An understudied species: Understanding therapists and their impact on psychotherapy

Inauguraldissertation zur Erlangung der Doktorwürde (Dr. rer. nat.) im Fach Psychologie, Fachbereich I an der Universität Trier

Vorgelegt im

Februar 2019

von

Dirk Zimmermann

Gutachter:

Prof. Dr. Wolfgang Lutz

Prof. Dr. Julian A. Rubel

Dissertationsort: Trier

DANKSAGUNG

Mein Dank gilt meinem Doktorvater Prof. Dr. Wolfgang Lutz, der mich in meiner fachlichen Entwicklung unterstützt und sein Vertrauen in mich gesetzt hat. Seine Tür stand immer offen, wenn es Probleme oder Fragen gab. Weiterhin war die Ermöglichung der vielen nationalen sowie internationalen Kongressen äußerst hilfreich, um meinen Horizont zu erweitern und Ideen für die Forschung zu generieren.

Darüber hinaus möchte ich Prof. Dr. Julian Rubel danken, der mich vom Beginn meiner wissenschaftlichen Karriere von der wissenschaftlichen Hilfskraft bis zur Promotion immer fachlich unterstützt hat. Nicht zuletzt ist aus einem fachlichen Mentor sogar ein guter Freund geworden.

Ein weiterer Dank geht an Dr. Nadine Kasten. Bereits im Bachelorstudiengang war sie Mentor in methodischen Fragen und Betreuerin meiner Bachelorarbeit. In der Disputation unterstützt sie mich als Protokollantin. Aus der wissenschaftlichen Begleitung ist darüber hinaus eine sehr gute Freundin geworden.

Mein Dank gilt auch allen Kollegen, die mich auf dem Entstehungsweg dieser Arbeit begleitet haben. Es hat mir viel Spaß gemacht, in den unterschiedlichsten Projekten mit Euch zu arbeiten und auch in den Pausen viele tolle und bereichernde Gespräche mit euch zu führen.

Außerdem möchte ich meiner Familie und meinen Freunden danken, die mir dabei geholfen haben den Blick auf das Wesentliche im Leben nicht zu verlieren.

Ein ganz besonderer Dank gilt meiner Frau Lisa und meiner Tochter Laila. Lisa, dir danke ich, für die emotionale Unterstützung über alle Jahre, den Glauben an mich und die Kraft, die du mir gibst. Ihr beide seid für mich der größte Antrieb – jeden Tag!

Table of contents

Abstract

A huge number of clinical studies and meta-analyses have shown that psychotherapy is effective on average. However, not every patient profits from psychotherapy and some patients even deteriorate in treatment. Due to this result and the restricted generalization of clinical studies to clinical practice, a more patient-focused research strategy has emerged. The question whether a particular treatment works for an individual case is the focus of this paradigm. The use of repeated assessments and the feedback of this information to therapists is a major ingredient of patient-focused research. Improving patient outcomes and reducing dropout rates by the use of psychometric feedback seems to be a promising path. Therapists seem to differ in the degree to which they make use of and profit from such feedback systems. This dissertation aims to better understand therapist differences in the context of patient-focused research and the impact of therapists on psychotherapy. Three different studies are included, which focus on different aspects within the field:

Study I (Chapter 5) investigated how therapists use psychometric feedback in their work with patients and how much therapists differ in their usage. Data from 72 therapists treating 648 patients were analyzed. It could be shown that therapists used the psychometric feedback for most of their patients. Substantial variance in the use of feedback (between 27% and 52%) was attributable to therapists. Therapists were more likely to use feedback when they reported being satisfied with the graphical information they received. The results therefore indicated that not only patient characteristics or treatment progress affected the use of feedback.

Study II (Chapter 6) picked up on the idea of analyzing systematic differences in therapists and applied it to the criterion of premature treatment termination (dropout). To answer the question whether therapist effects occur in terms of patients' dropout rates, data from 707 patients treated by 66 therapists were investigated. It was shown that approximately six percent of variance in dropout rates could be attributed to therapists, even when initial impairment was controlled for. Other predictors of dropout were initial impairment, sex, education, personality styles, and treatment expectations.

Study III (Chapter 7) extends the dissertation by investigating the impact of a transfer from one therapist to another within ongoing treatments. Data from 124 patients who agreed to and experienced a transfer during their treatment were analyzed. A significant drop in patient-rated as well as therapist-rated alliance levels could be observed after a transfer. On average, there seemed to be no difficulties establishing a good therapeutic alliance with the new therapist, although differences between patients were observed. There was no increase in symptom

severity due to therapy transfer. Various predictors of alliance and symptom development after transfer were investigated. Impacts on clinical practice were discussed.

Results of the three studies are discussed and general conclusions are drawn. Implications for future research as well as their utility for clinical practice and decision-making are presented.

List of all authored and co-authored publications

Zimmermann, D., Lutz, W., Reiser, M., Boyle, K., Schwartz, B., Schilling, V., Deisenhofer, A.-K., & Rubel, J. (2019). What happens when the therapist leaves? The impact of therapy transfer on the therapeutic alliance and symptoms. *Clinical Psychology & Psychotherapy, 26*(1), 135-145.

Schilling, V.N.L.S, Lutz, W., Hofmann, S.G., Zimmermann, D., Wolter, K., & Stangier, U. (in press). Loving Kindness Meditation zur Behandlung der chronischen Depression. *Zeitschrift für Klinische Psychologie und Psychotherapie.*

Müller, V. N., Boyle, K., Zimmermann, D., Weinmann-Lutz, B., Rubel, J. A., & Lutz, W. (2018). What is individually tailored mental health care? *Revista Argentina de Clínica Psicológica, 27(2*), 157-181.

Deisenhofer, A. K., Delgadillo, J., Rubel, J. A., Böhnke, J. R., Zimmermann, D., Schwartz, B., & Lutz, W. (2018). Individual treatment selection for patients with posttraumatic stress disorder. *Depression and anxiety, 35(6*), 541-550.

Rubel, J. A., Zimmermann, D., Müller, V., & Lutz, W. (2017). Qualitätssicherung in der Psychotherapie. *PPmP-Psychotherapie· Psychosomatik· Medizinische Psychologie, 67(09/10)*, 436-448.

Lutz, W., Zimmermann, D., Müller, V. N., Deisenhofer, A. K., & Rubel, J. A. (2017). Randomized controlled trial to evaluate the effects of personalized prediction and adaptation tools on treatment outcome in outpatient psychotherapy: study protocol. *BMC psychiatry, 17(1)*, 306.

Rubel, J. A., Zimmermann, D., Deisenhofer, A. K., Müller, V., & Lutz, W. (2017). Nutzung von psychometrischem Feedback als empirische Unterstützung des Supervisionsprozesses bei Ausbildungstherapien. *Zeitschrift für Klinische Psychologie und Psychotherapie, 46(2),* 83-95.

Zimmermann, D., Rubel, J., Page, A. C., & Lutz, W. (2017). Therapist effects on and predictors of non-consensual dropout in psychotherapy. *Clinical (Zimmermann, Rubel, Page, & Lutz, 2017) Psychology & Psychotherapy, 24(2),* 312-321.

Lutz, W., Rubel, J., Schiefele, A. K., Zimmermann, D., Böhnke, J. R., & Wittmann, W. W. (2015). Feedback and therapist effects in the context of treatment outcome and treatment length. *Psychotherapy Research*, *25*(6), 647-660.

Rubel, J., Lutz, W., Kopta, S. M., Köck, K., Minami, T., Zimmermann, D., & Saunders, S. M. (2015). Defining early positive response to psychotherapy: An empirical comparison between clinically significant change criteria and growth mixture modeling. *Psychological assessment*, *27*(2), 478-488.

List of publications included in the cumulative dissertation

1. Study I

   Rubel, J. A., Zimmermann, D., Deisenhofer, A. K., Müller, V., & Lutz, W. (2017). Nutzung von psychometrischem Feedback als empirische Unterstützung des Supervisionsprozesses bei Ausbildungstherapien. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *46*(2), 83-95. doi: 10.1026/1616-3443/a000413

2. Study II

   Zimmermann, D., Rubel, J., Page, A. C., & Lutz, W. (2017). Therapist Effects on and Predictors of Non-Consensual Dropout in Psychotherapy. *Clinical Psychology & Psychotherapy*, *24*(2), 312-321. doi: 10.1002/cpp.2022

3. Study III

   Zimmermann, D., Lutz, W., Reiser, M., Boyle, K., Schwartz, B., Schilling, V., Deisenhofer, A.-K., & Rubel, J. (2019). What happens when the therapist leaves? The impact of therapy transfer on the therapeutic alliance and symptoms. *Clinical Psychology & Psychotherapy*, *26*(1), 135-145. doi: 10.1002/cpp.2336

1. Introduction

Psychotherapy is a well-established field in many health care systems around the world and a great many research studies have been published in the last decades (Lambert, 2013a). A huge effort has been undertaken to prove the positive effects of psychotherapeutic treatments (Barth et al., 2013; Cuijpers, van Straten, Andersson, & van Oppen, 2008; Lambert, 2013b). Different psychotherapy approaches help for most people for many different mental disorders (Lambert, 2013b). However, not every patient profits from treatment, even if a highly efficacious treatment was delivered. This is one reason for the development of patient-focused research (Castonguay, Barkham, Lutz, & McAleavey, 2013). This research strategy focuses on the individual patient and his progress in an ongoing treatment (Lutz, Jong, & Rubel, 2015). Repeated assessments of patients' impairment with psychometric questionnaires and feedback of this information to therapists is an attempt to reduce the number of patients not profiting from therapy (Lambert et al., 2003). Therapists are often unaware of patients' negative developments (Hannan et al., 2005; Hatfield, McCullough, Frantz, & Krieger, 2010). Feedback systems are a promising tool to enhance therapists' awareness of patients' developments and a way to improve clinical decision-making (Castonguay et al., 2013). The use of psychometric feedback seems to improve clinical outcomes especially in patients not developing well in treatment (Knaup, Koesters, Schoefer, Becker, & Puschner, 2009; Shimokawa, Lambert, & Smart, 2010). However, therapists seem to differ in the degree to which they profit from such feedback systems (De Jong, van Sluis, Nugter, Heiser, & Spinhoven, 2012). Understanding differences in therapists and their impact on therapy has long been of less interest than proving the positive effects of therapy (Baldwin & Imel, 2013; Barkham, Lutz, Lambert, & Saxon, 2017). Understanding therapist differences in the context of patient-focused research is the main objective of this dissertation. Therefore, this dissertation summarizes three studies that were designed to fill research gaps in the context of outpatient psychotherapy.

Study I (Chapter 5) investigated how therapists use psychometric feedback in their work with patients and how much therapists differ in their usage. Data from 72 therapists treating 648 patients were analyzed. It could be shown that therapists used the psychometric feedback for most of their patients. Substantial variance in the use of feedback (between 27% and 52%) was attributable to therapists. Therapists were more likely to use feedback when they reported being satisfied with the graphical information they received. The results therefore indicated that not only patient characteristics or treatment progress affected the use of feedback.

Study II (Chapter 6) picked up on the idea of analyzing systematic differences in therapists and applied it to the criterion of premature treatment termination (dropout). To answer the question whether therapist effects occur in terms of patients' dropout rates data from 707 patients treated by 66 therapists were investigated. It was shown that approximately six percent of variance in dropout rates could be attributed to therapists, even when initial impairment was controlled for. Other predictors of dropout were initial impairment, sex, education, personality styles, and treatment expectations.

Study III (Chapter 7) extends the dissertation by investigating the impact of a transfer from one therapist to another within ongoing treatments. Data from 124 patients who agreed to and experienced a transfer during their treatment were analyzed. A significant drop in patient-rated as well as therapist-rated alliance levels could be observed after a transfer. On average, there seemed to be no difficulties establishing a good therapeutic alliance with the new therapist, although differences between patients were observed. There was no increase in symptom severity due to therapy transfer. Various predictors of alliance and symptom development after transfer were investigated. Impacts on clinical practice were discussed.

All three studies are depicted in chapters five to seven. Chapter two provides a common theoretical background, which was the basis of the studies. This leads to the deduction of the research questions in chapter three. To facilitate the understanding and interpretation of the studies, chapter four provides the reader with the methodological specialties all studies have in common. Finally, a general discussion of the studies is presented in chapter 8. Future research areas and practical implications are considered.

2. Theoretical background

2.1 Psychotherapy

One can assume that psychological interventions have long been practiced in human history (Jackson, 1999). Medics, spiritual healers, philosophers, and ordinary people have used psychological methods to promote healing in others. A popular healer in the 18th and early 19th century was Franz Mesmer (Kiesewetter, 2010). He developed his animal magnetism to cure hysteria. In 1774, he instructed his patient Franziska Österlin swallow a prepared object containing iron and put magnets on various spots on her body. She felt mysterious fluid running through her body and reported relief of her hysterical symptoms. Although Mesmer struggled to get accepted within the scientific community throughout his career, his work influenced the development of hypnosis and the treatment of mental illness (Jackson, 1999). The term "psycho-therapeutics" was firstly mentioned by Daniel Hack Tuke in 1872 in his work "Illustrations of the Influence of the Mind upon the Body in Health and Disease designed to elucidate the Action of the Imagination" (as cited in Shamdasani, 2005). In 1896, the first scientific journal had the term psychotherapy in its name (Shamdasani, 2005). Sigmund Freud greatly influenced modern psychotherapy with his development of psychoanalysis (Freedheim et al., 1992). Freud started using hypnosis to get access to unconscious parts of his patients' minds to cure them of neurotic symptoms. The discovery of patients' unconscious motives and wishes was seen as the key element to symptom relief. Freud later stopped using hypnosis and worked with free association and the interpretation of dreams. Freud had a strong vision and was convinced that his theories and treatment approach were correct and the only option. His psychoanalysis can be seen as the first school of psychotherapy. Freud had many students who adapted his treatment approach and turned away from original psychoanalysis. Alfred Adler developed his individual psychotherapy, Carl Jung designed his analytic psychology, Melanie Klein her object relations theory, just to name a few. The growing field of analytic therapies led to many different therapeutic schools (Freedheim et al., 1992). Alongside the growth of analytic therapy, behaviorism developed in the 1920s. Behavior and feelings were explained with learning principles. The application of classical conditioning, operant conditioning and social learning theory led to behavioral therapy in the 1950s and 1960s (Lambert, 2013a). Phobias and other mental disorders were treated with this approach. Among the foremost of the popular representatives were Joseph Wolpe, Burrhus Frederick Skinner, and Hans Eysenck. With Albert Bandura, Aaron Beck, and Albert Ellis came a shift and focus on cognitive processes within treatment. Maladaptive thoughts and beliefs were seen as causal for the

development of mental disorders such as depression (Freedheim et al., 1992). During the 1980s and 1990s both behavioral and cognitive methods were combined and cognitive behavioral therapy was born. Besides the development of psychoanalytic therapy, behavioral therapy and all of their offspring, other streams such as family therapy or humanistic approaches have been developed (Freedheim et al., 1992). A comprehensive description of the history of psychotherapy is beyond the scope of this dissertation. Based on the degree of abstraction, an estimated 250 to 600 different treatment approaches exist today (Goldfried & Wolfe, 1996; Kazdin, 2000; Wampold, 2001). For this dissertation a rather broad definition of psychotherapy will be used:

> "Psychotherapy is a primarily interpersonal treatment that is based on psychological principles and involves a trained therapist and a client who has a mental disorder, problem, or complaint; it is intended by the therapist to be remedial for the client's disorder, problem, or complaint; and it is adapted or individualized for the particular client and his or her disorder, problem, or complaint." (Wampold, 2001, p. 2-3)

This definition includes most therapeutic approaches and does not rely on a specific therapeutic theory. However, treatments without an interpersonal focus such as bibliotherapy or internet therapy are excluded.

2.2 Efficacy and Effectiveness

In the early 1950s, Eysenck (1952) received major attention for his conclusion that there is no support that psychotherapy facilitates recovery from neurotic disorders. Although, the paper was criticized by many other researchers, the provocative questioning of the positive effects of psychotherapy led to a tremendous increase of research within the field of psychotherapy (Lambert, 2013a). A relatively new field had to justify its existence. More and more studies were conducted to underpin the hypothesis that psychotherapy helps people to overcome mental disorders and improve personal well-being. Many of the studies were randomized controlled trials (RCTs), which assure strong internal validity. The first big meta-analysis by Smith and Glass (1977) included 375 RCTs in which psychotherapy was compared to an untreated or different therapy group. They concluded that the typical patient is better off than 75% of untreated individuals on average. In the last decades, hundreds of studies have been carried out, including different treatment approaches in the treatment of various disorders. Meta-analyses show strong evidence for the positive effects of psychotherapy (Barth et al., 2013; Cuijpers et al., 2008; Cuijpers et al., 2011; Leichsenring & Leibing, 2003). Also, a combined treatment with pharmacotherapy seems to be superior to pharmacological treatment alone for various

disorders (Cuijpers et al., 2014). There is evidence that psychotherapy has a prophylactic effect and that it reveals its superiority in comparison to medications especially in long-term comparisons (Hollon, Stewart, & Strunk, 2006; Imel, Malterer, McKay, & Wampold, 2008).

To date, the gold standard of research within the field of psychotherapy is the RCT, which is known as efficacy research (Lutz & Böhnke, 2010). Guidelines that inform policy makers and practitioners which treatment to offer often rely on RCTs (Bandelow, Lichte, Rudolf, Wiltink, & Beutel, 2014; DGPPN et al., 2009; National Institute for Clinical Excellence, 2009). To get the highest evidence grade within a guideline, different independent RCTs are necessary. However, the focus on RCTs has been criticized for various reasons. One major issue concerns external validity (Rothwell, 2005). Results from RCTs cannot simply be generalized to real-world settings. RCTs are often conducted with specifically trained therapists in sites with a strong research perspective. Therapists often receive additional supervision and strict in- and exclusion criteria for patients results in a selective patient population (Fairburn et al., 2015; Mulder, Boden, Carter, Luty, & Joyce, 2017; Storch et al., 2015). In order to overcome this problem, an increasing number of clinical trials have been carried out under routine care conditions (Lambert, 2013b; Lambert & Ogles, 2004). This research approach is known as effectiveness research (Lutz & Böhnke, 2010). Although, the efficacy and effectiveness of psychotherapy has been demonstrated and effect sizes are comparable to pharmacological treatments, at least two issues remain:

First, clinical trials provide evidence on which treatment helps for which disorder under certain circumstances on average. Neither a single nor a hundred RCTs can answer the question whether a specific treatment helps for the individual case. However, as treatment is not helpful for every patient, knowledge about how a particular treatment works for the individual case seems rather important.

Second, the need for standardization in clinical trials narrows the focus to treatment ingredients. However, as psychotherapy is an interpersonal exchange between therapist and patient, the individual characteristics of the therapist have long been neglected. However, therapist differences and how patients and therapists match (or do not) may be partly responsible for the many cases that do not respond to psychotherapy.

The following two sections tackle these issues. The subsequent section describes the shift from focusing on treatment to focusing on the individual patient within the context of patient-focused research. The last section of the introduction deals with the therapists and what we know about their impact on therapy.

2.3 Patient-focused research

About 40% to 70% of patients benefit in RCTs, whereas the effects are lower in effectiveness studies (Lambert, 2013b). In other words, 30% to 60% of patients do not profit from psychotherapy. Approximately 5% to 10% of patients even deteriorate over the course of treatment (Hansen, Lambert, & Forman, 2002; Lambert, 2013b). These numbers dramatically underpin that a focus on specific treatments or treatment ingredients is short-sighted. In this context, Howard and colleagues proposed patient-focused research as an idiographic research strategy to overcome the above-mentioned shortcomings of treatment-focused research (Howard, Moras, Brill, Martinovich, & Lutz, 1996; Lutz, 2002). Patient-focused research is a paradigm that attempts to close the scientist-practitioner gap by implementing scientific knowledge in clinicians' everyday practice. One main goal is to improve the individual patient's outcome by tracking the patient's progress over the course of treatment and feeding this information back to the clinician. This approach is comparable to physicians doing lab testing and taking vital signs to manage physical diseases such as diabetes (Lambert, 2010). Based on large data sets and studying patterns of patient change, predictions can be derived for the individual patient. A prediction for the individual case is therefore based on datasets from former studies or data from routine care (Lutz, Zimmermann, Müller, Deisenhofer, & Rubel, 2017). Thus, aggregated patient data is used to model the expected recovery curve for the individual patient. Deviations from the expected recovery curve help to identify whether a patient is still making the expected progress in treatment. On this basis, clinicians are provided with tools to support their clinical decisions with actual ongoing research data (Castonguay et al., 2013). These tools can help identify patients who are at risk for treatment failure and help clinicians to prevent negative developments in treatment.

Therapists overestimate the progress patients make and are often not able to detect when a patient is deteriorating (Hannan et al., 2005; Hatfield et al., 2010). These findings emphasize the need to track each individual case and help clinicians identify which patient is not on track. The positive effects of the usage of continuous psychometric feedback to therapists have been documented in a number of studies (Lambert et al., 2001; Reese, Norsworthy, & Rowlands, 2009) and meta-analyses (Lambert & Shimokawa, 2011; Lambert, Whipple, & Kleinstäuber, 2018; Østergård, Randa, & Hougaard, 2018). It has been shown that the positive effects are often larger for patients with a risk of treatment failure (Lambert & Shimokawa, 2011). The most recent meta-analysis found a small to medium overall effect in favor of the feedback condition (Østergård et al., 2018). Lambert and colleagues (2018) reported small to medium effect sizes for patients with a risk of treatment failure. Another recent meta-analysis found no

positive effects for the usage of psychometric feedback in general, but reported a small effect for patients with a risk for treatment failure (Kendrick et al., 2016). However, the authors concluded that the included studies were of low quality and that no solid assumption could currently be drawn.

Considering the mixed results regarding feedback effects, more effort has been made to understand how and when feedback works. Therefore, studies on mediators and moderators have been carried out. Some of the early feedback studies found a relation between feedback and therapy length. Patients with a high risk for treatment failure stayed in treatment longer, while cases that developed well had shorter therapies (Lambert et al., 2003). In the same vein, a more recent study in a naturalistic setting found that patients in the feedback condition had fewer therapy sessions (Delgadillo et al., 2017). Shimokawa and colleagues (2010) found that this result is not consistent throughout the literature. However, a more recent meta-analysis again underpinned the finding that patients in the feedback condition seem to have fewer sessions (Kendrick et al., 2016). More research is needed to understand when feedback has an effect on treatment lengths and how it affects treatment outcomes.

There also seem to be differences in the way therapists deal with feedback. It could be shown that only 50% of therapists were able to use feedback in a way that could improve outcome (Simon, Lambert, Harris, Busath, & Vazquez, 2012). Similarly, De Jong and colleagues (2012) found that therapists differ in the degree to which they profited from a feedback intervention. Therapists with a higher commitment to feedback as well as female therapists used feedback more often. The frequency of the usage of feedback was related to better outcomes for patients with a risk for treatment failure (De Jong et al., 2012). Similarly, Lutz et al. (2015) found that therapists' attitudes toward feedback were related to treatment outcome. Another moderator of the feedback effect may be the setting where treatment takes place. The feedback effect has been found to be stronger in counseling settings compared to psychiatric settings (Østergård et al., 2018). Research on feedback effects provides mixed results and knowledge on mediators and moderators is limited. Therefore, further studies are needed to identify the underlying processes.

2.4 Therapists

The focus on RCTs and on the comparison between different therapy approaches lead to a neglect of the therapist. However, therapists do not only differ in terms of clinical training or the theoretical approach they adhere to. They differ in age, sex, race, education, therapeutic styles, well-being, attitudes, beliefs, expectations, world view, idea of man, individual caseload

and so on and so forth. There is growing evidence that therapists perform differently, which, in turn, produces varying patient outcomes. Early studies have shown that some therapists had better outcomes than others, irrespective of the therapeutic approach, and that some therapists even produce consistently negative outcomes (Lafferty, Beutler, & Crago, 1989; Lambert, 1989; Luborsky et al., 1986). Crits-Christoph and colleagues (1991) published the first meta-analysis that investigated different factors, which might lead to differences in therapist efficacy. They concluded that the use of a treatment manual as well as more experience lead to smaller therapist differences. The finding that the use of treatment manuals seems to reduce the variability between therapists is not surprising and reflects the focus of treatment-focused research (see above). Studies are designed to reduce variability between both patients and therapists. However, ignoring the therapist variable when analyzing differences between treatments can lead to flawed conclusions (Lutz & Barkham, 2015). A given difference between treatment A and treatment B could simply reflect the fact that therapists performed differently in both treatments. Even in studies in which the same therapists deliver different treatments, one cannot exclude the influence of therapists' allegiance, which might explain outcome differences (Falkenström, Markowitz, Jonker, Philips, & Holmqvist, 2013).

Past editions of the Handbook of Psychotherapy and Behavior Change reviewed the relation between therapist characteristics and patient outcomes. Features of the therapist have been organized into four quadrants: (1) observable traits (e.g., age, ethnicity), (2) observable states (e.g., professional background, therapist interventions), (3) inferred traits (e.g., personality style, attitudes), and (4) inferred states (e.g., therapeutic relationship, expectancies) (Beutler et al., 2004; Beutler, Machado, & Allstetter Neufeldt, 1994). Although dozens of studies that analyze variables within the four quadrants have been reviewed, no clear picture on these variables and their relationships has emerged. One example of unclear results pertains to research on therapist age and age similarity between therapist and patient. One study found that clients within the same age range as their therapists reported less distress, less social isolation, and less public system dependency after treatment compared to clients who had older therapists (Dembo, Iklé, & Ciarlo, 1983). In contrast, a review analyzing similarities in counselling found no studies that reported a positive relation between age similarity and outcome (Atkinson & Schein, 1986). Beutler and colleagues (2004) concluded that there is little research to suggest that therapist age or age similarity contributes significantly to treatment outcome. One reason for the mixed results could be the complex interactions between different therapist characteristics with client features, the setting, or the type of treatment (Beutler et al., 1994).

The latest edition of the Handbook of Psychotherapy and Behavior Change made a shift from reviewing relationships within the four quadrants to a focus on the differential effectiveness of therapists (Baldwin & Imel, 2013). They defined therapist effects as "the effect of a given therapist on patient outcomes as compared to another therapist" (Baldwin & Imel, 2013, p. 259-260). For this dissertation the term therapist effects will be extended to also include effects on dropout: The effect of a given therapist on patient outcomes or patient dropouts as compared to another therapist. Advances in statistical methodologies and computing performances have made it possible to more easily analyze the systematic impact of therapists on patient outcomes. Baldwin & Imel (2013) found an average therapist effect of 5% of outcome variance based on 46 studies with 1,281 therapists and 14,519 patients. Large differences have been found between the underlying single studies. Some studies reported therapist effects equal to 0%, while others reported therapist effects of up to 55%. Another recent study analyzing eight different datasets with a total of 48,648 patients treated by 1,800 therapists revealed therapist effects between 2.7% and 10.2%, with an average of 5.7% (Schiefele et al., 2017), which is comparable to the results found by Baldwin and Imel (2013). Therapist effects have been found to be larger in naturalistic settings (approximately 7%) than in RCTs. (approximately 3%), which is consistent with findings from earlier studies (Baldwin & Imel, 2013; Crits-Christoph et al., 1991). Although a growing body of research suggests that therapists differ in their effectiveness, little is known about the factors that might explain therapist differences. Research on therapist effects has mainly looked at patient outcomes, whereas therapist differences in dropout rates have hardly been studied to date.

Another issue in the field of therapist research is the transfer from one therapist to another within an ongoing treatment due to reasons of parental leave, turnover, sick leave, and so on. This phenomena is common in naturalistic settings and rarely the case in RCTs, which still dominate the research field today. Therapist transfers might have negative effects on patient outcomes and/or dropout. Empirical evidence on how therapist transfers affect patients is sparse (Sauer, Rice, Richardson, & Roberts, 2017).

3. Research questions

Section 2.3. focused on the research strategy patient-focused research and gave a short summary of important findings within this field. One major trend in patient-focused research is psychometric feedback to therapists, which is meant to improve clinical decision-making and patient outcomes. Section 2.4. focused on therapists and summarized findings of therapists' differential effectiveness. There are many unanswered questions in both research areas, which this dissertation tries partly to overcome.

Not much is known about the working mechanisms of psychometric feedback and their use in clinical practice. Adding to that, there is not much evidence as to whether therapists differ in their use of feedback. Study I combined both research on therapists and research on feedback. The following research questions were addressed in study I:

Study I

1. How do therapists use psychometric feedback for their patients?
2. Are there differences in the use of feedback between therapists and if yes, to what extent?

As pointed out in the previous chapter, there is growing evidence pointing to differential therapist effectiveness with regard to patient outcomes. However, there is little evidence on how dropouts are affected by the means of therapist differences. Therefore, study II focused on that understudied topic by addressing the following research questions:

Study II

1. Are there therapist effects on dropout rates and if yes, to what extent?
2. Which variables can predict dropout in outpatient psychotherapy?

The previous chapter described how therapists have long been neglected as the majority of past research has been treatment-focused. A common clinical phenomena such as a transfer from one therapist to another within an ongoing treatment has rarely been studied. Study III attempted to understand how changing therapists affects patients and therefore addressed the following research questions:

Study III

1. Is there a drop in the alliance with the new therapist after a transfer?
2. Does a therapist transfer have a negative impact on patients' impairment?
3. How do alliance and symptoms develop after a therapist transfer?

4. Methodological aspects

The statistical approach of all three studies included the same methodology: Multilevel modeling. Therefore, this chapter explains why multilevel modeling is needed, when to apply this method and which requirements are necessary.

Most of the common statistics (e.g., linear regression) require a random sample drawn from one population where cases are independent of each other (Hox, Moerbeek, & van de Schoot, 2010). This assumption is often violated, which the following example adapted from Eid, Gollwitzer, and Schmitt (2017) attempts to explain: Imagine a psychologist wants to examine aggression in pupils. Is the level of aggression associated with sex, peer-popularity within their class, and their intellectual ability? The question is, whether the measured values on the dependent variable (level of aggression) are independent from each other. One can assume that aggression values within one class are more similar than between different classes. Also, aggression values might be more similar in one school, than between different schools. Some of the investigated schools might have different psycho-social climates. In a school with rough and tense atmosphere, pupils might me more aggressive than pupils in a friendly and harmonious school. In this example, we have no random sample from one population. Instead, we have a hierarchical random sampling. First a random sample of schools was selected, then a random sample of classes, finally a random sample of pupils.

The situation is similar in most clinical contexts. If psychotherapy is studied, patients are usually treated by different therapists and a therapist usually has a caseload of different patients. Analyzing patients' data without accounting for the aggregation of therapists violates the requirement that the sample was randomly drawn and that patients are independent of each other. One can assume that patients treated by the same therapist are more similar than patients from different therapists. The therapist has the same level of experience, uses the same approach, has the same habits, and so on. Therefore, patients treated by the same therapist are exposed to the same therapist features. The same is true, if data is analyzed on a session level. If, for instance, symptom severity is studied over the course of treatment, one can assume that session n and n+1 from patient A are more similar than session n from patient A compared to session n + 1 from patient B.

The natural structure of psychotherapy research data is characterized by multiple sessions from multiple patients treated by the same therapists (see Figure 1).

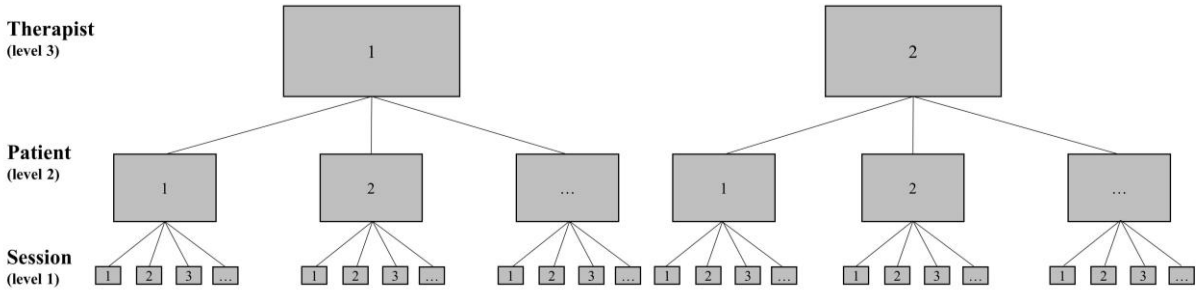

Figure 1. Illustration of a three-level nested data structure where sessions are nested within patients and patients are nested within therapists.

Statistically speaking, the session observations are nested within a higher-level unit (patient) and the patient observations are also nested within a higher-level unit (therapist). Therefore, neither session observations nor patient observations are independent of one another. This relatedness of observations violates the assumption of independence, which is a basic assumption of most traditional statistical methods. Ignoring hierarchical data structure can lead to ecological fallacies and to flawed p-values (Hox et al., 2010). Multilevel modeling takes hierarchical data structures into account, allowing data to be analyzed on different levels at the same time (Hox et al., 2010; Raudenbush & Bryk, 2002). Also, interactions between predictors on different levels can be studied. Multilevel modeling can be conceptualized as a hierarchical system of regression equations. To gain a better understanding of this approach, a three-level model with variables on different levels will be presented below. Imagine a researcher wants to study whether the change of symptoms in patients within a treatment is associated with the experience of the therapist. Accordingly, the model specifications for the different levels are:

Level 1: $Y_{ijk} = \beta_{0jk} + \beta_{1jk}(\text{Session}_{ijk}) + \varepsilon_{ijk}$

Level 2: $\beta_{0jk} = \gamma_{00k} + b_{0jk}$

$\beta_{1jk} = \gamma_{10k}$

Level 3: $\gamma_{00k} = \delta_{000}$

$\gamma_{10k} = \delta_{100} + \delta_{101}(\text{Experience}_k) + V_{10k}$

The session i is nested within the patient j, who is nested within the therapist k. Y_{ijk} is the symptom impairment in session i, for patient j, treated by therapist k. Session_{ijk} is a predictor on level one (i.e., a linear change of symptoms from session to session). Each of the level one regression coefficients (β_{0jk}; β_{1jk}) is explained by separate equations on level two. Again, regression coefficients on level two (γ_{00k}; γ_{10k}) are explained by separate equations on level three.

Substituting all higher level terms into the lower level equation yields the following equation:

$Y_{ijk} = (\delta_{000} + b_{0jk}) + (\delta_{100} + \delta_{101} * (\text{Experience}_k) + V_{10k}) * (\text{Session}_{ijk}) + \varepsilon_{ijk}$

This equation expands to:

$Y_{ijk} = \delta_{000} + b_{0jk} + \delta_{100} * (\text{Session}_{ijk}) + \delta_{101} * (\text{Experience}_k) * (\text{Session}_{ijk}) + V_{10k} * (\text{Session}_{ijk}) + \varepsilon_{ijk}$

δ_{000} is the mean intercept among therapists and patients with a deviation term b_{0jk}, indicating that patients differ in their initial symptom impairment. The slope (i.e., the linear symptom change of patients over sessions) consists of three different parts: A mean slope among therapists and patients ($\delta_{100} * (\text{Session}_{ijk})$), a cross-level interaction term between the experience of the therapists and the session ($\delta_{101} * (\text{Experience}_k) * (\text{Session}_{ijk})$), as well as a deviation term ($V_{1k} * (\text{Session}_{ijk})$; i.e., the change in patients' symptoms over the sessions varies between therapists). Additionally, the equation contains a session-specific residual term (ε_{ijk}). This example model can deal with different research questions such as: How strong is change per session on average? Do therapists differ with regard to their change curves? What is the impact of therapists' experience on change? If questions on differences between therapists are studied, the variation terms are analyzed. Multilevel modeling is able to partition the total variability of the dependent variable into different components.

In a two level model (e.g., patients nested within therapists), there is variance within patients on level 1 and variance between therapists on level 2. The variance of level 2 divided by the total variance is termed the intraclass correlation coefficient (ICC; Raudenbush & Bryk, 2002). The ICC ranges from 0 to 1, with higher values indicating that a greater proportion of the variance in the dependent variable is due to differences between therapists. The coefficient multiplied by 100 provides an estimation of the therapist effect (Baldwin & Imel, 2013). Correspondingly, the ICC in a three-level model can be estimated by dividing the variance component of interest (e.g., therapist variance on level three) by the total variance (Lutz, Leon, Martinovich, Lyons, & Stiles, 2007). Calculating therapist effects with dichotomized dependent variables (e.g., drop-out) is a little different as the variance of the error term is fixed to $\frac{\pi^2}{3}$ (Hox et al., 2010). In this case, the therapist effect can be calculated by dividing the variance between therapists by the total variance (i.e., variance between therapists plus error variance).

5. Study I

Rubel, J. A., Zimmermann, D., Deisenhofer, A. K., Müller, V., & Lutz, W. (2017). Nutzung von psychometrischem Feedback als empirische Unterstützung des Supervisionsprozesses bei Ausbildungstherapien. *Zeitschrift für Klinische Psychologie und Psychotherapie, 46*(2), 83-95. doi: 10.1026/1616-3443/a000413

Schwerpunktbeitrag

Nutzung von psychometrischem Feedback als empirische Unterstützung des Supervisions-prozesses bei Ausbildungstherapien

Julian A. Rubel, Dirk Zimmermann, Anne-Katharina Deisenhofer, Viola Müller und Wolfgang Lutz

Universität Trier

Zusammenfassung: *Theoretischer Hintergrund:* Trotz der gut belegten generellen Wirksamkeit von Psychotherapie wird die Zahl der Patienten, die nicht auf eine Therapie ansprechen oder sich sogar im Verlauf verschlechtern, auf etwa ein Drittel geschätzt. Da Therapeuten im Vergleich zu empirischen Algorithmen weniger gut in der Lage sind negative Entwicklungen zu entdecken oder zu prognostizieren, brauchen sie zusätzliche Unterstützung in Form von kontinuierlichen Rückmeldungen über den Fortschritt ihrer Patienten. Solche Feedbackinterventionen können als empirisch basierte Unterstützung des Supervisions- oder Intervisionsprozesses verstanden werden. Diese Interventionen haben ihre Wirksamkeit zur Reduktion therapeutischer Misserfolge in zahlreichen Einzelstudien und Meta-Analysen wiederholt zeigen können. *Fragestellung:* Wie nutzen Therapeuten ein solches Feedback in ihrer praktischen Arbeit und in welchem Ausmaß spielen Therapeutenfaktoren dabei eine Rolle? *Methode:* In der vorliegenden Studie wurden 72 Therapeuten für 648 ihrer Patienten unmittelbar nach der Therapie dazu befragt, wie sie das psychometrische Feedback genutzt haben. Therapeutenunterschiede wurden mittels Mehrebenenmodellen ermittelt. *Ergebnisse:* Es konnte gezeigt werden, dass Therapeuten für einen Großteil ihrer Patienten das Feedback verwendeten. In etwa einem Drittel der Fälle gab es den Therapeuten den Anstoß zusätzliche Hilfen (z. B. Supervision/Intervision) zu beanspruchen. Ähnlich einer personengestützten Supervision wurde für über die Hälfte der Patienten das Feedback genutzt um therapeutische Interventionen anzupassen. Ob und in welcher Form es genutzt wurde hing jedoch stark von dem Therapeuten ab. Je nach Verwendungsart konnten Therapeutenunterschiede zwischen 27 % und 52 % der Feedbacknutzung erklären. Die Wahrscheinlichkeit, dass Therapeuten die Nutzungsmöglichkeiten des Feedbacks anwendeten war größer, wenn diese angaben im Durchschnitt zufriedener mit den grafischen Rückmeldungen zu sein. *Schlussfolgerungen:* Diese Ergebnisse machen deutlich, dass es zu einem großen Teil nicht nur vom Therapieverlauf oder den Charakteristika der jeweiligen Patienten abhängt wie Therapeuten das Feedback nutzen, sondern auch von Variablen, die den Therapeuten betreffen.

Schlüsselwörter: Supervision, Ausbildung, psychometrisches Feedback, Therapeutenunterschiede, Qualitätssicherung, patientenorientierte Therapieforschung

Using Psychometric Feedback as Empirical Support for Supervision Processes in Psychotherapy Training

Abstract: *Background:* Despite the well-evidenced general effectiveness of psychotherapy, an estimated one third of patients do not respond to – or even deteriorate – during therapy. As therapists are less able to recognize and prognosticate negative developments in comparison with empirical algorithms, they require further support in the form of continuous feedback of their patients' progress. Such feedback interventions can be seen as empirically based support of the supervision and intervision process. These interventions have repeatedly demonstrated their effectiveness for the reduction of treatment failure in numerous single studies and meta-analyses. *Objective:* The current study investigates how therapists use feedback in their clinical work. Additionally, we quantified to what extent therapist factors influence the use feedback. *Method:* Therefore, 72 therapists were surveyed about how they used psychometric feedback immediately after having finished therapy with 648 of their patients. Differences between therapists were determined by employing multilevel models. *Results:* We were able to show that therapists used feedback for a large portion of their patients. In approximately one third of the cases, usage prompted the therapists to employ extra support (e. g., supervision, intervision). Similar to staff-based supervision, feedback was used to adjust therapeutic interventions for over half of the patients. However, when and how it was used was largely dependent on the individual therapist. Depending on the type of utilization, therapist differences were able to explain between 27 % and 52 % of feedback use. The probability that therapists used the possible applications of feedback was greater when they also indicated being more satisfied with the

Diese Arbeit wurde teilweise durch die Deutsche Forschungsgemeinschaft (DFG) unterstützt (LU 660/10 – 1; LU 660/8 – 1).

graphic feedback on average. *Conclusion:* These results underline that how therapists use feedback depends not only on therapy progress or specific patient characteristics, but also to a large degree on therapist variables.

Keywords: supervision, training, psychometric feedback, therapist differences, quality assurance, patient-oriented therapy research

Trotz der gut belegten generellen Wirksamkeit von Psychotherapie wird die Zahl der Patienten, die nicht auf die Therapie ansprechen oder sich in deren Verlauf sogar verschlechtern, auf ca. ein Drittel geschätzt (Lambert, 2013). Zusätzliche Befunde zeigen, dass Therapeuten selbstwertdienlichen Verzerrungen unterliegen und daher die Verschlechterungsraten ihrer Patienten häufig unterschätzen (Hannan et al., 2005; Hatfield, McCullough, Frantz & Krieger, 2010; Walfish, McAlister, O'Donnell & Lambert, 2012). Folglich ist eine frühzeitige Detektion von negativen Verläufen sowie eine daran orientierte Anpassung der Therapiestrategie indiziert (Lutz et al., 2014; Rubel et al., 2015). Allerdings stößt die traditionelle, auf Therapieschulen und Ansätze fokussierte Psychotherapieforschung dabei an ihre Grenzen (Lutz & Rubel, 2015). Eine Forschungsstrategie, die sich auf differenzielle Verläufe konzentriert, ist die *Patientenorientierte Therapieforschung* (POF; Howard, Moras, Brill, Martinovich & Lutz, 1996). Diese zeichnet sich durch kontinuierliche, psychometrische Erhebungen im Verlauf der Therapie und der unmittelbaren Rückmeldung dieser Ergebnisse an die behandelnden Therapeuten aus (z. B. Castonguay, Barkham, Lutz & McAleavy, 2013; Lambert, 2007). Die Rückmeldungen können sich sowohl auf den Fortschritt des Patienten bezüglich der Symptome und des Wohlbefindens beziehen (Ergebnisrückmeldung, engl. outcome feedback/progress feedback), als auch auf prozessrelevante Mechanismen wie die Therapiebeziehung (Prozessrückmeldungen; engl. process feedback; Douglas et al., 2015; Lucock et al., 2015; Miller, Duncan, Brown, Sorrell & Chalk, 2006). Psychometrisches Feedback gilt als vielversprechende Möglichkeit die Anzahl derjenigen Patienten, die nicht auf Psychotherapie ansprechen, weiter zu reduzieren, indem Therapeuten für diese Patientengruppe Warnsignale erhalten und so auf problematische Entwicklungen aufmerksam gemacht werden, welche ihnen ohne Feedback potenziell nicht bewusst gewesen wären (z. B. Lambert, 2007; Bar-Kalifa et al., 2016). In mehreren Metaanalysen konnte der positive Effekt solcher Warnsignale für „not-on-track" Patienten nachgewiesen werden (Lambert et al., 2003; Shimokawa, Lambert & Smart, 2010). Die Studien scheinen sich aber dahingehend zu unterschieden ob und wie intensiv Therapeuten psychometrische Rückmeldungen mit dem Patienten in den Sitzungen besprechen (Krägeloh, Czuba, Billington, Kersten & Siegert, 2015).

Auch für Patienten, die sich erwartungsgemäß oder besser entwickeln („on-track") wurden bereits positive Effekte eines kontinuierlichen Verlaufsfeedbacks für ein verbessertes Therapieergebnis nachgewiesen (Shimokawa et al., 2010; Amble, Gude, Stubdal, Andersen & Wampold, 2015). Diese fielen jedoch deutlich kleiner aus als für „not-on-track" Patienten. Potenziell zeigen sich Feedbackeffekte für diese Patientengruppe vor allem in Form einer besseren Stabilität der Therapieeffekte. So konnten Byrne, Hooke, Newnham und Page (2012) in einem stationären Setting für die Gruppe der „on-track" Patienten eine bedeutsam reduzierte Zahl an stationären Wiederaufnahmen zeigen, wenn deren Behandler mit Feedback über den Verlauf ihrer Patienten informiert wurden.

Psychometrisches Feedback kann im Rahmen von Ausbildungstherapien folglich als eine Art automatisierte und empirisch gestützte Supervision verstanden werden. Ähnlich einer persönlichen Super- oder Intervision erhalten Therapeuten durch psychometrische Feedbackinterventionen zusätzliche Informationen und neue Denkanstöße, die sie zur Optimierung ihrer laufenden Therapien nutzen können. Auch können die psychometrischen Rückmeldungen innerhalb der Supervision als zusätzliche Informationsquelle vom Supervisor genutzt werden. Neben der klinischen Rückmeldung des Therapeuten, bekommt der Supervisor somit zusätzlich einen direkten Einblick in die Selbstauskunftsperspektive des Patienten (Worthen & Lambert, 2007). Insgesamt bestätigen bisherige Studien die Effektivität von Feedback (z. B. Lambert et al., 2003). Die mittlere Effektstärke zwischen Feedback und *treatment as usual* (TAU) liegt bei $d = .52$ (95 % Konfidenzintervall (KI): .30 – .72; Shimokawa et al., 2010). Der Feedbackeffekt kann durch die Hinzunahme sogenannter klinischer Unterstützungshilfen (*clinical support tools;* CST) erhöht werden (z. B. Whipple et al., 2003). Neben den Ergebnisrückmeldungen erhält der Therapeut auch Feedback über Belastungen in innertherapeutischen sowie außertherapeutischen Bereichen und ihm werden zusätzlich Problemlösehilfen vorgestellt (White et al., 2015). In ihrer Ursprungsform basieren diese Problemlösehilfen auf einem zusätzlichen Fragebogen zur Einschätzung der Signalfälle in Bezug auf fünf Bereiche (Lambert, Bailey, Kimball & Shimokawa, 2007): Therapiebeziehung, soziale Unterstützung, Lebensereignisse, Motivation und Medikation. Je nachdem in welchem dieser Bereiche der Signalpatient eine besondere Belastung aufweist, erhalten Therapeuten Empfehlungen zum Umgang mit diesen Problemen. Für Bedingungen in denen zusätzlich zum Feedback CST eingesetzt wurden liegt die Effektstärke

im Vergleich zur Standardbehandlung ohne Feedback bei d = .70 (KI: .52–.87). In diesem Zusammenhang konnte eine Studie von White und Kollegen (2015) zeigen, dass ein relativ hoher Anteil an „not-on-track" Patienten (29 %) klinisch relevante Belastungen im Bereich der sozialen Unterstützung aufweist. Bei 19 % der „not-on-track" Patienten erhielten deren Therapeuten außerdem Warnsignale für die Therapiebeziehung und/oder Therapiemotivation.

Die anfänglichen Feedbackstudien wurden jedoch vornehmlich in Settings mit relativ kurzen Interventionen und wenig belastetsten Patienten durchgeführt (z. B. Beratungszentren; Newnham & Page, 2010; Poston & Hansen, 2010; Shimokawa et al., 2010). Aktuellere Studien untersuchten die Effekte von Feedback bei vergleichsweise stärker belasteten Patienten (De Jong, van Sluis, Nugter, Heiser & Spinhoven, 2012; Simon, Lambert, Harris, Busath & Vazquez, 2012), in stationären psychosomatischen Einrichtungen (Probst et al., 2013; Byrne et al., 2012), bei Patienten mit Essstörungen (Simon et al., 2013) und bei Patienten mit längeren Therapien (35 Wochen oder mehr; De Jong et al., 2014). Diese aktuelleren Studien identifizierten Variablen, welche die Wirksamkeit von Feedback moderieren (Lutz, De Jong & Rubel, 2015). Diese Befunde stehen in Einklang mit der Contextual Feedback Intervention Theory, welche die positiven Effekte von Feedback in der Psychotherapie theoretisch erklären soll (Sapyta, Riemer und Bickman, 2005). Laut dieser Theorie sind die Effekte von Feedback nicht per se immer zu erwarten, sondern werden von verschiedenen Einflussfaktoren moderiert: Neben der Glaubwürdigkeit der Feedbackgebenden Quelle werden hier die Richtung des Feedbacks (positiv vs. negativ), der Inhalt des Feedbacks (deskriptive Informationen vs. weiterführende Empfehlungen) und das Feedbackformat (direkt vs. zeitverzögert; einmalig vs. wiederholt; Status- vs. Verlaufsfeedback) als entscheidende Einflussfaktoren diskutiert. In Einklang mit dieser Theorie beobachteten beispielsweise Bickman, Kelley, Breda, de Andrade und Riemer (2011) einen positiven Zusammenhang zwischen der Häufigkeit, mit der Therapeuten Feedback über den Verlauf ihrer Patienten erhalten und dem Therapieergebnis. Eine höhere Feedback-Frequenz hing mit einem größeren Effekt des Feedbacks zusammen. Neben der Häufigkeit der gegebenen Rückmeldungen zeigt auch die Gestaltung des Feedbacks einen Einfluss auf die Wirkung dieser Intervention. So zeigte ein Feedbacksystem, das auf mehrere Domänen psychischen Funktionierens (Wohlbefinden und Symptombelastung) abzielte bessere Erfolge für Risikopatienten, als ein eindimensionales Feedbacksystem (nur Wohlbefinden; Dyer, Hooke & Page, 2016).

Neben verschiedenen Charakteristika des jeweiligen Feedbacksystems werden mittlerweile auch Therapeutenmerkmale als moderierende Einflussfaktoren des Feedbackeffekts diskutiert. Insgesamt gibt es nur sehr wenige Studien, die sich mit der Feedbacknutzung in Kombination mit Therapeutenmerkmalen und Therapeutenunterschiedenen auseinandersetzen (De Jong et al, 2012; Simon et al, 2012; Lutz, Rubel et al, 2015). In Bezug auf das Therapieergebnis gibt es zahlreiche Studien, die zeigen konnten, dass sich Therapeuten hinsichtlich ihrer Effektivität systematisch unterscheiden (z. B. Baldwin & Imel, 2013; Lutz, Leon, Martinovich, Lyons & Styles, 2007; Saxon & Barkham, 2012; Schiefele et al, 2016). Im Großteil dieser Studien wird der Anteil der Ergebnisvarianz, der auf Unterschiede zwischen Therapeuten zurückzuführen ist, auf ca. 5–8 % geschätzt.

Auch im Bereich der Feedbackforschung gibt es erste Befunde, die Therapeutenunterschiede nahelegen. So konnten De Jong und Kollegen (2012) einen differentiellen Feedbackeffekt für sich zu Beginn der Therapie negativ entwickelnde Patienten nur dann finden, wenn deren Therapeuten berichteten das Feedback in der Behandlung des Patienten genutzt zu haben. Die Nutzung wiederum war bei weiblichen Therapeuten sowie bei Therapeuten mit einer höheren Motivation für die Feedbacknutzung zu beobachten, während die Erfahrung der Therapeuten keinen Einfluss hatte. Dementsprechend beobachten auch Simon et al. (2012), dass nur 50 % der Therapeuten in ihrer Studie dazu in der Lage waren Feedback so einzusetzen, dass es das Ergebnis ihrer Patienten verbessern konnte. Für die andere Hälfte machte es keinen Unterschied, ob die Therapeuten Feedback erhielten oder nicht. Lutz, Rubel et al. (2015) berichten überdies einen generellen Einfluss der Einstellung der Therapeuten und Patienten gegenüber psychometrischer Rückmeldungen und Feedback auf das Therapieergebnis. Patienten, die im Rahmen einer Feedbackstudie eine positive Einstellung gegenüber Qualitätssicherungsmaßnahmen mittels psychometrischer Fragebögen angaben, zeigten ein signifikant besseres Therapieergebnis als Patienten mit durchschnittlicher oder negativer Einstellung. Die Einstellung der Therapeuten wurde in dieser Studie als Index aus deren Zufriedenheit mit dem Feedbacksystem und der Häufigkeit der Feedbacknutzung gebildet. Dabei zeigten vor allem Patienten von Therapeuten, die nicht mit dem Feedbacksystem zufrieden waren und dieses trotzdem häufig und ungezielt nutzten um Modifikationen der Therapie vorzunehmen, die schlechtesten Therapieeffekte. Besonders effektiv waren die Therapien für Patienten von zufriedenen Therapeuten, die das Feedback sehr gezielt nur für eine spezifische Modifikation ihres Vorgehens nutzten.

Zusammenfassend lässt sich konstatieren, dass die Untersuchung der Feedbacknutzung in Kombination mit Therapeutenunterschieden bisher wenig Beachtung in

der Forschung gefunden hat, wobei erste Ergebnisse auf differenzielle Effekte hinweisen. Die vorliegende Studie untersucht erstmalig die Feedbacknutzung von Psychotherapeuten in Ausbildung. Da bisher vergleichbare Daten in diesem Kontext fehlen, können wir unsere Fragestellungen und Erwartungen lediglich auf die Befunde eines Modellprojekts der Techniker Krankenkasse (TK-Studie) zum Qualitätsmonitoring in der ambulanten Psychotherapie gründen (Lutz, Böhnke, Köck & Bittermann, 2011; Lutz, Wittmann, Böhnke, Rubel & Steffanowski, 2013). Vor diesem Hintergrund erwarten wir, dass Therapeuten für einen Großteil ihrer Patienten (ca. 2/3) Feedback nutzen. Als häufigste Art der Nutzung erwarten wir, wie in der TK-Studie, das Besprechen des Feedbacks mit dem Patienten. In einem zweiten Schritt soll getestet werden, inwieweit die Nutzung des Feedbacks von patienten- oder therapeutenspezifischen Faktoren abhängt. Dementsprechend soll die Frage untersucht werden, ob Therapeuten für jeden Patienten individuell entscheiden ob sie das Feedback nutzen oder ob diese Entscheidung eher von Therapeutenvariablen, wie deren Einstellung zum Feedback, abhängt. Es wird erwartet, dass die Nutzung von Feedback systematisch zwischen Therapeuten variiert und ein signifikanter Anteil der Nutzung pro Patient auf Therapeutenmerkmale zurückzuführen ist.

Methode

Stichprobe

Die Stichprobe umfasst 648 Patienten, welche von 72 Therapeuten in Ausbildung (Schwerpunkt Verhaltenstherapie) behandelt wurden. Eine Übersicht der demographischen Charakteristiken der Patienten kann in Tabelle 1 eingesehen werden. Die durchschnittliche psychometrische Belastung auf dem Global Severity Index des Brief Symptom Inventory (Franke, 2000) liegt bei 1,23 (*SD* = 0,65) und liegt damit im Bereich der für eine ambulante Stichprobe erwartet werden kann. Nach dem Strukturierten Klinischen Interview für DSM-IV (Wittchen, Zaudig & Fydrich, 1997) erhalten die meisten der Patienten als Primärdiagnose nach DSM-IV eine depressive Störung (50,4 %) oder eine Angststörung (29,4 %). 190 Patienten erfüllen die Kriterien für eine Diagnose (29,3 %) während 443 Patienten (68,4 %) zwei oder mehr Diagnosen erhalten. Von den Patienten mit mehreren Diagnosen erfüllen 188 Patienten (42,4 %) sowohl die Diagnose einer Angststörung als auch die Diagnose einer depressiven Störung.

Neun Therapeuten sind männlich (12,5 %) und 63 Therapeutinnen weiblich (87,5 %). Zum Zeitpunkt des ersten Patientenkontaktes liegt das Durchschnittsalter der Therapeuten bei 29,7 Jahren (*SD* = 5,00). Diese Stichprobe umfasst alle Therapeuten, die zwischen 2010 und 2016 Therapien in der Institutsambulanz angeboten haben und für die mindestens Daten von fünf verschiedenen Patienten in der Datenbank vorliegen. Dieses Einschlusskriterium (mindestens fünf Patienten pro Therapeut) wurde angelegt um Therapeutenunterschiede quantifizieren zu können.

Rückmeldungen

Die Symptome der Patienten wurden unmittelbar vor Beginn jeder Sitzung mittels „Touchscreen"-Eingabemasken erhoben. Hierzu wurde eine 11-Item Kurzversion des SCL-90-R (Franke, 1995) eingesetzt, dessen Gesamtmittelwert hohe Korrelationen mit dem GSI zeigen konnte (r = .91; Lutz, Tholen, Schürch & Berking, 2006). Nach jeder Sitzung gaben Patienten und Therapeuten im Berner Stundenbogen (Flückiger, Regli, Zwahlen, Hostettler & Caspar, 2010) an, wie sehr sie in der heutigen Sitzung die vier Wirkfaktoren nach Grawe (Ressourcenaktivierung, Problemaktualisierung, Problembewältigung, Motivationale Klärung) realisiert gesehen haben und wie gut sie die Qualität der therapeutischen Beziehung empfanden.

Unmittelbar nach jeder Sitzung wurden den Therapeuten die auf diese Weise erhobenen Daten ebenfalls am Touchscreen-Computer grafisch aufbereitet zurückgemeldet. Das Feedback umfasste also den Symptomverlauf des Patienten sowie die Entwicklung der Wirkfaktoren und Beziehungseinschätzungen aus Patienten- und Therapeutenperspektive. Neben diesem sitzungsspezifischen Feedback unmittelbar nach jeder Sitzung war es den Therapeuten außerdem jederzeit möglich sich, in einer zu diesem Zweck programmierten Software, die Verläufe ihrer Patienten einzusehen. Während die Rückmeldungen nach jeder Sitzung routinemäßig vorgenommen wurden und von den Therapeuten nicht abgestellt werden konnten, erfolgte die Nutzung der Feedbacksoftware auf freiwilliger Basis.

Instrumente

Fragebogen zur Nutzung von Rückmeldungen. Nach jeder Behandlung wurden Therapeuten dazu befragt, wie sie die psychometrischen Rückmeldungen im Verlauf der Therapie genutzt haben. Dazu wurde den Therapeuten eine Liste mit elf Möglichkeiten vorgelegt, für die diese jeweils angeben sollten, ob Sie diese Anpassung vorgenommen haben oder nicht. Die Liste bestand aus dem Eingangssatz „Ich habe aufgrund der Rückmeldungen …"

Tabelle 1. Demographische Daten der Patienten

	M (SD) oder *n* (%)
Sitzungen	37.19 (18.37)
Alter	36.96 (12.84)
Weiblich	411 (63.4)
Höchster Schulabschluss	
Hauptschule	140 (21,6)
Realschule	179 (27,6)
Abitur oder Fach	300 (46,3)
Sonstiges	29 (4,5)
Familienstand	
ledig	347 (53,5)
verheiratet	195 (30,1)
getrennt lebend	23 (3,5)
geschieden	70 (10,8)
verwitwet	13 (2,0)
Arbeitsunfähig	130 (20,1)
GSI	1.23 (.65)
Komorbiditäten	
1 Diagnose	190 (29.3)
>1 Diagnose	443 (68.4)
Primärdiagnosen (nach SKID-I)	
Depression	304 (46.9)
Dysthymie	23 (3.5)
Agoraphobie und Panikstörung	46 (7.1)
Belastungs- od. Anpassungsstörung	84 (13)
Andere Angststörung	60 (9,3)
Essstörung	19 (2,9)
Andere Störung	93 (14.4)
Keine Diagnose	15 (2,3)

Anmerkungen: GSI= Global Severity Index des Brief Symptom Inventory, *M* = Mittelwert; *SD* = Standardabweichung; SKID = Strukturiertes Klinisches Interview für DSM-IV.

und den folgenden elf Optionen: „... keine Veränderungen meiner Behandlung vorgenommen", „... meine therapeutischen Interventionen versucht anzupassen", „... versucht, die therapeutische Beziehung zu verbessern", „... das Therapieende vorbereitet", „... die Ressourcen des Patienten / der Patientin versucht zu fördern", „... die Therapiemotivation des Patienten / der Patientin versucht zu erhöhen", „... die interpersonalen Probleme des Patienten / der Patientin besprochen", „... die Ergebnisse in den Fragebögen mit dem Patienten / der Patientin besprochen", „... neue Hausaufgaben erprobt", „... zusätzliche Hilfe hinzugezogen (Intervision / Supervision, Literatur, Weiterbildung, etc.)", oder „... die Sitzungsintervalle variiert".

Auswirkungen des Feedbacks. Um die globalen Auswirkungen der Rückmeldungen pro Patient zu erfassen, wurden Therapeuten am Ende jeder Behandlung gebeten die Frage „Wie hilfreich waren für Sie – im Hinblick auf diesen Patienten – die Informationen, die Sie aus den Status- und Verlaufsrückmeldungen erhalten haben?" auf einer fünfstufigen Likert-Skala (von „2" – „sehr hilfreich" bis „-2" – „nicht hilfreich") zu beantworten. Darüber hinaus sollten Therapeuten einschätzen, ob sich die psychometrischen Rückmeldungen positiv oder negativ auf

die therapeutische Beziehung ausgewirkt haben („Haben sich die grafischen Rückmeldungen auf die Beziehung zum / zur Patienten / In ausgewirkt?"). Auch hier bestand das Antwortformat aus einer fünfstufigen Likert-Skala (von „2" – „sehr positiv" bis „-2" – „sehr negativ"). Da diese beiden Fragen erst zu einem späteren Zeitpunkt in die Routineerhebungen aufgenommen wurden als die oben beschriebenen Fragen zur Feedbacknutzung, liegen diese Einschätzungen nur von 52 der 72 Therapeuten vor.

Darüber hinaus wurden für jeden Patienten weitere störungsübergreifende und störungsspezifische Fragebögen eingesetzt. Eine detaillierte Beschreibung des Standardvorgehens zur Indikationsstellung, Therapieverlaufs- und Ergebnismessung in der Forschungsambulanz der Universität Trier wurde an anderer Stelle bereits ausführlich dargestellt (Köck & Lutz, 2012).

Datenanalyse

Therapeutenunterschiede wurden über die Intraklassen-Korrelation (ICC) in hierarchisch linearen Modellen (HLM) ermittelt. Dabei wurden je nach abhängiger Variable entweder klassische HLMs für kontinuierliche Variablen oder binär-logistische HLMs für binäre Variablen gerechnet (z. B. Hox, Moerbeek & van de Schoot, 2010). In beiden Fällen wurden Patienten auf Level 1 und Therapeuten auf Level 2 aufgenommen, um die genestete Datenstruktur zu berücksichtigen. Alle Auswertungen wurden mit der Statistiksoftware R Version 3.25 (R Core Team, 2016) sowie dem Paket lme4 (Bates, Maechler, Bolker & Walker, 2014) berechnet. Bei den HLMs mit kontinuierlichen Variablen wurde die Varianz auf Level 2 durch die Gesamtvarianz (Level 1 Varianz + Level 2 Varianz) geteilt, um die ICC (entspricht dem Therapeuteneffekt) zu bestimmen. In Modellen mit binären abhängigen Variablen ist der Fehlerterm (Level 1 Varianz) vorab mit $\sigma_e^2 = \frac{\pi^2}{3} = 3,29$ fixiert. Zur Berechnung der ICC wird hierzu die Level 2 Varianz durch die Gesamtvarianz (3,29 + Level 2 Varianz) geteilt. Patienten, für die Therapeuten keine Informationen zur Feedbacknutzung angegeben haben, wurden aus der jeweiligen Analyse ausgeschlossen.

Ergebnisse

Feedbacknutzung

Abbildung 1 zeigt die relative Häufigkeiten der Feedbacknutzung in den jeweiligen Kategorien. Für über die Hälfte der Patienten (51,8 %) geben die Therapeuten an das Feedback genutzt zu haben um ihre therapeutischen Interventionen anzupassen. Die zweit- und dritthäufigste Reaktion auf das Feedback bestand in der Arbeit an der therapeutischen Beziehung (31,2 %) und im Hinzuziehen zusätzlicher Hilfen (30,7 %). Am seltensten nutzten Therapeuten das Feedback um neue Hausaufgaben zu erproben (7 %). Für 19,1 % der Patienten geben die Therapeuten an, keine Anpassungen vorgenommen zu haben.

Auf die Frage, wie hilfreich die Rückmeldungen für diesen Patienten waren, geben die Therapeuten im Mittel an, dass es „eher hilfreich" gewesen sei (M = 0,99; SD = 0,70; N = 360).

Die Auswirkungen der grafischen Rückmeldungen bewerten die Therapeuten im Durchschnitt zwischen „weder positiv noch negativ" und „eher positiv" (M = 0,64; SD = 0,63; N = 359).

Tabelle 2 enthält die Korrelationen zwischen den Häufigkeiten mit denen Therapeuten die jeweilige Nutzungsmöglichkeit angewendet haben sowie deren mittleren Nützlichkeitseinschätzungen. Die Tabelle macht deutlich, dass Therapeuten, die im Durchschnitt angeben zufriedener mit den grafischen Rückmeldungen gewesen zu sein, auch mit einer höheren Wahrscheinlichkeit die jeweiligen Modifikationen vorgenommen haben. Die höchsten Korrelationen der Nützlichkeitseinschätzungen zeigen sich mit Versuchen die therapeutischen Interventionen anzupassen (r = .553), dem Besprechen interpersonaler Probleme der Patienten (r = .520) und dem generellen Besprechen der Fragebogenergebnisse (r = .518). Je höher die wahrgenommene Nützlichkeit der Rückmeldungen, desto niedriger war die Wahrscheinlichkeit, dass Therapeuten angeben keine Anpassungen auf der Grundlage des Feedbacks vorgenommen zu haben (r = -.616) und desto positiver ist auch der wahrgenommene Einfluss der Rückmeldungen auf die Therapiebeziehung (r = .677).

Therapeuten- und Patienteneinflüsse

Tabelle 3 schlüsselt den Anteil der Gesamtvarianz in der Feedbacknutzung auf, der auf Patienten- (PE) und Therapeutenmerkmale (TE) zurückzuführen ist. Die AIC-Differenz gibt den Unterschied des AIC im Modell ohne und mit Therapeutenunterschieden an. Je höher dieser Wert, desto besser passt das Modell, welches nicht nur Patientenunterschiede sondern auch Therapeutenunterschiede zulässt. Gleiches gilt für die χ^2-Differenz. Ein signifikanter Wert hier zeigt an, dass signifikante Therapeutenvarianz vorliegt und somit das Modell mit Therapeutenunterschieden besser auf die Daten passt. Außer für die Nutzung der Kategorie „Sonstiges" zeigt sich für alle Fragen

Abbildung 1. Relative Häufigkeit der Feedbacknutzung in der Selbstauskunft.

ein signifikanter Anteil an Therapeutenvarianz. Den größten signifikanten Therapeutenvarianzanteil zeigt die Anpassung „neue Hausaufgaben erprobt" (52,26 %). Die Nutzung von Hausaufgaben als Konsequenz des Feedbacks scheint also eine eher therapeutenspezifische Maßnahme zu sein. Der zweithöchste Therapeuteneffekt ergibt sich für die Kategorie „keine Anpassungen vorgenommen" (49,60 %). Ob also überhaupt Feedback genutzt wird, hängt ungefähr zu 50 % von Therapeutenmerkmalen, wie zum Beispiel der Einstellung gegenüber Qualitätssichernder Maßnahmen, ab. Dies deckt sich mit den in Tabelle 2 berichteten Korrelationen. Je häufiger Therapeuten angeben, keine Anpassungen vorgenommen zu haben, desto weniger hilfreich nehmen sie die Rückmeldungen wahr (r = -0,616***) und desto mehr negative Auswirkungen auf die Therapiebeziehung nehmen sie wahr (r = -0,467***). Am wenigsten von Therapeutenvariablen beeinflusst scheint die Nutzung der Rückmeldungen zur Verbesserung der Therapiebeziehung (27,37 %). Ob ein Therapeut also auf der Grundlage der Rückmeldungen versucht die therapeutische Beziehung zu verbessern, hängt deutlich stärker von patientenspezifischen Faktoren ab (zu etwas mehr als zwei Drittel), als von therapeutenspezifischen.

Auch die Einschätzung, wie hilfreich die Verlaufsrückmeldungen für die Behandlung eines Patienten waren und ob sich diese auf die Therapiebeziehung ausgewirkt haben, ist stark durch Therapeutenunterschiede beeinflusst (TE = 39,47 % bzw. 34,16 %).

Diskussion

Die vorliegende Studie gibt einen Überblick über die Nutzung von psychometrischen Rückmeldungen durch Psychotherapeuten in Ausbildung. Demnach passen diese auf der Grundlage des Feedbacks vor allem ihre therapeutischen Interventionen an, versuchen die Therapiebeziehung zu verbessern und ziehen zusätzliche Hilfen zu Rate (Atzil-Slonim et al., 2015). Darüber hinaus zeigen sich hohe Korrelationen zwischen der Häufigkeit mit der das Feedback genutzt wird und der wahrgenommenen Nützlichkeit des Feedbacks. Dieser Zusammenhang betont die besondere Rolle, die der Einstellung der Therapeuten gegenüber psychometrischen Erhebungen und Feedback zukommt, die in verschiedenen aktuellen Studien zur Wirkung von und Einstellung gegenüber Rück-

Tabelle 2. Korrelationen der Häufigkeiten mit denen Therapeuten im Mittel die Nutzungsmöglichkeiten angewendet haben sowie mittlere Nützlichkeitseinschätzungen und Einschätzung der Auswirkung auf die Therapiebeziehung

Ich habe aufgrund der Rückmeldungen ...	keine	Interventionen angepasst	Beziehung verbessert	Therapieende vorbereitet	Ressourcen gefördert	Motivation gefördert	interpers. Probleme besprochen	diese mit dem Pat. besprochen	neue Hausaufgaben erprobt	zusätzliche Hilfe hinzugezogen	Sitzungsintervalle variiert	Sonstiges	Nützlichkeit Insgesamt
Interventionen angepasst	-.684***A												
Beziehung verbessert	-.457***A	.387**A											
Therapieende vorbereitet	-.360**A	.457***A	.481***A										
Ressourcen gefördert	-.473***A	.550***A	.468***A	.429***A									
Motivation gefördert	-.355**A	.385**A	.546***A	.548***A	.645***A								
interpersonale Probleme besprochen	-.411***A	.518***A	.478***A	.590***A	.700***A	.555***A							
Diese mit dem Pat. besprochen	-.327**A	.319**A	.259*A	.516***A	.430***A	.473***A	.447***A						
neue Hausaufgaben erprobt	-.139[A]	.313**A	.348**A	.394**A	.521***A	.333**A	.586***A	.332**A					
zusätzliche Hilfe hinzugezogen	-.417***A	.270*A	.236*A	.273*A	.224[A]	.178[A]	.293*A	.213[A]	.317**A				
Sitzungsintervalle variiert	-.293*A	.347**A	.365**A	.464***A	.316**A	.172[A]	.479***A	.189[A]	.597***A	.477***A			
Sonstiges	-.132[A]	-.016[A]	.054[A]	-.094[A]	-.020[A]	-.082[A]	.044[A]	.271*A	.067[A]	.061[A]	-.016[A]		
Nützlichkeit Insgesamt	-.616***B	.553***B	.440**B	.363**B	.436**B	.332*B	.520***B	.518***B	.330*B	.296*B	.358**B	.209[B]	
Auswirkung auf Therapiebeziehung	-.467***B	.484***B	.308*B	.298*B	.277*B	.284*B	.517***B	.383**B	.400**B	.371**B	.240[B]	.122[B]	.677***B

Anmerkungen: [A]$N = 72$; [B]$N = 52$; * $p < 0{,}05$; ** $p < 0{,}01$; *** $p < 0{,}001$.

Tabelle 3. Patienten- und Therapeuteneffekte in Bezug auf die Nutzung psychometrischen Feedbacks

Ich habe aufgrund der Rückmeldungen ...	PE	TE	AIC-Diff.	χ^2-*Differenz*
... keine Veränderungen meiner Behandlung vorgenommen.	51,40 %	49,60 %	89,83	91,85***
... meine therapeutischen Interventionen versucht anzupassen.	59,04 %	40,96 %	106,83	108,79***
... versucht, die therapeutische Beziehung zu verbessern.	72,63 %	27,37 %	43,51	45,53***
... das Therapieende vorbereitet.	55,47 %	44,53 %	86,53	88,51***
... die Ressourcen des Patienten / der Patientin versucht zu fördern	59,57 %	40,43 %	88	90.02***
... die Therapiemotivation des Patienten / der Patientin versucht zu erhöhen	62,55 %	37,45 %	44,55	46,60***
... die interpersonalen Probleme des Patienten / der Patientin besprochen	65,90 %	34,10 %	65,41	67,44***
... die Ergebnisse in den Fragebögen mit dem Patienten / der Patientin besprochen.	54,60 %	45,40 %	111,33	113,32***
... neue Hausaufgaben erprobt.	47,73 %	52,26 %	35,84	37,85***
... zusätzliche Hilfe hinzugezogen (Intervision / Supervision, Literatur, Weiterbildung, etc.).	57,20 %	42,80 %	99,24	101,22***
... die Sitzungsintervalle variiert.	55,34 %	44,66 %	46,46	48,50***
... Sonstiges.	65,59 %	34,41 %	0,88	2,91
Wie hilfreich waren für Sie – im Hinblick auf diesen Patienten – die Informationen, die sie aus den Status- und Verlaufsrückmeldungen erhalten haben?	60,53 %	39,47 %	58,571	59,23***
Haben sich die grafischen Rückmeldungen auf die Beziehung zum Patienten ausgewirkt?	65,83 %	34,17 %	68,665	69,51***

Anmerkungen: PE = Patienteneffekt; TE = Therapeuteneffekt; AIC-Diff. = Differenz der AIC-Werte zwischen einem Modell mit und ohne Therapeutenunterschieden (positive Werte indizieren eine Überlegenheit des Modells mit Therapeutenunterschieden). * $p < 0{,}05$; ** $p < 0{,}01$; *** $p < 0{,}001$.

meldesystemen propagiert wurde (Simon et al., 2012; De Jong et al., 2012; Lutz, Rubel et al., 2015). Diese spiegelt sich auch in einem hohen Therapeuteneffekt in Bezug auf die spezifische Nutzung des Feedbacks und der wahrgenommenen Nützlichkeit dessen, wider. Diese Therapeutenunterschiede liegen mit ca. 27–52 % deutlich über den Werten, die man in Bezug auf das Therapieergebnis (ca. 5–8 %; Baldwin & Imel, 2013; Schiefele et al., 2016), die Therapiedauer (ca. 9 %; Lutz, Rubel et al., 2015), die Abbruchrate (ca. 6–13 %; Saxon, Barkham, Foster & Parry, 2016; Zimmermann, Rubel, Page & Lutz, 2017) oder der Rate von Verschlechterungen (ca. 10 %; Saxon, Barkham et al., 2016) findet. Die Entscheidung ob und wie das Feedback genutzt wird, wird folglich nur zu einem gewissen Teil fallorientiert entschieden. Einen großen Anteil nehmen bei diesem Entscheidungsprozess auch therapeutenspezifische Aspekte ein. Diese Therapeutenunterschiede können zukünftiger Forschung als potenzielle Erklärungen dafür dienen, warum nur für einen Teil der Therapeuten positive Feedbackeffekte nachweisbar waren (Simon et al., 2012). Vor diesem Hintergrund könnte es möglich werden zu identifizieren welche Nutzungsstrategien derjenige Teil der Therapeuten anwendet, der Feedback erfolgreich einsetzt.

In Bezug auf die konkrete Nutzung gibt es derzeit lediglich eine deutsche Studie mit der die vorliegenden Befunde verglichen werden können. In einem Modellprojekt der Techniker Krankenkasse (TK) zum Qualitätsmonitoring wurde den Therapeuten der gleiche Katalog an Nutzungsmöglichkeiten vorgelegt wie in der gegenwärtigen Studie (Lutz et al., 2011; Lutz et al., 2013). Vergleicht man die jeweilige Nutzungshäufigkeit mit denen, die im TK-Projekt berichtet werden, zeigt sich für einen Großteil der Optionen eine weitgehende Übereinstimmung der Antworten in den unterschiedlichen Stichproben. Substanzielle Unterschiede zeigen sich nur in fünf der elf Nutzungsmöglichkeiten. Insgesamt gibt es nur einen geringeren Teil der Ausbildungstherapeuten, der angibt keine Modifikationen vorgenommen zu haben (ca. 19 % vs. 30 %). Ausbildungstherapeuten ziehen häufiger zusätzliche Hilfen wie Supervision heran (ca. 31 % vs. 8 %), versuchen häufiger ihre therapeutischen Interventionen anzupassen (ca. 52 % vs. 31 %) und versuchen häufiger die Therapiebeziehung zu verbessern (ca. 31 % vs. 9,9 %). Dem hingegen bespricht ein deutlich höherer Anteil an Therapeuten im TK-Projekt, als Reaktion auf das Feedback, die Fragebogenergebnisse mit den Patienten (ca. 67 % vs. 28 %).

Ein erster Erklärungsansatz für die Nutzungsunterschiede sind die Erfahrungsunterschiede der Therapeuten. Die TK Studie untersuchte approbierte Therapeuten, wohingegen in der vorliegenden Studie Ausbildungskandidaten betrachtet wurden. Erfahrenere Therapeuten könnten ein größeres Vertrauen in die eigenen diagnostischen und therapeutischen Fähigkeiten haben, weshalb Feedback das klinische Urteil dieser Therapeuten nur schwerer beeinflussen kann. Es gibt jedoch Befunde die zeigen, dass dieses klinische Urteil gerade in Bezug auf die Identifikation und Prognose negativer Entwicklungen bei erfahrenen Therapeuten mangelhaft ist (Hannan et al., 2005; Hatfield et al., 2010).

Neben der Erfahrung liefert die unterschiedliche Geschlechterverteilung der beiden Studien eine mögliche

Antwort. Die erhöhte Feedbacknutzung zur Modifikation der Therapie könnte durch den deutlich höheren Anteil an weiblichen Therapeuten in der vorliegenden Studie (87,5 % vs. 54,5 % TK Studie) erklärt werden, da frühere Forschungsbefunde die erhöhte Feedback Nutzung von weiblichen Therapeuten belegt (De Jong et al., 2012).

Darüber hinaus lassen sich Abweichungen auf unterschiedliche therapeutische Ausrichtungen zurückführen. Die approbierten Therapeuten der TK Studie arbeiteten auf der Basis der drei Richtlinienverfahren Verhaltenstherapie, tiefenpsychologisch fundierte Psychotherapie und analytische Psychotherapie. Die Ausbildungskandidaten der aktuellen Studie arbeiteten mit dem Schwerpunkt Verhaltenstherapie. Wenn davon auszugehen ist, dass die Verhaltenstherapie aufgrund ihrer wissenschaftlichen Wurzeln psychometrischem Feedback positiver eingestellt ist, so kann dies zusätzlich die erhöhte Feedbacknutzung in der vorliegenden Studie erklären. In diesem Kontext konnte in einer englischen Studie gezeigt werden, dass Therapeuten Feedback eher nutzen wenn die Therapeuten das Empfinden haben, dass das System mit ihrem therapeutischen Ansatz kompatibel ist (Lucock et al., 2015).

In einer amerikanischen Studie gaben über die Hälfte der Therapeuten (55.6 %) an, bei einer rückgemeldeten Verschlechterung des Patienten Psychopharmaka zu empfehlen (Hatfield et al., 2010). Darüber hinaus berichteten die meisten der Therapeuten bei einer Verschlechterung die Frequenz der Sitzungen zu erhöhen (33.3 %), mehr Informationen zu sammeln (33.3 %), die Verschlechterung mit dem Patienten zu besprechen (30.6 %), mit Kollegen in Austausch zu treten (27.8 %) sowie den Therapieansatz zu verändern (27.8). Auffallend ist die hohe Empfehlungsrate von Psychopharmaka, was durch kulturelle Unterschiede in der Versorgung psychisch kranker Menschen begründet sein könnte. Die vorliegende Studie kann dazu keine Aussage treffen, da dieser Aspekt nicht betrachtet wurde. Zusätzlich ähneln die amerikanischen Befunde denen der TK Studie mehr als den Ergebnissen der aktuellen Studie. Dies könnte mit der unterschiedlichen Zusammensetzung der Therapeutenstichproben erklärt werden. Im Gegensatz zur aktuellen Studie handelte es sich in der amerikanischen Studie und in der TK Studie um erfahrene Psychotherapeuten mit unterschiedlichen Therapierichtungen (Kognitive Verhaltenstherapie, Psychoanalyse & eklektische Psychotherapie). Welcher dieser konfundierten Aspekte in welchem Maß die beschriebenen Unterschiede bedingt, lässt sich in der vorliegenden Studie jedoch nicht differenzieren und sollte in zukünftiger Forschung untersucht werden.

Die Anforderungen, die im Therapieverlauf an einen Psychotherapeuten gestellt werden, sind höchst unterschiedlich, wobei ein adaptives und flexibles Reagieren auf die Bedürfnisse des Patienten notwendig ist, um den Therapieverlauf zu optimieren. Psychometrisches Feedback kann dabei helfen solche Veränderungen frühzeitig wahrzunehmen und entsprechende Anpassungen vorzunehmen (z. B. Lambert, 2007). Es kann als eine Art automatisierte, empirisch fundierte Unterstützung des Supervisionsprozesses verstanden werden. Die oben beschriebenen Ergebnisse weisen darauf hin, dass es Gemeinsamkeiten und Unterschiede zwischen Ausbildungskandidaten und berufserfahrenen Therapeuten im Umgang mit Feedback gibt. Insgesamt reagieren die beiden Gruppen auf eine ähnliche Art und Weise, indem das Feedback mit einem Experten besprochen wird. Die Ausbildungskandidaten wenden sich an einen erfahrenen Therapeuten, wohingegen erfahrene Therapeuten direkt mit ihren Patienten in Kontakt treten, die als Experten ihrer eigenen Situation angesehen werden können. Dies schreibt den Supervisoren in den Ausbildungsinstituten eine wichtige Rolle im Rahmen der Feedbacknutzung zu, da diese die weitere Anpassung aufgrund des Feedbacks stark beeinflussen. Diese Ergebnisse stimmen mit Befunden überein, die Supervision, neben der direkten klinischen Praxis mit den Patienten, als wichtigste Einflussquelle in der Entwicklung von Therapeuten betonen (Orlinsky, Botermans & Rønnestad, 2001). Zukünftige Forschung sollte an dieser Stelle ansetzen und die Rolle der Supervisoren einer gesonderten Analyse unterziehen. Eine Limitation des vorliegenden Datensatzes ist es, dass alle Therapeuten von mehreren Supervisoren supervidiert wurden. Es liegt also keine eindeutige Hierarchisierung der Daten im Sinne eines 3-Ebenen Modells (Patienten geschachtelt in Therapeuten, geschachtelt in Supervisoren) vor, die es ermöglichen würde den Einfluss der Supervisoren auf die Feedbacknutzung zu quantifizieren. Auch wenn eine solche eindeutige Zuordnung im Sinne der Analyse von Supervisoreneffekten wünschenswert wäre, ist sie praktisch kaum zu realisieren. Therapeuten müssen im Laufe ihrer Ausbildung von mindestens drei verschiedenen Supervisoren supervidiert werden. Diese Vorgabe verhindert es im Ausbildungskontext ausreichend große Datensätze zu sammeln, in denen jeweils nur die Fälle eines Therapeuten bei einem spezifischen Supervisor inkludiert wären. Zudem würden selbst dann „Spill-over" Effekte durch die Supervision bei anderen, nicht eingeschlossenen Supervisoren, den Therapeuteneinfluss wahrscheinlich über- und den Supervisoreneinfluss unterschätzen.

Eingedenk der gefundenen differentiellen Effekte von Feedback sollte die Implementierung von Feedbacksystemen stets gut vorbereitet und mit entsprechenden Therapeutenschulungen zum Umgang mit Feedback einhergehen. Die Ergebnisse weisen zusätzlich darauf hin, dass neben den Therapeuten auch die Supervisoren eine Schu-

lung erhalten sollten, um den angemessenen Umgang mit Feedback entsprechend an ihre Ausbildungskandidaten weiter geben zu können. Nur wenn Therapeuten und Supervisoren Feedback als integralen Bestandteil ihrer Praxis und als empirisch gestützte Erweiterung des Supervisionsprozesses verstehen und nicht als oktroyierten Fremdkörper, können sie dieses gewinnbringend einsetzen und ihre Patienten glaubhaft vom zusätzlichen Nutzen überzeugen (Strauss et al., 2015).

Limitationen

Die Generalisierbarkeit und die Implikationen der vorliegenden Studie werden durch bestimmte Faktoren eingeschränkt. Zum einen ist es durch das Fehlen einer Kontrollgruppe, für welche Verlaufserhebungen vorgenommen werden, diese jedoch nicht an die Therapeuten rückgemeldet werden, nicht möglich den Einfluss der jeweiligen Modifikation auf den Feedbackeffekt zu quantifizieren. Der Feedbackeffekt zielt spezifisch auf Patienten ab, die sich zunächst nicht verbessern oder sogar verschlechtern (z. B. Lambert, 2007). Ohne eine wie oben beschriebene Kontrollgruppe ist es jedoch nicht möglich solche Patienten zu identifizieren, die sich zunächst nicht verbessern. Bisherige Forschung hat bisher immer einen dieser beiden Aspekte vernachlässigt. Entweder es lag keine adäquate Kontrollgruppe vor, die eine Abschätzung des Feedbackeffektes möglich macht oder es wurde nicht erfasst was Therapeuten konkret mit dem Feedback gemacht haben. Zukünftige Studien sollten beide dieser Aspekte berücksichtigen, um die erfolgreichsten Strategien der Feedbacknutzung identifizieren zu können. Dadurch könnten nicht nur für die Ergebnisforschung relevante Erkenntnisse generiert werden. Auch die Auswirkungen der Feedbacknutzung auf den Therapieprozess könnten auf diese Weise untersucht werden. Aus den vorliegenden Analysen wissen wir, dass die Therapeuten der untersuchten Stichprobe im Schnitt eher positive Auswirkungen der Rückmeldungen auf die Therapiebeziehung wahrnehmen und das Feedback sogar oftmals zum Anlass nehmen die Allianz zu verbessern. Zukünftige Studien sollten mittels Stundenbögen, welche unterschiedliche therapeutische Wirkfaktoren erfassen (z. B. Mander et al., 2015; Rubel, Rosenbaum & Lutz, 2017), untersuchen wie sich die unterschiedliche Feedbacknutzung zum Beispiel auf die vom Patienten wahrgenommene Beziehung niederschlägt (Falkenström, Finkel, Sandell, Rubel & Holmqvist, 2017).

Darüber hinaus arbeiteten alle in diesem Datensatz untersuchten Therapeuten nach dem gleichen therapeutischen Ansatz (VT) und absolvieren ihre Ausbildung an einem universitären Institut. Sowohl zwischen verschiedenen therapeutischen Ansätzen, als auch innerhalb der verhaltenstherapeutischen Ausbildungsinstitute ist eine große Heterogenität in Bezug auf die Einstellung gegenüber qualitätssichernder Maßnahmen denkbar. Für einen Ausbildungskandidaten im Bereich der tiefenpsychologisch fundierten Therapie nimmt ein symptomorientiertes Verlaufsfeedback, vor dem Hintergrund der unterschiedlichen Störungstheorien, wahrscheinlich einen anderen Stellenwert ein als für einen Verhaltenstherapeuten. Durch diese Selbstselektionsprozesse könnten in der untersuchten Therapeutenstichprobe vor allem solche Therapeuten enthalten sein, die eine größere Affinität gegenüber Forschung und qualitätssichernder Maßnahmen haben. Die Ergebnisse sollten daher nur mit Vorsicht auf andere Schulen und Ausbildungskontexte übertragen werden.

Eine weitere Einschränkung stellt die Quelle dar, aus der sich die Auskunft über die Feedbacknutzung in der vorliegenden Studie speist. Mögliche soziale Erwünschtheitserwartungen seitens der Ausbildungsteilnehmer könnten die Auskünfte bezüglich der Feedbacknutzung verzerrt haben. In zukünftigen Studien sollte verstärkt auf Maße zurückgegriffen werden, die weniger anfällig für solche Effekte sind. In einem Kontext in dem Therapeuten das Feedback über ein Onlineportal zur Verfügung gestellt bekommen, stellen die Zeiten/Dauern zu denen Therapeuten die Feedbackreports der jeweiligen Patienten anschauen, ein Beispiel für ein solches Maß dar.

In diesem Zusammenhang sollte weiterhin beachtet werden, dass die Items, welche zur Abfrage der Feedbacknutzung und der Einstellung gegenüber Qualitätssichernder Maßnahmen eingesetzt wurden, nicht Teil eines psychometrisch validierten Fragebogens waren. Für diese Zwecke liegen bislang keine standardisierten Maße vor. Die Items wurden daher in Anlehnung an die gestellten Fragen im TK-Projekt ausgewählt. Dieses Vorgehen erlaubt zwar einen besonders guten Vergleich mit den Ergebnissen der TK-Studie, macht es jedoch schwierig die Ergebnisse mit anderen Datensätzen zu vergleichen.

Darüber hinaus muss die Schätzung der Therapeuteneffekte mit einer gewissen Vorsicht interpretiert werden. Für die vorgestellten Analysen wurden nur Therapeuten eingeschlossen, die mindestens fünf Patienten im Datensatz haben. Fünf Patienten pro Therapeut ist zwar bereits eine untere Grenze, die in vielen früheren Studien zu Therapeuteneffekten nicht erreicht werden konnte (Baldwin & Imel, 2013), trotzdem sind für eine möglichst genaue Schätzung solcher Effekte oftmals noch größere Datensätze notwendig. Aktuelle Studien zeigen jedoch, dass die Punktschätzung des Therapeuteneffektes mit Datensätzen, wie dem in der vorliegenden Studie genutzten, relativ genau möglich ist (Schiefele et al., 2016).

Fazit

Die Nutzung psychometrischer Rückmeldungen hängt stark von den behandelnden Therapeuten und deren Einstellung gegenüber solchen qualitätssichernden Maßnahmen ab. Bereits in der Ausbildung angehender Therapeuten sollte die Wichtigkeit solcher Maßnahmen als Ergänzung des klinischen Eindrucks betont werden, um kontinuierliche Qualitätssicherung zum integralen Bestandteil der persönlichen Praxis zu machen.

Literatur

Amble, I., Gude, T., Stubdal, S., Andersen, B. J. & Wampold, B. E. (2015). The effect of implementing the Outcome Questionnaire-45.2 feedback system in Norway: A multisite randomized clinical trial in a naturalistic setting. *Psychotherapy Research, 25,* 669 – 677.

Atzil-Slonim, D., Bar-Kalifa, E., Rafaeli, E., Lutz, W., Rubel, J., Schiefele, A. K. et al. (2015). Therapeutic bond judgments: Congruence and incongruence. *Journal of Consulting and Clinical Psychology, 83,* 773 – 784.

Bar-Kalifa, E., Atzil-Slonim, D., Rafaeli, E., Peri, T., Rubel, J. & Lutz, W. (2016). Therapist-client agreement in assessments of clients' functioning. *Journal of Consulting and Clinical Psychology, 84,* 1127 – 1134.

Baldwin, S. A. & Imel, Z. E. (2013). Therapist effects: Findings and methods. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 258 – 297). New York, NY: John Wiley & Sons, Inc.

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R package version, 1* (7), 1 – 23.

Bickman, L., Kelley, S. D., Breda, C., Andrade, A. R. de & Riemer, M. (2011). Effects of routine feedback to clinicians on mental health outcomes of youths: Results of a randomized trial. *Psychiatric Services, 62,* 1423 – 1429.

Byrne, S. L., Hooke, G. R., Newnham, E. A. & Page, A. C. (2012). The effects of progress monitoring on subsequent readmission to psychiatric care: A six-month follow-up. *Journal of Affective Disorders, 137,* 113 – 116.

Castonguay, L. G., Barkham, M., Lutz, W. & McAleavey, A. A. (2013). Practice-oriented research: Approaches and applications. In M. J. Lambert (Ed.). *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 85 – 133). New York, NY: John Wiley & Sons, Inc.

De Jong, K., Sluis, P. van, Nugter, M. A., Heiser, W. J. & Spinhoven, P. (2012). Understanding the differential impact of outcome monitoring: Therapist variables that moderate feedback effects in a randomized clinical trial. *Psychotherapy Research, 22,* 464 – 474.

De Jong, K., Timman, R., Hakkaart-Van Roijen, L., Vermeulen, P., Kooiman, K., Passchier, J. et al. (2014). The effect of outcome monitoring feedback to clinicians and patients in short and long-term psychotherapy: A randomized controlled trial. *Psychotherapy Research, 24,* 629 – 639.

Douglas, S. R., Jonghyuk, B., Andrade, A. R. V. de, Tomlinson, M. M., Hargraves, R. P. & Bickman, L. (2015). Feedback mechanisms of change: How problem alerts reported by youth clients and their caregivers impact clinician-reported session content. *Psychotherapy Research, 25,* 678 – 693.

Dyer, K., Hooke, G. R. & Page, A. C. (2016). Effects of providing domain specific progress monitoring and feedback to therapists and patients on outcome. *Psychotherapy Research, 26,* 297 – 306.

Falkenström, F., Finkel, S., Sandell, R., Rubel, J. A. & Holmqvist, R. (2017). Dynamic models of individual change in psychotherapy process research. *Journal of Consulting and Clinical Psychology, 85,* 537 – 549.

Flückiger, C., Regli, D., Zwahlen, D., Hostettler, S. & Caspar, F. (2010). Der Berner Patienten- und Therapeutenstundenbogen 2000. *Zeitschrift für Klinische Psychologie und Psychotherapie, 39,* 71 – 79.

Franke, G. H. (1995). *SCL-90-R. Die Symptom-Checkliste von Derogatis – Deutsche Version.* Göttingen: Beltz Test.

Franke, G. H. (2000). *Brief symptom inventory (BSI) von LR Derogatis* (Kurzform der SCL-90-R). Göttingen: Hogrefe.

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K. et al. (2005). A lab test and an algorithm for identifying patients at risk for treatment failure. *Journal of Clinical Psychology, 61,* 155 – 163.

Hatfield, D., McCullough, L., Frantz, S. H. B. & Krieger, K. (2010). Do we know when our clients get worse? An investigation of therapists'ability to detect negative client change. *Clinical Psychology and Psychotherapy, 17,* 25 – 32.

Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z. & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist, 51,* 1059 – 1064.

Hox, J. J., Moerbeek, M. & Schoot, R. van de (2010). *Multilevel analysis: Techniques and applications.* New York: Routledge.

Köck, K. & Lutz, W. (2012). Therapieverlaufs- und Ergebnismessung im Rahmen einer patientenorientierten Psychotherapieforschung. *PiD – Psychotherapie im Dialog, 13* (01), 10 – 16.

Krägeloh, C. U., Czuba, K. J., Billington, D. R., Kersten, P. & Siegert, R. J. (2015). Using feedback from patient-reported outcome measures in mental health services: A scoping study and typology. *Psychiatric Services*, *66,* 224 – 241.

Lambert, M. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research, 17,* 1 – 14.

Lambert, M. J. (2013). The Efficacy and Effectiveness of Psychotherapy. In M. J. Lambert (Ed.). *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 169 – 218). New York, NY: John Wiley & Sons, Inc.

Lambert, M. J., Bailey, R. J., Kimball, K. & Shimokawa, K. (2007). *Clinical Support Tools Manual – Brief Version – 40:* OQ-Measures LLCC.

Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L. & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice, 10,* 288 – 301.

Lucock, M., Halstead, J., Leach, C., Barkham, M., Tucker, S., Randal, C. et al. (2015). A mixed-method investigation of patient monitoring and enhanced feedback in routine practice: Barriers and facilitators. *Psychotherapy Research, 25,* 633 – 646.

Lutz, W., Böhnke, J. R., Köck, K. & Bittermann, A. (2011). Diagnostik und psychometrische Verlaufsrückmeldungen im Rahmen eines Modellprojektes zur Qualitätssicherung in der ambulanten Psychotherapie. *Zeitschrift für Klinische Psychologie und Psychotherapie, 40,* 283 – 297.

Lutz, W., De Jong, K. & Rubel, J. (2015). Patient-focused and feedback research in psychotherapy: Where are we and where do we want to go? *Psychotherapy Research: Journal of the Society for Psychotherapy Research, 25,* 625 – 632.

Lutz, W., Hofmann, S. G., Rubel, J., Boswell, J. F., Shear, M. K., Gorman, J. M. et al. (2014). Patterns of early change and their relationship to outcome and early treatment termination in pa-

tients with panic disorder. *Journal of consulting and clinical psychology, 82,* 287.

Lutz, W., Leon, S. C., Martinovich, Z., Lyons, J. S. & Stiles, W. B. (2007). Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology, 54,* 32 – 39.

Lutz, W. & Rubel, J. (2015). Warum die Psychotherapieforschung in Deutschland ausgebaut werden sollte/könnte/müsste. *PPmP-Psychotherapie· Psychosomatik· Medizinische Psychologie, 65,* 161 – 162.

Lutz, W., Rubel, J., Schiefele, A.-K., Zimmermann, D., Böhnke, J. R. & Wittmann, W. W. (2015). Feedback and Therapist effects in the context of treatment outcome and treatment length. *Psychotherapy Research, 25,* 647 – 660.

Lutz, W., Tholen, S., Schürch, E. & Berking, M. (2006). Die Entwicklung, Validität und Reliabilität von Kurzformen gängiger psychometrischer Instrumente zur Evaluation des therapeutischen Fortschrittes in Psychotherapie und Psychiatrie: *Diagnostica, 52,* 11 – 25.

Lutz, W., Wittmann, W., Böhnke, J., Rubel, J. & Steffanowski, A. (2013). Das Modellprojekt der TK zum Qualitätsmonitoring in der ambulanten Psychotherapie aus Sicht der Evaluatoren–Ein Plädoyer für mehr Psychotherapieforschung in Deutschland. *PPmP – Psychotherapie· Psychosomatik· Medizinische Psychologie, 63,* 225 – 228.

Mander, J., Schlarb, A., Teufel, M., Keller, F., Hautzinger, M., Zipfel, S. et al. (2015). The individual therapy process questionnaire: Development and validation of a revised measure to evaluate general change mechanisms in psychotherapy. *Clinical Psychology & Psychotherapy, 22,* 328 – 345.

Miller, S. D., Duncan, B. L., Brown, J., Sorrell, R. & Chalk, M. B. (2006). Using formal client feedback to improve retention and outcome: Making ongoing, real-time assessment feasible. *Journal of Brief Therapy, 5* (1), 5 – 22.

Newnham, E. A. & Page, A. C. (2010). Bridging the gap between best evidence and best practice in mental health. *Clinical Psychology Review, 30,* 127 – 142.

Orlinsky, D. E, Botermans, J.-F. & Rønnestad, M. H. (2001). Towards an empirically grounded model of psychotherapy training: Four thousand therapists rate influences on their development. *Australian Psychologist, 36,* 139 – 148.

Poston, J. M. & Hanson, W. E. (2010). Meta-analysis of psychological assessment as a therapeutic intervention. *Psychological Assessment, 22,* 203 – 212.

Probst, T., Lambert, M. J., Loew, T. H., Dahlbender, R. W., Göllner, R. & Tritt, K. (2013). Feedback on patient progress and clinical support tools for therapists: Improved outcome for patients at risk of treatment failure in psychosomatic in-patient therapy under the conditions of routine practice. *Journal of Psychosomatic Research, 75,* 255 – 261.

R Core Team (2016). *R: A language and environment for statistical computing.* Wien: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Rubel, J., Lutz, W., Kopta, S. M., Köck, K., Minami, T., Zimmermann, D. et al. (2015). Defining early positive response to psychotherapy: An empirical comparison between clinically significant change criteria and growth mixture modeling. *Psychological assessment, 27,* 478 – 488.

Rubel, J. A., Rosenbaum, D. & Lutz, W. (2017). Patients' in-session experiences and symptom change: Session-to-session effects on a within- and between-patient level. *Behaviour Research and Therapy, 90*, 58 – 66.

Sapyta, J., Riemer, M. & Bickman, L. (2005). Feedback to clinicians: Theory, research, and practice. *Journal of Clinical Psychology, 61,* 145 – 153.

Saxon, D. & Barkham, M. (2012). Patterns of therapist variability: therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology, 80,* 535 – 546.

Saxon, D., Barkham, M., Foster, A. & Parry, G. (2016). The Contribution of Therapist Effects to Patient Dropout and Deterioration in the Psychological Therapies. *Clinical Psychology & Psychotherapy, 24, 575* – 588.

Schiefele, A. K., Lutz, W., Barkham, M., Rubel, J., Böhnke, J., Delgadillo, J. et al. (2016). Reliability of therapist effects in practice-based psychotherapy research: A guide for the planning of future studies. *Administration and Policy in Mental Health and Mental Health Services Research.* Advance online publication. https://doi.org/10.1007/s10488-016-0736-3

Shimokawa, K., Lambert, M. J. & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology, 78,* 298 – 311.

Simon, W., Lambert, M. J., Busath, G., Vazquez, A., Berkeljon, A., Hyer, K. et al. (2013). Effects of providing patient progress feedback and clinical support tools to psychotherapists in an inpatient eating disorders treatment program: A randomized controlled study. *Psychotherapy Research, 23,* 287 – 300.

Simon, W., Lambert, M. J., Harris, M. W., Busath, G. & Vazquez, A. (2012). Providing patient progress information and clinical support tools to therapists: Effects on patients at risk of treatment failure. *Psychotherapy Research, 22,* 638 – 647.

Strauss, B. M., Lutz, W., Steffanowski, A., Wittmann, W. W., Boehnke, J. R., Rubel, J. et al. (2015). Benefits and challenges in practice-oriented psychotherapy research in Germany: The TK and the QS-PSY-BAY projects of quality assurance in outpatient psychotherapy. *Psychotherapy Research, 25,* 32 – 51.

Walfish, S., McAlister, B., O'Donnell, P. & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports, 110,* 639 – 644.

Whipple, J. L., Lambert, M. J., Vermeersch, D. A., Smart, D. W., Nielsen, S. L. & Hawkins, E. J. (2003). Improving the effects of psychotherapy: The use of early identification of treatment and problem-solving strategies in routine practice. *Journal of Counseling Psychology, 50* (1), 59 – 68.

White, M. M., Lambert, M. J., Ogles, B. M., Mclaughlin, S. B., Bailey, R. J. & Tingey, K. M. (2015). Using the Assessment for Signal Clients as a feedback tool for reducing treatment failure. *Psychotherapy Research, 25,* 724 – 734.

Wittchen, H. U., Zaudig, M. & Fydrich, T. (1997). *Skid. Strukturiertes klinisches Interview für DSM-IV. Achse I: Psychische Störungen.* Göttingen: Hogrefe.

Worthen, V. E. & Lambert, M. J. (2007). Outcome oriented supervision: Advantages of adding systematic client tracking to supportive consultations. *Counselling and Psychotherapy Research*, *7* (1), 48 – 53.

Zimmermann, D., Rubel, J., Page, A. C. & Lutz, W. (2017). Therapist Effects on and Predictors of Non-Consensual Dropout in Psychotherapy. *Clinical Psychology & Psychotherapy, 24,* 312 – 321.

Dr. Julian Rubel
Klinische Psychologie und Psychotherapie
Fachbereich I – Psychologie
Universität Trier
Am Wissenschaftspark 25 + 27
54296 Trier
rubel@uni-trier.de

6. Study II

Zimmermann, D., Rubel, J., Page, A. C., & Lutz, W. (2017). Therapist effects on and predictors of non-consensual dropout in psychotherapy. *Clinical Psychology & Psychotherapy, 24*(2), 312-321. doi: 10.1002/cpp.2022

Clinical Psychology and Psychotherapy
Clin. Psychol. Psychother. **24**, 312–321 (2017)
Published online 10 May 2016 in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/cpp.2022

Therapist Effects on and Predictors of Non-Consensual Dropout in Psychotherapy

Dirk Zimmermann,[1]* Julian Rubel,[1] Andrew C. Page[2] and Wolfgang Lutz[1]
[1] *Clinical Psychology and Psychotherapy, Department of Psychology, University of Trier, Crawley, Australia*
[2] *Clinical Psychology and Psychotherapy, Department of Psychology, University of Western Australia, Crawley, Australia*

Background: Whereas therapist effects on outcome have been a research topic for several years, the influence of therapists on premature treatment termination (dropout) has hardly been investigated. Since dropout is common during psychological treatment, and its occurrence has important implications for both the individual patient and the healthcare system, it is important to identify the factors associated with it.
Method: Participants included 707 patients in outpatient psychotherapy treated by 66 therapists. Multilevel logistic regression models for dichotomous data were used to estimate the impact of therapists on patient dropout. Additionally, sociodemographic variables, symptoms, personality style and treatment expectations were investigated as potential predictors.
Results: It was found that 5.7% of variance in dropout could be attributed to therapists. The therapist's effect remained significant after controlling for patient's initial impairment. Furthermore, initial impairment was a predictor of premature termination. Other significant predictors of dropout on a patient level were male sex, lower education status, more histrionic and less compulsive personality style and negative treatment expectations.
Conclusions: The findings indicate that differences between therapists influence the likelihood of dropout in outpatient psychotherapy. Further research should focus on variables, which have the potential to explain these inter-individual differences between therapists (e.g., therapist's experience or self-efficacy). Copyright © 2016 John Wiley & Sons, Ltd.

Key Practitioner Messages:
- There are substantial differences between therapists concerning their average dropout rates.
- At the patient level, higher initial impairment, male sex, lower education, less compulsive personality style, more histrionic personality style and low treatment expectations seem to be risk factors of non-consensual treatment termination.
- Psychometric feedback during the course of treatment should be used to identify patients who are at risk for dropout.

Keywords: Therapist Effects, Dropout, Mixed Effect Logistic Regression, Treatment Expectation, Personality Styles, Cognitive Behavioural Therapy

The effectiveness and cost-efficiency of psychological treatments have been documented in many randomized controlled trials as well as in effectiveness studies; however, some patients do not respond to psychotherapy, and others may even deteriorate over the course of treatment or drop out (Lambert & Ogles, 2004). Difficulties in identifying non-responders and deteriorators might result from premature discontinuation of treatment (dropout), which negatively affects data collection. Patients not progressing well early in treatment have the tendency to dropout (e.g., Lutz *et al.*, 2014). A recent meta-analysis focusing on psychotherapy patients examined dropouts amongst 83 834 adult patients. The authors calculated a mean dropout rate of 19.7%, which ranged between studies from 0 to 74.2% (Swift & Greenberg, 2012). Thus, dropout is a common phenomenon in psychotherapy.

Despite dropout's ubiquity and importance, there is little consensus about its definition. Some researchers use the therapist judgment of the appropriateness of termination to classify dropout. For example, Brogan, Prochaska, and Prochaska (1999) characterized discontinuing prior to 10 sessions against therapist's advice as dropout, whereas mutually agreed termination prior to 10 sessions was

*Correspondence to: Dirk Zimmermann, M.Sc., Clinical Psychology and Psychotherapy, Department of Psychology, University of Trier, D-54286 Trier, Germany
E-mail: zimmermann@uni-trier.de

defined as an appropriate termination. Similarly, Horner and Diamond (1996) defined dropout as termination initiated by the patient against the advice of the therapist. Other researchers have defined dropout as failure to attend a predefined treatment duration indicated by either a specific number of sessions or months (e.g., Beckham, 1992; Gunderson *et al.*, 1989). Others have used non-completion of the treatment protocol as a definition (e.g., Maher *et al.*, 2010). Again, other studies have defined dropout when patients missed two consecutive sessions or the last scheduled session (e.g., Kolb *et al.*, 1985; Hatchett, Han, & Cooker, 2002). As a result of these heterogeneous definitions, dropout rates vary between different studies, and findings have to be interpreted with caution and respect to the criteria used (Hatchett & Park, 2003). Each definition has advantages and disadvantages. Duration-based operationalization relies on a subjective interpretation of an appropriate therapy length. This might be clear in randomized clinical trials, where treatment length is usually defined within the treatment protocol but more difficult for naturalistic samples. In effectiveness studies with varying treatment lengths, it seems difficult to predefine a minimum number of sessions as a standard. Early-treatment responders might be classified as dropouts, while patients with no progress or deterioration and longer-term treatments might be classified as completers. In contrast to duration-based operationalization, the definition of dropout by therapist judgments is based on clinical decisions, which can also be biased. However, therapist judgments of dropout are face-valid and take the experience of the therapist with his or her patient into account (Garb, 2005). This definition is often used within naturalistic studies (e.g., Richards & Borglin, 2011; Self, Oates, Pinnock-Hamilton, & Leach, 2005) and is therefore used within the present study (refer below).

Several studies report an association between dropout and worse treatment outcomes (e.g., Delgadillo *et al.*, 2014; Richards & Borglin, 2011). For example, a follow-up study of day treatment programs showed that dropouts (defined as non-completers) had higher hospitalization rates than completers (Karterud *et al.*, 2003). Aside from these negative consequences for the individual patient, treatment dropout strains the health system on several levels (Barrett *et al.*, 2008). Sledge, Moras, Hartley, and Levine (1990) pointed out that patient dropouts could lead to an inefficient use of clinical treatment personnel. Moreover, considering that many patients conceivably continue to suffer from mental disorders after dropping out of treatment, dropout leads to an increase of mental health and social care costs and can reduce productivity as well as quality of life (Druss & Rosenheck, 1999; Sainsbury Centre for Mental Health, 2003).

Given the potential negative consequences associated with dropout, the identification of potential predictors and risk factors could be important in developing decision support tools to prevent dropout (e.g., Lambert, 2007; Lutz *et al.*, 2015). Swift and Greenberg (2012) found that age had a small effect of $d = 0.16$ in Cohen's terms (Cohen, 1988). Compared with older patients, younger patients showed an increased risk for dropout. Additionally, lower education was associated with more dropout with a small to medium effect ($d = 0.29$). No significant association could be found for gender, race, marital status or employment status. Several studies reported an impact of patients' diagnoses on dropout (McMurran, Huband, & Overton, 2010). It seems more likely that patients with a personality disorder drop out than patients without a personality disorder. Accordingly, Karterud *et al.* (2003) found that day-treatment patients fulfilling many personality disorder criteria had a higher probability for premature termination. Additionally, patients' treatment expectations have repeatedly been shown to influence the likelihood of dropout (e.g., Aubuchon-Endsley & Callahan, 2009; Greenberg, Constantino, & Bruce, 2006). Furthermore, several studies report a positive association between initial severity of distress or low global functioning levels and dropout rate (e.g., Bower *et al.*, 2013; Karterud *et al.*, 2003; Ogrodniczuk *et al.*, 2008).

While patient variables which influence the likelihood of dropout have been the focus of many studies, research on therapist characteristics connected to dropout is rare. This lack of research on therapist effects on dropout is stunning given the increasing interest in therapist differences in outcome (e.g., Baldwin & Imel, 2013; Crits-Christoph *et al.*, 1991; Kim, Wampold, & Bolt, 2006; Lutz *et al.*, 2007; Saxon & Barkham, 2012). Therapist effects can be defined as the proportion of variance in the outcome, which can be explained by therapist differences (Baldwin & Imel, 2013). On average, about 5–8% of differences in patient outcomes have been able to be attributed to differences between therapists (e.g., Baldwin & Imel, 2013). Swift and Callahan (2011) showed that patients who were educated about the positive relationship between the number of psychotherapy sessions and the probability of recovery (i.e., the dose–effect model; Howard, Kopta, Krause, & Orlinsky, 1986) stayed longer in treatment and had lower dropout rates. This result suggests that therapists' effects could influence patients' dropout rate even though the therapist effect on dropout was not directly investigated within this study. One methodological confound has, however, typically been ignored in dropout research: the nested data structure. Ignoring the nested data structure (i.e., patients are nested within therapists) might lead to biased parameter estimates (Hox, 2010). So far, only one study systematically investigated therapist effects on dropout rates with multilevel models (Huppert *et al.*, 2014).

In a sample of 350 patients with panic disorder receiving cognitive behavioural therapy from 17 therapists, Huppert *et al.* (2014) found no significant differences in

dropout rates between therapists. It is, however, premature to accept the conclusion that there are no therapist effects on dropout, given the very small number of therapists included in the study and the marginal significance ($p = 0.07$) of therapist effects. To address these issues, the present study investigates two research questions in a large database ($n = 707$) with multilevel logistic regression models: (1) Do therapist effects on dropout in psychotherapy exist, and if yes, is the magnitude comparable to therapist effects on outcome ranging from 5 to 8%? (2) Which patient variables can predict dropout in outpatient psychotherapy? Based on past research on predictors of dropout, we expect to find a significant influence of age and education on dropout. Younger patients and patients with lower education are expected to drop out more likely. No significant relations are expected for other sociodemographic variables like gender, marital status or employment status. Furthermore, we assume that interpersonal problems, personality styles as well as treatment expectations are associated with dropout. Patients having more interpersonal problems or personality styles in the clinical range, as well as patients with low treatment expectations, are expected to have higher dropout rates.

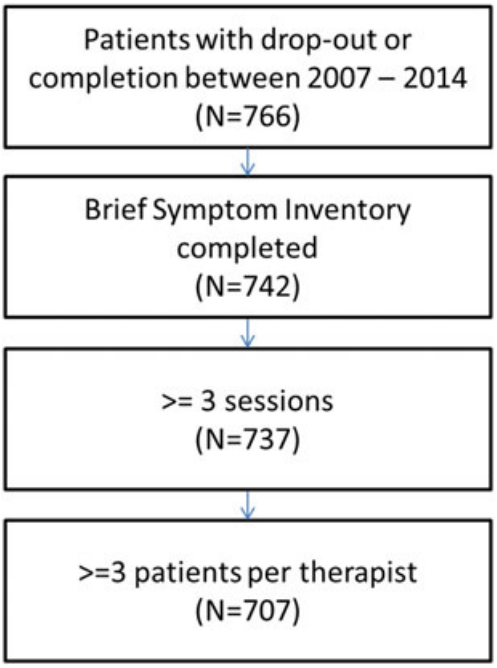

Figure 1. Flowchart of selected patients for this study. [Colour figure can be viewed at wileyonlinelibrary.com]

METHODS

Patients

The analyses were based on a sample that comprised 707 patients treated at the Trier University Outpatient Clinic in Southwest Germany between 2007 and 2014. The patients were included if they completed the Brief Symptom Inventory (BSI) at intake and attended at least an intake interview, a structured diagnostic interview (Wittchen, Wunderlich, Gruschwitz, & Zaudig, 1997) and the first treatment session. Additionally, only those patients whose therapists had seen at least three patients in the data set were analysed (refer to Figure 1) to provide a minimum reliability in estimating therapists' mean dropout rates (Baldwin & Imel, 2013). The mean age of the patients was 35.93 ($SD = 12.7$), of whom 63.4% were women. Based on structured clinical interviews for Diagnostic and Statistical Manual of Mental Disorders fourth version (DSM-IV) (Wittchen *et al.*, 1997), 97.0% ($n = 686$) had at least one diagnosis. Of all patients with a diagnosis, 42.3% ($n = 299$) had a major depressive disorder as the primary diagnosis, 5.7% ($n = 40$) had a dysthymic disorder, 12.7% ($n = 90$) had an adjustment disorder, 15.7% ($n = 111$) had an anxiety disorder, 2.1% ($n = 15$) had an eating disorder, and the remainder (18.5%; $n = 131$) had other diagnoses. While 26.7% ($n = 189$) of the patients had only one diagnosis, the majority (70.3%; $n = 497$) had two or more comorbid diagnoses. The International Diagnostic Checklist for Personality Disorders (Bronisch, Hiller, Mombour, & Zaudig, 1996) was adopted for the diagnosis of personality disorders. One-hundred and twenty-seven (18.0%) patients fulfilled the criteria for a personality disorder. Of the 705 (99.7%) patients for which marital status was available, 398 (56.5%) were single, 195 (27.7%) were married, 103 (14.6%) were divorced or lived alone, and nine (1.3%) were widowed. In terms of education, 172 (24.3%) patients had attended secondary modern school, 202 (28.6%) had attended junior high school, 293 (41.4%) had general qualification for university entrance, 14 (2.0%) were still in school, and 24 (3.4%) had no graduation. One-hundred and fourteen (16.1%) patients were not working because of sick leave at the beginning of treatment (information missing for $n = 14$).

Therapists

The therapist sample consisted of 66 individuals. The mean number of the patients per therapist was 10.71 ($SD = 6.32$), ranging from 3 to 30. All the therapists in this study participated in a 3-year (full-time) or 5-year (part-time) post-graduate clinical training program with a cognitive behavioural therapy focus. All the therapists held their master's degree in clinical psychology. They had at least 1 year of training before seeing their first patient in the clinic. Fifty-six (84.8%) therapists were women, and the average age at the beginning of the study was 29.44 years ($SD = 5.65$).

Treatment

The therapists were supervised every four sessions by a senior professional therapist and were supported by a feedback system monitoring patient outcomes on a session-by-session basis (Lutz, 2002). Treatment sessions were typically provided on a weekly basis with a length of 50 min. The overall average treatment length for all 707 patients was 31.00 ($SD = 20.62$) sessions with $M = 39.53$ ($SD = 17.28$) for completers and $M = 13.06$ ($SD = 14.70$) sessions for the patients with no consensual termination.

MEASURES

Dropout

Dropout was assessed via the therapist's evaluation at the end of each treatment. If the ending was planned and consensual, therapy was considered as completed. Otherwise, if a patient stayed away from treatment against the recommendation of the therapist, treatment was considered as dropout. Examples for dropouts were the patient did not show up again and was not available for any further contact or the patient told the therapist that he or she was not interested in therapy anymore although the therapist advised continuation.

Initial impairment

Initial impairment was assessed using the BSI (Franke, 2000; German translation by Derogatis, 1975). This 53-item questionnaire is a short version of the Symptom Checklist-90-Revised and assesses nine subscales with the following dimensions: somatization, obsessive–compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation and psychoticism. Symptom statements are self-rated on a five-point Likert scale ranging from 0 (not at all) to 4 (extremely). For this study, the Global Severity Index (GSI) was used, which had a very good internal consistency ($\alpha = 0.96$)

Interpersonal problems

The amount of interpersonal distress was assessed using a German short version of the Inventory of Interpersonal Problems (Thomas, Brähler, & Strauß, 2011; German short version by Horowitz *et al.*, 1988). This measure consists of eight scales, each with four items asking about difficulties in different interpersonal situations. The scale ranges from 0 (no problems) to 4 (very strong problems). To calculate the level of interpersonal problems, the mean score over all the items was calculated, which showed a good internal consistency of $\alpha = 0.86$.

Treatment expectations

Two single items concerning treatment expectations were used in this study. At the beginning of therapy, the patients rated the item: 'How difficult will it be for you to attend psychotherapy?' on a six-point Likert scale ranging from 1 ('it will be very easy for me') to 6 ('it will be impossible'). The second item asked the patient: 'How confident are you that psychotherapy will help you in dealing with your problems?'. This item was answered on a four-point Likert scale ranging from 1 ('not at all confident') to 4 ('very confident').

Personality style

Personality style was assessed using the personality style and disorder inventory (PSSI-K; Kuhl & Kazén, 1997). This instrument is a short form of the PSSI (140 items) consisting of 54 items, belonging to 14 dimensions each measured with four items. The PSSI-K assesses non-pathological personality styles based on DSM-IV personality disorder descriptions. Although this instrument aims on measuring personality in general populations, high values in any dimension suggest the existence of a personality disorder (Kuhl & Kazén, 1997). Statements are self-rated on a scale ranging from 0 ('not at all') to 3 ('totally agree'). All but one dimension showed good internal consistencies between $\alpha \geq 0.70$ and $\alpha \leq 0.77$, except the narcissism scale with $\alpha = 0.58$.

Data analysis strategy

Since a dichotomous-dependent variable (dropout) was investigated with patients (i.e., level 1) nested within therapists (i.e., level 2), a multilevel logistic regression model was used (Hox, 2010). The data were analysed with the software R version 3.1.2 (R Core Team, 2014) and the package lme4 (Bates, Maechler, Bolker, & Walker, 2013).

To estimate the variance attributed to the therapists, the variance partition coefficient for a logistic model was calculated (also refer to Lewis *et al.*, 2010):

$$VPC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

For a multilevel logistic regression model, the variance for the error term σ_e^2 is fixed to be $\frac{\pi^2}{3}$, whereas the variance of the random intercept σ_u^2 is estimated. In two-level multilevel models, the total variability is split into variation *between* level two units (e. g., between therapists) and *within* level two units (e.g., within therapists). Therapist effects are calculated by dividing the proportion of therapist variance (i.e., variation between therapists) through

the total variance (variation between plus variation within therapists; Moineddin, Matheson, & Glazier, 2007).

In a first step, we investigated whether therapist differences had a significant impact on dropout. Therefore, an unconditional model (without predictors, but with a random intercept), allowing the therapists to differ in their average dropout rates, was calculated. This model was compared with a model with a fixed intercept (using a likelihood ratio test) to examine if the therapist variance was larger than zero (Hox, 2010).

In the second analytic step, it was tested whether a significant therapist effect still remained when the model controlled for initial patient impairment. Therefore, the grand-mean centered initial impairment (GSI of the BSI) was entered as a predictor on the patient level. The estimation of the therapist effect in this model is thus the therapist effect for averagely impaired patients (Lutz *et al.*, 2007; Saxon & Barkham, 2012; Wampold & Brown, 2005). Again, this model was compared with a model with a fixed intercept using a likelihood ratio test.

In a next step, a random intercept/random slope model was tested against a random intercept/fixed slope model to check if the impact of initial impairment on dropout differs between the therapists. Then, in order to identify additional predictors besides initial impairment, a consecutive variable selection approach based on three blocks of variables was applied (sociodemographic variables, interpersonal problems and personality styles and treatment expectations). Prior to inclusion into the prediction model, all interval-scaled predictor variables were grand-mean centred, and ordinal scaled variables were dummy coded (Enders & Tofighi, 2007). Initial impairment (GSI) was included in all three variable blocks because based on previous research, initial distress was expected to have a strong influence on dropout, and we were interested only in predictors, which explain additional variance beyond impairment. Therefore, models for each dimension were calculated with the GSI plus all available variables in this block (refer to Tables 2-4). Based on these block-wise analyses, only predictors, which showed a liberal $p \leq 0.10$, were included in the final model (refer to Table 5). To increase power, the liberal significance level was chosen for the blocks, while in the final model, only predictors with a $p \leq 0.05$ were interpreted as statistically significant.

Since data were collected in a naturalistic setting, some missing values did exist in the investigated predictor variables. Therefore, we ran the final model with listwise deletion as well as with multiple imputations and compared the results. The multiple imputation analysis was conducted using the R package Amelia2 based on five different imputed datasets (Honaker, King, & Blackwell, 2011).

RESULTS

Therapist effects on dropout

The average dropout rate per therapist was 33.2% with an interquartile range of 19.0 to 50.0%. The variance of the random intercept was 0.20, resulting in estimation for the variance partition coefficient (therapist effect) of 5.72%. The therapist effect differed significantly from zero ($\chi^2(1,\ 707) = 6.87,\ p < 0.001$). When controlling for the initial impairment, the estimated level two variance was 0.22, resulting in a therapist effect of 6.21% which was significant ($\chi^2(1, 707) = 7.50, p < 0.001$).

The results for the fixed effect part of the model are shown in Table 1. An intercept of −0.778 could be translated into the dropout probability when inserted into the regression equation:

$$\frac{e^{(-.078)}}{1 + e^{(-0.78)}} = 0.\,314$$

Thus, the probability of a dropout for an averagely impaired patient is 31.4%. With one additional unit on the GSI, the dropout probability increases to:

$$\frac{e^{(-0.78+0.33)}}{1 + e^{(-0.78+.033)}} = \ 0.\,390$$

To test whether the interrelation between the GSI and the dropout criterion differs between the therapists, a random slope model was calculated. The model with a varying slope, for the GSI did not fit better than the intercept only model with $\chi^2(1, 707) = 0.218$, $p = 0.14$. That is, the impact of the initial impairment on dropout did not differ between the therapists.

Impact of socio-demographic variables

The model examining socio-demographic variables included age, sex, education, marital status, sick leave status and sex match between therapist and patient. Age was standardized at the grand mean. Sex of the patient was coded with 0 = female and 1 = male. The education variable was dummy-coded with the reference category 'low education'. 'Education middle' was coded with 1 if a patient attended junior high school

Table 1. Fixed effects on dropout

	Estimate	Std. error	*z* value	Pro.(>|z|)
Intercept	−0.778	0.105	−7.395	0.000***
GSI	0.331	0.117	2.823	0.005**

Note. ***$p < 0.001$., **$p < 0.01$.

Table 2. Fixed effects on dropout with socio-demographic variables

	Estimate	Std. error	*z* value	Pro.(>|z|)
Intercept	0.041	0.362	0.112	0.910
GSI	0.252	0.126	1.992	0.046**
Age	−0.015	0.008	−1.791	0.073*
Sex	0.550	0.237	2.321	0.020**
Education middle	−0.200	0.226	−0.884	0.376
Education high	−0.774	0.228	−3.390	0.001***
Marital Status	−0.248	0.224	−1.107	0.268
Sick leave	0.171	0.239	0.715	0.475
Sex match	−0.269	0.239	−1.126	0.260

Note. ***$p < 0.001$., **$p < 0.05$., *$p < 0.10$.

(otherwise = 0), and 'education high' was coded with 1 if a patient had a general qualification for university entrance (otherwise = 0). Marital status was coded with 1 if a patient was married, otherwise with 0. The variable of the sick leave status was coded with 1 if a patient was on sick leave at the beginning of therapy. Finally, a variable which stands for the match of the sex of both therapist and patient was included. A zero was coded when both patient and therapist were of the same sex, and a one was coded when they were of different sexes.

As shown in Table 2, age, sex and 'education high' were significant predictors of patients' dropout probability. Younger patients, men compared with women and patients with 'lower education' compared with patients with 'higher education' showed significantly higher dropout probabilities. The marginal R^2_{GLMM} for this model was 5.83%, and the conditional R^2_{GLMM} for this model was 12.13% (Nakagawa & Schielzeth, 2013).

Impact of personality style and interpersonal problems

To investigate interpersonal problems and personality styles, the mean score of the Inventory of Interpersonal Problems as well as all 14 scales of the PSSI-K were entered into one prediction model. As can be seen in Table 3, two scales of the PSSI-K predicted dropout significantly. More compulsive patients were associated with a lower probability for dropping out of treatment, while more histrionic patients were more likely to dropout. For this model, the marginal R^2_{GLMM} was 7.74%, and the conditional R^2_{GLMM} was 17.54%.

Impact of treatment expectations

Two single items measuring treatment expectations have been entered into a model predicting dropout. Both items showed a significant relation with attrition (refer to Table 4). The patients with more difficulties attending psychotherapy were less likely to drop out. The patients with more confidence that psychotherapy will help them with their problems showed a lower probability of dropping out of treatment. The marginal R^2_{GLMM} for this model was 4.17%, and the conditional R^2_{GLMM} for this model was 11.27%.

Table 3. Fixed effects on personality style and interpersonal problems

	Estimate	Std. error	*z* value	Pro.(>|z|)
Intercept	−0.920	0.137	−6.727	0.000***
GSI	0.355	0.246	1.442	0.149
IIP-32	0.199	0.294	0.677	0.498
Antisocial	0.126	0.188	0.672	0.501
Paranoid	0.203	0.229	0.888	0.375
Schizoid	0.224	0.220	1.018	0.309
Self-insecure	−0.298	0.240	−1.241	0.215
Compulsive	−0.323	0.168	−1.919	0.055*
Schizotypal	0.157	0.172	0.913	0.361
Rhapsodic	−0.065	0.242	−0.268	0.789
Narcissism	0.035	0.201	0.176	0.860
Negativism	0.042	0.222	0.189	0.850
Dependent	0.123	0.193	0.638	0.523
Borderline	−0.320	0.203	−1.580	0.114
Histrionic	0.489	0.219	2.232	0.026**
Depressive	0.182	0.253	0.721	0.471
Selfless	−0.064	0.183	−0.350	0.726

Note. ***$p < 0.001$., **$p < 0.05$., *$p < 0.10$.

Final model with patient level predictors

In the final prediction model including all significant predictors from the previous models, all variables beside age remained significant (refer to Table 5). The marginal R^2_{GLMM} for the final model was 10.55%, and the conditional R^2_{GLMM} for this model was 22.71%. The results remained mainly the same when estimating the final model using multiple imputation of missing values instead of list-wise deletion (refer to Table 6).

Table 4. Fixed effects on treatment expectations

	Estimate	Std. error	*z* value	Pro.(>|z|)
Intercept	−0.835	0.112	−7.477	0.000***
GSI	0.359	0.125	2.884	0.004**
Treatment expectation 1	−0.210	0.093	−2.245	0.025*
Treatment expectation 2	−0.426	0.127	−3.342	0.001***

Note. ***$p < 0.001$., **$p < 0.01$., *$p < 0.05$.

Table 5. Fixed effects on final model

	Estimate	Std. error	z value	Pro.(>|z|)
Intercept	−0.577	0.421	−1.373	0.170
GSI	0.396	0.146	2.713	0.007**
Age	−0.005	0.009	−0.599	0.549
Sex	0.458	0.210	2.181	0.029*
Education middle	−0.102	0.258	−.397	0.691
Education high	−0.590	0.257	−2.294	0.022*
Compulsive	−0.390	0.151	−2.573	0.010*
Histrionic	0.395	0.156	2.539	0.011*
Treatment expectation 1	−0.254	0.106	−2.399	0.016*
Treatment expectation 2	−0.457	0.143	−3.192	0.001**

Note. $^{**}p < 0.01$., $^{*}p < 0.05$.

Table 6. Fixed effects on final model with multiple imputed values

	Estimate	Std. error	z value	Pro.(>|z|)
Intercept	−0.278	0.375	−0.765	0.444
GSI	0.384	0.129	2.984	0.003**
Age	−0.010	0.008	−1.272	0.204
Sex	0.435	0.188	2.311	0.021*
Education middle	−0.186	0.227	−0.817	0.414
Education high	−0.699	0.229	−3.049	0.002**
Compulsive	−0.398	0.141	−2.812	0.005**
Histrionic	0.363	0.152	2.396	0.017*
Treatment expectation 1	−0.186	0.095	−1.960	0.050*
Treatment expectation 2	−0.366	0.131	−2.804	0.005**

Note. $^{**}p < 0.01$., $^{*}p < 0.05$.

DISCUSSION

The estimate of the therapist effect on dropout within this study was 5.72% (and 6.21% when initial impairment was taken into account). This is in the range of therapist effects on patients' outcome, which is estimated to be 5–8% (e. g., Baldwin & Imel, 2013; Crits-Christoph *et al.*, 1991; Kim *et al.*, 2006; Lutz *et al.*, 2007; Saxon & Barkham, 2012). Compared with therapeutic alliance, which is one of the most studied variables in psychotherapy, which accounts for about 5% in patient's outcomes, approximately 6% variance in patient's dropout seems quite much (Baldwin & Imel, 2013).

The final prediction model within this study included seven significant predictors, which were associated with a higher dropout rate: higher initial impairment, male sex, lower education, lower compulsive and higher histrionic personality styles, less difficulties in attending psychotherapy and lower confidence that psychotherapy will help (treatment expectation).

Swift and Greenberg (2012) reported mixed results for socio-demographic variables and their impact of dropout. In line with the finding of Swift and Greenberg (2012) and our hypothesis, lower education was associated with an increased dropout risk. Contrary to our hypothesis, no significant association with age could be found. Also, sex had a significant impact on dropout in the way that men showed an increased dropout probability compared with women. Socioeconomic status and related factors like education level have been repeatedly shown to have an association with dropout in both older and more recent research (Baekeland & Lundwall, 1975; Wierzbicki & Pekarik, 1993; Swift & Greenberg, 2012). Our study strengthened the finding that education is a robust predictor of dropout which should be examined more closely and should be considered when studying dropout. The impact of age and sex on dropout is mixed within the literature. While Wierzbicki and Pekarik (1993) could not find any significant relation between sex, age and dropout, it was found for age in the latest large systematic review (Swift & Greenberg, 2012).

In this study, the compulsive and histrionic personality dimensions were significantly associated with dropout. Other personality dimensions as well as interpersonal problems showed no meaningful relationship to dropout. It was shown that more compulsive patients tended to dropout less often. This might be explained with the construct of compulsivity. People with compulsive tendencies are characterized as conscientious, tidy and perfectionist (4th ed., text rev.; DSM-IV-TR; American Psychiatric Association, 2000). In a therapy setting, such patients might be more hardworking and reliable. This could lead to a higher commitment to therapy and the wish to stick conscientiously to the treatment. However, more histrionic patients had an increased risk for dropout. Amongst other things, histrionic people are characterized by liveliness and impulsivity (DSM-IV-TR). Both the beginning and termination of psychotherapeutic treatment could be affected by these traits. Therefore, a premature dropout might be explained with spontaneous behaviours (e.g., spontaneous intention to quit therapy), which will more likely occur in histrionic patients. In line with this finding, Fernandez-Montalvo and López-Goñi (2010) showed in a treatment for cocaine addiction that more compulsive patients were more likely to complete the treatment, while more histrionic patients were more likely to dropout. In their study, all patients diagnosed with a comorbid histrionic personality disorder showed premature treatment termination. It is puzzling that other personality styles such as borderline, which could also be characterized by liveliness and impulsivity, were not predictive in the model. To increase the power of finding such an effect, we checked a model, which contained only the GSI and the borderline scale. Still, the borderline scale was not significant ($p = 0.76$). Both personality

dimensions (histrionic and compulsive) remained significant in the final prediction model, indicating that they each explain a specific variation in dropout.

This study found significant relations for both items on treatment expectations with dropout. The result that the patients with more confidence that psychotherapy will help them with their problems showed lower dropout probabilities seems intuitive. The patients with high treatment expectations are confident that further therapy will help them in relieving their distress. A premature termination would lead to cognitive dissonance which people like to avoid (Cooper, 2012). Cognitive dissonance theory might also explain the contra-intuitive finding that amongst patients with more difficulties in attending psychotherapy, dropout is less likely. When people decide to begin therapy and have large difficulties in attending the treatment, the investment was quite high. After investing much effort in beginning therapy, the decision to premature termination would lead to cognitive dissonance. Another explanation could be that patients having more difficulties attending psychotherapy have more psychological strain and therefore more need for treatment; however, higher initial impairment measured with the symptom-based BSI was associated with a higher likelihood of dropout, which is contradictory to this explanation.

The impact of patients' treatment expectations on outcome is small but robust (Constantino *et al.*, 2011). However, findings on patients' treatment expectations and their relation on dropout are mixed. Tsai, Ogrodniczuk, Sochting, and Mirmiran (2014) could not find any relationship between outcome expectations at the beginning of treatment and completion status. Also, in the treatment of patients with post-traumatic stress disorder, neither therapy motivation nor treatment expectations significantly discriminated between completers and dropouts (van Minnen, Arntz, & Keijsers, 2002). In contrast, Simpson and Joe (1993) found, in a drug abuse treatment, a significant relationship between treatment expectations and dropout.

How can we reduce the dropout rates? Swift, Greenberg, Whipple, and Kominiak (2012) pointed out six different strategies for a therapist to reduce the likelihood of a dropout. For instance, many patients have unrealistic expectations about duration and recovery. Therapists should therefore educate their patients about average therapy length and goal attainment. The results from this study underpin these recommendations, since treatment expectations significantly predicted dropout.

Another strategy could be the implementation of routinely monitoring individual treatment outcomes to reduce dropout rates (e.g., Lutz *et al.*, 2015). Future research could investigate whether different feedback strategies which focus more closely to the identified risk factors (e.g., treatment expectation or personality styles) could help to avoid dropout. There are other factors which might play a role in dropout. For example, Swift and Greenberg (2012) found some evidence that therapist's experience is important for reducing dropout. Investigating this variable to check whether experience could explain some of the therapist variance in dropout seems fruitful. This study could not examine this question because the therapists were all in their clinical training and all had similar experience in conducting therapy. Another example is the therapeutic relationship. This variable is a well-studied construct which should be analysed with different methods. Training therapists in repairing alliance ruptures occurring in the treatment might also reduce dropout rates (Safran, Muran, & Eubanks-Carter, 2011).

Aside from the strengths of the present paper, there are also some limitations. One could argue that the differences between the therapists are due to non-randomized assigning of the patients to the therapists, as it was the case in this study. The therapists with more severely impaired patients or with more difficult patients might face higher dropout rates. However, it could be shown that the therapist variance was not reduced when initial impairment was introduced as a covariate.

Some methodological issues limit the scope of this study. The sample size of 707 patients nested within 66 therapists ($M = 10.7$ patients per therapist) might be small when investigating therapist effects (Moineddin *et al.*, 2007). Further studies should replicate these findings with an increased number of patients treated per therapist and more therapists to get more precise estimates of the therapist effect.

Another limitation of this study could be seen in the dropout criteria applied. In this sample, no specific reasons about why the patients decided to quit the intended therapy were investigated. There is a lack of research identifying reasons why the patients leave treatment in a non-consensual way. Future research should take into account the reasons for why therapists or patients intent to quit treatment.

There is a vast body of literature studying the effects of ethnicity or cultural differences in psychotherapy (e. g., Flaskerud, 1990; Markowitz, Spielman, Sullivan, & Fishman, 2000; Owen, Imel, Adelson, & Rodolfa, 2012; Ruglass *et al.*, 2015). As a limitation of this study, no data on ethnicity neither for the patients nor for the therapists were available. Therefore, it could not be tested for a potential matching effect which was for example found by Owen *et al.* (2012).

The present study found substantial therapist effects on dropout in outpatient psychotherapy. Future research should investigate the effect of therapists in different settings and with an expansion of the applied methods. Further therapist- and patient–therapist interaction variables should be studied to clarify which variables are associated with the differences in dropout rates observed between therapists.

REFERENCES

American Psychiatric Association. (2000). DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders (4th ed. text rev.). Washington, DC: Author.

Aubuchon-Endsley, N. L., & Callahan, J. L. (2009). The hour of departure: Predicting attrition in the training clinic from role expectancies. *Training and Education in Professional Psychology, 3*, 120–126.

Baekeland, F., & Lundwall, L. (1975). Dropping out of treatment: A critical review. *Psychological Bulletin*, *82*(5), 738.

Baldwin, S. A., & Imel, Z. E. (2013). Therapist effects: Findings and methods. In Lambert, M. J. (Ed.), Bergin and Garfield's handbook of psychotherapy and behavior change (6th ed. pp. 258–297). New Jersey: John Wiley & Sons.

Barrett, M. S., Chua, W. J., Crits-Christoph, P., Gibbons, M. B., & Thompson, D. (2008). Early withdrawal from mental health treatment: Implications for psychotherapy practice. *Psychotherapy: Theory, Research, Practice, Training*, *45*(2), 247. DOI:10.1037/a0016184.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-5. http://CRAN.R-project.org/package=lme4.

Beckham, E. E. (1992). Predicting patient dropout in psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, *29*(2), 177–182.

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

Brogan, M. M., Prochaska, J. O., & Prochaska, J. M. (1999). Predicting termination and continuation status in psychotherapy using the transtheoretical model. *Psychotherapy: Theory, Research, Practice, Training*, *36*(2), 105. DOI:10.1037/h0087773.

Bronisch, T., Hiller, W., Mombour, W., & Zaudig, M. (1996). International diagnostic checklists for personality disorders according to ICD-10 and DSM–IV-IDCL-P. Seattle, WA: Hogrefe and Huber Publishers.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). New Jersey: Lawrence Erlbaum.

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.

Cooper, J. (2012). Cognitive dissonance theory. In Van Lange, P. A. M., Kruglanski, A. W., & Higgins, E. T. (Ed.), The handbook of theories of social psychology (ed. pp. 377–398) Vol. 1,. London: Sage.

Crits-Christoph, P., Baranackie, K., Kurcias, J., Beck, A., Carroll, K., Perry, K., … Zitrin, C. (1991). Meta-analysis of therapist Effects in psychotherapy outcome studies. *Psychotherapy Research*, *1*(2), 81–91.

Delgadillo, J., McMillan, D., Lucock, M., Leach, C., Ali, S., & Gilbody, S. (2014). Early changes, attrition, and dose–response in low intensity psychological interventions. *British Journal of Clinical Psychology, 53*, 114–130.

Derogatis, L. R. (1975). Brief Symptom Inventory. Baltimore: Clinical Psychometric Research.

Druss, B. G., & Rosenheck, R. A. (1999). Patterns of health care costs associated with depression and substance abuse in a national sample. *Psychiatric Services*, *50*(2), 214–218.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological Methods*, *12*(2), 121. DOI:10.1037/1082-989X.12.2.121.

Fernandez-Montalvo, J., & López-Goñi, J. J. (2010). Comparison of completers and dropouts in psychological treatment for cocaine addiction. *Addiction Research & Theory, 18*(4), 433–441.

Flaskerud, J. H. (1990). Matching client and therapist ethnicity, language, and gender: A review of research. *Issues in Mental Health Nursing*, *11*(4), 321–336.

Franke, G. H. (2000). Brief Symptom Inventory (BSI) von L. R. Derogatis (Kurzform der SCL-90-R) – Deutsche Version. Göttingen: Beltz.

Garb, H. N. (2005). Clinical judgment and decision making. *Annual Review of Clinical Psychology, 1*, 67–89.

Greenberg, R. P., Constantino, M. J., & Bruce, N. (2006). Are patient expectations still relevant for psychotherapy process and outcome? *Clinical Psychology Review, 26*, 657–678.

Gunderson, J. G., Prank, A. F., Ronningstam, E. F., Wachter, S., Lynch, V. J., & Wolf, P. J. (1989). Early discontinuance of borderline patients from psychotherapy. *The Journal of Nervous and Mental Disease*, *177*(1), 38–42.

Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data. *Journal of Statistical Software*, *45*(7), 1–47.

Hatchett, G. T., Han, K., & Cooker, P. G. (2002). Predicting premature termination from counseling using the Butcher Treatment Planning Inventory. *Assessment*, *9*, 156–163.

Hatchett, G. T., & Park, H. L. (2003). Comparison of four operational definitions of premature termination. *Psychotherapy: Theory, Research, Practice, Training*, *40*(3), 226.

Horner, M. S., & Diamond, D. (1996). Object relations development and psychotherapy dropout in borderline outpatients. *Psychoanalytic Psychology, 13*(2), 205. DOI:10.1037/h0079648.

Horowitz, L., Rosenberg, A. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). Inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56*(6), 885–892.

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist*, *41*(2), 159.

Hox, J. J. (2010). Multilevel analysis: Techniques and applications (2. ed.). New York: Routledge.

Huppert, J. D., Kivity, Y., Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2014). Therapist effects and the outcome–alliance correlation in cognitive behavioral therapy for panic disorder with agoraphobia. *Behaviour Research and Therapy, 52*, 26–34.

Karterud, S., Pedersen, G., Bjordal, E., Brabrand, J., Friis, S., Haaseth, Ø., … Urnes, Ø. (2003). Day treatment of patients with personality disorders: experiences from a Norwegian treatment research network. *Journal of Personality Disorders*, *17*(3), 243–262.

Kim, D. M., Wampold, B. E., & Bolt, D. M. (2006). Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychotherapy Research*, *16*(2), 161–172.

Kolb, D. L., Beutler, L. E., Davis, C. S., Crago, M., & Shanfield, S. B. (1985). Patient and therapy process variables relating to dropout and change in psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, *22*(4), 702.

Kuhl, J., & Kazén, M. (1997). Persönlichkeits-Stil-und Störungs-Inventar: (PSSI). Göttingen: Hogrefe.

Lambert, M. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research*, *17*(1), 1–14.

Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In Lambert, M. J. (Ed.), Bergin and Garfield's handbook of psychotherapy and behavior change (5th ed. pp. 139–193). New York: John Wiley & Sons.

Lewis, M., Morley, S., van der Windt, D. A. W. M., Hay, E., Jellema, P., Dziedzic, K., & Main, C. J. (2010). Measuring practitioner/therapist effects in randomised trials of low back pain and neck pain interventions in primary care settings. *European Journal of Pain*, *14*, 1033–1039.

Lutz, W. (2002). Patient-focused psychotherapy research and individual treatment progress as scientific groundwork for an empirically based clinical practice. *Psychotherapy Research*, *12*(3), 251–272.

Lutz, W., Hofmann, S. G., Rubel, J., Boswell, J. F., Shear, M. K., Gorman, J. M., … Barlow, D. H. (2014). Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. *Journal of Consulting and Clinical Psychology*, *82*(2), 287–297.

Lutz, W., Leon, S. C., Martinovich, Z., Lyons, J. S., & Stiles, W. B. (2007). Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology*, *54*(1), 32. DOI:10.1037/0022-0167.54.1.32.

Lutz, W., Rubel, J., Schiefele, A.-K., Zimmermann, D., & Böhnke, J. R. (2015). Therapist and feedback effects in the context of treatment outcome and length. *Psychotherapy Research* Advance online publication. DOI:10.1080/10503307.2015.1053553.

Maher, M., Huppert, J. D., Chen, H., Duan, N., Foa, E. B., Liebowitz, M. R., & Simpson, H. B. (2010). Moderators and predictors of response to cognitive-behavioral therapy augmentation of pharmacotherapy in obsessive–compulsive disorder. *Psychological Medicine*, *40*(12), 2013–2023.

Markowitz, J. C., Spielman, L. A., Sullivan, M., & Fishman, B. (2000). An exploratory study of ethnicity and psychotherapy outcome among HIV-positive patients with depressive symptoms. *The Journal of Psychotherapy Practice and Research*, *9*(4), 226–231.

McMurran, M., Huband, N., & Overton, E. (2010). Non-completion of personality disorder treatments: A systematic review of correlates, consequences, and interventions. *Clinical Psychology Review*, *30*(3), 277–287.

Moineddin, R., Matheson, F. I., & Glazier, R. H. (2007). A simulation study of sample size for multilevel logistic regression models. *BMC Medical Research Methodology*, *7*(1), 34. DOI:10.1186/1471-2288-7-34.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142.

Ogrodniczuk, J. S., Joyce, A. S., Lynd, L. D., Piper, W. E., Steinberg, P. I., & Richardson, K. (2008). Predictors of premature termination of day treatment for personality disorder. *Psychotherapy and Psychosomatics*, *77*, 365–371.

Owen, J., Imel, Z., Adelson, J., & Rodolfa, E. (2012). 'No-show': Therapist racial/ethnic disparities in client unilateral termination. *Journal of Counseling Psychology*, *59*(2), 314–320.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Richards, D. A., & Borglin, G. (2011). Implementation of psychological therapies for anxiety and depression in routine practice: Two year prospective cohort study. *Journal of Affective Disorders*, *133*, 51–60.

Ruglass, L. M., Pedersen, A., Cheref, S., Hu, M. C., & Hien, D. A. (2015). Racial differences in adherence and response to combined treatment for full and subthreshold post-traumatic stress disorder and alcohol use disorders: A secondary analysis. *Journal of Ethnicity in Substance Abuse*, 1–15.

Safran, J. D., Muran, J. C., & Eubanks-Carter, C. (2011). Repairing alliance ruptures. *Psychotherapy*, *48*(1), 80–87.

Sainsbury Centre for Mental Health (2003). The economic and social costs of mental illness. London: Sainsbury Centre for Mental Health.

Saxon, D., & Barkham, M. (2012). Patterns of therapist variability: Therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology*, *80*(4), 535. DOI:10.1037/a0028898.

Self, R., Oates, P., Pinnock-Hamilton, T., & Leach, C. (2005). The relationship between social deprivation and unilateral termination (attrition) from psychotherapy at various stages of the health care pathway. *Psychology and Psychotherapy: Theory, Research and Practice*, *78*(1), 95–111.

Simpson, D. D., & Joe, G. W. (1993). Motivation as a predictor of early dropout from drug abuse treatment. *Psychotherapy: Theory, Research, Practice, Training*, *30*(2), 357. DOI:10.1037/0033-3204.30.2.357.

Sledge, W. H., Moras, K., Hartley, D., & Levine, M. (1990). Effect of time-limited psychotherapy on patient dropout rates. *American Journal of Psychiatry*, *147*(10), 1341–1347.

Swift, J. K., & Callahan, J. L. (2011). Decreasing treatment dropout by addressing expectations for treatment length. *Psychotherapy Research*, *21*(2), 193–200.

Swift, J. K., & Greenberg, R. P. (2012). Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, *80*(4), 547. DOI:10.1037/a0028226.

Swift, J. K., Greenberg, R. P., Whipple, J. L., & Kominiak, N. (2012). Practice recommendations for reducing premature termination in therapy. *Professional Psychology: Research and Practice*, *43*(4), 379. DOI:10.1037/a0028291.

Thomas, A., Brähler, E., & Strauß, B. (2011). IIP-32: Entwicklung, Validierung und Normierung einer Kurzform des Inventars zur Erfassung interpersonaler Probleme. *Diagnostica*, *57*(2), 68–83.

Tsai, M., Ogrodniczuk, J. S., Sochting, I., & Mirmiran, J. (2014). Forecasting success: Patients' expectations for improvement and their relations to baseline, process and outcome variables in group cognitive–behavioural therapy for depression. *Clinical Psychology and Psychotherapy*, *21*, 97–107.

Van Minnen, A., Arntz, A., & Keijsers, G. P. J. (2002). Prolonged exposure in patients with chronic PTSD: Predictors of treatment outcome and dropout. *Behaviour Research and Therapy*, *40*(4), 439–457.

Wampold, B. E., & Brown, G. S. J. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology*, *73*(5), 914–923.

Wierzbicki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology: Research and Practice*, *24*(2), 190–195.

Wittchen, H.-U., Wunderlich, U., Gruschwitz, S., & Zaudig, M. (1997). SKID-I. Strukturiertes Klinisches Interview für DSM-IV. Achse I: Psychische Störungen. Interviewheft. Göttingen: Hogrefe.

7. Study III

Zimmermann, D., Lutz, W., Reiser, M., Boyle, K., Schwartz, B., Schilling, V., Deisenhofer, A.-K., & Rubel, J. (2019). What happens when the therapist leaves? The impact of therapy transfer on the therapeutic alliance and symptoms. *Clinical Psychology & Psychotherapy*, *26*(1), 135-145. doi: 10.1002/cpp.2336

Received: 14 February 2018 | Revised: 21 September 2018 | Accepted: 21 September 2018

DOI: 10.1002/cpp.2336

RESEARCH ARTICLE

WILEY

What happens when the therapist leaves? The impact of therapy transfer on the therapeutic alliance and symptoms

Dirk Zimmermann | Wolfgang Lutz | Michelle Reiser | Kaitlyn Boyle | Brian Schwartz | Viola N.L.S. Schilling | Anne-Katharina Deisenhofer | Julian A. Rubel

Clinical Psychology and Psychotherapy, University of Trier, Trier, Germany

Correspondence
Dirk Zimmermann, MSc, Clinical Psychology and Psychotherapy, Department of Psychology, University of Trier, 54286 Trier, Germany.
Email: zimmermann@uni-trier.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Numbers: LU 660/10-1 and LU 660/8-1

Background: The therapeutic alliance is an important factor in psychotherapy, affecting both therapy processes and outcome. Therapy transfers may impair the quality of the therapeutic alliance and increase symptom severity. The aim of this study is to investigate the impact of patient transfers in cognitive behavioural therapy on alliance and symptoms in the sessions after the transfer.

Method: Patient- and therapist-rated therapeutic alliance and patient-reported symptom severity were measured session-to-session. Differences in the levels of alliance and symptom severity before (i.e., with the original therapist) and after (i.e., with the new therapist) the transfer session were analysed. The development of alliance and symptom severity was explored using multilevel growth models.

Results: A significant drop in the alliance was found after the transfer, whereas no differences were found with regard to symptom severity. After an average of 2.93 sessions, the therapeutic alliance as rated by patients reached pretransfer levels, whereas it took an average of 5.05 sessions for therapist-rated alliance levels to be at a similar level as before the transfer. Inter-individual differences were found with regard to the development of the therapeutic alliance over time.

Conclusions: Therapy transfers have no long lasting negative effects on either symptom impairment or the therapeutic alliance.

KEYWORDS

case transfer, multilevel modelling, symptom severity, therapeutic alliance, therapy transfer

1 | INTRODUCTION

Therapy or case transfers from one therapist to another are common in clinical practice (Wapner, Klein, Friedlander, & Andrasik, 1986). They occur as a result of turnover, sick leave, or parental leave. In training clinics, where therapists graduate and therefore leave the institute, therapy transfers occur on a regular basis (Cox, 2017). Despite the clinical relevance in most psychotherapeutic settings, the impact of therapy transfers has rarely been studied. So far, there is little empirical evidence on whether and/or how patients are affected by such transfers (for a notable exception, see Sauer, Rice, Richardson, & Roberts, 2017).

Flowers and Booraem (1995) conducted one of a few quantitative studies investigating the association between transferring patients and therapy outcomes. They compared patients treated in a training clinic who stayed with their therapists for the entirety of their treatment with those who were transferred to another trainee. At the end of treatment, no significant differences were found between these groups with regard to general contentment or diagnostic outcomes. However, although the patients who were not transferred showed a linear improvement of contentment over time, the pattern of contentment scores varied for the transferred group. The transferred patients showed an initial improvement of contentment followed by a decrease that subsequently improved again over the final set of sessions. These results suggest that adverse effects of transfers on patients might be temporary. However, patient development immediately after a transfer was not analysed.

Another recent quantitative study investigated the impact of client attachment and gender on therapy transfers and client outcomes (Sauer et al., 2017). The symptom changes of 35 adult clients receiving psychotherapy were analysed over four pretransfer and four posttransfer sessions. Results suggested that men were likely to experience an increase in psychological distress posttransfer whereas this association was not found for women. Another finding was that attachment orientation predicted symptom change posttransfer. With higher levels of avoidant attachment styles, clients were more likely to show an increase of symptoms following the transfer to a different therapist. The authors concluded that both male clients and clients with avoidant attachment styles may have more struggles to disclose with the new therapist leading to an increase in psychological distress.

Despite the few quantitative studies focusing on client outcomes, most of the existing research focused on theoretical considerations and on qualitative strategies (e.g., Penn, 1990). For example, forced transfer is described as narcissistic mortification comparable with abandonment by a lover and in violation of the implied rule that the therapist should care for the patient unconditionally, leading to a loss of trust and safety (Siebold, 1991). Such an event could lead to higher dropout rates and diminish the success of therapy (Trimboli & Keenan, 2010). The therapeutic alliance between therapist and patient seems to be the critical factor explaining these findings.

So far, no study investigated the impact of therapy transfer on therapeutic alliance, although the therapeutic alliance is one of the most studied and robust predictors of treatment outcome (e.g., Falkenström, Finkel, Sandell, Rubel, & Holmqvist, 2017; Flückiger, Del Re, Wampold, Symonds, & Horvath, 2012; Nevid, Ghannadpour, & Haggerty, 2017). Bordin (1979) introduced the concept of a pan-theoretical therapeutic alliance assuming that every treatment approach shares the ingredient of a therapeutic alliance between therapist and patient. His definition of a positive alliance involves consensus regarding therapy goals and tasks and an affective interpersonal bond, which includes mutual trust, acceptance, and confidence, between patient and therapist (Horvath & Luborsky, 1993).

In cognitive behavioural therapy (CBT), the alliance is an important ingredient enabling therapists to more effectively apply their therapeutic techniques (such as confrontation and Socratic questioning; e.g., Castonguay, Goldfried, Wiser, Raue, & Hayes, 1996; Rubel, Rosenbaum, & Lutz, 2017). The therapeutic alliance is not only important for a constructive therapeutic process but also directly influences therapeutic outcome. The strength of the therapeutic alliance is a known predictor of treatment outcome, accounting for up to 10% of variance (Horvath, Del Re, Flückiger, & Symonds, 2011; Lambert, 1992). Meta-analyses report a mean effect size from $r = 0.22$ in homogeneous samples (Martin, Garske, & Davis, 2000) to $r = 0.26$ in heterogeneous samples (Horvath & Symonds, 1991). The alliance–outcome association is described as robust and independent of therapeutic orientation and disorder-specific problems (Castonguay, Constantino, & Grosse Holtforth, 2006). Flückiger et al. (2012) confirm the solid association with correlations from $r = 0.21$ to 0.38, irrespective of study design (randomized controlled trial vs. naturalistic), use of therapeutic manuals, outcome specification, and therapeutic orientation. Moreover, the association between alliance and outcome is substantially greater in comparison with other variables such as adherence to manualized treatment or therapeutic competence (Webb, DeRubeis, & Barber, 2010).

Key Practitioner Message

- Therapy transfers affect the therapeutic alliance in the short term; however, pretransfer alliance levels can be restored in three to five sessions after the transfer.
- Therapy transfers seem to have no negative impact on patients' symptom impairment on average.
- There are inter-individual differences between patients concerning the impact of a therapy transfer on the perceived quality of the therapeutic alliance.

To date, a few studies have examined the alliance–outcome association on a session-to-session level, suggesting that patient-rated alliance predicts subsequent symptom change and vice versa, indicating a reciprocal causal model (Falkenström, Ekeblad, & Holmqvist, 2016; Falkenström, Granström, & Holmqvist, 2013; Rubel et al., 2017; Sasso, Strunk, Braun, DeRubeis, & Brotman, 2016; Zilcha-Mano, Dinger, McCarthy, & Barber, 2014). In general, stronger alliances are associated with better outcomes (Horvath et al., 2011; Zilcha-Mano et al., 2014).

Due to the fact that studies examining the effect of therapist transfers are lacking, the development of alliance and symptom severity after a transfer is unclear. Therefore, this study aims at investigating the impact of a therapy transfer on both symptoms and alliance. Alliance is dynamic, as its elements and quality change over time (Horvath & Greenberg, 1994; Zilcha-Mano & Errázuriz, 2017). There are two main research areas studying the development of alliance over the course of treatment. On the one hand, there is evidence for a linear trend of alliance development displaying a consistent strengthening of the alliance over the course of treatment (e.g., Fitzpatrick, Iwakabe, & Stalikas, 2005; Rubel, Lutz, & Schulte, 2015; Sauer, Lopez, & Gormley, 2003). On the other hand, some studies describe a nonlinear U-shaped trend or rupture-resolution processes (e.g., Gelso & Carter, 1994; Horvath & Luborsky, 1993; Kivlighan Jr, Kline, Gelso, & Hill, 2017; Rubel, Bar-Kalifa, et al., 2018a; Safran & Muran, 2000). Both patterns of alliance development have received attention and lend support to the impact of the selection of the time point of assessment during treatment and the consideration of individual differences in alliance development.

In some alliance studies, brief psychotherapy is examined with only eight to 16 sessions (e.g., Sauer et al., 2017), whereas others have examined the alliance in long-term psychotherapy (Barber, Connolly, Crits-Christoph, Gladis, & Siqueland, 2009; Hersoug, Høglend, Havik, Lippe, & Monsen, 2009). There might be differences in the quality of the therapeutic alliance depending on the time spent with one therapist before the patient is referred to another therapist. The longer a therapy lasts, the more time has passed to influence the development of the alliance. A strong connection between patient and therapist might have an impact on a subsequent therapy transfer. Therefore, the number of sessions attended up to the transfer session is a potentially relevant predictor of posttransfer development.

Depending on how well-established the alliance was up to the time point of the transfer, alliance quality differs, irrespective of a linear or U-shaped pattern. The transfer may be perceived differently depending on the quality of the alliance the patient had with his original therapist. A patient, who already had a strong therapeutic alliance might have a better ability to build strong relationships, for example, because of a more agreeable or open personality style. This patient might therefore start with a higher initial alliance and display a stronger increase over time. Another reason to investigate pretransfer alliance levels is methodological. Patients with a tendency to score on questionnaires in a particular way (e.g., central tendency) are likely to show a consistent response pattern across different therapists (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). One could therefore expect that a patient with a higher rating before transfer provides higher ratings after the transfer as well. Thus, the alliance level before the case transfer may be a potential predictor of the development of the alliance after the transfer. In sum, the number of session after transfer, the transfer session, and the pretransfer alliance level were taken into account as predictors of posttransfer development of the therapeutic alliance.

1.1 | Aim of this study

To the best of our knowledge, this is the first study investigating the impact of therapy transfer on both symptoms and alliance in a large dataset with a longitudinal design. The aim of this study was to investigate the impact of therapy transfers during treatment on alliance and symptom severity and the development of these variables after the transfer. Due to the fact that the therapeutic alliance develops over the course of treatment, a transfer to a new therapist may negatively impact the alliance to the initial therapist. This event may influence alliance ratings after the transfer. As a new relationship has to be built with the new therapist, lower alliance ratings should be expected after the transfer. However, this pattern might be different for patients who experience the transfer as a loss, combined with a feeling of sadness and a difficulty to build a good relationship with the new therapist. This might be contrasted by patients who interpret the loss as abandonment, associated with feelings of anger, followed by an idealization of the new therapist. This might lead to higher alliance ratings after the transfer and/or a strong increase in the alliance over time. On the basis of the assumption of an alliance–outcome association, the transfer may influence symptom severity as well.

First, it is hypothesized that patient- and therapist-rated alliance drop significantly after a transfer. Second, an increase of symptom severity is expected post-transfer. The development of posttransfer alliance and symptom severity is investigated with a longitudinal multilevel modelling approach taking the number of session after transfer, pretransfer alliance levels, and the transfer session into account as potential predictors.

2 | METHOD

2.1 | Therapist sample

The therapist sample stemmed from the outpatient clinic of a German university. All therapists participated in a 3- to 5-year postgraduate clinical training programme to become licensed clinical psychologists in CBT. Before treating their first patient at the outpatient clinic, therapists gained at least 1,000 hr of clinical experience. Therapies were guided by CBT manuals; however, treatment was not strictly manualized and included integrative and interpersonal aspects. Every fourth session was supervised by a senior clinician. Sixty-three initial therapists (84.1% female) treated one to seven patients, and most therapists had only one patient in the sample (32 of 63). Seventy-five subsequent therapists (85.3% female) treated one to four patients with 44 therapists having only one patient in the sample; 36 of 63 initial therapists in the sample who transferred patients to a new therapist also received patients from other therapists (i.e., 36 of 63 therapists were both initial and subsequent therapist in the sample). The experience level of initial and new therapists is comparable also because both initial and subsequent therapists were part of the sample, and reasons for transfer were not only end of training (which could have resulted in an unbalanced sample with more experienced therapists in the initial therapist sample). Therapists after a transfer were selected on availability and on a clinical decision made by the management board of the outpatient training clinic.

2.2 | Therapist transfer

Between 2007 and 2016, a case transfer during treatment was documented for 130 patients. Only patients who received at least four sessions with the initial therapist were considered for this study, leaving a total of 124 patients fulfilling this criterion. A transfer was defined as a change of the therapist during ongoing treatment and the continuation of therapy with the subsequent therapist (at least one therapy session).

2.3 | Patient sample

In the transfer group (N = 124), 63.7% patients were female. The mean age was 35.94 years (SD = 12.79), and patients had an average of 2.27 (SD = 1.24) diagnoses. Patients were assessed with the Structural Clinical Interview for DSM-IV (Wittchen, Wunderlich, Gruschwitz, & Zaudig, 1997); 46.8% of the patients were primarily diagnosed with major depression, 17.8% with an anxiety disorder, 6.5% of the patients had a post-traumatic stress disorder, 4.0% an adjustment disorder, 3.2% an eating disorder, and 2.4% of the patients were diagnosed with dysthymia; 13.7% of the patients had other diagnoses and 5.6% did not have a diagnosis; 4.0% of the patients were diagnosed with at least one personality disorder based on clinical evaluations. This study was approved by the ethical board of the University of Trier.

The mean value of the Global Severity Index[1] was 1.40 (SD = 0.63) at the beginning of therapy. Treatments consisted of 53.8 sessions on average (range: 10–110). All descriptives were compared with all other patients who had at least four sessions with their therapist but did not experience a transfer during treatment (see Table 1). All values were comparable for both groups except for the amount of sessions. On

[1]The Global Severity Index is the average perceived distress caused by psychological and physical symptoms. It is the global value of the German version of the Brief Symptom Inventory (Franke, 2000), a short version of the Symptom Checklist-90-Revised (Derogatis, 1992), which assesses symptoms on nine subscales.

TABLE 1 Comparison of patients with a transfer (N = 124) and without a transfer (N = 1536) via Welch two sample t-test resp. Pearson's chi-squared test for sex

Variables	Transfers	No transfers	Significance
Female	63.7%	63.2%	0.983
Age	35.94 (12.79)	36.24 (12.78)	0.801
Amount of diagnoses	2.27 (1.24)	2.15 (1.17)	0.300
Treatment duration	53.84 (18.28)	31.75 (17.83)	0.000***
GSI	1.40 (0.63)	1.30 (0.71)	0.106

Note. GSI: Global Severity Index. Standard errors appear in parentheses.

*p < 0.05.

**p < 0.01.

***p < 0.001.

average, patients who experienced a transfer received more treatment sessions.

2.4 | Instruments

2.4.1 | Bern Post Session Report

The therapeutic alliance was measured by the German version of the Bern Post Session Report 2000 (Regli & Grawe, 2000), patient- and therapist-rated versions. The questionnaire assessed four common factors of psychotherapy (Grawe, 1998) and the quality of the therapeutic alliance. The patient-rated version has 14 items; the therapist-rated version 12 items. All items are answered on a 7-point bipolar scale with values ranging from −3 (*not at all*) to +3 (*yes, exactly*). Both versions of the questionnaire show good internal consistencies (for patients, α between 0.87 and 0.92; for therapists, α between 0.84 and 0.87) and have been validated and applied in several previous studies (e.g., Flückiger, Grosse Holtforth, Znoj, Caspar, & Wampold, 2013; Rubel et al., 2017). In this analysis, only the therapeutic alliance subscales were used. The patient version consists of four items (e.g., "Today I felt comfortable with the therapist."), and the therapist version consists of three items (e.g., "The patient and I are getting along well."). The questionnaire was administered after each session and filled out on touchscreen devices.

2.4.2 | Hopkin's Symptom Checklist

Symptom severity was measured with the German version of the Hopkin's Symptom Checklist-11 (Lutz, Tholen, Schürch, & Berking, 2006). It is a short version of the Symptom Checklist-90-Revised (Derogatis, 1992) that assesses perceived distress caused by psychological and physical symptoms during the past 7 days. The mean value indicates overall symptom severity. All items are answered on a 4-point Likert scale ranging from 1 (*not at all*) to 4 (*extremely*). The questionnaire was developed and validated by Lutz et al. (2006) and found to be an economical but sensitive process and outcome measure with good reliability values (Cronbach's α between 0.81 and 0.91) and high correlations with the full version (r = 0.89) and the German version of the Brief Symptom Index-53 (Franke, 2000; r = 0.91). The Hopkin's Symptom Checklist-11 was administered before each session and filled out on touchscreen devices.

2.5 | Statistical analyses

The data analyses were conducted with IBM SPSS Statistics 22 and the free statistical software R, using package "lme4" (Bates et al., 2015). First, paired t-tests were used to investigate whether there was a significant drop in the level of therapeutic alliance or a significant increase in symptom severity between the transfer session (last session with initial therapist) and the first session with the subsequent therapist. Subsequently, changes in the therapeutic alliance and symptom severity were investigated in the sessions after the transfer using multilevel growth modelling (e.g., Raudenbush & Bryk, 2002). In this study, the five sessions after the transfer were assessed to analyse the development of the alliance (Flückiger et al., 2013). This rather arbitrary number was chosen to further look at the direct impact of the transfer and to assess enough sessions to get a good picture of posttransfer development. Three sessions are the minimum requirement to investigate linear developments (Hox, Moerbeek, & van de Schoot, 2010). Therefore, at least three sessions should be investigated. One could argue that six or more sessions with one therapist is already quite long and that huge alliance differences should not be expected on an aggregated level. Another reason to take five and not more sessions into account was missing data; treatment termination increases with an increasing number of posttransfer sessions.

Multilevel models are generalizations of linear models that should be applied if data are nested (e.g., treatment sessions nested in patients and patients in therapists; Hox et al., 2010). An advantage of multilevel modelling is that predictors can be studied on different levels of data. Furthermore, interactions between levels can be investigated. In this study, individual repeated measures of alliance and symptom severity (level one) were nested within patients (level two) nested within therapists (level three). A potential predictor on level one was the number of session after transfer. To facilitate interpretation of the regression weights, the first session with the new therapist was coded 0, the second session 1, the third session 2, and so on. Potential predictors on level two were the pretransfer levels of alliance/symptom severity and transfer session (i.e., the last treatment session with the initial therapist). All level two variables were grand-mean centred.

Model fit was evaluated using the Akaike information criterion (AIC;e.g., Long, 2012; Vrieze, 2012) with lower values indicating better model fit. A stepwise growth model selection with AIC as the criterion was conducted to find the model that best predicted the development of patient- and therapist-rated alliance and symptom severity during the five sessions after the transfer. For each dependent variable (patient-rated alliance, therapist-rated alliance, and symptom severity), independent models were built. The number of session after transfer, pretransfer therapeutic alliance and symptom severity levels, and transfer session were included in the models as predictors. Cross-level interactions were also tested.

First, a two-level random intercept model was calculated to account for differences in therapeutic alliance and symptom severity in the sessions after the transfer attributable to patients. This model was compared with a three-level-model including the nesting of patients within therapists. The best-fitting unconditional means model

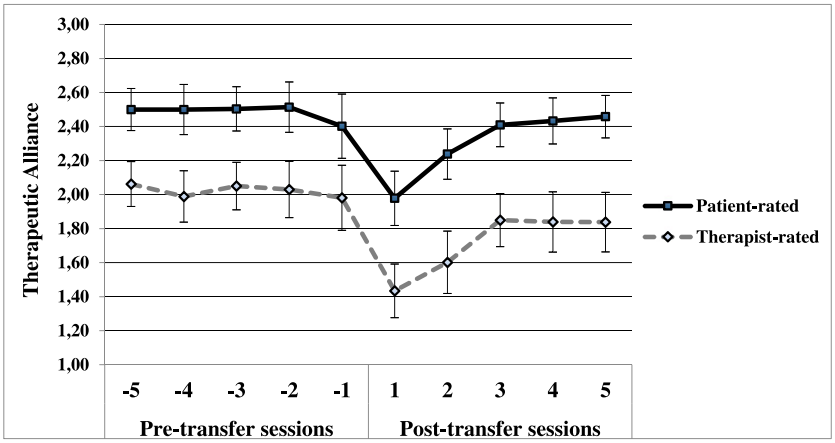

FIGURE 1 Observed therapist- and patient-rated alliance curves five sessions before and five sessions after the transfer with 95% confidence intervals [Colour figure can be viewed at wileyonlinelibrary.com]

(UMM; two-level-model vs. three-level-model) without containing any predictors was used to calculate the intraclass correlation (ICC; Model 1). For patient-rated alliance and symptom severity, a two-level-model was best fitting, whereas for therapist-rated alliance, a three-level-model was best fitting.[2] The ICC estimates the variance in alliance and symptom severity that could be attributed to patients (i.e., how much variance after the transfer can be attributed to differences between patients) and therapists (i.e., how much variance after the transfer of the patients can be attributed to differences between therapists). The ICC was calculated by dividing the variance of Level 2 or if applicable Level 3 variance by the total variance.

Second, a model was built, which included the level one predictor number of session after transfer to test for a growth in therapeutic alliance and symptom severity, respectively (Model 2). This model was compared with a model allowing the growth of therapeutic alliance and symptom severity to vary between patients (Model 3; random intercept and slope model). The best-fitting model (Model 2 vs. Model 3) was further used to investigate predictors on level two (pretransfer therapeutic alliance and symptom severity levels and transfer session). Therefore, predictors were entered into the model separately, and model fit was compared with the model without the predictor (Models 4–5). Finally, interactions between session and significant level two predictors of Models 4–5 were tested to check whether growth was moderated by level two predictors (Model 6). To distinguish between patient-rated alliance, therapist-rated alliance, and symptom severity, models were labelled with additional letters. Models with patient-rated alliance as the dependent variable were marked with an "a" (e.g., Model 2a), models predicting therapist-rated alliance were marked with a "b," and models predicting symptom severity were marked with a "c."

3 | RESULTS

There was a significant difference between the patient-rated therapeutic alliance one session before and one session after the transfer, $t(120) = 4.40$, $p < 0.000$, with higher values (i.e., a better therapeutic alliance) before the transfer. The mean difference was 0.42. There was also a significant difference of 0.55 scores between the therapist-rated therapeutic alliance before and after the transfer, $t(119) = 5.23$, $p < 0.000$, with higher values before the transfer (see Figure 1). No significant difference was observed between symptom severity levels before and after the transfer, $t(122) = -1.65$, $p = 0.10$ (see Figure 2).

3.1 | Results for patient-rated therapeutic alliance

The average score of the patient-rated therapeutic alliance after the transfer was 2.28 (*SD* = 0.06). An ICC coefficient of $\rho = 0.68$ was calculated with the UMM (Model 1a), which indicated a variability of 68% between patients regarding their therapeutic alliance ratings after the transfer. The model comparisons and the values for each model are shown in Table 2. The best fit (AIC = 914.9, marked in bold in Table 2) was reached by Model 6a with the predictors number of session after transfer on level one and pretransfer alliance level on level two as well as their cross-level interaction in a random intercept and random slope model. The model's equations were

$$\begin{aligned} \text{Level:1} \quad & y_{ij} = \beta_{0j} + \beta_{1j}(\text{number of session after transfer}_{ij}) + \varepsilon_{ij}, \\ \text{Level:2} \quad & \beta_{0j} = \gamma_{00} + \gamma_{01}(\text{pretransfer alliance level}_j) + b_{0j}, \\ & \beta_{1j} = \gamma_{10} + \gamma_{11}(\text{pretransfer alliance level}_j) + b_{1j}. \end{aligned}$$

Patient-rated alliance after the transfer increased significantly with the amount of time having passed after the transfer (i.e., increase of 0.11 scores in the therapeutic alliance per additional session). A higher pretransfer alliance level was associated with better alliance after the transfer. Patients developed in various ways during the five sessions after the transfer and, therefore, differed vastly with regard to their individual fitted curves. Some patients reported considerably lower therapeutic alliance levels than the average after the transfer. Some patients' alliance ratings rapidly increased in the five sessions after the transfer whereas others showed less rapid increases. On average, it took 2.93 sessions to reach the average pretransfer therapeutic alliance score. Taking the 95% confidence interval into account, 1.28 (lower boundary) to 4.61 (upper boundary) sessions were necessary to reach pretransfer alliance levels.

3.2 | Results for therapist-rated therapeutic alliance

The mean value of the therapist-rated therapeutic alliance after the transfer was 1.69 (*SD* = 0.08). The ICC of the UMM (Model 1b)

[2]The variance on the therapist level was estimated to be zero for patient-rated alliance and symptom severity.

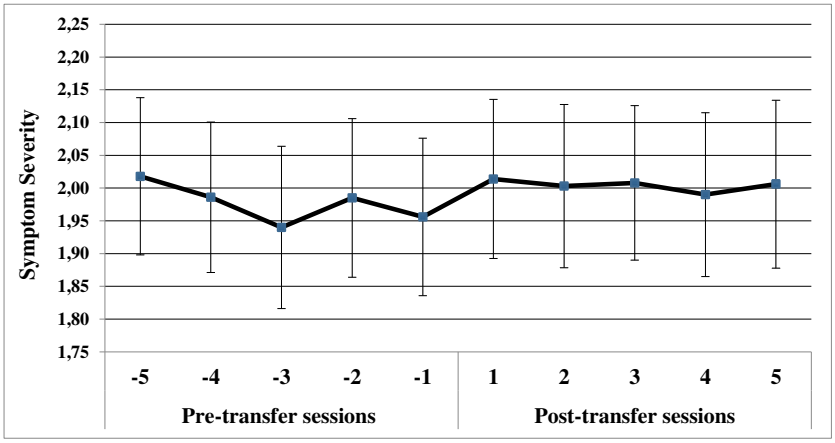

FIGURE 2 Observed curve of symptom severity five sessions before and five sessions after the transfer with 95% confidence intervals [Colour figure can be viewed at wileyonlinelibrary.com]

indicates that 24% of variance of the alliance after the transfer was due to differences between patients and that 36% are due to differences between therapists. The model comparison and values for each model are shown in Table 3. Best model fit (AIC = 1309.0, marked in bold in Table 3) was reached by Model 4b with the number of session after transfer and pretransfer therapeutic alliance predictors and random intercepts for both patient and therapist. The model's equations were

$$\text{Level 1: } \gamma_{ijk} = \beta_{0jk} + \beta_{1jk}\,(\text{number of session after transfer}_{ijk}) + \varepsilon_{ijk},$$
$$\text{Level 2: } \beta_{0jk} = \gamma_{00k} + \gamma_{01k}\,(\text{pretransfer alliance level}_{jk}) + b_{0jk},$$
$$\beta_{1ik} = \gamma_{10k},$$
$$\text{Level 3: } \gamma_{00k} = \delta_{000} + V_{00k},$$
$$\gamma_{10k} = \delta_{100},$$
$$\gamma_{01k} = \delta_{200}.$$

The therapist-rated alliance after the transfer increased significantly with the amount of time having passed after the transfer, a higher pretransfer alliance level, and a longer treatment duration. For the therapist-rated alliance, model fit did not improve when considering inter-individual differences in patient development over the five sessions after the transfer. On average, it took 5.05 sessions for the new therapists to reach the average pretransfer therapeutic alliance levels. Taking the 95% confidence interval into account, 2.97 (lower boundary) to 7.14 (upper boundary) sessions were necessary to reach pretransfer alliance levels.

3.3 | Results for symptom severity

The mean value of symptom severity after the transfer was 2.00 (*SD* = 0.06). The ICC of the UMM (Model 1c) was $\rho = 0.85$, indicating 85% of variability of symptom severity was due to differences between patients. The model comparisons and values for each model are shown in Table 4. Best model fit (AIC = 350.7, marked in bold in Table 4) was reached by Model 4c with number of session after transfer and pretransfer symptom severity, as predictors in a random intercept and a random slope model. The model's equations were

TABLE 2 Multilevel models for patient-rated alliance over five sessions after the transfer

Model	1a	2a	3a	4a	5a	6a
Fixed effects						
Intercept	2.28 (0.06)***	2.07 (0.07)***	2.07 (0.08)***	2.07 (0.07)***	2.07 (0.08)***	2.07 (0.07)***
Number of session after transfer	—	0.11 (0.01)***	0.11 (0.02)***	0.11 (0.02)***	0.11 (0.02)***	0.11 (0.02)***
Pretransfer therapeutic alliance	—	—	—	0.29 (0.05)***	—	0.36 (0.07)***
Transfer session	—	—	—	—	0.01 (0.00)	—
Pretransfer therapeutic alliance × Number of session after transfer	—	—	—	—	—	−0.02 (0.02)
Random effects						
Level 1						
Residual	0.21 (0.46)	0.18 (0.43)	0.14 (0.38)	0.14 (0.38)	0.14 (0.38)	0.14 (0.38)
Level 2						
Intercept	0.45 (0.67)	0.45 (0.67)	0.67 (0.82)	0.53 (0.73)	0.65 (0.80)	0.52 (0.72)
Number of session after transfer	—	—	0.02 (0.12)	0.02 (0.12)	0.02 (0.12)	0.02 (0.12)
AIC	1,048.7	977.4	944.0	915.1	945.0	**914.9**
BIC	1,061.8	994.9	970.3	**945.7**	975.6	949.9

Note. AIC: Akaike information criterion; BIC: Bayesian information criterion. Standard errors appear in parentheses. Best model fits are marked in bold.

*$p < 0.05$.

**$p < 0.01$.

***$p < 0.001$.

TABLE 3 Multilevel models for therapist-rated alliance over five sessions after the transfer

Model	1b	2b	3b	4b	5b	6b
Fixed effects						
Intercept	1.69 (0.09)***	1.51 (0.09)***	1.51 (0.09)***	1.51 (0.09)***	1.51 (0.09)***	1.51 (0.09)***
Number of session after transfer	—	0.09 (0.02)***	0.09 (0.02)***	0.09 (0.02)***	0.09 (0.02)***	0.09 (0.02)***
Pretransfer therapeutic alliance	—	—	—	0.18 (0.06)**	—	0.22 (0.07)**
Transfer session	—	—	—	—	0.00 (0.01)	—
Pretransfer therapeutic alliance × Number of session after transfer	—	—	—	—	—	−0.02 (0.02)
Random effects						
Level 1						
Residual	0.39 (0.62)	0.37 (0.61)	0.35 (0.59)	0.37 (0.61)	0.37 (0.61)	0.37 (0.61)
Level 2						
Intercept	0.24 (0.49)	0.24 (0.49)	0.57 (0.76)	0.23 (0.48)	0.25 (0.50)	0.23 (0.47)
Number of session after transfer	—	—	0.00 (0.01)	—	—	—
Level 3						
Intercept	0.36 (0.60)	0.35 (0.59)	0.32 (0.56)	0.32 (0.56)	0.35 (0.59)	0.32 (0.56)
Number of session after transfer	—	—	0.01 (0.09)	—	—	—
AIC	1,339.7	1,315.7	1,318.1	**1,309.0**	1,317.7	1,309.9
BIC	1,357.2	1,337.5	1,357.3	**1,335.1**	1,343.8	1,340.4

Note. AIC: Akaike information criterion; BIC: Bayesian information criterion. Standard errors appear in parentheses. Best model fits are marked in bold.

*$p < 0.05$.

**$p < 0.01$.

***$p < 0.001$.

TABLE 4 Multilevel model results for symptom severity over five sessions after the transfer

Model	1c	2c	3c	4c	5c	6c
Fixed effects						
Intercept	2.00 (0.06)***	2.01 (0.06)***	2.01 (0.06)***	2.01 (0.03)***	2.01 (0.06)***	2.01 (0.03)***
Number of session after transfer	—	−0.01 (0.01)	−0.01 (0.01)	−0.01 (0.01)	−0.01 (0.01)	−0.01 (0.01)
Pretransfer symptom severity	—	—	—	0.84 (0.04)***	—	0.86 (0.04)***
Transfer session	—	—	—	—	−0.00 (0.00)	—
Pretransfer symptom severity × Number of session after transfer	—	—	—	—	—	−0.01 (0.01)
Random effects						
Level 1						
Residual	0.07 (0.27)	0.07 (0.27)	0.06 (0.25)	0.06 (0.25)	0.06 (0.25)	0.06 (0.25)
Level 2						
Intercept	0.39 (0.62)	0.39 (0.63)	0.42 (0.65)	0.08 (0.27)	0.42 (0.65)	0.08 (0.28)
Number of session after transfer	—	—	0.00 (0.07)	0.00 (0.07)	0.01 (0.07)	0.00 (0.07)
AIC	550.7	551.9	540.1	**350.7**	541.4	351.5
BIC	563.9	569.5	566.5	**381.4**	572.2	386.7

Note. AIC: Akaike information criterion; BIC: Bayesian information criterion. Standard errors appear in parentheses. Best model fits are marked in bold.

*$p < 0.05$.

**$p < 0.01$.

***$p < 0.001$.

Level 1: $\gamma_{ij} = \beta_{0j} = \beta_{1j}(\text{number of session after transfer}_{ij}) + \varepsilon_{ij}$.

Level 2:
$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{pretransfer symptom severity level}_j) + b_{0j},$$
$$\beta_{1j} = \gamma_{10} + \gamma_{1j}.$$

Symptom severity after the transfer was significantly higher with a higher pretransfer level of symptom severity.

4 | DISCUSSION

The present study aimed to elucidate the potential impact of therapy transfer on the quality of the therapeutic alliance and symptom severity. First, it was hypothesized that the therapeutic alliance, rated by both patients and therapists, would show a significant drop in the first

session with the new therapist. Second, an increase of symptom severity was expected. The first hypothesis was confirmed, as a significant drop in alliance ratings was observed. However, no significant increase of symptom severity was found.

The development of alliance and symptom severity was investigated in an explorative multilevel modelling approach. For both the patient- and therapist-rated therapeutic alliance, patients showed an increase over time after the transfer session. This indicates that the observed drop in the therapeutic alliance was able to be restored. It was shown that an average of three to five sessions were necessary to reach the average pretransfer level of therapeutic alliance. Patients and therapists differed in their perception of alliance development after the transfer. Alliance development differed substantially between patients (i.e., a random slope for number of session after transfer), but this was not the case for therapists (i.e., fixed slope for number of session after transfer). This means, in an average case, patient-rated alliance reaches the level as before within three sessions. However, there are differences between patients, indicating that this is not true for every case. Some patients might benefit from a therapy transfer in terms of the therapeutic alliance, whereas others might not restore the same strong alliance they had with the previous therapist.

The point in time during therapy when the transfer occurred had no effect on the therapeutic alliance. How long the treatment lasts with the old therapist does not seem to have an impact of the development of the alliance with the new therapist.

The best-fitting model for patient-rated alliance included a cross-level interaction between number of session after transfer and pretransfer therapeutic alliance. Patients with lower alliance levels before the transfer had lower alliance scores after the transfer and showed a stronger increase over time compared with patients with high alliance scores. However, the impact of this interaction was rather small. Although the AIC was slightly better compared with Model 4a (914.9 vs. 915.1), this was not the case when taking the BIC into account. Therefore, this result should be interpreted with great caution. The significant main effect of pretransfer therapeutic alliance indicated that the alliance level with the previous therapist is not independent of the alliance level with the subsequent therapist. This result might reflect that patients differ with regard to how they interpret certain constructs such as the therapeutic alliance. From a clinical perspective, it could mean that some patients might feel better in relations than others. From a statistical perspective, this finding could simply underpin the fact that some patients rate the alliance lower than others because of a different "anchors" they have in mind.

The development of the therapist-rated alliance is unique in that the ratings before and after the transfer were provided by different therapists. In our view, it is important to consider both perspectives of the therapeutic alliance despite this limitation. The pretransfer therapeutic alliance level was a significant predictor of the alliance in the session after the transfer, so the new therapist's perception of the alliance seems not to be independent of the perception of his or her predecessor. This finding suggests that patients bring their own "ability to form a therapeutic alliance" along, no matter who has previously treated them. Possibly, personality traits such as agreeableness or extraversion could explain differences in the ability to form a therapeutic alliance. The importance of the separate consideration of both perspectives is underlined by the fact that patient- and therapist-rated alliance levels were predicted best by different models. To reach pretransfer therapeutic alliance levels, it took an average of three to five sessions with a more rapid increase of the patient- than the therapist-rated alliance. These results are in line with other findings stating that therapist ratings of the alliance are often more conservative than the patient's perspective (Atzil-Slonim et al., 2015; Rubel, Zilcha-Mano, et al., 2018b).

To compare the development of alliance levels after a transfer with the beginning of the therapy with the original therapist, additional analyses were carried out. Therefore, the first five sessions of patient- and therapist-rated alliance at the beginning of treatment were analysed. The model for predicting therapist-rated alliance at the beginning of treatment had a lower intercept (1.35 vs. 1.51; i.e., patients began treatment with a slightly lower therapist-rated alliance compared to a "start after a transfer"). The rate of change was slightly higher (0.13 vs. 0.09) for the increase of alliance ratings by session (i.e., alliance ratings increased with 0.13 scores per session). For the patient-rated alliance, a slightly different picture arose: Patients also had a slightly lower alliance at the beginning of treatment compared with the first posttransfer session (intercept 1.95 vs. 2.07); however, increase of alliance per session was almost the same as during the posttransfer phase (0.10 vs. 0.11). The individual slopes for patient-rated alliance (i.e., the individual slope of the first sessions of treatment correlated with posttransfer session slopes) showed an association of medium strength ($r = 0.39$, $p < 0.001$). These results suggest that a new alliance has to be built both at the beginning of therapy and after a transfer to a new therapist. There seem to be no major difficulties in the development of the alliance after a transfer. The positive association between individual patient-rated alliance slopes shows that patients tend to develop comparably in both situations. However, it also shows that these developments differ for the same patient, indicating that a therapy transfer is a critical event within a treatment.

Patients' previous experiences may also have an influence on the impact of transfers. Patients who have already had more than one therapy may be more flexible and open to a transfer. However, these patients may also perceive a transfer even more negatively, as they are unwilling to adapt to another therapist yet again. Wapner et al. (1986) examined demographic and clinical factors that may help to predict successful transfer outcomes in a psychology training clinic using archival clinical data. Only the history of previous therapy experience showed a trend effect ($p < 0.065$). Patients who had a previous therapy were more likely to have a successful transfer. The null effects associated with the majority of patient, therapist, and treatment variables evaluated by the authors suggested that many of the purported treatment and demographic factors discussed in the literature did not help to predict successful transfer cases. Clark, Robertson, Keen, and Cole (2011) found similar results, namely, that many patient and therapist variables were not significantly correlated with successful transfers. Only the number of sessions missed prior to transfer and the number of therapy transfer sessions were significantly correlated. They concluded that at least four cotherapy transfer sessions (i.e., sessions with both the old and the new therapist present) are appropriate to optimize successful transfers. Another factor to enhance successful

therapy transfer might be clinical supervision. Although results suggest that transfers do not have negative impacts on average and that a good therapeutic alliance can be developed with the new therapist, there are cases where this is not the case. Motivating the original therapist to discuss the transfer with his or her patient and to educate his or her patient that a therapy transfer may also have beneficial effects (e.g., another perspective on the problem) could increase the chances of a successful transfer. Future studies should further enlighten which factors facilitate successful transfers.

As ruptures in the therapeutic alliance can lead to drop out and poorer therapy outcomes, routine outcome monitoring of both the alliance and the outcome may help to prevent these negative effects (Lutz et al., 2015; Lutz, Zimmermann, Müller, Deisenhofer, & Rubel, 2017). Especially in the situation of a therapy transfer, monitoring the therapeutic alliance may help to identify a negative impact on the patient at an early stage. This information could be used to discuss the issue with the patient and repair a potential rupture in the relationship (Safran & Muran, 2000). Future studies should focus on alliance ruptures and their impact on subsequent therapy transfers.

The hypothesis that a transfer to a new therapist leads to a significant increase of symptom severity immediately after the transfer could not be confirmed. This is in line with findings from Sauer et al. (2017) who did not find any direct negative effect for successful transfers. There were no significant overall effects on symptom severity. However, the results of the explorative multilevel modelling indicated that patients differed in their development after the transfer. Although the fixed effects showed an overall constant symptom severity level over time, the random effect model revealed that patients differ in their development after the transfer. The range of the individual slopes was between −0.14 and 0.12 suggesting that some patients improved (i.e., 72 patients with a negative slope), whereas other patients deteriorated (i.e., 52 patients with a positive slope). In effect, inter-individual differences were observed with regard to the impact of a transfer on symptom severity.

The best-fitting model predicting symptom severity after the transfer included the number of session after transfer and pretransfer symptom severity. In contrast to the models predicting alliance, there was no average linear improvement over time after the transfer, although developments between patients differed (i.e., a random slope for number of session after transfer). At a first glance, this finding might be in contrast to previous findings regarding dose–effect relations in psychotherapy (e.g., Howard, Kopta, Krause, & Orlinsky, 1986). However, there is evidence that patients having a high dosage (i.e., long treatment durations) show smaller average improvement rates per session (Baldwin, Berkeljon, Atkins, Olsen, & Nielsen, 2009). Taking the relatively long treatment durations into account, this might explain why, on average, no linear improvement was found. From research on patterns of change, we know that patients differ in terms of their trajectories over the course of treatment (Rubel et al., 2015). Taken together, these arguments might explain why, on average, no increase or decrease of symptom severity was observed in the sessions after the transfer but that patients showed positive and negative developments on an individual level.

Pretransfer symptom severity predicted symptom severity after the transfer in that way that higher impairment before the transfer was associated with higher impairment after the transfer. This finding strengthens results found by Sauer et al. (2017) in which Outcome Questionnaire-45 scores after a transfer were significantly predicted by the measure's pretransfer levels. Independent of therapy transfer, the finding that previous symptom levels predict subsequent symptom levels is well-known from process-outcome studies (Falkenström et al., 2017).

4.1 | Limitations

One limitation of this investigation was the studied sample. On the one hand, the studied subgroup had longer treatments on average than those who had no therapy transfer. The probability of a transfer to another therapist increases with the length of treatment as causes of therapy transfer (e.g., therapist finishes clinical training) become more likely. Also, the treatment length was rather high compared with most manualized cognitive behavioural treatments. A generalization of the results to patients with shorter treatments and to manualized treatments should therefore be made with caution. On the other hand, all patients who experienced a therapy transfer agreed on a continuation of the treatment with another therapist. This may lead to a bias in that more patients were in the sample, who were more open to a new therapist and less afraid of a change in the therapeutic process. These patients may therefore have been better able to establish a new alliance quickly and have been less likely to experience an increase of symptom severity. For those patients who did not agree to a transfer and left therapy, findings may be different. Unfortunately, no data were collected on patients who refused a therapy transfer. It may be beneficial to identify dropouts caused by upcoming transfers to evaluate the potential negative consequences of therapy transfer and the phenomenon of dropout to its full extent (Zimmermann, Rubel, Page, & Lutz, 2017). Another limitation is that no data were collected on the timing with which the original therapist informed the patient of a potential upcoming therapy transfer. Patients might respond differently, depending on whether they were informed many sessions before a transfer or immediately before a transfer.

Another potential limitation of the generalization of the results is the transfer from a CBT to another CBT therapist. Patients may not only attach to a therapist but also attach to the therapeutic model. A transfer to a therapist working within a similar treatment model might help the patient to adapt to the new therapist. This may be different, when a patient transfers from a CBT to a psychodynamic therapist, for instance.

The literature suggests that the quality of the therapeutic alliance with patients suffering from certain disorders such as personality disorders (e.g., borderline) is lower compared with Axis I disorders and that the association between alliance and outcome is less strong (Spinhoven, Giesen-Bloo, van Dyck, Kooiman, & Arntz, 2007). Consequently, transfers may have a potentially more detrimental impact on the alliance and symptom severity for some group of patients than for others. Conducting subsample analyses for different groups of diagnoses was not feasible in this study due to the small sample size. Additional studies should take these different patient groups into account to explore potential moderating effects of diagnoses on the impact of transfer processes.

Therapy transfers in training clinics occur on a regular basis as therapists graduate and leave the institute. It was therefore advantageous to investigate forced therapy transfers in this setting, as this sample is more easily accessible in training clinics. However, therapies are supervised at the university training clinic, and upcoming therapy transfers are discussed with the supervisor to facilitate a good transition from the old to the new therapist. This practice may help to protect the patient from a potential negative impact. Further, patients seeking treatment in a training clinic are willing to be treated by trainees and participate in research, which raises the question whether this population differs from those patients seeking treatment in private practices. Therapy transfer in other settings may have different impacts on patients and should therefore be investigated in further studies. The time frame of the study was almost a decade. Although unlikely, potential cohort effects in patients/therapists can therefore not be excluded.

5 | CONCLUSION

This study is one of the first to analyse the impact of therapy transfer on alliance and symptom severity. The development of patients' therapeutic alliance and symptoms after a transfer was studied on a session-to-session basis. In general, therapy transfer influenced the alliance ratings of patients and therapists. A significant drop in patient- and therapist-rated alliance was found post-transfer. The development of alliance after the transfer was differentially predicted for patient- and therapist-rated alliance; however, it is important to consider variability between patients. The alliance was restored within three to five sessions after a transfer. Overall, transfers had no effect on symptom severity; however, the findings suggest that patients' reactions to transfers differ. In summary, patients' inter-individual differences after a therapy transfer are notable and warrant further examination. Qualitative studies could help to gain a better understanding of different alliance developments.

ORCID

Dirk Zimmermann http://orcid.org/0000-0002-8764-6343

REFERENCES

Atzil-Slonim, D., Bar-Kalifa, E., Rafaeli, E., Lutz, W., Rubel, J., Schiefele, A.-K., & Peri, T. (2015). Therapeutic bond judgments: Congruence and incongruence. *Journal of Consulting and Clinical Psychology*, *83*(4), 773–784. https://doi.org/10.1037/ccp0000015

Baldwin, S. A., Berkeljon, A., Atkins, D. C., Olsen, J. A., & Nielsen, S. L. (2009). Rates of change in naturalistic psychotherapy: Contrasting dose–effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology*, *77*(2), 203. https://doi.org/10.1037/a0015235–211.

Barber, J. P., Connolly, M. B., Crits-Christoph, P., Gladis, L., & Siqueland, L. (2009). Alliance predicts patients' outcome beyond in-treatment change in symptoms. *Personality Disorders: Theory, Research, and Treatment*, *1*, 80–89. https://doi.org/10.1037/1949-2715.S.1.80

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., & Singmann, H. (2015). lme4: Linear mixed-effects models using Eigen and S4, 2014. *R Package Version*, *1*(4).

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice*, *16*(3), 252–260. https://doi.org/10.1037/h0085885

Castonguay, L. G., Goldfried, M. R., Wiser, S., Raue, P. J., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression: A study of unique and common factors. *Journal of Consulting and Clinical Psychology*, *64*(3), 497–504. https://doi.org/10.1037/0022-006X.64.3.497

Castonguay, L. G., Constantino, M. J., & Grosse Holtforth, M. (2006). The working alliance: Where are we and where should we go? *Psychotherapy*, *43*(3), 271–279. https://doi.org/10.1037/0033-3204.43.3.271

Clark, P., Robertson, J. M., Keen, R., & Cole, C. (2011). Outcomes of client transfers in a training setting. *The American Journal of Family Therapy*, *39*(3), 214–225. https://doi.org/10.1080/01926187.2010.531650

Cox, M. L. (2017). The role of the therapeutic alliance on the successful outcome of transfers in marriage and family therapy cases. All Theses and Dissertations. Advance Online Publication. https://doi.org/10.1002/eat.20303. 39, 677–684.

Derogatis, L. R. (1992). SCL-90-R: Administration, scoring and procedures manual for the R(evised) version and other instruments of the psychopathology rating scale series: Clinical psychometric research.

Falkenström, 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. https://doi.org/10.1037/ccp0000119–751.

Falkenström, F., Finkel, S., Sandell, R., Rubel, J. A., & Holmqvist, R. (2017). Dynamic models of individual change in psychotherapy process research. *Journal of Consulting and Clinical Psychology*, *85*(6), 537–549. https://doi.org/10.1037/ccp0000203

Falkenström, F., Granström, F., & Holmqvist, R. (2013). Working alliance predicts psychotherapy outcome even while controlling for prior symptom improvement. *Psychotherapy Research*, *24*(2), 146–159. https://doi.org/10.1080/10503307.2013.847985

Fitzpatrick, M. R., Iwakabe, S., & Stalikas, A. (2005). Perspective divergence in the working alliance. *Psychotherapy Research*, *15*(1–2), 69–80. https://doi.org/10.1080/10503300512331327056

Flowers, J. V., & Booraem, C. D. (1995). The cost of client transfer. *Journal of Counseling & Development*, *73*(5), 564–566. https://doi.org/10.1002/j.1556-6676.1995.tb01795.x

Flückiger, C., Del Re, A. C., Wampold, B. E., Symonds, D., & Horvath, A. O. (2012). How central is the alliance in psychotherapy? A multilevel longitudinal meta-analysis. *Journal of Counseling Psychology*, *59*(1), 10–17. https://doi.org/10.1037/a0025749

Flückiger, C., Grosse Holtforth, M., Znoj, H. J., Caspar, F., & Wampold, B. E. (2013). Is the relation between early post-session reports and treatment outcome an epiphenomenon of intake distress and early response? A multi-predictor analysis in outpatient psychotherapy. *Psychotherapy Research*, *23*(1), 1–13. https://doi.org/10.1080/10503307.2012.693773

Franke, G. H. (2000). *Brief Symptom Inventory von Derogatis (BSI)*. Göttingen: Beltz.

Gelso, C. J., & Carter, J. A. (1994). Components of the psychotherapy relationship: Their interaction and unfolding during treatment. *Journal of counseling psychology*, *41*, 296–306. https://doi.org/10.1037/0022-0167.41.3.296

Grawe, K. (1998). *Psychologische therapie*. Göttingen: Hogrefe.

Hersoug, A. G., Høglend, P., Havik, O., von der Lippe, A., & Monsen, J. (2009). Therapist characteristics influencing the quality of alliance in long-term psychotherapy. *Clinical Psychology & Psychotherapy*, *16*(2), 100–110. https://doi.org/10.1002/cpp.605

Horvath, A. O., & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy. *Journal of Consulting and Clinical Psychology*, *61*(4), 561–573. https://doi.org/10.1037/0022-006X.61.4.561

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, *38*(2), 139–149. https://doi.org/10.1037/0022-0167.38.2.139

Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy*, *48*(1), 9–16. https://doi.org/10.1037/a0022186

Horvath, A. O., & Greenberg, L. S. (1994). *The working alliance: Theory, research, and practice*. New York: John Wiley & Sons.

Howard, K. I., Kopta, M., Krause, M., & Orlinsky, D. (1986). The dose-effect response in psychotherapy. *American Psychologist*, *41*(2), 159–164. https://doi.org/10.1037/0003-066X.41.2.159

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications*. Routledge.

Kivlighan, D. M. Jr., Kline, K., Gelso, C. J., & Hill, C. E. (2017). Congruence and discrepancy between working alliance and real relationship: Variance decomposition and response surface analyses. *Journal of Counseling Psychology*, *64*(4), 394. https://doi.org/10.1037/cou0000216–409.

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

Long, J. D. (2012). *Longitudinal data analysis for the behavioral sciences using R*. London: Sage.

Lutz, W., Tholen, S., Schürch, E., & Berking, M. (2006). Die Entwicklung, Validität und Reliabilität von Kurzformen gängiger psychometrischer Instrumente zur Evaluation des therapeutischen Fortschrittes in Psychotherapie und Psychiatrie. *Diagnostica*, *52*, 11–25. https://doi.org/10.1026/0012-1924.52.1.11

Lutz, W., Rubel, J., Schiefele, A.-K., Zimmermann, D., Böhnke, J. R., & Wittmann, W. W. (2015). Feedback and therapist effects in the context of treatment outcome and treatment length. *Psychotherapy Research*, *25*(6), 647–660. https://doi.org/10.1080/10503307.2015.1053553

Lutz, W., Zimmermann, D., Müller, V. N. L. S., Deisenhofer, A.-K., & Rubel, J. A. (2017). Randomized controlled trial to evaluate the effects of personalized prediction and adaptation tools on treatment outcome in outpatient psychotherapy: Study protocol. *BMC Psychiatry*, *17*(1), 306. https://doi.org/10.1186/s12888-017-1464-2

Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, *68*(3), 438–450. https://doi.org/10.1037/0022-006x.68.3.438

Nevid, J. S., Ghannadpour, J., & Haggerty, G. (2017). The role of gender as a moderator of the alliance-outcome link in acute inpatient treatment of severely disturbed youth. *Clinical Psychology & Psychotherapy*, *24*(2), 528–533. https://doi.org/10.1002/cpp.2025

Penn, L. (1990). When the therapist must leave: Forced termination of psychodynamic therapy. *Professional Psychology: Research and Practice*, *21*(5), 379–384. https://doi.org/10.1037/0735-7028.21.5.379

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879. https://doi.org/10.1037/0021-9010.88.5.879–903.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. London: Sage.

Regli, D., & Grawe, K. (2000). Die Berner Stundenbogen Version 2000–ein Forschungsinstrument. Unpublished Manuscript. In *Universität Bern*.

Rubel, J. A., Bar-Kalifa, E., Atzil-Slonim, D., Schmidt, S., & Lutz, W. (2018a). Congruence of therapeutic bond perceptions and its relation to treatment outcome: Within- and between-dyad effects. *Journal of Consulting and Clinical Psychology*, *86*(4), 341–353. https://doi.org/10.1037/ccp0000280

Rubel, J. A., Zilcha-Mano, S., Feils-Klaus, V., & Lutz, W. (2018b). Session-to-session effects of alliance ruptures in outpatient CBT: Within- and between-patient associations. *Journal of Consulting and Clinical Psychology*, *86*(4), 354–366. https://doi.org/10.1037/ccp0000286

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

Rubel, J. A., Rosenbaum, D., & Lutz, W. (2017). Patients' in-session experiences and symptom change: Session-to-session effects on a within- and between-patient level. *Behaviour Research and Therapy*, *90*, 58–66. https://doi.org/10.1016/j.brat.2016.12.007

Safran, J. D., & Muran, J. C. (2000). *Negotiating the therapeutic alliance*. New York: Guilford Press.

Sasso, K. E., Strunk, D. R., Braun, J. D., DeRubeis, R. J., & Brotman, M. A. (2016). A re-examination of process-outcome relations in cognitive therapy for depression: Disaggregating within-patient and between-patient effects. *Psychotherapy Research*, *26*(4), 387–398. https://doi.org/10.1080/10503307.2015.1026423

Sauer, E. M., Lopez, F. G., & Gormley, B. (2003). Respective contributions of therapist and client adult attachment orientations to the development of the early working alliance: A preliminary growth modeling study. *Psychotherapy Research*, *13*(3), 371–382. https://doi.org/10.1093/ptr/kpg027. https://doi.org/10.1093/ptr/kpg027

Sauer, E. M., Rice, K. G., Richardson, C. M. E., & Roberts, K. E. (2017). Influence of client attachment and gender on therapy transfers: A multilevel examination. *Training and Education in Professional Psychology*, *11*(1), 33–40. https://doi.org/10.1037/tep0000134

Siebold, C. (1991). Termination: When the therapist leaves. *Clinical Social Work Journal*, *19*(2), 191–204. https://doi.org/10.1007/bf00755046

Spinhoven, P., Giesen-Bloo, J., van Dyck, R., Kooiman, K., & Arntz, A. (2007). The therapeutic alliance in schema-focused therapy and transference-focused psychotherapy for borderline personality disorder. *Journal of Consulting and Clinical Psychology*, *75*(1), 104. https://doi.org/10.1037/0022-006x.75.1.104–115.

Trimboli, F., & Keenan, C. W. (2010). Supervising student-therapists: Suggestions for the transfer of long-term psychotherapy patients. *Psychoanalytic Psychology*, *27*(3), 319–329. https://doi.org/10.1037/a0015484

Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*, *17*(2), 228–243. https://doi.org/10.1037/a0027127

Wapner, J. H., Klein, J. G., Friedlander, M. L., & Andrasik, F. J. (1986). Transferring psychotherapy clients: State of the art. *Professional Psychology: Research and Practice*, *17*(6), 492–496. https://doi.org/10.1037//0735-7028.17.6.492

Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, *78*(2), 200–211. https://doi.org/10.1037/a0018912

Wittchen, H.-U., Wunderlich, U., Gruschwitz, S., & Zaudig, M. (1997). *SKID I. Strukturiertes Klinisches Interview für DSM-IV. Achse I: Psychische Störungen*. Göttingen: Hogrefe.

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

Zilcha-Mano, S., & Errázuriz, P. (2017). Early development of mechanisms of change as a predictor of subsequent change and treatment outcome: The case of working alliance. *Journal of Consulting and Clinical Psychology*, *85*(5), 508–520. https://doi.org/10.1037/ccp0000192

Zimmermann, D., Rubel, J., Page, A. C., & Lutz, W. (2017). Therapist effects on and predictors of non-consensual dropout in psychotherapy. *Clinical Psychology & Psychotherapy*, *24*(2), 312–321. https://doi.org/10.1002/cpp.2022

How to cite this article: Zimmermann D, Lutz W, Reiser M, et al. What happens when the therapist leaves? The impact of therapy transfer on the therapeutic alliance and symptoms. *Clin Psychol Psychother*. 2019;26:135–145. https://doi.org/10.1002/cpp.2336

8. General discussion

The studies in the current dissertation represent three important contributions to the field of patient-focused research and research on therapists. Study I concentrated on the use of psychometric feedback and on differences between therapists. Study II focused on therapist effects concerning dropout rates and the prediction of dropout. Study III investigated the understudied phenomena of therapist transfer in an ongoing treatment. All studies used the same statistical approach, namely multilevel modeling, to analyze data and to estimate therapist differences. Data stemmed from the outpatient clinic at the University of Trier.

Study I expands the literature on psychometric feedback and therapist effects with a combination of both research areas. It was one of the first studies to investigate the differentiated usage of psychometric feedback in therapists undergoing clinical training. The frequency of use within therapists was compared to results from a study with experienced therapists (Lutz, Rubel et al., 2015). Additionally, systematic differences between therapists were analyzed. Results suggested that the use of feedback is not only patient-specific, but also therapist-specific. Some therapists seem to make use of the feedback information more frequently than others. The degree of therapist effects regarding the different ways feedback is used seemed to vary (Rubel, Zimmermann, Deisenhofer, Müller, & Lutz, 2017).

Study II enriches the field with findings concerning therapist effects on treatment dropout and the prediction of dropout. This study is one of a few analyzing therapist effects on the dichotomized outcome dropout using logistic multilevel modeling. Results suggested that therapists differ in their dropout rates, even when controlling for initial impairment. Patient dropout can be predicted early on, with higher initial impairment, male sex, lower education, less compulsive personality style, more histrionic personality style, and low treatment expectation being associated with a higher dropout likelihood (Zimmermann et al., 2017).

Study III broadens our knowledge on the understudied phenomenon of therapist transfer during ongoing treatment. Alliance and impairment were studied with multilevel models in the sessions after the transfer. Results suggested that therapist transfers have no direct negative effects on patients. No difficulties in alliance development with the subsequent therapist could be found. However, there seemed to be inter-individual differences between patients regarding how transfers affected impairment and alliance (Zimmermann et al., 2019).

The following sections draw general conclusions from the findings based on the three studies. Future research directions are summarized and limitations are discussed.

8.1. General conclusions and future research

Some general conclusions and future directions can be derived from considering the results of all three studies simultaneously. The main finding of all studies seems to be: the therapist matters. As pointed out above, studying the therapist has attracted less attention than conducting RCTs in efficacy research. Therefore, it seems rather important to focus on therapists and their impact on psychotherapy, which was the aim of this dissertation.

Study I suggested that therapists differ in their use of feedback information. Differences between therapists accounted for 27% of variance in the statement that therapists tried to improve the therapeutic relationship because of the psychometric feedback they received. The therapist effect was as high as 52% of variance in the statement that therapists assigned different homework due to the psychometric feedback. These findings imply that the decision on what to do with feedback information is not only patient- but also highly therapist-specific. Feedback systems were designed to help clinicians track patients' progress and to hinder treatment failure and dropout (Castonguay et al., 2013). Future studies need to identify which therapist behaviors affected by the psychometric feedback are those helping to prevent negative effects. Can general recommendations be given on what to do with the information provided by psychometric feedback or are case-specific modifications needed to improve outcome and to reduce dropout rates?

Study II focused on therapist effects on patient dropout. Deduced from findings on therapist effects on patient outcomes, it was assumed that therapists systematically vary in their dropout rates. A variance portion of approximately 6% was attributable to therapists in this study, which is comparable to numbers found in patient outcomes, which are usually between 5-8% (Lutz & Barkham, 2015). Studies by other research groups following study II confirmed therapist effects on patient dropout. A study investigating 10,521 patients treated by 85 therapists in a naturalistic setting found therapist effects accounted for 12.6% of dropout variance (Saxon, Barkham, Foster, & Parry, 2017). Therapist dropout rates varied widely between 1.2% and 73.2%. In the same vein, a study investigating 10,147 patients treated by 481 therapists in a college counseling center found a therapist effect of 9.51% (Xiao, Castonguay et al., 2017). Another recent paper found varying therapist effects depending on the stage of treatment. While little therapist variance was found for early treatment attendance, over 40% of client attendance in a later therapy phase could be attributed to therapists (Xiao, Hayes, Castonguay, McAleavey, & Locke, 2017). A study that investigated therapist effects on nonattendance and took racial/ethnic disparities into account found that 14% of variance in client's nonattendance was

attributable to therapists (Kivlighan, Jung, Berkowitz, Hammer, & Collins, 2018). Some therapists had higher dropout rates compared to others when treating patients coming from a racial/ethnic minority in comparison to white clients.

The research body on therapist effects on dropout is growing and the finding that therapists differ in their dropout rates seems to be confirmed in different settings. However, there are still a number of unanswered questions. Most of the studies in that area stem from naturalistic settings which usually come with the advantage of larger data sets and a lot of patients treated by the same therapist. This is a methodological requisite when analyzing therapist effects (Schiefele et al., 2017). Due to the methodological need for large sample sizes and a need for many patients treated by the same therapist, it is difficult to get good estimates of the therapist effects on dropout in RCTs. Although estimates of the therapist effects can be biased in small samples, neglecting the therapist variable might distort the estimated treatment effects (Jong, Moerbeek, & van der Leeden, 2010). Therefore, the therapist variable should always be included in the analyses.

There have been efforts to better understand the factors associated with therapist effects on patient outcomes. It seems as though therapist effects become larger when patients are more distressed (Saxon & Barkham, 2012). One recent study found that therapists' burnout level might explain part of the therapist effect (Delgadillo, Saxon, & Barkham, 2018). However, research on factors explaining therapist effects on dropout is sparse and there seems to be no direct link between findings on treatment outcome and findings on dropout. Therapists who have more deteriorating patients are not the same therapists who have higher dropout rates (Saxon et al., 2017). Another study found no relation between therapist's treatment length and outcomes (Lutz, Rubel et al., 2015). In other words, therapists who have better outcomes are not those with longer or shorter treatment lengths. Future research is needed to understand why therapists differ in their dropout rates. Qualitative research might add insight as to why patients leave treatment prematurely and which factors might be attributable to therapists.

Another aspect for future research is the combination of the different studies. The large differences between therapists concerning the use of psychometric feedback might, on the one hand, explain why feedback interventions do not help every therapist (De Jong et al., 2012). On the other hand, differences in the use of feedback might explain why therapists differ in their dropout rates as found in study II. Feedback interventions were designed to improve outcomes and to prevent treatment dropouts (Castonguay et al., 2013). Those therapists not taking advantage of the feedback system might be the same therapists who face higher dropout rates.

One might deduce further research questions when combining findings from both study I and study II. Are those therapists not taking advantage of the psychometric feedback the same therapists who have to face higher dropout rates? Could differences in the feedback usage explain therapist effects on patient dropout?

A further combination of both studies is feedback to therapists on patient dropout risk early in treatment. Based on findings from study II and with the help of the assessment of crucial patient characteristics, an individual risk level for premature treatment termination can be calculated. The feasibility and benefits of feeding this information back to therapists is being tested in an ongoing study within the outpatient clinic at the University of Trier (Lutz et al., 2017).

Study III investigated the impact of therapy transfer during ongoing treatment. Although no negative effects were found on average, some patients had difficulties building a good therapeutic alliance with the subsequent therapist or even dropped out of treatment. Subsequent studies need to investigate the features of patients who struggle with therapy transfers. If shared features can be found, one might be able to identify patients at risk at an early stage. An upcoming therapy transfer could then be discussed in supervision or intervision to avoid negative effects. Furthermore, training programs could be developed to improve therapist's ability to detect patients at risk. Therapists could be trained in how to prepare patients for detachment and to motivate patients to engage in a new relationship with the subsequent therapist.

Future research is needed to study therapist effects with regard to therapy transfer. One might assume that therapists differ in their ability to prepare patients for an upcoming transfer. This might result in either more patients disagreeing to a transfer or in more patients having problems building a good relationship with the subsequent therapist. Problems engaging with the new therapist might be expressed via symptom impairment or dropout. To detect these differences, larger data sets are necessary. A large number of "leaving therapists" with an adequate number of patients per therapist need to be studied to derive precise estimates of potential therapist effects (Schiefele et al., 2017). The same is true when studying the subsequent therapists. There might also be differences in the ability of subsequent therapists to engage patients who just ended the therapeutic relationship with their former therapist. For example, therapists might feel misvalued when a patient makes comparisons to his preceding therapist.

Regardless of the strengths the discussed studies have, they are not free of limitations. General limitations of the studies are discussed in the subsequent section.

8.2. General limitations

All studies were based on data from the outpatient clinic at the University of Trier. Therapists were trainees in a three- to five-year postgraduate program. Therefore, generalization of the results to other settings must be made with caution. Therapists were relatively young and unexperienced in comparison to other mental health service organizations. Treatments of trainee therapists were supervised every fourth session. This extensive support from experienced clinicians is again not necessarily similar to other settings. All therapies were assisted with a modern and extensive feedback system (Lutz et al., 2017). This setting is therefore not completely comparable to other outpatient clinics.

Another limitation all three studies share is the sample size. Although the studied samples were rather large compared to most RCTs, sample sizes are still small when investigating therapist effects (Schiefele et al., 2017). Estimates of the therapist effects from study I and study II must be interpreted with a certain amount of caution. The sample size in study III was adequate to study the impact of therapy transfer on the patient level, however, much larger data sets are necessary to study effects on the therapist level.

A major limitation of study III is the agreed-upon therapist transfer. All patients consented to continue treatment after the therapist left. However, no information was available on which patients decided to discontinue treatment after being informed that a therapist transfer would ensue. The result that a therapy transfer has no negative effects on average, may not hold if patients are included who disagreed to a therapy transfer.

All studies were based on data from a naturalistic setting within a patient-focused research strategy. No control groups were used in the studies. All results were correlational in nature and therefore causal inferences cannot be drawn.

8.3. Concluding remarks

The aim of this dissertation was to add knowledge to the field of patient-focused research by studying therapists and their impact on psychotherapy. Despite the above mentioned limitations to the generalizability of the results, one can conclude that focusing on therapists might help to further improve psychotherapeutic treatments. Understanding the factors in which therapists differ in the use of feedback might help to improve future feedback interventions. Understanding the impact of therapists on patient dropout can help to identify patients at risk and help to prevent dropouts. Understanding that therapy transfers do not seem to have direct negative effects on patients is relieving. All three studies are more or less starting points in their

specific fields. More research is needed to investigate whether findings can be replicated in different settings and to better understand therapist effects. Differences in therapists can be seen as an opportunity to further improve treatments and to strengthen good therapist behaviour in clinical training.

9. References

Atkinson, D. R., & Schein, S. (1986). Similarity in counseling. *The Counseling Psychologist*, *14*(2), 319–354.

Baldwin, S. A., & Imel, Z. E. (2013). Therapist Effects: Findings and Methods. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (6th ed., pp. 258–297). New York: Wiley & Sons.

Bandelow, B., Lichte, T., Rudolf, S., Wiltink, J., & Beutel, M. (2014). *S3-Leitlinie Angststörungen*. Berlin: Springer-Verlag.

Barkham, M., Lutz, W., Lambert, M. J., & Saxon, D. (2017). Therapist effects, effective therapists, and the law of variability. In L. G. Castonguay & C. E. Hill (Eds.), *How and why are some therapists better than others?: Understanding therapist effects* (pp. 13–36). Washington, DC: American Psychological Association. https://doi.org/10.1037/0000034-002

Barth, J., Munder, T., Gerger, H., Nüesch, E., Trelle, S., Znoj, H., . . . Tsai, A. C. (2013). Comparative Efficacy of Seven Psychotherapeutic Interventions for Patients with Depression: A Network Meta-Analysis. *PLoS Medicine*, *10*(5), e1001454. https://doi.org/10.1371/journal.pmed.1001454

Beutler, L. E., Machado, P. P. P., & Allstetter Neufeldt, S. (1994). Therapist Variables. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 229–296). New York: Wiley.

Beutler, L. E., Malik, M., Alimohamed, S., Harwood, T. M., Talebi, H., Noble, S., & Wong, E. (2004). Therapist variables. In M. Lambert (Ed.), *Handbook of Psychotherapy and Behavior Change* (5th ed., pp. 227–306). New York: Wiley.

Castonguay, L., Barkham, M., Lutz, W., & McAleavey, A. (2013). Practice-Oriented Research. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (6th ed., pp. 85–133). New York: Wiley & Sons.

Crits-Christoph, P., Baranackie, K., Kurcias, J., Beck, A., Carroll, K., Perry, K., . . . Zitrin, C. (1991). Meta-Analysis of Therapist Effects in Psychotherapy Outcome Studies. *Psychotherapy Research*, *1*(2), 81–91. https://doi.org/10.1080/10503309112331335511

Cuijpers, P., Geraedts, A. S., van Oppen, P., Andersson, G., Markowitz, J. C., & van Straten, A. (2011). Interpersonal Psychotherapy for Depression: A Meta-Analysis. *American Journal of Psychiatry*, *168*(6), 581–592. https://doi.org/10.1176/appi.ajp.2010.10101411

Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds, C. F. (2014). Adding psychotherapy to antidepressant medication in depression and anxiety disorders: a meta-analysis. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, *13*(1), 56–67. https://doi.org/10.1002/wps.20089

Cuijpers, P., van Straten, A., Andersson, G., & van Oppen, P. (2008). Psychotherapy for depression in adults: A meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology*, *76*(6), 909–922. https://doi.org/10.1037/a0013075

De Jong, K., van Sluis, P., Nugter, M. A., Heiser, W. J., & Spinhoven, P. (2012). Understanding the differential impact of outcome monitoring: Therapist variables that moderate feedback effects in a randomized clinical trial. *Psychotherapy Research*, *22*(4), 464–474. https://doi.org/10.1080/10503307.2012.673023

Delgadillo, J., Overend, K., Lucock, M., Groom, M., Kirby, N., McMillan, D., . . . Jong, K. de. (2017). Improving the efficiency of psychological treatment using outcome feedback technology. *Behaviour Research and Therapy*, *99*, 89–97. https://doi.org/10.1016/j.brat.2017.09.011

Delgadillo, J., Saxon, D., & Barkham, M. (2018). Associations between therapists' occupational burnout and their patients' depression and anxiety treatment outcomes. *Depression and Anxiety*, *35*(9), 844–850. https://doi.org/10.1002/da.22766

Dembo, R., Iklé, D. N., & Ciarlo, J. A. (1983). The influence of client–clinician demographic match on client treatment outcomes. *Journal of Psychiatric Treatment & Evaluation*.

DGPPN, B. Ä.K., KBV, AWMF, AkdÄ, BPtK, BApK, . . . DGRW. (2009). S3-Leitlinie/Nationale Versorgungs Leitlinie Unipolare Depression-Langfassung 1. Auflage. Version 5. Retrieved from www.depression.versorgungsleitlinien.de/

Eid, M., Gollwitzer, M., & Schmitt, M. (2017). *Statistik und Forschungsmethoden*. Weinheim: Beltz.

Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology*, *16*, 319–324.

Fairburn, C. G., Bailey-Straebler, S., Basden, S., Doll, H. A., Jones, R., Murphy, R., . . . Cooper, Z. (2015). A transdiagnostic comparison of enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy in the treatment of eating disorders. *Behaviour Research and Therapy*, *70*, 64–71. https://doi.org/10.1016/j.brat.2015.04.010

Falkenström, F., Markowitz, J. C., Jonker, H., Philips, B., & Holmqvist, R. (2013). Can psychotherapists function as their own controls? Meta-analysis of the crossed therapist design in comparative psychotherapy trials. *The Journal of Clinical Psychiatry*, *74*(5), 482–491. https://doi.org/10.4088/JCP.12r07848

Freedheim, D. K., Freudenberger, H. J., Kessler, J. W., Messer, S. B., Peterson, D. R., Strupp, H. H., & Wachtel, P. L. (1992). *History of psychotherapy: A century of change*. Washington: American Psychological Association.

Goldfried, M. R., & Wolfe, W. (1996). Psychotherapy practice and research: Repairing a strained alliance. *American Psychologist*, *51*, 1007–1016. https://doi.org/10.1037/0003-066X.51.10.1007

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology*, *61*(2), 155–163. https://doi.org/10.1002/jclp.20108

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and it's implication for treatment delivery systems. *Clinical Psychology: Science and Practice*, *9*, 329–343. https://doi.org/10.1093/clipsy.9.3.329

Hatfield, D., McCullough, L., Frantz, S. H. B., & Krieger, K. (2010). Do we know when our clients get worse? an investigation of therapists' ability to detect negative client change. *Clinical Psychology & Psychotherapy*, *17*(1), 25–32. https://doi.org/10.1002/cpp.656

Hollon, S. D., Stewart, M. O., & Strunk, D. (2006). Enduring effects for cognitive behavior therapy in the treatment of depression and anxiety. *Annual Review of Psychology*, *57*, 285–315. https://doi.org/10.1146/annurev.psych.57.102904.190044

Howard, K. I., Moras, K., Brill, P., Martinovich, Z., & Lutz, W. (1996). The evaluation of psychotherapy: American Psychologist, *52*, 1059–1064.

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications*: Routledge.

Imel, Z. E., Malterer, M. B., McKay, K. M., & Wampold, B. E. (2008). A meta-analysis of psychotherapy and medication in unipolar depression and dysthymia. *Journal of Affective Disorders*, *110*(3), 197–206. https://doi.org/10.1016/j.jad.2008.03.018

Jackson, S. W. (1999). *Care of the psyche: A history of psychological healing*: Yale University Press.

Jong, K. de, Moerbeek, M., & van der Leeden, R. (2010). A priori power analysis in longitudinal three-level multilevel models: an example with therapist effects. *Psychotherapy Research*, *20*(3), 273–284. https://doi.org/10.1080/10503300903376320

Kazdin, A. E. (2000). *Psychotherapy for children and adolescents: Directions for research and practice*: Oxford University Press.

Kendrick, T., El-Gohary, M., Stuart, B., Gilbody, S., Churchill, R., Aiken, L., . . . Jong, K. de. (2016). Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. *The Cochrane Library*, *7*. https://doi.org/10.1002/14651858.CD011119.pub2

Kiesewetter, K. (2010). *Franz Anton Mesmers Leben und Lehre: mit einer Vorgeschichte des Mesmerismus, Hypnotismus und Somnambulismus*. Leipzig: Bohmeier.

Kivlighan, D. M., Jung, A.-K., Berkowitz, A. B., Hammer, J. H., & Collins, N. M. (2018). 'To show or No-show'? Therapist Racial-ethnic Disparities in Clients' Nonattendance in Therapy. *Journal of College Student Psychotherapy*, *6*(1), 1–13. https://doi.org/10.1080/87568225.2018.1424597

Knaup, C., Koesters, M., Schoefer, D., Becker, T., & Puschner, B. (2009). Effect of feedback of treatment outcome in specialist mental healthcare: meta-analysis. *The British Journal of Psychiatry*, *195*(1), 15–22.

Lafferty, P., Beutler, L. E., & Crago, M. (1989). Differences between more and less effective psychotherapists: A study of select therapist variables. *Journal of Consulting and Clinical Psychology*, *57*(1), 76.

Lambert, M. J. (2013a). Introduction and Historical Overview. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (6th ed., pp. 3–20). New York: Wiley & Sons.

Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 139–193). New York, NY: Wiley.

Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis: Clinical Psychology: Science and Practice, *10*, 288–301. https://doi.org/10.1093/clipsy.bpg025

Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A., Nielsen, S. L., & Hawkins, E. J. (2001). The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research*, *11*, 49–68. https://doi.org/10.1080/713663852

Lambert, M. J. (1989). The individual therapist's contribution to psychotherapy process and outcome. *Clinical Psychology Review*, *9*(4), 469–485.

Lambert, M. J. (2010). *Prevention of treatment failure: The use of measuring, monitoring, and feedback in clinical practice*: American Psychological Association.

Lambert, M. J. (2013b). The Efficacy and Effectiveness of Psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (6th ed., pp. 169–218). New York: Wiley & Sons.

Lambert, M. J., & Shimokawa, K. (2011). Collecting client feedback. *Psychotherapy*, *48*(1), 72–79. https://doi.org/10.1037/a0022238

Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy (Chicago, Ill.)*, *55*(4), 520–537. https://doi.org/10.1037/pst0000167

Leichsenring, F., & Leibing, E. (2003). The effectiveness of psychodynamic and cognitive behavior therapy in the treatment of personality disorders: A meta-analysis: American Journal of Psychiatry, *160*, 1223–1232. https://doi.org/10.1176/appi.ajp.160.7.1223

Luborsky, L., Crits-Christoph, P., McLellan, A. T., Woody, G., Piper, W., Liberman, B., . . . Pilkonis, P. (1986). Do Therapists vary much in their Success? Findings from Four Outcome Studies. *American Journal of Orthopsychiatry*, *56*(4), 501–512.

Lutz, W. (2002). Patient-focused psychotherapy research and individual treatment progress as scientific groundwork for an empirical based clinical practice: Psychotherapy Research, *12*, 251–273. https://doi.org/10.1080/713664389

Lutz, W., & Böhnke, J. R. (2010). Psychotherapieforschung: Verläufer, Prozesse, Ergebnisse und Qualitätssicherung. In W. Lutz (Ed.), *Lehrbuch Psychotherapie* (pp. 49–70). Bern: Hans Huber.

Lutz, W., & Barkham, M. (2015). Therapist Effects. In R. L. Cautin & S. O. Lilienfeld (Eds.), *The encyclopedia of clinical psychology* (pp. 1–6). Chichester, West Sussex, Malden, MA: John Wiley and Sons, Inc. https://doi.org/10.1002/9781118625392.wbecp109

Lutz, W., Jong, K. de, & Rubel, J. (2015). Patient-focused and feedback research in psychotherapy: Where are we and where do we want to go? *Psychotherapy Research*, *25*(6), 625–632. https://doi.org/10.1080/10503307.2015.1079661

Lutz, W., Leon, S. C., Martinovich, Z., Lyons, J. S., & Stiles, W. B. (2007). Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology*, *54*(1), 32–39. https://doi.org/10.1037/0022-0167.54.1.32

Lutz, W., Rubel, J., Schiefele, A.-K., Zimmermann, D., Böhnke, J. R., & Wittmann, W. W. (2015). Feedback and therapist effects in the context of treatment outcome and treatment length. *Psychotherapy Research*, *25*(6), 647–660. https://doi.org/10.1080/10503307.2015.1053553

Lutz, W., Zimmermann, D., Müller, V. N. L. S., Deisenhofer, A.-K., & Rubel, J. A. (2017). Randomized controlled trial to evaluate the effects of personalized prediction and adaptation tools on treatment outcome in outpatient psychotherapy: study protocol. *BMC Psychiatry*, *17*(1), 306. https://doi.org/10.1186/s12888-017-1464-2

Mulder, R., Boden, J., Carter, J., Luty, S., & Joyce, P. (2017). Ten month outcome of cognitive behavioural therapy v. interpersonal psychotherapy in patients with major depression: a randomised trial of acute and maintenance psychotherapy. *Psychological Medicine*, *47*(14), 2540–2547. https://doi.org/10.1017/S0033291717001106

National Institute for Clinical Excellence. (2009). The guidelines manual. *London: NiCe*.

Østergård, O. K., Randa, H., & Hougaard, E. (2018). The effect of using the Partners for Change Outcome Management System as feedback tool in psychotherapy-A systematic review and meta-analysis. *Psychotherapy Research*, 1–18. https://doi.org/10.1080/10503307.2018.1517949

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. London: Sage.

Reese, R. J., Norsworthy, L. A., & Rowlands, S. R. (2009). Does a continuous feedback system improve psychotherapy outcome? *Psychotherapy*, *46*(4), 418. https://doi.org/10.1037/a0017901

Rothwell, P. M. (2005). External validity of randomised controlled trials:"to whom do the results of this trial apply?". *The Lancet*, *365*(9453), 82–93. https://doi.org/10.1016/S0140-6736(04)17670-8

Rubel, J. A., Zimmermann, D., Deisenhofer, A.-K., Müller, V., & Lutz, W. (2017). Nutzung von psychometrischem Feedback als empirische Unterstützung des Supervisionsprozesses bei Ausbildungstherapien. *Zeitschrift Für Klinische Psychologie Und Psychotherapie*. (46), 83–95. https://doi.org/10.1026/1616-3443/a000413

Sauer, E. M., Rice, K. G., Richardson, C. M. E., & Roberts, K. E. (2017). Influence of client attachment and gender on therapy transfers: A multilevel examination. *Training and Education in Professional Psychology*, *11*(1), 33–40. https://doi.org/10.1037/tep0000134

Saxon, D., & Barkham, M. (2012). Patterns of therapist variability: Therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology*, *80*(4), 535–546. https://doi.org/10.1037/a0028898

Saxon, D., Barkham, M., Foster, A., & Parry, G. (2017). The Contribution of Therapist Effects to Patient Dropout and Deterioration in the Psychological Therapies. *Clinical Psychology & Psychotherapy*, *24*(3), 575–588. https://doi.org/10.1002/cpp.2028

Schiefele, A.-K., Lutz, W., Barkham, M., Rubel, J., Böhnke, J., Delgadillo, J., . . . Lambert, M. J. (2017). Reliability of Therapist Effects in Practice-Based Psychotherapy Research: A Guide for the Planning of Future Studies. *Administration and Policy in Mental Health*, *44*(5), 598–613. https://doi.org/10.1007/s10488-016-0736-3

Shamdasani, S. (2005). 'Psychotherapy': the invention of a word. *History of the Human Sciences*, *18*(1), 1–22. https://doi.org/10.1177/0952695105051123

Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, *78*(3), 298–311. https://doi.org/10.1037/a0019247

Simon, W., Lambert, M. J., Harris, M. W., Busath, G., & Vazquez, A. (2012). Providing patient progress information and clinical support tools to therapists: Effects on patients at risk of treatment failure. *Psychotherapy Research*, *22*(6), 638–647. https://doi.org/10.1080/10503307.2012.698918

Storch, E. A., Lewin, A. B., Collier, A. B., Arnold, E., Nadai, A. S. de, Dane, B. F., . . . Murphy, T. K. (2015). A randomized controlled trial of cognitive-behavioral therapy versus treatment as usual for adolescents with autism spectrum disorders and comorbid anxiety. *Depression and Anxiety*, *32*(3), 174–181. https://doi.org/10.1002/da.22332

Wampold, B. E. (2001). *The great psychotherapy debate. Models, methods, and findings*. New Jersey: Lawrence Erlbaum Associates.

Xiao, H., Castonguay, L. G., Janis, R. A., Youn, S. J., Hayes, J. A., & Locke, B. D. (2017). Therapist effects on dropout from a college counseling center practice research network. *Journal of Counseling Psychology*, *64*(4), 424. https://doi.org/10.1037/cou000020

Xiao, H., Hayes, J. A., Castonguay, L. G., McAleavey, A. A., & Locke, B. D. (2017). Therapist effects and the impacts of therapy nonattendance. *Psychotherapy, 54*(1), 58–65. https://doi.org/10.1037/pst0000103

Zimmermann, D., Lutz, W., Reiser, M., Boyle, K., Schwartz, B., Schilling, V. N. L. S., . . . Rubel, J. A. (2019). What happens when the therapist leaves? The impact of therapy transfer on the therapeutic alliance and symptoms. *Clinical Psychology & Psychotherapy*, *26*(1), 135–145. https://doi.org/10.1002/cpp.2336

Zimmermann, D., Rubel, J., Page, A. C., & Lutz, W. (2017). Therapist effects on and predictors of non-consensual dropout in psychotherapy. *Clinical Psychology & Psychotherapy*, *24*(2), 312–321. https://doi.org/10.1002/cpp.2022

Eidesstattliche Erklärung

Ich versichere, dass ich meine Dissertation ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Trier, den

Nachname: Zimmermann Vorname: Dirk

Unterschrift: _____