also reported stronger agreement with the tasks and goals of the intervention later on. Additionally, expectations were positively associated with mid-treatment bond-ratings in the IF-condition (IF: $b = 0.303$, $p < .001$), but not in the SF-condition. On average, individuals in the IF-condition that reported more optimistic expectations before treatment also reported a stronger bond.

Baseline symptom severity. In both treatment conditions, baseline symptom severity (M1) was negatively associated with mid-treatment task ratings (IF: $b = -0.381$, $p = .002$; SF: $b = -0.530$, $p = .002$). On average, individuals with higher baseline symptom severity reported weaker agreement with tasks and goals of the intervention later on.

Depressive symptom changes prior to alliance ratings. Symptom changes during the first two treatment periods (M1-M2: Expressive Writing; M3-M4: Behavioral Activation) were negatively associated with mid-treatment task ratings (M1-M2, IF: $b = -0.749$, $p = .001$; SF: $b = -0.560$, $p = .034$; M3-M4, IF: $b = -0.693$, $p < .001$; SF: $b = -0.918$, $p = .041$). On average, individuals reporting more favorable symptom courses during these modules also reported stronger agreement with task and goals of treatment later on.

Additionally, symptom changes during behavioral activation (M3-M4) were negatively associated with bond-ratings in the SF-condition (SF: $b = -0.594$, $p = .015$), but not in the IF-

condition. On average, individuals in the SF-condition reporting more favorable symptom courses during M3-M4 also reported a stronger bond.

Discussion

The current study is the first to investigate and compare mechanisms of change in IBI for depression with varying levels of guidance (individualized vs. standardized feedback) on the basis of the theory of contributing factors proposed by Lambert (1992). The results indicate significant associations of the common factor working alliance and extra-therapeutic factors (perceived psychosocial stress) with outcome in IBI for depression. By contrast, neither expectations nor the uptake of specific tools or modules were directly linked to outcome.

Previous Symptom Change

The study revealed that previous symptom changes are the most consistent predictors of subsequent symptom changes. On average, individuals displaying more favorable symptom courses during earlier modules showed less favorable symptom courses in later modules and vice versa. This mirrors findings from meta-analyses on the efficacy of psychotherapy, demonstrating that individuals who experience larger initial improvements in their symptoms may improve less afterward and that patients who start out with more severe symptoms have a more potential for improvement (e.g., Bower et al., 2013).

Relationship factors

Overall, this study yielded an interesting pattern regarding the *common factor* of working alliance. First, we replicated findings from face-to-face studies (e.g., Zilcha-Mano et al., 2014) in that initial symptom severity and early symptom change predicted later task ratings.

Second, working alliance ratings were significant predictors of subsequent symptom change for patients that received individual feedback (IF-condition). However, only higher

task-ratings had a beneficial influence on depressive symptoms. Regarding bond ratings, associations pointed in the opposite direction: Patients who perceived the bond to their counselor as stronger, showed higher depression change scores in the first two weeks after termination, indicating lessened improvement or deterioration immediately after the end of treatment. On the one hand, these findings are in line with Berger's (2017) review on therapeutic alliance in IBI in that it is mostly the task/goal-component of the alliance ratings that predicts outcome. On the other hand, it seems counterintuitive, that higher bond ratings are associated with detrimental symptom developments at the end of treatment. Nonetheless, there are indications from multiple studies that patients see the termination phase as a critical time point in psychotherapy that is associated with feelings of loss and pain (Knox et al., 2011). In IBI there is a very short termination phase and some patients may experience the loss of their therapist as abrupt – especially if they experienced a positive bond during treatment. This might be associated with depressive symptom increase. Interestingly, in a consensus statement on possibly adverse side effects stemming from IBI, multiple researchers did not consider "withdrawal symptoms" from psychotherapy or psychotherapists as a possibility (Rozental et al., 2014). Therefore, future research should investigate whether this phenomenon can be replicated and whether it generalizes over different forms of IBI and different patient populations. It seems reasonable to assume that termination should be addressed more explicitly during treatment and should be prepared more thoroughly to avoid detrimental outcomes.

Expectations towards treatment

Patient's expectations did not predict outcomes directly. This contrasts suggestions by Lambert (1992) who proposed a small but significant influence of patients' expectations. A meta-analysis on the contributions of expectations on outcome in psychotherapy found small effects ($d = 0.24$) (Constantino, Arnkoff, Glass, Ametrano, & Smith, 2011). This effect estimate

was based on a summary of studies, which usually did not control for other variables and might thus have overestimated the contribution of expectations. For example, Patterson, Anderson, and Wei (2014) controlled expectation-outcome associations for the influence of working alliance and found expectations to be mostly related to therapeutic alliance. Only one in four facets of the construct (expectation for counselor expertise) was independently related to outcome. These results are in line with our finding that expectations predicted working alliance, and working alliance predicted some aspects of symptom change.

Specific factors

Regarding specific factors, this study did not find associations between the uptake of any specific tools, login durations or login frequencies and outcome. Results were consistent over both conditions. These findings may be interpreted in support of the dodo-bird verdict which proposes that all bona fide treatments yield comparable effects due to the contributions of common factors (e.g., Baardseth et al., 2013). Nevertheless, it is important to stress that the assessed indicators of uptake in this trial are only proxies of the target behavior. While the use of the intervention and its tools were objectively tracked, it is unknown what the patients did "offline." For example, some patients might have used the information they received from the intervention and applied techniques such as positive behavioral activation utilizing a paper-pencil calendar instead of online tools. Additionally, only the quantity of uptake was assessed, while content *quality* and *compliance* with instructions were not measured. For instance, it was not possible to control whether entries in thought protocols indeed contained alternative views of situations that would lead to positive emotional consequences. Finally, in both conditions only cognitive-behavioral techniques were offered, limiting the generalizability of results over other forms of treatments (e.g., psychodynamic, mindfulness-based or interpersonal interventions). A suggestion for future studies addressing the uptake of specific treatment

components in IBI would be to assess other treatment approaches as well and to include quality and compliance ratings as moderators of dose-outcome associations.

Extra-Therapeutic Factors

The results of the current trial further showed that changes in extra-therapeutic factors (perceived external psychosocial stressors) predicted outcomes at most time points in both treatment arms. On average, patients that reported alleviation in stressful life circumstances during treatment also showed larger depressive symptom improvements at multiple time points. This finding is not surprising, given the large body of evidence supporting the role of stress in the etiology of depression (for a review of the literature see Pizzagalli, 2014). The relative importance of life circumstances outside of therapy that is found in this trial on IBI is consistent with theoretical propositions of Lambert (1992) and empirical results from Cuijpers et al. (2012). Both researchers suggest that extra-therapeutic changes are the most critical factors contributing to outcomes in face-to-face psychotherapy. The current trial suggests a similar pattern in IBI for depression and points to the importance of considering patients extra-therapeutic life circumstances during online therapy. However, the current study relies on self-assessments of the perceived burden by external stressors, which might also show some overlap with symptoms of depression themselves. While perceptions of external stressors might change due to participation in IBI regardless of objective changes in external stressors, it is unlikely that the assessed burden (i.e., from work, education, family or financial situation) is sufficiently explained by perceptions alone. Therefore, it is reasonable to assume that a substantial portion of variance in this variable indicates real changes in stressful life circumstances. High correlations between objective assessments of stressors (i.e. financial circumstances, familial obligations), stress perceptions, and biological markers of stress support this conclusion (Bull, Almond, Christensen, & Fenech, 2014; Ursache, Noble, & Blair, 2015). Another limitation

stems from the questionnaire's focus on a limited set of adverse conditions, while previous studies also emphasized the importance of positive life events (e.g., Blonski, Conradi, Oldehinkel, Bos, & de Jonge, 2016). For that reason, future studies should include more diverse uplifts and hassles while relying on more objective assessments.

Differences between IBI with Individualized and Standardized Feedback

One of the assets of the present study is the possibility to investigate differences between IBI with individualized and standardized feedback regarding the underlying mechanisms of change. While results within IF and SF conditions revealed similar patterns, there were some differences regarding the influence of working alliance ratings and perceived stress.

In contrast to the IF condition, neither bond nor goal/task ratings of the working alliance were predictive of symptom change in the SF condition. A possible explanation lies in difficulties of patients to rate bond-items in the SF condition. It is unclear, whether patients validly rate the "alliance to the program" or if they feel obliged to answer these questions, even if they are not sure how to interpret them. Incidentally, 24 participants of the SF condition complained about the working alliance questionnaire in the open-comments section at the end of treatment. These patients indicated that they had difficulties answering adequately, given the lack of interaction with clinicians. There are reports from multiple studies on working alliance in IBI suggesting that the construct in IBI may differ from the one in face-to-face therapy and that different working alliance measures are needed (Berger, 2017). The same review also reported that some participants in IBIs were confused by working alliance questionnaires. Since this seems to be an overarching problem, future studies should analyze the validity of the instrument's facets for assessing bond in unguided IBI more thoroughly, applying qualitative methods (e.g. focus groups, thinking-aloud-techniques).

Second, the results indicate differences in the influence of perceived stress on outcome between the two conditions. While changes in stress seem to be influential in both treatment conditions, the initial level of perceived burden due to external stressors assessed at baseline only showed significant associations with changes in depressive symptoms in the SF condition. On average, participants with high initial stress levels showed lower improvements or stronger deteriorations in their depressive symptoms at multiple time points during the SF treatment. That was not the case in the IF condition. This finding is consistent with the "buffer-hypothesis" of social support on stress that was established through meta-analyses in the past (Smith, Fernengel, Holcroft, Gerald, & Marien, 1994). Participants in the IF condition received more individual contact and human support, while participants in the SF condition typically only received support in the form of standardized psychoeducation. Social support and "formal help" in the form of psychotherapy share similarities in the way they may buffer stress (Barker & Pistrang, 2002). Thus it is to be expected that the initial stress levels of participants play a lesser role in the IF condition. In practical terms, it seems especially important for unguided interventions to assess patient's initial stress levels and to offer them tools to cope with their environmental stressors that compensate for the missing "social buffer" in the form of a human online counselor.

Limitations and directions for future research

There are some general limitations to be considered when interpreting the results. Only patients with mild-to-moderate symptom severity were included in this trial, thus limiting the generalizability of findings. This may have resulted in low covariance between the assessed predictors and outcome and could have reduced the probability of detecting significant pathways (restriction of range). Further, this trial lacks a passive (i.e. waitlist) control group. However, summarizing qualities of good randomized controlled trials researching variables

associated with treatment outcome Kraemer (2002) highlights the need for a comparison group that may be both active or passive. Additionally, for the variables under investigation in this trial (uptake of treatment components, expectations towards treatment, therapeutic alliance) it is difficult to investigate them within an untreated sample.

Another limitation pertains to the frequency of assessments. While symptoms were assessed before and after each treatment phase, other variables such as stress and therapeutic alliance were only assessed once or twice, thus possibly neglecting fine-grained changes and impeding interpretations about chronology and causality. It is important to emphasize that it is challenging to establish causality within a single trial (Kazdin, 2007). Further studies should aim to confirm these findings using more intensive and methodological diverse assessment strategies. While this study sheds light on the associations of essential constructs representing the four categories proposed by the four-factor-model (Lambert, 1992; Miller et al., 1996), the included variables only represent a limited set of proxies of these broad categories. For example, common factors such as therapists' or programs empathy, genuineness or cultural appropriateness might complement the results of this study.

It is critical to note that most trials on the topic of contributing factors, including the one at hand, focused on guidance through written messages exchanged between clinician and patient. It is noteworthy that the results of studies featuring guidance through (video-)chats, telephone or other means of communication might deviate from the findings presented here. These alternative forms of interaction might be more similar to face-to-face communication due to their synchronicity or the availability of visual or acoustic cues (Dennis, Fuller, & Valacich, 2008).

Consequently, given the overall lack of research on mechanisms of change in IBI, the most important challenge for future research on this topic is the replication and extension of our

findings in more diverse settings of IBI (e.g., blended-treatments, telephone-guided treatments, mobile-based treatments) and within more diverse and less restricted groups of patients.

Conclusion

The current study emphasizes the relevance of researching contributing factors in IBI for depression. On the one hand, therapeutic alliance and changes in external life circumstances were identified as influential predictors of symptom change. On the other hand, further factors proposed as influential in face-to-face-treatments by Lambert (1992), such as expectations towards treatment and proxies of specific treatment component uptake showed no significant contribution to symptom change. While this study does not explicitly address the question whether the same mechanisms of change are in place in IBI and face-to-face therapy, the results indicate that there is a lot of variance in symptom change left to be explained and that the mechanisms suggested for face-to-face treatments might not be replicated in every form of IBI. This finding calls for additional studies on this topic in the future, moving beyond mainly considering the contributing factors suggested for face-to-face interventions. Moreover, the differential patterns between the standardized and the semi-standardized feedback condition found in this trial, underline the fact that "IBI" stands for a heterogeneous group of interventions and the quality and quantity of "human" guidance in this form of treatment has to be addressed more explicitly in upcoming studies on the mechanisms contributing to the success or failure of online therapy.

References

Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Almlöv, J., Carlbring, P., Berger, T., Cuijpers, P., & Andersson, G. (2009). Therapist Factors in Internet-Delivered Cognitive Behavioural Therapy for Major Depressive Disorder. *Cognitive Behaviour Therapy*, *38*(4), 247–254. https://doi.org/10.1080/16506070903116935

Baardseth, T. P., Goldberg, S. B., Pace, B. T., Wislocki, A. P., Frost, N. D., Siddiqui, J. R., … Wampold, B. E. (2013). Cognitive-behavioral therapy versus other therapies: Redux. *Clinical Psychology Review*, *33*(3), 395–405. https://doi.org/10.1016/j.cpr.2013.01.004

Baldwin, S. A., Murray, D. M., Shadish, W. R., Pals, S. L., Holland, J. M., Abramowitz, J. S., … Watson, J. (2011). Intraclass correlation associated with therapists: estimates and applications in planning psychotherapy research. *Cognitive Behaviour Therapy*, *40*(1), 15–33. https://doi.org/10.1080/16506073.2010.520731

Barker, C., & Pistrang, N. (2002). Psychotherapy and social support: integrating research on psychological helping. *Clinical Psychology Review*, *22*(3), 363–381.

Berger, T. (2017). The therapeutic alliance in internet interventions: A narrative review and suggestions for future research. *Psychotherapy Research*, *27*(5), 511–524. https://doi.org/10.1080/10503307.2015.1119908

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.

Blonski, S. C., Conradi, H. J., Oldehinkel, A. J., Bos, E. H., & de Jonge, P. (2016). Associations Between Negative and Positive Life Events and the Course of Depression: A Detailed Repeated-Assessments Study. *Journal of Nervous*, *204*(3), 175–180. https://doi.org/10.1097/NMD.0000000000000445

Boettcher, J., Renneberg, B., & Berger, T. (2013). Patient Expectations in Internet-Based Self-Help for Social Anxiety. *Cognitive Behaviour Therapy*, *42*(3), 203–214. https://doi.org/10.1080/16506073.2012.759615

Bower, P., Kontopantelis, E., Sutton, A., Kendrick, T., Richards, D. A., Gilbody, S., … Liu, E. T.-H. (2013). Influence of initial severity of depression on effectiveness of low intensity interventions: meta-analysis of individual patient data. *British Medical Journal*, *346*, 540–551. https://doi.org/10.1136/bmj.f540

Bull, C. F., Almond, T., Christensen, H., & Fenech, M. F. (2014). 29. Perceived stress and poor nutrition in individuals caring for family members with dementia, and controls, is associated

with elevated serum cortisol and increased chromosomal instability. *Brain, Behavior, and Immunity*, *40*, e9. https://doi.org/10.1016/j.bbi.2014.06.049

Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Christensen, H., Griffiths, K. M., Mackinnon, A. J., & Brittliffe, K. (2006). Online randomized controlled trial of brief and full cognitive behaviour therapy for depression. *Psychological Medicine*, *36*(12), 1737–1746. https://doi.org/10.1017/S0033291706008695

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, *6*(4), 330–351. https://doi.org/10.1037/1082-989X.6.4.330

Constantino, M. J., Arnkoff, D. B., Glass, C. R., Ametrano, R. M., & Smith, J. Z. (2011). Expectations. *Journal of Clinical Psychology*, *67*(2), 184–192. https://doi.org/10.1002/jclp.20754

Constantino, M. J., Coyne, A. E., Luukko, E. K., Newkirk, K., Bernecker, S. L., Ravitz, P., & McBride, C. (2017). Therapeutic Alliance, Subsequent Change, and Moderators of the Alliance-Outcome Association in Interpersonal Psychotherapy for Depression. *Psychotherapy*, *54*(2), 125–135. https://doi.org/10.1037/pst0000101

Cuijpers, P., Driessen, E., Hollon, S. D., van Oppen, P., Barth, J., & Andersson, G. (2012). The efficacy of non-directive supportive therapy for adult depression: A meta-analysis. *Clinical Psychology Review*, *32*(4), 280–291. https://doi.org/10.1016/j.cpr.2012.01.003

Dennis, A. R., Fuller, R. M., & Valacich, J. S. (2008). Media, tasks, and communication processes: A theory of media synchronicity. *MIS Quarterly*, *32*(3), 575–600.

Donkin, L., Hickie, I. B., Christensen, H., Naismith, S. L., Neal, B., Cockayne, N. L., & Glozier, N. (2013). Rethinking the Dose-Response Relationship Between Usage and Outcome in an Online Intervention for Depression: Randomized Controlled Trial. *Journal of Medical Internet Research*, *15*(10). https://doi.org/10.2196/jmir.2771

El Alaoui, S., Hedman, E., Kaldo, V., Hesser, H., Kraepelien, M., Andersson, E., … Lindefors, N. (2015). Effectiveness of Internet-Based Cognitive-Behavior Therapy for Social Anxiety Disorder in Clinical Psychiatry. *Journal of Consulting and Clinical Psychology*, *83*(5), 902–914. https://doi.org/10.1037/a0039198

El Alaoui, S., Ljotsson, B., Hedman, E., Svanborg, C., Kaldo, V., & Lindefors, N. (2016). Predicting Outcome in Internet-Based Cognitive Behaviour Therapy for Major Depression: A Large Cohort Study of Adult Patients in Routine Psychiatric Care. *Plos One*, *11*(9), e0161191. https://doi.org/10.1371/journal.pone.0161191

Falkenstrom, F., Ekeblad, A., & Holmqvist, R. (2016). Improvement of the Working Alliance in One Treatment Session Predicts Improvement of Depressive Symptoms by the Next Session. *Journal of Consulting and Clinical Psychology*, *84*(8), 738–751. https://doi.org/10.1037/ccp0000119

Falkenström, F., Hatcher, R. L., & Holmqvist, R. (2015). Confirmatory Factor Analysis of the Patient Version of the Working Alliance Inventory--Short Form Revised. *Assessment*, *22*(5), 581–593. https://doi.org/10.1177/1073191114552472

Falkenström, F., Hatcher, R. L., Skjulsvik, T., Larsson, M. H., & Holmqvist, R. (2015). Development and validation of a 6-item working alliance questionnaire for repeated administrations during psychotherapy. *Psychological Assessment*, *27*(1), 169–183. https://doi.org/10.1037/pas0000038

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, *55*(4), 316–340. https://doi.org/10.1037/pst0000172

Hatcher, R. L., & Gillaspy, J. A. (2006). Development and validation of a revised short version of the working alliance inventory. *Psychotherapy Research*, *16*(1), 12–25. https://doi.org/10.1080/10503300500352500

Hautzinger, M., Keller, F., & Kühner, C. (2006). *Beck-Depressionsinventar - BDI-II*. Göttingen, Germany: Hogrefe.

Hedman, E., Ljotsson, B., & Lindefors, N. (2012). Cognitive behavior therapy via the Internet: a systematic review of applications, clinical efficacy and cost-effectiveness. *Expert Review of Pharmacoeconomics & Outcomes Research*, *12*(6), 745–764. https://doi.org/10.1586/ERP.12.67

Holländare, F., Gustafsson, S. A., Berglind, M., Grape, F., Carlbring, P., Andersson, G., … Tillfors, M. (2016). Therapist behaviours in internet-based cognitive behaviour therapy (ICBT) for depressive symptoms. *Internet Interventions*, *3*(Supplement C), 1–7. https://doi.org/10.1016/j.invent.2015.11.002

Johansson, R., & Andersson, G. (2012). Internet-based psychological treatments for depression. *Expert Review of Neurotherapeutics*, *12*(7), 861–870. https://doi.org/10.1586/ern.12.63

Karyotaki, E., Riper, H., Twisk, J., Hoogendoorn, A., Kleiboer, A., Mira, A., … Cuijpers, P. (2017). Efficacy of Self-guided Internet-Based Cognitive Behavioral Therapy in the Treatment of Depressive Symptoms A Meta-analysis of Individual Participant Data. *Jama Psychiatry*, *74*(4), 351–359. https://doi.org/10.1001/jamapsychiatry.2017.0044

Kazdin, A. E. (2005). Treatment outcomes, common factors, and continued neglect of mechanisms of change. *Clinical Psychology-Science and Practice*, *12*(2), 184–188. https://doi.org/10.1093/clipsy/bpi023

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. In *Annual Review of Clinical Psychology* (Vol. 3, pp. 1–27). Palo Alto: Annual Reviews.

Knox, S., Adrians, N., Everson, E., Hess, S., Hill, C., & Crook-Lyon, R. (2011). Clients' perspectives on therapy termination. *Psychotherapy Research: Journal of the Society for Psychotherapy Research*, *21*(2), 154–167. https://doi.org/10.1080/10503307.2010.534509

Königbauer, J., Letsch, J., Doebler, P., Ebert, D., & Baumeister, H. (2017). Internet- and mobile-based depression interventions for people with diagnosed depression: A systematic review and meta-analysis. *Journal of Affective Disorders*, *223*, 28–40. https://doi.org/10.1016/j.jad.2017.07.021

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and Moderators of Treatment Effects in Randomized Clinical Trials. *Archives of General Psychiatry*, *59*(10), 877–883. https://doi.org/10.1001/archpsyc.59.10.877

Lambert, M. J. (1992). Psychotherapy outcome research: Implications for integrative and eclectical therapists. In Norcross, J. C. & Goldfried, M. R. (Eds.), *Handbook of psychotherapy integration* (pp. 94–129). New York: Basic Books.

Lemmens, L. H. J. M., Müller, V. N. L. S., Arntz, A., & Huibers, M. J. H. (2016). Mechanisms of change in psychotherapy for depression: An empirical update and evaluation of research aimed at identifying psychological mediators. *Clinical Psychology Review*, *50*, 95–107. https://doi.org/10.1016/j.cpr.2016.09.004

Lin, C.-Y., Updegraff, J. A., & Pakpour, A. H. (2016). The relationship between the theory of planned behavior and medication adherence in patients with epilepsy. *Epilepsy & Behavior: E&B*, *61*, 231–236. https://doi.org/10.1016/j.yebeh.2016.05.030

Ljótsson, B., Hesser, H., Andersson, E., Lackner, J. M., El Alaoui, S., Falk, L., … Hedman, E. (2014). Provoking symptoms to relieve symptoms: A randomized controlled dismantling study of exposure therapy in irritable bowel syndrome. *Behaviour Research and Therapy*, *55*(Supplement C), 27–39. https://doi.org/10.1016/j.brat.2014.01.007

Löwe, B., Spitzer, R. L., Zipfel, S., & Herzog, W. (2003). *Gesundheitsfragebogen für Patienten (PHQ). Komplettversion und Kurzform. [Patient Health Questionnaire (PHQ-D). Complete and Shortened Version.]* (2nd ed.). Karlsruhe, Germany: Pfizer.

Marcus, D. K., O'Connell, D., Norris, A. L., & Sawaqdeh, A. (2014). Is the Dodo bird endangered in the 21st century? A meta-analysis of treatment comparison studies. *Clinical Psychology Review*, *34*(7), 519–530. https://doi.org/10.1016/j.cpr.2014.08.001

Mausbach, B. T., Moore, R. C., Davine, T., Cardenas, V., Bowie, C. R., Ho, J., … Patterson, T. L. (2013). The use of the Theory of Planned Behavior to Predict Engagement in Functional Behaviors in Schizophrenia. *Psychiatry Research*, *205*(1–2), 36–42. https://doi.org/10.1016/j.psychres.2012.09.016

Mendez, R. D. R., Rodrigues, R. C. M., Cornélio, M. E., Gallani, M. C. B. J., & Godin, G. (2010). Development of an instrument to measure psychosocial determinants of physical activity behavior among coronary heart disease patients. *Revista Da Escola de Enfermagem Da USP*, *44*(3), 584–596. https://doi.org/10.1590/S0080-62342010000300006

Miller, S. D., Duncan, B. L., & Hubble, M. A. (1996). *Escape from Babel: Toward a Unifying Language for Psychotherapy Practice* (1 edition). New York: W. W. Norton & Company.

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York, NY, USA: Routledge.

Munder, T., Wilmers, F., Leonhart, R., Linster, H. W., & Barth, J. (2010). Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. *Clinical Psychology & Psychotherapy*, *17*(3), 231–239. https://doi.org/10.1002/cpp.658

Nordgreen, T., Havik, O. E., Ost, L. G., Furmark, T., Carlbring, P., & Andersson, G. (2012). Outcome predictors in guided and unguided self-help for social anxiety disorder. *Behaviour Research and Therapy*, *50*(1), 13–21. https://doi.org/10.1016/j.brat.2011.10.009

Patterson, C. L., Anderson, T., & Wei, C. (2014). Clients' Pretreatment Role Expectations, the Therapeutic Alliance, and Clinical Outcomes in Outpatient Therapy. *Journal of Clinical Psychology*, *70*(7), 673–680. https://doi.org/10.1002/jclp.22054

Pizzagalli, D. A. (2014). Depression, Stress, and Anhedonia: Toward a Synthesis and Integrated Model. In T. D. Cannon & T. Widiger (Eds.), *Annual Review of Clinical Psychology, Vol 10* (Vol. 10, pp. 393–423). Palo Alto: Annual Reviews.

Richards, D., & Richardson, T. (2012). Computer-based psychological treatments for depression: A systematic review and meta-analysis. *Clinical Psychology Review*, *32*(4), 329–342. https://doi.org/10.1016/j.cpr.2012.02.004

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. https://doi.org/10.1016/j.invent.2014.02.001

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research Online*, *8*(8), 23–74.

Schneider, A. J., Mataix-Cols, D., Marks, I. M., & Bachofen, M. (2005). Internet-guided self-help with or without exposure therapy for phobic and panic disorders - A randomised controlled trial. *Psychotherapy and Psychosomatics*, *74*(3), 154–164. https://doi.org/10.1159/000084000

Sieverink, F., Kelders, S., Poel, M., & Gemert-Pijnen, L. van. (2017). Opening the Black Box of Electronic Health: Collecting, Analyzing, and Interpreting Log Data. *JMIR Research Protocols*, *6*(8), e156. https://doi.org/10.2196/resprot.6452

Smith, C. E., Fernengel, K., Holcroft, C., Gerald, K., & Marien, L. (1994). Meta-Analysis of the Associations between Social Support and Health Outcomes. *Annals of Behavioral Medicine*, *16*(4), 352–362. https://doi.org/10.1093/abm/16.4.352

Steyer, R., Eid, M., & Schwenkmezger, P. (1997). Modeling true intraindividual change: True change as a latent variable. *Methods of Psychological Research*, *2*, 21–33.

Ursache, A., Noble, K. G., & Blair, C. (2015). Socioeconomic Status, Subjective Social Status, and Perceived Stress: Associations with Stress Physiology and Executive Functioning. *Behavioral Medicine (Washington, D.C.)*, *41*(3), 145–154. https://doi.org/10.1080/08964289.2015.1024604

van Buuren, S., & Groothuis-Oudshoorn, C. G. M. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry*, *14*(3), 270–277. https://doi.org/10.1002/wps.20238

Whitton, A. E., Proudfoot, J., Clarke, J., Birch, M.-R., Parker, G., Manicavasagar, V., & Hadzi-Pavlovic, D. (2015). Breaking Open the Black Box: Isolating the Most Potent Features of a Web and Mobile Phone-Based Intervention for Depression, Anxiety, and Stress. *JMIR Mental Health*, *2*(1). https://doi.org/10.2196/mental.3573

WHO (Ed.). (2008). *The Global Burden of Disease: 2004 Update*. World Health Organization.

Wilcox, R. R. (2003). *Applying Contemporary Statistical Techniques*. San Diego, CA, United States of America: Elsevier.

Zilcha-Mano, S., Dinger, U., McCarthy, K. S., & Barber, J. P. (2014). Does Alliance Predict Symptoms Throughout Treatment, or Is It the Other Way Around? *Journal of Consulting and Clinical Psychology*, *82*(6), 931–935. https://doi.org/10.1037/a0035141

Table 1
Socio-demographic and clinical sample characteristics

Variable		Total	IF	SF	χ^2	*p*
socio-demographic characteristics						
age [a]	*M* (*SD*)	45.7 (11.3)	45.7 (11.8)	45.8 (10.7)	0.166 [b]	.868
female gender	*n* (%)	714 (65.6)	360 (64.9)	354 (66.3)	0.246	.620
education					5.673	.339
no certificate	*n* (%)	7 (0.6)	4 (0.7)	3 (0.6)		
lower secondary	*n* (%)	66 (6.1)	34 (6.1)	32 (6.0)		
secondary school	*n* (%)	263 (24.2)	122 (22.0)	141 (26.4)		
trade school	*n* (%)	245 (22.5)	118 (21.3)	127 (23.8)		
college	*n* (%)	162 (14.9)	88 (15.9)	74 (13.9)		
university	*n* (%)	346 (31.8)	189 (34.1)	157 (29.4)		
marital status					1.354	.716
single	*n* (%)	373 (34.3)	196 (35.3)	177 (33.1)		
married	*n* (%)	561 (51.5)	286 (51.5)	275 (51.5)		
divorced	*n* (%)	141 (12.9)	66 (11.9)	75 (14.0)		
widowed	*n* (%)	14 (1.3)	7 (1.3)	7 (1.3)		
clinical baseline characteristics						
PHQ-9	*M* (*SD*)	11.8 (3.4)	11.9 (13.4)	11.7 (3.5)	1.044 [b]	.297
SCID diagnosis					3.194	.670
current major depression	*n* (%)	458 (42.1)	247 (44.5)	211 (39.5)		
remitted depressive episode	*n* (%)	285 (26.2)	137 (24.7)	148 (27.7)		
dysthymic disorder	*n* (%)	90 (8.3)	46 (8.3)	44 (8.2)		
double-depression	*n* (%)	58 (5.3)	27 (4.9)	31 (5.8)		
bipolar or NOS [c]	*n* (%)	65 (6.0)	32 (5.8)	33 (6.2)		
no current/past affective disorder	*n* (%)	133 (12.2)	66 (11.9)	67 (12.5)		
Expectations [a]	*M* (*SD*)	10.0 (4.0)	9.9 (4.0)	10.1 (4.0)	0.462 [b]	.644
PHQ-S [a]	*M* (*SD*)	9.5 (3.2)	9.5 (3.2)	9.5 (3.2)	0.144 [b]	.886

Note. $N = 1089$. IF = Individualized Feedback ($n = 555$); SF = Standardized Feedback ($n = 534$); NOS = not otherwise specified. [a] variables have some missing values: age $n = 1081$; Expectations $n = 1081$; PHQ-S: $n = 1067$. [b] *t*-test for independent samples. [c] individuals with bipolar disorders were included only if they were not experiencing current mania/hypomania

Table 2

Regressive Paths for Each Treatment Condition

Path - Group IF		*b*	[95% CI]	*p*	b_{stdxy} *(SE)*	R^2
ΔDEP_{M3-M1} ON	$\Delta STRESS_{Post-Pre}$	0.156	[0.044, 0.268]	.006**	0.200 (0.057)	.05
ΔDEP_{M5-M3} ON	ΔDEP_{M3-M1}	-0.177	[-0.346, -0.008]	.040*	-0.172 (0.086)	.09
	$\Delta STRESS_{Post-Pre}$	0.168	[0.048, 0.288]	.006**	0.208 (0.061)	
ΔDEP_{M7-M5} ON	ΔDEP_{M5-M3}	-0.264	[-0.446, -0.082]	.005**	-0.312 (0.093)	.17
	$\Delta STRESS_{Post-Pre}$	0.196	[0.094, 0.298]	< .001***	0.289 (0.052)	
$\Delta DEP_{POST-M7}$ ON	ΔDEP_{M7-M5}	-0.423	[-0.599, -0.247]	< .001***	-0.439 (0.090)	.10
	$\Delta STRESS_{Post-Pre}$	0.133	[0.031, 0.235]	.010*	0.203 (0.052)	
	$TASK_{M5}$	-0.089	[-0.136, -0.042]	< .001***	-0.307 (0.024)	
	$BOND_{M5}$	0.051	[0.010, 0.092]	.014*	0.188 (0.021)	
$BOND_{M5}$ ON	EXP_{PRE}	0.303	[0.148, 0.458]	< .001***	0.226 (0.079)	.06
$TASK_{M5}$ ON	EXP_{PRE}	0.278	[0.141, 0.415]	< .001***	0.221 (0.070)	.17
	DEP_{M1}	-0.381	[-0.746, -0.016]	.041*	-0.161 (0.186)	
	ΔDEP_{M3-M1}	-0.749	[-1.186, -0.312]	.001***	-0.258 (0.223)	
	ΔDEP_{M5-M3}	-0.693	[-1.063, -0.323]	< .001***	-0.246 (0.189)	
Path - Group SF		*b*	[95% CI]	*p*	b_{stdxy} *(SE)*	
ΔDEP_{M3-M1} ON	$\Delta STRESS_{Post-Pre}$	0.124	[0.018, 0.230]	.022*	0.186 (0.054)	.04
ΔDEP_{M5-M3} ON	ΔDEP_{M3-M1}	-0.257	[-0.500, -0.014]	.038*	-0.184 (0.124	.12
	$STRESS_{PRE}$	0.148	[0.013, 0.283]	.032*	0.141 (0.069)	
	$\Delta STRESS_{Post-Pre}$	0.184	[0.039, 0.329]	.013*	0.198 (0.074)	
ΔDEP_{M7-M5} ON	ΔDEP_{M5-M3}	-0.294	[-0.455, -0.133]	< .001***	-0.363 (0.082)	.17
$\Delta DEP_{POST-M7}$ ON	ΔDEP_{M7-M5}	-0.242	[-0.414, -0.070]	.006**	-0.290 (0.088)	.27
	$STRESS_{PRE}$	0.157	[0.057, 0.257]	.002**	0.221 (0.051)	
	$\Delta STRESS_{Post-Pre}$	0.244	[0.150, 0.338]	< .001***	0.388 (0.048)	
$BOND_{M5}$ ON	ΔDEP_{M5-M3}	-0.594	[-1.074, -0.114]	.015*	-0.190 (0.245)	.08
$TASK_{M5}$ ON	EXP_{PRE}	0.224	[0.104, 0.344]	< .001***	0.192 (0.061)	.22
	DEP_{M1}	-0.530	[-0.857, -0.203]	.002**	-0.246 (0.167)	
	ΔDEP_{M3-M1}	-0.560	[-1.077, -0.043]	.034*	-0.180 (0.264)	
	ΔDEP_{M5-M3}	-0.918	[-1.232, -0.604]	< .001***	-0.410 (0.160)	

Note. Only paths with $p < .050$ are shown. For all estimated paths see supplementary Table B.3. * $p < .050$; ** $p < .010$; *** $p < .001$. b /b_{stdxy} = un-/standardized regression weight; DEP = PHQ-9 depression score; STRESS = PHQ-Stress score; EXP = expectations; BOND = bond-component of the Working Alliance Inventory; TASK = task-component of the Working Alliance Inventory. Measurement occasions are indexed.

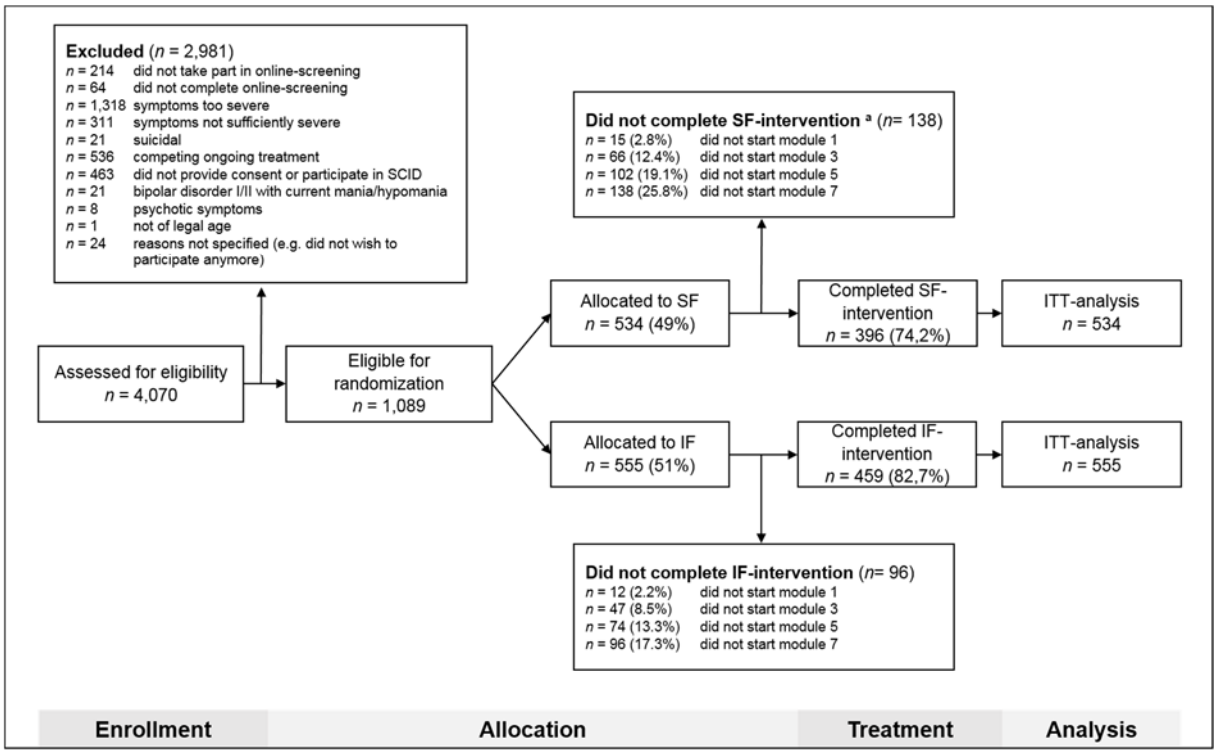


Figure 1. Flowchart in accordance with CONSORT guidelines. SF = standardized feedback condition; IF = individualized feedback condition; SCID = structured clinical interview (section for affective and psychotic disorders).

[a] Completers were defined as participants who at least started with the respective treatment module.

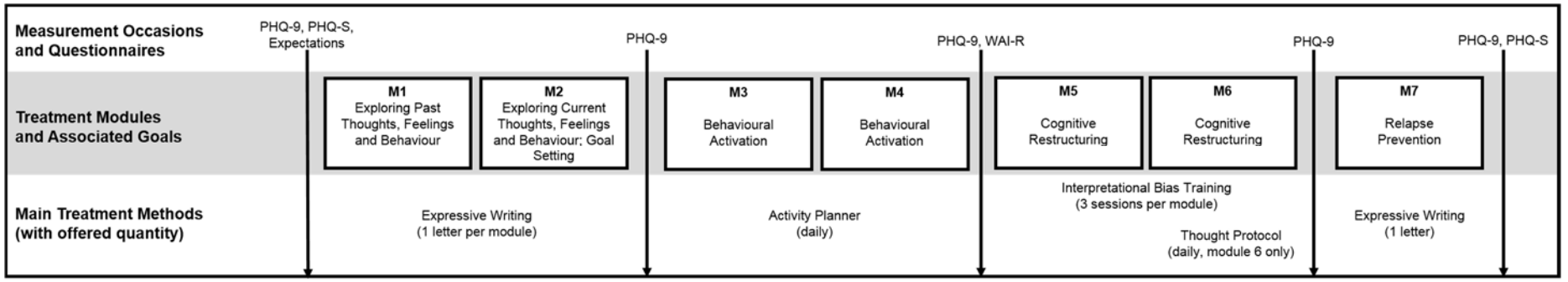


Figure 2. Treatment goals and methods of the intervention as well as questionnaires and their measurement occasions included in the analyses. PHQ-9 = Patient-Health Questionnaire – Subscale assessing depressive symptoms; PHQ-S = Patient-Health Questionnaire – Subscale assessing psychosocial stressors; WAI-R = Revised Working Alliance Inventory.

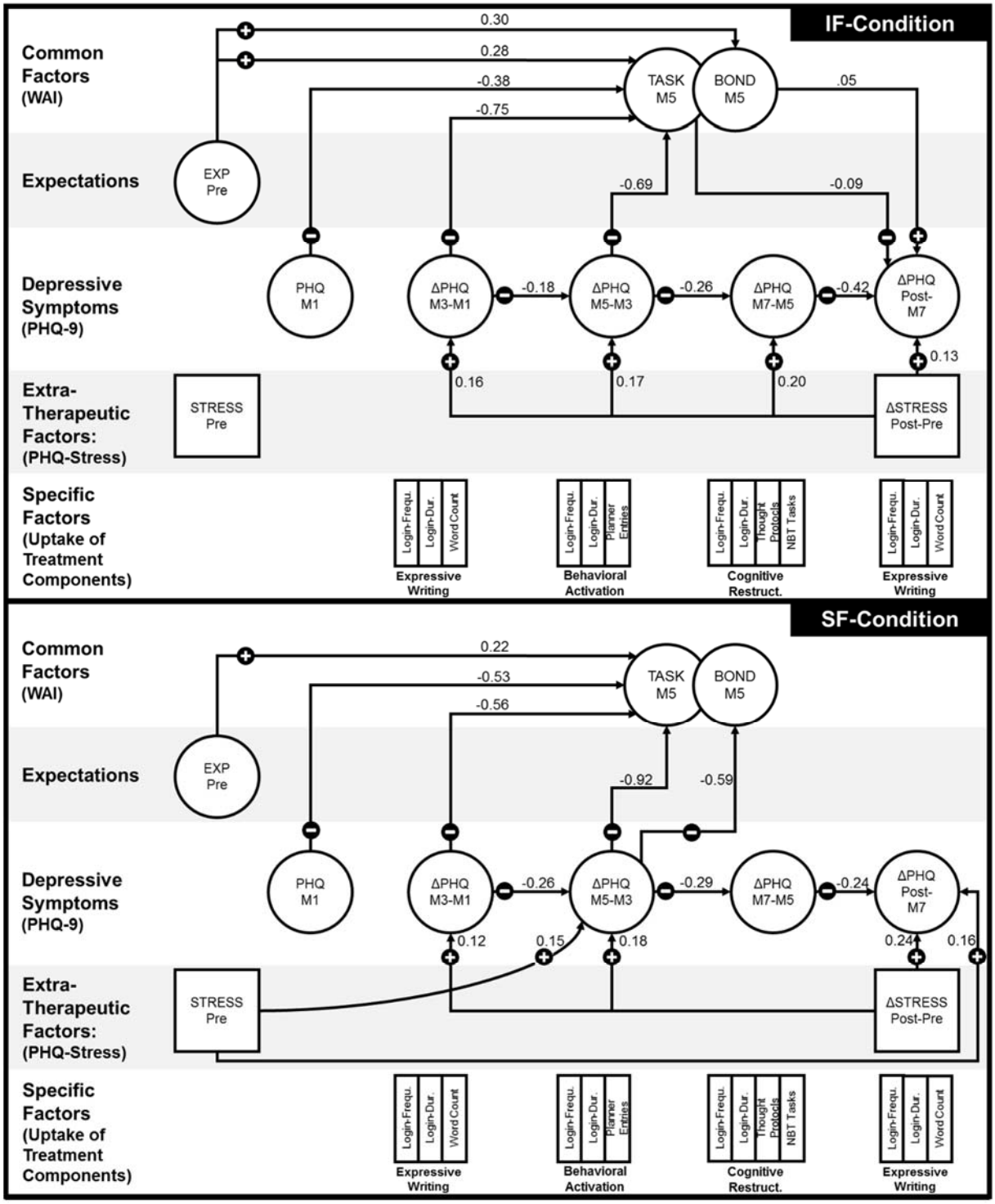


Figure 3. Visual depiction of significant paths in the model based on multiple imputations with respective unstandardized regression weights. SF = standardized feedback condition; IF = individualized feedback condition; PHQ-9 = Patient-Health Questionnaire – Subscale assessing depressive symptoms; PHQ-Stress = Patient-Health Questionnaire – Subscale assessing psychosocial stressors; EXP = expectations; M1 through M7 = measurement occasions prior to respective modules; Δ = difference scores; "+" and "-" illustrate positive and negative regression weights, respectively.

Supplementary materials

The following supplementary materials related to this article are available:

- *Appendix A*. Aid with interpreting change scores.
- *Appendix B*. Supplementary results on model comparisons with different measurement invariance restriction (Table B1), results on general mean change (Table B2) and results on non-significant regression weights (Table B3).

CHAPTER 4

HOW INDIVIDUALS CHANGE DURING INTERNET-BASED INTERVENTIONS FOR DEPRESSION: A RANDOMIZED CONTROLLED TRIAL COMPARING STANDARDIZED AND INDIVIDUALIZED FEEDBACK.

A revised version of the following paper was published in the journal *Brain and Behavior*:

Zagorscak P, Heinrich M, Bohn J, Stein J, Knaevelsrud C. How individuals change during internet-based interventions for depression: A randomized controlled trial comparing standardized and individualized feedback. Brain Behav. 2020;10:e01484. https://doi.org/10.1002/brb3.1484

How Individuals Change During Internet-Based Interventions for Depression: A Randomized Controlled Trial Comparing Standardized and Individualized Feedback.

Pavle Zagorscak, Manuel Heinrich, Johannes Bohn, Jana Stein and Christine Knaevelsrud
Freie Universität Berlin

Author Note

Pavle Zagorscak, M.Sc., Department of Clinical-Psychological Intervention, Freie Universität Berlin; Manuel Heinrich, M.Sc., Department of Clinical-Psychological Intervention, Freie Universität Berlin; Johannes Bohn, M.Sc., Department of Clinical-Psychological Intervention, Freie Universität Berlin; Jana Stein, M.Sc., Department of Clinical-Psychological Intervention, Freie Universität Berlin; Christine Knaevelsrud, Ph.D., Department of Clinical-Psychological Intervention, Freie Universität Berlin.

Correspondence

Correspondence concerning this article should be addressed to Pavle Zagorscak, Department of Clinical-Psychological Intervention, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany. E-mail: pav.zag@fu-berlin.de

Acknowledgements

The German public health insurance company “Techniker Krankenkasse” funded this trial.

Abstract

Background: Standardized and individualized internet-based interventions (IBI) for depression yield significant symptom improvements. However, change patterns during standardized or individualized IBI are unknown. Identifying subgroups that experience different symptom courses during IBI and their characteristics is vital for improving response.

Methods: Mildly to moderately depressed individuals according to self-report ($N = 1089$) were randomized to receive module-wise feedback that was either standardized or individualized by a counselor within an otherwise identical cognitive-behavioral IBI for depression (seven modules over six weeks). Depressive symptoms were assessed at baseline and before each module (Patient-Health-Questionnaire; PHQ-9). Other individual characteristics (self-report) and presence of an affective disorder (structured clinical interview) were assessed at baseline. Growth-mixture modeling was used to identify and compare subgroups with discernable change patterns and associated client variables across conditions.

Results: Model comparisons suggest equal change patterns in both conditions. Across conditions a group of immediate (62.5%) and a group of delayed improvers (37.4%) were identified. Immediate improvers decreased their PHQ-9 score by 5.5 points from pre to post, with 33% of improvement occurring before treatment commenced. Delayed improvers were characterized by stable symptom severity during the first two modules and smaller overall symptom decrease (3.2 points). Higher treatment expectations, a current major depressive disorder (interview) and lower social support were associated with delayed improvement.

Conclusion: IBI for depression with individualized and with standardized feedback lead to comparable patterns of change. Expectation management and bolstering of social support are promising strategies for individuals that are at risk for delayed improvement.

Keywords: depression; patterns of change; internet-based interventions; growth mixture modeling; social support; expectations

How Individuals Change During Internet-Based Interventions for Depression: A Randomized Controlled Trial Comparing Standardized and Individualized Feedback

The World Health Organization (2017) identified depression as the leading cause of disability worldwide. Even in high-income countries, only one in five depressed individuals receives adequate treatment (Thornicroft et al., 2017). Different researchers (e.g., Kazdin, 2018) proposed Internet-based interventions (IBI) as one approach to circumvent individual-level barriers like problems with transportation, inconvenient treatment hours and locations (i.e., long distance from home) and fear of stigma that impede the uptake of evidence-based treatments (Harvey & Gumport, 2015). Meta-analyses confirm the efficacy of standardized and individualized IBI for depression (e.g., Karyotaki et al., 2018, 2017). However, research on why, when and how individuals improve throughout IBI with varying levels of individualization is lacking.

Identifying individuals who improve during IBI and their sociodemographic and clinical characteristics is a prerequisite for offering interventions that are tailored to the needs of specific populations and thus might increase response rates (Khan, Faucett, Lichtenberg, Kirsch, & Brown, 2012; Manen et al., 2015; Mueller et al., 2018). Moreover, learning about the particular point during treatment (and the associated intervention elements) at which certain individuals change is essential to advance the understanding of the underlying mechanisms of change (Klein & Kotov, 2016; Silberschatz, 2015). Growth mixture modeling (GMM) is a statistical approach that addresses these questions. It explores whether populations with heterogeneous symptom trajectories contain distinct homogenous subgroups (e.g., Jung & Wickrama, 2008; B. Muthén, 2006).

While GMM has been regularly used to investigate depressive symptom courses during face-to-face psychotherapy (e.g., Rubel, Lutz, & Schulte, 2015), there has only been a limited number of trials on this topic in IBI for depression (Batterham et al., 2018; Lutz et al., 2017; Sunderland, Wong, Hilvert-Bruce, & Andrews, 2012). Sunderland et al. (2012) and

Batterham et al. (2018) found two discernable symptom trajectories during IBI, with 75 to 81% of individuals showing improvement and the remainder showing no or low symptom improvements. Divergently, Lutz et al. (2017) found three distinct groups of depressed individuals. One group improved immediately after baseline assessment (45%), another after being randomized to the intervention and registered on the website (39%), and a third showed early symptom deterioration (16%). The differing number of identified subgroups might be due to significant differences in study design and interventions under research. Lutz et al. (2017) focused exclusively on symptom change during the first four weeks of IBI, while Sunderland et al. (2012) and Batterham et al.(2018) aimed to explore the heterogeneity in symptom trajectories beyond the early stages of treatment. In addition, the studies differed with regard to the provided intervention. While Lutz et al. (2017) and Sunderland et al. (2012) focused on individuals with depression and anxiety, Batterham et al. (2018) treated depressive symptom load as secondary outcome in an intervention focusing on reducing suicidal thoughts. Another critical difference between the three studies pertains to the level of individualization offered. Batterham et al. (2018) and Sunderland et al. (2012) evaluated a self-guided treatment (i.e., standardized, without regular guidance or feedback by clinicians) and Lutz et al. (2017) provided more severely depressed individuals with additional guidance (individualized weekly e-mail support).

Since the intensity of guidance is considered to be one of the most central moderators of outcome in IBI for depression (e.g., Johansson & Andersson, 2012), more research is necessary to assess the influence of contact quantity and quality on patterns of change. Consequently, the current study investigates depressive symptom courses and their associations with pre-interventional client characteristics in an individualized form (IF-condition: feedback individualized by a counselor; contact-on-demand) and a standardized form (SF-condition: standardized feedback; contact-on-demand) of the same IBI for depression within a randomized controlled trial. To our knowledge, this is the first study exploring a) if individualization of

feedback leads to quantitatively and qualitatively different patterns of change and b) if these change patterns show diverging associations with participants' characteristics in a large clinical sample of adults provided with IBI for depression.

Method

Design and Sample

Clients in this two-arm assessor-blind randomized controlled trial were recruited nationwide in Germany between March 2014 and March 2015 from the client-base of a German public health care provider. The trial was prospectively registered (URL https://www.anzctr.org.au (ID: ACTRN12614000312640) and approved by the Research Ethics Committee of Freie Universität Berlin. Informed consent was obtained from all clients.

Only non-suicidal individuals with mild to moderate depression (Beck Depression Inventory-II score between 14 and 28; score ≤ 1 on suicide item) were included. Individuals with current mania, hypomania or psychosis as assessed during a telephone-administered structured clinical interview for DSM IV (SCID-I, sections A through F; Wittchen, Zaudig, & Fydrich, 1997) were excluded. A previous publication comparing the efficacy of the two treatment arms describes the recruitment strategy in more detail (Zagorscak, Heinrich, Sommer, Wagner, & Knaevelsrud, 2018). Overall, $N = 1089$ individuals participated in the intervention. The mean age of the sample was 45.7 ($SD = 11.3$) years; 65.6% were female. A majority of individuals was married (51.5%), employed (88.2%) and highly educated (69.2% finished college-preparatory school). Table 1 displays baseline sample characteristics. No significant differences between study conditions were found regarding any clinical or sociodemographic variables (all p-values > .05).

Treatment

Clients were randomly assigned to one of two variants of an IBI for depression. Both conditions offered the same psychoeducation and intervention tools in seven modules (M1-M7). In particular, clients completed two expressive writing tasks (M1-M2, one week),

behavioral activation through a daily planner (M3-M4, two weeks), cognitive restructuring through thought protocols and interpretational bias training (M5-M6, two weeks) as well as relapse prevention (M7, one week). Clients received either standardized feedback (n_{SF} = 534) or feedback individualized by a counselor (40 counselors, 21 holding a bachelor's degree and 19 holding a master's degree in psychology, n_{IF} = 555). Feedback was offered via written messages within a password-protected internet platform after completion of each module. Clients in both intervention groups could receive contact on demand in case of technical problems or specific questions concerning the intervention.

Measures

Categorical diagnoses of affective disorders (e.g., current or past major depressive disorder (MDD), current dysthymia) were obtained from telephone-administered structured clinical interviews (SCID-I, sections A through F; Wittchen et al., 1997).

Depressive symptom burden was assessed with the nine-item Patient Health Questionnaire-9 (PHQ-9; Kroenke, Spitzer, & Williams, 2001).

The tendency for perseverative thinking was measured using the 15-item Perseverative Thinking Questionnaire (PTQ; Ehring et al., 2011).

Expectations were assessed with five seven-point semantic differentials (Mendez, Rodrigues, Cornélio, Gallani, & Godin, 2010). The original item wording was slightly adapted to address expectations in the specific IBI context (e.g., "For me, participation in the IBI during the next six weeks would be "beneficial" to "harmful").

Perceived social support was measured using the respective 8-item subscale of the Berlin-Social Support Scale (BSSS; Schulz & Schwarzer, 2003).

Several socio-demographic characteristics were assessed, that is, age and gender, level of education, employment, marital status and history of psychotherapeutic treatment.

Clients completed the PHQ-9 at baseline and after the completion of the intervention, as well as at the beginning of each week (at the beginning of M1, M3, M4, M5, M6, and M7). All other variables were assessed during baseline assessment only.

Statistical Analysis

Overall, the analysis aimed to identify subgroups of clients with different patterns of change in depressive symptoms as measured with the PHQ-9 in the IF- and SF-condition. The analysis had to consider that the two conditions might differ with regard to the number of change patterns and shape of the derived trajectories. GMM with latent base specifications was used for this purpose. The PHQ-9 measurement structured the change process. A detailed description of the modeling process is available in the supporting online information (Appendix C). In short, the modeling process comprised three steps: *First*, the optimal number of classes for each condition was determined separately using single-group GMM. *Second,* to test for potential differences in change trajectories between conditions, multi-group GMM was used. *Third*, potential predictors of class membership, initial symptom load and inter-individual differences in overall symptom change were included directly into the model (Asparouhov & Muthén, 2014). Model selection was based on information criteria (AIC, aBIC, BIC and CAIC). All models were estimated using M*Plus* 8.1 (L. K. Muthén & Muthén, 2017). Missing data were dealt with using FIML (depressive symptom load) and single-value imputation (predictor variables).

Results

Number of Trajectory Classes and Patterns of Change

Single-group GMM. The single-group analyses pointed towards a two-class solution in both intervention arms. Visual inspection indicated that the derived change patterns of both conditions showed considerable similarities. An illustration of the estimated change patterns of both classes together with estimated parameters separately for each intervention arm can be obtained from the supporting online information (Appendix C: Figure C1, Table C1).

Multi-group GMM. The multi-group analysis provided further evidence for the similarity of the derived classes across the two intervention conditions. All information criteria favored the more parsimonious model assuming no differences in change patterns between conditions (Appendix C: Table C2). This result supports the notion that the intervention conditions do not differ regarding the number of classes, class sizes and change trajectories. Therefore, class characterizations based on this constrained model are reported in the following.

Class 1 (delayed improvers) comprises 37.5% of all randomized individuals. The average trajectory was marked by a low decrease in depressive symptoms (average initial symptom-load: 12.4 points on the PHQ-9; average pre-to-post decrease by 3.4 points, see Figure 1). The growth factor loadings of the first two measurement occasions were not significant (λ = -0.13, p = .298 and λ = .18, p = .220) indicating a rather stable average symptom load during early stages of the intervention. In other words, delayed improvers showed no early change in reference to overall improvement. The residual variances (i.e., scatter of the observed variables around the predicted curves) of delayed improvers were relatively large throughout the intervention ranging from 5.5 points (post-treatment) to 9.1 points (baseline). The numbers indicate that the observed individual trajectories are marked by ups and downs scattered around the individually predicted curve.

Class 2 (immediate improvers) was the larger class and comprises 62.5% of the clients. The average symptom decrease in this class was larger than in Class 1 (5.5 points on the PHQ-9) while the average initial symptom-load was similar (11.2 points). In contrast to class 1, immediate improvers went through a significant proportion of their average symptom improvement immediately after the initial screening. The growth factor loadings indicate that 33% of the average overall improvement had already occurred before intervention commenced (Slope-loading at M1: λ = 0.33, p < .001). Immediate improvers showed the largest residual variances early during intervention ranging from 4.09 (M3) to 5.12 (pre-assessment). The residual variances decreased towards the end of the intervention ranging from 1.44 (M6) to 3.02

(post-assessment) indicating more stable symptom trajectories at this stage than in Class 1. Table 2 summarizes estimated parameters for both classes, the average change trajectories are illustrated in Figure 1.

Predictors of Class Membership and Symptom Course

Models were compared on the basis of information criteria. Results favored the use of a model that constrains the associations of predictors with initial symptom-load and with the amount of symptom improvement to be equal across conditions and classes. For detailed results on model comparisons see the supporting online information (Appendix C, Table C2).

Initial Level of Depressive Symptom Load across Classes. When compared to individuals who did not receive any diagnosis in the SCID-I, those who fulfilled the diagnostic criteria for MDD (SCID), b = 1.93, 95% CI [1.33, 2.52], $p < .001$, or double depression, b = 2.29, 95% CI [1.41, 3.18], $p < .001$ reported higher baseline depressive symptoms. Further, individuals with more severe perseverative thinking also reported higher baseline symptom severity, b = 0.07, 95% CI [0.05, 0.10], $p < .001$. In contrast, individuals with higher expectations towards the intervention showed lower initial depressive burden, b = -0.06, 95% CI [-0.11, -0.01], p = .012.

Amount of Overall Symptom Improvement across Classes. On average, larger depressive symptom improvements over the course of the intervention were reported by individuals who fulfilled the criteria for MDD, b = -1.26, 95% CI [-1.97, -0.56], $p < .001$, and by individuals who reported more severe perseverative thinking b = -0.04, 95% CI [-0.06, -0.01], p = .008. In contrast, unemployed individuals experienced less improvement throughout the intervention, when compared to employed individuals, b = 0.67, 95% CI [0.01, 1.34], p = .048.

Predictors of Class Membership. A current MDD diagnosis (SCID-I), expectations and perceived social support were statistically significant predictors of class membership. Individuals with higher initial expectations, OR = 1.09, 95% CI [1.01, 1.17], p = .020, and

individuals with a current MDD diagnosis, *OR* = 2.68, 95% CI [1.34, 5.32], $p = .005$, showed increased odds of being classified as *delayed improvers*. Individuals with higher perceived social support showed increased odds of being classified as *immediate improvers*, *OR* = 1.08, 95% CI [1.04, 1.12], $p < .001$. All predictor variables and their associations with slope, intercept and class membership are summarized in Table 3.

Discussion

The current study is the first to investigate and compare qualitatively and quantitatively discernable patterns of change in IBI for depression with varying levels of feedback-individualization.

Across conditions, the study identified two groups of individuals that showed distinct average change patterns. Nearly two-thirds of individuals randomized in the current trial belonged to an *immediate improver* class. Interestingly, this class size corresponds with response rates in previous studies on IBI for depression, which were summarized to range between 55% and 96% in a recent meta-analysis (Königbauer et al., 2017). The depressive symptom change in this class is characterized by significant improvements after the initial screening phase with 33% of overall improvement taking place before the beginning of the first treatment module. On average, this class improved by 5.5 PHQ-9 points overall, which is considered to be clinically significant change according to measure-specific conventions (Titov et al., 2011).

In contrast, individuals in the second class were *delayed improvers* (37.4% of the sample). Depressive symptom change in this class is characterized by smaller symptom improvement (by 3.2 PHQ-9 points) and by an initial treatment phase marked by stagnant symptom severity. These results complement the study by Lutz et al. (2017) in that they stress the importance of changes before and during early phases of IBI for depression and their association with overall outcome. Divergent from findings of the study at hand and two other studies on IBI for depression (Batterham et al., 2018; Sunderland et al., 2012), Lutz et al. (2017)

found three discernable classes. However, two of them differed with regard to change patterns during two separate periods prior to treatment (after screening and after registration on the study website). This distinction is not suitable for the present study design since screening and registration happened simultaneously.

Importantly, the results suggest that whether written feedback was individualized by a counselor or fully-standardized did not influence the number of discernable subgroups or associated change patterns in otherwise identical intervention arms. These results are consistent with a recent meta-analysis on the efficacy of IBI for individuals diagnosed with depression which did not find the presence of guidance to be a meaningful moderator of intervention success overall (Königbauer, Letsch, Doebler, Ebert, & Baumeister, 2017). Moreover, the current study extends the research by showing that not only the amount of pre- to post-changes is equal, but the average change patterns follow the same trajectories as well. Conversely, earlier meta-analyses on pre-post changes during IBI for depression found feedback quantity and quality to be an essential contributor to treatment success (e.g., Johansson & Andersson, 2012; Richards & Richardson, 2012). Here, it is important to note that the study at hand investigated module-wise change and change-associated subgroups between two treatment conditions, which only differed in the degree feedback was individualized. Our study thus represents an encouragement to use GMM for the investigation of change patterns across more dissimilar forms of contact in IBI (e.g., guidance by telephone vs. standardized written guidance), which might result in divergent conclusions.

Regarding individuals' characteristics associated with depressive symptoms and class membership, the results show that individuals who fulfill the criteria for MDD in a structured clinical interview show heightened baseline depressive symptom severity and larger improvement over time. That is not surprising, given that PHQ-9 items are derived from the DSM-IV criteria for depression (Kroenke et al., 2001). Furthermore, the finding is consistent with meta-analyses on psychotherapy for depressive patients highlighting that the expected

pre-post effect sizes (the amount of improvement) are lower for subclinical patients than those for individuals that fulfill the diagnostic criteria for MDD (Cuijpers, Karyotaki, et al., 2014; Cuijpers, Koole, et al., 2014). Interestingly, the presence of an MDD diagnosis is also associated with heightened odds of membership in the *delayed improver* class. In contrast to an individual that reports mild to moderate depressive symptoms on a questionnaire (PHQ-9) only, an individual that further fulfills all criteria for a current MDD diagnosis might have a more complex symptom and comorbidity profile that decreases the probability of fast response to treatment (Melchior et al., 2016). While perseverative thinking was not associated with class membership, individuals with high levels of perseverative thinking reported more severe depressive symptoms at baseline and increased improvement. This finding is in line with several studies highlighting the importance of perseverative thinking for the prediction of symptoms of anxiety and depression (e.g., Spinhoven, van Hemert, & Penninx, 2018).

Regarding sociodemographic and psychosocial variables, the results demonstrate that unemployed individuals report lower symptom improvement than employed individuals, which is in accordance with previous results on the relationship between socioeconomic risk factors and depressive symptoms (e.g., Arias-de la Torre, Vilagut, Martín, Molina, & Alonso, 2018). Moreover, individuals with higher perceived social support exhibit higher odds of being classified as an *immediate improver*. Apart from established cross-sectional associations of social support and depression (e.g., Gariépy, Honkaniemi, & Quesnel-Vallée, 2016), previous studies demonstrated, that individuals with low social support profit less from short-term treatments and might benefit from treatment extension (Lindfors, Ojanen, Jääskeläinen, & Knekt, 2014). These findings stress that providers of IBI might increase response rates by identifying individuals with low social support and either improve their access to social resources or offer more extended treatment.

Finally, higher expectations were associated with lower baseline scores, a finding congruent with previous studies on baseline expectation-symptom associations (e.g., Cohen,

Beard, & Björgvinsson, 2015). These correlations may be explained through hopelessness that increases with depressive symptom severity (e.g., Horwitz, Berona, Czyz, Yeguez, & King, 2017) and dampens expectations for improvement (through treatment). Higher expectations were further associated with membership in the delayed improver class. Given that the contents and procedures of IBIs are still mostly unknown to the public (Apolinário-Hagen, Vehreschild, & Alkoudmani, 2017), some clients may have unrealistic expectations towards treatment. As a consequence, initial disappointment might reduce the probability of experiencing rapid improvement (Greer, 1980). Overall, these findings stress the importance of assessing expectations in IBI for depression in order to react to expectations that might be either unrealistic or pessimistic. While a recent study highlighted that expectations might change during treatment (Vîslă, Flückiger, Constantino, Krieger, & Holtforth, 2018), there are no studies on how expectations develop through the course of IBI for depression. Thus, future studies should assess expectations at multiple time points to further explore the expectation symptom-course interplay.

Some other directions for future research can be derived from the limitations of the present study design. The findings of this study pertain to cognitive-behavioral IBI that utilizes written feedback (standardized vs. individualized) and includes mildly to moderately depressed individuals only. Patterns of change and associated individual characteristics might differ in other populations, in forms of treatment that apply other qualities or quantities of feedback or use treatment techniques that might entail other change trajectories (e.g., interpersonal or psychodynamic treatments). Furthermore, this study is limited to exploring change patterns derived using PHQ-9 sum scores. While this is standard in clinical research and practice, it might cover up relevant changes on the symptom level (i.e., cognitive symptom change during modules targeting cognitive restructuring). Thus, a fruitful direction for future studies would be more symptom-oriented modeling of depression and depression change (e.g., Heinrich, Zagorscak, Eid, & Knaevelsrud, 2018).

Conclusion

Individualizing feedback did not influence patterns of change when compared to standardized feedback, and a majority of clients showed immediate improvements in both treatment conditions. However, a smaller group was at risk of delayed and reduced improvements. Fruitful directions for clinicians aiming to increase improvements during IBI are expectation management, treatment extension and a bolstering of socially supportive relationships.

References

Apolinário-Hagen, J., Vehreschild, V., & Alkoudmani, R. M. (2017). Current Views and Perspectives on E-Mental Health: An Exploratory Survey Study for Understanding Public Attitudes Toward Internet-Based Psychotherapy in Germany. *JMIR Mental Health*, *4*(1), e8. https://doi.org/10.2196/mental.6375

Arias-de la Torre, J., Vilagut, G., Martín, V., Molina, A. J., & Alonso, J. (2018). Prevalence of major depressive disorder and association with personal and socio-economic factors. Results for Spain of the European Health Interview Survey 2014–2015. *Journal of Affective Disorders*, *239*, 203–207. https://doi.org/10.1016/j.jad.2018.06.051

Asparouhov, T., & Muthén, B. (2014). Auxiliary Variables in Mixture Modeling: Three-Step Approaches Using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 329–341. https://doi.org/10.1080/10705511.2014.915181

Batterham, P. J., Spijker, B. A. J. van, Mackinnon, A. J., Calear, A. L., Wong, Q., & Christensen, H. (2018). Consistency of trajectories of suicidal ideation and depression symptoms: Evidence from a randomized controlled trial. *Depression and Anxiety*. https://doi.org/10.1002/da.22863

Cohen, M., Beard, C., & Björgvinsson, T. (2015). Examining patient characteristics as predictors of patient beliefs about treatment credibility and expectancies for treatment outcome. *Journal of Psychotherapy Integration*, *25*(2), 90–99. https://doi.org/10.1037/a0038878

Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & van Straten, A. (2014). The effects of psychotherapies for major depression in adults on remission, recovery and improvement: a meta-analysis. *Journal of Affective Disorders*, *159*, 118–126. https://doi.org/10.1016/j.jad.2014.02.026

Cuijpers, P., Koole, S. L., Dijke, A. van, Roca, M., Li, J., & Reynolds, C. F. (2014). Psychotherapy for subclinical depression: meta-analysis. *The British Journal of Psychiatry*, *205*(4), 268–274. https://doi.org/10.1192/bjp.bp.113.138784

Ehring, T., Zetsche, U., Weidacker, K., Wahl, K., Schönfeld, S., & Ehlers, A. (2011). The Perseverative Thinking Questionnaire (PTQ): Validation of a content-independent measure of repetitive negative thinking. *Journal of Behavior Therapy and Experimental Psychiatry*, *42*(2), 225–232. https://doi.org/10.1016/j.jbtep.2010.12.003

Gariépy, G., Honkaniemi, H., & Quesnel-Vallée, A. (2016). Social support and protection from depression: systematic review of current findings in Western countries. *The British Journal of Psychiatry*, *209*(4), 284–293. https://doi.org/10.1192/bjp.bp.115.169094

Greer, F. L. (1980). Prognostic Expectations and Outcome of Brief Therapy: *Psychological Reports*. https://doi.org/10.2466/pr0.1980.46.3.973

Harvey, A. G., & Gumport, N. B. (2015). Evidence-based psychological treatments for mental disorders: Modifiable barriers to access and possible solutions. *Behaviour Research and Therapy*, *68*, 1–12. https://doi.org/10.1016/j.brat.2015.02.004

Heinrich, M., Zagorscak, P., Eid, M., & Knaevelsrud, C. (2018). Giving G a Meaning: An Application of the Bifactor-(S-1) Approach to Realize a More Symptom-Oriented Modeling of the Beck Depression Inventory-II. *Assessment*, 1073191118803738. https://doi.org/10.1177/1073191118803738

Horwitz, A. G., Berona, J., Czyz, E. K., Yeguez, C. E., & King, C. A. (2017). Positive and Negative Expectations of Hopelessness as Longitudinal Predictors of Depression, Suicidal Ideation, and Suicidal Behavior in High-Risk Adolescents. *Suicide and Life-Threatening Behavior*, *47*(2), 168–176. https://doi.org/10.1111/sltb.12273

Johansson, R., & Andersson, G. (2012). Internet-based psychological treatments for depression. *Expert Review of Neurotherapeutics*, *12*(7), 861–870. https://doi.org/10.1586/ern.12.63

Jung, T., & Wickrama, K. a. S. (2008). An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling. *Social and Personality Psychology Compass*, *2*(1), 302–317. https://doi.org/10.1111/j.1751-9004.2007.00054.x

Karyotaki, E., Ebert, D. D., Donkin, L., Riper, H., Twisk, J., Burger, S., … Cuijpers, P. (2018). Do guided internet-based interventions result in clinically relevant changes for patients with depression? An individual participant data meta-analysis. *Clinical Psychology Review*, *63*, 80–92. https://doi.org/10.1016/j.cpr.2018.06.007

Karyotaki, E., Riper, H., Twisk, J., Hoogendoorn, A., Kleiboer, A., Mira, A., … Cuijpers, P. (2017). Efficacy of Self-guided Internet-Based Cognitive Behavioral Therapy in the Treatment of Depressive Symptoms A Meta-analysis of Individual Participant Data. *Jama Psychiatry*, *74*(4), 351–359. https://doi.org/10.1001/jamapsychiatry.2017.0044

Kazdin, A. E. (2018). Expanding mental health services through novel models of intervention delivery. *Journal of Child Psychology and Psychiatry*, *Advance online publication*. https://doi.org/10.1111/jcpp.12937

Khan, A., Faucett, J., Lichtenberg, P., Kirsch, I., & Brown, W. A. (2012). A Systematic Review of Comparative Efficacy of Treatments and Controls for Depression. *PLoS ONE*, *7*(7). https://doi.org/10.1371/journal.pone.0041778

Klein, D. N., & Kotov, R. (2016). Course of depression in a 10-year prospective study: Evidence for qualitatively distinct subgroups. *Journal of Abnormal Psychology*, *125*(3), 337–348. https://doi.org/10.1037/abn0000147

Königbauer, J., Letsch, J., Doebler, P., Ebert, D., & Baumeister, H. (2017). Internet- and mobile-based depression interventions for people with diagnosed depression: A systematic review and meta-analysis. *Journal of Affective Disorders*, *223*, 28–40. https://doi.org/10.1016/j.jad.2017.07.021

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9 - Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lindfors, O., Ojanen, S., Jääskeläinen, T., & Knekt, P. (2014). Social support as a predictor of the outcome of depressive and anxiety disorder in short-term and long-term psychotherapy. *Psychiatry Research*, *216*(1), 44–51. https://doi.org/10.1016/j.psychres.2013.12.050

Lutz, W., Arndt, A., Rubel, J., Berger, T., Schröder, J., Späth, C., … Moritz, S. (2017). Defining and Predicting Patterns of Early Response in a Web-Based Intervention for Depression. *Journal of Medical Internet Research*, *19*(6). https://doi.org/10.2196/jmir.7367

Manen, J. G. V., Horn, E. K., Stijnen, T., Timman, R., Busschbach, J. J. V., & Verheul, R. (2015). Tailoring psychotherapy in patients with personality disorders: Matching the level of psychological strengths to the level of stabilizing versus destabilizing psychotherapy. *Personality and Mental Health*, *9*(2), 133–149. https://doi.org/10.1002/pmh.1281

Melchior, H., Schulz, H., Kriston, L., Hergert, A., Hofreuter-Gätgens, K., Bergelt, C., … Watzke, B. (2016). Symptom change trajectories during inpatient psychotherapy in routine care and their associations with long-term outcomes. *Psychiatry Research*, *238*, 228–235. https://doi.org/10.1016/j.psychres.2016.02.046

Mendez, R. D. R., Rodrigues, R. C. M., Cornélio, M. E., Gallani, M. C. B. J., & Godin, G. (2010). Development of an instrument to measure psychosocial determinants of physical activity behavior among coronary heart disease patients. *Revista Da Escola de Enfermagem Da USP*, *44*(3), 584–596. https://doi.org/10.1590/S0080-62342010000300006

Mueller, V. N. L. S., Boyle, K., Zimmermann, D., Weinmann-Lutz, B., Rubel, J. A., & Lutz, W. (2018). What Is Individually Tailored Mental Health Care? *Revista Argentina De Clinica Psicologica*, *27*(2), 157–169. https://doi.org/10.24205/03276716.2018.1062

Muthén, B. (2006). The potential of growth mixture modelling. *Infant and Child Development*, *15*(6), 623–625. https://doi.org/10.1002/icd.482

Muthén, L. K., & Muthén, B. O. (2017). MPlus (Version 8). Los Angeles, CA: Muthén & Muthén.

Richards, D., & Richardson, T. (2012). Computer-based psychological treatments for depression: A systematic review and meta-analysis. *Clinical Psychology Review*, *32*(4), 329–342. https://doi.org/10.1016/j.cpr.2012.02.004

Rubel, J., Lutz, W., & Schulte, D. (2015). Patterns of change in different phases of outpatient psychotherapy: a stage-sequential pattern analysis of change in session reports. *Clinical Psychology & Psychotherapy*, *22*(1), 1–14. https://doi.org/10.1002/cpp.1868

Schulz, U., & Schwarzer, R. (2003). Soziale Unterstützung bei der Krankheitsbewältigung: Die Berliner Social Support Skalen (BSSS). *Diagnostica*, *49*(2), 73–82. https://doi.org/10.1026//0012-1924.49.2.73

Silberschatz, G. (2015). Improving the yield of psychotherapy research. *Psychotherapy Research: Journal of the Society for Psychotherapy Research*, *27*(1), 1–13. https://doi.org/10.1080/10503307.2015.1076202

Spinhoven, P., van Hemert, A. M., & Penninx, B. W. (2018). Repetitive negative thinking as a predictor of depression and anxiety: A longitudinal cohort study. *Journal of Affective Disorders*, *241*, 216–225. https://doi.org/10.1016/j.jad.2018.08.037

Sunderland, M., Wong, N., Hilvert-Bruce, Z., & Andrews, G. (2012). Investigating trajectories of change in psychological distress amongst patients with depression and generalised anxiety disorder treated with internet cognitive behavioural therapy. *Behaviour Research and Therapy*, *50*(6), 374–380. https://doi.org/10.1016/j.brat.2012.03.005

Thornicroft, G., Chatterji, S., Evans-Lacko, S., Gruber, M., Sampson, N., Aguilar-Gaxiola, S., … Kessler, R. C. (2017). Undertreatment of people with major depressive disorder in 21 countries. *The British Journal of Psychiatry*, *210*(2), 119–124. https://doi.org/10.1192/bjp.bp.116.188078

Titov, N., Dear, B. F., McMillan, D., Anderson, T., Zou, J., & Sunderland, M. (2011). Psychometric comparison of the PHQ-9 and BDI-II for measuring response during treatment of depression. *Cognitive Behaviour Therapy*, *40*(2), 126–136. https://doi.org/10.1080/16506073.2010.550059

Vîslă, A., Flückiger, C., Constantino, M. J., Krieger, T., & Holtforth, M. G. (2018). Patient characteristics and the therapist as predictors of depressed patients' outcome expectation over time: A multilevel analysis. *Psychotherapy Research*, *0*(0), 1–14. https://doi.org/10.1080/10503307.2018.1428379

Wittchen, H. U., Zaudig, M., & Fydrich, T. (1997). *Strukturiertes Klinisches Interview für DSM-IV Achse I: Psychische Störungen (SKID-I). [Structured Clinical Interview for DSM-IV. (SCID-I)]*. Göttingen, Germany: Hogrefe.

World Health Organization (Ed.). (2017). *Depression and Other Common Mental Disorders: Global Health Estimates*. Geneva, Switzerland: World Health Organization. Retrieved from http://www.un.org/apps/news/story.asp?NewsID=56230#.WLRpQBB3xBw

Zagorscak, P., Heinrich, M., Sommer, D., Wagner, B., & Knaevelsrud, C. (2018). Benefits of Individualized Feedback in Internet-Based Interventions for Depression: A Randomized Controlled Trial. *Psychotherapy and Psychosomatics*, *87*(1), 32–45. https://doi.org/10.1159/000481515

Table 1

Socio-Demographic and Clinical Sample Characteristics at Baseline

Variable		Total	IF	SF	χ^2	*p*
Socio-Demographic Characteristics						
Age †	*M* (*SD*)	45.7 (11.3)	45.7 (11.8)	45.8 (10.7)	0.166 ‡	.868
Female Sex	*n* (%)	714 (65.6)	360 (64.9)	354 (66.3)	0.246	.620
Unemployed	*n* (%)	127 (11.8)	59 (10.6)	68 (12.7)	1.169	.280
Lower level of formal education ¶	*n* (%)	336 (30.9)	160 (28.8	176 (33.3)	5.673	.140
Marital status					1.301	.521
Single	*n* (%)	373 (34.3)	196 (35.3)	177 (33.1)		
Married	*n* (%)	561 (51.5)	286 (51.5)	275 (51.5)		
Widowed/Divorced	*n* (%)	155 (14.2)	73 (13.2)	82 (15.4)		
Clinical Baseline Characteristics						
PHQ-9	*M* (*SD*)	11.8 (3.4)	11.9 (13.4)	11.7 (3.5)	1.044 ‡	.297
SCID-I diagnosis					3.194	.670
Current MDD	*n* (%)	458 (42.1)	247 (44.5)	211 (39.5)		
Remitted MDD	*n* (%)	285 (26.2)	137 (24.7)	148 (27.7)		
Dysthymia	*n* (%)	90 (8.3)	46 (8.3)	44 (8.2)		
Double Depression	*n* (%)	58 (5.3)	27 (4.9)	31 (5.8)		
Bipolar or NOS §	*n* (%)	65 (6.0)	32 (5.8)	33 (6.2)		
No Affective Disorder Diagnosis (current/past)	*n* (%)	133 (12.2)	66 (11.9)	67 (12.5)		
Expectations †	*M* (*SD*)	10.0 (4.0)	9.9 (4.0)	10.1 (4.0)	0.462 ‡	.644
Perseverative Thinking †	*M* (*SD*)	37.4 (8.7)	37.2 (8.8)	37.6 (8.7)	0.739 ‡	.460
Perceived Social Support	*M* (*SD*)	25.7 (5.0)	25.5 (5.0)	25.8 (5.0)	1.051 ‡	.294

Note. N = 1089. IF, Individualized Feedback (n = 555); SF, Standardized Feedback (n = 534); PHQ-9, Patient Health Questionnaire-9; SCID-I, Structured Clinical Interview for DSM-IV; NOS, not otherwise specified. † variables have some missing values: age, n = 1081; expectations, n = 1081; PHQ-9, n = 1067. ‡ *t*-test for independent samples. § individuals with bipolar disorders were included only if they were not experiencing current mania/hypomania. ¶ "lower" category encompasses no certificate or certificates from lower secondary/secondary school, "higher" category encompassing certificates from trade school/college-preparatory school, college or university.

Table 2

Model Parameters of the Single-Group and the Constrained Multi-Group GMM Model

Parameter	Multi Group Model	
	Delayed Improvers	Immediate Improvers
λ_{k1}	0	0
λ_{k2}	-0.13 (0.13) ^{NS}	0.33 (0.05)
λ_{k3}	0.18 (0.15) ^{NS}	0.55 (0.04)
λ_{k4}	0.57 (0.12)	0.84 (0.03)
λ_{k5}	0.80 (0.07)	0.90 (0.02)
λ_{k6}	1.03 (0.07)	1.02 (0.02)
λ_{k7}	1.12 (0.05)	1.11 (0.02)
λ_{k8}	1	1
μ_{Ik}	12.39 (0.33)	11.23 (0.28)
μ_{Sk}	-3.41 (0.76)	-5.54 (0.37)
$\psi_{Sk,Ik}$	-2.23 (0.87)	-3.42 (1.04)
ψ_{Ik}	5.75 (0.73)	6.27 (0.78)
ψ_{Sk}	10.43 (2.48)	7.07 (1.15)
Var(ε_{i1k})	8.91 (1.60)	5.12 (0.96)
Var(ε_{i2k})	7.79 (1.94)	5.35 (1.33)
Var(ε_{i3k})	6.51 (1.46)	4.09 (0.87)
Var(ε_{i4k})	8.21 (1.14)	2.21 (0.30)
Var(ε_{i5k})	8.31 (1.71)	2.45 (0.32)
Var(ε_{i6k})	5.50 (1.17)	1.44 (0.30)
Var(ε_{i7k})	5.77 (1.08)	1.19 (0.22)
Var(ε_{i8k})	9.12 (1.56)	3.02 (0.40)

Note. λ_{kt} = class and time specific growth factor-loading, where k = refers to the class and t to the measurement occasions. μ_{Ik} and μ_{Sk} = mean of the intercept and slope, respectively. ψ_{Ik} and ψ_{Sk} = variance of the intercept and slope. $\psi_{Sk,Ik}$ = covariance between slope and intercept. Var(ε_{itk}) = residual variance at the corresponding measurement occasion t. All parameters significant with $p <$ .05 if not indicated otherwise. ^{NS} = non-significant.

Table 3

Association Between Predictor Variables and Class-Specific Intercepts, Slopes and Class-Membership

Predictor	Intercept		Slope		Predict Delayed Response Class-Membership		
	β (*SE*)	*p*	β (*SE*)	*p*	β (*SE*)	*p*	OR [95% CI]
Age	0.01 (0.01)	.554	-0.01 (0.01)	.456	-0.02 (0.01)	.158	0.99 [0.96, 1.01]
Male Sex	0.26 (0.20)	.181	-0.26 (0.24)	.275	0.14 (0.24)	.540	1.16 [0.73, 1.84]
Higher Degree of Formal Education (Ref. Lower)†	-0.10 (0.22)	.641	-0.21 (0.27)	.433	-0.27 (0.24)	.258	0.76 [0.47, 1.22]
Marital Status (Ref.: Single)							
Married	-0.21 (0.23)	.368	-0.29 (0.29)	.313	0.01 (0.27)	.963	1.01 [0.60, 1.71]
Widowed/Divorced	0.29 (0.33)	.375	-0.64 (0.40)	.109	0.09 (0.34)	.786	1.10 [0.56, 2.14]
Prior Psychotherapy (Ref.: No Prior Psychotherapy)	0.16 (0.19)	.412	-0.01 (0.23)	.970	0.39 (0.22)	.076	1.48 [0.96, 2.29]
Unemployed (Ref.: Employed)	-0.39 (0.29)	.181	0.67 (0.34)	.048	-0.10 (0.33)	.764	0.90 [0.47, 1.74]
SCID-I Diagnosis (Ref.: No Affective Disorder)							
Current MDD	1.93 (0.30)	< .001	-1.26 (0.36)	< .001	0.98 (0.35)	.005	2.68 [1.34, 5.32]
Dysthymia	0.55 (0.39)	.152	0.32 (0.45)	.479	0.52 (0.46)	.262	1.67 [0.68, 4.12]
Remitted MDD	-0.16 (0.31)	.606	0.16 (0.36)	.663	0.49 (0.36)	.174	1.64 [0.81, 3.32]
Double Depression	2.29 (0.45)	< .001	-0.28 (0.55)	.605	0.09 (0.55)	.873	1.09 [0.37, 3.17]
MDD NOS	0.53 (0.47)	.260	-0.08 (0.57)	.889	-0.11 (0.52)	.828	0.89 [0.32, 2.49]
Perseverative Thinking	0.07 (0.01)	< .001	-0.04 (0.01)	.008	-0.00 (0.01)	.790	1.00 [0.97, 1.02]
Expectations	-0.06 (0.02)	.012	-0.03 (0.03)	.329	0.08 (0.04)	.020	1.09 [1.01, 1.17]
Perceived Social Support	-0.02 (0.02)	.346	-0.05 (0.03)	.069	-0.07 (0.02)	< .001	0.93 [0.89, 0.96]

Note. Beta-regression weights as estimated in a model with class and group invariance. Ref., category used as reference; MDD, Major Depressive Disorder; NOS, not otherwise specified. † "lower" category encompasses no certificate or certificates from lower secondary or secondary school and is contrasted against the "higher" category encompassing certificates from trade school/college-preparatory school, college or university.

Figures

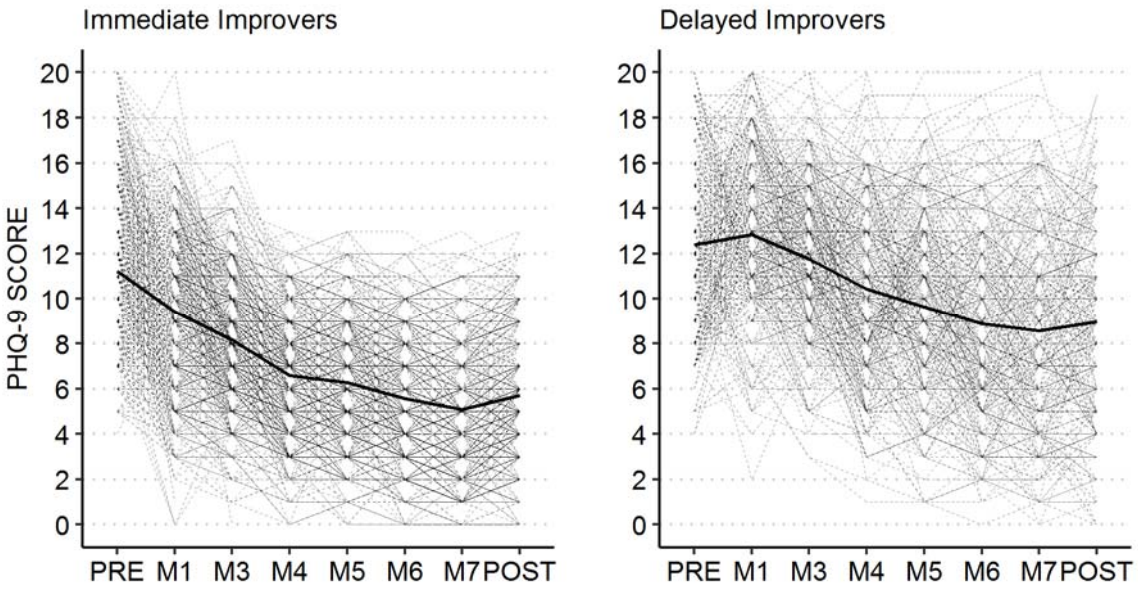


Figure 1. Observed patterns of change in the immediate improver class and the delayed improver class. Bold line depicts average change pattern.

Supplementary materials

The following supplementary materials related to STUDY 3 are available in the APPENDIX:

- *Appendix C*. Detailed description of the statistical approach.

CHAPTER 5

DISCUSSION, OUTLOOK, AND CONCLUSION

In this chapter I will summarize and discuss the results of each of the three studies and portray their core findings (CHAPTERS 5.1 – 5.3). The discussion of each study features three sections:

In the first section, I seek to embed the results into a ***broader context*** of findings on IBI in general and – if applicable – of findings on face-to-face psychotherapy. Due to the fact that all studies featured a comparison between a guided (individualized; IF) and an unguided (standardized; SF) feedback condition, this aspect will be reviewed more extensively within the discussion of the first study. In the second section, I will highlight ***specific problems*** of our own studies and problems within the respective research fields in order to suggest comprehensive solutions of these problems in ***future studies*** within the third section.

Finally, in CHAPTER 5.4 I will present a more generalized outlook on the future of research and practice regarding the application of "new media" and the internet as a tool to provide relief for individuals with mental disorders in general and depression in particular.

5.1. Study 1: Benefits of Individualized Feedback in IBI

5.1.1 Summary of Findings

The first study investigated differences between the SF and the IF condition concerning changes on clinical (depression, anxiety, perseverative thinking) and psychosocial (emotional self-efficacy, quality of life, and perceived social support) variables. The between-condition effects were estimated immediately after the intervention as well as 3, 6, and 12 months after participants finished the program. The results revealed significant within-group pre-post effects across conditions on all variables in the program-intended direction. The two conditions did not differ on any of the self-reported measures and at none of the measurement occasions. However, it is worth mentioning that the dropout rates were significantly higher in the SF condition than in the IF condition (25.8% vs. 17.3%). The core result on the primary outcome of depressive symptoms (Beck Depression Inventory – II) is visualized in FIGURE 5.1. This figure shows the high overlap in symptom change between both treatment arms.

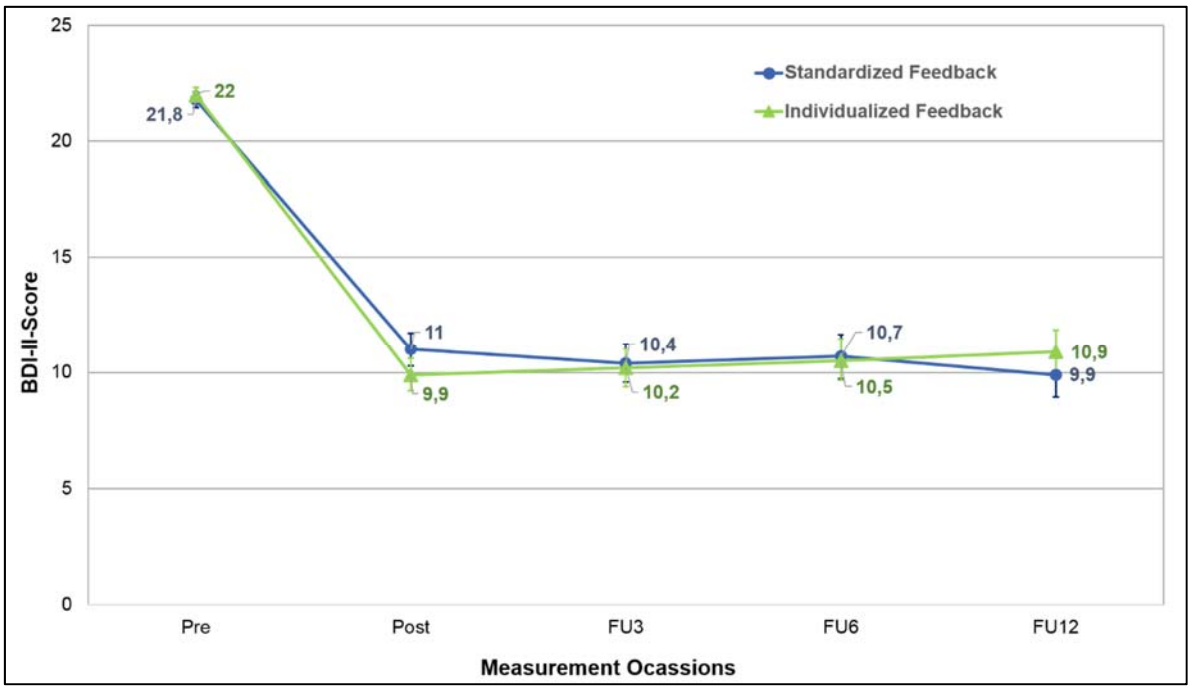


FIGURE 5.1. Depressive Symptom Scores for the Standardized Feedback-Condition (blue) and the Individualized Feedback-Condition (green) across all measurement occasions. Vertical lines indicate 95%-confidence intervals for each score (all respective intervals overlap between groups). FU, follow-up measurement with adjacent number indicating months after post-assessment. BDI-II, Beck-Depression-Inventory II.

5.1.2 Discussion

These findings are in conflict with results from most reviews and meta-analyses (e.g., Johansson & Andersson, 2012; Richards & Richardson, 2012) and question the superiority of "guided" over "unguided" treatments. A potential equivalence of unguided and guided IBI bears significant implications for clinical practice. Since qualified personnel that may offer "guidance" is limited and costly, it seems reasonable to focus resources on potentially more cost-effective unguided treatments in order to provide broader patient populations with easily-accessible and quickly-available treatment (e.g., Romero-Sanchiz et al., 2017). Given the fundamentality of this notion, it is important to question its generalizability and to critically evaluate the foundations it is based on.

Context. Studies comparing guided and unguided IBI exist for a variety of mental disorders. To present an unbiased collection of studies representative of the field, I performed a literature search[13] on "ISI Web of Science", on "guided" and "unguided" interventions. Note, that this search was not limited to depression. It produced 50 search results, from which I excluded study protocols, offline-studies or studies on other topics than psychological interventions. The results are summarized in TABLE 5.1. While this approach to literature provides only a simplified overview, it reveals that across disorder categories the differences in experimental comparisons of guided and unguided interventions are mostly non-significant. Apart from two meta-analyses on guidance in IBI for depression with significant shortcomings (see SECTION 1.6.2 for a review), none of the reviewed studies yield evidence for the superiority of "guided" over "unguided" IBI for any disorder category.

[13] The exact search terms were: TITLE: (guid*) AND TITLE: (unguid*) Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, ESCI. Date of search: 28th of December 2018.

TABLE 5.1 (CONT. ON FOLLOWING PAGE)
Results of randomized controlled studies and meta-analyses on guided and unguided IBI for a variety of mental disorders

Study	Sample	Target Disorder	Interventions	Pre-Post Group Comparisons	Mode of Contact with Participants
Berger et al. (2011)	*n* = 81	Social Phobia	1) guided iCBT 2) iCBT + contact on-demand (written first, phone calls second) 3) unguided iCBT	*Social Phobia Symptoms:* 1 = 2 = 3	Written (on-demand phone calls in one treatment arm)
Ciuca et al. (2018)	*n* = 111	Panic Disorder	1) iCBT guided 2) iCBT unguided 3) WLC	*Symptoms of Panic Disorder:* 1 = 2 > 3 *Diagnostic Status:* 1 > 2 1 > 3 2 > 3	real-time audio-video communication
Furmark et al. (2009)	*n* = 235	Social Phobia	1) guided iCBT 2) bibliotherapy with discussion group 3) internet-delivered applied relaxation + online discussion group 4) bibliotherapy alone	*Social Phobia Symptoms:* 1 = 2 = 3 = 4	Written (clinician + discussion group)
Ivanova et al. (2016)	*n* = 152	Panic Disorder Social Phobia	1) Online ACT guided 2) Online ACT unguided 3) WLC	*General and Social Anxiety Symptoms* 1 = 2 > 3	Written (clinician)
Kass et al. (2014)	*n* = 151	Binge Eating Disorder	1) Online-psychoeducation with guided discussion group 2) Online-psychoeducation	*Binge Eating Symptoms* 1 = 2 *Weight Concerns* 1 > 2	Written (discussion group)
Romero-Sanchiz et al. (2017)	*n* = 296	Depression	1) iCBT guided 2) iCBT unguided 3) TAU	*Cost-Effectiveness* 1 = 2 > 3	Written
Sundstrom et al. (2016)	*n* = 80	Alcohol Use	1) iCBT + chat counseling 2) iCBT + e-mail counseling 3) iCBT unguided	*Self-Reported Alcohol Consumption:* 1 = 2 > 3	Written

Meta-Analysis	Sample	Target Disorder	Interventions	Result regarding guidance	Mode of Contact with Participants
Küster et al. (2016)	$k = 20$	Post-traumatic stress disorder	Meta-Analysis	*Post-traumatic stress:* No differences between guided and unguided interventions	Not reported, but mostly written (expressive writing or CBT with written messages)
Riper et al. (2014)	$k = 16$	Alcohol Misuse	Meta-Analysis	*Alcohol Misuse:* No differences between guided and unguided interventions	14 Written 1 SMS 1 Written + Phone

Note. iCBT, internet-based cognitive-behavioral therapy; TAU, treatment as usual; WLC, waitlist control group. The table is adapted from a response letter sent during the publication process of STUDY 1.

However, due to methodological shortcoming in the current literature, neither does it follow that "unguided" IBI is superior nor that both treatment modalities are equivalent.

For example, a sensitivity analysis (see also TABLE 1.4, p. 29) demonstrates that three of the five studies with experimental comparisons of different levels of guidance in IBI for *depression* have only been sufficiently powered to detect large between-group effects of $d \geq .73$ (Andersson et al., 2013; Berger, Caspar, et al., 2011; Vernmark et al., 2010). Only two studies were appropriately designed to detect small to medium effects of $d \geq .36$ or $d \geq .5$, respectively (Kelders et al., 2015; Titov, 2011). As is evident in TABLE 5.1., most studies on IBI for *other mental disorders* did not feature larger sample sizes either. Given that these studies compared the same interventions with and without an additive component (i.e., guidance) it is implausible that the effects are *anything but* small (Bell, Marcus, & Goodlad, 2013; Cuijpers, Cristea, Karyotaki, Reijnders, & Hollon, 2019). Thus, it remains unclear whether there may be smaller differences between "guided" and "unguided" IBI.

The trial presented in this thesis is the only sufficiently powered study on IBI for depression to draw reliable conclusions on this question – but a singular study on a specific intervention with narrow inclusion criteria is far from sufficient to generalize regarding such an important question. This is partly due to specific problems with the definition of "guidance"

and the general heterogeneity in what is considered "guidance" in IBI (see also SECTION 1.6.2 for a summary).

Specific problems. The way contact, support or guidance for patients is provided within a comprehensive trial cannot be reduced to single words like "guided" or "unguided". The following BOX 5.1 provides details on how guidance was provided in the IBI that this thesis is based on ("TK-DepressionsCoach"). It illustrates the different facets of (human or automated) contact that are to be accounted for.

BOX 5.1. DETAILS ON THE QUALITY AND QUANTITY OF HUMAN CONTACT IN THE INTERVENTION TK-DEPRESSIONSCOACH

All participants in the TK-DepressionsCoach had contact per phone with an insurance representative and subsequently completed an hour-long structured clinical telephone interview by a student assistant from our research group. Participants may have even been in contact with a member of the IT-company responsible for hosting the intervention, if they had (technical) questions during initial symptom screening and registration. Further, during the intervention all participants were able to contact members of the research team via written messages in case of technical problems or were contacted by members of the research team by phone in case of symptom exacerbation. All participants received friendly, automated e-mail reminders if they did not login after a certain period of time or if a new treatment module was available for them. This is already a significant amount of (human or automated) contact – regardless of the condition a participant was randomized to.

Evidently, the conditions also differed. Individuals in the IF condition received module-wise feedback (six times) based on pre-written text modules that were adapted by counselors carrying a bachelor's or master's degree in psychology to fit the participants situation or the specific work product he or she submitted as therapeutic homework (e.g., completed thought protocols). Only a few paragraphs of feedback to clients had to be written without pre-typed text modules (e.g., summarizing an expressive writing task). Over the course of the entire intervention, counselors invested 120 – 180 minutes per client to provide feedback (about 15-45 minutes per feedback). Please see APPENDIX 5.2

for an excerpt from the treatment manual that illustrates how individualized feedback is written (Zagorscak et al., 2014; manual available upon request).

At the same time, individuals in the SF condition received completely automated feedback ("one size fits all") with generic encouragements and attempts at predicting possible difficulties.

Importantly, the feedback section comprises only about 15% of the overall text that is provided. Psychoeducation (including videos), therapeutic tools and instructions on how to use them are standardized and the same for every participant. All these texts, including the text modules used for individualized feedback are authored by the same group of researchers (Zagorscak et al., 2014).

This extensive description serves to demonstrate three things: *First*, the absolute difference between the compared conditions in our trial is small. Thus, the lack of differences between our treatment arms should not be used to draw conclusions on differences between more dissimilar forms of guidance in IBI.

Second, the level of detail in the description is necessary to determine what kind of "guidance" (if any) a participant in IBI received. Yet, no published paper (including our own) features such an extensive description. However, only detailed information on *who* contacted participants *when,* in what *frequency,* and with what kind of *aim* (e.g., technical assistance vs. therapeutic guidance) allows for meaningful discussions and (meta-analytic) comparisons of studies on "guidance". A solution for future studies might be a standardized form that details necessary information on guidance in every trial on IBI. Such a short checklist could complement publications as supplementary material. An example that is derived from my own review of the literature is provided in APPENDIX 5.1.

Third, the quantity (invested time per client) and quality (aim and scope) of feedback in our IF condition is consistent with what other studies considered to be "guided" IBI (Pihlaja et al., 2018). Similarly, our SF condition is consistent with what other trials defined as "unguided" IBI, since most trials on "unguided" IBI offered contact-on-demand and researchers contacted

participants when symptoms deteriorated (e.g., Berger, Caspar, et al., 2011, 2011; Ciuca et al., 2018; Kelders et al., 2015). When examining the descriptions of „guidance“ of our intervention and other guided self-help or minimal contact treatments (see SECTION 1.5.1 for definitions), it becomes clear, that this form of human contact is quite different from the one patient and therapist experience within face-to-face psychotherapy. It thus seems reasonable to expand on the *meaning* of guidance in IBI for patients.

For example, the intervention manual of the TK-DepressionsCoach explicitly instructs counselors to give encouraging, resource-focused, empathetic and non-confrontative feedback (see APPENDIX 5.2). This is due to the fact, that the asynchronous mode of communication is not suited for a faster-paced back-and-forth between patient and therapist, which is necessary for Socratic Questioning and other therapeutic techniques central to cognitive-behavioral therapy (and other schools of psychotherapy). Furthermore, the structure of guided self-help IBI restricts the number of interactions between counselor and client. In addition, the guidance is mostly focused on module-specific tasks. Nevertheless, there is some room for correcting the clients as the following fictitious BOX 5.2 illustrates, where a counselor gives feedback on a client's thought protocol:

BOX 5.2. FICTITIOUS EXAMPLE OF COUNSELOR FEEDBACK ON THOUGHT PROTOCOLS (BASED ON THE ORIGINAL GERMAN TEXT MODULES)

"In Ihren Gedankenprotokollen hatten Sie eine Situation berichtet, wo Ihre Kolleginnen Ihre Einladung zu einem Feierabendbier ausgeschlagen haben. Ich habe gesehen, dass Sie den alternativen Gedanken "Sie wollen einfach nichts mit mir zu tun haben" formuliert haben. Denken Sie daran, dass „alternative Gedanken" das Potential eröffnen können, auch alternative Gefühle zu empfinden. Wie Sie an Ihrem Gedankenprotokoll erkennen können, würde Ihr „alternativer Gedanke" genauso zu „Traurigkeit" führen, wie Ihr ursprünglicher Gedanke „Für mich hat nie jemand Zeit". Fällt Ihnen noch ein anderer Gedanke zur Erklärung der Reaktion Ihrer Kolleginnen ein, der möglicherweis mit einem anderen Gefühl einhergegangen wäre? Ich möchte Sie einladen noch ein weiteres Gedankenprotokoll zu dieser Situation auszufüllen."

If a client reacts to such a correction, the counselor in our treatment has the opportunity to give feedback on the revised thought protocol *a week later and no more than twice* in total, given that cognitive restructuring takes place during the final two weeks. Accordingly, the potential for "trial and error" is limited. This structure is common to all guided self-help or minimal contact interventions (e.g., Cuijpers, Kleiboer, Karyotaki, & Riper, 2017). Thus, school-specific elements (i.e., cognitive-behavioral or psychodynamic techniques) are mostly contained to the automated or standardized psychoeducational parts or the provided online-tools. In contrast, the central role of human contact in these interventions is to encourage participants to keep up their work and provide them with empathy. In that aspect, the aim of the interaction is more similar to non-directive supportive face-to-face treatments (e.g., Cuijpers et al., 2012).

For a comprehensive representation of the meaning of human interaction in IBI, it is important to consider the perspective of the participants as well. When asked about their experiences with the TK-DepressionsCoach, more than half of the participants (N_{SF} = 289; N_{IF} = 280) left free-form comments after completing the intervention. Large proportions of participants (N_{IF} = 147, 52%; N_{SF} = 115, 40%) wrote to thank the creators of the intervention, the public insurance company (or the counselor, in the guided variant) for offering treatment and support. A closer look at the contents of their comments emphasizes what they took from the provided guidance – and what was missing (see BOX 5.3 a-c)[14]:

[14] All comments printed in this chapter were slightly adapted to guarantee anonymity of participants without changing the meaning of the content. They were not translated to English to maintain the general style of the comments.

BOX 5.3. THREE EXAMPLES OF COMMENTS BY PARTICIPANTS IN THE TK-DEPRESSIONSCOACH CONCERNING THE QUALITY AND QUANTITY OF HUMAN CONTACT

a) *„Ich hätte gerne nach jeder Woche auch Stellung zu meinen Erlebnissen und Erfahrungen abgegeben damit es zu einem Dialog zwischen mir und meiner Beraterin kommt. Ansonsten nehme ich diese Zeit als eine sehr bereichernde Erfahrung und Ergänzung zu meinem Wissen um meine Situation mit.“*

b) *“Danke für die aufmunternden Worte als Antworten auf meine Schreiben!“*

c) *„Der DepressionsCoach gibt wirklich viele interessante Denkanstöße. Es ist auf eine ganz andere Art hilfreich als eine ambulante Therapie. Ich fände die Kombination aus persönlichem Gespräch und Online-Arbeit gut.“*

Concordantly, about 10% (N = 27) in the IF and 31% (N = 86) in the SF condition criticized the intervention for being too “impersonal”. Other participants explicitly wished for more phone calls as an add-on to the written communication (N_{IF} = 12; N_{SF} = 16).

Interestingly, a small and not precisely quantifiable share of participants in the SF condition were not aware that they did *not* receive feedback individualized by a counselor. This became evident when they were called due to symptom exacerbation and inquired whether the person calling was the same as the one writing the feedback (which was in fact fully-standardized). An ambivalence about whether or not the letters are written by a human and whether someone reads the client’s work product is nicely summarized by the comment of a participant in the SF-condition (BOX 5.4):

BOX 5.4. EXAMPLE OF A COMMENT BY A SF-PARTICIPANT IN THE TK-DEPRESSIONSCOACH CONCERNING AN AMBIVALENCE ABOUT HUMAN SUPPORT

„Ich bin [IT-Beruf, anonymisiert] und verstehe die Mechanismen hinter einem solchen Projekt. Eventuell liest das niemand, und eventuell schreibt mir auch niemand. Vielleicht dreht sich das Pilot-Projekt in Wirklichkeit um völlig automatisierte Behandlung ***zwinker***

Na ja, so krass wird es auch nicht sein, aber ich habe diese Fragen [Working Alliance Inventory; Anm. d. Autors] nicht beantwortet, weil sie (in meinen Augen) nicht beantwortbar sind.“

This illusion of human presence might be fostered by a trustworthy authority recommending a well-designed intervention (the participant's health insurance company), and an unguided condition providing generic but warm, encouraging and empathetic feedback. Further, the clinical interview, the contact-on-demand option and the phone call after symptom deterioration by a clinical expert might have further contributed to a feeling of being "looked after", even in the SF-condition. This is consistent with ideas by Mulder, Murray and Rucklidge (2017), who argued that participants in IBI form symbolic connections with the intervention which have repercussions, comparable to the disembodied relationships formed with beloved books or movies. The authors summarized that "although very different from the traditional image of a real-time healing encounter, engagement with an e-therapy website (laden with evidence of benevolent therapeutic intention) might nonetheless constitute a meaningful relationship with therapeutic benefits." (Mulder et al., 2017, p. 6). Indeed, a first study that addressed this question through interviews with 9 individuals found preliminary evidence for the notion that participants assign relationship-associated qualities to an unguided mental health app, such as feeling a supporting partnership with the inanimate treatment tool (Berry, Salter, Morris, James, & Bucci, 2018).

In summary, there are a number of conclusions to be drawn from the findings of our study, the practical experiences in carrying it out, and the participants' comments:

- *First,* well-designed unguided interventions with pre-treatment clinical interviews and contact-on-demand options have the potential to be as efficacious as comparable interventions with feedback individualized by a therapist. Nevertheless, higher drop-out rates are to be expected when feedback is fully-standardized.
- *Second*, guidance in text-based guided self-help or minimal-contact IBI is *not* comparable to the human interaction in school-specific face-to-face psychotherapy for it has a different quantity and quality (different aims, different forms of dialogue). The

aims of guidance seem to be more similar to those of non-directive supportive treatments.

Future studies. Since the *first conclusion* is based primarily on our trial (due to similar trials being underpowered), it is necessary to conduct other sufficiently powered trials that compare different qualities and quantities of guidance with more diverse target populations (e.g., other mental disorders, more severely affected individuals, individuals with comorbidities). Most importantly, the circumstances under which this trial was conducted were those of a typical *efficacy* study. Among other aspects of quality control, all counselors were under the employ of the university and a strict adherence to the treatment manual was ensured by introductory courses and regular supervision for all employees.

In an earlier review on studies of IBI *effectiveness* Andersson and Hedman (2013) found only two open trials on IBI for depression that featured no control condition. Similar effectiveness studies without appropriate control groups have been published since (Hadjistavropoulos et al., 2016; Newby, Mewton, & Andrews, 2017). Only one large-scale study (N = 639) in Great Britain compared the effectiveness of two usually unguided IBI for depression in routine care ("MoodGYM" and "Beating the Blues") with added technical support[15] against a control condition ("care as usual"). The authors did not find any significant differences between any of the three conditions and reported low adherence and engagement with the IBI offered (Gilbody et al., 2015). Nevertheless, they cautioned in their conclusion that "support in [this] trial did not involve detailed explanations of cognitive behavioral therapy and did not involve detailed review of homework or tasks between sessions. The [IBI] was therefore a form of supported self-help but was not one that was guided by a clinician." (Gilbody et al., 2015, p. 11). It seems reasonable to assume that without the circumstances common in studies

[15] Technicians tried to call the participants weekly (average: 6-7 calls per patient) for providing a few minutes of additional support. However, only an average number of 3-4 calls were taken per patient for the entire duration of the intervention.

on the efficacy of IBI (e.g., clinical interviews, contact-on-demand, strictly manualized procedures, symptom tracking and reactions to exacerbation), providing guidance by clinicians would become more important. Thus, different qualities and quantities of guidance should be investigated within future effectiveness studies.

The *second conclusion* in the paragraph above regarding the meaning and nature of guidance and the formation of a symbolic relationship with an IBI are not yet empirically researched. They require more qualitative approaches and detailed inquisitions of the participants' experiences within IBI (e.g., Berry et al., 2018). They are necessary to determine what participants need from an intervention and whether the perception of human presence might be enough for some participants, whereas others need more care and human attention in order to adhere to and profit from IBI. Thus, a major challenge for future studies on IBI is to identify whether certain types of individuals (e.g., depending on personality, clinical or psychosocial variables) might reliably profit from less resource-intensive and immediately available unguided IBI, whereas other individuals are better served with guided treatments or face-to-face psychotherapy, even if these incur costs and might require waiting times. This notion bears some similarities with the idea of personalized medicine (for a summary see Foster, Petrie, Mischoulon, & Fava, 2019). However, for this approach to work, further studies on interventional, clinical, biological and psychosocial predictors of symptom change in IBI is required.

5.2 Study 2: Contributing factors in IBI for depression

5.2.1 Summary of Findings

The second study investigated the contributions of different variables to in-treatment depressive symptom change in standardized and individualized IBI for depression. The variables investigated corresponded to the four-factor-model (Lambert, 1992; Miller et al., 1996). Results show that reductions of extra-therapeutic stressors during the intervention and high agreements on tasks and goals of treatment are associated with depressive symptom reductions in both conditions. The study yielded interesting differences between the two conditions: On the one hand, the level of baseline stressors was only predictive of in-treatment symptom change in the standardized feedback condition, possibly indicating a stress-buffer effect of guidance. On the other hand, bond ratings were associated with symptom deterioration during the final week of treatment in the individualized feedback condition only, which might indicate "parting pains". In contrast to assumptions made by the four-factor-model, baseline expectations and the uptake of specific treatment components were not significantly associated with depressive symptom change in this trial. See FIGURE 5.2. for an illustration of findings.

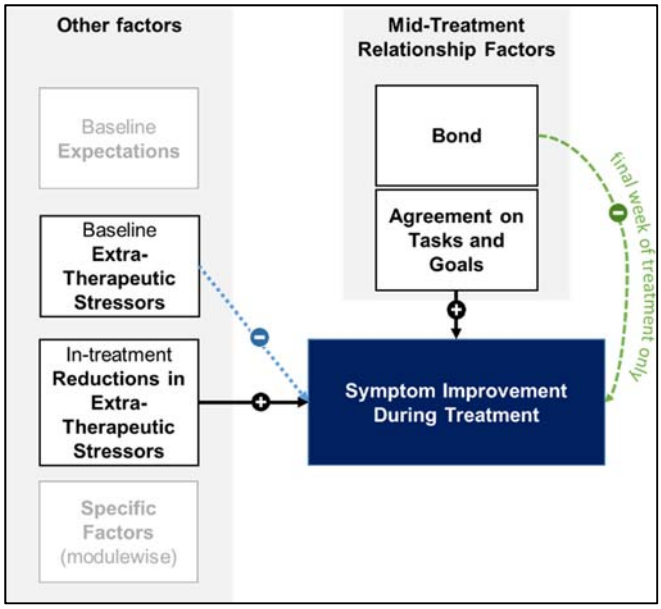


FIGURE 5.2. Simplified illustration of factors contributing to symptom change in IBI for depression. Black lines indicate findings valid across conditions; the green line indicates findings specific for the individualized feedback condition; the blue line indicates findings specific for the standardized feedback condition. "-" and "+" indicate negative and positive associations, respectively. No significant associations were found for factors in grey boxes.

5.2.2 Discussion

The findings of this study are in line with previous results, in that they emphasize the associations of therapeutic alliance with treatment outcome (Flückiger et al., 2018; Pihlaja et al., 2018). At the same time, they confirm that agreement on tasks and goals might be more relevant for symptom change in IBI than bond ratings (Berger, 2017). Concerning the relevance of expectations as contributing factor, multiple studies on IBI for social anxiety found significant associations with outcome (Boettcher et al., 2013; El Alaoui et al., 2015; Hedman et al., 2012; Nordgreen et al., 2012). In contrast, a significant expectation-outcome association has only been reported in one study on IBI for depression (El Alaoui et al., 2016). While it is possible that predictors of outcome differ in dependence of the mental disorder of an individual or the treatment offered (Wolitzky-Taylor, Arch, Rosenfield, & Craske, 2012), the contradiction to our results might also be due to the fact that the other studies did not control for the influence of other potential contributing factors, such as working alliance (e.g., Patterson, Anderson, & Wei, 2014). Overall, this highlights the strength of our multi-factor approach.

Another strength of our study is its focus on the uptake of treatment-specific tools and extra-therapeutic stressors as possible contributors to module-wise symptom change. This has not been investigated in other studies on IBI. Thus, our results will be contextualized more broadly and compared to factors known to contribute to outcome measures in face-to-face-psychotherapy.

Context. Knowledge on the importance of *common vs. specific* factors in face-to-face psychotherapy is mostly derived from comparative trials of different treatment approaches (e.g., psychodynamic vs. cognitive-behavioral) and from treatment component studies ("dismantling studies"). Meta-analyses of comparative treatment studies have mostly concluded on the equivalence of different psychotherapeutic approaches or have only found small differences between differing treatments (e.g., Baardseth et al., 2013; Tolin, 2010). Similarly, dismantling studies that compared treatments with added specific components to those without these

components, yielded small or non-significant between-group effects (for meta-analyses, see Ahn & Wampold, 2001; Bell et al., 2013; Cuijpers, Cristea, et al., 2019). Both the results from treatment component studies and from comparative meta-analyses have been used in support of the notion that all schools of psychotherapy are mostly effective due to the same mechanisms of change ("common factors") and that specific factors only play a minor role (e.g., Wampold, 2015).

Apart from these studies on the broader categories of "common" and "specific" factors, there have been numerous trials on the contribution of *therapeutic alliance* and *expectations*. Meta-analyses show significant correlations with outcome in face-to-face psychotherapy for both variables, with larger associations reported for therapeutic alliance (r = .28, 95% CI [.26, .30]; d = .58) than for expectations (r = .12, 95% CI [0.04, 0.20]; d = .24) (Constantino, Vîslă, Coyne, & Boswell, 2018; Flückiger et al., 2018).

Evidence for the importance of *extra-therapeutic factors* can be derived from work by Cuijpers et al. (2012). This meta-analysis compared the effects of non-directive supportive treatments to specific treatments and passive control groups. It was estimated that extra-therapeutic factors contribute 33.3% to overall outcome (for a detailed description of this study see SECTION 1.6.3).

Despite the fact that meta-analyses on all these topics exist, the evidence for the relevance of different contributing factors in face-to-face psychotherapy is still questionable for a couple of reasons, which also apply to similar studies on IBI:

First, it is problematic to infer similar mechanisms of change for all therapeutic approaches from the fact that they yield comparable outcomes. Alternative explanations for this finding may exist, e.g., while one therapeutic approach may change behaviors first, another might change interpersonal interactions or thoughts. These facets of change then interact with each other and subsequently influence the therapeutic interaction with the therapist (Cuijpers, Reijnders, et al., 2019). Inappropriately, such complex aspects are usually reduced to (pre-post

differences of) sum scores of a given questionnaire and thus do not allow for valid conclusions about the pathways to change (e.g., Heinrich, Zagorscak, Eid, & Knaevelsrud, 2018; Kazdin, 2007).

Second, most studies investigating predictors of change display methodological or design flaws that impede reliable conclusions on the importance of the factors under research. In a recent review on this topic Cuijpers, Reijnders et al. (2019) concluded that "[p]sychotherapies may work through techniques that are specific to each therapy or through factors that all therapies have in common, but currently, there is insufficient evidence to enable either common factors or specific factors to explain how therapies work." (p. 19). The authors justified this conclusion by summarizing that most studies on contributing factors are correlational, observational and do not investigate temporal associations, possible mediators or dose-response relationships. As other papers have previously pointed out, these shortcomings impede interpretations about causality and thus do not validly inform the debate on mechanisms of change in psychotherapy (Kazdin, 2005, 2007).

In summary, there are a large number of studies on the correlational association of the four factors proposed by Lambert (1992) and Miller et al. (1996) with outcome in face-to-face psychotherapy, which have been summarized in multiple meta-analyses (e.g., Constantino et al., 2018; Cuijpers et al., 2012; Flückiger et al., 2018). Based on this evidence, the strongest associations with outcome seem to exist with extra-therapeutic factors and the common factor of therapeutic alliance. Smaller associations are reported for expectations and specific factors or techniques. Nevertheless, the current state of research does not permit reliable conclusions on their relative importance due to methodological and design problems of the conducted studies. The following points represent recommendations by several different research groups to overcome the limitations in the study of mechanisms of change as discussed above (for a detailed summary, see Cuijpers, Cristea, et al., 2019; Cuijpers, Reijnders, et al., 2019; Kazdin, 2005, 2007; Mulder et al., 2017; Sieverink, Kelders, Poel, & Gemert-Pijnen, 2017; Silberschatz,

2017). These suggestions offer important guidelines for future research on both IBI and face-to-face psychotherapy:

- Clinical trials require multiple measurement occasions to establish temporal associations of factors contributing to change (Kazdin, 2007).
- The specific variable of interest must not solely be derived from observational designs (Kazdin, 2005), but should rather be manipulated to control for possible confounders.
- Instead of focusing on single predictors of change, studies on mechanisms of change need to include possible mediators. This requires statistical methods that allow for the inclusion of multiple variables and are suited for modeling change over time (Cuijpers, Reijnders, et al., 2019).
- The included variables need to be derived from and supported by theoretical frameworks (Cuijpers, Reijnders, et al., 2019).
- Dose-response relationships need to be investigated (Kazdin, 2007). That requires small-grained investigations of the amount of different treatment components a specific individual receives or the experimental manipulation of that amount through dismantling designs.
- Patient populations are not homogenous. Thus, it is necessary to identify and investigate subgroups with discernable characteristics on mechanisms of change to improve psychotherapy (Silberschatz, 2017).
- Large sample sizes are imperative for the investigation of small effects (e.g., dismantling studies or other comparisons of active treatments) (Cuijpers, Cristea, et al., 2019).

Again, this list highlights the potential of IBI for psychotherapy research. Within IBI studies it is easier to track all activities of participants within an intervention in order to investigate dose-response relationships. Due to the fact that all activities take place within an online environment, it is easier to implement questionnaires for repeated assessments. Further,

given the lower cost of providing treatments with reduced human contact, it is less resource-intensive to acquire large samples of individuals participating in treatment. Finally, the high amount of standardization and control reduces possible confounders. Even experimental tests of factors that are very difficult to manipulate in face-face-psychotherapy are more easily changed in IBI. For example, concerning possible improvements on research on therapeutic alliance, Cuijpers, Reijnders et al. (2019), caution that it might be challenging from an ethical and organizational standpoint to reliably manipulate aspects of therapeutic alliance or therapists' behavior. Interestingly, the common comparison of different quantities and qualities of guidance in IBI entails such a manipulation.

At the same time, in spite of the available potential, previous IBI studies have rarely fulfilled many of the listed requirements. This emphasizes the strengths and qualities of the three studies in this thesis that has carefully manipulated "guidance" as an important aspect of treatment within an experimental design. Further, we investigated multiple possible factors contributing to symptom change on the basis of a theoretical framework while accounting for temporal ordering of change and possible dose-response relationships.

Specific problems. Apart from the many strengths of our trial, some specific problems and questions remain. They are related to the use of the working alliance inventory, the operationalization of the theoretical model and the temporal modeling of changes.

First, the instrument we used to asses therapeutic alliance is a version of the Working Alliance Inventory (WAI; Hatcher & Gillaspy, 2006; Munder, Wilmers, Leonhart, Linster, & Barth, 2010) as adapted by Berger, Boettcher, and Caspar (2014) for the IBI setting. The WAI is the most commonly used instrument for assessing this construct in trials of face-to-face psychotherapy and IBI (Berger, 2017; Flückiger et al., 2018).

The original instrument assesses the three sub-facets bond (e.g., „I believe my therapist likes me"), agreement on goals (e.g., "My therapist and I agree on what is important for me to work on.") and agreement on tasks (e.g., "I believe the way we are working with my problem

is correct."). However, it is important to note that multiple studies have shown that the task and goal component are not differentiable in factor analyses (Falkenström, Hatcher, & Holmqvist, 2015; Falkenström, Hatcher, Skjulsvik, Larsson, & Holmqvist, 2015; Munder et al., 2010).

The adapted version by Berger, Böttcher and Caspar (2014) differs in that it assesses agreement with tasks and goals *of the intervention* rather than agreement with tasks and goals *set personally with the therapist* (e.g., "I believe the way I am working within the online-program on my problem is correct."). In our own trial, for items assessing the bond component, the term "therapist" was replaced with "counselor" (IF) or "The Team of the Online-Intervention" (SF). While these adaptations seem appropriate for the IBI-setting, they already highlight a change in meaning to the original instrument and might impede comparisons between them.

Further, the validity of the general construct in the IBI context, and of bond-ratings in particular, is questioned by the fact that participants in previous studies reported problems with answering WAI-questions due to the fact that they do not feel they know their therapists (Berger, 2017). In our study, about 5-10% (N_{SF} = 24; N_{IF} = 11) of the participants' open comments at the end of treatment addressed problems with the WAI, either highlighting the participants' refusal to answer the questions deemed inappropriate or trying to explain why they were reporting poor mid-treatment ratings (see BOX 5.4 in SECTION 5.1.2).

Second, we chose the four-factor-model as a theoretical basis for our study due to its large significance in face-to-face psychotherapy and its broad influence on discourse on contributing factors in the past. Thus, we sought to highlight the similarities between face-to-face psychotherapy and IBI. However, contrary to what the model predicts for the face-to-face setting, expectations and specific factors were not associated with symptom change in our analysis. Further, while Lambert (1992) proposed that the four factors explain all of the variance in outcome, our model explained only a small proportion, which is exemplified by low R^2 values (between .05 and .17 depending on the measurement occasion, see CHAPTER 3/TABLE 2, p. 92).

It is important to note, that the amount of explained variance in psychological research is generally low due to the complexity of human thoughts, emotions and behavior. For example, the therapeutic alliance was summarized to account for 7% of outcome across empirical face-to-face psychotherapy trials (Koole & Tschacher, 2016), whereas the four-factor-model suggests (based on a non-empirical review of the literature) relationship factors to account for 40% of the outcome (Lambert, 1992).

Nevertheless, the low R^2-values found in our trial indicate that there are probably other variables predictive of outcome that have not been taken into account. One reason might lie in the particular operationalization of the four factors in STUDY 2. It should be kept in mind that our trial did not assess all facets of the four-factor-model. Instead, we used the model as a basis to derive variables that might be particularly representative of the four factors. For instance, we assumed therapeutic alliance ratings to represent the factor "relationship" or the uptake of specific treatment components to represent the factor "specific technique". Adding further indicators for the four factors may contribute to a higher explained variance in symptom improvement.

Reviews and meta-analyses on all these constructs highlight a wide variety of definitions and operationalizations (e.g., Constantino et al., 2018; Cuijpers, Reijnders, et al., 2019; Flückiger et al., 2018; Wampold, 2015). While it is impossible to discuss them all in the context of this thesis, I want to highlight a few important gaps in our operationalization in order to suggest ways to close them in the context of the next section.

While the largest quantity of studies exists on therapeutic alliance, other *relationship factors* have been discussed and partly empirically confirmed as being correlated to outcome as well (for a summary see Wampold, 2015). Rather than being related to patients' perceptions of the therapeutic alliance, they are related to the therapists' behavior (e.g., warmth, empathy, encouragement, positive regard). Even though a meta-analysis has shown these factors to be highly correlated and oftentimes indistinguishable from each other (Nienhuis et al., 2018), it

seems reasonable to quantify therapists' behaviors in more detail in future studies (e.g, Paxling et al., 2013).

With regard to *extra-therapeutic factors*, our operationalization (extra-therapeutic stress as assessed with the PHQ-S) was quite coarse, given that the entire life of the client is considered to be "extra-therapeutic" (Thomas, 2006). This pertains to critical life events, daily hassles and uplifts and to a lesser degree to the client's stable traits (e.g., personality, inner strengths) and changes in external resources (e.g., socioeconomic, personal relationships), which have been previously linked to well-being and depression (Luhmann, Hofmann, Eid, & Lucas, 2012; Maybery, Neale, Arentz, & Jones-Ellis, 2007; McIntosh, Gillanders, & Rodgers, 2010; Vinkers et al., 2014).

Similarly, our approach to *specific factors* was limited to quantitative indicators of the uptake of specific treatment components (e.g., number of completed thought protocols). However, we were neither able to track how much of the specific techniques individuals were applying in their everyday life (as they were supposed to) nor if the uptake was as intended by the treatment rationale (e.g., if the thought protocols indeed contained alternative thoughts that might have led to improved emotional reactions).

Finally, while the operationalization of *expectations* seems somewhat straightforward, our questionnaire was limited to outcome expectations alone. A review of expectations in the setting of psychotherapy research stresses that other facets of this construct exist, such as treatment credibility or expectations towards the general utility of treatment (Constantino, Arnkoff, Glass, Ametrano, & Smith, 2011). However, the authors highlight that these constructs are highly-correlated and difficult to differentiate. Thus, while different options of assessing expectations and related facets exist (e.g., Patterson et al., 2014), it is unclear whether they might have incremental benefits over assessing expectations as done in STUDY 2.

Apart from the way constructs were assessed, it is also important to consider how often they were assessed. Following the recommendations on how to conduct studies on mechanisms

of change in psychotherapy (e.g.; Kazdin, 2007; see above for summary), we evaluated symptom-change and specific-treatment uptake during each module. Further, we considered the temporal ordering of important variables when modeling predictive paths (i.e., previous symptom change → mid-treatment alliance ratings → subsequent symptom change).

However, we failed to apply this design to all the variables in our model due to having otherwise overburdened participants with large batteries of weekly online questionnaires. Thus, expectations were assessed only at the beginning of treatment and extra-therapeutic stressor could only be considered with regard to baseline ratings and pre-post change.

Future applications. As a direction for future studies, it seems most reasonable to follow the solid foundation STUDY 2 has built, while overcoming some of its methodological flaws. In particular, I suggest a more comprehensive operationalization of the four-factor model with a broader set of variables, assessed on more measurement occasions and with a broader variety of methods.

For example, *extra-therapeutic factors* should consider baseline traits of the client (personality, inner resources). Further, daily hassles and uplifts as well as other extra-therapeutic changes (e.g., changing jobs or relationships) are best assessed with diary methods applied in individuals' everyday lives (e.g., ambulatory assessment) and would thus offer a more complete picture of relevant events and changes outside of the therapeutic setting. Recent studies demonstrated that these variables can be reliably and validly obtained by using smartphone apps that present individuals with questionnaires multiple times a day (Y. Chan et al., 2019; Starr & Hershenberg, 2017).

Mobile diary-tools might also be used to protocol whether participants were applying the *specific therapeutic techniques* outside of the online environment. Diary and tracking data could be complemented by ratings of the therapist on the quality and quantity with which specific treatment components have been used during the last session (or in the participant's homework).

Regarding *relationship factors*, text-analysis tools may be applied to the therapists' feedback in order to identify the amount of warmth, empathy, encouragement, or positive regard. Appropriate text-analytic methods have already been applied on written feedbacks in IBI (e.g., Paxling et al., 2013) or texts written by patients (e.g., Rosenbach & Renneberg, 2015). Consequently, it seems reasonable to include these methods within a more comprehensive study on contributing factors. Concerning the therapeutic alliance, the reliance on the WAI which was derived from the face-to-face setting should be reconsidered. The design of questionnaires that are specific to the IBI context might be necessary to obtain valid data on the perceived alliance in this setting (for an example see Berry et al., 2018).

As to the factor of *expectations,* it is important to broaden the scope from considering baseline expectations to evaluating changes of expectations over the course of treatment. The plausible assumption that expectations and therapeutic outcome interact with each other during treatment has been demonstrated in recent trials on in-treatment expectation change (Vîslă, Flückiger, Constantino, Krieger, & Holtforth, 2018).

Concordantly, independent of specific constructs, all variables of interest should be assessed multiple times and the analysis should ideally focus on facets of constructs or even single-symptoms where appropriate (for an example see Heinrich et al., 2018). In the context of depression for instance, this would allow for investigations on whether specific techniques do indeed change specific facets of the disorder (e.g., does anhedonia change after behavioral activation; are feelings of guilt decreased after cognitive bias modification). This type of research is not possible with the current focus on sum scores of questionnaires.

All recommendations for a future study on the basis of the four-factor-model and our own experiences are summarized in FIGURE 5.3.

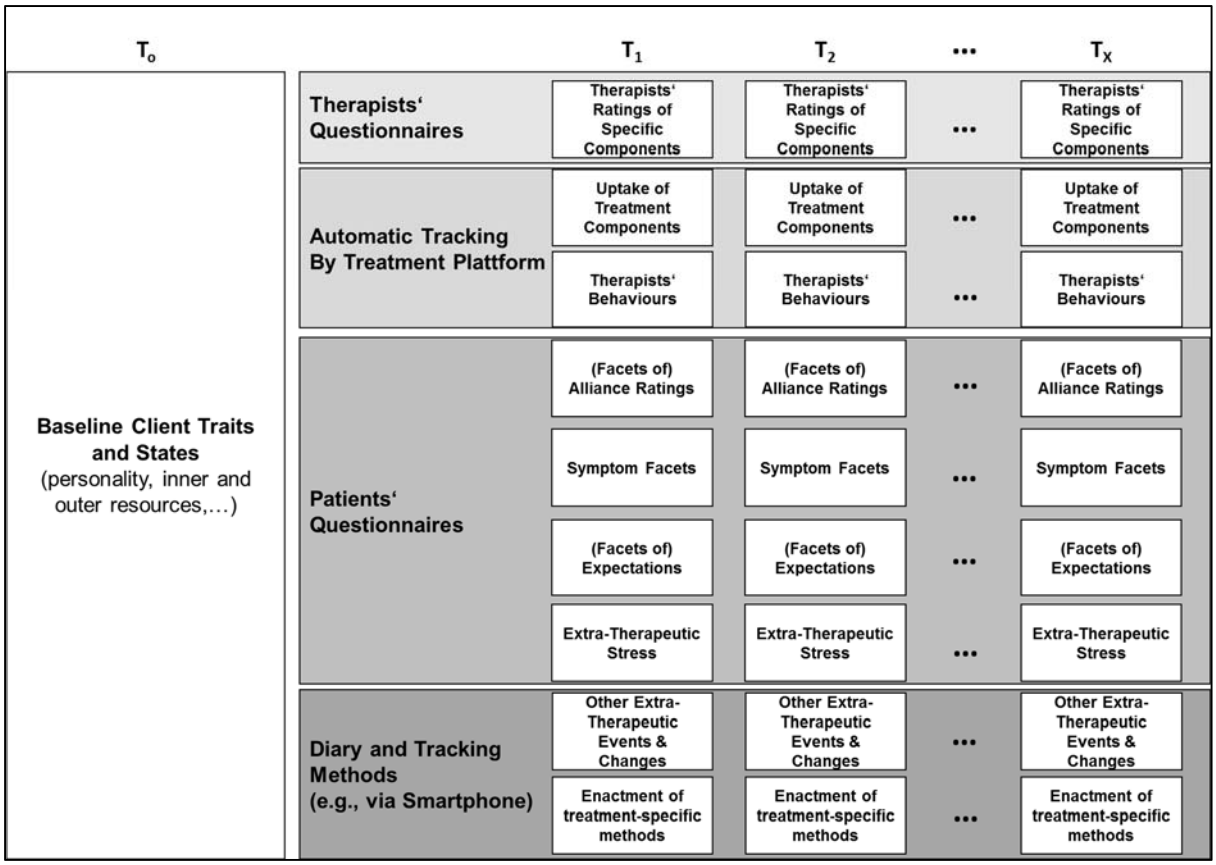


FIGURE 5.3. Illustration of a possible research design to study mechanisms of change in IBI.

5.3 Study 3: Patterns of change

5.3.1 Brief Summary of Findings

The third study investigated whether there are subgroups of individuals participating in standardized and individualized IBI for depression showing discernable patterns of symptom change. The study found that most individuals (62.5%) can be classified as "immediate improvers" with substantial improvement before the start of treatment. Another class (37.4%) of individuals was labeled "delayed improvers" for their symptoms did not change up until week three of treatment. Overall, these individuals displayed slightly smaller change in depressive symptoms than the class of "immediate improvers". Individuals with higher perceived social support had higher odds of being classified as "immediate improvers". In contrast, individuals fulfilling the criteria for MDD in a structured clinical interview (SCID-I) and individuals with high outcome expectations had higher odds of being classified as "delayed improvers". No differences between the standardized and individualized feedback condition emerged concerning the number of discernable classes or the resulting change patterns. FIGURE 5.4 illustrates the core findings of STUDY 3.

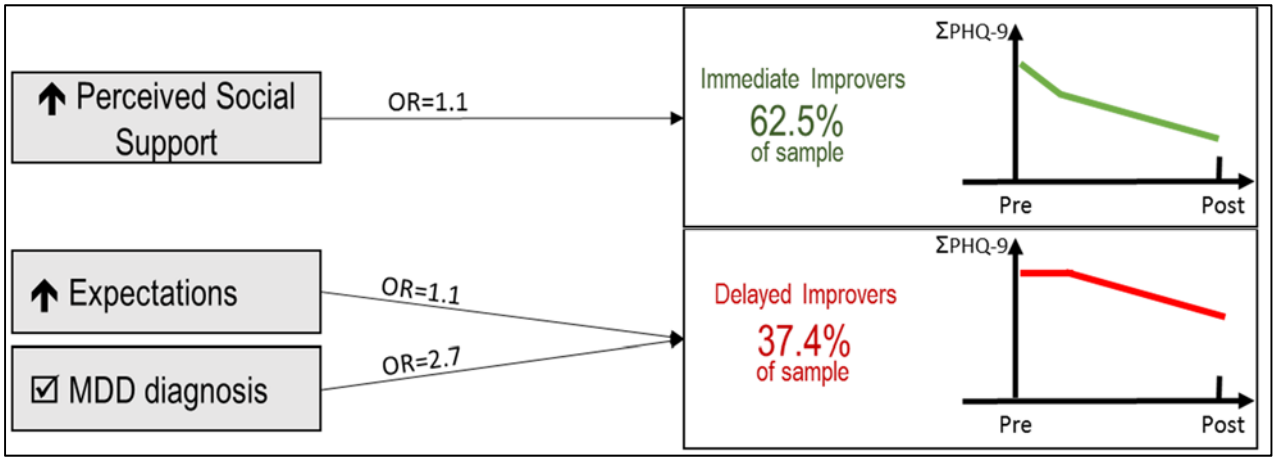


FIGURE 5.4. Simplified illustration of discernable classes of symptom change in IBI for depression with predictors of class membership. OR, odds ratios indicating the increase in odds of belonging to the respective class if the variable in question is increased by one unit of measurement (social support, expectations) or is categorized as "present" (MDD diagnosis).

5.3.2 Discussion

Overall, results of STUDY 3 complement findings of other studies on IBI for depression (for a detailed review see CHAPTER 3 and SECTION 1.6.4). Three studies have also identified two distinct classes of symptom change (Batterham et al., 2017, 2018; Sunderland et al., 2012). In these studies, a majority of participants with early or more profound improvements and a minority of individuals with delayed or stagnant symptom change were identified. While these studies targeted insomnia or suicidal ideation (Batterham et al., 2017, 2018) or reported on generalized levels of psychological distress (Sunderland et al., 2012), we investigated depressive symptom change during IBI for depression. Furthermore, we modeled multiple time periods during the entire duration of IBI, rather than solely focusing on pre-post-change (Batterham et al., 2017, 2018) or early change (Lutz et al., 2017).

Regarding predictors of class membership, our study confirmed previous findings showing that more severely depressed participants or participants with a confirmed disorder of MDD have increased odds of membership in less favorable classes[16] (Batterham et al., 2017, 2018; Lutz et al., 2017; Sunderland et al., 2012). Other findings on predictors of membership in less favorable classes were not consistent across the previously conducted trials on IBI for depression and include younger age, unemployment, and not being in a relationship (Batterham et al., 2017, 2018). Given that strong positive associations exist between interpersonal relationships and social support (Feeney & Collins, 2015), our finding on perceived social support being predictive of immediate improvement seems consistent with the finding that individuals in a relationship tend to be in a more favorable class (Batterham et al., 2018). While we did investigate age and employment status, we did not find them to be predictive of class membership. We rather found that unemployment was a *general* predictor of symptom

[16] The classes have been labeled differently in the trials. The term "less or least favorable class" was used as a descriptive umbrella, that stands for the the minority class in all trials that was characterized by mostly stable or deteriorating symptoms over time (see TABLE 1.5 in CHAPTER 1 for details).

deterioration over time without predictive power for class membership. This also highlights an important difference between our work and previous studies; we separately considered general predictors of initial symptom load and symptom change across classes. Given that there are few studies on IBI with significant differences in trial design, our results will be compared to findings from studies on patterns of change that have been published for the setting of face-to-face-psychotherapy.

Broader context. While there are some studies that have applied Growth Mixture Modeling (GMM) to face-to-face psychotherapy, the method is still relatively rarely used. For example, a recent summary on in-treatment depressive symptom change reported on six trials (Altmann et al., 2015). The number of discernable classes varied widely from two to five classes (Cuijpers et al., 2005; Gueorguieva, Mallinckrodt, & Krystal, 2011; Hunter, Muthén, Cook, & Leuchter, 2010; Keller & Hautzinger, 2007; Lutz et al., 2009; Pöhlmann et al., 2008).

When comparing the number of resulting classes across trials it is important to note that two studies reporting four or more classes investigated in-treatment and post-treatment changes of up to 18 months after treatment, potentially enabling a broader heterogeneity in symptom courses in the long term (Pöhlmann, et al., 2008; Cuijpers et al., 2005). Given the length of our IBI, it seems most appropriate to compare our findings to those from face-to-face psychotherapy studies that investigated symptom change within a comparable timeline.

A recent study summarized findings from such trials, investigating symptom change across disorder categories during the *first eight weeks* of face-to-face psychotherapy. The authors identified eight trials that found one to five discernable classes (Koffmann, 2018). As only two of these trials were primarily focused on depressive symptom change (Keller & Hautzinger, 2007; Lutz et al., 2009), I will review these two trials in detail.

First, Keller and Hautzinger (2007) investigated depressive symptom trajectories in a German sample of 341 inpatients and outpatients diagnosed with MDD. These patients were randomized to either receive medication, medication and CBT, or CBT only. The study design

featured baseline and post-assessments (after eight weeks), as well as weekly assessments of depressive symptoms (IDS; Hautzinger, 2003). Across treatment conditions the authors found four discernable classes of patients with symptom courses labeled as "late response" (5.9% of sample), "quick response" (2.7%), "slow and steady improvement" (65.2%), and "stagnant symptoms" (26.2%).

Second, Lutz et al. (2009), reanalyzed data from a sample of 250 outpatients diagnosed with MDD in the United States that scored 14 or above on the Hamilton Rating Scale for Depression (Elkin, Parloff, Hadley, & Autry, 1985). Patients were randomized to one of four conditions (CBT, Interpersonal Psychotherapy, psychopharmacological treatment with "clinical management"[17], or pill placebo with "clinical management"). Symptoms were assessed with the Beck Depression Inventory at intake, as well as four and eight weeks after the beginning of treatment. Based on these measurements, Lutz et al. (2009) found three different patterns of change across conditions: "moderate to severe depression with moderate early improvement" (19.8% of the sample), "moderate to severe depression with rapid early improvement" (61.1%), and "mild to moderate depression with moderate improvement" (19.1%). As is evident from the labels, these classes differed not only with regard to patterns of change but also with regard to initial symptom severity. Consequently, the authors reported initial symptom severity to be a significant predictor of class membership. The three classes were distributed evenly across the four conditions, indicating that symptom change patterns are not dependent on allocation. While no other predictors of class membership were investigated, the authors speculate that treatment motivation and expectancy might predict membership in the class that experiences rapid improvement (Lutz et al., 2009).

[17] According to the original study design "clinical management" entails meetings with psychiatrists and pharmacotherapists that provide a generally supportive atmosphere as well as bi-weekly appointments in order to obtain blood levels of the substance under research (Imipramine; same procedure as pill placebo group).

In summary, both trials found a class with rapid early improvement. However, the size of this class differed substantially between trials (2.7% vs. 61.1%). Furthermore, the two studies differed with regard to class enumeration and the resulting patterns of symptom change. When integrating findings from the face-to-face and the IBI-setting, the only consistent finding seems to be that higher baseline symptom severity is associated with membership in less favorable classes of symptom change. Class numbers, other associated predictors and patterns of symptom change are inconsistent both between settings (IBI vs. face-to-face psychotherapy) and within each of the respective settings.

In search for possible reasons for the lack of consistency in GMM treatment studies in general, Koffmann (2018) summarized that "there are important sample and design differences among the studies. Sample sizes differed dramatically, patient populations varied, lengths of treatment differed, and the treatments offered ran the gamut from treatment as usual to various manualized approaches, among many other differences. In short, no study replicated any other study" (p. 9). However, the differences between the trials might be due to specific methodological problems that affect our trial as well as other studies on patterns of change. Apart from these problems, the following section will also highlight restrictions that arise from the choice of questionnaires and measurement occasions in our trial.

Specific problems. *First*, all studies on patterns of change have to be considered exploratory in nature. That is due to the fact that multiple decisions have to be made by the researchers. For example, assumptions about differences and similarities between hypothesized classes might exist that influence the way a researcher formulates the model that is tested: Do different treatment groups display the same patterns of change or could their data be merged? Do researchers assume that the change follows a certain form and are reasonable alternative models tested (e.g. quadratic or linear symptom change)? Are the within-class variance–covariance structures (which stand for variability of the curve-specific parameters, e.g. deviations of observed values from the estimated change curve) the same across all classes and

measurement occasions? All these decisions might lead to different restrictions imposed on the compared models and might explain divergence between our trial and others.

Additionally, a wide array of possible indices and tests can be considered when deciding between models suggesting different numbers of classes (e.g., *Akaike Information Criteria*, AIC; *Corrected AIC*, CAIC; *Bayes Information Criteria*, BIC; *sample-size adjusted BIC*, aBIC; *Vuo-Long-Mendel Likelihood Ratio-Test*, VLM-LRT, entropy). While all studies reviewed by Koffmann (2018) use BIC for this decision, the number and choice of other indices that are considered vary between studies. That is problematic, given that different indices might arrive at different conclusions (e.g., Diallo, Morin, & Lu, 2016). In that case, the final decision on the number of derived classes has to include visual inspection of patterns of change in different models and a judgement based on theoretical considerations about the probable number of classes.

Second, so far all studies on GMM have considered only symptom change on the basis of sum scores of questionnaires. The limits of this approach have already been discussed with regard to mechanisms of change (see SECTION 5.2.2), but they apply to the idea of modeling symptom courses in IBI as well. As an illustrative example, some participants in our IBI reported that the modules on behavioral activation helped them with their anhedonia. At the same time, they felt that they led a tedious life in comparison to others due to the fact they were not able to protocol many activities in their daily planner. In a hypothetical scenario, these participants might have reported reduced anhedonia and increased feelings of worthlessness at the same time. The sum score of a depression screening instrument, such as the one applied in our trial (PHQ-9) would have indicated a symptom pattern characterized by stagnation for this participant during modules on behavioral activation, falsely suggesting that nothing changed.

Third, patterns of depressive symptom change have been modeled on the basis of measurements in weekly or broader intervals. This is potentially problematic, given previous research that has shown that individuals tend to recollect unpleasant emotions (i.e., anger,

anxiety, sadness) more vividly than pleasant ones in retrospective assessments (Urban, Charles, Levine, & Almeida, 2018). The effect was especially pronounced in individuals with (a history of) depression, questioning the validity of retrospective symptom surveys in this specific population (Urban et al., 2018; Wenze, Gunthert, & German, 2012). Further, the retrospective approach might conceal important differences between individuals that may arise during shorter periods of time. One illustrative example concerns the investigation of "chronotypes" in depression:

Previous studies highlighted that depressive symptoms are dependent on the time of day and that individuals belong to different "chronotypes" (e.g. "evening types", "eveningness" vs. "morning types", "morningness") (e.g., Jankowski, 2016). These chronotypes are likely to show distinct symptom courses during the day.

Discerning symptom trajectory classes according to chronotypes and investigating socio-demographic and clinical characteristics of individuals in these classes has potential for improving treatment. For example, studies showed that "morning types" responded better to treatment with antidepressants than "evening types" (Corruble et al., 2014) and that "eveningness" was related to insomnia, suicidality, more severe depressive symptoms and non-remission of MDD (J. W. Y. Chan et al., 2014; Müller, Kundermann, & Cabanel, 2016). Thus, patients might profit from treatment methods that are tailored to their chronotype (e.g., coping with a circadian rhythm that is shifted towards the evening). However, chronotypes have mostly been identified through the use of questionnaires that assess "morningness-eveningness" as a trait (e.g., Horne & Östberg, 1976), instead of tracking the symptoms themselves over the course of the day, which would likely be more reliable (e.g., Jankowski, 2016). That is due to the fact that this approach requires investigations of symptom developments through multiple measurements in shorter periods of time (i.e., multiple times a day), which are usually resource-intensive and obtrusive, especially when done in paper-pencil format.

In summary, current shortcomings in GMM-based studies on patterns of change in treatment settings offline and online concern inconsistencies in restrictions imposed on models for class enumeration across trials and a sole reliance on sum scores of symptoms that are assessed relatively infrequently.

Future applications. I will briefly outline some directions for future studies that might contribute to overcoming the three previously raised issues and thus augment the findings derived from our trial.

Concerning the exploratory nature of GMM, while the method will always have a certain potential for subjective influences of researchers, Ram and Grimm (2009) have argued that "the method should be a guided and constrained exploration of the data" (p.568). The authors made a first step towards unifying the heterogeneous field of GMM research by suggesting a path diagram of procedures that studies should follow when implementing GMM. In particular, they suggested the formulation of a-priori hypotheses of expected class numbers and their symptom developments. In addition, they provide guidelines on how to perform model specification, model estimation and model selection (details on these steps can be obtained from the original publications and from STUDY 3/APPENDIX C). Similar guidelines and instructions were provided by other authors (e.g., Diallo et al., 2016; Jung & Wickrama, 2008; Koffmann, 2018; Muthén, Brown, Leuchter, & Hunter, 2010). Koffmann (2018), for instance, further proposed that comparable restrictions on models should be imposed across trials and commended a certain unity in applying the BIC as a criterion for deciding between different models of change in recent studies.

Importantly, the multiple different steps of the analysis and the flexibility in decision-making within GMM emphasize the need for more transparency. A-priori hypotheses and each step of the analysis (including every decision and its basis) should be detailed in an appendix to papers published in the future. This was done for STUDY 3 (see APPENDIX C), but most other

trials have not provided the necessary information to evaluate the similarities and differences to other studies in the field.

Apart from the theoretical propositions on how to perform GMM, future studies should also base their methodology on more recent simulation studies that have highlighted assets and pitfalls of different approaches to GMM. Specifically, Diallo et al. (2016) investigated the consequences of different forms of restrictions imposed on the models. They concluded that more unrestricted models should be preferred in most circumstances, since unnecessary restrictions tend to lead to inflated class numbers (e.g., Diallo et al., 2016; Peugh & Fan, 2012). However, unrestricted models oftentimes do not converge and require larger computational resources than models with restrictions. Thus, the authors suggested a number of steps that can be taken if these models do not converge properly (for a detailed discussion see Diallo et al., 2016).

Further, different simulation studies have investigated the performance of different model fit indices within GMM under different circumstances (e.g., Diallo et al., 2016; Nylund, Asparouhov, & Muthén, 2007; Tofighi & Enders, 2008). Overall, they discourage the use of the AIC. While two studies reported the best results for the BIC in most scenarios (Nylund et al., 2007; Tofighi & Enders, 2008), none of the information criteria was consistently superior to others across simulation studies and conditions. Consequently, Diallo et al. (2016) summarized that "each of the four remaining [information criteria] (CAIC, BIC, SCAIC, SBIC) outperformed the others within specific design conditions, and thus all appear useful and complementary" (p. 526). Given the exploratory nature of most GMM studies, it might be difficult to know a-priori which of the information criteria is most useful. I recommend considering (and reporting) multiple information criteria when selecting statistical models.

Concerning the focus on sum scores in previous studies, future studies should focus on modeling of change on the level of single symptoms. Decomposing the sum score in patterns

of change research might help to disentangle the true effects of an intervention and thus help with identifying both the strengths and weaknesses of a given treatment module. In the hypothetical example from the previous section where both anhedonia and self-worth were reduced during behavioral activation, a symptom-focused approach might have confirmed that the behavioral activation module works as intended, i.e. reduces anhedonia. At the same time, it would be revealed that this module might have side-effects, such as reduced feeling of self-worth. Accordingly, actions could be taken to decrease this side-effect, e.g., by improved instructions for participants who have only few entries in their daily planner. An interesting example of a more symptom-oriented approach to the Beck Depression Inventory-II (and the Patient-Health-Questionnaire-9) is discussed elsewhere (Heinrich et al., 2018).

Concerning the issue of infrequent assessments, future studies should expand our findings by assessing symptom courses with more fine-grained temporal resolution to investigate phenomena where short-term changes are relevant, such as chronotypes and their reaction to treatment modules in IBI. This approach is very promising, given that recent studies suggested that ecological momentary assessment (e.g., asking for an individual's depression-related symptoms multiple times a day through ambulatory assessment) might lead to more unbiased results on depressive symptom severity versus retrospective questionnaires (Moore, Depp, Wetherell, & Lenze, 2016; Wenze & Miller, 2010).

5.4. Conclusion and Outlook

The three studies presented in this doctoral thesis contribute knowledge on mechanisms of change in IBI for mild to moderate depression and highlight ways to improve future IBI-related research and practice.

STUDY 1 showed that well-designed unguided interventions with initial phone calls and written contact-on-demand options have the potential to be as efficacious as (guided) interventions featuring individualized feedback. STUDY 2 demonstrated that changes in extra-therapeutic stressors and agreement on tasks and goals are important predictors of symptom change in IBI for depression. Further, this study showed that initial extra-therapeutic stress levels have negative effects on in-treatment depressive symptom development in unguided IBI only, whereas higher bond ratings were predictive of symptom deterioration in the final week of treatment in guided IBI only. STUDY 3 identified two discernable classes of symptom change in IBI for depression with about two-thirds of individuals responding immediately and profoundly to the offered treatment ("immediate responders") – partly even before it commenced – and about one-third showing a delayed and weakened response ("delayed responders"). Individuals with a diagnosis of MDD based on a structured clinical interview, with low perceived social support, and with high expectations were more likely to experience delayed response.

On a general level, the results highlight the utility of fully-standardized IBI, which seems to have been underestimated on the basis of previously conducted trials and meta-analyses. While there are slight differences in the importance of different factors contributing to symptom change between the experimental conditions investigated in our trial (STUDY 2), they differed neither concerning overall efficacy (STUDY 1) nor on any features related to patterns of change (STUDY 3). Thus, the findings encourage further studies on the efficacy and effectiveness of IBI that differ in the quality and quantity of human contact. Careful and

empirically-based discussions of the costs and benefits of different forms of IBI need to take place in the future, in order to determine which IBI is best offered to which individual.

On a more specific level, our studies pinpoint ways to improve the yield of both guided and unguided IBI for depression. STUDY 1 shows that individuals profit from guidance in IBI through lowered dropout rates. At the same time, results from STUDY 2 indicate that individuals with a strong bond to their counselor might suffer from "parting pains" as is observable through an increase in depressive symptoms at the end of treatment. Currently, IBI are usually short and the end of treatment is quite abrupt. Thus, a more thorough preparation of the discontinuation of treatment and the offering of tools to cope well with farewell might be helpful.

Additionally, STUDIES 2 and 3 underscore the value of obtaining baseline diagnostic information from participants in IBI. This allows providing a treatment experience that is better tailored to their specific situation. Firstly, individuals with low social support and high expectations tended to show delayed response to treatment in STUDY 2, thus suggesting the provision of tools for realistic expectation management (e.g., Ekberg et al., 2015) or treatment approaches that are better suited to individuals with low social support (e.g., Lindfors, Ojanen, Jääskeläinen, & Knekt, 2014), respectively. Secondly, participants with high levels of extra-therapeutic stress at the beginning of treatment might profit more from guided IBI. The results of STUDY 2 suggest that the human contact provided in guided IBI has a "stress buffering" effect on these individuals.

Overall, the findings of all three studies are derived from a large randomized controlled trial (N = 1089), that recruited self-selected and externally-selected (by insurance representatives) individuals drawn from the client base of a large public health care company. These aspects testify to the validity of conclusions drawn in this thesis. However, it is important to note that strict inclusion and exclusion criteria and our study design resulted in a sample of non-suicidal individuals with mild to moderate symptom severity that participated in mostly text-based IBI. These limitations reduce the generalizability of our findings to other forms of

IBI and differently-structured clinical populations. Importantly, these are common trial-specific restrictions that can be overcome in future studies with different target populations or variations of interventions under research. Nonetheless, this thesis highlights the potential of IBI to study mechanisms of change due to the ease of acquiring large sample sizes, implementing multiple assessments and controlling numerous variables through the tracking of all events within the internet-platform (e.g., login-duration, uptake of specific treatment components, text-based interaction). Concordantly, the suggestions for future research in the previous chapter have alluded to the fact that continuous technological developments might lead to studies that make use of an even broader scope of diagnostic tools.

Specifically, I argued that future studies should implement more frequent assessments and consider a broader set of assessment methods in order to offer a more comprehensive picture of the individuals' thoughts, emotions, and behaviors both within and outside of IBI. For that purpose, I suggested the use of smartphones (ambulatory assessment). A growing number of such studies are already being done, albeit not necessary to investigate mechanisms of change in IBI for depression (e.g., BinDhim et al., 2015; Y. Chan et al., 2019). Notably, the possibilities far exceed the use of ambulatory assessment. For example, researchers have started to exploit the sensory capabilities of smartphones and wearables (e.g., smart-watches) to track a wide variety of variables (e.g., movement and GPS-location, app use, sleep patterns, heart rate) (e.g., Ben-Zeev, Scherer, Wang, Xie, & Campbell, 2015). The expansion of possible sources of data are tempting for research and practice and it seems promising to take advantage of them to understand, predict, and change human behavior in the future within IBI and beyond. Nevertheless, there are ethical challenges inherent in this approach and researchers should keep in mind that the accumulation of large amounts of data bear a growing potential for their misuse as well – while the utility for patients is yet to be conclusively demonstrated in studies in the future (Zagorscak & Knaevelsrud, 2019).

As an overall conclusion, this thesis demonstrated the similarities between standardized and individualized feedback in IBI for depression concerning efficacy and patterns of change. At the same time the results highlight the importance of individual characteristics when deciding on which type of IBI to offer in what composition. Individuals' socio-demographic and clinical features influence both the mechanisms of change and symptom courses in IBI for depression, and they have differential effects depending on whether the feedback is standardized or individualized.

More fine-grained research that offers a more comprehensive picture of participants' thoughts, feelings and behaviors in IBI is needed to offer interventions that are individually tailored. Significant technological developments are likely to help research and practice in this endeavor, as long as the possible pitfalls of amassing data are carefully weighed against the benefits.

REFERENCES FOR CHAPTER 1 AND CHAPTER 5

(References for each paper listed separately in chapters 2-4)

Aderka, I. M., Nickerson, A., Boe, 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. https://doi.org/10.1037/a0026455

Ahn, H. N., & Wampold, B. E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology*, *48*(3), 251–257. https://doi.org/10.1037//0022-0167.48.3.251

Altmann, U., Steyer, R., Kramer, D., Steffanowski, A., Wittmann, W. W., von Heymann, F., … Strauss, B. (2015). Typical patterns of depressive disorders during outpatient psychotherapy and their prediction. *Zeitschrift Fur Psychosomatische Medizin Und Psychotherapie*, *61*(2), 156–172. https://doi.org/10.13109/zptm.2015.61.2.156

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th Revised edition). Washington, D.C: American Psychiatric Association Publishing.

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. https://doi.org/10.1080/16506070903318960

Andersson, G., & Hedman, E. (2013). Effectiveness of Guided Internet-Based Cognitive Behavior Therapy in Regular Clinical Settings. *Verhaltenstherapie*, *23*(3), 140–148. https://doi.org/10.1159/000354779

Andersson, G., Sarkohi, A., Karlsson, J., Bjarehed, J., & Hesser, H. (2013). Effects of Two Forms of Internet-Delivered Cognitive Behaviour Therapy for Depression on Future Thinking. *Cognitive Therapy and Research*, *37*(1), 29–34. https://doi.org/10.1007/s10608-012-9442-y

Andersson, G., & Titov, N. (2014). Advantages and limitations of Internet-based interventions for common mental disorders. *World Psychiatry*, *13*(1), 4–11. https://doi.org/10.1002/wps.20083

Arias-de la Torre, J., Vilagut, G., Martín, V., Molina, A. J., & Alonso, J. (2018). Prevalence of major depressive disorder and association with personal and socio-economic factors. Results for Spain of the European Health Interview Survey 2014–2015. *Journal of Affective Disorders*, *239*, 203–207. https://doi.org/10.1016/j.jad.2018.06.051

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. https://doi.org/10.1371/journal.pone.0098118

Asarnow, J., & Ougrin, D. (2017). From efficacy to pragmatic trials: does the dodo bird verdict apply? *Lancet Psychiatry*, *4*(2), 84–85. https://doi.org/10.1016/S2215-0366(16)30404-7

Baardseth, T. P., Goldberg, S. B., Pace, B. T., Wislocki, A. P., Frost, N. D., Siddiqui, J. R., … Wampold, B. E. (2013). Cognitive-behavioral therapy versus other therapies: Redux. *Clinical Psychology Review*, *33*(3), 395–405. https://doi.org/10.1016/j.cpr.2013.01.004

Barth, J., Munder, T., Gerger, H., Nüesch, E., Trelle, S., Znoj, H., … Cuijpers, P. (2016). Comparative Efficacy of Seven Psychotherapeutic Interventions for Patients with Depression: A Network Meta-Analysis. *FOCUS*, *14*(2), 229–243. https://doi.org/10.1176/appi.focus.140201

Bathje, G. J., Kim, E., Rau, E., Bassiouny, M. A., & Kim, T. (2014). Attitudes toward Face-to-Face and Online Counseling: Roles of Self-Concealment, Openness to Experience, Loss of Face, Stigma, and Disclosure Expectations among Korean College Students. *International Journal for the Advancement of Counselling*, *36*(4), 408–422. https://doi.org/10.1007/s10447-014-9215-2

Batterham, P. J., Christensen, H., Mackinnon, A. J., Gosling, J. A., Thorndike, F. P., Ritterband, L. M., … Griffiths, K. M. (2017). Trajectories of change and long-term outcomes in a randomised controlled trial of internet-based insomnia treatment to prevent depression. *BJPsych Open*, *3*(5), 228–235. https://doi.org/10.1192/bjpo.bp.117.005231

Batterham, P. J., Spijker, B. A. J. van, Mackinnon, A. J., Calear, A. L., Wong, Q., & Christensen, H. (2018). Consistency of trajectories of suicidal ideation and depression symptoms: Evidence from a randomized controlled trial. *Depression and Anxiety*. https://doi.org/10.1002/da.22863

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. https://doi.org/10.1016/j.invent.2014.08.003

Beattie, A., Shaw, A., Kaur, S., & Kessler, D. (2009). Primary-care patients' expectations and experiences of online cognitive behavioural therapy for depression: a qualitative study. *Health Expectations*, *12*(1), 45–59. https://doi.org/10.1111/j.1369-7625.2008.00531.x

Beesdo-Baum, K., Knappe, S., Einsle, F., Knothe, L., Wieder, G., Venz, J., … Bergmann, A. (2018). Wie häufig werden Patienten mit depressiven Störungen in der hausärztlichen Praxis erkannt? *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, *61*(1), 52–64. https://doi.org/10.1007/s00103-017-2662-2

Bell, E. C., Marcus, D. K., & Goodlad, J. K. (2013). Are the parts as good as the whole? A meta-analysis of component treatment studies. *Journal of Consulting and Clinical Psychology*, *81*(4), 722–736. https://doi.org/10.1037/a0033004

Ben-Zeev, D., Scherer, E. A., Wang, R., Xie, H., & Campbell, A. T. (2015). Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health. *Psychiatric Rehabilitation Journal*, *38*(3), 218–226. https://doi.org/10.1037/prj0000130

Berger, T. (2017). The therapeutic alliance in internet interventions: A narrative review and suggestions for future research. *Psychotherapy Research*, *27*(5), 511–524. https://doi.org/10.1080/10503307.2015.1119908

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 (Chicago, Ill.)*, *51*(2), 207–219. https://doi.org/10.1037/a0032527

Berger, T., Caspar, F., Richardson, R., Kneubühler, B., Sutter, D., & Andersson, G. (2011). Internet-based treatment of social phobia: A randomized controlled trial comparing unguided with two types of guided self-help. *Behaviour Research and Therapy*, *49*(3), 158–169. https://doi.org/10.1016/j.brat.2010.12.007

Berger, T., Hammerli, K., Gubser, N., Andersson, G., & Caspar, F. (2011). Internet-based treatment of depression: a randomized controlled trial comparing guided with unguided self-help. *Cognitive Behaviour Therapy*, *40*(4), 251–266. https://doi.org/10.1080/16506073.2011.616531

Berger, T., Stolz, T., & Schulz, A. (2013). Internetbasierte geleitete Selbsthilfeansätze bei Angststörungen und Depressionen. *E-Beratungsjournal.Net*, *9*(2), 1–16.

Berry, K., Salter, A., Morris, R., James, S., & Bucci, S. (2018). Assessing Therapeutic Alliance in the Context of mHealth Interventions for Mental Health Problems: Development of the Mobile Agnew Relationship Measure (mARM) Questionnaire. *Journal of Medical Internet Research*, *20*(4), e90. https://doi.org/10.2196/jmir.8252

Beutler, L. E. (2002). The Dodo Bird Is Extinct. *Clinical Psychology: Science and Practice*, *9*(1), 30–34. https://doi.org/10.1093/clipsy/9.1.30

Bhattacharya, R., Kelley, G., & Bhattacharjee, S. (2012). Long-Term Follow-up Effects of Computerized or Internet-Based Cognitive Behavioral Therapy for Depression and Anxiety: A Meta-Analysis. *Value in Health*, *15*(4), A82–A82.

BinDhim, N. F., Shaman, A. M., Trevena, L., Basyouni, M. H., Pont, L. G., & Alhawassi, T. M. (2015). Depression screening via a smartphone app: cross-country user characteristics and feasibility. *Journal of the American Medical Informatics Association*, *22*(1), 29–34. https://doi.org/10.1136/amiajnl-2014-002840

Birnbaum, M. L., Rizvi, A. F., Correll, C. U., Kane, J. M., & Confino, J. (2017). Role of social media and the Internet in pathways to care for adolescents and young adults with psychotic disorders and non-psychotic mood disorders. *Early Intervention in Psychiatry*, *11*(4), 290–295. https://doi.org/10.1111/eip.12237

Blanco, C., Okuda, M., Markowitz, J. C., Liu, S.-M., Grant, B. F., & Hasin, D. S. (2010). The Epidemiology of Chronic Major Depressive Disorder and Dysthymic Disorder: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *The Journal of Clinical Psychiatry*, *71*(12), 1645–1656. https://doi.org/10.4088/JCP.09m05663gry

Boettcher, J., Berger, T., & Renneberg, B. (2012). Does a Pre-Treatment Diagnostic Interview Affect the Outcome of Internet-Based Self-Help for Social Anxiety Disorder? A Randomized

Controlled Trial. *Behavioural and Cognitive Psychotherapy*, *40*(5), 513–528. https://doi.org/10.1017/S1352465812000501

Boettcher, J., Renneberg, B., & Berger, T. (2013). Patient Expectations in Internet-Based Self-Help for Social Anxiety. *Cognitive Behaviour Therapy*, *42*(3), 203–214. https://doi.org/10.1080/16506073.2012.759615

Bonabi, H., Müller, M., Ajdacic-Gross, V., Eisele, J., Rodgers, S., Seifritz, E., … Rüsch, N. (2016). Mental Health Literacy, Attitudes to Help Seeking, and Perceived Need as Predictors of Mental Health Service Use: A Longitudinal Study. *The Journal of Nervous and Mental Disease*, *204*(4), 321. https://doi.org/10.1097/NMD.0000000000000488

Boss, L., Lehr, D., Berking, M., Riper, H., Schaub, M. P., & Ebert, D. D. (2015). Evaluating the (cost-)effectiveness of guided and unguided Internet-based self-help for problematic alcohol use in employees - a three arm randomized controlled trial. *Bmc Public Health*, *15*, 1043. https://doi.org/10.1186/s12889-015-2375-0

Bower, P., Kontopantelis, E., Sutton, A., Kendrick, T., Richards, D. A., Gilbody, S., … Liu, E. T.-H. (2013). Influence of initial severity of depression on effectiveness of low intensity interventions: meta-analysis of individual patient data. *British Medical Journal*, *346*, 540–551. https://doi.org/10.1136/bmj.f540

BPtK (Ed.). (2018). *Ein Jahr nach der Reform der Psychotherapie-Richtlinien. Wartezeiten 2018*. Berlin: Editor.

Carlbring, P., Andersson, G., Cuijpers, P., Riper, H., & Hedman-Lagerlöf, E. (2018). Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. *Cognitive Behaviour Therapy*, *47*(1), 1–18. https://doi.org/10.1080/16506073.2017.1401115

Carroll, L. (1922). *Alice's Adventures in Wonderland*. Boston, MA: International Pocket Library.

César González-Blanch, & Laura Carral-Fernández. (2017). CAGE UP DODO, PLEASE! THE TALE OF ALL PSYCHOTHERAPIES BEING EQUALLY EFFECTIVE. *Papeles Del Psicólogo*, *38*(2), 94–106. https://doi.org/10.23923/pap.psicol2017.2828

Chan, J. W. Y., Lam, S. P., Li, S. X., Yu, M. W. M., Chan, N. Y., Zhang, J., & Wing, Y.-K. (2014). Eveningness and Insomnia: Independent Risk Factors of Nonremission in Major Depressive Disorder. *Sleep*, *37*(5), 911–917. https://doi.org/10.5665/sleep.3658

Chan, Y., So, S. H., Mak, A. D. P., Siah, K. T. H., Chan, W., & Wu, J. C. Y. (2019). The temporal relationship of daily life stress, emotions, and bowel symptoms in irritable bowel syndrome—Diarrhea subtype: A smartphone-based experience sampling study. *Neurogastroenterology & Motility*, *0*(0), e13514. https://doi.org/10.1111/nmo.13514

Chang, C. (2008). Increasing mental health literacy via narrative advertising. *Journal of Health Communication*, *13*(1), 37–55. https://doi.org/10.1080/10810730701807027

Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., … Geddes, J. R. (2018). Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *The Lancet*, *391*(10128), 1357–1366. https://doi.org/10.1016/S0140-6736(17)32802-7

Ciuca, A. M., Berger, T., Crişan, L. G., & Miclea, M. (2018). Internet-based treatment for panic disorder: A three-arm randomized controlled trial comparing guided (via real-time video sessions) with unguided self-help treatment and a waitlist control. PAXPD study results. *Journal of Anxiety Disorders*, *56*, 43–55. https://doi.org/10.1016/j.janxdis.2018.03.009

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Conner, K. O., Copeland, V. C., Grote, N. K., Koeske, G., Rosen, D., Reynolds, C. F., & Brown, C. (2010). Mental Health Treatment Seeking Among Older Adults With Depression: The Impact of Stigma and Race. *American Journal of Geriatric Psychiatry*, *18*(6), 531–543. https://doi.org/10.1097/JGP.0b013e3181cc0366

Constantino, M. J., Arnkoff, D. B., Glass, C. R., Ametrano, R. M., & Smith, J. Z. (2011). Expectations. *Journal of Clinical Psychology*, *67*(2), 184–192. https://doi.org/10.1002/jclp.20754

Constantino, M. J., Vîslă, A., Coyne, A. E., & Boswell, J. F. (2018). A meta-analysis of the association between patients' early treatment outcome expectation and their posttreatment outcomes. *Psychotherapy*, *55*(4), 473–485. https://doi.org/10.1037/pst0000169

Corruble, E., Frank, E., Gressier, F., Courtet, P., Bayle, F., Llorca, P.-M., … Gorwood, P. (2014). Morningness–eveningness and treatment response in major depressive disorder. *Chronobiology International*, *31*(2), 283–289. https://doi.org/10.3109/07420528.2013.834924

Cosci, F., Fava, G. A., & Sonino, N. (2015). Mood and Anxiety Disorders as Early Manifestations of Medical Illness: A Systematic Review. *Psychotherapy and Psychosomatics*, *84*(1), 22–29. https://doi.org/10.1159/000367913

Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie*, *58*(7), 376–385.

Cuijpers, P., Cristea, I. A., Karyotaki, E., Reijnders, M., & Hollon, S. D. (2019). Component studies of psychological treatments of adult depression: A systematic review and meta-analysis. *Psychotherapy Research*, *29*(1), 15–29. https://doi.org/10.1080/10503307.2017.1395922

Cuijpers, P., Driessen, E., Hollon, S. D., van Oppen, P., Barth, J., & Andersson, G. (2012). The efficacy of non-directive supportive therapy for adult depression: A meta-analysis. *Clinical Psychology Review*, *32*(4), 280–291. https://doi.org/10.1016/j.cpr.2012.01.003

Cuijpers, P., Kleiboer, A., Karyotaki, E., & Riper, H. (2017). Internet and mobile interventions for depression: Opportunities and challenges. *Depression and Anxiety*, *34*(7), 596–602. https://doi.org/10.1002/da.22641

Cuijpers, P., Reijnders, M., & Huibers, M. J. H. (2019). The Role of Common Factors in Psychotherapy Outcomes. *Annual Review of Clinical Psychology*, *15*(1), Advance Online Publication. https://doi.org/10.1146/annurev-clinpsy-050718-095424

Cuijpers, P., van Lier, P. A. C., van Straten, A., & Donker, M. (2005). Examining differential effects of psychological treatment of depressive disorder: An application of trajectory analyses. *Journal of Affective Disorders*, *89*(1–3), 137–146. https://doi.org/10.1016/j.jad.2005.09.001

de Jonge, P., Wardenaar, K. J., Lim, C. C. W., Aguilar-Gaxiola, S., Alonso, J., Andrade, L. H., … Scott, K. (2018). The cross-national structure of mental disorders: results from the World Mental Health Surveys. *Psychological Medicine*, *48*(12), 2073–2084. https://doi.org/10.1017/S0033291717003610

DGPPN (Ed.). (2017). *S3-Leitlinie/Nationale VersorgungsLeitlinie Unipolare Depression Kurzfassung* (2nd ed.). Berlin: Editor.

Diallo, T. M. O., Morin, A. J. S., & Lu, H. (2016). Impact of Misspecifications of the Latent Variance–Covariance and Residual Matrices on the Class Enumeration Accuracy of Growth Mixture Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(4), 507–531. https://doi.org/10.1080/10705511.2016.1169188

Driessen, E., Cuijpers, P., de Maat, S. C. M., Abbass, A. A., de Jonghe, F., & Dekker, J. J. M. (2010). The efficacy of short-term psychodynamic psychotherapy for depression: a meta-analysis. *Clinical Psychology Review*, *30*(1), 25–36. https://doi.org/10.1016/j.cpr.2009.08.010

Drozd, F., Vaskinn, L., Bergsund, H. B., Haga, S. M., Slinning, K., & Bjørkli, C. A. (2016). The Implementation of Internet Interventions for Depression: A Scoping Review. *Journal of Medical Internet Research*, *18*(9), e236. https://doi.org/10.2196/jmir.5670

Ekberg, S., Barnes, R. K., Kessler, D. S., Mirza, S., Montgomery, A. A., Malpass, A., & Shaw, A. R. G. (2015). Relationship between Expectation Management and Client Retention in Online Cognitive Behavioural Therapy. *Behavioural and Cognitive Psychotherapy*, *43*(6), 732–743. https://doi.org/10.1017/S1352465814000241

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. https://doi.org/10.1371/journal.pone.0124258

El Alaoui, S., Ljotsson, B., Hedman, E., Svanborg, C., Kaldo, V., & Lindefors, N. (2016). Predicting Outcome in Internet-Based Cognitive Behaviour Therapy for Major Depression: A Large

Cohort Study of Adult Patients in Routine Psychiatric Care. *Plos One*, *11*(9), e0161191. https://doi.org/10.1371/journal.pone.0161191

Elkin, I., Parloff, M. B., Hadley, S. W., & Autry, J. H. (1985). NIMH Treatment of Depression Collaborative Research Program: Background and Research Plan. *Archives of General Psychiatry*, *42*(3), 305–316. https://doi.org/10.1001/archpsyc.1985.01790260103013

Ernst & Young GmbH. (2017). *Online-Nutzung in Deutschland. Ergebnisse einer Befragung von 1.400 Verbrauchern.* Eschborn: Editor. Retrieved from https://www.ey.com/Publication/vwLUAssets/ey-online-nutzung-in-deutschland-juni-2017/$FILE/ey-online-nutzung-in-deutschland-juni-2017.pdf

Falkenström, F., Hatcher, R. L., & Holmqvist, R. (2015). Confirmatory Factor Analysis of the Patient Version of the Working Alliance Inventory--Short Form Revised. *Assessment*, *22*(5), 581–593. https://doi.org/10.1177/1073191114552472

Falkenström, F., Hatcher, R. L., Skjulsvik, T., Larsson, M. H., & Holmqvist, R. (2015). Development and validation of a 6-item working alliance questionnaire for repeated administrations during psychotherapy. *Psychological Assessment*, *27*(1), 169–183. https://doi.org/10.1037/pas0000038

Feeney, B. C., & Collins, N. L. (2015). A New Look at Social Support: A Theoretical Perspective on Thriving Through Relationships. *Personality and Social Psychology Review*, *19*(2), 113–147. https://doi.org/10.1177/1088868314544222

Ferrari, A. J., Charlson, F. J., Norman, R. E., Patten, S. B., Freedman, G., Murray, C. J. L., … Whiteford, H. A. (2013). Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010. *PLoS Med*, *10*(11), e1001547. https://doi.org/10.1371/journal.pmed.1001547

Fine, T. H., Contractor, A. A., Tamburrino, M., Elhai, J. D., Prescott, M. R., Cohen, G. H., … Calabrese, J. R. (2013). Validation of the telephone-administered PHQ-9 against the in-person administered SCID-I major depression module. *Journal of Affective Disorders*, *150*(3), 1001–1007. https://doi.org/10.1016/j.jad.2013.05.029

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, *55*(4), 316–340. https://doi.org/10.1037/pst0000172

Foster, S. L., Petrie, S. R., Mischoulon, D., & Fava, M. (2019). Personalized Medicine. In B. G. Shapero, D. Mischoulon, & C. Cusin (Eds.), *The Massachusetts General Hospital Guide to Depression: New Treatment Insights and Options* (pp. 109–121). Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-97241-1_8

Friborg, O., Martinsen, E. W., Martinussen, M., Kaiser, S., Øvergård, K. T., & Rosenvinge, J. H. (2014). Comorbidity of personality disorders in mood disorders: A meta-analytic review of 122 studies from 1988 to 2010. *Journal of Affective Disorders*, *152–154*, 1–11. https://doi.org/10.1016/j.jad.2013.08.023

Furmark, T., Carlbring, P., Hedman, E., Sonnenstein, A., Clevberger, P., Bohman, B., … Andersson, G. (2009). Guided and unguided self-help for social anxiety disorder: randomised controlled trial. *British Journal of Psychiatry*, *195*(5), 440–447. https://doi.org/10.1192/bjp.bp.108.060996

Gilbody, S. M., Littlewood, E., Hewitt, C., Brierley, G., Tharmanathan, P., Araya, R., … White, D. (2015). Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): large scale pragmatic randomised controlled trial. *The BMJ*, *351*. https://doi.org/10.1136/bmj.h5627

Gilbody, S. M., Whitty, P. M., Grimshaw, J. M., & Thomas, R. E. (2003). Improving the detection and management of depression in primary care. *BMJ Quality & Safety*, *12*(2), 149–155. https://doi.org/10.1136/qhc.12.2.149

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

Gueorguieva, R., Mallinckrodt, C., & Krystal, J. H. (2011). Trajectories of Depression Severity in Clinical Trials of Duloxetine Insights Into Antidepressant and Placebo Responses. *Archives of General Psychiatry*, *68*(12), 1227–1237. https://doi.org/10.1001/archgenpsychiatry.2011.132

Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. *Bmc Psychiatry*, *10*, 113. https://doi.org/10.1186/1471-244X-10-113

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. https://doi.org/10.1016/j.janxdis.2016.05.006

Hajebi, A., Motevalian, A., Amin-Esmaeili, M., Hefazi, M., Radgoodarzi, R., Rahimi-Movaghar, A., & Sharifi, V. (2012). Telephone versus face-to-face administration of the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, for diagnosis of psychotic disorders. *Comprehensive Psychiatry*, *53*(5), 579–583. https://doi.org/10.1016/j.comppsych.2011.06.001

Hallgren, M., Kraepelien, M., Öjehagen, A., Lindefors, N., Zeebari, Z., Kaldo, V., & Forsell, Y. (2015). Physical exercise and internet-based cognitive-behavioural therapy in the treatment of depression: randomised controlled trial. *The British Journal of Psychiatry: The Journal of Mental Science*, *207*(3), 227–234. https://doi.org/10.1192/bjp.bp.114.160101

Hansen, A. H., & Kristoffersen, A. E. (2016). The use of CAM providers and psychiatric outpatient services in people with anxiety/depression: a cross-sectional survey. *BMC Complementary and Alternative Medicine*, *16*. https://doi.org/10.1186/s12906-016-1446-9

Hantsoo, L., & Epperson, C. N. (2015). Premenstrual Dysphoric Disorder: Epidemiology and Treatment. *Current Psychiatry Reports*, *17*(11), 87. https://doi.org/10.1007/s11920-015-0628-3

Harvey, A. G., & Gumport, N. B. (2015). Evidence-based psychological treatments for mental disorders: Modifiable barriers to access and possible solutions. *Behaviour Research and Therapy*, *68*, 1–12. https://doi.org/10.1016/j.brat.2015.02.004

Hatcher, R. L., & Gillaspy, J. A. (2006). Development and validation of a revised short version of the working alliance inventory. *Psychotherapy Research*, *16*(1), 12–25. https://doi.org/10.1080/10503300500352500

Hedman, E., Andersson, E., Ljótsson, B., Andersson, G., Andersson, E., Schalling, M., … Rück, C. (2012). Clinical and genetic outcome determinants of Internet- and group-based cognitive behavior therapy for social anxiety disorder. *Acta Psychiatrica Scandinavica*, *126*(2), 126–136. https://doi.org/10.1111/j.1600-0447.2012.01834.x

Heinrich, M., Zagorscak, P., Eid, M., & Knaevelsrud, C. (2018). Giving G a Meaning: An Application of the Bifactor-(S-1) Approach to Realize a More Symptom-Oriented Modeling of the Beck Depression Inventory-II. *Assessment*, 1073191118803738. https://doi.org/10.1177/1073191118803738

Hines, D. A., Douglas, E. M., & Mahmood, S. (2010). The effects of survey administration on disclosure rates to sensitive items among men: A comparison of an internet panel sample with a RDD telephone sample. *Computers in Human Behavior*, *26*(6), 1327–1335. https://doi.org/10.1016/j.chb.2010.04.006

Horne, J. A., & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology*, *4*, 97–110.

Hunter, A. M., Muthén, B. O., Cook, I. A., & Leuchter, A. F. (2010). Antidepressant response trajectories and quantitative electroencephalography (QEEG) biomarkers in major depressive disorder. *Journal of Psychiatric Research*, *44*(2), 90–98. https://doi.org/10.1016/j.jpsychires.2009.06.006

Ivanova, E., Lindner, P., Ly, K. H., Dahlin, M., Vernmark, K., Andersson, G., & Carlbring, P. (2016). Guided and unguided Acceptance and Commitment Therapy for social anxiety disorder and/or panic disorder provided via the Internet and a smartphone application: A randomized controlled trial. *Journal of Anxiety Disorders*, *44*, 27–35. https://doi.org/10.1016/j.janxdis.2016.09.012

Jacobi, F., Höfler, M., Siegert, J., Mack, S., Gerschler, A., Scholl, L., … Wittchen, H.-U. (2014). Twelve-month prevalence, comorbidity and correlates of mental disorders in Germany: the Mental Health Module of the German Health Interview and Examination Survey for Adults (DEGS1-MH). *International Journal of Methods in Psychiatric Research*, *23*(3), 304–319. https://doi.org/10.1002/mpr.1439

Jacobi, F., Uhmann, S., & Hoyer, J. (2011). How frequent is negative outcome in outpatient psychotherapy? Results from a CBT outpatient clinic. *Zeitschrift Für Klinische Psychologie Und Psychotherapie*, *40*(4), 246–256. https://doi.org/10.1026/1616-3443/a000122

Jankowski, K. S. (2016). Morningness-eveningness and depressive symptoms: Test on the components level with CES-D in Polish students. *Journal of Affective Disorders*, *196*, 47–53. https://doi.org/10.1016/j.jad.2016.02.015

Johansson, R., & Andersson, G. (2012). Internet-based psychological treatments for depression. *Expert Review of Neurotherapeutics*, *12*(7), 861–870. https://doi.org/10.1586/ern.12.63

Johansson, R., Ekbladh, S., Hebert, A., Lindstrom, M., Moller, S., Petitt, E., … Andersson, G. (2012). Psychodynamic Guided Self-Help for Adult Depression through the Internet: A Randomised Controlled Trial. *Plos One*, *7*(5), e38021. https://doi.org/10.1371/journal.pone.0038021

Johansson, R., Sjoberg, E., Sjogren, 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. https://doi.org/10.1371/journal.pone.0036905

Jung, T., & Wickrama, K. a. S. (2008). An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling. *Social and Personality Psychology Compass*, *2*(1), 302–317. https://doi.org/10.1111/j.1751-9004.2007.00054.x

Kamenov, K., Twomey, C., Cabello, M., Prina, A. M., & Ayuso-Mateos, J. L. (2017). The efficacy of psychotherapy, pharmacotherapy and their combination on functioning and quality of life in depression: a meta-analysis. *Psychological Medicine*, *47*(3), 414–425. https://doi.org/10.1017/S0033291716002774

Karyotaki, E., Ebert, D. D., Donkin, L., Riper, H., Twisk, J., Burger, S., … Cuijpers, P. (2018). Do guided internet-based interventions result in clinically relevant changes for patients with depression? An individual participant data meta-analysis. *Clinical Psychology Review*, *63*, 80–92. https://doi.org/10.1016/j.cpr.2018.06.007

Karyotaki, E., Riper, H., Twisk, J., Hoogendoorn, A., Kleiboer, A., Mira, A., … Cuijpers, P. (2017). Efficacy of Self-guided Internet-Based Cognitive Behavioral Therapy in the Treatment of Depressive Symptoms. A Meta-analysis of Individual Participant Data. *Jama Psychiatry*, *74*(4), 351–359. https://doi.org/10.1001/jamapsychiatry.2017.0044

Kass, A. E., Trockel, M., Safer, D. L., Sinton, M. M., Cunning, D., Rizk, M. T., … Taylor, C. B. (2014). Internet-based preventive intervention for reducing eating disorder risk: A randomized controlled trial comparing guided with unguided self-help. *Behaviour Research and Therapy*, *63*, 90–98. https://doi.org/10.1016/j.brat.2014.09.010

Kazdin, A. E. (2005). Treatment outcomes, common factors, and continued neglect of mechanisms of change. *Clinical Psychology-Science and Practice*, *12*(2), 184–188. https://doi.org/10.1093/clipsy/bpi023

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. In *Annual Review of Clinical Psychology* (Vol. 3, pp. 1–27). Palo Alto: Annual Reviews.

Kelders, S. M., Bohlmeijer, E. T., Pots, W. T. M., & van Gemert-Pijnen, J. E. W. C. (2015). Comparing human and automated support for depression: Fractional factorial randomized controlled trial. *Behaviour Research and Therapy*, *72*, 72–80. https://doi.org/10.1016/j.brat.2015.06.014

Keller, F., & Hautzinger, M. (2007). Klassifikation von Verlaufskurven in der Depressionsbehandlung. *Zeitschrift Für Klinische Psychologie Und Psychotherapie*, *36*(2), 83–92. https://doi.org/10.1026/1616-3443.36.2.83

Kenter, R. M. F., Cuijpers, P., Beekman, A., & van Straten, A. (2016). Effectiveness of a Web-Based Guided Self-help Intervention for Outpatients With a Depressive Disorder: Short-term Results From a Randomized Controlled Trial. *Journal of Medical Internet Research*, *18*(3). https://doi.org/10.2196/jmir.4861

Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S.-L. T., … Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine*, *32*(6), 959–976. https://doi.org/10.1017/S0033291702006074

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. https://doi.org/10.1001/archpsyc.62.6.593

Kessler, R. C., & Bromet, E. J. (2013). The Epidemiology of Depression Across Cultures. *Annual Review of Public Health*, *34*(1), 119–138. https://doi.org/10.1146/annurev-publhealth-031912-114409

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H.-U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*, *21*(3), 169–184. https://doi.org/10.1002/mpr.1359

Khan, A., Faucett, J., Lichtenberg, P., Kirsch, I., & Brown, W. A. (2012). A Systematic Review of Comparative Efficacy of Treatments and Controls for Depression. *PLoS ONE*, *7*(7). https://doi.org/10.1371/journal.pone.0041778

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. https://doi.org/10.1080/16506073.2014.921834

Klein, J. P., Berger, T., Schroeder, J., Spaeth, 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. https://doi.org/10.1159/000445355

Klein, J. P., Knaevelsrud, C., Bohus, M., Ebert, D. D., Gerlinger, G., Günther, K., … Hauth, I. (2018). Internetbasierte Selbstmanagementinterventionen: Qualitätskriterien für ihren Einsatz in Prävention und Behandlung psychischer Störungen. *Der Nervenarzt*, *89*(11), 1277–1286. https://doi.org/10.1007/s00115-018-0591-4

Koch, W., & Frees, B. (2017). ARD/ZDF-Onlinestudie 2017: Neun von zehn Deutschen online. *Media Perspektiven*, *9*. Retrieved from http://www.ard-zdf-onlinestudie.de/files/2017/Artikel/917_Koch_Frees.pdf

Koffmann, A. (2018). Has growth mixture modeling improved our understanding of how early change predicts psychotherapy outcome? *Psychotherapy Research*, *28*(6), 829–841. https://doi.org/10.1080/10503307.2017.1294771

Königbauer, J., Letsch, J., Doebler, P., Ebert, D., & Baumeister, H. (2017). Internet- and mobile-based depression interventions for people with diagnosed depression: A systematic review and meta-analysis. *Journal of Affective Disorders*, *223*, 28–40. https://doi.org/10.1016/j.jad.2017.07.021

Koole, S. L., & Tschacher, W. (2016). Synchrony in Psychotherapy: A Review and an Integrative Framework for the Therapeutic Alliance. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00862

Küster, A., Niemeyer, H., & Knaevelsrud, C. (2016). Internet-based interventions for posttraumatic stress: A meta-analysis of randomized controlled trials. *Clinical Psychology Review*, *43*, 1–16. https://doi.org/10.1016/j.cpr.2015.11.004

Kutcher, S., Bagnell, A., & Wei, Y. (2015). Mental health literacy in secondary schools: a Canadian approach. *Child and Adolescent Psychiatric Clinics of North America*, *24*(2), 233–244. https://doi.org/10.1016/j.chc.2014.11.007

Lambert, M. J. (1992). Psychotherapy outcome research: Implications for integrative and eclectical therapists. In Norcross, J. C. & Goldfried, M. R. (Eds.), *Handbook of psychotherapy integration* (pp. 94–129). New York: Basic Books.

Lappalainen, P., Langrial, S., Oinas-Kukkonen, H., Tolvanen, A., & Lappalainen, R. (2015). Web-Based Acceptance and Commitment Therapy for Depressive Symptoms With Minimal Support: A Randomized Controlled Trial. *Behavior Modification*, *39*(6), 805–834. https://doi.org/10.1177/0145445515598142

Laursen, T. M., Musliner, K. L., Benros, M. E., Vestergaard, M., & Munk-Olsen, T. (2016). Mortality and life expectancy in persons with severe unipolar depression. *Journal of Affective Disorders*, *193*, 203–207. https://doi.org/10.1016/j.jad.2015.12.067

Lewis, V. M., Williams, K., KoKo, C., Woolmore, J., Jones, C., & Powell, T. (2017). Disability, depression and suicide ideation in people with multiple sclerosis. *Journal of Affective Disorders*, *208*, 662–669. https://doi.org/10.1016/j.jad.2016.08.038

Lindfors, O., Ojanen, S., Jääskeläinen, T., & Knekt, P. (2014). Social support as a predictor of the outcome of depressive and anxiety disorder in short-term and long-term psychotherapy. *Psychiatry Research*, *216*(1), 44–51. https://doi.org/10.1016/j.psychres.2013.12.050

Ljótsson, B., Hesser, H., Andersson, E., Lackner, J. M., El Alaoui, S., Falk, L., … Hedman, E. (2014). Provoking symptoms to relieve symptoms: A randomized controlled dismantling study of exposure therapy in irritable bowel syndrome. *Behaviour Research and Therapy*, *55*(Supplement C), 27–39. https://doi.org/10.1016/j.brat.2014.01.007

Lucassen, M. F., Stasiak, K., Samra, R., Frampton, C. M., & Merry, S. N. (2017). Sexual minority youth and depressive symptoms or depressive disorder: A systematic review and meta-analysis of population-based studies. *Australian & New Zealand Journal of Psychiatry*, *51*(8), 774–787. https://doi.org/10.1177/0004867417713664

Luhmann, M., Hofmann, W., Eid, M., & Lucas, R. E. (2012). Subjective well-being and adaptation to life events: a meta-analysis. *Journal of Personality and Social Psychology*, *102*(3), 592–615. https://doi.org/10.1037/a0025948

Lutz, W., Arndt, A., Rubel, J., Berger, T., Schröder, J., Späth, C., … Moritz, S. (2017). Defining and Predicting Patterns of Early Response in a Web-Based Intervention for Depression. *Journal of Medical Internet Research*, *19*(6). https://doi.org/10.2196/jmir.7367

Lutz, W., Stulz, N., & Köck, K. (2009). Patterns of early change and their relationship to outcome and follow-up among patients with major depressive disorders. *Journal of Affective Disorders*, *118*(1), 60–68. https://doi.org/10.1016/j.jad.2009.01.019

Manen, J. G. V., Horn, E. K., Stijnen, T., Timman, R., Busschbach, J. J. V., & Verheul, R. (2015). Tailoring psychotherapy in patients with personality disorders: Matching the level of psychological strengths to the level of stabilizing versus destabilizing psychotherapy. *Personality and Mental Health*, *9*(2), 133–149. https://doi.org/10.1002/pmh.1281

Marcus, D. K., O'Connell, D., Norris, A. L., & Sawaqdeh, A. (2014). Is the Dodo bird endangered in the 21st century? A meta-analysis of treatment comparison studies. *Clinical Psychology Review*, *34*(7), 519–530. https://doi.org/10.1016/j.cpr.2014.08.001

Maybery, D. J., Neale, J., Arentz, A., & Jones-Ellis, J. (2007). The Negative Event Scale: Measuring frequency and intensity of adult hassles. *Anxiety, Stress, & Coping*, *20*(2), 163–176. https://doi.org/10.1080/10615800701217654

McIntosh, E., Gillanders, D., & Rodgers, S. (2010). Rumination, goal linking, daily hassles and life events in major depression. *Clinical Psychology & Psychotherapy*, *17*(1), 33–43. https://doi.org/10.1002/cpp.611

Melchior, H., Schulz, H., Kriston, L., Hergert, A., Hofreuter-Gätgens, K., Bergelt, C., … Watzke, B. (2016). Symptom change trajectories during inpatient psychotherapy in routine care and their associations with long-term outcomes. *Psychiatry Research*, *238*, 228–235. https://doi.org/10.1016/j.psychres.2016.02.046

Meyer, B., Bierbrodt, J., Schröder, J., Berger, T., Beevers, C. G., Weiss, M., … Klein, J. P. (2015). Effects of an Internet intervention (Deprexis) on severe depression symptoms: Randomized controlled trial. *Internet Interventions*, *2*(1), 48–59. https://doi.org/10.1016/j.invent.2014.12.003

Miller, S. D., Duncan, B. L., & Hubble, M. A. (1996). *Escape from Babel: Toward a Unifying Language for Psychotherapy Practice* (1st edition). New York: W. W. Norton & Company.

Miranda, R., Soffer, A., Polanco-Roman, L., Wheeler, A., & Moore, A. (2015). Mental Health Treatment Barriers Among Racial/Ethnic Minority Versus White Young Adults 6 Months After Intake at a College Counseling Center. *Journal of American College Health*, *63*(5), 291–298. https://doi.org/10.1080/07448481.2015.1015024

Moore, R. C., Depp, C. A., Wetherell, J. L., & Lenze, E. J. (2016). Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. *Journal of Psychiatric Research*, *75*, 116–123. https://doi.org/10.1016/j.jpsychires.2016.01.011

Mulder, R., Murray, G., & Rucklidge, J. (2017). Common versus specific factors in psychotherapy: opening the black box. *The Lancet Psychiatry*, *4*(12), 953–962. https://doi.org/10.1016/S2215-0366(17)30100-1

Müller, M. J., Kundermann, B., & Cabanel, N. (2016). Eveningness and poor sleep quality independently contribute to self-reported depression severity in psychiatric inpatients with affective disorder. *Nordic Journal of Psychiatry*, *70*(5), 329–334. https://doi.org/10.3109/08039488.2015.1112832

Munder, T., Wilmers, F., Leonhart, R., Linster, H. W., & Barth, J. (2010). Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. *Clinical Psychology & Psychotherapy*, *17*(3), 231–239. https://doi.org/10.1002/cpp.658

Musliner, K. L., Munk-Olsen, T., Eaton, W. W., & Zandi, P. P. (2016). Heterogeneity in long-term trajectories of depressive symptoms: Patterns, predictors and outcomes. *Journal of Affective Disorders*, *192*, 199–211. https://doi.org/10.1016/j.jad.2015.12.030

Muthén, B., Brown, H., Leuchter, A., & Hunter, A. (2010). General approaches to analysis of course: applying growth mixture modeling to randomized trials of depression. In P. Shrout, K. Keyes, & K. Ornstein (Eds.), *Causality and Psychopathology: Finding the Determinants of Disorders and their Cures* (1st edition, pp. 159–178). Oxford: Oxford University Press.

Naslund, J. A., Grande, S. W., Aschbrenner, K. A., & Elwyn, G. (2014). Naturally Occurring Peer Support through Social Media: The Experiences of Individuals with Severe Mental Illness Using YouTube. *PLOS ONE*, *9*(10), e110171. https://doi.org/10.1371/journal.pone.0110171

Newby, J. M., Mackenzie, A., Williams, A. D., McIntyre, K., Watts, S., Wong, N., & Andrews, G. (2013). Internet cognitive behavioural therapy for mixed anxiety and depression: a randomized controlled trial and evidence of effectiveness in primary care. *Psychological Medicine*, *43*(12), 2635–2648. https://doi.org/10.1017/S0033291713000111

Newby, J. M., Mewton, L., & Andrews, G. (2017). Transdiagnostic versus disorder-specific internet-delivered cognitive behaviour therapy for anxiety and depression in primary care. *Journal of Anxiety Disorders*, *46*, 25–34. https://doi.org/10.1016/j.janxdis.2016.06.002

Nienhuis, J. B., Owen, J., Valentine, J. C., Black, S. W., Halford, T. C., Parazak, S. E., … Hilsenroth, M. (2018). Therapeutic alliance, empathy, and genuineness in individual adult psychotherapy: A meta-analytic review. *Psychotherapy Research*, *28*(4), 593–605. https://doi.org/10.1080/10503307.2016.1204023

Nobis, S., Lehr, D., Ebert, D. D., Baumeister, H., Snoek, F., Riper, H., & Berking, M. (2015). Efficacy of a web-based intervention with mobile phone support in treating depressive symptoms in adults with type 1 and type 2 diabetes: a randomized controlled trial. *Diabetes Care*, *38*(5), 776–783. https://doi.org/10.2337/dc14-1728

Nordgreen, T., Havik, O. E., Ost, L. G., Furmark, T., Carlbring, P., & Andersson, G. (2012). Outcome predictors in guided and unguided self-help for social anxiety disorder. *Behaviour Research and Therapy*, *50*(1), 13–21. https://doi.org/10.1016/j.brat.2011.10.009

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(4), 535–569. https://doi.org/10.1080/10705510701575396

Orlinsky, D. E., & Howard, K. I. (1987). A generic model of psychotherapy. *Journal of Integrative & Eclectic Psychotherapy*, *6*(1), 6–27.

Patterson, C. L., Anderson, T., & Wei, C. (2014). Clients' Pretreatment Role Expectations, the Therapeutic Alliance, and Clinical Outcomes in Outpatient Therapy. *Journal of Clinical Psychology*, *70*(7), 673–680. https://doi.org/10.1002/jclp.22054

Paxling, B., Lundgren, S., Norman, A., Almlöv, J., Carlbring, P., Cuijpers, P., & Andersson, G. (2013). Therapist Behaviours in Internet-Delivered Cognitive Behaviour Therapy: Analyses of E-Mail Correspondence in the Treatment of Generalized Anxiety Disorder. *Behavioural and Cognitive Psychotherapy*, *41*(3), 280–289. https://doi.org/10.1017/S1352465812000240

Pennant, M. E., Loucas, C. E., Whittington, C., Creswell, C., Fonagy, P., Fuggle, P., … Kendall, T. (2015). Computerised therapies for anxiety and depression in children and young people: A systematic review and meta-analysis. *Behaviour Research and Therapy*, *67*, 1–18. https://doi.org/10.1016/j.brat.2015.01.009

Peugh, J., & Fan, X. (2012). How Well Does Growth Mixture Modeling Identify Heterogeneous Growth Trajectories? A Simulation Study Examining GMM's Performance Characteristics. *Structural*

Equation Modeling-a Multidisciplinary Journal - STRUCT EQU MODELING, *19*, 204–226. https://doi.org/10.1080/10705511.2012.659618

Pihlaja, S., Stenberg, J.-H., Joutsenniemi, K., Mehik, H., Ritola, V., & Joffe, G. (2018). Therapeutic alliance in guided internet therapy programs for depression and anxiety disorders – A systematic review. *Internet Interventions*, *11*, 1–10. https://doi.org/10.1016/j.invent.2017.11.005

Pöhlmann, K., Altmann, U., Butter, A., Israel, M., Keller, A., Schellong, J., … Joraschky, P. (2008). Differenzielle Therapieverläufe in der stationären Psychotherapie: Eine störungsübergreifende clusteranalytische Differenzierung von Verlaufstypen auf der Basis des Merkmals Depressivität. In W. Kirch, B. Badura, & H. Pfaff (Eds.), *Prävention und Versorgungsforschung* (pp. 195–216). Berlin, Heidelberg: Springer.

Ram, N., & Grimm, K. J. (2009). Methods and Measures: Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development*, *33*(6), 565–576. https://doi.org/10.1177/0165025409343765

Richards, D., & Richardson, T. (2012). Computer-based psychological treatments for depression: A systematic review and meta-analysis. *Clinical Psychology Review*, *32*(4), 329–342. https://doi.org/10.1016/j.cpr.2012.02.004

Riper, H., Blankers, M., Hadiwijaya, H., Cunningham, J., Clarke, S., Wiers, R., … Cuijpers, P. (2014). Effectiveness of Guided and Unguided Low-Intensity Internet Interventions for Adult Alcohol Misuse: A Meta-Analysis. *Plos One*, *9*(6), e99912. https://doi.org/10.1371/journal.pone.0099912

Romero-Sanchiz, P., Nogueira-Arjona, R., Garcia-Ruiz, A., Luciano, J. V., Garcia Campayo, J., Gili, M., … Mayoral Cleries, F. (2017). Economic evaluation of a guided and unguided internet-based CBT intervention for major depression: Results from a multi-center, three-armed randomized controlled trial conducted in primary care. *Plos One*, *12*(2), e0172741. https://doi.org/10.1371/journal.pone.0172741

Rosenbach, C., & Renneberg, B. (2015). Remembering rejection: specificity and linguistic styles of autobiographical memories in borderline personality disorder and depression. *Journal of Behavior Therapy and Experimental Psychiatry*, *46*, 85–92. https://doi.org/10.1016/j.jbtep.2014.09.002

Rosenzweig, S. (1936). Some Implicit Common Factors in Diverse Methods of Psychotherapy. *American Journal of Orthopsychiatry*, *6*(3), 412–415. https://doi.org/10.1111/j.1939-0025.1936.tb05248.x

Salk, R. H., Hyde, J. S., & Abramson, L. Y. (2017). Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms. *Psychological Bulletin*, *143*(8), 783–822. https://doi.org/10.1037/bul0000102

Samoocha, D., Bruinvels, D. J., Elbers, N. A., Anema, J. R., & van der Beek, A. J. (2010). Effectiveness of Web-based Interventions on Patient Empowerment: A Systematic Review and Meta-analysis. *Journal of Medical Internet Research*, *12*(2). https://doi.org/10.2196/jmir.1286

Schlagert, H. S., & Hiller, W. (2017). The predictive value of early response in patients with depressive disorders. *Psychotherapy Research*, *27*(4), 488–500. https://doi.org/10.1080/10503307.2015.1119329

Schouler-Ocak, M., Aichberger, M. C., Penka, S., Kluge, U., & Heinz, A. (2015). Psychische Störungen bei Menschen mit Migrationshintergrund in Deutschland. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, *58*(6), 527–532. https://doi.org/10.1007/s00103-015-2143-4

Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review*, *17*(8), 881–901. https://doi.org/10.1016/S0272-7358(97)00056-1

Sheeber, L. B., Seeley, J. R., Feil, E. G., Davis, B., Sorensen, E., Kosty, D. B., & Lewinsohn, P. M. (2012). Development and Pilot Evaluation of an Internet-Facilitated Cognitive-Behavioral Intervention for Maternal Depression. *Journal of Consulting and Clinical Psychology*, *80*(5), 739–749. https://doi.org/10.1037/a0028820

Sieverink, F., Kelders, S., Poel, M., & Gemert-Pijnen, L. van. (2017). Opening the Black Box of Electronic Health: Collecting, Analyzing, and Interpreting Log Data. *JMIR Research Protocols*, *6*(8), e156. https://doi.org/10.2196/resprot.6452

Silberschatz, G. (2017). Improving the yield of psychotherapy research. *Psychotherapy Research*, *27*(1), 1–13. https://doi.org/10.1080/10503307.2015.1076202

Smith, E. C., & Grawe, K. (2003). What makes psychotherapy sessions productive? A new approach to bridging the gap between process research and practice. *Clinical Psychology & Psychotherapy*, *10*(5), 275–285. https://doi.org/10.1002/cpp.377

So, M., Yamaguchi, S., Hashimoto, S., Sado, M., Furukawa, T. A., & McCrone, P. (2013). Is computerised CBT really helpful for adult depression?-A meta-analytic re-evaluation of CCBT for adult depression in terms of clinical implementation and methodological validity. *Bmc Psychiatry*, *13*, 113. https://doi.org/10.1186/1471-244X-13-113

Solomon, D., & Adams, J. (2015). The use of complementary and alternative medicine in adults with depressive disorders. A critical integrative review. *Journal of Affective Disorders*, *179*, 101–113. https://doi.org/10.1016/j.jad.2015.03.031

Späth, C., Hapke, U., Maske, U., Schröder, J., Moritz, S., Berger, T., … Klein, J. P. (2017). Characteristics of participants in a randomized trial of an Internet intervention for depression (EVIDENT) in comparison to a national sample (DEGS1). *Internet Interventions*, *9*, 46–50. https://doi.org/10.1016/j.invent.2017.05.003

Spek, V., Cuijpers, P., Nykicek, I., Riper, H., Keyzer, J., & Pop, V. (2007). Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. *Psychological Medicine*, *37*(3), 319–328. https://doi.org/10.1017/S0033291706008944

Starr, L. R., & Hershenberg, R. (2017). Depressive Symptoms and the Anticipation and Experience of Uplifting Events in Everyday Life. *Journal of Clinical Psychology*, *73*(10), 1442–1461. https://doi.org/10.1002/jclp.22447

Steele, L. S., Daley, A., Curling, D., Gibson, M. F., Green, D. C., Williams, C. C., & Ross, L. E. (2016). LGBT Identity, Untreated Depression, and Unmet Need for Mental Health Services by Sexual Minority Women and Trans-Identified People. *Journal of Women's Health*, *26*(2), 116–127. https://doi.org/10.1089/jwh.2015.5677

Stenzel, N., & Berking, M. (2012). Wirkfaktoren in der Psychotherapie. In M. Berking & W. Rief (Eds.), *Klinische Psychologie und Psychotherapie für Bachelor: Band II: Therapieverfahren Lesen, Hören, Lernen im Web* (pp. 149–164). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-25523-6_14

Stulz, N., & Lutz, W. (2007). Multidimensional patterns of change in outpatient psychotherapy: the phase model revisited. *Journal of Clinical Psychology*, *63*(9), 817–833. https://doi.org/10.1002/jclp.20397

Sunderland, M., Wong, N., Hilvert-Bruce, Z., & Andrews, G. (2012). Investigating trajectories of change in psychological distress amongst patients with depression and generalised anxiety disorder treated with internet cognitive behavioural therapy. *Behaviour Research and Therapy*, *50*(6), 374–380. https://doi.org/10.1016/j.brat.2012.03.005

Sundstrom, C., Gajecki, M., Johansson, M., Blankers, M., Sinadinovic, K., Stenlund-Gens, E., & Berman, A. H. (2016). Guided and Unguided Internet-Based Treatment for Problematic Alcohol Use - A Randomized Controlled Pilot Trial. *Plos One*, *11*(7), e0157817. https://doi.org/10.1371/journal.pone.0157817

Szajewska, H. (2018). Evidence-Based Medicine and Clinical Research: Both Are Needed, Neither Is Perfect. *Annals of Nutrition & Metabolism*, *72 Suppl 3*, 13–23. https://doi.org/10.1159/000487375

Sztein, D. M., Koransky, C. E., Fegan, L., & Himelhoch, S. (2018). Efficacy of cognitive behavioural therapy delivered over the Internet for depressive symptoms: A systematic review and meta-analysis. *Journal of Telemedicine and Telecare*, *24*(8), 527–539. https://doi.org/10.1177/1357633X17717402

Thomas, M. L. (2006). The Contributing Factors of Change in a Therapeutic Process. *Contemporary Family Therapy*, *28*(2), 201–210. https://doi.org/10.1007/s10591-006-9000-4

Thornicroft, G., Chatterji, S., Evans-Lacko, S., Gruber, M., Sampson, N., Aguilar-Gaxiola, S., … Kessler, R. C. (2017). Undertreatment of people with major depressive disorder in 21 countries. *The British Journal of Psychiatry*, *210*(2), 119–124. https://doi.org/10.1192/bjp.bp.116.188078

Titov, N. (2011). Internet-delivered psychotherapy for depression in adults. *Current Opinion in Psychiatry*, *24*(1), 18–23. https://doi.org/10.1097/YCO.0b013e32833ed18f

Titov, N., Andrews, G., Davies, M., McIntyre, K., Robinson, E., & Solley, K. (2010). Internet Treatment for Depression: A Randomized Controlled Trial Comparing Clinician vs. Technician Assistance. *Plos One*, *5*(6), e10939. https://doi.org/10.1371/journal.pone.0010939

Titov, N., Dear, B. F., McMillan, D., Anderson, T., Zou, J., & Sunderland, M. (2011). Psychometric comparison of the PHQ-9 and BDI-II for measuring response during treatment of depression. *Cognitive Behaviour Therapy*, *40*(2), 126–136. https://doi.org/10.1080/16506073.2010.550059

Tofighi, D., & Enders, C. K. (2008). Identifying the correct number of classes in growth mixture models. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (Vol. 2007, pp. 317–341). Charlotte, North Carolina: Information Age Publishing.

Tolin, D. F. (2010). Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. *Clinical Psychology Review*, *30*(6), 710–720. https://doi.org/10.1016/j.cpr.2010.05.003

Tolin, D. F. (2014). Beating a Dead Dodo Bird: Looking at Signal vs. Noise in Cognitive-Behavioral Therapy for Anxiety Disorders. *Clinical Psychology: Science and Practice*, *21*(4), 351–362. https://doi.org/10.1111/cpsp.12080

Tolliver, B. K., & Anton, R. F. (2015). Assessment and treatment of mood disorders in the context of substance abuse. *Dialogues in Clinical Neuroscience*, *17*(2), 181–190.

Uher, R., Muthén, B., Souery, D., Mors, O., Jaracz, J., Placentino, A., … McGuffin, P. (2010). Trajectories of change in depression severity during treatment with antidepressants. *Psychological Medicine*, *40*(8), 1367–1377. https://doi.org/10.1017/S0033291709991528

Urban, E. J., Charles, S. T., Levine, L. J., & Almeida, D. M. (2018). Depression history and memory bias for specific daily emotions. *PLOS ONE*, *13*(9), e0203574. https://doi.org/10.1371/journal.pone.0203574

Vallejo, M. A., Jordán, C. M., Díaz, M. I., Comeche, M. I., & Ortega, J. (2007). Psychological Assessment via the Internet: A Reliability and Validity Study of Online (vs Paper-and-Pencil) Versions of the General Health Questionnaire-28 (GHQ-28) and the Symptoms Check-List-90-Revised (SCL-90-R). *Journal of Medical Internet Research*, *9*(1). https://doi.org/10.2196/jmir.9.1.e2

van Spijker, B. A., Werner-Seidler, A., Batterham, P. J., Mackinnon, A., Calear, A. L., Gosling, J. A., … Christensen, H. (2018). Effectiveness of a Web-Based Self-Help Program for Suicidal Thinking in an Australian Community Sample: Randomized Controlled Trial. *Journal of Medical Internet Research*, *20*(2), e15. https://doi.org/10.2196/jmir.8595

Vandeleur, C. L., Fassassi, S., Castelao, E., Glaus, J., Strippoli, M.-P. F., Lasserre, A. M., … Preisig, M. (2017). Prevalence and correlates of DSM-5 major depressive and related disorders in the community. *Psychiatry Research*, *250*, 50–58. https://doi.org/10.1016/j.psychres.2017.01.060

Vernmark, K., Lenndin, J., Bjarehed, J., Carlsson, M., Karlsson, J., Oberg, J., … Andersson, G. (2010). Internet administered guided self-help versus individualized e-mail therapy: A randomized trial of two versions of CBT for major depression. *Behaviour Research and Therapy*, *48*(5), 368–376. https://doi.org/10.1016/j.brat.2010.01.005

Vinkers, C. H., Joëls, M., Milaneschi, Y., Kahn, R. S., Penninx, B. W. J. H., & Boks, M. P. M. (2014). Stress Exposure Across the Life Span Cumulatively Increases Depression Risk and Is Moderated by Neuroticism. *Depression and Anxiety*, *31*(9), 737–745. https://doi.org/10.1002/da.22262

Vîslă, A., Flückiger, C., Constantino, M. J., Krieger, T., & Holtforth, M. G. (2018). Patient characteristics and the therapist as predictors of depressed patients' outcome expectation over time: A multilevel analysis. *Psychotherapy Research*, Advance Online Publication. https://doi.org/10.1080/10503307.2018.1428379

Wallin, E. E. K., Mattsson, S., & Olsson, E. M. G. (2016). The Preference for Internet-Based Psychological Interventions by Individuals Without Past or Current Use of Mental Health Treatment Delivered Online: A Survey Study With Mixed-Methods Analysis. *JMIR Mental Health*, *3*(2). https://doi.org/10.2196/mental.5324

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry*, *14*(3), 270–277. https://doi.org/10.1002/wps.20238

Wampold, B. E., Flückiger, C., Del Re, A. C., Yulish, N. E., Frost, N. D., Pace, B. T., … Hilsenroth, M. J. (2017). In pursuit of truth: A critical examination of meta-analyses of cognitive behavior

therapy. *Psychotherapy Research: Journal of the Society for Psychotherapy Research*, *27*(1), 14–32. https://doi.org/10.1080/10503307.2016.1249433

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. N. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "'all must have prizes.'" *Psychological Bulletin*, *122*(3), 203–215. https://doi.org/10.1037/0033-2909.122.3.203

Wentzel, J., van der Vaart, R., Bohlmeijer, E. T., & van Gemert-Pijnen, J. E. W. C. (2016). Mixing Online and Face-to-Face Therapy: How to Benefit From Blended Care in Mental Health Care. *JMIR Mental Health*, *3*(1). https://doi.org/10.2196/mental.4534

Wenze, S. J., Gunthert, K. C., & German, R. E. (2012). Biases in affective forecasting and recall in individuals with depression and anxiety symptoms. *Personality & Social Psychology Bulletin*, *38*(7), 895–906. https://doi.org/10.1177/0146167212447242

Wenze, S. J., & Miller, I. W. (2010). Use of ecological momentary assessment in mood disorders research. *Clinical Psychology Review*, *30*(6), 794–804. https://doi.org/10.1016/j.cpr.2010.06.007

WHO. (2018). ICD-11 - Coding Tools. Retrieved September 21, 2018, from https://icd.who.int/ct11_2018/icd11_mms/en/

Wijnen, B. F., Lokman, S., Leone, S., Evers, S. M., & Smit, F. (2018). Complaint-Directed Mini-Interventions for Depressive Symptoms: A Health Economic Evaluation of Unguided Web-Based Self-Help Interventions Based on a Randomized Controlled Trial. *Journal of Medical Internet Research*, *20*(10), e10455. https://doi.org/10.2196/10455

Wolitzky-Taylor, K. B., Arch, J. J., Rosenfield, D., & Craske, M. G. (2012). Moderators and non-specific predictors of treatment outcome for anxiety disorders: A comparison of cognitive behavioral therapy to acceptance and commitment therapy. *Journal of Consulting and Clinical Psychology*, *80*(5), 786–799. https://doi.org/10.1037/a0029418

Wong, K. P., Bonn, G., Tam, C. L., & Wong, C. P. (2018). Preferences for Online and/or Face-to-Face Counseling among University Students in Malaysia. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00064

World Health Organization (Ed.). (2017). *Depression and Other Common Mental Disorders: Global Health Estimates*. Geneva, Switzerland: World Health Organization. Retrieved from http://www.un.org/apps/news/story.asp?NewsID=56230#.WLRpQBB3xBw

Zagorscak, P., & Knaevelsrud, C. (2019). Online-Therapie. In S. Schneider & J. Margraf (Eds.), *Lehrbuch der Verhaltenstherapie, Band 3. Psychologische Therapie bei Indikationen im Kindes- und Jugendalter.* (2nd ed., pp. 233–247). Berlin, Heidelberg: Springer.

Zagorscak, P., Sommer, D., Haug, R., & Knaevelsrud, C. (2014). *Arbeiten mit dem TK-DepressionsCoach: Theorie- und Praxismanual für die beraterische Arbeit mit depressiven Online-Klientinnen. [Working with TK-DepressionsCoach: Theory and Practice of Counseling Depressed Clients Online]*. Unpublished Manual, Department of Clinical-Psychological Intervention, Freie Universität Berlin, Germany.

APPENDIX

The following materials supplement the doctoral thesis:

APPENDICES FOR STUDY 1

- *Online Table 1*. Change predicted by antidepressant medication, additional e-mail contacts and pre-interventional symptom load.
- *Online Table 2.* Estimated within-group changes for both primary outcome measures under various NMAR conditions.
- *Online Table 3*. Rates of reliable change, remission and recovery under various NMAR conditions.
- *Online Table 4*. Means and standard deviations for each outcome at each measurement occasion.

APPENDICES FOR STUDY 2

- *Appendix A*. Aid with interpreting change scores
- *Appendix B*. Supplementary results on model comparisons with different measurement invariance restriction (Table B1), results on general mean change (Table B2) and results on non-significant regression weights (Table B3).

APPENDICES FOR STUDY 3

- *Appendix C*. Detailed description of the statistical approach.

APPENDICES FOR CHAPTER 5

- *Appendix 5.1.* Scaffold for providing information about guidance in IBI
- *Appendix 5.2.* Excerpts from the treatment manual of the TK-DepressionsCoach

APPENDICES FOR STUDY 1

ONLINE TABLE 1

Change predicted by antidepressant medication, additional e-mail contacts and pre-interventional symptom load

Group	Predictor	Pre score		Pre- to post-treatment		Pre-treatment to 3-month follow-up		Pre-treatment to 6-month follow-up		Pre-treatment to 12-month follow-up	
		b (SE)	*p*	*b* (SE)	*p*	*b* (SE)	*p*	*b* (SE)	*p*	*b* (SE)	*p*
						Beck Depression Inventory-II					
CoD	AD	0.12 (0.38)	.743 NS	-0.28 (0.74)	.701 NS	0.67 (0.87)	.443 NS	0.98 (1.14)	.393 NS	2.26 (1.08)	.036 NS
	contacts	-	-	0.24 (0.19)	.198 NS	0.40 (0.24)	.091 NS	1.01 (0.35)	.004 NS	0.67 (0.33)	.039 NS
	pre-BDI	-	-	-0.64 (0.08)	< .001	-0.73 (0.10)	< .001	-0.75 (0.12)	< .001	-0.53 (0.12)	< .001
IC	AD	0.28 (0.37)	.440 NS	0.16 (0.74)	.829 NS	-1.20 (0.81)	.138 NS	-1.21 (0.84)	.151 NS	0.62 (1.02)	.548 NS
	contacts	-	-	0.22 (0.24)	.362 NS	-0.07 (0.27)	.782 NS	0.40 (0.25)	.109 NS	0.10 (0.33)	.754 NS
	pre-BDI	-	-	-0.62 (0.09)	< .001	-0.59 (0.10)	< .001	-0.71 (0.11)	<.001	-0.71 (0.12)	< .001
						Patient Health Questionnaire - 9					
CoD	AD	0.50 (0.33)	.132 NS	-0.31 (0.37)	.406 NS	-0.16 (0.47)	.732 NS	0.05 (0.55)	.926 NS	1.31 (0.58)	.023 NS
	contacts	-	-	0.14 (0.10)	.168 NS	0.29 (0.15)	.044 NS	0.60 (0.19)	.001	0.39 (0.15)	.009 NS
	pre-PHQ-9	-	-	-0.60 (0.05)	< .001	-0.66 (0.07)	< .001	-0.77 (0.07)	< .001	-0.70 (0.08)	< .001
IC	AD	0.44 (0.31)	.159 NS	0.05 (0.35)	.897 NS	-0.45 (0.40)	.258 NS	-0.64 (0.46)	.168 NS	0.12 (0.53)	.823 NS
	contacts	-	-	0.13 (0.13)	.314 NS	-0.05 (0.13)	.684 NS	0.22 (0.15)	.136 NS	0.02 (0.16)	.905 NS
	pre-PHQ-9	-	-	-0.72 (0.05)	< .001	-0.74 (0.06)	< .001	-0.80 (0.07)	< .001	-0.80 (0.07)	< .001

Note. $N = 1089$. Bonferroni-correction was applied to keep familywise errors rate at .05, therefore, only paths with $p < .001$ were considered as statistically significant. Unstandardized parameter estimates are reported. Both models including the predictor variables provided a good overall model fit, PHQ-9: average $\chi^2(df = 2) = 1.264$, $SD = 1.019$, average RMSEA = .005, $SD_{RMSEA} = .011$, average CFI = 1.000, $SD_{CFI} = .001$, average SRMR = .007, $SD_{SRMR} = .003$; BDI-II: average $\chi^2(df = 2) = 2.316$, $SD = 1.131$, average RMSEA = .017, $SD_{RMSEA} = .017$, average CFI = 1.000, $SD_{CFI} = .001$, average SRMR = .008, $SD_{SRMR} = .002$. CoD, Contact on Demand group; IC, Individual Counseling group; AD, antidepressant medication; Contacts, number of unplanned e-mail contacts during the intervention. NS = non-significant.

ONLINE TABLE 2

Estimated within-group changes for both primary outcome measures under various NMAR conditions

Outcome	Group	Pre- to post-treatment			Pre-treatment to 3-month follow-up			Pre-treatment to 6-month follow-up			Pre-treatment to 12-month follow-up		
		M_{LC} (SE_{LC})	*p*	*d* [95% CI]	M_{LC} (SE_{LC})	*p*	*d* [95% CI]	M_{LC} (SE_{LC})	*p*	*d* [95% CI]	M_{LC} (SE_{LC})	*p*	*d* [95% CI]
Beck Depression Inventory-II													
MAR	IC	-12.06 (0.36)	< .001	-1.53 [-1.69, -1.36]	-11.75 (0.42)	< .001	-1.40 [-1.58, -1.21]	-11.46 (0.43)	< .001	-1.35 [-1.52, -1.17]	-11.05 (0.49)	< .001	-1.17 [-1.35, -0.99]
	CoD	-10.81 (0.37)	< .001	-1.37 [-1.51, -1.22]	-11.38 (0.44)	< .001	-1.35 [-1.51, -1.19]	-11.13 (0.51)	< .001	-1.16 [-1.34, -0.98]	-11.95 (0.49)	< .001	-1.35 [-1.56, -1.15]
+ 3 BDI-II	IC	-11.69 (0.36)	< .001	-1.47 [-1.63, -1.31]	-11.31 (0.42)	< .001	-1.33 [-1.51, -1.15]	-11.03 (0.43)	< .001	-1.29 [-1.46, -1.12]	-10.58 (0.5)	< .001	-1.12 [-1.29, -0.94]
	CoD	-10.27 (0.37)	< .001	-1.3 [-1.44, -1.15]	-10.75 (0.44)	< .001	-1.27 [-1.43, -1.11]	-10.45 (0.51)	< .001	-1.08 [-1.26, -0.91]	-11.26 (0.49)	< .001	-1.27 [-1.47, -1.07]
+ 6 BDI-II	IC	-11.33 (0.37)	< .001	-1.4 [-1.56, -1.24]	-10.87 (0.43)	< .001	-1.25 [-1.43, -1.07]	-10.59 (0.44)	< .001	-1.22 [-1.38, -1.05]	-10.12 (0.5)	< .001	-1.05 [-1.22, -0.87]
	CoD	-9.74 (0.38)	< .001	-1.2 [-1.35, -1.06]	-10.13 (0.45)	< .001	-1.17 [-1.33, -1.01]	-9.77 (0.52)	< .001	-0.99 [-1.16, -0.82]	-10.57 (0.5)	< .001	-1.16 [-1.35, -0.97]
+ 9 BDI-II	IC	-10.97 (0.38)	< .001	-1.31 [-1.46, -1.15]	-10.43 (0.44)	< .001	-1.16 [-1.33, -0.99]	-10.16 (0.45)	< .001	-1.13 [-1.29, -0.97]	-9.65 (0.51)	< .001	-0.97 [-1.13, -0.8]
	CoD	-9.2 (0.39)	< .001	-1.09 [-1.23, -0.96]	-9.50 (0.46)	< .001	-1.05 [-1.2, -0.9]	-9.09 (0.53)	< .001	-0.89 [-1.05, -0.73]	-9.87 (0.51)	< .001	-1.04 [-1.22, -0.86]
Patient Health Questionnaire-9													
MAR	IC	-5.22 (0.19)	< .001	-1.20 [-1.32, -1.08]	-5.30 (0.22)	< .001	-1.14 [-1.28, -1.00]	-4.86 (0.25)	< .001	-0.95 [-1.08, -0.81]	-4.74 (0.28)	< .001	-0.87 [-1.01, -0.73]
	CoD	-4.47 (0.20)	< .001	-1.04 [-1.16, -0.92]	-4.58 (0.25)	< .001	-0.93 [-1.06, -0.79]	-4.62 (0.27)	< .001	-0.85 [-0.99, -0.71]	-5.13 (0.28)	< .001	-1.00 [-1.16, -0.84]
+ 2 PHQ-9	IC	-4.98 (0.19)	< .001	-1.15 [-1.27, -1.02]	-5.00 (0.23)	< .001	-1.07 [-1.20, -0.93]	-4.57 (0.25)	< .001	-0.89 [-1.03, -0.76]	-4.43 (0.28)	< .001	-0.81 [-0.95, -0.67]
	CoD	-4.11 (0.20)	< .001	-0.96 [-1.08, -0.84]	-4.16 (0.25)	< .001	-0.83 [-0.96, -0.7]	-4.17 (0.27)	< .001	-0.76 [-0.89, -0.63]	-4.67 (0.28)	< .001	-0.91 [-1.06, -0.75]
+ 4 PHQ-9	IC	-4.74 (0.2)	< .001	-1.07 [-1.19, -0.95]	-4.71 (0.23)	< .001	-0.98 [-1.11, -0.84]	-4.28 (0.26)	< .001	-0.82 [-0.95, -0.69]	-4.12 (0.29)	< .001	-0.74 [-0.88, -0.61]
	CoD	-3.75 (0.2)	< .001	-0.85 [-0.97, -0.73]	-3.75 (0.26)	< .001	-0.73 [-0.85, -0.60]	-3.71 (0.28)	< .001	-0.66 [-0.78, -0.53]	-4.21 (0.29)	< .001	-0.79 [-0.94, -0.64]
+ 6 PHQ-9	IC	-4.49 (0.20)	< .001	-0.97 [-1.09, -0.85]	-4.42 (0.24)	< .001	-0.87 [-1.00, -0.74]	-4.00 (0.26)	< .001	-0.74 [-0.87, -0.61]	-3.81 (0.29)	< .001	-0.66 [-0.8, -0.53]
	CoD	-3.4 (0.22)	< .001	-0.73 [-0.84, -0.61]	-3.33 (0.27)	< .001	-0.61 [-0.73, -0.49]	-3.26 (0.29)	< .001	-0.55 [-0.67, -0.43]	-3.75 (0.30)	< .001	-0.67 [-0.8, -0.53]

Note. CoD, Contact on Demand group; IC, Individual Counseling group; MAR, missing at random; BDI-II, Beck Depression Inventory II; PHQ-9, Patient Health Questionnaire 9.

ONLINE TABLE 3

Rates of reliable change, remission and recovery under various NMAR conditions

	Pre- to post-treatment			pre to 3-month follow-up			pre to 6-month follow-up			pre to 12-month follow-up		
	IC	CoD	Total	IC	CoD	Total	IC	CoD	Total	IC	CoD	Total
MAR imputation												
Improvement	312.3[b] (56.3)	268.9 (50.4)	581.2 (53.4)	320.0 (57.7)	282.6 (52.9)	602.6 (55.3)	303.9 (54.8)	284.3 (53.2)	588.3 (54.0)	314.1 (56.6)	304.8 (57.1)	618.9 (56.8)
No change	233.3 (42.0)	254 (47.6)	487.3 (44.7)	223.6 (40.3)	231.6 (43.4)	455.3 (41.8)	223.6 (40.3)	218.8 (41.0)	442.5 (40.6)	210.6 (37.9)	207.2 (38.8)	417.8 (38.4)
Deterioration	9.4 (1.7)	11.1 (2.1)	20.5 (1.9)	11.4 (2.1)	19.8 (3.7)	31.2 (2.9)	27.5 (4.9)	30.8 (5.8)	58.3 (5.4)	30.4 (5.5)	22.0 (4.1)	52.3 (4.8)
asymptomatic	151.4 (27.4)	115.4 (21.7)	266.8 (24.6)	165.3 (29.9)	147.5 (27.7)	312.8 (28.8)	153.7 (27.8)	156.5 (29.4)	310.2 (28.6)	159.2 (28.8)	181.9 (34.1)	341.1 (31.4)
RCSC	251 (60.8)	209.2 (55.6)	460.2 (58.3)	257 (62.2)	220.7 (58.7)	477.7 (60.5)	248.9 (60.3)	227.3 (60.5)	476.2 (60.4)	253.9 (61.5)	240 (63.3)	493.9 (62.6)
MAR Imputation + 2 PHQ-9 points												
Improvement	299.8 (54)	251.1 (47)	550.9 (50.6)	305.3 (55)	265 (49.6)	570.2 (52.4)	290.5 (52.3)	266.5 (49.9)	557 (51.1)	300.9 (54.2)	284.7 (53.3)	585.6 (53.8)
No change	244.3 (44)	269.1 (50.4)	513.4 (47.1)	236.2 (42.6)	245.6 (46)	481.8 (44.2)	234.9 (42.3)	232.1 (43.5)	467 (42.9)	221.5 (39.9)	223.6 (41.9)	445.1 (40.9)
Deterioration	11 (2)	13.7 (2.6)	24.7 (2.3)	13.6 (2.5)	23.4 (4.4)	37 (3.4)	29.6 (5.3)	35.4 (6.6)	65 (6)	32.7 (5.9)	25.7 (4.8)	58.3 (5.4)
asymptomatic	137.2 (24.8)	99.4 (18.6)	236.6 (21.8)	148.7 (26.9)	127 (23.8)	275.8 (25.4)	140.1 (25.3)	136.3 (25.6)	276.4 (25.5)	144.3 (26.1)	159.7 (30)	304 (28)
RCSC	240.3 (58.2)	195 (51.9)	435.4 (55.2)	245.5 (59.4)	207.7 (55.2)	453.3 (57.5)	238.1 (57.7)	213.8 (56.9)	451.8 (57.3)	241.9 (58.6)	224.8 (59.8)	466.6 (59.1)
MAR Imputation + 4 PHQ-9 points												
Improvement	288.4 (52)	234.9 (44)	523.2 (48)	291.8 (52.6)	247.3 (46.3)	539.1 (49.5)	278.3 (50.1)	249.5 (46.7)	527.8 (48.5)	287.5 (51.8)	264.8 (49.6)	552.3 (50.7)
No change	253.1 (45.6)	280.7 (52.6)	533.7 (49)	246 (44.3)	256.7 (48.1)	502.7 (46.2)	243.7 (43.9)	242.4 (45.4)	486.1 (44.6)	230.9 (41.6)	237.7 (44.5)	468.6 (43)
Deterioration	13.6 (2.5)	18.5 (3.5)	32 (2.9)	17.3 (3.1)	30 (5.6)	47.3 (4.3)	33.1 (6)	42.1 (7.9)	75.1 (6.9)	36.6 (6.6)	31.5 (5.9)	68.1 (6.3)
asymptomatic	132.5 (24)	93.9 (17.6)	226.4 (20.8)	141.4 (25.6)	118.6 (22.3)	260 (23.9)	132.6 (24)	122.1 (22.9)	254.7 (23.5)	135.5 (24.5)	142.7 (26.8)	278.2 (25.6)
Recovery [a]	228.2 (55.25)	179 (47.6)	407.2 (51.6)	232.8 (56.37)	191 (50.8)	423.8 (53.7)	224.4 (54.33)	198.2 (52.7)	422.5 (53.5)	228.5 (55.33)	207.6 (55.2)	436.1 (55.3)
MAR Imputation + 6 PHQ-9 points												
Improvement	280.6 (50.6)	224 (41.9)	504.6 (46.3)	281.7 (50.8)	235.1 (44)	516.8 (47.5)	270.3 (48.7)	235.4 (44.1)	505.7 (46.4)	277.8 (50.1)	248.6 (46.6)	526.3 (48.3)
No change	256.3 (46.2)	283 (53)	539.2 (49.5)	249.6 (45)	258.6 (48.4)	508.2 (46.7)	245.5 (44.2)	245.8 (46)	491.2 (45.1)	234.1 (42.2)	244.1 (45.7)	478.1 (43.9)
Deterioration	18.2 (3.3)	27 (5.1)	45.2 (4.2)	23.7 (4.3)	40.3 (7.5)	64 (5.9)	39.2 (7.1)	52.9 (9.9)	92.1 (8.5)	43.2 (7.8)	41.4 (7.8)	84.5 (7.8)
asymptomatic	131.6 (23.8)	92.7 (17.4)	224.3 (20.7)	139.8 (25.3)	115 (21.6)	254.8 (23.5)	131.2 (23.7)	118.8 (22.3)	249.9 (23)	133.2 (24.1)	138.2 (25.9)	271.4 (25)
RCSC	219.6 (53.2)	169.3 (45)	389 (49.3)	223.1 (54)	179.1 (47.6)	402.2 (51)	215.5 (52.2)	184.2 (49)	399.8 (50.7)	218 (52.8)	191.9 (51)	409.9 (52)

Note. CoD , Contact on Demand group; IC, Individual Counseling group; NMAR, not missing at random; MAR, missing at random; RCSC, reliable change significant change.

ONLINE TABLE 4

Means and standard deviations for each outcome at each measurement occasion

		Pre-Treatment		Post-treatment		3-month FU		6-month FU		12-month FU	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
BDI-II	IC	21.987	3.947	9.932	7.656	10.242	8.246	10.529	8.139	10.940	9.089
	CoD	21.802	3.986	10.997	7.641	10.427	8.003	10.676	9.198	9.854	8.815
PHQ-9	IC	11.886	3.353	6.668	3.731	6.591	4.024	7.027	4.415	7.147	4.786
	CoD	11.670	3.471	7.205	4.002	7.091	4.552	7.052	4.831	6.541	4.678
GAD-7	IC	9.804	3.585	5.910	3.298	6.09	4.071	6.241	4.168	5.874	3.807
	CoD	9.554	3.480	6.328	3.475	6.045	3.886	6.104	4.066	5.788	3.825
PTQ	IC	37.159	8.752	28.739	11.871	27.364	12.766	27.829	12.154	26.68	12.443
	CoD	37.591	8.684	29.918	11.561	27.767	13.147	26.744	13.290	25.859	13.350
ESWE	IC	15.008	2.704	16.118	2.541	15.694	2.729	15.66	2.690	15.684	2.857
	CoD	15.198	2.711	15.784	2.792	15.679	2.897	15.611	2.872	15.902	2.994
WHO-5	IC	6.507	3.442	11.758	5.711	11.656	5.966	11.478	6.082	11.373	6.002
	CoD	6.770	3.749	11.260	5.794	11.582	6.033	12.023	6.067	11.891	6.115
BSSS	IC	25.533	4.972	26.987	4.922	26.327	5.332	26.187	5.514	26.009	5.717
	CoD	25.832	5.018	26.602	5.056	26.319	5.429	26.484	5.443	26.902	5.263

Note. All results are pooled across 100 imputed data sets; CoD, Contact on Demand group; IC, Individual Counseling group; BDI-II, Beck Depression Inventory; BSSS, Berlin Social Support Scale perceived social support; ESWE, Scale to Assess Emotional Self-Efficacy; GAD-7, General Anxiety Disorder – 7-item scale; PHQ-9, Patient Health Questionnaire-9; PTQ, Perseverative Thinking Questionnaire; WHO-5, Well Being Index.

APPENDICES FOR STUDY 2

APPENDIX A. Supplementary Information on Interpreting Change Scores

Overall, interpreting change scores and their relationships to predictor variables is challenging. To ease the interpretation of the coefficients reported in the paper, we aim to clarify the interpretation with some illustrative (fictitious) examples.

Interpreting Change Scores Derived from PHQ-9 Ratings

The model depicting changes in depression in this paper implies that

- *change scores < 0* ("negative change scores") indicate symptom *improvement*.

  For example, Whitney's PHQ-9 sum score (assessing severity of depressive symptoms) at the beginning of treatment (module 1; M1) is 12. Two weeks later (module 3; M3) her score fell to 8. Consequently, her change score is $\Delta_{M3\text{-}M1,\ Whitney} = 8 - 12 = -4$. She improved by 4 points.

- *change scores > 0* ("positive change scores") indicate symptom *deterioration*.

  For example, John's PHQ-9 sum score at module 1 is 12. Two weeks later (module 3; M3) his score increased to 16. Consequently, his change score is $\Delta_{M3\text{-}M1,\ Whitney} = 16 - 12 = 4$. He deteriorated by 4 points.

It is important to note, that wording like "*higher change score"* or *"lower change score"* refers to neither improvement nor deterioration, but to the numerical meaning of the scores, that is, -1 is higher than -3 and 1 is higher than -3 ($-3 < -1 < 1 < 3$) and so on.

Change Scores Regressed on "Simple" Predictors

General remarks. A depression change score can be regressed on a predictor. In our example, higher values of this predictor are indicative of higher symptom severity on the predictor variable (e.g. higher scores in the PHQ-stress module indicate more perceived stress). A positive regression weight of $b_{PRE\text{-}STRESS} = 0.157$ in this instance indicates that increases in reported PHQ-S scores by one unit of the questionnaire correspond to a change score that is heightened by .157 units (of the PHQ-9 depressive symptom questionnaire).

Illustrative example. James and Whitney are expected to *improve* regarding their depressive symptoms between Module 3 and 5. That is, their change score is negative, based on estimates derived from all other predictors [$\alpha_{M5\text{-}M3} = -1$ for both individuals]. However, Whitney's perceived stress at pre-assessment ($STRESS_{PRE} = 1$) is one unit higher compared to

James' ($STRESS_{PRE} = 0$). Consequently, Whitney [$E(\Delta_{M5\text{-}M3,\ Whitney}) = \alpha_{M5\text{-}M3} + STRESS_{PRE,\ Whitney} \times b_{PRE\text{-}STRESS} = -1 + 1 \times .157 = -0.843$] will have an expected change score that is .157 units higher compared to James [$E(\Delta_{M5\text{-}M3,\ James}) = \alpha_{M5\text{-}M3} + STRESS_{PRE,\ James} \times b_{PRE\text{-}STRESS} = -1 + 0 \times .157 = -1.000$]. In summary, both individuals are expected to improve, but Whitney is expected to *improve less* given her heightened stress levels.

In contrast, Daniel and Mary are expected to *deteriorate* regarding their depressive symptoms between Module 3 and 5. That is, their change score is positive, based on estimates derived from all other predictors [$\alpha_{M5\text{-}M3} = 1$]. However, Daniel's perceived stress at pre-assessment ($STRESS_{PRE,\ Daniel} = 1$) is one unit higher compared to Mary's ($STRESS_{PRE,\ Mary} = 0$). Consequently, Daniel [$E(\Delta_{M5\text{-}M3,\ Daniel}) = \alpha_{M5\text{-}M3} + STRESS_{PRE,\ Daniel} \times b_{PRE\text{-}STRESS} = 1 + 1 \times .157 = 1.157$] will have an expected change score that is .157 units higher compared to Mary's [$E(\Delta_{M5\text{-}M3,\ Mary}) = \alpha_{M5\text{-}M3} + STRESS_{PRE,\ Mary} \times b_{PRE\text{-}STRESS} = 1 + 0 \times .157 = 1.0$]. In summary, both individuals are expected to deteriorate, but Daniel is expected to *deteriorate more* given his heightened stress levels.

Change Scores Regressed on Change Scores

General remarks. When relating change-scores to change-scores, the interpretation is more complex than in the examples above. For illustrative purposes, let us assume that change in depression is a significant predictor of later change in depression ($b_{M3\text{-}M1} = -0.177$). A negative regression coefficient indicates that higher change scores on the predictor are associated with lower change scores on the dependent variable.

Illustrative example. Again we take Whitney and James as example. Both individuals are expected to *improve* regarding their depressive symptoms between Module 3 and 5. That is, their change score is negative, based on estimates derived from all other predictors [$\alpha_{M5\text{-}M3} = -1$ for both individuals]. However, there are differences between Whitney and James regarding their previous symptom changes (changes between modules 1 and 3; $\Delta_{M3\text{-}M1}$). While both individuals improved during the first modules, James ($\Delta_{M3\text{-}M1,\ James} = -2$) showed bigger improvements compared to Whitney ($\Delta_{M3\text{-}M1,\ Whitney} = -1$). Since James showed a stronger improvement between M3 and M1, we expect a higher subsequent change score for James [$E(\Delta_{M5\text{-}M3,\ James}) = \alpha_{M5\text{-}M3} + \Delta_{M3\text{-}M1,\ James} \times b_{M3\text{-}M1} = -1 + -2 \times -0.177 = -.646$] than for Whitney [$E(\Delta_{M5\text{-}M3,\ Whitney}) = \alpha_{M5\text{-}M3} + \Delta_{M3\text{-}M1,\ Whitney} \times b_{M3\text{-}M1} = -1 + -1 \times -0.177 = -.823$]. In summary, both

individuals are expected to improve, but James is expected to *improve less* given that he already improved more during the previous time period.

Again, Daniel and Marry show a different change pattern. Both individuals are expected to *deteriorate* regarding their depressive symptoms between Module 3 and 5. That is, their expected change score is positive, based on estimates derived from all other predictors [$\alpha_{M5\text{-}M3}$ = 1 for both individuals]. During the first two modules Daniel ($\Delta_{M3\text{-}M1,\ Daniel}$ = 2) showed a stronger increase in symptom load than Mary ($\Delta_{M3\text{-}M1,\ Mary}$ = 1), therefore, we expect a lower change score for Daniel (1 + 2 × -0.177 = 0.646) than for Mary (1 + 1 × -0.177 = 0.823). In summary, both individuals are expected to deteriorate, but Mary is expected to *deteriorate more*, given that she deteriorated less during the previous time period.

Describing the Results in Relation to Change Score Means

As highlighted by Hauk et al. (2016) it can be helpful to take the average change score into account to assess the extent of changes in the majority of individuals (we assume that the mean change score is indicative for the direction of change in the majority of individuals and not confounded by a subset of influential outliers). In the previous example, both change score means are negative. This indicates that the majority of individuals improved during both time periods (between measurement occasions M1-M3 and M3-M5).

In this case, the most accurate description of the results would be:

- *"Individuals that showed higher change scores during the first time period have lowered expected change scores during the consecutive time period."*

However, this description is somewhat difficult to understand and given the negative mean of both change scores (indicating improvements in the majority of individuals during both time periods) it seems reasonable to adapt the following wording to describe the results while keeping in mind that this wording is simplified:

- Example 1: *"On average, individuals that showed stronger improvements during the first time period improved less during the consecutive time period."*
- Example 2: *"On average, individuals that showed more favorable symptom courses (i.e., stronger improvements or less deterioration) during the first time period, showed less favorable symptom courses (i.e., stronger deterioration or less improvement) during the consecutive time period."*

Again, this example shows that the interpretation of beta-coefficients related to change scores as predictors of change scores depends on the value of the beta-coefficient (positive or negative) *and* on the mean of both change scores. An illustration on how the interpretations of change scores depend on beta-coefficients and means/intercepts is given in FIGURE A1 on the following page.

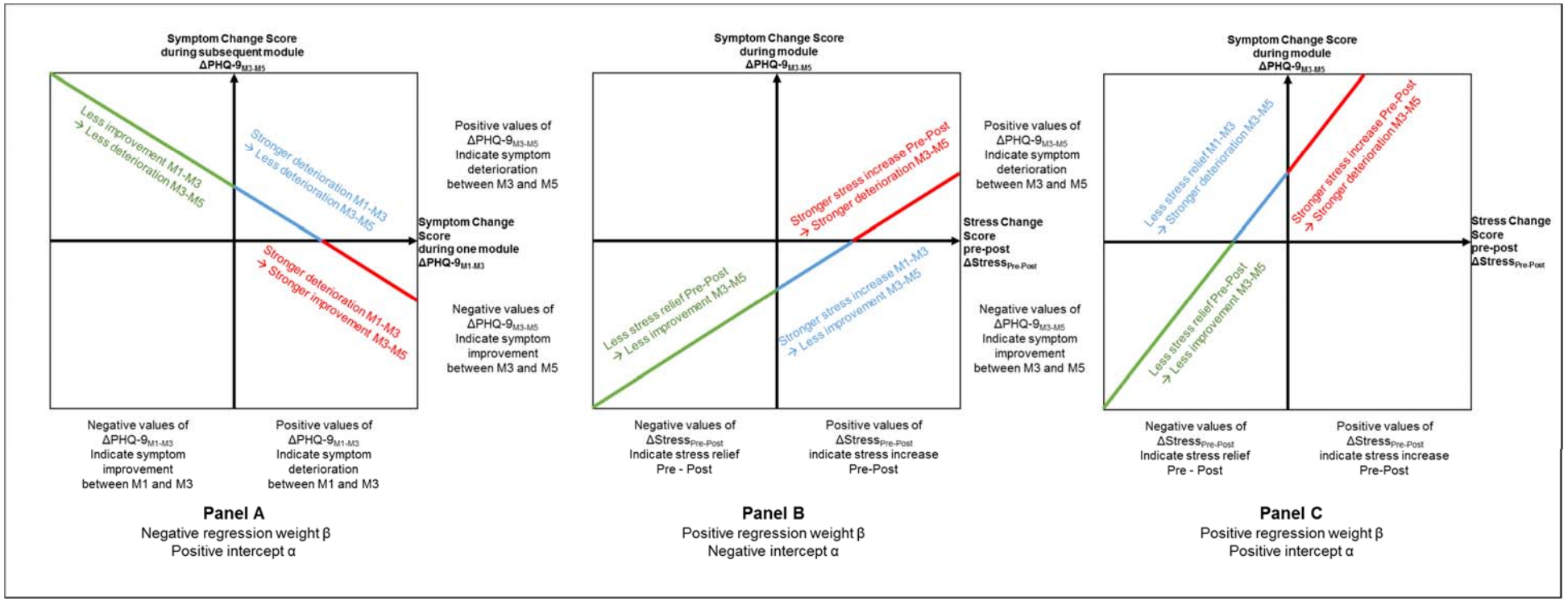


FIGURE A1. Different fictitious constellations of intercepts and regression weights and their consequences for interpreting the regression of change scores on change scores (illustrated in different colors). Panel A depicts the regression of depressive symptom change on previous depressive symptom change. The resulting regression weight is negative. The effect depicted in this panel can be summarized as "More favorable symptom courses in one module are associated with less favorable symptom developments in the next module". Panel B and C depict the regression of depressive symptom change on changes in stress. The resulting regression weights are positive. The differences in the intercepts between Panel B (negative) and Panel C (positive) can result in different interpretations depending on the specific change score values of individuals. Regardless, the effects depicted in both panel B and C can be summarized as "More favorable stress symptom courses are associated with more favorable depressive symptom courses".

APPENDIX B. Supplementary Tables

TABLE B.1

Estimated Model-Fit For the Different Models

Model	Name	χ^2 (df)		RMSEA		CFI		SRMR		Change in Fit (Δ)		
		M	SD	*M*	*SD*	*M*	*SD*	*M*	*SD*	RMSEA	CFI	SRMR
1	depression only - no invariance restrictions	192.899 (100)	16.597	.041	.004	.989	.002	.033	.002	--	--	--
2	Model 1 + equal loadings across groups and time	232.533 (118)	17.403	.042	.003	.986	.002	.037	.002	.001	-.003	.004
3	Model 2 + equal intercepts across groups and time	269.725 (136)	17.766	.042	.003	.984	.002	.037	.002	.000	-.002	.000
4	Model 3 + other psychological constructs	896.217 (486)	29.530	.039	.001	.971	.002	.038	.001	--	--	--
5	Model 4 + other psychological constructs with equal loadings across groups and time	899.611 (492)	28.898	.039	.001	.972	.002	.038	.001	.000	.001	.000
6	Model 5 + other psychological constructs with equal intercepts across groups and time	912.466 (498)	29.928	.039	.001	.971	.002	.038	.001	.000	-.001	.000
7	Model 6 + objective indicators	1570.851 (914)	36.819	.036	.001	.958	.002	.033	.001	-.003	-.013	-.005

Note. Fit indices for the different models, as well as change in RMSEA, CFI and SRMR after adding invariance restrictions. All results are pooled across the 100 imputed data-sets.

TABLE B.2

Within-Group and Between-Group Differences in Mean-Change and respective Standardized Effect Measures

	IF-Condition				SF-Condition				Between-Group Differences			
	M (*SE*)	*p*	d_w	[95%CI]	*M* (*SE*)	*p*	d_w	[95%CI]	*M* (*SE*)	*p*	d_b	[95%CI]
DEP_{M1}	1.201 (0.018)	< .001	---	---	1.202 (0.02)	< .001	---	---	---	---	---	---
Δ(M3-M1)	-0.161 (0.016)	< .001	-0.577	[-0.702, -0.452]	-0.125 (0.016)	< .001	-0.511	[-0.654, -0.368]	0.036 (0.023)	.111	0.137	[-0.032, 0.306]
Δ(M5-M3)	-0.196 (0.018)	< .001	-0.686	[-0.849, -0.523]	-0.192 (0.021)	< .001	-0.560	[-0.699, -0.421]	0.004 (0.027)	.885	0.012	[-0.155, 0.179]
Δ(M7-M5)	-0.112 (0.016)	< .001	-0.461	[-0.608, -0.314]	-0.157 (0.018)	< .001	-0.571	[-0.712, -0.430]	-0.045 (0.024)	.054	-0.175	[-0.353, 0.003]
Δ(POST-M7)	0.025 (0.016)	.117	0.108	[-0.029, 0.245]	0.094 (0.017)	< .001	0.410	[0.257, 0.563]	0.069 (0.023)	.003	0.297	[0.103, 0.491]

Note. Mean derived from the model assuming strong measurement invariance across groups and measurement occasions. Within-group effects were calculated in terms of standardized response means (dw) that is, dividing the mean LD by the group-specific standard deviation of the LD. Between-group effect sizes were estimated by dividing the difference in mean LD scores by the pooled standard deviation. Δ(M3-M1) = Change between Module 1 and Module 3, negative values indicate that individuals improved on average. d_w /d_b = standardized within/between group difference.

TABLE B.3 (CONT. ON FOLLOWING PAGE)

Significant and non-significant regression weights in both treatment conditions.

		IF-Condition				SF-Condition			
Path		*b*	[95% CI]	p_{indi}	b_{std} (SE)	*b*	[95% CI]	p_{indi}	b_{std} (SE)
ΔDEP_{M3-M1} ON	EXP_{PRE}	0.016	[-0.035, 0.067]	.538	0.037 (0.060)	0.026	[-0.031, 0.083]	.359	0.070 (0.076)
	$STRESS_{PRE}$	0.034	[-0.076, 0.144]	.544	0.039 (0.064)	0.008	[-0.100, 0.116]	.884	0.011 (0.072)
	$\Delta STRESS_{POST-PRE}$	0.156	[0.044, 0.268]	.006**	0.200 (0.071)	0.124	[0.018, 0.230]	.022*	0.186 (0.084)
	$LOGDUR_{M1M2}$	-0.004	[-0.012, 0.004]	.306	-0.068 (0.067)	-0.002	[-0.010, 0.006]	.628	-0.038 (0.079)
	$LOGINS_{M1M2}$	0.013	[-0.028, 0.054]	.518	0.041 (0.063)	-0.003	[-0.054, 0.048]	.895	-0.011 (0.084)
	$LETSIZE_{M1M2}$	-0.003	[-0.015, 0.009]	.633	-0.031 (0.064)	0.001	[-0.015, 0.017]	.877	0.013 (0.082)
ΔDEP_{M5-M3} ON	ΔDEP_{M3-M1}	-0.177	[-0.346, -0.008]	.040*	-0.172 (0.083)	-0.257	[-0.500, -0.014]	.038*	-0.184 (0.089)
	EXP_{PRE}	0.011	[-0.044, 0.066]	.697	0.024 (0.063)	-0.016	[-0.075, 0.043]	.593	-0.031 (0.058)
	$STRESS_{PRE}$	0.019	[-0.093, 0.131]	.745	0.020 (0.063)	0.148	[0.013, 0.283]	.032*	0.141 (0.064)
	$\Delta STRESS_{POST-PRE}$	0.168	[0.048, 0.288]	.006**	0.208 (0.074)	0.184	[0.039, 0.329]	.013*	0.198 (0.078)
	$LOGDUR_{M1M2}$	-0.001	[-0.009, 0.007]	.763	-0.023 (0.075)	-0.002	[-0.012, 0.008]	.641	-0.032 (0.069)
	$LOGINS_{M1M2}$	-0.004	[-0.047, 0.039]	.868	-0.011 (0.065)	0.028	[-0.021, 0.077]	.260	0.065 (0.058)
	$LETSIZE_{M1M2}$	0.009	[-0.005, 0.023]	.219	0.087 (0.070)	-0.009	[-0.027, 0.009]	.349	-0.063 (0.067)
	$LOGDUR_{M3M4}$	0.000	[-0.004, 0.004]	.947	-0.005 (0.070)	-0.001	[-0.007, 0.005]	.704	-0.030 (0.079)
	$LOGINS_{M3M4}$	-0.018	[-0.038, 0.002]	.084	-0.144 (0.082)	-0.024	[-0.051, 0.003]	.086	-0.143 (0.081)
	$PLANFREQ_{M3M4}$	0.000	[-0.002, 0.002]	.845	-0.014 (0.074)	0.000	[-0.002, 0.002]	.759	0.025 (0.082)
ΔDEP_{M7-M5} ON	ΔDEP_{M5-M3}	-0.264	[-0.446, -0.082]	.005**	-0.312 (0.106)	-0.294	[-0.455, -0.133]	< .001***	-0.363 (0.094)
	EXP_{PRE}	-0.011	[-0.064, 0.042]	.675	-0.030 (0.071)	-0.001	[-0.060, 0.058]	.983	-0.002 (0.071)
	$STRESS_{PRE}$	0.011	[-0.079, 0.101]	.811	0.014 (0.060)	-0.021	[-0.135, 0.093]	.715	-0.025 (0.067)
	$\Delta STRESS_{POST-PRE}$	0.196	[0.094, 0.298]	< .001***	0.289 (0.076)	0.082	[-0.024, 0.188]	.132	0.109 (0.072)
	$TASK_{M5}$	-0.034	[-0.089, 0.021]	.223	-0.114 (0.092)	-0.002	[-0.073, 0.069]	.945	-0.007 (0.099)
	$BOND_{M5}$	0.007	[-0.040, 0.054]	.780	0.024 (0.085)	-0.017	[-0.060, 0.026]	.428	-0.066 (0.084)
	$LOGDUR_{M3M4}$	-0.001	[-0.005, 0.003]	.771	-0.031 (0.105)	0.004	[-0.002, 0.010]	.205	0.136 (0.109)
	$LOGINS_{M3M4}$	-0.008	[-0.030, 0.014]	.484	-0.072 (0.103)	0.020	[-0.007, 0.047]	.156	0.144 (0.100)
	$PLANFREQ_{M3M4}$	0.001	[-0.001, 0.003]	.118	0.115 (0.072)	-0.001	[-0.003, 0.001]	.519	-0.051 (0.079)
	$LOGDUR_{M5M6}$	-0.001	[-0.007, 0.005]	.746	-0.035 (0.110)	-0.003	[-0.013, 0.007]	.486	-0.085 (0.123)
	$LOGINS_{M5M7}$	0.011	[-0.014, 0.036]	.392	0.075 (0.088)	0.015	[-0.024, 0.054]	.444	0.072 (0.096)
	$NBTFREQ_{M5M7}$	-0.001	[-0.007, 0.005]	.637	-0.038 (0.081)	-0.003	[-0.009, 0.003]	.316	-0.088 (0.087)
	$PROTFREQ_{M5M7}$	-0.001	[-0.015, 0.013]	.836	-0.015 (0.071)	-0.004	[-0.026, 0.018]	.723	-0.028 (0.079)

Path		IF-Condition				SF-Condition			
		b	[95% CI]	p_{indi}	b_{std} (SE)	*b*	[95% CI]	p_{indi}	b_{std} (SE)
$\Delta DEP_{POST-M7}$ ON	ΔDEP_{M7-M5}	-0.423	[-0.599, -0.247]	< .001***	-0.439 (0.090)	-0.242	[-0.414, -0.070]	.006**	-0.290 (0.104)
	EXP_{PRE}	0.008	[-0.041, 0.057]	.737	0.023 (0.069)	0.036	[-0.015, 0.087]	.173	0.102 (0.074)
	$STRESS_{PRE}$	0.087	[-0.005, 0.179]	.065	0.118 (0.063)	0.157	[0.057, 0.257]	.002**	0.221 (0.072)
	$\Delta STRESS_{POST-PRE}$	0.133	[0.031, 0.235]	.010	0.203 (0.076)	0.244	[0.150, 0.338]	< .001***	0.388 (0.071)
	$TASK_{M5}$	-0.089	[-0.136, -0.042]	< .001***	-0.307 (0.084)	0.012	[-0.053, 0.077]	.727	0.038 (0.109)
	$BOND_{M5}$	0.051	[0.010, 0.092]	.014*	0.188 (0.077)	-0.021	[-0.064, 0.022]	.339	-0.098 (0.101)
	$LOGINS_{M5M7}$	0.010	[-0.012, 0.032]	.327	0.072 (0.073)	-0.001	[-0.030, 0.028]	.947	-0.006 (0.085)
	$LOGDUR_{M5M6}$	-0.002	[-0.008, 0.004]	.500	-0.060 (0.089)	-0.002	[-0.008, 0.004]	.478	-0.071 (0.099)
	$NBTFREQ_{M5M7}$	0.004	[-0.002, 0.010]	.109	0.123 (0.076)	-0.001	[-0.007, 0.005]	.841	-0.018 (0.092)
	$PROTFREQ_{M5M7}$	-0.007	[-0.021, 0.007]	.318	-0.078 (0.078)	-0.006	[-0.024, 0.012]	.482	-0.052 (0.074)
	$LOGINS_{M7POST}$	-0.002	[-0.016, 0.012]	.752	-0.028 (0.089)	-0.008	[-0.020, 0.004]	.229	-0.113 (0.094)
	$LOGDUR_{M7POST}$	0.000	[-0.004, 0.004]	.903	-0.012 (0.098)	0.002	[-0.004, 0.008]	.579	0.062 (0.110)
	$LETSIZE_{M7}$	-0.005	[-0.023, 0.013]	.605	-0.036 (0.071)	-0.015	[-0.040, 0.010]	.235	-0.090 (0.076)
BOND ON	EXP_{PRE}	0.303	[0.148, 0.458]	< .001***	0.226 (0.057)	0.120	[-0.056, 0.296]	.181	0.074 (0.055)
	DEP_{M1}	0.025	[-0.396, 0.446]	.908	0.010 (0.085)	0.094	[-0.382, 0.570]	.699	0.031 (0.081)
	ΔDEP_{M3-M1}	-0.108	[-0.614, 0.398]	.675	-0.035 (0.083)	0.529	[-0.243, 1.301]	.180	0.121 (0.087)
	ΔDEP_{M5-M3}	-0.146	[-0.581, 0.289]	.511	-0.048 (0.073)	-0.594	[-1.074, -0.114]	.015*	-0.190 (0.077)
TASK ON	EXP_{PRE}	0.278	[0.141, 0.415]	< .001***	0.221 (0.053)	0.224	[0.104, 0.344]	< .001***	0.192 (0.051)
	DEP_{M1}	-0.381	[-0.746, -0.016]	.041*	-0.161 (0.080)	-0.530	[-0.857, -0.203]	.002**	-0.246 (0.078)
	ΔDEP_{M3-M1}	-0.749	[-1.186, -0.312]	.001***	-0.258 (0.077)	-0.560	[-1.077, -0.043]	.034*	-0.180 (0.088)
	ΔDEP_{M5-M3}	-0.693	[-1.063, -0.323]	< .001***	-0.246 (0.066)	-0.918	[-1.232, -0.604]	< .001***	-0.410 (0.067)

Note. * $p < .050$; ** $p < .010$; *** $p < .001$. b /b_{stdxy} = un-/standardized regression weight; DEP = PHQ-9 depression score; STRESS = PHQ-Stress score; EXP = expectations; BOND = bond-component of the Working Alliance Inventory; TASK = task-component of the Working Alliance Inventory. Measurement occasions are indexed.

APPENDICES FOR STUDY 3

APPENDIX C. Detailed description of the statistical approach.

In the first section of the supporting information we present a detailed description of the applied statistical methods and the steps taken during the analysis. The second section of the supplement offers a more detailed explanation and extension of the results of the single-group, the multi-group and the predictor analysis.

Statistical Methods

Analysis. We used Growth Mixture Models (GMM; Muthén & Shedden, 1999; Ram & Grimm, 2009; please see Morin for an introduction to GMM) with latent base-specification (first and last slope-factor loading fixed at 0 and 1, respectively) to investigate whether the data set contains subgroups of individuals characterized by similar average change trajectories in depressive symptom load (Morin et al., 2011; Morin, Maiano, Marsh, Nagengast, & Janosz, 2013). The latent base specification allows for the modeling of non-linear change (Morin et al., 2011). We used PHQ-9 measurements obtained at the beginning of each specific treatment module to structure the change process. Importantly, the current data set contains data from two separate treatment conditions. Each treatment condition might lead to different patterns of symptom changes in participating patients. Consequently, the number of classes and the quantitative features of these classes (e.g., different amount of change) might differ across conditions. Therefore, a modeling approach considering these conditions seemed necessary (Muthén, Brown, Leuchter, & Hunter, 2008).

In a first step, we analyzed each treatment condition separately to account for the possibility that both treatment conditions show a different number of classes (single-group GMM). As suggested by Diallo, Morin, and Lu (2016), we used a rather unrestricted GMM for class-enumeration. Means (μ_{Sk} and μ_{Ik}; *I* represents the intercept-factor and *S* represents the slope-factor), variances (ψ_{Ik} and ψ_{Sk}), and co-variances ($\psi_{Ik,Sk}$), were estimated class-specific (k = 1, …, K, K = number of classes), while slope-loadings (λ_{tk}; $t = 1, \ldots, T$; T = number of

measurement occasions) and the residual terms of the indicators were estimated class and time-specific (ε_{itk}). Only if estimation difficulties occurred (best-likelihood not replicated, negative variances, excessively small classes) and the information criteria did not favor a specific model, more restrictive models were tested. These models included constraints on the residual terms (ε_{ik}, ε_{it} or ε_{i}). Recent simulation studies have shown, that unrestricted models are beneficial in identifying the appropriate number of classes and help to avoid small classes which might represent artefacts of unnecessary restrictions on variance or co-variance parameters of the model (Diallo et al., 2016, 2017; Peugh & Fan, 2012). Following the suggestions by several groups of authors, the optimal number of classes for each treatment condition was determined on the basis of several statistical information criteria (*Akaike Information Criteria*, AIC; *Corrected AIC*, CAIC; *Bayes Information Criteria*, BIC; *sample-size adjusted* BIC, aBIC; *Vuo-Long-Mendel Likelihood Ratio-Test*, VLM-LRT), the appropriateness of the estimated model parameters as well as the interpretability, distinctiveness and sizes of derived classes to decide for the optimal number of classes (Masyn, 2013; Meyer & Morin, 2016; Morin et al., 2011). In line with recommendations from recent studies, we decided to do class enumeration using an unconditional GMM, that is without inclusion of covariates (Diallo et al., 2017; Masyn, 2013).

In a second step, we specified a multi-group GMM taking the known membership of individuals to two distinct treatment conditions and the number of change patterns derived in the single group-analysis into account (g = SF or g = IF). The multi-group approach allows for the testing of potential differences in the configuration of change patterns and class sizes across treatment conditions. In other words, the MG approach provides answers to the question as to whether the provision of qualitatively different feedback (i.e., individualized vs. standardized) leads to differences in average change of classes. We used a model with the optimal number of classes for each treatment condition as determined by the single-class analysis with group and class-specific means (μ_{Skg} and μ_{Ikg}), variances (ψ_{Ikg} and ψ_{Skg}), co-variances ($\psi_{Ikg,Skg}$) as baseline-model. Furthermore, slope loadings (λ_{tkg}) and residuals (ε_{itkg}) were estimated separately and

specifically for each treatment condition, class and measurement occasion. Class sizes were allowed to vary across treatment conditions. We compared this model against a model assuming the same configuration of means, variances, co-variances and residuals for corresponding classes across treatment conditions with equal class sizes across treatment conditions (μ_{Sk}, μ_{Ig} ψ_{Ik}, ψ_{Sk}, $\psi_{Ik,Sk}$, λ_{tk}, ε_{itk}). We used information criteria to decide on whether the unrestricted or restricted model should be favored.

In a third step, predictor variables for slope and intercept, as well as class-membership were included directly into the model. Therefore, the models consider inaccuracy of class-assignment properly. In a baseline model, all regression weights were allowed to vary group- and class-specific (slope regression-weights: $\beta_{xSkg;}$ intercept regression-weights: β_{xIkg}) which is comparable to including an interaction term into multiple regression. A more restricted model assuming equal regression weights across classes and treatment conditions (βxI and βxS) followed this baseline model. We evaluated information criteria to decide which model should be favored. Given the exploratory fashion of the analysis, we favored parsimony when comparing models.

All models were estimated with M*plus* 8.1 (Muthén & Muthén, 1998-2017; normal distribution, 10'000 initial starts, 500 stage optimizations and 500 initial stage iterations). Missing data on the PHQ-9 scores were dealt with using the full information maximum likelihood estimation procedure, while missing values on predictor variables were replaced using a single value imputation.

Results of Single-Group GMM

Individualized Feedback-Condition. Models with 1 to 3 classes with class-specific, means (μ_{Sk} and μ_{Ik}), slope (ψ_{Sk}) and intercept variances (ψ_{Ik}), co-variances (ψ_{SkIk}) as well as class- and time-specific factor loadings (λ_{tk}) and residuals (ε_{itk}) converged properly. The best likelihood replicated several times. All parameter estimates were within a plausible range (no variances smaller than zero, no inflated standard errors). However, the best likelihood value of

the 4-class solution could not be replicated even after increasing the number of random starts. Since BIC and CAIC reached a plateau at the 3-class solution, no further models with additional restrictions models were estimated and class enumeration was performed considering the statistically sound models with one to three classes (all information criteria are summarized in Supp. Table C2). The AIC and sBIC were uninformative and decreased continuously from the 1-class to the 3-class-soluation. While the VLM-LRT favored a 3-class solution over a 2-class-soluation (still significant after adding a third class), the BIC and CAIC favored a 2-class-solution. Visual inspection of the derived change patterns showed large similarities of two classes in the 3-class solution. Therefore, we favored the 2-class model as the most parsimonious one.

Class 1 (*delayed improvers*) comprises 46.0% of the participants. The average trajectory is marked by an average improvement of μ_{S1} = -3.5, which is below the cut-off suggested for reliable change using the PHQ-9 (Titov et al., 2011). Class 2 (*immediate improvers*) comprises 54.0% of the participants. The average symptom improvement in this class was μ_{S2} = -6.2 points. Interestingly, these two classes differ in their early symptom development. *Immediate improvers* showed a significant proportion of their average symptom improvement immediately after the SCID-I interview and prior to treatment uptake at MI (λ_{12} = 0.266, $p < .001$), while no such changes occurred in the class of delayed improvers (λ_{11} = 0.071, p = .626). Both classes showed not only considerable heterogeneity in initial symptom-load (ψ_{I1} = 6.1, ψ_{I2} = 7.9) but in changes throughout the intervention (ψ_{S1} = 9.3, ψ_{S2} = 9.4). Supp. Table C1 summarizes all estimated parameters. Supp. FIGURE C1 (Panel IF-condition) visualizes the derived change patterns.

Standardized Feedback-Condition. The 1- and 2-class solution with class-specific means (μ_{Sk} and μ_{Ik}), slope (ψ_{Sk}) and intercept variances (ψ_{Ik}), co-variances (ψ_{SkIk}) as well as class- and time-specific factor loadings (λ_{tk}) and residuals (ε_{itk}) converged properly. The best likelihood replicated several times and all parameter estimates were in a plausible range. The

best likelihood value of the 3-class solution did not replicate. In order to determine whether a more parsimonious 3-class model fits the data better, additional restrictions were added. A model with class-specific but time-unspecific residuals did not converge and the best likelihood value could not be replicated (ε_{ik}). However, models with error variances constrained to be equal across classes and measurement occasions (ε_{i}) as well as across classes but specific for each measurement occasion (ε_{it}) converged and were considered during the model selection process. The statistical information criteria (AIC, BIC, SA-BIC, CAIC, see Supp. Table C2) pointed to the more complex 2-class model. Therefore, we selected this model as optimal solution for the SF group (see FIGURE C1).

Class 2 (*immediate improvers*) comprises 64.2% of the participants. The average improvement in the class was μ_{S2} = -5.5 points on the PHQ-9, which is descriptively larger than the 5 points reliable change benchmark of the PHQ-9 (Titov et al., 2011). Class 1 (*delayed improvers*) comprises 35.8% of the participants. The average improvement throughout the interventions was smaller (μ_{S1} = -2.4) than in class one. As in the IF condition, the classes showed considerable difference in the symptom course at early treatment periods. Individuals in the *immediate improvers* class started to show a significant proportion of their average symptom improvement immediately after the SCID-I interview and prior to starting to treatment uptake at MI (λ_{12} = 0.33, $p < .001$) while no beneficial changes occurred in class 1 (λ_{11} = -0.48, $p < .001$; λ_{21} = -0.33, $p < .001$). In contrast, the trajectory showed a slight increase in symptom load followed by improvements later throughout the intervention. Again, both classes showed not only significant heterogeneity in initial symptom-load (ψ_{I1} = 6.3, ψ_{I2} = 6.2) but in changes throughout the intervention (ψ_{S1} = 7.2, ψ_{S2} = 5.6). Supp. Table C1 summarizes all estimated parameters. Supp. Figure C1 (Panel SF-Condition) visualizes the change patterns.

Results of Multi-Group GMM

The single group analysis pointed towards two classes in each treatment condition. Based on visual inspection, the two classes of both groups showed considerable similarities. *First*, both treatment conditions comprised a class of individuals showing constant improvement throughout the whole treatment. *Second,* we observed a class of individuals with delayed and lower overall-improvements in both treatment groups. Based on these results, we specified a two class-model for each treatment condition. *In a first step*, we allowed class-sizes, (μ_{Skg} and μ_{Ikg}), slope (ψ_{Skg}) intercept variances (ψ_{Ikg}) and co-variances ($\psi_{Skg,Ikg}$) to vary across groups and classes. Additionally, factor loadings (λ_{tkg}) and residuals (ε_{itkg}) were allowed to vary across treatment conditions, classes and time-points. The model converged and the best likelihood values were replicated several times. To check for differences between the treatment conditions, we estimated a more constrained model assuming that not only the numbers of classes are equal, but the configuration of class-specific average change patterns and their sizes were constrained to be equal across the different treatment conditions as well (e.g. there is no condition-specificity, μ_{Sk}, μ_{Ik} ψ_{Ik}, ψ_{Sk}, $\psi_{Ik,Sk}$, λ_{tk}, ε_{itk}). All information criteria pointed towards the more parsimonious model assuming no condition-specificity. This suggests that the patterns of change between both treatment conditions are highly similar.

Results of Multi-Group Modeling, including descriptions of the two classes are provided in the manuscript.

Results of the Predictor Analysis

First, we estimated an unconstrained model estimating the classes within each treatment condition freely (λ_{tkg}, ψ_{Skg}, ψ_{Ikg}, $\psi_{Skg,Ikg}$, ε_{itkg}). Predictors of slopes (β_{xSg}) and intercepts (β_{xIg}) were allowed to vary across treatment conditions, while they were restricted to be equal across condition-specific classes. A model constraining shape of change trajectories and regression weights of predictor variables to be equal across classes and treatment conditions (β_{xg} and $\beta_{xg,}$ respectively) was estimated in a second step. Again, both models resulted in proper parameter

estimates and the best likelihood value was replicated several times. The shape of the derived trajectories did not change. All information criteria clearly favored the most parsimonious model with equal regressive relations across classes and treatment conditions. Therefore, it seems reasonable to assume, that symptom change in the two treatment arms does not only follow the same trajectories but also shares the same predictors. We present a detailed discussion of the estimated paths in the manuscript.

In order to illustrate and contextualize the meaning of the OR in this setting, we would like to provide a comparative example: Current MDD is a dichotomous variable, but expectations are measured with sums of longer scales. Fulfilling the criteria for a current MDD increases the odds of being in the delayed responder class by a factor of 2.8 ($OR_{MDD} = 2.8$). In order to achieve such an increase in the odds, the individually reported expectations need to be raised by 12 units of measurement ($OR_{EXP} = 1.09$; $1.09^{12} = 2.81$).

SUPP. TABLE C1

Model Parameters of the Single-Group GMM and the constrained Multi-Group GMM models.

Parameter	IF-Condition		SF-Condition		MG-Model	
	Delayed Improvers	Immediate Improvers	Delayed Improvers	Immediate Improvers	Delayed Improvers	Immediate Improvers
	Est (SE)	Est (SE)	Est (SE)	Est (SE)	Est (SE)	Est (SE)
λ_{k1}	0	0	0	0	0	0
λ_{k2}	0.07 (0.15)^{NS}	0.27 (0.05)	-0.48 (0.15)	0.33 (0.07)	-0.13 (0.13)^{NS}	0.33 (0.05)
λ_{k3}	0.34 (0.14)	0.54 (0.04)	-0.32 (0.15)	0.57 (0.07)	0.18 (0.15)^{NS}	0.55 (0.04)
λ_{k4}	0.60 (0.14)	0.83 (0.03)	0.34 (0.16)	0.86 (0.04)	0.57 (0.12)	0.84 (0.03)
λ_{k5}	0.78 (0.11)	0.92 (0.03)	0.78 (0.13)	0.87 (0.03)	0.80 (0.07)	0.90 (0.02)
λ_{k6}	1.01 (0.09)	1.02 (0.02)	0.98 (0.12)	1.06 (0.04)	1.03 (0.07)	1.02 (0.02)
λ_{k7}	1.08 (0.07)	1.09 (0.02)	1.13 (0.09)	1.17 (0.03)	1.12 (0.05)	1.11 (0.02)
λ_{k8}	1	1	1	1	1	1
μ_{Ik}	12.35 (0.27)	11.37 (0.28)	11.78 (0.31)	11.41 (0.21)	12.39 (0.33)	11.23 (0.28)
μ_{Sk}	-3.46 (0.49)	-6.23 (0.45)	-2.37 (0.54)	-5.45 (0.32)	-3.41 (0.76)	-5.54 (0.37)
$\psi_{Sk,Ik}$	-2.50 (0.95)	-5.80 (1.30)	-0.74 (0.81)	-2.99 (0.90)	-2.23 (0.87)	-3.42 (1.04)
ψ_{Ik}	6.06 (0.88)	7.87 (1.17)	6.27 (1.03)	6.15 (0.86)	5.75 (0.73)	6.27 (0.78)
ψ_{Sk}	9.30 (1.91)	9.43 (1.36)	7.24 (1.95)	5.64 (1.04)	10.43 (2.48)	7.07 (1.15)
Var(ε_{i1k})	6.40 (1.26)	4.51 (0.99)	8.61 (1.72)	6.21 (1.07)	8.91 (1.60)	5.12 (0.96)
Var(ε_{i2k})	7.47 (1.45)	6.46 (1.23)	3.65 (1.20)	7.79 (1.41)	7.79 (1.94)	5.35 (1.33)
Var(ε_{i3k})	5.33 (0.76)	4.45 (0.66)	4.74 (1.16)	4.43 (1.18)	6.51 (1.46)	4.09 (0.87)
Var(ε_{i4k})	7.28 (1.33)	1.99 (0.34)	8.35 (1.40)	2.40 (0.34)	8.21 (1.14)	2.21 (0.30)
Var(ε_{i5k})	6.92 (1.43)	1.86 (0.42)	9.01 (3.59)	3.53 (0.75)	8.31 (1.71)	2.45 (0.32)
Var(ε_{i6k})	4.72 (0.95)	0.84 (0.35)	6.88 (2.24)	1.91 (0.33)	5.50 (1.17)	1.44 (0.30)
Var(ε_{i7k})	6.13 (1.26)	1.03 (0.41)	5.11 (1.27)	1.12 (0.31)	5.77 (1.08)	1.19 (0.22)
Var(ε_{i8k})	8.88 (1.44)	2.25 (0.79)	8.44 (2.50)	3.85 (0.63)	9.12 (1.56)	3.02 (0.40)

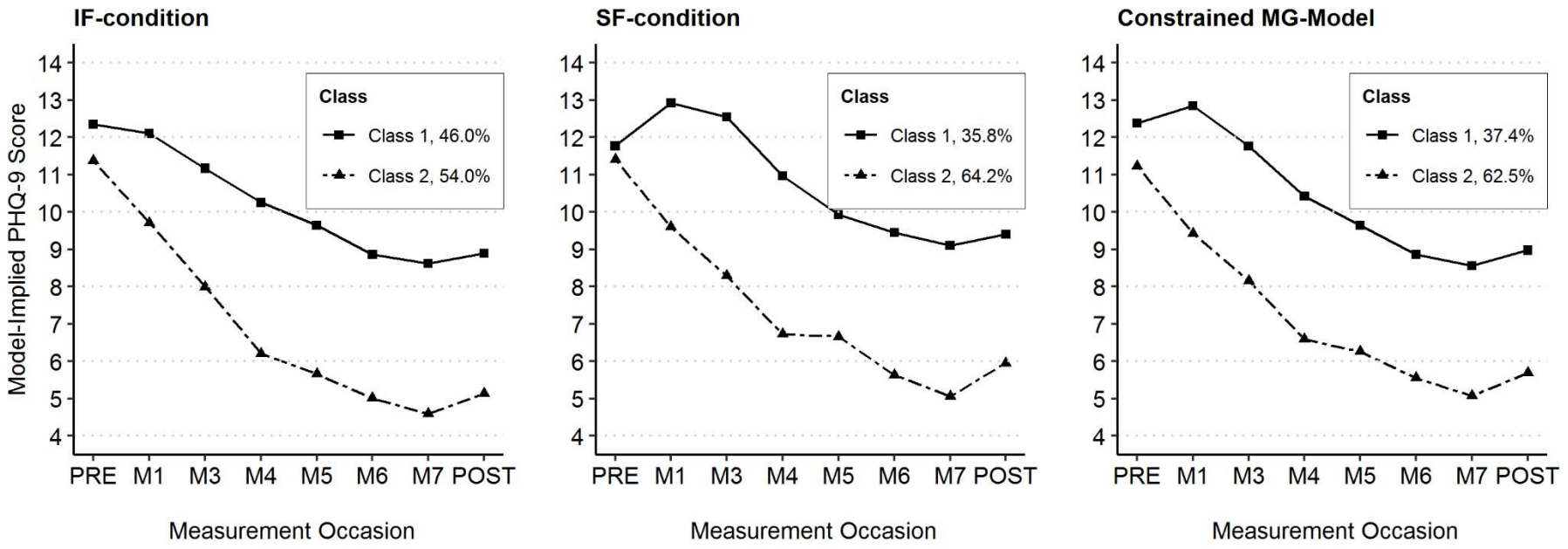
Note. λ_{kt} = class and time specific growth factor-loading, where k = refers to the class and t to the measurement occasions. μ_{Ik} and μ_{Sk} = mean of the intercept and slope, respectively. ψ_{Ik} and ψ_{Sk} = variance of the intercept and slope. $\psi_{Sk,Ik}$ = covariance between slope and intercept. Var(ε_{itk}) = residual variance at the corresponding measurement occasion t. All parameters significant with $p < .05$ if not indicated otherwise. ^{NS} = non-significant.

SUPP. TABLE C2

Information Criteria and Entropy for Each of the Estimated Models.

Model	#par	logL	scaling	AIC	BIC	aBIC	CAIC	VLM-LRT	entropy
IF-Group [a]									
1 class	19	-9584.172	1.294	19206	19288	19228	19320	---	---
2 classes	39	-9417.485	1.369	18913	19081	18958	19147	.014	0.556
3 classes	59	-9356.155	1.195	18830	19085	18898	19184	.013	0.547
SF-Group									
1 class	19	-9013.258	1.590	18065	18146	18086	18178	---	---
2 classes	39	-8861.870	1.425	17802	17969	17845	18035	.011	0.571
3 classes [b]	36	-8915.305	1.275	17903	18057	17942	18118	.186	0.487
3 classes [c]	43	-8888.899	1.422	17864	18048	17911	18121	.538	0.650
MG-Group									
2 classes	79	-19033.990	1.392	38226	38620	38370	38699	---	0.782
2 classes, constrained	40	-19068.615	1.463	38217	38417	38290	38457	---	0.769
MG-Group incl. Predictors									
2 classes	85	-18912.895	1.221	37996	38420	38150	38505	---	0.787
2 classes, constrained	213	-18811.701	1.191	38049	39113	38436	39326	---	0.823

Note. [a] Models were estimated with class specific means, variances, covariance's and class- and time-specific loadings and variances if not stated otherwise. [b] additional constraint: ε_i [c] additional constraints added: ε_{it}



SUPP. FIGURE C1. Estimated Change Pattern of the Single-Group Analysis (IF-Condition and SF) Condition and the Constrained Multi-Group Model

References for Appendix C

Diallo, T. M. O., Morin, A. J. S., & Lu, H. (2016). Impact of Misspecifications of the Latent Variance–Covariance and Residual Matrices on the Class Enumeration Accuracy of Growth Mixture Models. *Structural Equation Modeling: a Multidisciplinary Journal*, *23*, 507–531. https://doi.org/10.1080/10705511.2016.1169188

Diallo, T. M. O., Morin, A. J. S., & Lu, H. (2017). The impact of total and partial inclusion or exclusion of active and inactive time invariant covariates in growth mixture models. *Psychological Methods*, *22*, 166–190. https://doi.org/10.1037/met0000084

Masyn, K. E. (2013). Latent Class Analysis and Finite Mixture Modeling. *The Oxford Handbook of Quantitative Methods*, 551.

Meyer, J. P., & Morin, A. J.S. (2016). A person-centered approach to commitment research: Theory, research, and methodology. *Journal of Organizational Behavior*, *37*, 584–612. https://doi.org/10.1002/job.2085

Morin, A. J. S., Maiano, C., Marsh, H. W., Nagengast, B., & Janosz, M. (2013). School life and adolescents' self-esteem trajectories. *Child Development*, *84*, 1967–1988. https://doi.org/10.1111/cdev.12089

Morin, A. J. S., Maïano, C., Nagengast, B., Marsh, H. W., Morizot, J., & Janosz, M. (2011). General Growth Mixture Analysis of Adolescents' Developmental Trajectories of Anxiety: The Impact of Untested Invariance Assumptions on Substantive Interpretations. *Structural Equation Modeling: a Multidisciplinary Journal*, *18*, 613–648. https://doi.org/10.1080/10705511.2011.607714

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, *55*, 463–469.

Muthén, B., Brown, H., Leuchter, A., & Hunter, A. (2008). General approaches to analysis of course: applying growth mixture modeling to randomized trials of depression medication. *Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures. Washington, DC: American Psychiatric Publishing*, 159–178.

Muthén, L. K., & Muthén, B. (1998-2017). *Mplus User's Guide.: Eighth Edition.* Los Angeles, CA: Muthén & Muthén.

Peugh, J., & Fan, X. (2012). How Well Does Growth Mixture Modeling Identify Heterogeneous Growth Trajectories? A Simulation Study Examining GMM's Performance Characteristics. *Structural Equation Modeling: a Multidisciplinary Journal*, *19*, 204–226. https://doi.org/10.1080/10705511.2012.659618

Ram, N., & Grimm, K. J. (2009). Growth Mixture Modeling: A Method for Identifying Differences in Longitudinal Change Among Unobserved Groups. *International Journal of Behavioral Development*, *33*, 565–576. https://doi.org/10.1177/0165025409343765

Titov, N., Dear, B. F., McMillan, D., Anderson, T., Zou, J., & Sunderland, M. (2011). Psychometric comparison of the PHQ-9 and BDI-II for measuring response during treatment of depression. *Cognitive Behaviour Therapy*, *40*, 126–136. https://doi.org/10.1080/16506073.2010.550059

APPENDICES FOR CHAPTER 5

Appendix 5.1 Scaffold for providing information about guidance in IBI

1. General information		
Type of Treatment	☐ CBT ☐ Psychodynamic ☐ Interpersonal ☐ Mindfulness-Based ☐ Cognitive Bias Modification ☐ Non-directive supportive ☐ Other →SPECIFY	
Target Group	☐ Specific mental disorders →SPECIFY ☐ Metal disorders, transdiagnostic →SPECIFY ☐ Specific somatic disorders →SPECIFY ☐ Other →SPECIFY	
Average Duration of Intervention	[ ] weeks	
Number of treatment modules/topics or sessions	☐ Fixed, [ ] modules/sessions ☐ Variable, [ ] modules/sessions on average	
Structure/ chronology of intervention	☐ Fixed ☐ Unstructured ☐ Adaptive, tailored or variable →SPECIFY	
Standardization of Treatment	☐ Completely individualized, with manual →CONTINUE WITH SECTION 2 ☐ Completely individualized, without manual →CONTINUE WITH SECTION 2 ☐ Completely standardized ☐ Semi-standardized, →please estimate the share of standardized treatment content	
Standardized Elements	☐ General psychoeducation ☐ Instructions on the use of therapeutic techniques ☐ Therapeutic tools/techniques ☐ Feedback on completed tasks ☐ Other →specify	
2. Automated Reminders		
Existence of automated Reminders	☐ No →CONTINUE WITH SECTION 3 ☐ Login-Reminders ☐ Reminders of new treatment content ☐ Other →SPECIFY	
Communication channel of reminders	☐ Written e-mails ☐ Messages sent to phone ☐ Other →SPECIFY	
3. Contact Before Treatment		
Contact before treatment	☐ No →CONTINUE WITH SECTION 4 ☐ Face-to-face ☐ Telephone ☐ Online, written ☐ Online, video-/audio	

	☐ Other →SPECIFY	
Purpose of contact before treatment	☐ Diagnostics, approx. [ ] minutes ☐ Technical/organizational instruct., approx. [ ] minutes ☐ Therapeutic, approx. [ ] minutes; →SPECIFY ☐ Other, approx. [ ] minutes; →SPECIFY	
4. Planned Contact During Treatment		
Planned Contact during treatment	☐ No →CONTINUE WITH SECTION 5 ☐ Yes	
Frequency of planned contact	☐ Singular ☐ Weekly ☐ Bi-weekly ☐ Other →SPECIFY	
Purpose of planned contact	☐ Technical/organizational assistance ☐ Encouragement/motivation ☐ Feedback on completed tasks ☐ Comprehensive therapeutic interaction comparable to face-to-face psychotherapy ☐ Other →SPECIFY	
Invested Time for providers of contact	☐ Approx. [ ] minutes per interaction/module ☐ Approx. [ ] minutes per patient (total)	
Fixed/stable provider for the entire treatment	☐ Yes ☐ No	
Qualifications of providers of contact	If diverse, please indicate share in brackets: ☐ Psychologists (B.Sc.) ☐ Psychologists (M.Sc.) ☐ Licensed psychotherapists ☐ Psychiatrists ☐ Primary care practicioners ☐ Technicians ☐ Study nurses ☐ Other →SPECIFY	
5. Contact-on-Demand During Treatment		
contact-on-demand during treatment	☐ No →END ☐ Yes, explicitly encouraged ☐ Yes, as existing option	
Communication channel of contact-on-demand	☐ Face-to-face ☐ Telephone ☐ Online, written ☐ Online, video-/audio ☐ Other →SPECIFY	
Purpose of contact-on-demand	☐ Encouragement ☐ Feedback ☐ Technical assistance ☐ Organizational assistance (e.g., pausing treatment) ☐ Issues with clinical relevance (e.g., instructions on how to improve work on therapeutic tasks) ☐ Other →SPECIFY	

APPENDIX 5.2 Excerpt from treatment manual

The following excerpt is taken from the unpublished treatment manual by Zagorscak et al. (2014; pp. 52-56):

Beispielhafter Klientinnenbrief nach Modul 1:

Zum ersten Mal so ein Tief wie im Moment hatte ich nach der Scheidung von meinem Mann vor 5 Jahren. Es lief ja schon lange nicht mehr zwischen uns, eigentlich haben wir uns nur die ganze Zeit gestritten. Wenn die Kinder nicht gewesen wären, wär es glaube ich schon viel früher auseinandergegangen, aber wir haben eben versucht, heile Welt zu spielen, obwohl es die nie gab. Und die Kinder haben das natürlich trotzdem mitbekommen. Ich mache mir Vorwürfe, dass ich es nicht geschafft habe, Ihnen eine funktionierende Familie zu bieten, vor allem die Große hat ganz schön viel abbekommen. Peter ist damals ausgezogen und hat uns sitzen lassen. Die erste Zeit war ich einfach nur wütend und konnte es nicht fassen. Ich bin dann mit den Kindern zu meinen Eltern, weil ich es in dem Haus nicht mehr ausgehalten habe. Ich glaube wenn meine Mutter nicht gewesen wäre, hätte ich die Zeit damals nicht überstanden. Aber damals war mein Vater auch schon ein Pflegefall und meine Mutter völlig überlastet. Deswegen sind wir dann nach 4 Monaten doch zurück ins Haus. Als die Scheidung dann endlich durch war, bin ich zusammengeklappt. Ich hab den ganzen Tag im Schlafzimmer gesessen und geweint und gegrübelt, mir hat alles wehgetan. An arbeiten war nicht zu denken und mein Hausarzt empfahl mir einen Klinikaufenthalt. Dazu konnte ich mich aber nicht aufraffen, irgendjemand musste sich ja auch noch um die Kinder kümmern, Peter hatte sie nur jedes zweite Wochenende. Alles in allem war ich dann einen Monat krankgeschrieben. Als ich dann wieder angefangen habe zu arbeiten, waren erst alle ganz verständnisvoll, schon fast mitleidig. Mittlerweile ist davon aber nichts mehr übriggeblieben. Seit ungefähr einem Jahr gibt es nur noch Ärger und Stress mit den Kollegen. Vor Weihnachten habe ich ein Gespräch mit meinem Chef gesucht und da sagte er mir ins Gesicht, dass ich schlampig arbeiten und mich nicht ins Team einbringen würde und wenn ich so weiter machen würde, könne ich mir einen neuen Job suchen. Ich weiß dass er mich nicht so einfach kündigen kann, aber er wartet nur auf den kleinsten Anlass, um mich loszuwerden. Auch von den Kollegen bekomme ich keinerlei Unterstützung, das macht mich total fertig. Was mir geholfen hat nach der Scheidung und auch heute noch, das sind allem meine Mutter und meine Kinder, ohne die wäre ich glaube ich überhaupt nicht mehr aufgestanden, aber ich musste eben funktionieren. Die meisten meiner sogenannten Freunde haben sich nach der Scheidung von mir abgewendet, eigentlich ist mir nur noch eine richtige Freundin geblieben, mit ihr kann ich gut sehr gut reden, aber sie hat auch ihre eigenen Probleme. Ich versuche auch, mich durch Sport abzulenken, mache Nordic Walking oder gehe spazieren. Aber in letzter Zeit schaffe ich das kaum noch, ich bin einfach zu erschöpft.

- Critical Life-Event: Scheidung
- Mögliche soziale Ressource: Kinder
- „Denkmuster": Schuldgefühle
- Mögliche Ressource: Mutter
- Möglicher Problembereich: Belastung durch Pflege
- Behaviorale/ Körperliche Probleme: Rückzug, Müdigkeit, Weinen, Schmerzen
- Möglicher Problembereich: Arbeitsumfeld
- Ressource in früherer Episode: Mutter und Kinder
- Mögliche Ressource: Freundin
- Angenehme Aktivität: Nordic Walking, Spaziergang

4.1.2 Ziele und Inhalte des Feedbacks zu Modul 1

Der erste Brief ist eine zentrale Informationsquelle für die Beraterin. Hier öffnet sich die Klientin erstmalig. Durch den Fokus auf die Vergangenheit können relevante Lebensereignisse und –krisen identifiziert werden, aber auch damals hilfreiche Bewältigungsstrategien und Ressourcen, die möglicherweise auch aktuell wieder genutzt werden können. Das zentrale Ziel Ihres ersten Feedbacks ist zunächst der Aufbau einer Beziehung zur Klientin, um eine tragfähige Basis für die Zusammenarbeit für die nächsten Wochen herzustellen. Daher stellt eine empathisch und wertschätzend formulierte Zusammenfassung der Problemlage der Klientin den Hauptteil des Feedbacks dar. Wenn möglich sollten Sie sich darum bemühen, das Stigma der Depression zu reduzieren. Dies ist z.B. möglich indem Sie betonen, dass Sie die depressive Verstimmung während der vergangenen Lebenssituation gut nachvollziehen können und dass es vielen Menschen nach kritischen Lebensereignissen ähnlich ergehen kann. Diese ‚Depathologisierung' ist insbesondere dann wichtig, wenn die Klientin Schuld- und Schamgefühle rund um Ihre Beschwerden äußert. Gleichzeitig ist zu betonen, dass Depressionen überwindbar sind - im Idealfall anhand eines von der Klientin geschilderten Beispiels. Berücksichtigen Sie daher insbesondere, was der Klientin aus ihrer Sicht in vergangenen depressiven Episoden bzw. schwierigen Zeiten geholfen hat. Auch für spätere Rückmeldungen ist der erste Brief eine „Fundgrube": Achten Sie auf Äußerungen zu persönlichen Bewältigungsstrategien, wichtigen aktuellen Bezugspersonen sowie auf die Schilderung von behavioralen (z.B. Rückzug, wenig Aktivität) und kognitiven (z.B. Sorgen, Schuld, Grübelei) Facetten der Depression. Damit können Sie Ihre Feedbacks individualisieren. Verweisen Sie nach Möglichkeit auf die von der Klientin diesem Zusammenhang benannten Ressourcen (z.B. aktive Bemühungen zur Änderung der eigenen Lage, Sport, Unterstützung durch Partner, Familie oder Freunde). Zu Stärkung der Motivation ist es weiterhin wichtig, der Klientin eine realistische Hoffnung und eine Perspektive für die kommenden Wochen der Programmteilnahme zu vermitteln.

Tipp: Notizen sparen Zeit

Schreiben Sie sich Aspekte, die Sie später aufgreifen könnten, in Stichpunkten in die Notiz-Funktion des TK-DC.

Ihre Ziele beim ersten Feedback im Überblick

- Beziehungsaufbau/ Empathie vermitteln
- Wertschätzung der Leidensgeschichte und der Bewältigungsversuche
- Depathologisieren
- Perspektive bieten: Depression ist episodisch und endlich
- frühere Strategien zur Linderung/Remission betonen (Selbstwirksamkeit)
- Ressourcen identifizieren und betonen

4.1.3. Textbaustein zu Modul 1

Vielen Dank für Ihren ersten Brief, in dem Sie mir sehr offen über [Zusammenfassung der Situation] berichtet haben. Ihre [bewegende/ausführliche] Beschreibung hat mir geholfen, einen Einblick in Ihre damalige Lebenssituation zu bekommen. Ich kann gut nachempfinden, dass Sie sich [in dieser Zeit // in schwierigen Zeiten in ihrem Leben] [EMPFINDUNGEN einfügen] gefühlt haben. Besonders deutlich wurde für mich, dass [Eingehen auf als besonders belastend wahrgenommene Elemente] [WICHTIG dabei: Empathie vermitteln; Wertschätzung des damaligen Bewältigungs-verhaltens].

[Stärkung der Selbstwirksamkeit/ individuelle Kraftquellen]
[falls keine Kraftquellen/Bewältigungsstrategien berichtet: gesonderter Textbaustein]

Tipp: Ein Schritt zurück

Anfangs begehen Beraterinnen häufig den Fehler, dass Sie Urteile darüber fällen, welche Aspekte der Schilderung der Klientin funktional bzw. dysfunktional sind.

Daher ist es wichtig, sich der Begrenztheit des eigenen Urteils zu besinnen: Ihr einziger Einblick in das Leben der Klientin ist der vorliegende Brief. Betrachten Sie die Klientin als Expertin, bleiben Sie zurückhaltend. Nur so können Sie sichergehen, dass Sie wirklich die Ziele der Klientin verfolgen und ihr nicht ungewollt Ihre eigene Perspektive „überstülpen".

Ich empfand es als sehr ermutigend, wie Sie von [indiv. Strategien zur Besserung der Beschwerden; Kraftquellen zusammenfassen] schrieben. Das zeigt mir, dass es Ihnen selbstständig [und auch mit der Unterstützung anderer Menschen] gelingen kann, aus Phasen schlechterer Stimmung auch wieder herauszufinden. Ich bin überzeugt, dass Sie das wieder schaffen können, auch wenn es Kraft kostet.

Für die nächsten Tage ist es unser Ziel, diese Erkenntnisse [aus der Vergangenheit] für Ihre aktuelle Lage zu nutzen. Dazu ist es mir wichtig, noch einmal genauer hinzuschauen, welche unterschiedlichen Lebensbereiche von Ihren aktuellen Beschwerden betroffen sind. Wie sie vielleicht schon gemerkt haben, können sich depressive Symptome auf ganz unterschiedliche Arten äußern.

Auf den nächsten Seiten habe ich dazu weitere Informationen und einen zweiten Schreibauftrag für Sie vorbereitet.

[Bevorzugte Abschiedsformel

Name der Beraterin/des Beraters]

4.1.4. Beispielhaftes Feedback zu Modul 1

Liebe Frau Schröder,

Vielen Dank für Ihren ersten Brief, in dem Sie mir sehr offen über die Zeit der Trennung von Ihrem Partner berichtet haben. Ihre ausführliche Beschreibung hat mir geholfen, einen Einblick in Ihre damalige Lebenssituation zu bekommen. Ich kann gut nachempfinden, dass Sie sich in dieser Zeit sehr belastet gefühlt haben. Die Schilderung Ihres Zusammenbruchs, des Weinens, Grübelns und der körperlichen Schmerzen nach Monaten des Kämpfens hat mich sehr bewegt. Besonders deutlich wurde für mich, dass es Ihnen sehr schwer gefallen sein muss, wieder in das gemeinsame Haus, mit all den Erinnerungen an die Ehe und Ihren ehemaligen Partner, zurückzuziehen. Dazu kam noch, dass Sie sich um Ihre Kinder gekümmert haben und deshalb den Ihnen nahegelegten Klinikaufenthalt nicht wahrnehmen konnten.

Fokussierte, empathische Zusammenfassung

Ich empfand es als sehr ermutigend, wie Sie von Ihrem Umfeld, vor allem Ihrer Mutter und den beiden Kindern, als Unterstützung und Kraftquelle schrieben. Nach einer Trennung ist es oft schwierig, mit gemeinsamen Freunden in Kontakt zu bleiben, umso mehr freut es mich, dass Sie auch weiterhin eine gute Freundin haben, die in schwierigen Zeiten zu Ihnen steht. Es ist auch ein gutes Zeichen, dass es Ihnen schon gelungen ist, sich durch Nordic Walking und Spaziergänge abzulenken. Das zeigt mir, dass es Ihnen selbstständig und auch mit der Unterstützung anderer Menschen gelingen kann, aus Phasen schlechterer Stimmung wieder herauszufinden. Ich bin überzeugt, dass Sie das wieder schaffen können, auch wenn es Kraft kostet.

Milde, wertschätzende Konfonrtation mit den „Kosten" der Aufopferung

Fokus auf Ressourcen, Verstärkung ihrer Inanspruchnahme

Für die nächsten Tage ist es unser Ziel, diese Erkenntnisse auf Ihre aktuelle Lage zu übertragen. Dazu ist es mir wichtig, noch einmal genauer zu klären, welche unterschiedlichen Lebensbereiche von Ihren aktuellen Beschwerden betroffen sind. Wie Sie vielleicht schon gemerkt haben, können sich depressive Symptome auf ganz unterschiedliche Arten äußern.

Betonung des Erfolgs, Stärkung der Selbstwirksamkeits-erwartung

Auf den nächsten Seiten habe ich dazu weitere Informationen und einen neuen Schreibauftrag für Sie vorbereitet.

Überleitung zum nächsten Modul

Es grüßt Sie herzlich, Beraterin

LIST OF OWN PUBLICATIONS

BOOKS AND BOOK CHAPTERS

Schultze-Krumbholz, A., Zagorscak, P., Roosen-Runge, A. & Scheithauer H. (2018, 2nd ed.). *Medienhelden. Unterrichtsmaterialien zur Prävention von Cybermobbing.* München: Reinhardt.

Zagorscak P. & Knaevelsrud C. (2019, 2nd ed.). Online-Therapie. In S. Schneider & J. Margraf (Eds.), *Lehrbuch der Verhaltenstherapie. Band 3: Störungen im Kindes- und Jugendalter* (pp. 233-247). Berlin: Springer.

Schultze-Krumbholz, A., Zagorscak, P., & Scheithauer, H. (2017). A school-based cyberbullying preventive intervention approach: The Media Heroes program. In M. Campbell & S. Bauman (Eds.), *Reducing cyberbullying in schools: International evidence-based best practice* (pp. 145-158). San Diego, CA/Cambridge, MA, USA: Academic Press. doi: 10.1016/B978-0-12-811423-0.00011-0

Knaevelsrud, C., Küster, A., & Zagorscak, P. (2016). E-Mental-Health. In: I. Hauth, P. Falkai & A. Deister (Eds.) *Psyche Mensch Gesellschaft* (pp. 25-30). Berlin: Springer.

Renneberg, B. & Zagorscak, P. (2016). Sexuelle Orientierung aus psychologischer und medizinischer Sicht. In C. D. Classen, D. Richter & B. Łukańko (Eds.) *„Sexuelle Orientierung" als Diskriminierungsgrund: Regelungsbedarf in Deutschland und Polen?* (pp. 1-12). Tübingen: Mohr Siebeck.

Zagorscak P., Schultze-Krumbholz, A., & Scheithauer H. (2015). Cybermobbing. In Deutsche Gesetzliche Unfallversicherung (DGUV) (Ed.), *Taschenbuch 2015 -Sicherheitsbeauftragte im öffentlichen Dienst* (pp. 154-161). Wiesbaden: Universum. ISSN 0942-0274

Jäkel, A., Schultze-Krumbholz, A., Wölfer, R., Zagorscak, P. & Scheithauer, H. (2013). Il progetto Medienhelden: prevenzione del cyberbullismo e promozione delle competenze mediatiche nella sculoa, In: M.L. Genta, A. Brighi & A. Guarini (Eds.) *Cyberbullismo. Ricerche e strategie di intervento* (pp. 163-170). Rome (Italy): Franco Angeli..

Zagorscak, P. (2013). Medienhelden. In Landeskommission Berlin gegen Gewalt (Ed.), *Schöne neue Welt - total vernetzt! Fluch oder Segen?* (pp. 66–68). Berlin: Herausgeber.

Schultze-Krumbholz, A., Zagorscak, P., Siebenbrock, A. & Scheithauer H. (2012). *Medienhelden. Unterrichtsmaterialien zur Prävention von Cybermobbing.* München: Reinhardt.

PUBLICATIONS IN PEER-REVIEWED JOURNALS

Zagorscak, P., Heinrich, M., Sommer, D., Wagner, B. & Knaevelsrud, C. (2018). Reply to the Letter by Singer, Mitter, and Porsch Related to Our Paper "Benefits of Individualized Feedback in Internet-Based Interventions for Depression: A Randomized Controlled Trial". *Psychotherapy and Psychosomatics*, 1–2. doi:10.1159/000493394

Heinrich, M., Zagorscak, P., Eid, M. & Knaevelsrud, C. (2018). Giving G a Meaning: An Application of the Bifactor-(S-1) Approach to Realize a More Symptom-Oriented Modeling of the Beck Depression Inventory-II. *Assessment*. Advance online publication. doi: 10.1177/1073191118803738

Zagorscak, P., Schultze-Krumbholz, A., Heinrich, M., Wölfer, R. & Scheithauer, H. (2018). Efficacy of Cyberbullying Prevention on Somatic Symptoms—Randomized Controlled Trial

Applying a Reasoned Action Approach. *Journal of Research on Adolescence.* Advance online publication. doi: 10.1111/jora.12429

Zagorscak, P., Heinrich, M., Sommer, D., Wagner, B. & Knaevelsrud, C. (2018). Benefits of Individualized Feedback in Internet-Based Interventions for Depression: A Randomized Controlled Trial. *Psychotherapy and Psychosomatics, 87* (1), 32–45. doi:10.1159/000481515

Schultze-Krumbholz, A., Schultze, M., Zagorscak, P., Wölfer, R., & Scheithauer, H. (2016). Feeling cybervictims' pain - The effect of empathy training on cyberbullying. *Aggressive Behavior, 42*, 147-156. doi: 10.1002/ab.21613

Schultze-Krumbholz, A., Zagorscak, P., Wölfer, R., & Scheithauer, H. (2014). Das Medienhelden-Programm zur Förderung von Medienkompetenz und Prävention von Cybermobbing: Konzept und Ergebnisse aus der Evaluation. *Praxis der Kinderpsychologie und Kinderpsychiatrie, 63*(5), 379-394. doi: 10.13109/prkk.2014.63.5.379

Schultze-Krumbholz, A., Zagorscak, P., Wölfer, R., & Scheithauer, H. (2014). Prävention von Cybermobbing und Reduzierung aggressiven Verhaltens Jugendlicher durch das Programm Medienhelden: Ergebnisse einer Evaluationsstudie. *Diskurs Kindheits- und Jugendforschung, 9*(1), 61–79. doi: 10.3224/diskurs.v9i1.19083

Wölfer, R., Schultze-Krumbholz, A., Zagorscak, P., Jäkel, A., Göbel, K., & Scheithauer, H. (2014). Prevention 2.0: Targeting cyberbullying @ school. *Prevention Science, 15*, 879-887. doi: 10.1007/s11121-013-0438-y

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorgelegte Arbeit selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel verwendet habe. Die Arbeit ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Berlin, Januar 2019

Pavle Zagorscak