Psychotherapy Research

Accuracy of therapist perceptions of patients' alliance: Exploring the divergence

Armin Hartmann[a], Andreas Joos[a], David Elliot Orlinsky[b] & Almut Zeeck[a]

[a] Department of Psychosomatic Medicine and Psychotherapy, Center for Mental Disorders, Medical Center - University of Freiburg, Germany

[b] Department of Comparative Human Development, University of Chicago, Chicago, IL, USA
Published online: 07 Jul 2014.

CrossMark

Click for updates

PLEASE SCROLL DOWN FOR ARTICLE

EMPIRICAL PAPER

Accuracy of therapist perceptions of patients' alliance: Exploring the divergence

ARMIN HARTMANN[1], ANDREAS JOOS[1], DAVID ELLIOT ORLINSKY[2], & ALMUT ZEECK[1]

[1]Department of Psychosomatic Medicine and Psychotherapy, Center for Mental Disorders, Medical Center – University of Freiburg, Germany & [2]Department of Comparative Human Development, University of Chicago, Chicago, IL, USA

Abstract
Objective: The therapeutic alliance is a well-established predictor of psychotherapy outcome, yet much research has shown that therapists' and patients' views of the alliance can diverge substantially. Therapists systematically underestimate their patients' perceived level of alliance, and the correlation between therapist and patient estimates of patient alliance is only moderate. The present study explored the divergence between therapists' and patients' perspectives on patients' alliance experience, and its relations to therapists' concurrent work involvement and session process experiences. **Method:** Data from 98 patients treated by 26 therapists with psychodynamic psychotherapy were analyzed. **Results:** Therapist-patient divergence was significantly related to therapists' case-wise work involvement, but not to therapist's views of session process. The best predictor of therapist-patient divergence was therapists experiencing a "distressed practice" work involvement pattern. **Conclusion:** Although therapists' work involvement experiences are not commonly investigated, they can be a relevant predictor of therapy processes.

Keywords: alliance; psychoanalytic/psychodynamic therapy; process research

Assessment of the therapeutic alliance, and its development over time clearly have been central topics of psychotherapy research, and an extensive body of process-outcome research clearly supports the view that the patient's working alliance is a reliable predictor of treatment outcome (Flückiger, Del Re, Wampold, Symonds, & Horvath, 2012). However, studies have also found only moderate correspondence at best between therapists' views of their patient's alliance experience and the patients' own reports of their experience (Tryon, Blackwell, & Hammel, 2007). In their meta-analysis of 53 studies, Tryon et al. (2007) found that therapists significantly underestimate their patients' alliance ratings, with a standardized effect size of $d = 0.63$. The discrepancies were larger in shorter therapies, and in treatments of mild and moderate patient pathology (vs. severe pathology). Overall, the correlation between patients' and therapists' views was moderate ($r = .36$), with a shared variance of only 13%. Naturally,

considering their different roles, therapists and patients might be expected to have somewhat different perspectives on their therapeutic alliance (Bachelor, 2013), yet nevertheless it is clinically important that therapists have an accurate awareness of their patients' alliance experience, since the patient's alliance experience predicts success or failure (Horvath, Del Re, Flückiger, & Symonds, 2011; Orlinsky & Howard, 1975). As Krause and Lutz (2009) point out, therapists are responsible for managing the process of therapy, including knowing when problematic processes such as alliance ruptures have occurred.

Additionally, the differences between patient and therapist views of patient alliance have shown inconsistent relations to outcome. Some studies have reported no association (Fitzpatrick, Iwakabe, & Stalikas, 2005) while others have found mixed associations to session process and overall treatment outcome (Marmarosh & Kivlighan, 2012). Using

different alliance measures and comparing clients' and therapists' views, Bachelor (2013) did find that some differences in single components of the alliance were related to outcome.

Some of these inconsistencies may be due to differences in the way that alliance has been measured. When evaluating therapists' ratings of the alliance, it is important to differentiate between (i) the therapist's own view of the alliance and (ii) the therapist's assessment of how the patient's alliance experience. One of the most often used measures, the *Helping Alliance Questionnaire* (HAQ; Alexander & Luborsky, 1986), has a therapist version, which asks therapists to evaluate how they think their patients view the alliance. By contrast, another commonly used measure, the *Working Alliance Inventory* (WAI; Horvath & Greenberg, 1989), asks therapists to evaluate the "ways you think or feel about your client," and therefore assesses the therapist's own perspective.[1]

However, these issues do not plausibly account for all of the significant divergence that has been observed between patient and therapist perspectives, and the fact of that divergence ought not to be discounted as an inconvenient methodological problem. Because patient alliance is an important predictor of therapeutic progress, and because therapists are clinically responsible for ensuring the best attainable outcome for patients, the divergence in therapists' and patients' views of the patient's alliance should be recognized as an independent variable worthy of study, especially with respect to how divergences are related to the therapeutic process and the work experience of therapists. As a step in this direction, the present study addressed these research questions: (i) How well do therapists know the current state of their patient's alliance experience? Specifically, is there a substantial and systematic level difference between patients' and therapists' perspectives? (ii) Do therapists differ with respect to how accurately they perceive their patients' views? (iii) What proportions of the variance in divergence between patients' and therapists' ratings of the patient's bond experience are attributable to therapist and patient characteristics (including work experience of therapists)?

We hypothesize two important sources of alliance development and perception: Therapist's experience of therapy and their processes of working through of these experiences. Exploration of these questions will focus on the extent that divergences in therapists' perceptions of their patients' alliance ratings are related (i) to therapists' experiences during therapy sessions (e.g., Orlinsky & Howard, 1977), (ii) to therapists' inter-session process experiences—their memories, thoughts, and feelings about therapy and

the patient during time between sessions (e.g., Hartmann, Orlinsky & Zeeck, 2011; Orlinsky & Geller, 1993), and (iii) to the overall quality of their "work involvement" reflected in how much the therapist experiences therapeutic work with the patient as a "healing involvement" and how much as a "stressful involvement" (Orlinsky & Rønnestad, 2005).

Before engaging these questions, it is important to clarify the definition of alliance to be used in this study. The concept of the therapeutic relationship has become increasingly complex over time. Bordin's (1979) therapeutic alliance concept comprised goal consensus, task consensus, and the interpersonal bond. This drew upon previous concepts (e.g., Freud's (1912) "transference" and Rogers's (1957) "therapist-offered conditions"), and complements later contributions (e.g., Gelso's (2009) "real relationship," and studies of patient expectations about treatment). In the "Generic Model of Psychotherapy" (Orlinsky, 2010; Orlinsky & Howard, 1986; Orlinsky, Rønnestad & Willutzki, 2004) the concept of "therapeutic bond"—similar to Bordin's view of the bond—refers specifically to relations between patient and therapist as *persons* (as distinct from the "therapeutic contract" which organizes their reciprocal *roles* as patient and therapist, and corresponds to Bordin's goal and task concepts). This has both communicative aspects (such as expressive and empathic rapport) and affective aspects (like mutual respect, trust and liking), which are usually seen by others as the core of the alliance (e.g., Crits-Christoph, Johnson, Connolly Gibbons, & Gallop, 2013). The present study focuses on therapists' and patients' view of the therapeutic *bond* rather than other facets of the therapeutic alliance, because it is the central and common construct in almost all above mentioned theoretical concepts of alliance.

Methods

Data

Data for this study were drawn from a project on psychotherapy process (Zeeck et al., 2002) with a primary focus on therapists' experiences of "work involvement" (Orlinsky & Rønnestad, 2005). As part of the study, patients' and therapists' views of the therapeutic bond were routinely monitored over the course of treatment.

Participants

The study sample includes 98 treatment cases, conducted by 26 psychodynamic psychotherapists of varying experience levels. All patients and therapists gave informed consent and the study was approved by

Table I. Patient sample.

	Outpatient (*n* = 49)		Day clinic (*n* = 49)	
	% (*n*)	*m* (SD)	% (*n*)	*m* (SD)
Age		40.7 (12.8)***		49.2 (9.8)***
% Female	72.3% (35)[a]		53% (26)	
SCL-90R: GSI		0.7 (0.5)*		1.2 (0.7)*
Main diagnosis				
Depression	81.6% (40)		44.9% (22)	
Anxiety disorder	2.0% (1)		12.2% (6)	
Somatoform disorder	10.2% (5)		12.2% (6)	
Adjustment disorder	2.0% (1)		10.2% (5)	
Eating disorders	4.1% (2)		4.1% (2)	
Other	0% (0)		16.2% (8)	
Treated by novices	34.7% (17)		65.3% (35)	
Treated by experienced	65.3% (35)		34.7% (17)	

Note. Significant differences between samples: **p < .05; ***p < .001.
[a]One missing value.

the local ethics committee. Half of the sample consisted of individual outpatient treatments (private practice) and the other half were individual treatments in a day clinic setting (university hospital).

Most of the 98 patients suffered from a major depressive disorder. In the day clinic setting, ICD-10 diagnoses were made by psychologists or psychiatrists based on an intake interview of one hour. In the outpatient setting, therapists documented ICD-10 diagnoses as part of the German approval procedure of the health insurance system. The day clinic patients showed a broader range of main diagnoses, a higher mean age, and a higher symptom severity (see Table I). Patients with a diagnosis of psychosis, substance dependency, bipolar disorder, antisocial personality disorder, dementia or cognitive impairment were excluded.

All therapists (*N* = 26) had a psychodynamic orientation. Approximately one half of the therapist sample consisted of therapists in training at an independent psychoanalytic training institute, who were at an advanced stage of their education and were getting supervision after every fourth session (in Germany, training as a psychotherapist is not part of

Table II. Therapist sample.

	Outpatient (*n* = 19)		Day clinic (*n* = 10)	
	% (*n*)	*m* (SD)	% (*n*)	*m* (SD)
In training	37% (7)		60% (6)	
Training completed	63% (12)		40% (4)	
% Female	84% (16)*		40% (4)*	
Years of experience		12.8 (7.7)***		4.8 (6.0)***

Note. Significant differences between samples: **p < .05; ***p < .001; three therapists treated patients in both settings.

the graduate university curriculum). Therapist characteristics are summarized in Table II.

Treatments

The day clinic sample consisted of 49 consecutive treatment cases conducted over a period of 9 months. The day clinic is part of the Clinic of Psychosomatic Medicine and Psychotherapy of the Freiburg University Clinic. Twice-weekly individual sessions of 45 minutes length were combined with a psychodynamic group (2/week), art therapy (2/week), body therapy (2/week), sessions with the nursing staff, physicians rounds, and a relaxation group (see Zeeck et al. (2002) for a detailed description of the program). Patients' ratings of the day clinic program components showed that individual psychotherapy sessions were valued as the most important treatment element, and were a good indicator of the quality of the overall process (Zeeck, Hartmann, & Orlinsky, 2004). All process and alliance measures refer to the individual treatment sessions in the day clinic setting. The day clinic sample excluded the starting phase of treatment (first three sessions equivalent to the first two weeks) where the processes of building the alliance predominates, and the closing phase of treatment (last four sessions equivalent to the last two weeks before discharge) where impending separation may influence the values of process and alliance scores (Hartmann, Orlinsky, Weber, Sandholz, & Zeeck, 2010). No other measures could be taken to parallelize the treatment processes of the two settings.

The outpatient sample consisted of sessions from the middle phase of another 49 treatments. Cases having more than three sessions per week were excluded to make the two samples comparable. Each therapist contributed 10 consecutive sessions

from two to four patients who agreed to participate in the study.

Measures

Helping Alliance Questionnaire. The therapeutic alliance was measured using the German version of the HAQ (Alexander & Luborsky, 1986; Bassler, Potratz, & Krauthauser, 1995; Luborsky, Barber, & Crits-Christoph, 1990). The HAQ has a patient and a therapist version and comprises 12 items reflecting two factors: (i) the perceived quality of the patient-therapist relationship and (ii) the patient's and therapist's respective expectations of a good outcome (for the patient's view also called "internal change"). As the latter is confounded with treatment outcome (Bassler et al., 1995; Hendriksen et al., 2010), only the "quality of the therapeutic relationship" (HAQ-Rel) factor was used in this study. It should be noted that only the patients rated their self-experienced alliance. Therapists were asked to take the patient's perspective and rated their view of the patient's alliance with the same questionnaire.

Session Evaluation Questionnaire. The *Session Evaluation Questionnaire* (SEQ; Stiles et al., 1994) measures the process and the impact of psychotherapy sessions with four scales, of 5 items each (Stiles, Gordon, & Lani, 2002): two factors of in-session process (*session depth* and *session flow*) and two of post-session impact (*positivity* and *arousal*). The present study used only three factors: session depth, session flow, and post-session positivity, because the scale for *arousal* showed insufficient reliability in European samples (Hafkenscheid, 2009; Hartmann et al., 2013). The SEQ is used to measure the therapist immediate session experience.

Therapist Work Involvement Scales. The *Therapist Work Involvement Scales* (TWIS) comprise a selection of 52 items from the *Development of Psychotherapist Common Core Questionnaire* (DPCCQ; Orlinsky et al., 1999) that were empirically found to describe the professional work experience of psychotherapists (Orlinsky & Rønnestad, 2005). The two orthogonal second-order factors show therapists' work experience to consist variable degrees of *Healing Involvement* (HI) and *Stressful Involvement* (SI), and these scales have demonstrated good convergent and discriminant validity and yielded satisfactory internal consistencies (Cronbach's alpha, α_{HI} = .74; α_{SI} = .66). (A detailed description of the TWIS scales is given by Nissen-Lie, Monsen, and Rønnestad (2010)). In the present sample the reliabilities were even better (α_{HI} = .82; α_{SI} = .93) (Hartmann et al., 2011). HI is defined by scores for

basic relational skills (e.g., empathy in the particular therapy), positive in-session feelings (e.g., "flow"), relational agency (e.g., feeling efficacious), and constructive coping when difficulties arise. SI is defined as frequent difficulties (e.g., insecurity about how to deal with the patients problems), negative in-session feelings (e.g., boredom, anxiety) and avoidant coping with difficulties (e.g., showing frustration to the patient). Orlinsky and Rønnestad (2005) and Nissen-Lie et al. (2010) provide further details. Hartmann, Schröder et al. (2010) translated the TWIS into German and adapted it for use as a case-focused repeated measurement (Hartmann et al., 2011).

Orlinsky and Rønnestad (2005, p. 82f.) defined cut-offs for low and high ranges of HI and SI for a sample of 3629 therapists. The resulting four combinations describe classes of practice patterns (see Figure 1). *Effective Practice* characterized therapists who experienced "much" HI and relatively "little" SI. Therapists who experienced "much" HI, but also "more than a little" SI were described as having a *Challenging Practice*, in which they appear to be positively engaged in their work but also clearly encountered difficulty with some patients. These two patterns accounted for 73% of the sample reported by Orlinsky and Rønnestad (2005). A third pattern defined by "little" SI, but also "not much" HI was called *Disengaged Practice* (17%), and a fourth pattern called *Distressing Practice* defined by more than a "little" SI and "not much" HI (10%).

In the present study, the TWIS were used to assess overall practice patterns by measuring therapist levels on HI and SI, and were administered casewise assuming that there is a substantial component of patient variance (patient level).
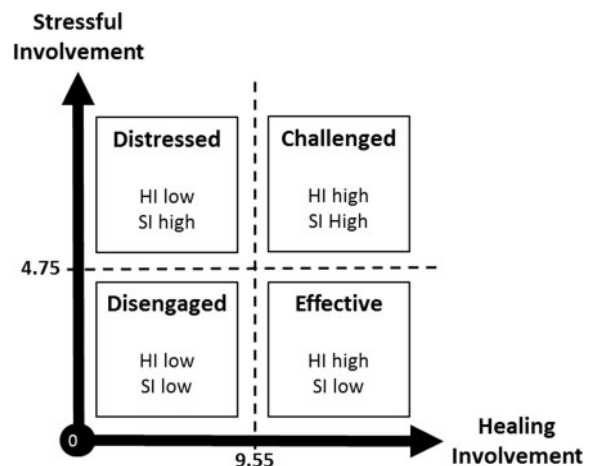

Figure 1. Practice patterns.
Note: Therapist practice patterns according to their Healing and Stressful Involvement, after Orlinsky and Rønnestadt (2005, p. 82f.).

Intersession Experience Questionnaire. The German version of the *Intersession Experience Questionnaire* (IEQ; Orlinsky & Geller, 1993; Hartmann, Orlinsky, Geller, & Zeeck, 2003; Hartmann et al., 2011) was used to assess how treatment is "processed" or internalized by patients and therapists between sessions. The IEQ contains 52 items assessing 8 dimensions of intersession experience: intensity (*frequency* and *duration*) of thoughts and feelings about the therapist or patient and the treatment; the contexts in which those thoughts and feelings arise (*relaxed* or *affectively loaded*); their content (*recreating the therapeutic dialogue, relationship fantasies, application of therapy*); and their emotional tone (*positive* and *negative emotions*).

Measurement Schedule, Data Management and Operationalization

All outpatient sample measures (10 sessions) were taken from middle phases of the therapies with various starting points, typically starting at session 34 (median). All day clinic treatment sample measures covered the middle working phase of the treatments, with up to 12 points of measurement. Session and intersession processes were measured continuously for the whole time span of observation. Bond and work involvement were measured every third session (see Table III).

As there was no simple match of session process (measured session by session) and alliance/work involvement (every third session), the means of the scores were computed for all remaining observations of each variable, omitting session-by-session variability and using values to represent the average of each process variable rated by patient or therapist (for example, the mean *session depth* of the selected sessions). The divergence of the therapist's from the patient's view of the therapeutic bond was operationalized as the difference (patient *minus*

therapist) between the aggregated HAQ-Rel scores, such that a positive difference indicates a higher patient score (i.e., an underestimation of the bond by the therapist). This difference score is the dependent variable for all subsequent analyses.

Data Analyses

Means, standard deviations and frequencies were computed for descriptive purposes. A linear regression and a scatter plot of patient alliance scores and therapists' perceived patient alliance were used to describe their relation. The main analysis of the divergence makes use of difference scores between therapist and patient ($\delta = P - T$; divergence of view—abbreviated as "DoV" in formulas and mixed modes).

The statistical analyses were performed with SAS-JMP V8 (SAS Institute Inc., 2009) and SPSS V20. To analyze possible predictors of the divergence, regressions were computed and in case of nested data (patients nested within therapists) this approach was complemented by mixed models (SPSS). The respective formulas are summarized in Table IV.

For advanced exploration of nonlinear relations between variables we used response surface models (SAS-JMP) as suggested by Marmarosh and Kivlighan (2012).[2] The following level 1 covariates were entered into the model: session process was represented by the SEQ-factors *positivity*, *flow*, and *depth*; inter-session process was represented by all available factors of the IEQ; scores for *Healing Involvement* and *Stressful Involvement* were used from the Therapist Work Involvement Scales.

Results

Divergence of Views

Overall, therapists clearly underestimated their patients' ratings of the therapeutic alliance. The

Table III. Measurement schedule.

Instruments	Setting/source	Sessions[a]													
	IO	n	+1	+2	+3	+4	+5	+6	+7	+8	+9	–	–	–	–
	DC	1	2	3	4	5	6	7	8	9	10	11	12	13	14+
SEQ, IEQ	P&T	√	√	√	√	√	√	√	√	√	√	(√)	(√)	(√)	(√)
HAQ	P&T	√			√			√			√			(√)	
TWIS	T only	√			√			√			√			(√)	
Selection/aggregation		–	–	–	DC	DC	DC	DC	DC	DC	DC	DC	DC	DC	(–)
		IO	IO	IO	IO	IO	IO	IO	IO	IO	IO	–	–	–	–

SEQ = Session Evaluation Questionnaire, filled out after the session; IEQ = Intersession Experience Questionnaire, immediately before session; HAQ = Helping Alliance Questionnaire; P&T = parallel forms for patients and therapists; TWIS = Therapist Work Involvement Scales, after every third session; DC = day clinic; IO = individual outpatient; Selection = measurement selected for aggregation for DC or IO treatments. Regular measurement schedule = √, measured only if day treatment duration allowed for (√).
[a]Measurement starts at session number "*n*" for outpatients (median no. of starting session = 34) and ends after 10 sessions, for day clinic patients is starts with session no. 1 continues with the same schedule until discharge.

Table IV. Formulas.

Formulas	Meaning of formulas
(1) $Y_{ij} = \beta_{0j} + \varepsilon_{ij}$	DoV in a dyad (patient i and therapist j; Y_{ij}) = intercept of DoV by therapist j (β_{0j}) + variation within therapists (ε_{ij})
(2) $\beta_{0j} = \gamma_{00} + u_{0j}$	Variation in intercepts of DoV by therapist j (β_{0j}) = level 2 fixed effect (γ_{00}) + between therapist error (u_{ij})
(3) $Y_{ij} = \gamma_{00} + u_{0j} + \varepsilon_{ij}$	Substitution of β_{0j} with Formula (2) in Formula (1)
(4) $Y_{ij} = \beta_{0j} + \beta_1(L1C_1) + \beta_2(L1C_2) + \dots + \beta_m(L1C_m) + \varepsilon_{ij}$	Extension of Formula (1) with up to m level 1 covariates ($L1C_i$)
(5) $Y_{ij} = \beta_{0j} + \beta_1(HI) + \beta_2(SI) + \varepsilon_{ij}$	Level-1 covariates in final model: healing (HI) and stressful involvement (SI) as predictors

DoV = Divergence of patient's and therapist's view of patient's alliance.

therapists' estimates of patients' bond scores was $M = 3.70$ ($SD = 0.55$), while patient's (presumably "true") mean was $M = 4.06$ ($SD = 0.58$). The difference between the two distributions expressed as effect size (mean difference/pooled SD) is $ES = 0.64$. In addition, a systematic difference was found in therapist "errors" depending on the level of patients' alliance rating (see Figure 2). If there was no divergence (δ) the points are located on the diagonal (dashed line). Therapists' overestimations are located above the diagonal, and underestimations below. The plot and the regression line (LR) show that therapists systematically *underestimated high alliance* ratings by patients and *overestimated low alliance* ratings.

Variance Components

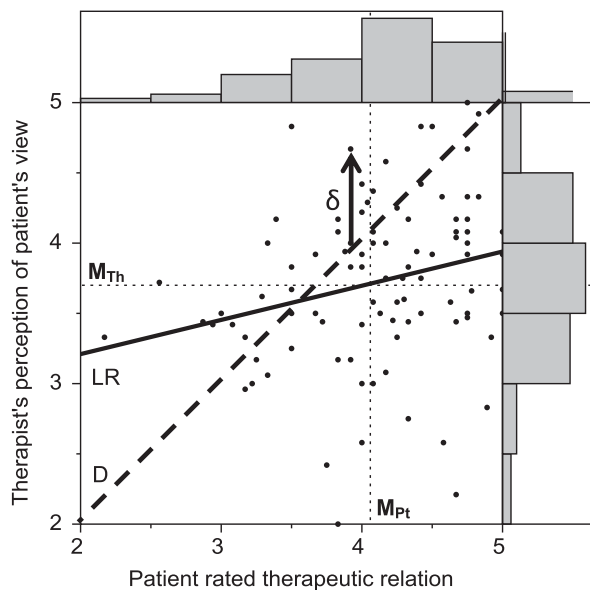The variance components (vc) in divergence scores were estimated with a mixed model containing no

Figure 2. Divergence of patients' alliance and therapists' perception thereof.
Note: M_{th} = mean of therapists' ratings; M_{pt} = mean of patients' ratings; D = diagonal; LR = linear regression; δ = deviation from diagonal.

covariates (null model) but a clear rank order of variance was observed in the variables of treatment setting, therapist, and patient. The results of the three level null model (treatment, therapist, patient) showed that the effects of treatment on divergence were indistinguishable from zero and not significant. Therefore, we present the details of the two level model only (see Table V).

The model estimates a mean divergence of 0.433 rating points in alliance scores, which is more than the result of the descriptive analysis (mean difference = 0.36). The intra class correlation coefficient indicates that 22.6% of the variance is due to therapists. Inspection of the 95% confidence interval (lower CI = 0.41; 0.41 > 0) shows that the intercept of this variance is significantly different from zero. Wald's Z test[3] yields $p < .06$ (one-sided $p < .03$). These figures show that most of the variance is due to processes in the individual patient–therapist dyad, presumably caused by patient or therapy process characteristics. Nevertheless more than one-fifth of the variance is due to therapist characteristics, so that multilevel models should be used.

Prediction of Divergences

All relations of process variables with divergence scores were first explored for bivariate relations. A mixed regression model was computed for session process, therapist inter-session process, and therapist work involvement scales, accounting for the nesting of patients (as random effects) within therapists. The results of the regression are shown in Table VI.

Of the therapist's SEQ scores, only in-session *depth* reached the defined level of significance. None of the variables describing therapist inter-session experience were significantly related to divergence. Therapist experiences of *Healing Involvement* and *Stressful Involvement* from the TWIS both reached the defined level of significance. A comprehensive model with session *depth* (SEQ), and HI and SI (TWIS), showed a non-significant relation for *depth* but significant associations for HI and SI

Table V. Variance components of two level mixed model (null model).

	Estimate of fixed effects					CI 95%	
Parameter	Estimate	SE	df	t	p	Lower	Upper
Intercept Y_{00}	.433	.010	25.464	4.356	<.000	.229	.638

	Estimates of covariance parameters				CI 95%	
Parameter	Estimate	SE	Wald's Z	p	Lower	Upper
Residual	.405	.065	6.22	<.000	.295	.555
Intercept (subject = therapist) variance	.118	.064	1.84	<.066	.041	.344

DoV = dependent variable. ICC = $\sigma^2_{\text{between}}/(\sigma^2_{\text{between}} + \sigma^2_{\text{within}}) = \frac{.118}{(.118+.405)} = .226$

scales. (This is not surprising given the strong correlation of depth with the other two variables [$r_{\text{depth-HI}} = .44$; $r_{\text{depth-SI}} = −.38$]). We concluded that best model includes only the HI and SI work involvement scales. The estimates of this final model are shown in Table VII.

The relationship between the divergence in HAQ ratings and the TWIS variables was further explored with response surface model. Taking into account the setting (day clinic vs. outpatient treatment) allowed for a satisfactory fit (see Table VIII).

The response surface can be visualized with two and three dimensional plots as follows (see Figure 3). The non-linear character of the relationship becomes evident in the three-dimensional plot. There was no effect on the divergence from low to middle Stressful Involvement (SI), but thereafter (values > 4) the divergence (*underestimation* of patient alliance level) increased steeply. By contrast, no marked nonlinear effect was observed with respect to therapists' experiences Healing Involvement (HI), only a continuously increasing *overestimation* of patients' alliance levels

by therapists with increasing HI. Thus, therapists who reported experiencing high levels of *Stressful Involvement* tended to underestimate their patients' alliance, seeing at worse than it was; whereas the more therapists reported experiencing Healing Involvement, the more they tended to over-estimate the patient's alliance experience, viewing it as better than it was. The "folded surface" is the result of the combination of both relations.

In the two-dimensional plot, the same relation is shown by the contour lines. The horizontal and vertical reference lines show the medians of the work involvement variables (SI and HI) and divide the plane into four quadrants corresponding to the four "practice patterns" defined by Orlinsky and Rønnestad (2005). Together with the density curves, the directing arrows show that each practice pattern comes with a specific divergence level. In the *Distressed Practice* pattern (high SI, low HI) therapists show a clear underestimation of the patients' alliance, whereas therapists experiencing an *Effective Practice* pattern (low SI, high HI) obviously

Table VI. Exploratory multivariate mixed models: estimates of fixed effects.

Process focus	Source	Numerator df	Denominator df	F	p
Session evaluation (SEQ)	Intercept	1	73.228	14.51	<.000
	SEQ in-session depth	1	90.481	4.55	<.036
	SEQ in-session flow	1	87.398	1.39	<.242
	SEQ post-session positivity	1	67.367	.25	<.617
Intersession Experience (IEQ)	Intercept	1	65.762	13.46	<.000
	IEQ A intensity	1	78.731	1.42	<.237
	IEQ C1 recreating dialogue	1	82.597	1.55	<.216
	IEQ C2 relationship fantasies	1	85.718	.20	<.655
	IEQ C3 applying therapy	1	82.320	.05	<.829
	IEQ D1 positive emotions	1	85.904	3.65	<.059
	IEQ D2 negative emotions	1	73.752	.00	<.995
Therapist work involvement (TWIS)	Intercept	1	49.069	10.35	<.002
	Healing involvement	1	53.031	10.56	<.002
	Stressful involvement	1	58.473	5.00	<.029

DoV = dependent variable.

Table VII. Estimates of final mixed model.

						CI 95%	
						Lower	Upper
			Estimates of fixed effects				
Parameter	Estimate	SE	df	T	p	Lower	Upper
Intercept Y_{00}	1.804	.561	49.069	3.216	<.002	.677	2.932
Healing involvement	−.208	.064	53.031	−3.250	<.002	−.337	−.080
Stressful involvement	.103	.046	58.473	2.236	<.029	.011	.196

			Estimates of covariance parameters			
					CI 95%	
					Lower	Upper
Parameter	Estimate	SE	Wald's Z	p	Lower	Upper
Residual	.362	.059	6.098	<.000	.263	.499
Intercept (subject = therapist)	.086	.056	1.523	<.128	.311	.311

DoV = dependent variable.

overestimated the patients' alliance. The patterns for both *Disengaged Practice* (low SI, low HI) and *Challenged Practice* (high SI, high HI) move along the zero-level-curve of divergence, meaning that therapists experiencing practice patterns in which SI and HI are balanced were at low risk of misperceiving the patients' alliance.

Discussion

We replicated the results of Tryon et al. (2007) by showing that there is a mean level difference between patient and therapist alliance ratings, which is almost identical to the *ES* = 0.63 found in their meta-analysis. Viewed overall, patients rated the alliance more positively than their therapists thought they did.

Statistical and Methodological Issues

The analyzed difference scores depend on the level of the patient's perceived alliance scores. If the patient endorses a maximal score, then a therapist's
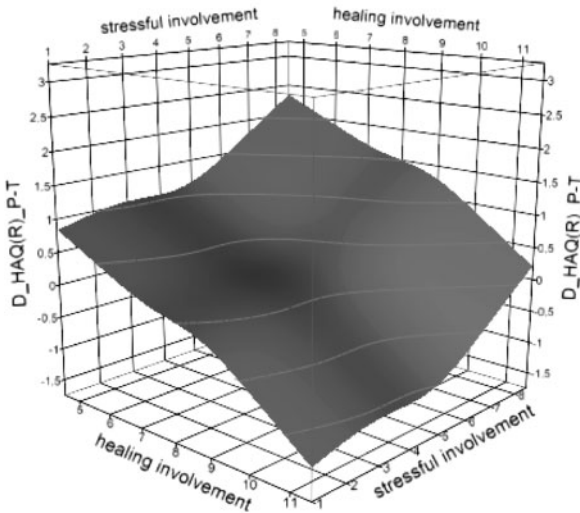
deviation can only be an under-estimation; or inversely, at the lower end of the scale, an over-estimation. Thus, a proportion of the variance of the divergences must be due to the "regression to mean" effect. Nevertheless, controlling for patients' alliance levels, therapist work involvement scores (HI and SI) explained an independent and substantial proportion of the variance. Difference scores have been criticized for unreliability and some researchers may object to using them as dependent variable. Yet, we follow recent methodological publications (e.g., Laird & Weems, 2011; Thomas & Zumbo, 2012) arguing that a general ban of difference scores is unjustified.

The mixed model analyses further identified therapist work involvement patterns (combinations of high or low HI scores with high or low SI scores) as a significant predictor of the divergence between therapists' ratings of their patients' alliance and their patients' own alliance ratings, in which the nature of the divergence was clearly associated with the level of the patient-rated quality of the alliance.
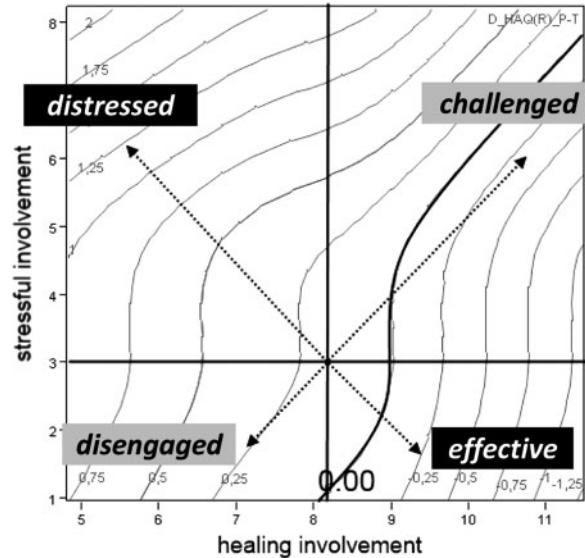
Table VIII. Fit of response surface model.

Source	df	Sum of squares	Mean squares	F	p
Model	7	13.43	1.92	4.60	.0002
Error	88	36.72	0.42		
Sum	95	50.16			

Source	N parameters	df	Sum of squares	F	p
Healing involvement & spline (4)	3	3	9.30	7.43	.0002
Stressful involvement & spline (4)	3	3	4.77	3.81	.0129
Treatment setting	1	1	3.35	8.03	.0057

Note. "& spline (4)" refers to a medium flexible, knotted spline fit, where the lambda value (number of knots the curve can "swing" around) is $\lambda = 4$. See SAS-JMP online help, handbook "Modeling and Multivariate Methods," Chapter 2, paragraph "Knotted Spline Effect."
Cave: As these models cannot account for hierarchically nested data, the *p* values must be interpreted with caution!

Figure 3. Response surface. (A) Three-dimensional projection and (B) two-dimensional top down projection.
Note: two-dimensional projection with curves indicating the height level of divergence (0.00 line bolded). Reference lines at medians of healing involvement (median = 8.2) and stressful involvement (median = 3.0). Arrows pointing to the practice patterns.

Being in a pattern of *Distressed Practice* (high SI, low HI) with a patient was accompanied by the largest divergence, namely a strong underestimation of the alliance by the therapist. Alternately, an *Effective Practice* pattern (high HI, low SI) was accompanied by (moderate) overestimation of the patient's alliance by therapists.

It is noteworthy that the other areas of process measurement (session process, session evaluation, and inter-session experience of therapists) were not predictive (except for session depth), although with relatively low statistical power in these analyses ($N =$ 26 therapists) it remains possible that subtler influences were not detected.

Additional analyses showed a bivariate, nonlinear response surface relation with therapists' experiences of the treatment case as a *Stressful Involvement* or a *Healing Involvement*. HI appeared related to the divergence in a simple linear way, while SI showed zero correlation in the low to middle range but a strong correlation in the higher range stress.

Viewed overall, the significance tests of the mixed models likely should be given first priority, because these models provide a statistically safer ground of inference. The response surface models complement the linear mixed models by offering insight into potential nonlinear relations.

Potential Clinical Implications

Although this was an initial and largely exploratory study, the findings already appear to have interesting and potentially clinically relevant if tentative implications. For example, knowing that therapists generally tend to underestimate the quality of their patients' alliance experience may provide encouragement for therapists to continue working with difficult or demanding patients. Therapists need to maintain positive work morale (Orlinsky, 2008; Orlinsky, Rønnestad, & Willutzki, 2004) if they are to convey a sense of hopefulness or "remoralization" to clients (Frank, 1993). The observed discrepancy in perceptions of the patient's alliance may reflect differences in their respective roles. Patients typically hope to get better with the help of their therapists and extend "credit" to them in advance of attaining that goal, whereas therapists may view an actual relationship in more "realistic" terms, based on the (possibly irrelevant) fact that they have treated multiple patients and so have a broader reference group in judging the alliance of any particular patient. It might be useful for therapists also to consider that their patients are likely to compare the therapist's typically attentive, empathic, and nonjudgmental attitudes (Orlinsky et al., 1996) with responses they've received from family members, friends, or colleagues.

Another factor that may lead therapists to underestimate their patients' alliance experience (at least for psychodynamically oriented therapists, as in this study) is that they may "discount" their patients' alliances as forms of transference, and may view highly positive patient alliances as reflecting an idealization process. If therapists don't accept a patient's positive alliance experience as valid, they

may not make full use of the potentials of the relationship; or, if they attribute it to a failure on their part, they may react defensively. Those therapists in particular who experience their work involvement as a *Distressing Practice* (high SI, low HI) might find some relief in knowing that their patients tend to have more positive views than they do of the alliance. Yet, this knowledge may not suffice to overcome *Distressing Practice* and we strongly recommend the considering of professional supervision and help.

By contrast, therapists whose work experience pattern is *Effective Practice* (low SI, high HI) may be cautioned against being too optimistic, since their patients tend to have less positive views of the therapeutic alliance than they do. In extreme cases, that optimism may reflect a therapist's narcissistic wish to feel like a "good therapist," leading them to neglect signs of a not-so-good relationship with the attendant risk of not recognizing and dealing with alliance ruptures (Safran & Muran, 2000; Safran, Muran, Eubanks-Carter, 2011) before negative consequences occur.

The response surface models showed that treatment setting also had a noticeable effect. Therapists working in the day clinic showed a smaller divergence (Estimate = −0.20, *SE* = .07, *t* = −2.78, *p* < .0066; see Table VIII), meaning that they underestimated the patient-rated alliance to a smaller degree—despite the fact that, on average, therapists in this setting were less experienced. Possibly working in a team with regular communication about patients in case conferences and in supervision helped these therapists to perceive the patient's alliance more accurately. An understanding of the discrepancy between therapists' and their patients' views of the patients' alliance might enable therapists to use self-observation of their work involvement in specific cases to provide cues for "off track" developments (Lambert, 2010) and thus complement other measures of treatment failure prevention.

Limitations

Although the sample comprised 98 therapies, the nesting of therapies in 26 therapists limited the statistical power of the investigation and its results. (Crits-Christoph et al. (2013) recommend 60 patients nested in each therapist for multilevel analysis!) Another limitation results from the fact that software available to compute response surface models cannot yet account for a nested data structure. Therapist and patient variance are not separable, so that significance tests of them may be misleading, and the results therefore, must be interpreted as exploratory and preliminary. Variance components obtained in a previous investigation (Zeeck et al., 2012) showed

that there is a strong variability between cases concerning HI and SI, suggesting that these variables are not trait-like but do have a substantial state-by-case component. Further investigation should explore temporal variations within cases of patterns of work involvement and alliance.

Finally, the sample comprised therapists who had only one therapeutic orientation (psychodynamic), in one region of Germany, which limits the generalizability of the results. Further investigation is needed to replicate the observed relation using samples of therapists with other theoretical orientations and in other clinical settings.

Funding

Notes

[1] The meta-analysis of Tryon et al. (2007) did not take this difference into account and integrated studies using all kinds of alliance measures and perspectives.

[2] Fitting a spline instead of a linear regression line is very common (two-dimensional data). The line of the spline can be set to be stiff or flexible. Response surface models are an extension of this method: they don't fit a "rope" on the floor but rather a "flying carpet" into the air. An inflexibly stiff fit of three-dimensional data generates a flat "shelf" floating in space and, at the other extreme, a fully flexible fit generates a "fluttering flag" in the wind of noisy data. The flexibility of the spline cannot be fitted. Its grade is chosen by the data analyst, and all significance tests thus depend on the analyst's intuition and ability to select a reasonable spline.

[3] Given that Wald's *Z* is a two-sided test but variances can never be lower than zero, it is justified to use it as a one-sided test and to divide the significance level probability by two.

References

Alexander, & Luborsky, L. (1986). The Penn Helping Alliance Scales. In L. S. Greenberg & W. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 325–366). New York, NY: Guilford Press.

Bachelor, A. (2013). Clients' and therapists' views of the therapeutic alliance: Similarities, differences and relationship to therapy outcome. *Clinical Psychology & Psychotherapy*, 20(2), 118–135. doi:10.1002/cpp.792

Bassler, M., Potratz, B., & Krauthauser, H. (1995). Der "Helping Alliance Questionnaire" (HAQ) von Luborsky [The "Helping Alliance Questionnaire" (HAQ) of Luborsky]. *Psychotherapeut*, 40, 23–32.

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

Crits-Christoph, P., Johnson, J. E., Connolly Gibbons, M. B., & Gallop, R. (2013). Process predictors of the outcome of group drug counseling. *Journal of Consulting and Clinical Psychology*, 81(1), 23–34. doi:10.1037/a0030101

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

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

Frank, J. D. (1993). *Persuasion and healing: A comparative study of psychotherapy* (3rd. ed.). Baltimore: Johns Hopkins University Press.

Freud, S. (1912). Zur Dynamik der Übertragung [The dynamics of transference]. *Zentralblatt für Psychoanalyse*, 2, 167–173.

Gelso, C. J. (2009). The time has come: The real relationship in psychotherapy research. *Psychotherapy Research*, 19, 278–282. doi:10.1080/10503300902777155

Hafkenscheid, A. (2009). The impact of psychotherapy sessions: Internal structure of the Dutch Session Evaluation Questionnaire (SEQ). *Psychology and Psychotherapy: Theory, Research and Practice*, 82(1), 99–111. doi:10.1348/147608308X380750

Hartmann, A., Leonhart, R., Hermann, S., Joos, A., Stiles, W. B., & Zeeck, A. (2013). Die Evaluation von Therapiesitzungen durch Patienten und Therapeuten: Faktorstruktur und Interpretation des SEQ-D [Patients' and therapists' evaluations of therapy sessions: Factor structure and interpretation of the SEQ-D]. *Diagnostica*, 59(1), 45–59. doi:10.1026/0012-1924/a000078

Hartmann, A., Orlinsky, D. E., Geller, J. D., & Zeeck, A. (2003). Der Inter-Session-Fragebogen (ISF) - Ein Instrument zur Erfassung von psychotherapierelevanten Prozessen zwischen Sitzungen [The Intersession Experience Questionnaire (IEQ) - A measure for therapy related processes between sessions]. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 53, 464–468.

Hartmann, A., Orlinsky, D. E., Weber, S., Sandholz, A., & Zeeck, A. (2010). Session and intersession experience related to treatment outcome in bulimia nervosa. *Psychotherapy: Theory/Research/Practice/Training*, 47, 355–370. doi:10.1037/a0021166

Hartmann, A., Orlinsky, D. E., & Zeeck, A. (2011). The structure of intersession experience in psychotherapy and its relation to the therapeutic alliance. *Journal of Clinical Psychology*, 67, 1044–1063. doi:10.1002/jclp.20826

Hartmann, A., Schröder, T., Orlinsky, D. E., & Zeeck, A. (2010, June 25). *Psychometric properties and application of the Therapist Work Involvement Scales: A short form of the Development of Psychotherapists Common Core Questionnaire*. Paper presented at the Society for Psychotherapy Research, 41st International Meeting, Asilomar, CA.

Hendriksen, M., Van, R., Peen, J., Oudejans, S., Schoevers, R., & Dekker, J. (2010). Psychometric properties of the Helping Alliance Questionnaire-I in psychodynamic psychotherapy for major depression. *Psychotherapy Research*, 20, 589–598. doi:10.1080/10503307.2010.493539

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36, 223–233. doi:10.1037/0022-0167.36.2.223

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

Krause, M. S., & Lutz, W. (2009). Process transforms inputs to determine outcomes: Therapists are responsible for managing process. *Clinical Psychology: Science and Practice*, 16(1), 73–81. doi:10.1111/j.1468-2850.2009.01146.x

Laird, R. D., & Weems, C. F. (2011). The equivalence of regression models using difference scores and models using separate scores for each informant: Implications for the study of informant discrepancies. *Psychological Assessment*, 23, 388–397. doi:10.1037/a0021926

Lambert, M. J. (2010). *Prevention of treatment failure: The use of measuring monitoring, and feedback in clinical practice* (1st ed.). Washington, D.C.: American Psychological Association.

Luborsky, L., Barber, J. P., & Crits-Christoph, P. (1990). Theory-based research for understanding the process of dynamic psychotherapy. *Journal of Consulting and Clinical Psychology, Special Series: Advances in Psychotherapy Process Research*, 58, 281–287.

Marmarosh, C. L., & Kivlighan, D. M. J. (2012). Relationships among client and counselor agreement about the working alliance, session evaluations, and change in client symptoms using response surface analysis. *Journal of Counseling Psychology*, 59, 352–367. doi:10.1037/a0028907

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, 627–646. doi:10.1080/10503307.2010.497633

Orlinsky, D. E. (2008). Die psychotherapeutische Beziehung, Soziale Unterstützung und die Heilende Energie des Therapeuten: Eine Neo-Durkheim'sche Perspektive. [The therapeutic relationship, social support, and the psychotherapist's healing energy: A neo-Durkheimian perspective]. In B. Röhrle & A-R. Laireiter (Eds.), *Soziale Unterstützung und Psychotherapie.* [Social support and psychotherapy.] Tübingen: dgvt-Verlag.

Orlinsky, D. E. (2010). The "Generic Model of Psychotherapy" after 25 years: Evolution of a research-based metatheory. *Journal of Psychotherapy Integration*, 19, 1–23.

Orlinsky, D. E., Ambühl, H., Rønnestad, M., Davis, J., Gerin, P., Davis, M., … SPR Collaborative Research Network. (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., & Geller, J. D. (1993). *Patients' representations of their therapists and therapy: New measures. Psychodynamic treatment research: A handbook for clinical practice* (pp. 423–466). New York, NY: Basic Books.

Orlinsky, D. E., & Howard, K. I. (1975). *Varieties of psychotherapeutic experience: Multivariate analyses of patients' and therapists' reports*. New York, NY: Teachers College Press.

Orlinsky, D. E., & Howard, K. I. (1977). The therapist's experience of psychotherapy. In A. Gurman & A. Razin, Eds., *Effective psychotherapy: A handbook of research*. New York, NY: Pergamon.

Orlinsky, D. E., & Howard, K. I. (1986). Process and outcome in psychotherapy. In S. L. Garfield & A. E. Bergin, Eds., *Handbook of psychotherapy and behavior change* (3rd ed., pp. 311–381). New York, NY: Wiley.

Orlinsky, D. E., & Rønnestad, M. H. (2005). *How psychotherapists develop: A study of therapeutic work and professional growth*. Washington, DC: American Psychological Association.

Orlinsky, D. E., Rønnestad, M. H., & Willutzki, U. (2004). Fifty years of psychotherapy process - outcome research: Continuity and change. In M. J. Lambert & D. R. Dupper (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 307–390). New York, NY: Wiley John + Sons.

Orlinsky, D. E., Willutzki, U., Meyerberg, J., Cierpka, M., Buchheim, P., Ambühl, H., & SPR-Collaborative Research Network. (1996). Die Qualität der therapeutischen Beziehung: Entsprechen gemeinsame Faktoren in der Psychotherapie gemeinsamen Charakteristika von PsychotherapeutInnen? *Psychotherapie Psychosomatik Medizinische Psychologie*, 46(3–4), 102–110.

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology*, 21(2), 95–103. doi:10.1037/h0045357

Safran, J. D. & Muran, J. C. (2000). *Negotiating the therapeutic alliance: A relational treatment guide*. New York, NY: Guilford.

Safran, J. D., Muran, J. C., & Eubanks-Carter, C. (2011). Repairing alliance ruptures. *Psychotherapy*, 48, 80–87. doi:10.1037/a0022140

SAS Institute Inc. (2009). *JMP® 8 User Guide* (Vol. 8). Cary, NC: Author.

Stiles, W., Reynolds, S., Hardy, G., Rees, A., Barkham, M., & Shapiro, D. (1994). Evaluation and description of psychotherapy sessions by clients using the Session Evaluation Questionnaire and the Session Impacts Scale. *Journal of Counseling Psychology*, 41, 175–185.

Stiles, W. B., Gordon, L. E., & Lani, J. A. (2002). Session evaluation and the Session Evaluation Questionnaire. In G. S. Tyron (Ed.), *Counseling based on process research: Applying what we know* (pp. 325–243). Boston, MA: Allyn & Bacon.

Thomas, D. R., & Zumbo, B. D. (2012). Difference scores from the point of view of reliability and repeated-measures ANOVA: In defense of difference scores for data analysis. *Educational and Psychological Measurement*, 72(1), 37–43. doi:10.1177/0013164411409929

Tryon, G. S., Blackwell, S. C., & Hammel, E. F. (2007). A meta-analytic examination of client–therapist perspectives of the working alliance. *Psychotherapy Research*, 17, 629–642. doi:10.1080/10503300701320611

Zeeck, A., Hartmann, A., & Orlinsky, D. E. (2004). Inter-Session-Prozesse. Ein vernachlässigtes Thema der Psychotherapieforschung [Intersession processes: A neglected field of psychotherapy research]. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 54, 236–242. doi:10.1055/s-2003-814790

Zeeck, A., Herzog, T., Kuhn, K., Hartmann, A., Scheidt, C., & Wirsching, M. (2002). Teilstationäre Psychotherapie: Settingbesonderheiten und Indikationsstellung am Beispiel der Freiburger Tagesklinik [Day clinic psychotherapy: Setting and indications using the example of the Freiburg Dayclinic]. *Psychotherapie Psychosomatik Medizinische Psychologie*, 52, 492–499. doi:10.1055/s-2002-36007

Zeeck, A., Orlinsky, D. E., Hermann, S., Joos, A., Wirsching, M., Weidmann, W., & Hartmann, A. (2012). Stressful involvement in psychotherapeutic work: Therapist, client and process correlates. *Psychotherapy Research*, 22, 543–555. doi:10.1080/10503307.2012.683345