



## Psychotherapist trainees' professional self-doubt and negative personal reaction: Changes during cognitive behavioral therapy and association with patient progress

Patrizia Odyniec, Thomas Probst, Jürgen Margraf & Ulrike Willutzki

To cite this article: Patrizia Odyniec, Thomas Probst, Jürgen Margraf & Ulrike Willutzki (2019) Psychotherapist trainees' professional self-doubt and negative personal reaction: Changes during cognitive behavioral therapy and association with patient progress, *Psychotherapy Research*, 29:1, 123-138, DOI: [10.1080/10503307.2017.1315464](https://doi.org/10.1080/10503307.2017.1315464)

To link to this article: <https://doi.org/10.1080/10503307.2017.1315464>



Published online: 26 Apr 2017.



[Submit your article to this journal](#)



Article views: 248



[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)

EMPIRICAL PAPER

## Psychotherapist trainees' professional self-doubt and negative personal reaction: Changes during cognitive behavioral therapy and association with patient progress

PATRIZIA ODYNIC<sup>1</sup>, THOMAS PROBST<sup>2</sup>, JÜRGEN MARGRAF<sup>3</sup>, & ULRIKE WILLUTZKI<sup>1</sup>

<sup>1</sup>*School of Psychology, University Witten/Herdecke, Witten, Germany;* <sup>2</sup>*Georg-Elias-Müller Institute for Psychology, Georg-August-University Göttingen, Göttingen, Germany* & <sup>3</sup>*Department of Psychology, Ruhr University Bochum, Bochum, Germany*

(Received 20 October 2016; revised 27 March 2017; accepted 30 March 2017)

### Abstract

**Objective:** This study examined psychotherapist trainees' experiences of "professional self-doubt" (PSD) and "negative personal reaction" (NPR) during cognitive behavioral therapy (CBT) and their associations with patients' symptoms and interpersonal problems. **Method:** Forty therapists treating 621 patients were analyzed. Patients' symptoms and interpersonal problems were collected repeatedly during therapy. Data about patients' interpersonal problems were available only for 106 patients and 18 therapists. Therapists' difficulties were assessed as trait-based (one assessment across all patients) and as state-based (repeated assessments for each individual patient) difficulties. Multilevel models were performed. **Results:** None of the trait-based difficulties correlated with the change of the patients' symptoms. Yet, more NPR at the trait-level predicted a more favorable change, whereas higher PSD at the trait-level showed an opposite effect on change of patients' interpersonal problems. Regarding state-based difficulties, PSD as well as NPR decreased significantly over the course of CBT. Patients whose therapists experienced PSD to increase during CBT were at risk of a less favorable patient progress regarding symptoms, whereas the change of interpersonal problems was not significantly associated with changes in therapists' difficulties. **Conclusion:** Patients' progress is associated with therapists' experiences of difficulties. Yet, trait- and state-based difficulties lead to different results.

**Keywords:** therapist effects; professional self-doubt; negative personal reaction; outcome; therapists' experience of difficulties

**Clinical or methodological significance of this article:** Associations between therapists' difficulties and patient-reported outcomes depended on whether therapists' difficulties were assessed once across all patients (trait-level) or for each individual patient repeatedly during CBT (state-level). Contrary to previous research on trait-level difficulties, the difficulty professional self-doubt (PSD) was associated with a less favorable course of patients' interpersonal problems, whereas the difficulty negative personal reaction (NPR) was associated with a more favorable progress. Moreover a patient-therapist-time contextualization (state-level) seems relevant for the assessment of complex therapist variables and an elaborated understanding of the therapist effect. Altogether, CBT trainees should not be scared of their negative personal reactions towards their patients and professional self-doubt of CBT trainees should be monitored and discussed (e. g., in supervision).

### Introduction

The person of the therapist has long been neglected as a factor contributing to psychotherapy effects (Wampold, Hollon, & Hill, 2011). However, a significant proportion of the outcome variance can be

attributed to the therapist (5–8%; Baldwin & Imel, 2013) and the therapist effect reaches small ( $d = .35$ ; in controlled studies) to medium ( $d = .55$ ; in naturalistic settings) effect sizes (Wampold, 2015). Therefore, it appears to be important to further

---

Correspondence concerning this article should be addressed to Patrizia Odyniec, Clinical Psychology and Psychotherapy, School of Psychology, University of Witten/Herdecke, D-58448 Witten, Germany. Email: [patrizia.odynec@uni-witten.de](mailto:patrizia.odynec@uni-witten.de)

investigate the characteristics of therapists being associated with psychotherapy process and outcome. According to an earlier review by Beutler et al. (2004) neither professional qualifications nor personal characteristics have been found to consistently predict psychotherapy process and outcome. Recent studies in this area therefore focused on more complex and interactive variables, like therapists' abilities to deal with challenging patient–therapist interactions (Anderson, Ogles, Patterson, Lambert, & Vermeersch, 2009), therapists' humor style and honesty (Yonatan-Leus, Tishby, Shefler, & Wiseman, 2017), therapists' attachment style (Degnan, Seymour-Hyde, Harris, & Berry, 2016), and on therapists' subjective experience of therapeutic work (Heinonen, Lindfors, Laaksonen, & Knekt, 2012; Schröder, Wiseman, & Orlinsky, 2009). In this context, therapists' experience of difficulties in therapeutic practice has also become an issue of increased interest (Odyniec, Victor, Berner, & Willutzki, 2016). It has been explored with regard to psychotherapists' self-perceived professional development (Black, Hardy, Turpin, & Parry, 2005; Sanness, 2012; Schröder et al., 2009) as well as their relationship to patients' perspective on psychotherapy process and outcome (e.g., Heinonen, Knekt, Jääskeläinen, & Lindfors, 2014; Nissen-Lie, Monsen, & Rønnestad, 2010; Nissen-Lie, Monsen, Ulleberg, & Rønnestad, 2013; Nissen-Lie, Havik, Høglend, Rønnestad, & Monsen, 2015). Later studies linking therapists' perspective to therapy process and outcome focused on two factors which represent challenging emotions therapists can experience during therapeutic practice: Therapists' professional self-doubt (PSD) about their efficacy as a therapist and negative personal reactions (NPR), such as reduced empathy and negative feelings towards patients. These studies revealed partly counterintuitive results: Nissen-Lie et al. (2010, 2013) reported in a naturalistic study with rather experienced psychodynamic-oriented therapists that higher PSD was associated with a better early patient-rated alliance and more improvement in patients' interpersonal problems; opposite effects were found for NPR. Whereas the detrimental impact of an NPR corresponds to other research on the relevance of empathy for psychotherapy process and outcome (Elliott, Bohart, Watson, & Greenberg, 2011), the positive association of PSD with therapy progress called for further explanation: Nissen-Lie et al. (2010, 2013) interpreted increased PSD as reflecting a more humble attitude and a willingness to critically evaluate one's work as a therapist, which may lead to a more flexible manner of dealing with obstacles in psychotherapy. These results suggest that it may be expedient in

psychotherapy training and supervision as well as in professional practice to foster an “error and doubt friendly” ambiance in order to facilitate therapists' self-reflection and improve therapy outcome. While such a conclusion has already been advocated in qualitative studies of professional development (Rønnestad & Skovholt, 2012; Skovholt & Jennings, 2004), replication studies are needed in order to explore whether the pattern Nissen-Lie et al. (2010, 2013) reported can be found in psychotherapists at different stages of professional development and with different theoretical orientations. Until now only therapists with an average professional experience of 10 years and a prevailing psychodynamic/psychoanalytic orientation have been analyzed in the context of such research questions (Nissen-Lie et al., 2010, 2013).

A follow-up study (Nissen-Lie, Rønnestad, et al., 2015) further explored the findings and took therapists' attitude towards themselves into account. The authors concluded that a protective and affectionate stance towards oneself in combination with a self-critical evaluation of one's own work seems to prepare the ground for therapists' higher “appropriate responsiveness.” “Appropriate responsiveness” (Stiles, 2013) is understood as a therapist's skill to constantly adjust to the patient's needs, states, and interpersonal style to optimize therapy outcome. It is not a specific therapeutic intervention, but an interplay of a therapist's interpersonal skills, techniques, and empathy (Hatcher, 2015). The idea of responsiveness as a therapist factor was also discussed in a recent study by Nissen-Lie et al. (2016), who found evidence that therapists are uniformly effective across patient outcome domains.

Assessing responsiveness is challenging because it varies by definition across patients and time. PSD and NPR may be interesting indicators of “appropriate responsiveness,” but earlier studies (Nissen-Lie et al., 2010, 2013; Nissen-Lie, Rønnestad, et al., 2015) did not take variability and adaptation of these therapists' difficulties into account as therapists rated their PSD/NPR only once for all their patients, assessing these factors as therapists' trait-based variables. This approach decontextualizes psychotherapy as it ignores dyadic effects specific to a particular therapist–patient combination and also the variation of therapists' experience with a specific patient over time. However, such a contextualization seems essential in order to more adequately account for therapists' responsiveness in research; at the same time it may further clarify the meaning and influence of therapists' difficulties (PSD and NPR).

To follow up and expand this line of research, cognitive behavioral therapy (CBT) trainees' (therapists') experience of difficulties in their work with

their patients were assessed in two different ways in the current study: In order to explore the replicability of previous research we looked at psychotherapists' difficulties in practice as *trait-based difficulties* by asking therapists once how they experience difficulties with their patients in general; expanding on this work we focused on therapists' difficulties as a *state-based variable* by assessing PSD/NPR repeatedly over the course of therapy for every patient of the therapist in question; finally we evaluated how the trait- and state-based assessments of the therapists' experience of difficulties are associated with the course of patient-reported outcomes (Greenhalgh, Long, & Flynn, 2005) during psychotherapy.

### Research Questions and Hypotheses

The following research questions were addressed:

- (1) Can the results of Nissen-Lie et al. (2010, 2013) and Nissen-Lie, Rønnestad, et al. (2015) be replicated in a different sample of therapists and patients, showing that PSD/NPR are associated with patient progress when PSD/NPR are assessed across patients and time (trait-based therapist difficulties; PSD\_T, NPR\_T)? Referring to the results of Nissen-Lie and colleagues, we hypothesized that PSD\_T would have a beneficial effect on patient progress whereas NPR\_T would be negatively associated with patient progress.
- (2) Do PSD and NPR vary between patients and change during the process of psychotherapy when PSD/NPR are assessed repeatedly with regard to a specific patient (state-based therapist difficulties; PSD\_S, NPR\_S)? We hypothesized that PSD\_S and NPR\_S vary between patients and change during the process of psychotherapy when PSD/NPR are assessed repeatedly with regard to a specific patient.
- (3) Are changes in PSD/NPR associated with patient progress when PSD and NPR are assessed repeatedly with regard to a specific patient (state-based therapist difficulties; PSD\_S, NPR\_S)? Again, we expected that PSD\_S would have a beneficial association with patient progress whereas NPR\_S would be negatively associated with the course of patient-reported outcomes.

### Method

#### Procedure

The procedure differed between patients and therapists.

**Patients:** The Global Severity Index (GSI) of the Brief Symptom Inventory (BSI; Derogatis, 1975; German translation: Franke, 2000) and interpersonal problems measured with the global scale of the Inventory of Interpersonal Problems (IIP; Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988; German translation: Horowitz, Alden, Kordy, & Strauß, 2000) were evaluated at the diagnostic assessment (2nd session), at the 5th session, the 10th session and from thereon after every 10th session and at the end of therapy. Informed written consent was obtained from all participating patients.

**Therapists:** Therapists' trait-based difficulties were recorded once when therapists were enrolled in the study. All therapists completed the Development of Psychotherapists Common Core Questionnaire (DPCCQ; Orlinsky et al., 1999) containing the Therapist Difficulties Scales (TDS; Orlinsky & Rønnestad, 2005). Therapist recruitment started in August 2012 for the first CBT cohort (2011) and in August 2013 for the second CBT cohort (2012). Trait-based difficulties were assessed after 12 months of clinical internship followed by at least 6 months of therapeutic work at the outpatient center (the therapists had the possibility to answer the DPCCQ including the trait-based difficulties within 1 year and were repeatedly reminded during this period). There was no systematic time relationship between the assessment of the trait-based difficulties and the assessment of the outcome measures or beginning/ending of the analyzed therapies. Therapists' state-based difficulties were assessed with the TDS-Individual Patient (TDS-IP; Willutzki, Hernandez Bark, Davis, & Orlinsky, 1997) questionnaire first after the 5th therapy session and from thereon at the same measurement points as the outcome measures. All therapists gave their informed written consent to participate in the study.

The measures are described in more detail below. Patients were allocated to therapists based on the patients' registration for a therapy and therapists' capacity. The assignment did not follow any particular selection criteria. The first available therapist took over the next patient from the waiting list. For a better understanding of the recruitment process, please see [Figure 1](#).

#### Participants

**Patients.** The total sample included 1049 patients seen by 40 therapists at the outpatient psychotherapy center of Ruhr University Bochum (Germany) between 2007 and 2014. Patients who did not meet the criteria for a mental disorder and those with serious substance abuse or acute crises (suicidality,

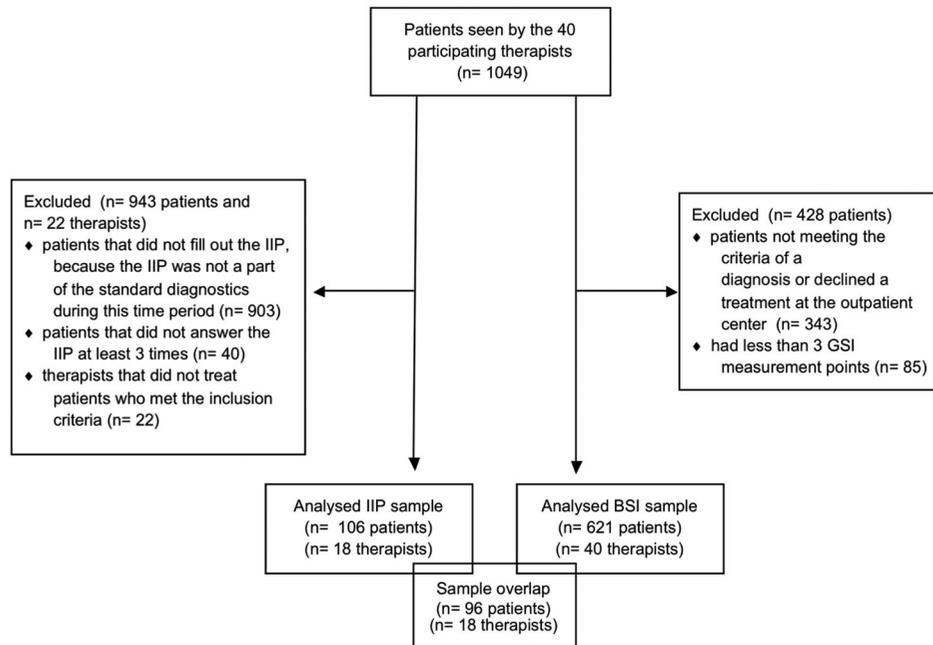


Figure 1. Flow diagram of the recruitment process.

psychosis or manic episodes requiring inpatient treatment) were excluded from treatment. Also, patients disagreeing with the general conditions of the outpatient center (treatment by psychotherapists in

training under supervision, video recording of every therapy session, participation in the standard diagnostic for research purposes) were excluded. Moreover, patients not completing the outcome

Table I. Sample description for the patients.

	BSI sample		IIP sample		Statistics
<i>N</i>	621		106		
Age					
<i>M (SD)</i>	37.93 (12.8)		37.08 (11.8)		$t(725) = 1.65$ $p = .53$
Gender					
Female	62%		64%		$\chi^2(1) = .13$ $p = .72$
Male	38%		36%		
Diagnosis					
Single	30.4%		33.3%		$\chi^2(1) = .99$ $p = .32$
Comorbidity	69.6%		66.7%		
Anxiety disorders	58.2%		74.5%		
Affective disorders	55.4%		71.7%		
Eating disorders	7.9%		12.2%		
Substance abuse disorders	5.6%		6.6%		
Somatic symptom disorders	5%		4.7%		
Personality disorders	3.3%		4.7%		
Schizophrenia disorders	1.3%		.9%		
Impulse control disorders	.6%		-		
Number of sessions					
<i>M (SD)</i>	35.18 (16.84)		42.22 (21.26)		$t(724) = -3.24$ $p = .002$
Pretreatment scores					
<i>M (SD)</i>	GSI 1.16 (.69)	IIP 1.40 (.65)	GSI 1.17 (.70)	IIP 1.64 (.50)	GSI $t(725) = -.25$ $p = .80$ IIP $t(233) = -3.40$ $p = .001$
Range	.00–3.38 73% $\geq .64$	.03–3.03	.00–3.11 76% $\geq .64$	.45–2.88	

Table II. Therapists' difficulties scales, assessed as trait-based (\_T) and state-based (\_S) variables in therapy for the BSI and IIP sample (descriptives and correlations of the scales).

	BSI sample		<i>r</i>	IIP sample		<i>r</i>	Statistics
	Mean ( <i>SD</i> )	Range		Mean ( <i>SD</i> )	Range		
PSD_T	1.52 (.94)	0–4	.67**	1.45 (.99)	0–4	.58*	
NPR_T	.62 (.60)	0–3		.50 (.43)	0–1.5		
PSD_S	1.01 (.90)	0–4.4	.69**	1.25 (1.06)	0–4.2	.69**	<i>t</i> (133) = -2.29 <i>p</i> = .02
NPR_S	.49 (.64)	0–3		.58 (.68)	0–3		<i>t</i> (691) = -1.24 <i>p</i> = .22

Note. PSD\_T = Professional self-doubt trait-based. NPR\_T = Negative personal reaction trait-based. PSD\_S = Professional self-doubt state-based. PSD\_S = Negative personal reaction state-based. “*r*” describes the correlations between PSD\_T and NPR\_T or between PSD\_S and NPR\_S in the different samples. The correlation of the state-based scales concerns the first assessment point (around the 5th session).

\*\**p* ≤ .01 (two-tailed).

\**p* ≤ .05 (two-tailed).

Table III. Therapist trait-based difficulties (rated across patients) and patient progress: Results of multilevel modeling analyses.

	BSI sample with GSI as the outcome measure Estimate (S.E)	IIP sample with GSI as the outcome measure Estimate (S.E)	IIP sample with IIP as the outcome measure Estimate (S.E)
Fixed effects			
Intercept	<b>1.15 (.04)***</b>	<b>1.26 (.09)***</b>	<b>1.60 (.06)***</b>
Time	<b>-.01 (.0008)***</b>	<b>-.006 (.002)***</b>	<b>-.003 (.0008)**</b>
PSD_T	.01 (.05)	.02 (.12)	.10 (.07)
NPR_T	.004 (.05)	-.02 (.14)	-.06 (.06)
PSD_T×Time	.0001 (.001)	.003 (.003)	<b>.002 (.001)*</b>
NPR_T×Time	-.001 (.001)	-.0001 (.002)	<b>-.002 (.001)*</b>
Model fit (-2LL)			
Unconditional null model	4666.283	846.441	786.649
Final model	4010.492	652.056	758.521

Note. PSD\_T = Professional self-doubt trait-based. NPR\_T = Negative personal reaction trait-based. BSI = Brief Symptom Inventory. IIP = Inventory of Interpersonal Problems. GSI = Global Severity Index of the Brief Symptom Inventory. -2LL = -2 Log Likelihood. Estimation method: maximum likelihood; bold characters indicate significant results. PSD\_T and NPR\_T were *z*-standardized.

\**p* ≤ .05.

\*\**p* ≤ .01.

\*\*\**p* ≤ .001.

measures at least at three measurement points were excluded to obtain reliable results with multilevel modeling. Consequently, of the 1049 patients 621 patients met the inclusion criteria and were analyzed regarding patients' symptomatology (BSI sample of the current study). Only for 106 of the 1049 patients data about their interpersonal problems (IIP sample of the present study) were available (see Figure 1 for flow diagram). We compared the BSI and IIP regarding their baseline variables (see Tables I and II).

*BSI sample (patients' symptomatology).* *N* = 621 patients filled in the BSI at least at 3 measurement points. See Table I for the sample description. This sample was used to analyze the research questions with the GSI as the outcome variable.

*IIP sample (patients' interpersonal problems).* *N* = 106 patients completed the IIP at least at 3 measurement points. The IIP was part of the diagnostic routine only for a short time interval at the outpatient center.

The sample is therefore substantially smaller than the BSI sample (see Table I for the sample description). Ninety-six patients of the 106 patients are also included in the BSI sample. The sample consisting of the *N* = 106 patients with at least 3 IIP assessments was used to analyze the research questions with the IIP as the outcome variable. Moreover, this sample was used to investigate the research questions with the GSI as the outcome measure in order to evaluate whether the same results can be obtained with the GSI as the outcome measure for the BSI sample and the IIP sample or whether the sample difference might be a confounder (see Tables III and V).

The diagnoses reported in Table I were based on the German version of the Structured Clinical Interview for Axis I and II DSM-IV Disorders (Fydrich, Renneberg, Schmitz, & Wittchen, 1997; Wittchen, Wunderlich, Gruschwitz, & Zaudig, 1997) and conducted by the therapist treating the patient.

Table IV. Variation of therapist state-based difficulties (specific patients at a particular time in therapy): Results of multilevel modeling analyses.

	BSI sample		IIP sample	
	PSD_S Estimate (S.E.)	NPR_S Estimate (S.E.)	PSD_S Estimate (S.E.)	NPR_S Estimate (S.E.)
Fixed effects				
Intercept	<b>1.15 (.09)***</b>	<b>.53 (.06)***</b>	<b>1.63 (.19)***</b>	<b>.81 (.16)***</b>
Time	<b>-.006 (.001)***</b>	<b>-.002 (.0008)*</b>	<b>-.01 (.003)***</b>	-.004 (.002)
Random effects				
Therapist intercept	<b>.26 (.06)***</b>	<b>.10(.03)***</b>	<b>.43 (.18)*</b>	<b>.39 (.16)*</b>
Patient intercept	<b>.37 (.04)***</b>	<b>.13 (.02)***</b>	<b>.70 (.14)***</b>	<b>.33 (.07)***</b>
Patient slope	<b>.0003 (.00005)***</b>	<b>.00006 (.00002)**</b>	<b>.0007 (.0002)***</b>	<b>.0003 (.00009)**</b>
Model fit (-2LL)				
Unconditional null model	3710.607	2501.963	1256.708	811.149
Final model	3610.966	2466.590	1168.509	746.601

Note. BSI = Brief Symptom Inventory. IIP = Inventory of Interpersonal Problems.

PSD\_S = Professional self-doubt state-based. NPR\_S = Negative personal reaction state-based. -2LL = -2 Log Likelihood. Estimation method: maximum likelihood; bold characters indicate significant results.

\* $p \leq .05$ .

\*\* $p \leq .01$ .

\*\*\* $p \leq .001$ .

**Therapists.** All CBT therapists held a master's or equivalent degree in psychology and worked at the outpatient center of Ruhr University Bochum (Germany). The majority of the therapists ( $n = 39$ ) were participants of the annually starting CBT training program of the cohorts 2011 and 2012. One therapist was a former trainee of the program, who continued treating patients at the outpatient center. All therapists spoke German on a native level. Therapists were either licensed in CBT or in training to become a licensed CBT psychotherapist. The

trainees were supervised on average every 4th session by a senior psychotherapist with at least 5 years of experience as licensed psychotherapist and at least 3 years of experience in teaching psychotherapy.

*BSI sample (patients' symptomatology).* The therapist sample consisted of  $n = 40$  psychologists, the majority being female (85%;  $n = 34$ ). Therapists' level of psychotherapy experience was rather low, as 97.5% ( $n = 39$ ) were still undergoing training to become licensed CBT psychotherapists or had

Table V. Therapist state-based difficulties (rated for each individual patient repeatedly) and patient progress: Results of multilevel modeling analyses.

	BSI sample with GSI as outcome measure Estimate (S.E.)	IIP sample with GSI as outcome measure Estimate (S.E.)	IIP sample with IIP as outcome measure Estimate (S.E.)
Fixed effects			
Intercept	<b>1.04 (.05)***</b>	<b>1.25 (.11)***</b>	1.60 (.08)***
Time	<b>-.01 (.001)***</b>	<b>-.01 (.002)***</b>	<b>-.005 (.002)**</b>
PSD_S	<b>.06 (.03)*</b>	-.02 (.04)	-.003 (.04)
NPR_S	-.02 (.04)	-.06 (.07)	-.08 (.06)
PSD_S $\times$ time	<b>.002 (.001)*</b>	.002 (.003)	.001 (.001)
NPR_S $\times$ time	-.0002 (.002)	-.0003 (.002)	.002 (.002)
Model fit (-2LL)			
Unconditional null model	4666.283	846.441	786.649
Final model	2328.732	584.595	447.900

Note. BSI = Brief Symptom Inventory. IIP = Inventory of Interpersonal Problems. PSD\_S = Professional self-doubt state-based. NPR\_S = Negative personal reaction state-based. GSI = Global Severity Index of the Brief Symptom Inventory. IIP = Inventory of Interpersonal Problems. BSI sample:  $n = 621$  patients and  $n = 40$  therapists. IIP sample:  $n = 106$  patients and  $n = 18$  therapists. -2LL = -2 Log Likelihood. Estimation method: maximum likelihood; bold characters indicate significant results. The term PSD\_S  $\times$  time in the IIP sample with the GSI as outcome measure was almost significant  $p = .07$ .

\* $p \leq .05$  \*\* $p \leq .01$ . \*\*\* $p \leq .001$ .

completed their training within the last year: The average psychotherapy experience was  $M = 1.7$  years ( $SD = 0.93$ ; range 1–5). The average patient caseload per therapist was  $M = 15.53$  ( $SD = 9.5$ , range 3–48). Most of the therapists (72.5%,  $n = 29$ ) were in a relationship, whereas 10 (25%) were single and 1 was married when filling out the DPCCQ (Orlinsky et al., 1999). Eighteen (62%) of the 29 therapists living in a partnership lived in one household with their partner, 11 (38%) did not. Almost all therapists ( $n = 37$ , 93%) were born in Germany, 3 (8%) were born in Poland, but grew up bilingually.

*IIP sample (patients' interpersonal problems).* The therapists in the IIP sample were the 17 CBT trainees and the one recently licensed psychotherapist ( $n = 18$ ), with the majority being female (78%;  $n = 14$ ) and with  $M = 2.11$  ( $SD = 1.08$ ; range 1–5) years of psychotherapy experience. On average, therapists treated  $M = 5.89$  patients ( $SD = 5.69$ ; range 1–20). Twelve (67%) therapists stated to be in a relationship, 5 (30%) were single and 1 therapist was married. Seven of the 12 (58%) therapists in a partnership lived together with their partners, whereas 5 (42%) had separate flats. All therapists were born in Germany.

## Treatment

All patients received a CBT-based treatment (Schulte & Eifert, 2002) at the outpatient center of Ruhr University Bochum. Treatment sessions typically took place on a weekly basis, lasting 50 min (see Table I for overall average treatment length) and were conducted under supervision. Treatments were based on individual case conceptualization. Therefore, therapists did not follow a specific manual, but supervisors discussed with the trainees the use of evidence-based interventions from CBT manuals.

## Measures

**Patient-reported outcome measures.** *Symptomatology.* General psychological symptom severity was assessed with the German Version of the BSI (Derogatis, 1975; German translation: Franke, 2000). This self-report questionnaire with 53 items is a short form of the Derogatis Symptom Checklist (SCL-90-R) and measures general psychological symptom severity (items are rates on a 5-point Likert scale). In the present study, we used the general symptom index (GSI), which is a common variable to represent patients' general psychopathology and regarded as sensitive to psychotherapy change (Ogles, Lambert, & Masters, 1996). Preceding studies confirmed the reliable psychometric

properties of the BSI with an internal consistency of .92 and a retest reliability of  $r_{tt}.90$  (Franke, 2000). In the BSI sample of the current study, Cronbach's alpha was .96 at the first assessment point ( $n = 621$  patients). Patients included in the analyses answered the BSI on average  $M = 5.17$  times during therapy (range 3–13;  $SD = 1.76$ ).

*Interpersonal problems.* Interpersonal problems were assessed with the German Version of the IIPs (Horowitz et al., 1988; German translation: Horowitz et al., 2000). The overall IIP score was calculated as the mean of all 64 items. Test-retest reliability ( $r_{tt}.98$ ), internal consistency ( $\alpha = .94$ ) and construct validity of the IIP have been demonstrated to be excellent (Horowitz et al., 2000). Internal consistency (Cronbach's alpha) at the first assessment point in the IIP sample of the present study was .93 ( $n = 106$  patients). The patients included in this sample answered the IIP on average  $M = 6.78$  times (range 3–12;  $SD = 2.50$ ) during CBT.

## Therapist measures

**Therapists' trait-based difficulties.** *Therapist Difficulties Scales (TDS).* As part of the DPCCQ (Orlinsky et al., 1999) all psychotherapists filled out the TDS (Orlinsky & Rønnestad, 2005). The TDS is a 19-item scale assessing difficulties therapists can encounter during their therapeutic work across patients, thus assessing these difficulties as a psychotherapist's trait-based variable. The core question is: "Presently, how often do you feel ..." followed by the different items, for example, "... lacking in confidence that you can have a beneficial effect on a patient?", answered on a six-point Likert-type scale ranging from 0 ("never") to 5 ("very often"). In a factor analysis, Orlinsky and Rønnestad (2005) found three factors: "PSD", "NPR", and "frustrating treatment case." The subscales can be interpreted as different origins of the perceived difficulties: therapist, patient, and circumstances. The TDS has been used in a number of different studies to investigate difficulties in therapeutic practice (Black et al., 2005; Heinonen et al., 2012; Heinonen, Knekt, et al., 2014; Heinonen, Lindfors, et al., 2014; Nissen-Lie et al., 2010, 2013; Nissen-Lie, Havik, et al., 2015; Nissen-Lie, Rønnestad, et al., 2015; Sanness, 2012; Schröder et al., 2009). The three factors have been reproduced in different samples (Nissen-Lie, 2001; Orlinsky & Rønnestad, 2005; Willutzki et al., 1997). Nissen-Lie et al. (2010) reproduced two (PSD and NPR) of the three factors in their sample of 70 therapists and focused their research on PSD and NPR as the most powerful predictors of psychotherapy process and outcome in terms of explained variance.

*Professional Self-Doubt, Trait-Based (PSD\_T)* is a subscale of the TDS. It consists of five items and assesses difficulties in practice that the therapist may attribute to himself or herself (e.g., “Presently, how often do you feel... unsure how best to deal effectively with a patient?”). The items included in this factor are based on the results of the International Study on the Development of Psychotherapists (ISDP; Orlinsky & Rønnestad, 2005). In our data (article in preparation), we found a factor structure similar to the results reported by Orlinsky and Rønnestad (2005) in terms of factor loadings (range .45–.75, p. 229). The PSD scale used in the present study includes the same items as in the ISDP study and comprises the following five items: (i) Lacking in confidence that you can have a beneficial effect on a patient; (ii) unsure how best to deal effectively with a patient; (iii) demoralized by your inability to find ways to help a patient; (iv) afraid that you are doing more harm than good in treating a patient; (v) unable to generate sufficient momentum to move therapy with a patient in a constructive direction. The factor obtained an acceptable internal consistency score in different samples of therapists: Orlinsky and Rønnestad (2005): Cronbach’s  $\alpha = .77$ ; in our sample: Cronbach’s  $\alpha = .80$  ( $n = 40$  therapists); for descriptives and the correlation of PSD\_T and NPR\_T in our samples see Table II).

*Negative Personal Reaction, Trait-Based (NPR\_T)* consists of four items from the TDS and assesses difficulties in practice that can rather be attributed to the patient. Following Orlinsky and Rønnestad (2005), the NPR\_T includes the following items: (i) Unable to find something to like or respect in a patient; (ii) unable to have much real empathy for a patient’s experiences; (iii) unable to withstand a patient’s emotional neediness; (iv) uneasy that personal values make it difficult to maintain an appropriate attitude. The factor obtained an acceptable internal consistency score in different samples of therapists: Orlinsky and Rønnestad (2005): Cronbach’s  $\alpha = .74$ ; in our sample: Cronbach’s  $\alpha = .67$  ( $n = 40$  therapists); for descriptives and the correlation between NPR\_T and PSD\_T in our samples see Table II).

**Therapists’ state-based difficulties.** *Professional Self-Doubt (PSD\_S)* and *Negative Personal Reaction (NPR\_S)* for a specific patient at a particular time in therapy. The TDS-IP was used to assess therapists’ difficulties with a specific patient at a particular time. The TDS-IP is a patient-specific adaptation (Willutzki et al., 1997) of the TDS, with a focus on the patient–therapist dyad and therapists’ state-based difficulties. The therapist was asked for each of his or her patients at several points of time over

the course of CBT about the frequency of experienced difficulties. Therefore, in the TDS-IP, the core question of the TDS was changed as follows: Instead of asking psychotherapists about current difficulties with patients *in general*, they were asked: “In your *current work with your patient* [‘name of the patient’] how often do you experience...?” Items included in the PSD\_S and NPR\_S are the same as in the PSD\_T and NPR\_T, respectively. The items of the TDS-IP are answered on a 6-point Likert scale ranging from min. 0 (“never”) to max. 5 (“very often”). Therapists answered the therapist difficulties scales for each of their patients several times: In the BSI sample, therapists rated their difficulties on average 2.98 times (range 1–11;  $SD = 2.02$ ); in the IIP sample, on average 5.11 times per patient during CBT (range 1–11;  $SD = 2.50$ ). See Table II for descriptives and correlations between PSD\_S and NPR\_S at the first assessment point (around the 5th session) for the BSI sample and the IIP sample respectively. Cronbach’s alpha for the state-based scales at the first assessment point reached .83 for PSD\_S and .74 for NPR\_S ( $n = 587$  patients).

No significant differences were found between our samples and the sample analyzed by Nissen-Lie et al. (2013) regarding the level of therapists’ trait-based difficulties (tested with *t*-tests for independent samples based on the data provided by Nissen-Lie et al. (2013); results are not shown in this article). Table II shows that, for the BSI sample, PSD on the state-level (PSD\_S) tended to be lower than on the trait-level (PSD\_T); this was not observed for NPR. The reported maximum of NPR\_T in the IIP sample was low ( $M = .5$  with a maximum of 1.5). Correlations between the difficulties scales in our samples were comparable to those described by Nissen-Lie et al. (2013).

## Data Analysis

All statistical analyses have been conducted with IBM SPSS 22.

**Therapist effect.** To estimate the therapist effect in our sample, we computed the intraclass correlation coefficient with the results from the unconditional null model of the BSI sample by dividing the random intercept of the therapist level through the sum of the random effects including the random intercepts at the patient and therapist levels as well as the residuum (see Heck, Thomas, & Tabata, 2014, p. 138).

**Multilevel modeling.** Linear multilevel models were performed to investigate the research questions described above. An advantage of multilevel

modeling is that it accounts for the nested structure of the data and the data dependence and therefore avoids an inflated Type I error (Raudenbush & Bryk, 2002). It also allows unequally distributed measurement waves like those in our dataset (Tasca & Gallop, 2009).

All multilevel models were three-level models: Assessments as level-1 nested within patients as level-2, which were nested within therapists as level-3. Full maximum likelihood estimation was used in all multilevel models. On the patient level, intercept and slope were allowed to vary randomly, and an unstructured variance-covariance matrix was selected. On the therapist level, only the intercept was allowed to vary randomly and, therefore, the identity variance-covariance matrix was applied. No random slope was added on the therapist level because of the relatively small sample of therapists. All described multilevel models were compared with unconditional null models regarding the model fit reflected by the  $-2LL$ , whereby lower values indicate a better fit. All statistical tests were conducted as two-tailed tests, and the significance value was set to  $p \leq .05$ .

**Multilevel models for therapists' trait-based difficulties (research question 1).** Three multilevel models analyzing therapists' trait-based difficulties were conducted. The NPR\_T and the PSD\_T scale were added as time-invariant covariates and the time variable (session number, first session set to 0) as predictor. The estimation for the "time" variable reflects the average change in the outcome measures from one session to the next session. Because of easier interpretation and lower fit indices (e.g.,  $-2$  Log Likelihood ( $-2LL$ ); Akaike Information Criterion (AIC)) we did not log-transform the time variable. The following fixed effects were evaluated: intercept, NPR\_T, PSD\_T, time, time\*NPR\_T, time\*PSD\_T. NPR\_T and PSD\_T were  $z$ -standardized for these analyses. Following Nissen-Lie et al. (2010, 2013), both difficulty scales were added to one and the same multilevel model to assess the specific contribution of each factor.

- *BSI sample:* One model with the BSI sample was conducted including the GSI as dependent variable.
- *IIP sample:* The next two models analyzing therapists' trait-based difficulties were conducted with the IIP sample. The first model included the IIP as dependent variable and the second included the GSI as dependent variable. The multilevel model with the IIP as the dependent variable did not converge with a random intercept and random slope at the patient level and a random

intercept at the therapist level. Since the fit indices were better for the random intercept only than for the random slope only model, we only kept the random intercept and omitted the random slope at the patient level. The multilevel model with the GSI as the dependent variable converged with a random intercept and random slope at the patient level and a random intercept at the therapist level also in this IIP sample.

**Multilevel models for therapists' state-based difficulties (research questions 2 and 3).** Seven multilevel models were performed to address the research questions two and three regarding *therapists' state-based difficulties*.

Research question 2: Four models tested whether PSD\_S and/or NPR\_S varied significantly over time. They all had time (session number, first session set to 0) as predictor.

- *BSI sample:* Two multilevel models tested whether PSD\_S and/or NPR\_S varied significantly over time for the BSI sample. The first model included PSD\_S as dependent variable and the second model NPR\_S as dependent variable.
- *IIP sample:* Two multilevel models were conducted with the IIP sample: One with PSD\_S as dependent variable and the other with NPR\_S as dependent variable.

Research question 3: Three further models evaluated, whether the change of the PSD\_S and/or NPR\_S ratings a therapist experienced for an individual patient over the course of therapy were associated with patient progress. In all models, the NPR\_S as well as the PSD\_S scale were added as time-varying covariates and the time variable (session number, first session set to 0) as predictor. The following main and interaction effects were evaluated: intercept, NPR\_S, PSD\_S, time, time\*NPR\_S, time\*PSD\_S. Again both difficulty scales were added to one model to assess the individual contribution of each scale.

- *BSI sample:* One multilevel model one model was performed with the BSI sample and the GSI as dependent variable.
- *IIP sample:* Two more models were performed with the IIP sample. One model with the IIP as dependent variable, and the other one with the GSI as dependent variable.

## Results

### Comparison of the BSI and the IIP sample

Comparing both samples, there was a significant effect for the IIP pretreatment score,  $t(233) = -3.4$

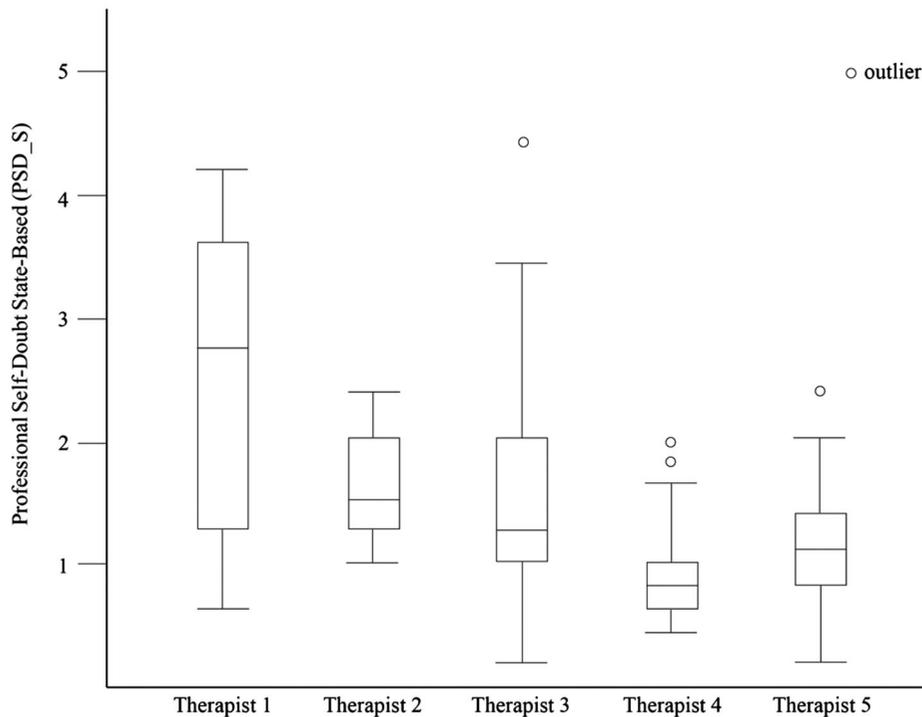


Figure 2. State-based professional self-doubt varies between therapists as well as within therapists at the first assessment in therapy (around 5th session).

$p = .001$  ( $g = .41$ ) and treatment length  $t(724) = -3.24$   $p = .002$  ( $g = .40$ ). The results indicate that patients in the IIP sample had more interpersonal problems at the beginning of the treatment and stayed longer in therapy than patients in the BSI sample. In addition the samples differed regarding therapists' experience of professional self-doubt (PSD\_S) at the first difficulties assessment  $t(133) = -2.29$   $p = .02$  ( $g = .26$ ). No other sample differences could be observed in terms of patients' age, gender, comorbidity, GSI pretreatment score and therapist's NPR (NPR\_S) towards the patient at the first state-based difficulties assessment (see Tables I and Table II).

### Therapist effect

The therapist effect in the BSI sample was 4.5%. We did not compute the therapist effect in the IIP sample because the sample size was too small to obtain reliable results.

### Therapists' Trait-Based Difficulties

- (1) Can the results of Nissen-Lie et al. (2010, 2013) and Nissen-Lie, Rønnestad, et al. (2015) be replicated in a different sample of therapists and patients, showing that PSD/

NPR are associated with patient progress when PSD/NPR are assessed across patients and time (trait-based therapists' difficulties; PSD\_T; NPR\_T)?

Neither PSD\_T nor NPR\_T had an effect on patients' symptom change (GSI) over time. This was the case for the BSI sample and the IIP sample (see Table III).

However, both therapist variables were relevant for the IIP change: NPR\_T showed a beneficial effect on patients' interpersonal problems over time ( $B = -.002$ ,  $p \leq .05$ ), while PSD\_T had the opposite effect ( $B = .002$ ,  $p \leq .05$ ): Higher therapists' NPR to patients at the trait-level as well as lower therapists' PSD at the trait-level were associated with stronger reductions in patients' interpersonal problems.

Patients significantly improved over the course of therapy on the GSI ( $B = -.01$ ,  $p \leq .001$ ) and the IIP ( $B = -.003$ ,  $p \leq .01$ ) when statistically controlling for NPR\_T and PSD\_T.

### Therapists' State-Based Difficulties

- (2) Do PSD and NPR vary between patients and change during the process of psychotherapy when PSD/NPR are assessed repeatedly with regard to a specific patient (state-based therapists' difficulties; PSD\_S; NPR\_S)?

All tested random effects for therapists' state-based difficulties scales in both samples were significant (see Table IV), indicating that PSD and NPR varied at the first assessment point between therapists (therapist level random intercept;  $p \leq .05$ ) as well as between patients (patient-level random intercept  $p \leq .001$ ) and changed differently between patients-over time (patient-level random slope;  $p \leq .01$ ).

Regarding the fixed effects for the BSI sample, PSD\_S ( $B = -.006$ ,  $p \leq .001$ ) and NPR\_S ( $B = -.002$ ,  $p \leq .05$ ) significantly declined over time. For the IIP sample, PSD\_S significantly decreased from one session to the other ( $B = -.01$ ,  $p \leq .001$ ), while there was only a trend for NPR\_S to decline over time ( $B = -.004$ ,  $p = .08$ ).

Figure 2 exemplifies how state-based PSD varies between therapists as well as within therapists at the first assessment in therapy (around 5th session). Results indicate that the contextualization of therapists' experience with a particular client over time is necessary, and both scales are sensitive to change in therapists' experience with a patient over time (see Table IV).

- (3) Are changes in PSD/NPR associated with patient progress when PSD and NPR are assessed repeatedly with regard to a specific patient (state-based therapist difficulties; PSD\_S; NPR\_S)?

As Table V shows, only PSD\_S ( $B = .06$ ,  $p \leq .05$ ) was significantly associated with patients' GSI at the beginning of the therapy, indicating that higher psychological symptom severity in a given patient was associated with higher PSD of this patient's therapist at the start of treatment. This effect was found in the BSI sample, but did not attain statistical significance for the IIP sample. There was also a significant interaction effect between PSD\_S and time ( $B = .002$ ,  $p \leq .05$ ) in the BSI sample. This indicates increases in PSD\_S for a specific patient over time to be associated with a less favorable treatment course on the GSI. This effect was not significant in the IIP sample, even though a tendency could be observed ( $B = .0002$ ,  $p = .07$ ). NPR\_S did not exert any effect on the GSI outcome measure (BSI sample and IIP sample), and neither NPR\_S nor PSD\_S affected the IIP outcome measure (IIP sample).

## Discussion

This study explored the impact of therapists' characteristics on the course of patient-reported outcomes during psychotherapy. We focused on therapists'

subjectively perceived difficulties—PSD and NPR—in their work in general (trait-based) and in their work with specific patients at a particular time in therapy (state-based) in a sample of CBT trainees. The paper had two aims: The first one was to investigate whether prior results of Nissen-Lie et al. (2010, 2013) and Nissen-Lie, Rønnestad, et al. (2015) addressing the impact of therapists' trait-based difficulties on the course of patient-reported outcomes can be replicated in a different sample; the second one was to look at variations of difficulties encountered by trainee psychotherapists during CBT with individual patients over time and their associations with patient progress.

Similar to the studies by Nissen-Lie et al. (2010, 2013) and Nissen-Lie, Rønnestad, et al. (2015), we found a significant relation between patients' changes in interpersonal problems and therapists' trait-based difficulties and no such relation regarding patients' symptomatology. Contrary to Nissen-Lie et al. (2013) a beneficial effect of trait-based NPR and an adverse one of trait-based PSD on the reduction of interpersonal problems could be shown. The contrary results in our study and Nissen-Lie et al.'s studies may be caused by the different therapist characteristics, considering the years of experience (trainees vs. 10 years of experience) and theoretical orientation (CBT vs. psychodynamic/psychoanalytic). This indicates that the difficulty scales may have different meanings for therapists at different career stages and are influenced by the therapist's theoretical background. Even though the level of trait-based difficulties in our sample was similar to that in Nissen-Lie et al. (2013; see Table II), it could be hypothesized that NPR towards patients might have a different function between experienced therapists as investigated by Nissen-Lie et al. and the rather inexperienced trainees analyzed in the present study. At the beginning of their psychotherapeutic work, trainees might be very reluctant to allow themselves to experience NPR towards their patients, since such feelings might not be helpful to build a positive image of oneself as a future therapist. Possibly, trainees with the trait not to avoid challenging feelings or cognitions such NPR can use this trait in favor of the psychotherapy process. The trait to avoid challenging feelings and cognitions has also been called "experiential avoidance" (Hayes, Wilson, Gifford, Follette, & Strosahl, 1996). Interestingly, Scherr, Herbert, and Forman (2015) found that therapists with less experiential avoidance intend to use exposure exercises more often than therapists with higher experiential avoidance in patients with obsessive-compulsive disorders. Exposure exercises are highly effective for patients with obsessive-compulsive disorders and

anxiety disorders in general (Abramowitz, 1996; Rosa-Alcázar, Sánchez-Meca, Gómez-Conesa, & Marín-Martínez, 2008) but also related to challenging feelings for the therapist (Schumacher et al., 2014). According to Hayes and Gelso (2001) and Gelso and Hayes (2007), a lack of fear of own unpopular reactions towards patients indicates a mature stance that is beneficial to therapeutic work and associated with countertransference management. A recent meta-analysis by Hayes, Gelso, and Hummel (2011) showed that countertransference management, including a therapist's openness to own countertransference, is positively associated with patients' outcome. In experienced therapists, on the other hand, higher NPR on a trait-level might reflect a negative therapeutic relationship. Future research is necessary to clarify whether NPR on a trait-level has different meanings for experienced and inexperienced therapists.

The fact that the developmental stage of the therapists was different in our study compared to the studies by Nissen-Lie et al. (2013) and Nissen-Lie, Rønnestad, et al. (2015) might also explain the diverging findings regarding trait-based PSD. Therapists with higher PSD at the trait-level had a detrimental effect on patients' reduction of interpersonal problems in the current study, but a beneficial effect in the studies by Nissen-Lie and colleagues. There is evidence that inexperienced therapists tend to feel more insecure about their own therapeutic skills and abilities than more experienced colleagues (Orlinsky & Rønnestad, 2005). In contrast, therapists at an advanced level of professional practice are regarded as generally confident, believing in their judgment and skills (Rønnestad & Skovholt, 2012). According to Stoltenberg, McNeill, and Delworth (1998) even these experienced therapists may still experience self-doubts, but do not impair their ability to function as a therapist. For experienced therapists as in Nissen-Lie et al.'s sample trait-based PSD might reflect an open self-reflective stance towards one's work grounded in a generally safe and self-confident basis associated with a humble attitude that guards against over-confidence. For the therapists in our sample, with just about 18 months in training and about 6–12 months working as a therapist with outpatients when filling in our measures, trait-based PSD may be a very irritating experience which interferes with their self-organization as a future therapist.

Apart from therapists' professional experience, the psychotherapeutic orientation should be taken into account. There is some evidence that theoretical orientation is associated with different personal attributes of the therapist such as personality, beliefs, and interpersonal style (Arthur, 2000, 2001; Buckman &

Barker, 2010; McLennan & Poznanski, 2004; Poznanski & McLennan, 2003). Studies found, for example, that therapists with a psychodynamic orientation tend to be less objective, more open to experience, and rather unstructured as well as more symbolic than CBT-oriented therapists (Heinonen & Orlinsky, 2013; Scragg, Bor, & Watts, 1999). Therapists with a psychoanalytic/psychodynamic orientation might have a stronger focus on working with their own challenging emotions. Therefore, the questions of the PSD and also of the NPR scale might be interpreted differently by CBT and psychoanalytic/psychodynamic therapists, and this could be related to the different findings between our study and the ones by Nissen-Lie and colleagues.

In addition, Nissen-Lie et al. (2010, 2013) and Nissen-Lie, Rønnestad, et al. (2015) used all items of the PSD and NPR scales that we used but added more items to the PSD and NPR subscales than recommended by Orlinsky and Rønnestad (2005) and used by us.

Altogether we consider the trait-based difficulties scales as important therapist characteristics that require further clarification in their particular meaning. This was also shown by the previous work of Nissen-Lie et al. where trait-based PSD was first found to be positively related with early patient-rated working alliance (2010). Later work, however, showed a negative association with patient-rated alliance levels over the course of therapy (Nissen-Lie, Havik, et al., 2015) and a positive association with the reduction of patients' interpersonal problems during and after therapy (Nissen-Lie et al., 2013).

Looking at the state-based difficulties, we found that the course of the symptoms becomes less favorable in case therapists' PSD increases during psychotherapy. Noteworthy, PSD measured on a state-level affected the course of symptoms (but not interpersonal problems), whereas PSD measured on a trait-level influenced the course of interpersonal problems (but not symptoms). Moreover, our results show that (i) therapists differ from each other concerning the level of PSD\_S and NPR\_S at the beginning of CBT, (ii) PSD\_S and NPR\_S are experienced differently between patients at the beginning of CBT, and (iii) the way in which therapists' experiences of PSD\_S as well as NPR\_S changes during CBT and varies between patients. Such situation-person interactions have long been discussed in interactional models of personality (Mischel, 2004), but are not often considered in psychotherapy research. In our opinion, patient-specific and time-sensitive assessment of therapists' experiences during therapeutic practice is necessary for a better understanding of the therapist effect. Findings of the study at hand (therapists' difficulties changed over therapy; trait-

based and state-based assessments of therapists' difficulties had a different impact on patient progress) demonstrate that the global assessment of therapists' experience—as undertaken with trait-based difficulties—leads to a reduction of existing variation and therefore measures a different aspect of therapists' self-assessed experience. To our knowledge, there is no other study assessing the difficulties therapists experience in their work with a particular patient over time (state-based difficulties). This seems essential in order to capture more complex therapist variables, like therapists' "appropriate responsiveness" (Stiles, 2013), and their impact on psychotherapy process and outcome. The ability to adapt effectively requires an enhanced sensitivity to patients' reactions and prevailing circumstances in the therapeutic process. The findings of our study indicate that it might be useful to assess this ongoing adaptation.

### Limitations

A major limitation of our study is that the results considering the IIP are based only on the small sample of 106 patients. This hampers the interpretation of the different results regarding the different outcome measures and did not allow us to control for other possible predictors. To ensure the validity of our results, we conducted the analysis regarding the psychological symptom severity (GSI) in the BSI sample of the current study and also in the IIP sample of 106 patients. We mostly found similar results in both samples. Comparing the two samples, we found significant differences in terms of IIP pre-treatment scores, treatment length and PSD\_S at the first difficulties assessment. Whereas the differences regarding therapist's PSD were rather small ( $g = .26$ ) the effect sizes for the IIP intake scores ( $g = .41$ ) and treatment length ( $g = .40$ ) were medium. This hampers the external validity of the results concerning the impact of therapists' trait-based difficulties on patients' interpersonal problems. Further research on larger samples is required to proof the general importance of the investigated therapist variables. Moreover the small sample size, did not allow us to control for patient predictors like treatment duration, patients' age or number of diagnoses. To our knowledge, neither did preceding studies of Nissen-Lie et al. Replications of the results regarding patients' interpersonal problems in more generalizable samples are therefore necessary.

The reported therapist effect of 4.5% in our sample should be interpreted with caution. Following the recommendations of Schiefele et al. (2016), our sample size is not able to generate a reliable estimate of the therapist effect.

In our replication study, we only focused on the relationship of therapists' trait-based difficulties with patient progress, which is only a small part of previous research in this area. In contrast to Nissen-Lie et al. (2010, 2013) we did not look at patient-rated working alliance. Besides, the working alliance, resource activation, problem activation, mastery, and clarification are important process variables and mechanisms of change in psychotherapy (Mander et al., 2013; Mander, Schlarb, et al., 2015; Mander, Jacob, et al., 2015) and their relation to therapists' difficulties should be investigated in future studies. The fact that we did not analyze general change mechanisms or more person variables of the therapists (such as self-affiliation, see Nissen-Lie, Rønnestad, et al., 2015) can be seen as a further shortcoming of the study at hand.

Another limitation of the research question on state-based difficulties lies in the fact that therapists did not answer the patient-related questionnaire as often as patients filled in the outcome measure.

No statement about causality can be made about state-based difficulties: On the one hand, it may be that therapists who increasingly doubt themselves during therapy have an adverse impact on patients' healing process, for example, by influencing the therapeutic alliance in a negative way or by weakening the patient's trust in the therapist's competence. On the other hand, increasing PSD may be a reaction to stagnation in psychotherapy. It is also possible that these processes are intertwined. Further research could add in-depth interviews or video analyses of therapy sessions as complementary data sources useful to clarify the longitudinal relationship between symptom change and PSD.

A general limitation of the study is that it is a naturalistic study. Patients received CBT of different contents and lengths. The data were collected at the outpatient center of Ruhr University Bochum where primarily young therapists are trained to become licensed psychotherapists. The general conditions of the outpatient center may operate as a potential selection bias as they may not be acceptable for a certain group of patients. Unfortunately, we cannot control for that as there are no data available on which patients refused a therapy because of the treatment conditions. Almost all therapies were conducted under supervision, which is supposed to help therapists to deal with the problems they encounter. It is possible that supervision had an influence on therapists' experience of difficulties and their way to deal with them: Supervisors may have had a specific focus on some difficulties while neglecting others; therapists may have addressed some difficulties during supervision while trying to

cope with others on their own. We do not know whether these influences from supervision had an impact on variations in therapists' experience of difficulties and patients' progress. The embeddedness of supervision in the psychotherapies analyzed here makes the findings less generalizable to general clinical practice. Future research could investigate whether and how supervision processes are related to the difficulties therapists experience in their therapeutic work.

### Conclusion and Implications

To sum up, our study was the first one trying to replicate the preceding findings of Nissen-Lie et al. (2010, 2013) in a different sample and the first one assessing therapists' difficulties with a particular patient repeatedly over the course of treatment. We could replicate an influence of therapists' trait-based difficulties on patients' interpersonal problems, but with diverge effects of the scales compared to the results of Nissen-Lie et al. (2010, 2013). And we demonstrated that a patient-therapist-time contextualization might be relevant for the assessment of more complex therapist variables and in consequence a more elaborate understanding of the therapist effect. Possible implications of the results of the present study for training and supervision might be that CBT trainees should not be scared of their NPR towards their patients and that increases in PSD of CBT trainees should be monitored and discussed.

### Funding

This study was funded by the University Witten/Herdecke (grant number: IFF2014-11).

### References

- Abramowitz, J. S. (1996). Variants of exposure and response prevention in the treatment of obsessive-compulsive disorder: A meta-analysis. *Behavior Therapy, 27*(4), 583–600. doi:10.1016/S0005-7894(96)80045-1
- Anderson, T., Ogles, B. M., Patterson, C. L., Lambert, M. J., & Vermeersch, D. A. (2009). Therapist effects: Facilitative interpersonal skills as a predictor of therapist success. *Journal of Clinical Psychology, 65*(7), 755–768. doi:10.1002/jclp.20583
- Arthur, A. R. (2000). The personality and cognitive-epistemological traits of cognitive-behavioural and psychoanalytic psychotherapists. *British Journal of Medical Psychology, 73*(2), 243–257. doi:10.1348/000711200160453
- Arthur, A. R. (2001). Personality, epistemology and psychotherapists' choice of theoretical model: A review and analysis. *European Journal of Psychotherapy & Counselling, 4*(1), 45–64. doi:10.1080/13642530110040082
- 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* (pp. 258–297). Hoboken, NJ: Wiley.
- Beutler, L. E., Malik, M., Alimohamed, S., Harwood, T. M., Talebi, H., Noble, S., ... Wong, E. (2004). Therapist variables. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 226–306). New York, NY: Wiley.
- Black, S., Hardy, G., Turpin, G., & Parry, G. (2005). Self-reported attachment styles and therapeutic orientation of therapists and their relationship with reported general alliance quality and problems in therapy. *Psychology and Psychotherapy: Theory, Research and Practice, 78*(3), 363–377. doi:10.1348/147608305X43784
- Buckman, J. R., & Barker, C. (2010). Therapeutic orientation preferences in trainee clinical psychologists: Personality or training? *Psychotherapy Research, 20*(3), 247–258. doi:10.1080/10503300903352693
- Degnan, A., Seymour-Hyde, A., Harris, A., & Berry, K. (2016). The role of therapist attachment in alliance and outcome: A systematic literature review. *Clinical Psychology & Psychotherapy, 23*(1), 47–65. doi:10.1002/cpp.1937
- Derogatis, L. R. (1975). *Brief symptom inventory*. Baltimore, MD: Clinical Psychometric Research.
- Elliott, R., Bohart, A. C., Watson, J. C., & Greenberg, L. S. (2011). Empathy. *Psychotherapy, 48*(1), 43–49. doi:10.1037/a0022187
- Franke, G. (2000). *BSI. Brief Symptom Inventory—Deutsche Version* [BSI. Brief Symptom Inventory: German Version]. Manual. Göttingen: Beltz.
- Fydrich, T., Renneberg, B., Schmitz, B., & Wittchen, H.-U. (1997). *Strukturiertes Klinisches Interview für DSM-IV, Achse II: Persönlichkeitsstörungen* [Structured Clinical Interview for DSM-IV-TR Axis II Disorders: Personality Disorders]. Göttingen: Hogrefe.
- Gelso, C., & Hayes, J. (2007). *Countertransference and the inner world of the psychotherapist: Perils and possibilities*. Mahwah, NJ: Erlbaum.
- Greenhalgh, J., Long, A. F., & Flynn, R. (2005). The use of patient reported outcome measures in routine clinical practice: Lack of impact or lack of theory? *Social Science & Medicine, 60*(4), 833–843. doi:10.1016/j.socscimed.2004.06.022
- Hatcher, R. L. (2015). Interpersonal competencies: Responsiveness, technique, and training in psychotherapy. *American Psychologist, 70*(8), 747–757. doi:10.1037/a0039803
- Hayes, J. A., & Gelso, C. J. (2001). Clinical implications of research on countertransference: Science informing practice. *Journal of Clinical Psychology, 57*(8), 1041–1051. doi:10.1002/jclp.1072
- Hayes, J. A., Gelso, C. J., & Hummel, A. M. (2011). Managing countertransference. *Psychotherapy, 48*(1), 88–97. doi:10.1037/a0022182
- Hayes, S. C., Wilson, K. G., Gifford, E. V., Follette, V. M., & Strosahl, K. (1996). Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting and Clinical Psychology, 64*(6), 1152. doi:10.1037/0022-006X.64.6.1152
- Heck, R. H., Thomas, S. L., & Tabata, L. N. (2014). *Multilevel and longitudinal modeling with IBM SPSS*. New York, NY: Routledge.
- Heinonen, E., Knekt, P., Jääskeläinen, T., & Lindfors, O. (2014). Therapists' professional and personal characteristics as predictors of outcome in long-term psychodynamic psychotherapy and psychoanalysis. *European Psychiatry, 29*(5), 265–274. doi:10.1016/j.eurpsy.2013.07.002
- Heinonen, E., Lindfors, O., Härkänen, T., Virtala, E., Jääskeläinen, T., & Knekt, P. (2014). Therapists' professional and personal characteristics as predictors of working alliance in short-term and long-term psychotherapies. *Clinical*

- Psychology & Psychotherapy*, 21(6), 475–494. doi:10.1002/cpp.1852
- Heinonen, E., Lindfors, O., Laaksonen, M. A., & Knekt, P. (2012). Therapists' professional and personal characteristics as predictors of outcome in short- and long-term psychotherapy. *Journal of Affective Disorders*, 138(3), 301–312. doi:10.1016/j.jad.2012.01.023
- Heinonen, E., & Orlinsky, D. E. (2013). Psychotherapists' personal identities, theoretical orientations, and professional relationships: Elective affinity and role adjustment as modes of congruence. *Psychotherapy Research*, 23(6), 718–731. doi:10.1080/10503307.2013.814926
- Horowitz, L. M., Alden, L. E., Kordy, H., & Strauß, B. (2000). *Inventar zur Erfassung interpersonaler Probleme: Deutsche Version. Manual (2., überarbeitete und neu normierte Auflage)* [Inventory of interpersonal problems (IIP-D): German Version. Manual (2nd edition)]. Göttingen: Beltz-Test GmbH.
- Horowitz, L. M., Rosenberg, S. 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. doi:10.1037/0022-006X.56.6.885
- Mander, J., Schlarb, A., Teufel, M., Keller, F., Hautzinger, M., Zipfel, S., ... Sammet, I. (2015). The individual therapy process questionnaire: Development and validation of a revised measure to evaluate general change mechanisms in psychotherapy. *Clinical Psychology & Psychotherapy*, 22(4), 328–345. doi:10.1002/cpp.1892
- Mander, J. V., Jacob, G. A., Götz, L., Sammet, I., Zipfel, S., & Teufel, M. (2015). Associations between Grawe's general mechanisms of change and Young's early maladaptive schemas in psychotherapy research: A comparative study of change processes. *Psychotherapy Research*, 25(2), 249–262. doi:10.1080/10503307.2014.889330
- Mander, J. V., Wittorf, A., Schlarb, A., Hautzinger, M., Zipfel, S., & Sammet, I. (2013). Change mechanisms in psychotherapy: Multiperspective assessment and relation to outcome. *Psychotherapy Research*, 23(1), 105–116. doi:10.1080/10503307.2012.744111
- McLennan, J., & Poznanski, J. (2004). Theoretical orientation and the person of the therapist: An Australian study. *Psychotherapy in Australia*, 10(3), 62–66.
- Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology*, 55, 1–22. doi:10.1146/annurev.psych.55.042902.130709
- Nissen-Lie, H. A. (2001). *The relationship between therapists' experiences of difficulties in practice and their interpersonal functioning: An empirical investigation*. Unpublished doctoral dissertation. University of Oslo, Department of Psychology.
- Nissen-Lie, H. A., Goldberg, S. B., Hoyt, W. T., Falkenström, F., Holmqvist, R., Nielsen, S. L., & Wampold, B. E. (2016). Are therapists uniformly effective across patient outcome domains? A study on therapist effectiveness in two different treatment contexts. *Journal of Counseling Psychology*. doi:10.1037/cou0000151
- Nissen-Lie, H. A., Havik, O. E., Høglend, P. A., Rønnestad, M. H., & Monsen, J. T. (2015). Patient and therapist perspectives on alliance development: Therapists' practice experiences as predictors. *Clinical Psychology & Psychotherapy*, 22(4), 317–327. doi:10.1002/cpp.1891
- Nissen-Lie, H. A., Monsen, J. T., & Rønnestad, M. H. (2010). Therapist predictors of early patient-rated working alliance: A multilevel approach. *Psychotherapy Research*, 20(6), 627–646.
- Nissen-Lie, H. A., Monsen, J. T., Ulleberg, P., & Rønnestad, M. H. (2013). Psychotherapists' self-reports of their interpersonal functioning and difficulties in practice as predictors of patient outcome. *Psychotherapy Research*, 23(1), 86–104. doi:10.1080/10503307.2012.735775
- Nissen-Lie, H. A., Rønnestad, M. H., Høglend, P. A., Havik, O. E., Solbakken, O. A., Stiles, T. C., & Monsen, J. T. (2015). Love yourself as a person, doubt yourself as a therapist? *Clinical Psychology & Psychotherapy*. doi:10.1002/cpp.1977
- Odyniec, P., Victor, P. P., Berner, A., & Willutzki, U. (2016). Difficulties in psychotherapeutic practice. *Psychotherapeut*, 61(3), 216–221. doi:10.1007/s00278-016-0105-4
- Ogles, B. M., Lambert, M. J., & Masters, K. S. (1996). *Assessing outcome in clinical practice*. Boston: Allyn & Bacon.
- Orlinsky, D. E., Ambühl, H., Rønnestad, M., Davis, J., Gerin, P., Davis, M., ... Wiseman, H. (1999). Development of psychotherapists: Concepts, questions, and methods of a collaborative international study. *Psychotherapy Research*, 9(2), 127–153. doi:10.1080/10503309912331332651
- Orlinsky, D. E., & Rønnestad, M. H. (2005). *How psychotherapists develop: A study of therapeutic work and professional growth*. Washington, DC: American Psychological Association.
- Poznanski, J. J., & McLennan, J. (2003). Becoming a psychologist with a particular theoretical orientation to counselling practice. *Australian Psychologist*, 38(3), 223–226. doi:10.1080/00050060310001707247
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Rønnestad, M. H., & Skovholt, T. M. (2012). *The developing practitioner: Growth and stagnation of therapists and counselors*. New York: Routledge.
- Rosa-Alcázar, A. I., Sánchez-Meca, J., Gómez-Conesa, A., & Marín-Martínez, F. (2008). Psychological treatment of obsessive-compulsive disorder: A meta-analysis. *Clinical Psychology Review*, 28(8), 1310–1325. doi:10.1016/j.cpr.2008.07.001
- Sanness, K. J. (2012). *Predictors and moderators of difficulties and coping for trauma therapists* (Doctoral Dissertation). The George Washington University, Washington.
- Scherr, S. R., Herbert, J. D., & Forman, E. M. (2015). The role of therapist experiential avoidance in predicting therapist preference for exposure treatment for OCD. *Journal of Contextual Behavioral Science*, 4(1), 21–29. doi:10.1016/j.jcbs.2014.12.002
- Schiefele, A.-K., Lutz, W., Barkham, M., Rubel, J., Böhnke, J., Delgado, J., ... Nielsen, S. L. (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*, 1–16. doi:10.1007/s10488-016-0736-3
- Schröder, T., Wiseman, H., & Orlinsky, D. (2009). "You were always on my mind": Therapists' intersession experiences in relation to their therapeutic practice, professional characteristics, and quality of life. *Psychotherapy Research*, 19(1), 42–53. doi:10.1080/10503300802326053
- Schulte, D., & Eifert, G. H. (2002). What to do when manuals fail? The dual model of psychotherapy. *Clinical Psychology: Science and Practice*, 9(3), 312–328. doi:10.1093/clipsy.9.3.312
- Schumacher, S., Gaudlitz, K., Plag, J., Miller, R., Kirschbaum, C., Fehm, L., ... Ströhle, A. (2014). Who is stressed? A pilot study of salivary cortisol and alpha-amylase concentrations in agoraphobic patients and their novice therapists undergoing in vivo exposure. *Psychoneuroendocrinology*, 49, 280–289. doi:10.1016/j.psyneuen.2014.07.016
- Scragg, P., Bor, R., & Watts, M. (1999). The influence of personality and theoretical models on applicants to a counselling psychology course: A preliminary study. *Counselling Psychology Quarterly*, 12(3), 263–270. doi:10.1080/09515079908254096
- Skovholt, T. M., & Jennings, L. (2004). *Master therapist: Exploring expertise in therapy and counseling*. Boston, MA: Allyn & Bacon.

- Stiles, W. B. (2013). The variables problem and progress in psychotherapy research. *Psychotherapy, 50*(1), 33–41. doi:10.1037/a0030569
- Stoltenberg, C. D., McNeill, B. W., & Delworth, U. (1998). *IDM supervision: An integrative developmental model for supervising counselors and therapists*. San Francisco, CA: Jossey-Bass.
- Tasca, G. A., & Gallop, R. (2009). Multilevel modeling of longitudinal data for psychotherapy researchers: I. The basics. *Psychotherapy Research, 19*(4–5), 429–437. doi:10.1080/10503300802641444
- Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry, 14*(3), 270–277. doi:10.1002/wps.20238
- Wampold, B. E., Hollon, S. D., & Hill, C. E. (2011). Unresolved questions and future directions in psychotherapy research. In J. C. Norcross, G. R. VandenBos, & D. K. Freedheim (Eds.), *History of psychotherapy: Continuity and change* (2nd ed., pp. 333–356). Washington, DC: American Psychological Association.
- Willutzki, U., Hernandez Bark, G., Davis, J., & Orlinsky, D. E. (1997, June). *Client outcomes as a function of the therapist's difficulties and coping strategies in the course of psychotherapy: Initial results*. Paper presented at the 28th Annual International Conference of the Society for Psychotherapy Research, Geilo, Norway.
- Wittchen, H.-U., Wunderlich, U., Gruschwitz, S., & Zaudig, M. (1997). *SKID I. Strukturiertes Klinisches Interview für DSM-IV. Achse I: Psychische Störungen* [Structured Clinical Interview for DSM-IV-TR Axis I Disorders: Psychological Disorders]. Göttingen: Beltz-Test.
- Yonatan-Leus, R., Tishby, O., Shefler, G., & Wiseman, H. (2017). Therapists' honesty, humor styles, playfulness, and creativity as outcome predictors: A retrospective study of the therapist effect. *Psychotherapy Research, 1*–10. doi:10.1080/10503307.2017.1292067