



## Love yourself as a therapist, doubt yourself as an institution? Therapist and institution effects on outcome, treatment length, and dropout

Anne-Katharina Deisenhofer, Miriam I. Hehlmann, Julian A. Rubel, Wolfgang Lutz, Brian Schwartz, Anne-Kathrin Bräscher, Hanna Christiansen, Lydia Fehm, Julia A. Glombiewski, Jens Heider, Sylvia Helbig-Lang, Andrea Hermann, Jürgen Hoyer, Tina In-Albon, Tania Lincoln, Jürgen Margraf, Anne Katrin Risch, Henning Schöttke, Lars Schulze, Rudolf Stark, Tobias Teismann, Julia Velten, Ulrike Willutzki, Gabriele Wilz, Michael Witthöft & Patrizia Odyniec

To cite this article: Anne-Katharina Deisenhofer, Miriam I. Hehlmann, Julian A. Rubel, Wolfgang Lutz, Brian Schwartz, Anne-Kathrin Bräscher, Hanna Christiansen, Lydia Fehm, Julia A. Glombiewski, Jens Heider, Sylvia Helbig-Lang, Andrea Hermann, Jürgen Hoyer, Tina In-Albon, Tania Lincoln, Jürgen Margraf, Anne Katrin Risch, Henning Schöttke, Lars Schulze, Rudolf Stark, Tobias Teismann, Julia Velten, Ulrike Willutzki, Gabriele Wilz, Michael Witthöft & Patrizia Odyniec (03 Jun 2024): Love yourself as a therapist, doubt yourself as an institution? Therapist and institution effects on outcome, treatment length, and dropout, *Psychotherapy Research*, DOI: [10.1080/10503307.2024.2352749](https://doi.org/10.1080/10503307.2024.2352749)

To link to this article: <https://doi.org/10.1080/10503307.2024.2352749>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 03 Jun 2024.



[Submit your article to this journal](#)






















[View related articles](#)



[View Crossmark data](#)

RESEARCH ARTICLE

# Love yourself as a therapist, doubt yourself as an institution? Therapist and institution effects on outcome, treatment length, and dropout

ANNE-KATHARINA DEISENHOFER <sup>1\*</sup>, MIRIAM I. HEHLMANN <sup>2</sup>, JULIAN A. RUBEL <sup>2</sup>, WOLFGANG LUTZ <sup>1</sup>, BRIAN SCHWARTZ <sup>1</sup>, ANNE-KATHRIN BRÄSCHER <sup>3</sup>, HANNA CHRISTIANSEN <sup>4,5,6</sup>, LYDIA FEHM <sup>7</sup>, JULIA A. GLOMBIEWSKI<sup>8</sup>, JENS HEIDER<sup>8</sup>, SYLVIA HELBIG-LANG<sup>9</sup>, ANDREA HERMANN <sup>10</sup>, JÜRGEN HOYER <sup>11</sup>, TINA IN-ALBON <sup>12</sup>, TANIA LINCOLN <sup>13</sup>, JÜRGEN MARGRAF<sup>5,14</sup>, ANNE KATRIN RISCH<sup>15</sup>, HENNING SCHÖTTKE <sup>2</sup>, LARS SCHULZE <sup>16</sup>, RUDOLF STARK <sup>11</sup>, TOBIAS TEISMANN<sup>17</sup>, JULIA VELTEN <sup>18</sup>, ULRIKE WILLUTZKI <sup>19</sup>, GABRIELE WILZ<sup>15</sup>, MICHAEL WITTHÖFT <sup>4</sup>, & PATRIZIA ODYNIEC <sup>19,20</sup>

<sup>1</sup>Clinical Psychology and Psychotherapy, Department of Psychology, University of Trier, Trier, Germany; <sup>2</sup>Institute of Psychology, Clinical Psychology and Psychotherapy, University of Osnabrück, Osnabrück, Germany; <sup>3</sup>Department of Clinical Psychology, Psychotherapy, and Experimental Psychopathology, University of Mainz, Mainz, Germany; <sup>4</sup>Department of Clinical Child and Adolescent Psychology, Philipps University Marburg, Marburg, Germany; <sup>5</sup>Child and Adolescent Psychotherapy Clinic (KJ-PAM), Marburg, Germany; <sup>6</sup>German Center for Mental Health (DZPG), Bochum-Marburg, Germany; <sup>7</sup>Center for Psychotherapy at the Institute of Psychology (ZPHU), Humboldt-Universität zu Berlin, Berlin, Germany; <sup>8</sup>Clinical Psychology and Psychotherapy of Adulthood, University of Kaiserslautern-Landau, Kaiserslautern, Germany; <sup>9</sup>Psychotherapy Training Program (PTA Hamburg) at Clinical Psychology and Psychotherapy, Faculty of Psychology and Movement Science, Universität Hamburg, Hamburg, Germany; <sup>10</sup>Psychotherapy and Systems Neuroscience, Department of Psychology, Justus Liebig University of Giessen, Giessen, Germany; <sup>11</sup>Clinical Psychology and Psychotherapy, Technische Universität Dresden, Dresden, Germany; <sup>12</sup>Clinical Child and Adolescent Psychology and Psychotherapy, University of Kaiserslautern-Landau, Kaiserslautern, Germany; <sup>13</sup>Clinical Psychology and Psychotherapy, Faculty of Psychology and Movement Science, Universität Hamburg, Hamburg, Germany; <sup>14</sup>Mental Health Research and Treatment Center, Ruhr University Bochum, Bochum, Germany; <sup>15</sup>Department of Counseling and Clinical Intervention, Institute of Psychology, Friedrich Schiller University Jena, Jena, Germany; <sup>16</sup>Clinical Psychology and Psychotherapy, Freie Universität Berlin, Berlin, Germany; <sup>17</sup>Mental Health Treatment and Research Center, Ruhr-University Bochum, Bochum, Germany; <sup>18</sup>Faculty of Psychology, Mental Health Research and Treatment Center, Clinical Psychology and Psychotherapy, Ruhr University Bochum, Bochum, Germany; <sup>19</sup>Department of Psychology and Psychotherapy, Witten/Herdecke University, Witten, Germany & <sup>20</sup>Department of Psychology, University of Kassel, Kassel, Germany

(Received 4 July 2023; revised 24 April 2024; accepted 2 May 2024)

\*Both A.-K. Deisenhofer and M. I. Hehlmann are acknowledged as first authors of this paper. A. Deisenhofer primarily conducted the analyses and authored the method and results sections, while M. I. Hehlmann contributed methodological ideas and authored the introduction and discussion sections.

Correspondence concerning this article should be addressed to Anne-Katharina Deisenhofer, Clinical Psychology and Psychotherapy, Department of Psychology, University of Trier, D-54296 Trier, Germany. Email: deisenhofer@uni-trier.de; Miriam Hehlmann, Institute of Psychology, Clinical Psychology and Psychotherapy, University of Osnabrück, D-49076 Osnabrück, Germany. Email: miriam.hehlmann@uni-osnabrueck.de

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

## Abstract

**Objective:** Research suggests that some therapists achieve better outcomes than others. However, an overlooked area of study is how institution differences impact patient outcomes independent of therapist variance. This study aimed to examine the role of institution and therapist differences in adult outpatient psychotherapy.

**Method:** The study included 1428 patients who were treated by 196 therapists at 10 clinics. Two- and three-level hierarchical linear regression models were employed to investigate the effects of therapists and institutions on three dependent patient variables: (1) symptom change, (2) treatment duration, and (3) dropout. Level three explanatory variables were tested.

**Results:** The results showed that therapist effects (TE) were significant for all three types of treatment outcome (7.8%–18.2%). When a third level (institution) was added to the model, the differences between therapists decreased, and significant institution effects (IE) were found: 6.3% for symptom change, 10.6% for treatment duration, and 6.5% for dropout. The exploratory analyses found no predictors able to explain the systematic variation at the institution level.

**Discussion:** TE on psychotherapy outcomes remain a relevant factor but may have been overestimated in previous studies due to not properly distinguishing them from differences at the institution level.

**Keywords:** therapist effect; institution effect; treatment outcome

**Clinical or methodological significance of this article:** The study demonstrates that institution effects (IE) are a relevant contributor to symptom change, treatment duration, and dropout probability in psychological therapies. Neglecting the institution level can lead to an overestimation of the therapist effect (TE). Therefore, future research should consider the inclusion of institution-level variables in study designs to gain a better understanding of the TE and other context variables, such as social deprivation, neighborhood effects, and organizational factors that may shape psychotherapy outcomes.

## Introduction

The field of psychotherapy research has long been interested in understanding the factors that contribute to successful treatment outcomes (Barkham & Lambert, 2021; Cuijpers et al., 2019). While some studies have delved into patient characteristics predictive of better or worse outcomes, such as motivation, comorbidity, level of distress, context, race/culture, and personality structure (Constantino et al., 2021; Wampold & Owen, 2021), others have examined the therapist as a variable influencing therapeutic effectiveness. So far, a robust and growing body of research indicates that after controlling for patient differences, on average, some therapists have significantly better outcomes than others (Johns et al., 2019; Wampold & Owen, 2021). The therapist effect (TE) exists in clinical trials (range 3–15%, Wampold & Owen, 2021) as well as naturalistic settings (range 5–7%, Wampold & Owen, 2021). Moreover, therapist differences have been observed regarding, e.g., treatment outcome (Schiefele et al., 2017), treatment length (Lutz et al., 2015), dropout rates (Zimmermann et al., 2017), and sudden gains (Deisenhofer et al., 2022). All of these studies adjusted for initial patient severity, so these effects should not be viewed solely as a proxy for the severity of a patient's distress. These findings may indicate that some therapists find it more challenging than others to conclude therapy regardless of case severity or that some therapists need more time to initiate change in patients regardless of their symptom severity. TE on treatment length can be

interpreted as systematic differences in treatment duration due to the therapists' individual concepts of how much treatment is enough or therapists' ability to maintain treatment and prevent patients from dropping out. The identification of therapists who generally provide longer or shorter treatments is an important issue in the context of scarce financial resources in mental health-care settings.

While the crucial role of institutions may not always be immediately apparent, it is imperative to recognize their significance in the context of psychotherapy. Comparisons can be made to other professional fields where institutions' influence on job performance has been extensively researched (Heskett, 2011). In areas such as finance and corporate leadership, it is common to conduct studies investigating how organizational culture, management practices, and workplace policies affect employee productivity and job satisfaction (Pritchard & Karasick, 1973; Sati, 2021). Accordingly, studies have identified a positive connection between a supportive organizational culture and improved job performance, intrinsic motivation, and job satisfaction (Radakovich, 2016). To summarize, a nurturing organizational culture has the potential to boost corporate performance by up to 30% (Heskett, 2011). In this context, it is also reasonable to assume that institutions such as psychosocial facilities or mental health treatment centers influence the therapeutic process and patient outcomes.

Despite the obviously important position of institutions, their significance in the context of

psychotherapy is a relatively unexplored area. This is understandable since historically, psychotherapy research has primarily centered around patients, with therapists being a relatively recent addition to the focus of study. Similarly, the fact that psychotherapy always takes place in a particular context has often been neglected, as outlined in the General Model of Psychotherapy (Howard et al., 1986; Orlinsky & Howard, 1978). Additionally, studies were often limited to one location, making it impossible to examine institution effects (IE).<sup>1</sup> Fortunately, there has been a promising increase in multi-center studies, which are still in the early stages (Flückiger et al., 2021). This development allows researchers to more closely examine possible differences between institutions as well as therapists, which can be both seen as part of a principle called variance (Barkham et al., 2017). According to this principle, it is natural that institutions as well as therapists differ in various aspects. However, if these differences have an impact on patient outcomes, as has already been shown for therapist differences, they take on important significance and should be explored for quality assurance purposes.

Initial studies have shown that differences between institutions do indeed affect patient outcomes. For example, a study from England analyzed treatment outcomes of 33,243 patients who were treated at 103 different sites and received either CBT or counseling (Pybis et al., 2017). Multilevel analyses identified an IE of 1.8%, with therapy type not being a predictor of outcome. In a recently published study using data from a national practice research network ( $N=2977$ ), TE (1.9%–11.1%) and IE (4.6%–7.6%) were identified, with variations depending on the definition of the dependent variable dropout (Xiao et al., 2023). Another study identified a significantly larger IE (8.2%) than TE (3.2%) in a dataset of 26,888 patients treated by 462 therapists at 30 clinics (Firth et al., 2019). Explanatory variables were able to explain variance between institutions, but not therapists, leading to a significant reduction in the IE to 1.9%, with no significant change in the TE. The addition of patient-level symptom severity and employment status as well as the percentage of white patients per clinic explained most of the institutions' outcome and overall patient variance (Firth et al., 2019). In addition, a review concluded that there was some evidence indicating that institution-level variables like organizational climate and culture could explain differences in patient outcome between clinics (Falkenström et al., 2018). Some studies have already been able to demonstrate the impact of organizational climate and interorganizational coordination on the quality and outcomes of human service

systems, but research in the area of mental health is still lacking (e.g., Glisson & Hemmelgarn, 1998; Hemmelgarn et al., 2006).

Although initial studies on the topic of institution differences in the context of psychotherapy have already been published, the topic is still in its early stages. A 2018 review examined the impact of organizational factors on the outcome of mental health treatments and was able to include 19 studies (Falkenström et al., 2018). The authors estimated that the number of studies is small compared to the likely importance of organizational effects on mental health treatments. They discuss a number of reasons, including that the examination of this research question requires complete data from a large number of clinics. Furthermore, only 3% of the 30 studies considered facilities as a random factor and 10% reported an estimate of treatment-center interactions. Seven studies stated that IE were not significant, but provided no further details. Out of the seven studies that reported statistically significant differences in treatment outcomes (with a maximum Cohen's  $d$  value greater than .80), only one study provided a statistical measurement of the impact of the center where the study was conducted. In summary, there is a growing number of multi-center studies, but the practice of reporting IE and estimating clinic differences has yet to be broadly established.

This study contributes to the expanding field of IE by addressing current concerns. First, existing IE studies have predominantly centered on the healthcare systems in the United States and England, underscoring the need to replicate these findings in a different healthcare context, such as Germany. Secondly, to our knowledge, this study is the first to simultaneously investigate the IE in conjunction with the TE across three distinct outcome parameters within a single dataset. This approach offers valuable insights into the interplay of these effects, enhancing our understanding of their dynamics. Accordingly, the study has the potential to replicate existing research findings in a different context, while also offering novel and unique perspectives to the field. However, given the relatively unexplored nature of IE predictors and the limitation of utilizing only aggregated patient-level data, we opted for an exploratory approach to examine potential predictors. Similarly, an exploratory approach was warranted to investigate the association between the residuals of all three dependent variables.

The primary objective of this study was to investigate both therapist and institution differences in adult treatment outcomes. We specifically examined symptom change, treatment duration, and patient dropout rates. Furthermore, we aimed to uncover

potential factors contributing to the differences observed between institutions. Given the limited existing literature on potential predictors of these differences and the absence of institutional-level data within the available dataset, our explorative analysis focused on aggregated patient-level variables. These included variables such as symptom severity, employment status, previous treatment history, and comorbidity. Our rationale for this approach aligns with the methodology employed by Firth et al. (2023), who uses aggregated patient-level variables to provide insights into average experiences, for instance, average symptom severity among the entire population served by a specific institution. The last aim was to discuss implications of the findings for further research, therapy practice, therapist training, and organizational development in university outpatient clinics.

## Methods

### Procedure and Participants

The data analyzed in this study stem from a project on the coordination of data collection and analysis at research and training outpatient clinics for psychotherapy (KODAP; Velten et al., 2017), which is a nationwide German research collaboration across 48 university outpatient clinics, 34 of which specialize in adult patients and 14 in child and adolescent patients. Founded in 2013, its main goal is to coordinate the aggregation of longitudinal treatment data from all participating outpatient clinics. In Germany, university outpatient clinics have a

unique infrastructure for research, training, and clinical care of adults, children, and adolescents. They routinely collect a large amount of data on therapy processes and outcomes as well as on patient and therapist characteristics. These data are subject to high quality standards, such as the implementation of structured clinical interviews and regular group supervision, and are thus not only useful for the therapeutic process, but also for answering various research questions. For more detailed information on the origins, structure, and goals of KODAP, see Margraf et al. (2021).

### Inclusion/Exclusion Criteria

Patient, therapist, and institution criteria were applied when selecting patients to be included in the analyses. These criteria aimed to produce a sample that would provide adequate sample sizes of patients, therapists, and institutions to produce robust estimates of effects at each level (Schiefele et al., 2017). For all analysis datasets, this resulted in the minimum requirements of four patients per therapist and four therapists per institution.

Patient flow is depicted in Figure 1. First, the original dataset ( $N_{\text{patients}} = 8545$ ) was cleaned by deleting duplicate cases and removing any cases that did not have a therapist ID assigned or had a therapist transfer. This resulted in a basic dataset of 7552 adult patients. Patients were included in the analyses (Analysis Sample 1) using treatment duration and dropout as dependent variables (1) if they had completed more than one treatment session, (2) if they

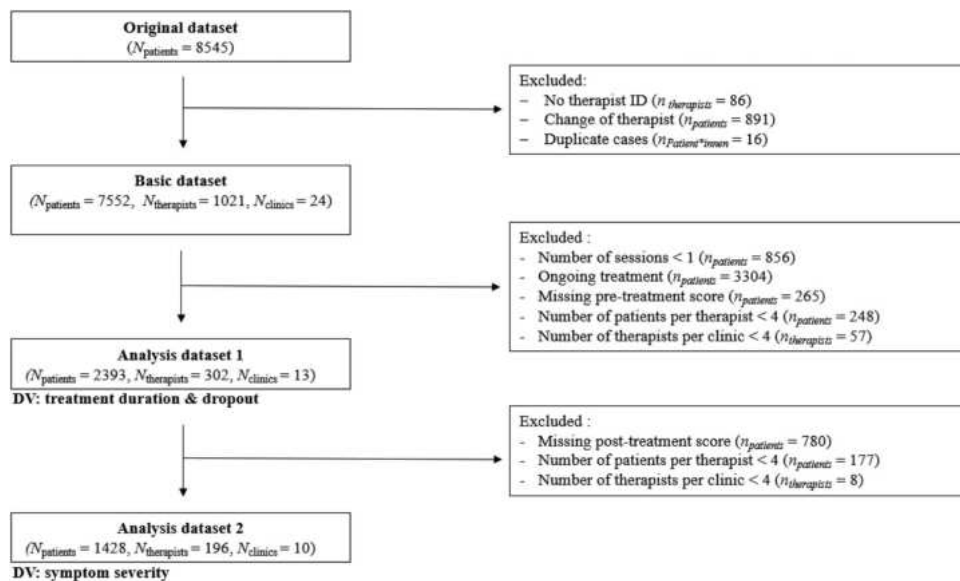


Figure 1. Flowchart showing the selection of patients for the analyses. Note: DV = dependent variable.

had ended treatment regularly<sup>2</sup> or dropped out of treatment prematurely and (3) if they had no missing values on the pre-treatment BSI to control for intake severity. These inclusion and exclusion criteria reduced the sample from 7552 patients to 2393 (Analysis Sample 1). To analyze the impact of institution and therapist differences on symptom severity, missing values for the post-treatment BSI were not allowed, further reducing the analysis sample to 1428 patients (Analysis Sample 2).

The sample size significantly decreased due to missing data, as illustrated in Figure 1. After removing the cases with missing values at post-treatment, only 59.67% of Analysis Sample 1 remained, resulting in a missing data percentage of 40.33%. It is important to note that due to the significant amount of missing data, we refrained from imputing values. This decision was influenced by the performance limitations of certain R packages, such as missForest (Stekhoven, 2013), which can effectively handle up to 30% missing values (Stekhoven & Bühlmann, 2012). Instead, we conducted a post hoc comparison between Analysis Samples 1 and 2 (see Table S1 in the supplemental material). The two datasets differed only regarding employment status, with more patients being incapable of working (29.6%) in Analysis Sample 1 compared to Analysis Sample 2 (26.1%). No significant differences were observed in the other variables.

### Patients

All psychological treatments were carried out between May 2018 and June 2021. In Analysis

Sample 1, patients had a mean age of 37.49 years ( $SD = 14.20$ ), were mostly female (64%), and the probability of premature treatment termination was 19% with a mean treatment duration of 33.41 ( $SD = 21.34$ ) sessions (see Table I). Thirty percent of patients were unable to work at the beginning of therapy. In addition, almost half of patients had already received previous treatment (47%). In most cases, diagnoses were based on a structured clinical interview (89.6%,  $n = 2144$ ). In some cases, a diagnostic checklist was used (7.7%,  $n = 184$ ) or only sections of a structured clinical interview were applied (0.6%,  $n = 15$ ). Five diagnostic assessments (0.2%) were based on therapists' clinical judgment. Affective disorders were the primary diagnosis for most patients (41.2%), followed by anxiety disorders (20.1%), and others (35.4%). Other diagnoses observed in the dataset included stress-related disorders (12.9%), somatoform disorders (5.1%), personality disorders (3.6%), eating disorders (3.5%), obsessive-compulsive disorder (3.1%), and other mental disorders (7.2%).<sup>3</sup> A small percentage of patients (3.3%) were not assigned a primary diagnosis in the dataset. Multiple diagnoses were given for 46.1% of patients ( $n = 1104$ ).

### Therapists

Patients were seen by a total of 302 therapists that all had a cognitive behavioral orientation, with 79% of therapists still in clinical training (see Table I). All therapists in training participated in a 3-year (full-time) or 5-year (part-time) postgraduate training program, as mandated by legal regulations. As part of the training program, trainees received a minimum of one year of training in an inpatient clinic before commencing patient care in outpatient clinics, which provided the data for this study. Furthermore, it is a mandatory requirement that every fourth session be supervised. Therapists in the dataset were mainly female (83%) and had a mean age of 30.40 years ( $SD = 5.40$ ). The mean number of patients per therapist was 10.12 and ranged between 4 and 35. The therapists included in the analyses worked at 13 different institutions, whereby the mean number of therapists per clinic was 36.39 (range 4 – 59).

### Measures

To assess symptom severity, the Brief Symptom Inventory (BSI; Franke, 2000; German translation of Derogatis et al., 1975), a self-report symptom inventory, was employed. Developed as a shorter version of the Symptom Checklist-90-R (SCL-90-

Table I. Overview of sample characteristics per level.

|         |                                 | <i>M (SD) or N (%)</i> |
|---------|---------------------------------|------------------------|
| Level 1 | Patients ( $N = 2393$ )         |                        |
|         | BSI pre <sup>1</sup>            | 1.04 (.63)             |
|         | BSI post <sup>1</sup>           | .63 (.57)              |
|         | BSI effect size <sup>1</sup>    | .68                    |
|         | Number of sessions <sup>2</sup> | 33.41 (21.34)          |
|         | Dropout                         | 443 (18.5)             |
|         | Age                             | 37.49 (14.20)          |
|         | Female                          | 1540 (64.4)            |
|         | Incapable of working            | 544 (29.6)             |
|         | Comorbidity                     | 1104 (46.1)            |
| Level 2 | Previous treatments             | 898 (47)               |
|         | Therapists ( $N = 302$ )        |                        |
|         | Age                             | 30.40 (5.40)           |
|         | Female                          | 1974 (82.5%)           |
| Level 3 | Patients per therapist          | 10.12 (4–35)           |
|         | Institutions ( $N = 13$ )       |                        |
|         | Therapists per clinic           | 36.39 (4–59)           |

Note. BSI = Brief Symptom Inventory.

<sup>1</sup>Values are based on analysis sample 2.

<sup>2</sup>Number of sessions after probation phase.

R; Derogatis, 1977), the BSI comprises 53 items and evaluates physical and psychological symptoms experienced by the individual in the past week. The instrument is composed of nine primary dimensions: *somatization*, *obsessive-compulsive*, *interpersonal sensitivity*, *depression*, *anxiety*, *hostility*, *phobic anxiety*, *paranoid ideation* and *psychoticism*. The items are scored on a 5-point Likert scale ranging from 0 (*not at all*) to 4 (*extremely*). The BSI's internal consistency has been found to be  $\alpha = .92$  and the retest-reliability to be  $r_{tt} = .90$  (Franke, 2000). In this study, the *Global Severity Index* (GSI) was calculated by averaging all items to measure patient outcome.

The dropout analysis in this study relied on institution case coding. This coding system assigns each case one of five codes: 0 for therapy in progress, 1 for regular termination of therapy, 2 for dropout, 3 for not reimbursed, and 4 for currently suspended. For the analyses, all cases coded as 2 were considered dropout cases. The dropout rate in the dataset (19%; see Figure 1) is consistent with the probability of premature treatment discontinuation identified in a recent meta-analysis (Swift & Greenberg, 2012).

### Explanatory Institution-level Variables

To investigate the impact of institution differences on treatment outcomes, potential predictors were examined via exploratory analysis. However, since the dataset lacked variables at the institution level, we derived clinic-level aggregates of patient-level variables from the original database of all patients ( $N = 7552$ ; basic dataset). This approach was selected to ensure that the aggregation process was not solely based on the analyzed samples and to best capture the actual composition of each institution's clinical population. The following patient-level variables were aggregated at the institution level: *symptom severity* at intake, assessed using the BSI; *employment status*<sup>4</sup>, dummy coded with 1 indicating incapacity to work and 0 indicating employable; *prior treatments*, dummy coded with 1 indicating previous treatments and 0 indicating no pre-treatments; and *comorbidity*, dummy coded with 1 indicating more than one diagnosis and 0 indicating only one diagnosis.

### Statistical Analyses

The purpose of this study was to investigate the impact of therapist and institution differences on three distinct types of treatment outcomes for adult psychotherapy outpatients: *symptom severity*, *treatment duration*, and *dropout*. To analyze the two continuous outcome variables, hierarchical linear regression models were employed, with patient *symptom severity*

and *treatment duration* (measured by the total number of sessions) serving as dependent variables in separate models. For the dichotomous variable *dropout*, a hierarchical logistic regression model was estimated accordingly. In all models, patients were situated at level 1, therapists at level 2, and institutions at level 3 (Formulas see supplemental material).

Each model started with a single level approach, followed by models with two and three levels. Two-level models were developed to examine therapist differences regarding patient outcome. Including the third level allowed the investigation of differences between institutions in terms of patient treatment outcomes and the extent to which these differences affected between-therapist variance. The statistical significance of each random effect (therapists and institutions) was determined by a chi-square difference test comparing the model with the respective random effect to the identical model without this random effect (Hox et al., 2010).

Multilevel models divide overall variability into individual components that are used to calculate the TE and IE. The three-level models divide total outcome variability into three components: variance between patients on level 1 ( $\sigma_e^2$ ), between therapists on level 2 ( $\sigma_t^2$ ), and between institutions on level 3 ( $\sigma_r^2$ ). To estimate the variance associated with therapist (TE) and institution differences (IE), the variance partition coefficient was calculated (VPC; e.g., Lewis et al., 2010). It divides the variance of the level of interest through the total variance, representing the proportion of variance in the dependent variable that can be attributed to the respective level of the model (Formulas see supplemental material). A higher VPC indicates larger differences between therapists or institutions with respect to the dependent variable being analyzed.

As initial patient severity notably influences outcomes (Okiishi et al., 2006) and patient risk is crucial to consider in TE models (Saxon & Barkham, 2012), all models were adjusted for pre-treatment BSI scores, centered around their grand mean, to address variations in patient severity scores before treatment initiation. This adjustment was particularly important, because the data came from a naturalistic context in which patients were not randomly assigned to therapists, but rather followed standard case assignment procedures at each clinic. It is important to note that the centering of the variable means that the reported TE and IE apply to patients with average initial severity.

Additionally, we explored whether there was a relationship between the TE and IE across the three different treatment outcomes (*symptom severity*, *treatment duration*, and *dropout*). To investigate this, we calculated the therapist and institution level

residuals for each model and examined their correlation. In this way, we could investigate whether, for example, therapists who achieve better or worse outcomes in comparison to the rest of the sample achieve them with a greater or fewer number of treatment sessions. It is also possible to examine whether therapists who have many dropouts in their caseload achieve better or worse therapy outcomes or whether their dropout rates are associated with treatment duration.

Finally, to explore potential explanations for institution differences, we tested the significance of aggregate institution-level variables including *symptom severity* at intake, *incapacity for work*, *previous treatments*, and *comorbidity*. To account for individual patient-level differences, we included corresponding patient-level variables in the models. In addition to assessing the *p*-values, we used the Akaike Information Criterion (AIC) to evaluate model fit including the possible predictors. Lower AIC values indicate a better fit (Hox et al., 2010).

All analyses were performed using the free software environment R version 4.2.2 (R Core Team, 2021). Analyses estimating hierarchical models were performed using the R package lme4 (Bates et al., 2015). The syntax used for the analyses conducted in this study is provided in the supplemental material.

## Results

### Descriptive Statistics

In Analysis Sample 2, the mean BSI score was 1.04 (*SD* = 0.63) at the beginning of treatment and 0.63 (*SD* = 0.57) after termination of therapy, which equals a pre–post effect size of Cohen’s *d* = 0.68

Table III. Therapist and institution effects for the three types of outcome.

|                    | 2-level model<br>TE | 3-level model |       |
|--------------------|---------------------|---------------|-------|
|                    |                     | TE            | IE    |
| Treatment duration | 18.2%               | 8.3%          | 10.6% |
| Dropout            | 17%                 | 10.5%         | 6.5%  |
| Symptom severity   | 7.8%                | .9%           | 6.3%  |

Note. BSI = Brief Symptom Inventory; All models are controlled initial impairment centered at the grand mean.  
\*\*\**p* = .001, \*\**p* < .01, \**p* < .05.

(95% *CI* [0.61, 0.76]) for the whole sample (see Table I).

For individual institutions, mean BSI scores ranged from 0.84 to 1.18 before treatment and from 0.47 to 0.91 after treatment (see Table II). Furthermore, the pre–post effect sizes of the samples were calculated and ranged between *d* = 0.27 and *d* = 1.05, which can be categorized as small to large effects (see Table II). Additionally, institutions differed regarding the percentage of patients incapable of working (range = 7.73% to 70.97%), having had previous treatments (range = 22.58% to 58.9%), having comorbid disorders (range = 13.81% to 48.22%), and average treatment length per clinic (range = 19.84–50.71 sessions, see Table II).

### Therapist and Institution Effects for the Three Types of Outcome

The two-level hierarchical linear regression models with symptom severity, treatment duration, and dropout as dependent variables yielded significant

Table II. Overview of the individual data sets of the institutions.

| Institutions                          | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    | 11    | 12    | 13    |
|---------------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Analysis dataset 1 ( <i>N</i> = 2393) |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <i>N</i> <sub>patients</sub>          | 31    | 31    | 36    | 36    | 45    | 73    | 181   | 205   | 284   | 314   | 351   | 371   | 435   |
| <i>N</i> <sub>therapists</sub>        | 4     | 5     | 5     | 5     | 7     | 13    | 30    | 27    | 37    | 40    | 25    | 45    | 59    |
| Patients per therapist ( <i>M</i> )   | 13.13 | 14.23 | 9.39  | 13.72 | 22.02 | 7.47  | 7.47  | 12.46 | 9.86  | 11.77 | 39.46 | 13.08 | 10.68 |
| Incapable of working (%)              | 70.97 | 45.16 | 33.33 | 38.89 | 40.00 | 19.18 | 7.73  | 25.37 | 24.30 | 28.66 | 35.33 | 24.53 | 30.35 |
| Previous treatments (%)               | 25.80 | 22.58 | 47.22 | 50.00 | 33.33 | 58.90 | 23.76 | 37.07 | 57.39 | 57.96 | 45.01 | 41.24 | 52.41 |
| Comorbidity (%)                       | 41.94 | 16.13 | 44.44 | 25.00 | 20.00 | 34.25 | 13.81 | 42.92 | 50.70 | 42.36 | 46.15 | 57.68 | 54.94 |
| Treatment duration ( <i>M</i> )       | 24.19 | 19.84 | 24.44 | 21.36 | 24.73 | 26.42 | 26.24 | 30.71 | 48.22 | 30.62 | 33.15 | 28.11 | 36.06 |
| Analysis dataset 2 ( <i>N</i> = 1428) |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <i>N</i> <sub>patients</sub>          | 25    | –     | 28    | –     | –     | 49    | 78    | 145   | 253   | 46    | 234   | 267   | 303   |
| <i>N</i> <sub>therapists</sub>        | 4     | –     | 4     | –     | –     | 9     | 14    | 23    | 33    | 8     | 23    | 36    | 43    |
| BSI pre ( <i>M</i> )                  | .84   | –     | .86   | –     | –     | .99   | .96   | .90   | 1.09  | 1.07  | 1.09  | 1.18  | .97   |
| BSI post ( <i>M</i> )                 | .57   | –     | .69   | –     | –     | .56   | .47   | .51   | .55   | .52   | .91   | .61   | .62   |
| Effect size ( <i>M</i> )              | .57   | –     | .30   | –     | –     | .70   | .84   | .72   | .93   | 1.05  | .27   | .91   | .63   |

Note. The table is based on analysis dataset 1 (*N* = 2393). BSI = Brief Symptom Inventory; Analysis dataset 2 (*N* = 1428) was used for the effect size calculation, so that corresponding pre and post values are also based on this sample; Due to missing values (67%–94%) in the BSI post values for institution 3, 4, and 5, these data sets were excluded from the analyses for symptom severity and effect sizes are not reported.



TE (see Table III), controlling for initial patient severity in all models. The TE for symptom severity was the smallest at 7.8%, followed by a TE of 17% for dropout and 18.2% for treatment duration. The inclusion of the therapist level resulted in a significantly better model fit in all models, as indicated by the chi-square difference test (e.g.,  $\chi^2_{\text{symptom severity}}(\text{df} = 1) = 7.69, p < .001^5$ ). In summary, the analyses revealed that a small but significant proportion of variance in patient outcome was due to therapist differences across all outcome variables.

Inclusion of the institution level in the analyses (third level) reduced the TE on symptom severity from 7.8% to 0.9%, a reduction of 88.5%. At the same time, a significant IE of 6.3% was found ( $\chi^2(\text{df} = 1) = 47.41, p < .001$ ). For dropout, the TE decreased from 17% to 10.5%, equaling a reduction of 38.2%. The associated IE was significant at 6.5% ( $\chi^2(\text{df} = 1) = 18.6, p < .001$ ). The TE on treatment duration decreased from 18.2% to 8.3%, a 54.3% reduction. The identified IE for treatment duration was again significant at 10.6%. For all dependent variables, the results showed that the inclusion of the third level substantially decreased variation at the therapist level.<sup>6</sup> Furthermore, examination of the differences between institutions revealed that third-level effects were present in all three models (see Table III).

### Associations Between Therapist and Institution Effects on the Three Dependent Variables

To determine whether there was a relationship between therapist differences in three types of

therapy outcome, the residuals of the three-level models were correlated. On level 2, the residuals of the dependent variables symptom severity and dropout were slightly positively correlated, but this correlation was not significant ( $r(194) = .05, p = .505$ ). Similarly, there was a slight positive correlation between the therapist-level residuals of the symptom severity and treatment duration models, but this correlation was not significant ( $r(194) = .11, p = .112$ ). The only significant correlation was found between the level 2 residuals of the treatment duration and dropout models, which were moderately negatively correlated ( $r(300) = -.43, p < .001$ ).

At the institution level, there was a negative correlation between the residuals of the three outcome models, but no significant relationship was found between them (symptom severity and dropout  $r(8) = -.56, p = .095$ ; symptom severity and treatment duration  $r(8) = -.23, p = .522$ ; treatment duration and dropout  $r(11) = -.42, p = .158$ ).

### Testing Possible Explanatory Institution-Level Variables

Although variables varied descriptively across institutions (see Table II), aggregate institution-level variables failed to explain institution-level variance in all models (see Table IV). Specifically, for the model with symptom severity as the dependent variable, addition of the institution-level variables did not reduce between-institution variance (IE) and resulted in worse model fit ( $\text{AIC}_{\text{with-predictors}} = 1858.516 > \text{AIC}_{\text{without-predictors}} = 1856.122$ ). In this

Table IV. Variables that account for the impact of institutions for the three different forms of outcome.

|   | Dropout     |          | Treatment duration |          | Treatment outcome (BSI) |          |
|---|-------------|----------|--------------------|----------|-------------------------|----------|
|   | Estimate    | <i>z</i> | Estimate           | <i>t</i> | Estimate                | <i>t</i> |
| Fixed effects:                          |             |          |                    |          |                         |          |
| Intercept                               | -1.63 (.17) | -9.32*** | 29.63 (2.17)       | 14.41*** | .60 (.04)               | 14.16*** |
| Level 1                                 |             |          |                    |          |                         |          |
| Initial impairment                      | .58 (.09)   | 6.74***  | 2.63 (.66)         | 3.96***  | .49 (.02)               | 24.19*** |
| Employment status                       | .26 (.12)   | 2.12*    | -5.22 (.93)        | -5.61*** | -                       | -        |
| Previous treatments                     | -           | -        | -2.90 (.86)        | -3.39*** | -                       | -        |
| Comorbidity                             | -           | -        | 4.91 (.86)         | 5.68***  | .07 (.03)               | 2.64**   |
| Level 3                                 |             |          |                    |          |                         |          |
| Initial impairment per institution      | -           | -        | -                  | -        | -                       | -        |
| Employment status per institution (%)   | -           | -        | -                  | -        | -                       | -        |
| Previous treatments per institution (%) | -           | -        | -                  | -        | -                       | -        |
| Comorbidity per institution (%)         | -           | -        | -                  | -        | -                       | -        |

Note. BSI = Brief Symptom Inventory; employment status is dummy coded with 0.5 representing incapacity to work and -0.5 employable; previous treatments is dummy coded with 0.5 representing previous treatments and -0.5 represents no pretreatments. Initial impairment centered at the grand mean.

\*\*\* $p = .001$  \*\* $p < .01$ , \* $p < .05$ .

model, after adjusting for initial impairment, the comorbidity variable was the sole significant predictor on the patient level. For the model with dropout as the dependent variable, the inclusion of institution-level variables reduced the IE from 6.5% to 3.97%, but the model fit indicated a deterioration ( $AIC_{\text{with-predictors}} = 2195.715 > AIC_{\text{without-predictors}} = 2194.558$ ). At the patient level, once again, one variable was significant, namely patient employment status. In the model with treatment duration as the dependent variable, adding the institution-level variables led to an improvement of model fit and a reduction of the IE from 10.6% to 9.2% ( $AIC_{\text{with-predictors}} = 21152.66 < AIC_{\text{without-predictors}} = 21173.37$ ) with all level one predictors included in the significant model.

### Discussion

The aim of the present study was to determine the impact of both TE and IE on three different psychotherapy outcomes (symptom severity, treatment length, and dropout) at German university outpatient clinics. The study's findings revealed that the TE was evident across all three forms of outcome, replicating previous research. Additionally, IE were detected for all three outcome measures. Importantly, the results showed that the TE for all three outcome measures was reduced, but persisted when the institution level was accounted for. Furthermore, there was no relationship between TE or IE on the three types of therapy outcomes, despite a significant negative correlation between treatment duration and dropout on the therapist level. Finally, none of the exploratory predictor variables (employment status, previous treatments, and comorbidity) aggregated at the institution level, were able to explain variance at the institution level in the multilevel models.

### Main Findings

**Therapist effects.** The results showed that therapists differed in terms of their patients' average symptom reduction (7.8%), treatment duration (18.2%), and dropout rates (17%), indicating variance in patient outcomes due to therapist differences. Incorporating the institutional level reduced therapist-related variance in symptom severity from 7.8% to 0.9%, an 88.5% reduction. Our study's TE on symptom severity falls within the range of previous findings, which vary between 3% (Baldwin & Imel, 2013) and 17.4% (Johns et al., 2019). In a comparable study by Firth et al. (2019), the inclusion of the institutional level also reduced the TE, but a

larger effect (3.4%) remained. The growing body of literature on TE has been heterogeneous, and insufficient sample sizes have been suggested as one possible cause, as discussed by Schiefele et al. (2017) and Wampold and Owen (2021). Our results support a further explanation of varying study outcomes, as different institutions also appear to have an impact on this effect.

Incorporating the institutional level, the TE on treatment length decreased from 18.2% to 8.3%, indicating a 54.3% reduction. According to Lutz et al. (2015), the only study we are aware of that examined this parameter, therapist differences accounted for 8.89% of the variance in treatment duration – half of our findings, but comparable to our results after including the institutional level. Ideally, the length of therapy should be based on achieving treatment goals and successfully addressing patients' concerns. However, the findings underscore that therapists may exhibit preferences for shorter or longer treatment durations independent of symptom severity. Further research is needed to better understand the factors that contribute to treatment duration variability.

For dropout, the TE reduced from 17% to 10.5%, resulting in a 38.2% decrease. This study found a greater TE on dropout compared to previous research; Zimmermann et al. (2017) noted a 5.7% variation due to therapists, while Saxon et al. (2017) identified a 12.6% variance. Varying dropout definitions and measurements across studies contribute to the issue (Xiao et al., 2023). Even participating clinics may have interpreted dropout differently due to unclear specifications in the present study. Nevertheless, the dropout rate of 19% in this dataset is consistent with the average probability of dropping out of treatment prematurely (Swift & Greenberg, 2012).

**Institution effects.** This study found that effect sizes at outpatient clinics ranged from small to large (Cohen's  $d = 0.27$ – $1.05$ ) in terms of their impact on patient symptomatology. In line with this, including the institution level in the multilevel analysis decreased the amount of variation explained by therapist differences, leading to the observation of significant IE. The reduction in therapist variance aligns with prior research (e.g., Firth et al., 2019; Xiao et al., 2023). Symptom severity variance attributed to institution differences was 6.3%, comparable to Firth et al. (2019)'s 8.2% in the UK. Other studies have found clinic effects of 2–5% and neighborhood effects of 1–2% (Firth et al., 2023), albeit in larger UK samples. Our findings, along with existing research, suggest that IE are present, despite standardized clinical training conditions in Germany, and ignoring these

institutional differences when analyzing TE on symptom severity can lead to biased results.

While the literature has completely overlooked treatment duration on higher levels, our study reveals comparable effects to other outcome variables: inclusion of the institution level reduces the TE on treatment duration, with a significant institutional effect of 10.6%. For dropout, 6.5% of patient-level variance was attributed to institutional differences, consistent with Xiao et al.'s (2023) findings. They observed TE and IE of 10% and 6.7% for attendance-based dropout, and 11.1% and 7.6% for therapist-rated dropout, respectively. These results emphasize the significant role of higher-level variables, such as TE and IE, in dropout variance.

In summary, our findings indicate that institutions in fact differ regarding average symptom reduction, treatment duration, and dropout of their patients. The results suggest that not considering institution differences can result in an overestimation of TE due to their potential entanglement with higher-level factors.

### **Correlation Between Therapist (TE) and Institution Effects (IE) on the Three Dependent Variables**

At the therapist level, the residuals of level 2 in the respective multilevel models were correlated to examine the relationships between symptom severity and both treatment duration and dropout, as well as between treatment duration and dropout. A similar approach was employed at the institution level, correlating the level 3 residuals of the respective models to investigate the associations between the IE of symptom severity, treatment duration, and dropout. The results indicated the absence of significant correlations between the residuals of the models that used symptom severity as the dependent variable, and both treatment duration and dropout. This applies to the residuals at both the therapist and institution levels. However, a moderate negative association ( $r = -.43$ ) was found between residuals on level 2 for treatment duration and dropout. This result suggests that therapists who provided shorter treatments on average also had more patients in their caseload who dropped out of therapy.

The lack of a correlation between treatment duration and symptom severity is in line with previous studies that also failed to establish an association between therapists' average effectiveness and treatment length (Lutz et al., 2015). It mirrors the assumption that treatment duration as such is not reflective of treatment quality. Ideally, treatment should end when the patient's problems have been sufficiently

addressed and resolved. Our findings contrast with studies suggesting that therapist effectiveness remains stable across different measures (e.g., Kraus et al., 2016; Nissen-Lie et al., 2017). While these studies propose that therapists who excel in one area tend to do so in others, our results indicate that this proficiency is not directly tied to treatment duration or dropout rates, warranting further investigation.

**Predictors of IE.** The present study investigated several aggregated institution-level variables (initial impairment, employment status, previous treatments, and comorbidity) as potential contributors to patient outcome variance, but was unable to explain the IE in any of the outcome variables. This means that institution-level variance, and thus differences between institutions, remain unexplained. Possible explanations may be related to underlying contextual and organizational factors, which were not fully captured in the present dataset. These will be briefly discussed for future directions.

Regarding *context factors*, it is conceivable that the institutions that provided data for the study are located in different neighborhoods (e.g., big city vs. rural area), which could explain some of the differences. Firth et al. (2019), for example, found that clinics with a higher proportion of minority patients had lower treatment outcomes on average, and identified differences between clinics that were partially linked to patients' baseline severity and employment status. These findings are supported by results of studies showing that patients from minority groups, such as adults with intellectual disabilities (Graser et al., 2022) and immigrants (Kobel et al., 2021), tend to have poorer treatment outcomes. It appears that patient-level variables, including minority affiliation and employment status, as well as neighborhood variables, are relevant factors when examining differences between institutions. Additional research has shown that patients residing in neighborhoods with high social deprivation tend to have poorer treatment outcomes (Clark, 2018; Finegan et al., 2020). However, neighborhood effects can largely be explained by social deprivation variables of the patient's context, such as income, employment status, education, and crime rate, which can also affect the patients themselves (Firth et al., 2023). In the present study, only employment status was available, thus recording whether individuals were capable of working at the start of therapy. The variable was significant at the patient level in the dropout and treatment duration models, with an inability to work at baseline associated with an increased likelihood of premature treatment termination and consequently shorter treatments. The

aggregation of this variable at the institution-level did not explain any further variance.

Previous research has overlooked the significance of *organizational structures* for psychotherapy outcomes (Falkenström et al., 2018). Unfortunately, our study couldn't explore factors like work climate, therapist burnout, or therapist training. Additionally, it is conceivable that institutions in our study may have varied policies on dropout management and treatment duration, influenced by factors like appointment reminders and payment structures. Understanding these organizational variations is essential to interpret institution-level findings. Future multi-center studies should focus on incorporating information on the organizational structures of individual institutions to gain further insight into their influence and to initiate corresponding measures for quality assurance.

### Limitations and Future Directions

When interpreting the study's results, it is important to consider several limitations. Firstly, the sample size of the dataset is relatively small compared to multi-center studies conducted in the UK (e.g., Firth et al., 2019; 2023), which may limit the reliability of the TE and IE. Additionally, the sample was reduced from Analysis dataset 1 to Analysis dataset 2 due to missing post-treatment BSI values. The post-hoc comparison of the two datasets shows no differences, but the potential for bias or unaccounted variation remains a limitation of our analyses. Moreover, although patient and therapist participation in completing research questionnaires was mandatory due to the structure of the outpatient centers, it is essential to acknowledge that the extent of engagement in data provision may have varied among individuals in both groups. This variation may have led to post-treatment measurements being predominantly provided by enthusiastic patients and therapists, potentially introducing a notable confounding factor when estimating institutional effects (Falkenström et al., 2018). Furthermore, the majority of the therapist sample comprises young females who are still in clinical training. Acknowledging this limitation is crucial, because utilizing non-diverse samples in research can lead to biased results that may unfairly impact minority groups. However, it should also be mentioned that the preponderance of young females in the context of clinical therapist training in Germany is unsurprising, and our sample accurately reflects this demographic reality. In addition, caution is advised when interpreting TE and IE on dropout, since the variance estimates cannot be solely

attributed to differences between therapists and institutions, and inconsistent dropout definitions between institutions may have contributed to this variance. Moreover, time represents a confounding variable for the dependent variable treatment duration, which was not accounted for in this study. Therefore, any associated effects should be interpreted with caution. Furthermore, the institutions in this study differ regarding their geographic locations and serve distinct patient populations with varying levels of stressors and resources. Unfortunately, the datasets did not provide sufficient information to illuminate the variance across institutions. Moreover, differences between therapists and institutions may be partly due to errors or inconsistencies in data coding or transmission (e.g., last value carried forward or therapies coded as completed although they were continued at another site), which may not reflect actual differences in psychotherapy outcomes. Additionally, the absence of a formal measure of sociodemographic and cultural factors at the institution level makes it impossible to draw firm conclusions about the association between structural and organizational factors, differences between institutions, and their impact on psychotherapy outcomes.

With regard to the predictor variables comorbidity and previous treatments, it should be mentioned that these are highly complex constructs that have been operationalized in different ways in previous research, depending on the research questions and patient populations (Constantino et al., 2021). Given the exploratory nature of our analyses of the predictor variables, the available data, and our statistical and conceptual approach, we believe that the binary summary of the variables was most appropriate for this manuscript. However, we cannot rule out the possibility that these constructs could still make a significant contribution to explaining institutional effects when operationalized differently.

Finally, it is essential to consider the possibility that the observed reduction of the TE when modeling the IE in our study may be an artifact, given that none of the included predictors effectively reduced unexplained variance. However, it is worth noting that the IE has been consistently identified in previous studies conducted across different health-care systems (e.g., Falkenström et al., 2018; Firth et al., 2019; Firth et al., 2023; Pybis et al., 2017), which suggests that it may indeed be a genuine phenomenon. Replicating our findings is a crucial step to validate the observed IE and understand its potential impact. Future research should focus on this aspect to confirm the robustness of our results.

The study's implications for future research suggest the need to include institutions as a random factor in analyses and report variance partition coefficients

(VPC) to enhance the reliability of estimates, while expanding our understanding of IE. Additionally, it is important to evaluate the potential influence of social deprivation, neighborhood effects, and organizational factors, such as climate and culture, and consider them in future studies in order to identify relevant factors that contribute to improved psychotherapeutic care. Accordingly, future studies should comprehensively assess the individual socioeconomic disadvantage of patients. Country-specific measures exist for this purpose, such as the indices of multiple deprivation (IMD; Ministry of Housing, Communities and Local Government, 2019) in England or the German index of socioeconomic deprivation (GISD; Michalski et al., 2022). These indices could be used to include the patients' area of residence and the facility's service area in the analyses to estimate the impact of these variables on psychotherapeutic care. Finally, collaboration between work and organizational psychologists and clinical psychologists could be beneficial to identify appropriate measurement instruments for the assessment of organizational factors in the psychotherapy clinic context.

### Conclusion

In conclusion, the following question could be asked: "Love yourself as a therapist, doubt yourself as an institution?" The results of this study suggest that facilities may need to be more critical of their impact on patient outcomes. Research has traditionally centered on therapists' skills and their influence on patient outcomes. However, our findings indicate that institutions may exert a greater influence on patient symptomatology, treatment duration, and dropout rates than previously assumed, underscoring the potential impact of institution-level factors on patient outcomes. Moreover, as therapist differences remain a relevant contributor to psychotherapy outcomes, the results highlight the importance of designing studies that consider the complex interplay between individual therapists and the institutions in which they work as well as sociodemographic and socioeconomic variables at the patient, therapist and institution levels. Consequently, it would be valuable to establish more research networks to be able to conduct multi-center studies on a regular basis. To gain a better understanding of the IE and how it arises, future research should focus on and plan for the inclusion of institution-level variables in study designs. In response to the question above, the present findings provide the following tentative conclusion: not only should therapists engage in self-critical reflection to improve therapy outcomes for their patients, but

institutions should follow suit and also be mindful of their impact.

### Acknowledgments

We thank Kaitlyn Poster for proofreading and Jay Lawrynowicz for his assistance with the tables. Furthermore, we would like to thank Nissen-Lie et al. (2017) for the article "Love yourself as a person, doubt yourself as a therapist," which served as an inspiration for our title.

### Disclosure Statement

No potential conflict of interest was reported by the author(s).

### Supplemental Data

Supplemental data for this article can be accessed online at <https://doi.org/10.1080/10503307.2024.2352749>.

### Notes

- <sup>1</sup> The term institution effect also includes related terms such as clinic or organizational effects, which have been examined in other studies.
- <sup>2</sup> Regular therapy termination is defined as a consensus between the patient and therapist to conclude therapy. In the German healthcare system, there are also session quotas that provide a framework for therapy duration. Within this framework, the length of therapy is individually tailored to the patient's needs. Therapy may therefore conclude either by mutual agreement of the patient and therapist or when the allocated number of sessions have been exhausted, as extensions are not always feasible.
- <sup>3</sup> for example sleep disorder, schizophrenia, etc.
- <sup>4</sup> Due to low frequencies, the following categories were merged into the category "incapable of working": "unable to work due to illness," "occupational disability pension," "retirement pension," and "other". Consolidation was done to ensure a more meaningful and reliable analysis with adequate representation in each group, leading to more robust conclusions.
- <sup>5</sup> Due to parsimony, not all results are presented, but can be obtained from the first author.
- <sup>6</sup> Table 2 shows that institutions 3 and 11 may be considered outliers based on their effect sizes of  $d = .30$  and  $d = .27$ . It could not be ruled out that this is due to errors in the data and that these outliers alone explained the effect found on the institution level. Therefore, the analyses were repeated without these institutions. Overall, the reanalysis led to a replication of the described findings.

### ORCID

Anne-Kathrin Bräscher  <http://orcid.org/0000-0001-5521-2535>

Miriam Hehlmann  <http://orcid.org/0000-0002-1069-0813>

Anne-kathrin Bräscher  <http://orcid.org/0000-0002-2621-5689>

Hanna Christiansen  <http://orcid.org/0000-0002-8104-0711>

Lydia Fehm  <http://orcid.org/0000-0002-0825-8466>

Andrea Hermann  <http://orcid.org/0000-0003-3186-7777>

Jürgen Hoyer  <http://orcid.org/0000-0002-1697-6732>

Tina In-albon  <http://orcid.org/0000-0002-2070-8458>

Tania Lincoln  <http://orcid.org/0000-0002-6674-2440>

Wolfgang Lutz  <http://orcid.org/0000-0002-5141-3847>

Julian A. Rubel  <http://orcid.org/0000-0002-9625-6611>

Henning Schöttke  <http://orcid.org/0000-0003-0049-0217>

Lars Schulze  <http://orcid.org/0000-0003-4720-2411>

Brian Schwartz  <http://orcid.org/0000-0003-4695-4953>

Rudolf Stark  <http://orcid.org/0000-0003-4299-5280>

Julia Velten  <http://orcid.org/0000-0002-6822-767X>

Ulrike Willutzki  <http://orcid.org/0000-0002-0149-4554>

Michael Witthöft  <http://orcid.org/0000-0002-4928-4222>

Patrizia Odyniec  <http://orcid.org/0000-0002-9444-7473>

## References

- 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). John Wiley & Sons.
- Barkham, M., & Lambert, M. J. (2021). The efficacy and effectiveness of psychological therapies. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (7th ed, pp. 135–190). Wiley.
- 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). American Psychological Association. <https://doi.org/10.1037/0000034-002>.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Clark, D. M. (2018). Realizing the mass public benefit of evidence-based psychological therapies: The IAPT program. *Annual Review of Clinical Psychology*, 14(1), 159–183. <https://doi.org/10.1146/annurev-clinpsy-050817-084833>
- Constantino, M. J., Boswell, J. F., Coyne, A. E., Swales, T. P., & Kraus, D. R. (2021). Effect of matching therapists to patients vs assignment as usual on adult psychotherapy outcomes: A randomized clinical trial. *JAMA Psychiatry*, <https://doi.org/10.1001/jamapsychiatry.2021.1221>
- Cuijpers, P., Reijnders, M., & Huibers, M. J. H. (2019). The role of common factors in psychotherapy outcomes. *Annual Review of Clinical Psychology*, 15(1), 207–231. <https://doi.org/10.1146/annurev-clinpsy-050718-095424>
- Deisenhofer, A.-K., Rubel, J. A., Bennemann, B., Aderka, I. M., & Lutz, W. (2022). Are some therapists better at facilitating and consolidating sudden gains than others? *Psychotherapy Research*, 32(3), 343–357. <https://doi.org/10.1080/10503307.2021.1921302>
- Derogatis, L. R. (1977). *SCL-90-R: Administration, scoring and procedures manual I*. Clinical Psychometric Research.
- Derogatis, L. R., Yevzeroff, H., & Wittelsberger, B. (1975). Social class, psychological disorder, and the nature of the psychopathologic indicator. *Journal of Consulting and Clinical Psychology*, 43(2), 183–191. <https://doi.org/10.1037/h0076514>
- Falkenström, F., Grant, J., & Holmqvist, R. (2018). Review of organizational effects on the outcome of mental health treatments. *Psychotherapy Research*, 28(1), 76–90. <https://doi.org/10.1080/10503307.2016.1158883>
- Finegan, M., Firth, N., & Delgadillo, J. (2020). Adverse impact of neighbourhood socioeconomic deprivation on psychological treatment outcomes: The role of area-level income and crime. *Psychotherapy Research*, 30(4), 546–554. <https://doi.org/10.1080/10503307.2019.1649500>
- Firth, N., Barkham, M., Delgadillo, J., Bell, A., & O’Cathain, A. (2023). The role of socioeconomic deprivation in explaining neighborhood and clinic effects in the effectiveness of psychological interventions. *Journal of Consulting and Clinical Psychology*, 91(2), 82–94. <https://doi.org/10.1037/ccp0000784>
- Firth, N., Saxon, D., Stiles, W. B., & Barkham, M. (2019). Therapist and clinic effects in psychotherapy: A three-level model of outcome variability. *Journal of Consulting and Clinical Psychology*, 87(4), 345–356. <https://doi.org/10.1037/ccp0000388>
- Flückiger, C., Paul, J., Hilpert, P., Vislä, A., Gómez Penedo, J., Probst, G. H., & Wampold, B. E. (2021). Estimating the reproducibility of psychotherapy effects in mood and anxiety disorders: The possible utility of multicenter trials. *World Psychiatry*, 20(3), 445–446. <https://doi.org/10.1002/wps.20901>
- Franke, G. (2000). *BSI. Brief Symptom Inventory—Deutsche Version* [BSI. Brief Symptom Inventory: German version]. Beltz.
- Glisson, C., & Hemmelgarn, A. (1998). The effects of organizational climate and interorganizational coordination on the quality and outcomes of children’s service systems. *Child Abuse & Neglect*, 22(5), 401–421. [https://doi.org/10.1016/S0145-2134\(98\)00005-2](https://doi.org/10.1016/S0145-2134(98)00005-2)
- Graser, J., Göken, J., Lyons, N., Ostermann, T., & Michalak, J. (2022). Cognitive-behavioral therapy for adults with intellectual disabilities: A meta-analysis. *Clinical Psychology: Science and Practice*, 29(3), 227–242. <https://doi.org/10.1037/cps0000077>
- Hemmelgarn, A. L., Glisson, C., & James, L. R. (2006). Organizational culture and climate: Implications for services and interventions research. *Clinical Psychology: Science and Practice*, 13(1), 73–89. <https://doi.org/10.1111/j.1468-2850.2006.00008.x>
- Heskett, J. (2011). *The culture cycle: Shaping the unseen force that will transform performance*. FT Press.

- Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist*, 41(2), 159–164. <https://doi.org/10.1037/0003-066X.41.2.159>
- Hox, J. J., Maas, C. J. M., & Brinkhuis, M. J. S. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Statistica Neerlandica*, 64(2), 157–170. <https://doi.org/10.1111/j.1467-9574.2009.00445.x>
- Johns, R. G., Barkham, M., Kellett, S., & Saxon, D. (2019). A systematic review of therapist effects: A critical narrative update and refinement to review. *Clinical Psychology Review*, 67, 78–93. <https://doi.org/10.1016/j.cpr.2018.08.004>
- Kobel, F., Erim, Y., & Morawa, E. (2021). Predictors for successful psychotherapy: Does migration status matter? *PLoS ONE*, 16(9), e0257387. <https://doi.org/10.1371/journal.pone.0257387>
- Kraus, D. R., Bentley, J. H., Alexander, P. C., Boswell, J. F., Constantino, M. J., Baxter, E. E., & Castonguay, L. G. (2016). Predicting therapist effectiveness from their own practice-based evidence. *Journal of Consulting and Clinical Psychology*, 84(6), 473–483. <https://doi.org/10.1037/ccp0000083>
- Lewis, M., Morley, S., Van Der Windt, D. A., 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(10), 1033–1039.
- 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>
- Margraf, J., Hoyer, J., Fydrich, T., In-Albon, T., Lincoln, T., Lutz, W., Schlarb, A., Schöttke, H., Willutzki, U., & Velten, J. (2021). The cooperative revolution reaches clinical psychology and psychotherapy: An example from Germany. *Clinical Psychology in Europe*, 3(1), e4459. <https://doi.org/10.32872/cpe.4459>
- Michalski, N., Reis, M., Tetzlaff, F., Herber, M., Kröll, L. E., Hövener, C., Nowossadeck, E., & Hoebel, J. (2022). *E-German Index of Socioeconomic Deprivation (GISD): Revision, update and application examples*. <https://doi.org/10.25646/10641>.
- Ministry of Housing, Communities and Local Government. The English Indices of Deprivation 2019. Statistical Release. (2019). Retrieved April 30, 2020, from [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/835115/IoD2019\\_Statistical\\_Release.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/835115/IoD2019_Statistical_Release.pdf)
- Nissen-Lie, H. A., Rønnestad, M. H., Høglend, P. A., Havik, O. E., Solbakken, O. A., Stiles, T. C., & Monsen, J. T. (2017). Love yourself as a person, doubt yourself as a therapist? *Clinical Psychology & Psychotherapy*, 24(1), 48–60. <https://doi.org/10.1002/cpp.1977>
- Okishi, J. C., Lambert, M. J., Eggett, D., Nielsen, L., Dayton, D. D., & Vermeersch, D. A. (2006). An analysis of therapist treatment effects: Toward providing feedback to individual therapists on their clients' psychotherapy outcome. *Journal of Clinical Psychology*, 62(9), 1157–1172. <https://doi.org/10.1002/jclp.20272>
- Orlinsky, D. E., & Howard, K. I. (1978). The relation of process to outcome in psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (2nd ed., pp. 283–330). Wiley.
- Pritchard, R. D., & Karasick, B. W. (1973). The effects of organizational climate on managerial job performance and job satisfaction. *Organizational Behavior and Human Performance*, 9(1), 126–146. [https://doi.org/10.1016/0030-5073\(73\)90042-1](https://doi.org/10.1016/0030-5073(73)90042-1)
- Pybis, J., Saxon, D., Hill, A., & Barkham, M. (2017). The comparative effectiveness and efficiency of cognitive behaviour therapy and generic counselling in the treatment of depression: Evidence from the 2nd UK National Audit of psychological therapies. *BMC Psychiatry*, 17(1), 215. <https://doi.org/10.1186/s12888-017-1370-7>
- Radakovich, P. S. (2016). The relationship between organizational culture, intrinsic motivation, and employee performance: A systematic review And meta-analysis. Wayne State University Dissertations. 1659. [https://digitalcommons.wayne.edu/oa\\_dissertations/1659](https://digitalcommons.wayne.edu/oa_dissertations/1659)
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. URL <https://www.R-project.org/>
- Sati, K. (2021). The role of factors in organizational behaviour. *International Journal of Educational Research and Studies*, 4(1), 1–456.
- 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., Delgado, J., Kopta, M., Schulte, D., Saxon, D., Nielsen, S. L., & 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 and Mental Health Services Research*, 44(5), 598–613. <https://doi.org/10.1007/s10488-016-0736-3>
- Stekhoven, D. J., & Bühlmann, P. (2012). MissForest-non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28(1), 112–118.
- Stekhoven, D. J., & Stekhoven, M. D. J. (2013). *Package 'missForest'*. R package version, 1, 21.
- Swift, J. K., & Greenberg, R. P. (2012). Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 80(4), 547–559. <https://doi.org/10.1037/a0028226>
- Velten, J., Hoyer, J., Margraf, J., Benecke, C., Berking, M., In-Albon, T., Lincoln, T., Lutz, W., Schlarb, A., Schöttke, H., Willutzki, U., & Hoyer, J. (2017). Methodenpapier zur Koordination der Datenerhebung und -auswertung an Hochschul- und Ausbildungsambulanzen für Psychotherapie (KODAP) [Method paper of a project for coordinating research at the outpatient university clinics for psychotherapy in Germany]. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 46(3), 169–175. <https://doi.org/10.1026/1616-3443/a000431>
- Wampold, B., & Owen, J. (2021). Therapist effects: History, methods, magnitude, and characteristics of effective therapists. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (7th ed., pp. 297–326). Wiley.
- Xiao, H., Castonguay, L. G., Hayes, J. A., Janis, R. A., & Locke, B. D. (2023). Reconstructing dropout: Building from multiple definitions, therapist effects, and center effects. *Psychotherapy Research*, 33(2), 146–157. <https://doi.org/10.1080/10503307.2022.2082897>
- 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>