S1: Rationale for the use of post-session questionnaires

The post-session questionnaires were designed as school-independent tools that allow the assessment of general factors/mechanisms in psychotherapy. The five factors used in this sample thus represent a collection that originated from empirically based and theoretically derived research: The patient-therapist relationship (the **alliance**) is one factor that is unanimously regarded as crucial for the success of therapy (Horvath, Del Re, Flückiger, & Symonds, 2011). Alliance may better predict therapy success than the use of specific techniques (Messer & Wampold, 2002; Norcross & Wampold, 2011). Concurrently, we measured alliance from both the patient and therapist perspective throughout the course of therapy. The concept of **self-efficacy** has also had a profound impact on the field of psychology (Bandura, 1977), and several studies demonstrate the importance of self-efficacy for positive outcomes in psychotherapy (e.g. Goldin et al., 2012; Tschacher et al., 2000; Williams, 2010). In addition to these variables capturing individual experiences during therapy sessions, we were also interested in the kind of interventions administered by therapists. Thus we included two factors that describe whether and how strongly two basic intervention strategies were used within a session: **Clarification** pertains to the focal change factor proposed by psychodynamic and humanistic therapies; **mastery** is a concept prominently featured in behavioral therapies. According to the dual concept of psychotherapy (Grawe, 2004), on which the therapies of the current sample were based, both factors are needed to optimally tailor interventions to a specific patient's needs (Grawe, 1997; 2004). The dual concept of psychotherapy encourages therapists to individually design treatments that satisfy patients' specific needs and predispositions. This aim is achieved by carefully assessing patients' plans (goals and needs that may be behaviorally manifest and/or indirectly inferred from other sources such as questionnaires or behavioral observation). This information allows formulating so-called plan analyses (Caspar, 2007), which provide the framework for individualized treatments.

S2: Sample Characteristics

Participants

A sample of 87 dyadic psychotherapy courses (49 of 87 patients were women; mean age = 34.5 years, *SD* = 10.3, all White Caucasian European ethnicity) was taken from a comprehensive database established at the outpatient psychotherapy clinic of the University of Bern, Switzerland. Patients admitted to this clinic received either individual or group psychotherapy. Patients who had psychotic disorders, severe substance dependence, or who received group psychotherapy treatment were excluded from this study. Patients were included only when they were enrolled in dyadic psychotherapy and belonged to one of the following diagnostic groups: affective disorders (31.9%), anxiety disorders (38.3%), other diagnoses (29.8%; with 12.5% adjustment disorders, 10.2% other disorders, 7.1% without a DSM-IV diagnosis). Approximately 10% of all patients suffered from personality disorders in addition to their main diagnosis. Comorbidity of Axis I disorders was 31.5% and predominantly found in patients with anxiety disorders (36% comorbid patients) and affective disorders (32%). Importantly, the study sample was representative of the database population (N = 474 cases): 28% anxiety disorders; 25% affective disorders; 9% adjustment disorder; 2.5% eating disorders; 10.8% no Axis I disorder; 6% personality disorders. All diagnoses were assessed prior to therapy using the Structured Clinical Interview (SCID; Wittchen, Wunderlich, Gruschwitz, & Zaudig, 1997) for the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV; American Psychiatric Association, 1994).

Patients were treated between 1996 and 2006 by therapists practicing integrative therapy with a combination of cognitive-behavioral, process-experiential, and interpersonal interventions (Grawe, 2004), which were based on individual case formulations (Caspar, 2007). Therapists were either experts (n = 6) or postgraduate psychologists in a psychotherapy training

Ramseyer, F., Kupper, Z., Caspar, F., Znoj, H., & Tschacher, W. (2014). Time Series Panel Analysis (TSPA) – Multivariate modeling of temporal associations in psychotherapy process. *Journal of Consulting and Clinical Psychology.*

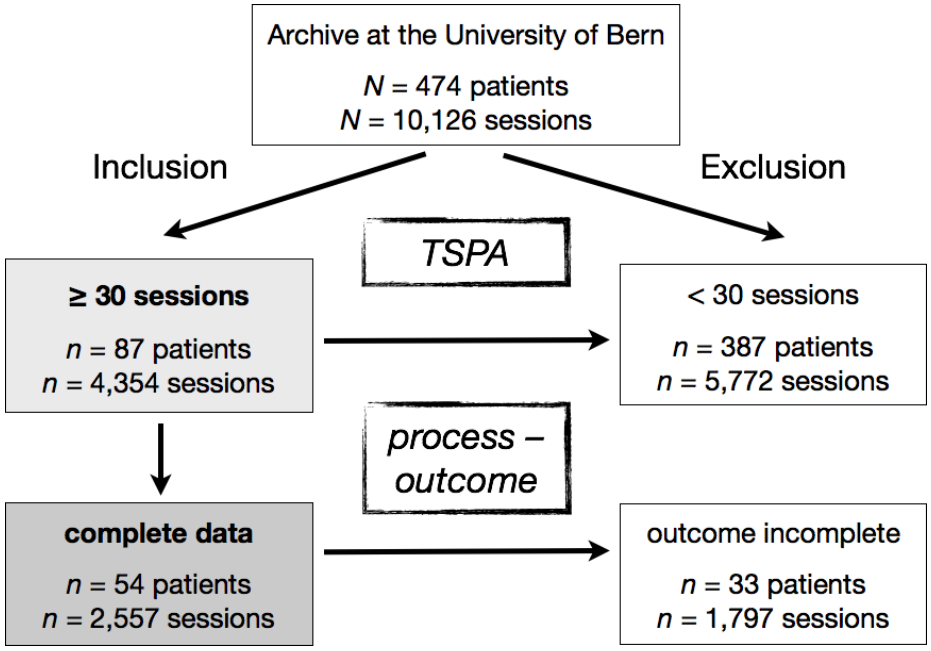

Figure W1. Flow chart of session selection criteria.

program (n = 50). Two experts had seven and four patients, respectively, while most therapists treated only a single patient. Only completed therapies with a minimum of 30 sessions were included in TSPA analyses. This additional inclusion criterion constituted a trade-off between large sample size and sufficient lengths of time series for TSPA analyses (see Methods section in main text). The 30-session criterion led to an overrepresentation of therapies with longer duration: The mean therapy duration of the entire sample (N = 474) was 21.4 sessions (SD = 20.3; MDN = 14; range = 1–112), whereas the selected therapies lasted an average of 50.0 sessions (SD = 18.6; MDN = 43; range = 30–112). These selection criteria also implied that early termination therapies were not considered in the current sample (flow chart, Fig. W1). Patients and therapists gave written informed consent in compliance with Swiss ethics legislation allowing scientific use of anonymized data.

Measures of therapeutic process

Post-session questionnaires. Versions of the Bern Post-Session Report (BPSR; Flückiger et al., 2010) were independently administered to patients (BPSR-P) and therapists (BPSR-T) after each therapy session. Session-report measures comprised a total of 22 (BPSR-P) and 27 (BPSR-T) items loading on five previously described psychotherapy factors (Tschacher et al., 2000; Tschacher, Ramseyer, & Grawe, 2007). These five rotated factors were derived from a previous sample of available therapies (N = 163) using principal component analysis on the averages of patients' and therapists' session report items. Averages of single items were used in order to control for the serial dependency stemming from repeated measurements within patients. The five factors explained 68.8 % of the total variance; 7 items had been excluded due to low communalities. (Statistical details of the present sample are reported in Table W1). Two factors captured the patient's view of therapy process: alliance (AL_p: 12 items; e.g. "Today, I felt comfortable in the relationship with my therapist"; Cronbach's α = .92 in the present sample) and patient self-efficacy (SE_p: 7 items; e.g. "I now feel more capable of solving my problems"; α = .89). One factor reflected the therapist's perspective on the alliance (AL_t: 11 items; e.g. "Today, I felt comfortable with the patient"; α = .85), and two factors assessed the therapist's actions during a session: clarification interventions (CL_t: 8 items; e.g. "Today, I have actively worked towards helping the patient to view his problems from a different angle"; α = .84) and

Table W1
Means (M), Standard Deviations (SD) and Intercorrelations of Process Variables (N = 4354) Reported by Patients (P) and Therapists (T).
Diagonal: α Reliability Estimates (Std α).

			Correlations					Corr. with CTI Effect Size
Variable	*M*	*SD*	1	2	3	4	5	*r*
1. Alliance P (AL_p)	1.94	0.62	(.92)					.30*
2. Self-Efficacy P (SE_p)	1.14	0.84	.54	(.90)				.39**
3. Alliance T (AL_t)	1.46	0.61	.35	.40	(.84)			.25
4. Clarification T (CL_t)	0.35	1.03	-.08	.03	.06	(.83)		-.04
5. Mastery T (MA_t)	0.42	1.27	.13	.22	.22	.12	(.83)	.05

Note. * $p < .05$; ** $p < .01$.

mastery (MA_t: 4 items; e.g. "In this session, I worked towards improving the patient's coping ability in difficult situations"; $\alpha = .84$). Internal consistency of BPSR scales in other samples range from $\alpha = .75$ to .88 as reported by Flückiger et al. (2010).

Measures of therapeutic success

Therapy outcome was assessed with various instruments based on patient self-reports, comprised of pre-to-post change measures of success (comparisons of pre-post assessments) and retrospective measures of success (evaluation at termination of therapy). As described in the methods section, we restricted the analysis and description of process-outcome associations to one single pre-to-post change measure, which quantified constructive thinking (Constructive Thinking Inventory, CTI; Epstein & Meier, 1989). According to Epstein's cognitive-experiential self-theory (CEST; Epstein, 1991), a person has both a intuitive-experiential and an analytical-rational mode of thinking. The intuitive-experiential mode is intimately associated with emotions, which makes it a good assessment dimension for psychotherapy patients. The CTI-K is a German translation of the original CTI (Epstein & Meier, 1989); it measures experiential intelligence, which is reflected in a person's tendency to automatically think in ways that are important for solving everyday problems, by categorizing the thoughts as constructive or destructive. The CTI-K consists of 46 items with three sub-scales: Emotional Coping EC, sample item "I worry a great deal about what other people think of me"; and Behavioral Coping, BC, sample item "I avoid challenges because it hurts too much when I fail", and the global score Global Constructive Thinking, GCT). In this study, we only report the global constructive thinking sub-scale (GCT), which consists of items from the two other scales and indicates general constructive thinking. Epstein (2001) reported internal consistency-reliability indices of .92 for GCT, .94 for EC and .84 for BC. Test-retest reliability indices were .86 for GCT, .90 for EC and .81 for BC.

The process-outcome sample was smaller than the sample for the prototypical model because CTI-scores were only available in $n = 54$ patients. The level-2 variable "therapist" contributed minimally to the sample's variance, with an intraclass correlation coefficient of .023, and was thus not considered in further analyses [intraclass correlation was calculated in a mixed model (unconditional means model) with patient's alliance rating as the dependent variable and patient nested in therapist as random variable].

S3: Results (background on sample data; traditional pre-post analyses)

Outcome

Process factors (post-session questionnaires) in the 87 available patients showed a distinct pattern of change, when comparing the beginning of therapy (average of initial five sessions) with the end of therapy (average of final five sessions): All factors were marked by a significant increase from initial to final stages (all $p < .001$; $d = 0.44 - 1.01$). The level of change in constructive thinking manifested a high effect size [$GCT_{pre} = 2.01$, $SD = 0.36$; $GCT_{post} = 2.40$; $SD = 0.49$; $t(53) = 6.66$; $p < .0001$; $d = 0.92$], which was unrelated to duration of treatment [$r(53) = .066$; $p = .634$], and positively associated with patient's alliance and self-efficacy assessed at the end of therapy [$r(53) = .37$ / .48; $p < .01$ / .001]. Compared to the archive's sample ($N = 474$), patients in the current selection both started ($d = –0.32$) and ended ($d = –0.19$) therapies with less constructive thinking. However, their change in CTI was higher than the one found in the archive ($d_{sample} = 0.91$; $d_{archive} = 0.74$). Further details are provided in Table W2.

Table W2
Means (M), Standard Deviations (SD), and Differences of Process Variables (N = 4354) Reported by Patients (P) and Therapists (T) During Initial 5 Sessions and Final 5 Sessions. Associations with CTI Effect Sizes (ES) (n = 54).

	Initial Phase (5 Sessions)		Corr. w. CTI ES	Final Phase (5 Sessions)		Corr. w. CTI ES	Difference Initial vs. Final Phase	
Variable	*M*	*SD*	*r*	*M*	*SD*	*r*	*p*	*d*
Alliance P (AL_P)	1.69	0.57	.16	2.08	0.58	.37**	<.0001	0.68
Self-Efficacy P (SE_P)	0.68	0.72	.08	1.43	0.76	.48***	<.0001	1.01
Alliance T (AL_T)	1.31	0.44	.02	1.56	0.52	.23	<.001	0.52
Clarification T (CL_T)	-0.16	0.69	-.09	0.19	0.83	-.11	<.001	0.44
Mastery T (MA_T)	-0.34	1.07	-.10	0.52	1.05	-.004	<.0001	0.80

Note. * $p < .05$; ** $p < .01$; *** $p < .001$.

References (supplementary online material)

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.) Washington, DC: Author.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191-215. doi:10.1037/0033-295X.84.2.191

Caspar, F. (2007). Plan analysis. In T. Eells (Ed.), *Handbook of psychotherapeutic case formulations* (2nd ed., pp. 251-289). New York, NY: Guilford.

Epstein, S., & Meier, P. (1989). Constructive thinking: A broad coping variable with specific components. *Journal of Personality and Social Psychology*, *57*, 332-350. doi: 10.1037/0022-3514.57.2.332

Epstein, S. (1991). Cognitive-experiential self-theory: An integrative theory of personality: In R. Curtis (Ed.), *The relational self: Convergences in psychoanalysis and social psychology* (pp. 111-137). New York, NY: Guilford Press.

Epstein, S. (2001). *CTI. Constructive Thinking Inventory. Professional manual*. Psychological Assessment Resources (PAR), Inc. Lutz, FL: PAR.

Flückiger, C., Regli, D., Zwahlen, D., Hostettler, S., & Caspar, F. (2010). Der Berner Patienten- und Therapeutenstundenbogen 2000 [The Bern Post Session Report 2000, patient and therapist versions]. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *39*, 71-79. doi:10.1026/1616-3443/a000015

Goldin, P. R., Ziv, M., Jazaieri, H., Werner, K., Kraemer, H., Heimberg, R. G., & Gross, J. J. (2012). Cognitive reappraisal self-efficacy mediates the effects of individual cognitive-behavioral therapy for social anxiety disorder. *Journal of Consulting and Clinical Psychology*, *80*, 1034-1040. doi:10.1037/a0028555

Grawe, K. (1997). Research-Informed psychotherapy. *Psychotherapy Research*, *7*, 1-19. doi: 10.1080/10503309712331331843

Grawe, K. (2004). *Psychological therapy*. Seattle, Toronto: Hogrefe.

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

Messer, S. B., & Wampold, B. E. (2002). Let's face facts: Common factors are more potent than specific therapy ingredients. *Clinical Psychology: Science and Practice*, *9*, 21-25. doi:10.1093/clipsy.9.1.21

Norcross, J. C., & Wampold, B. E. (2011). Evidence-based therapy relationships: Research conclusions and clinical practices. *Psychotherapy*, *48*, 98-102. doi:10.1037/a0022161

Tschacher, W., Baur, N., & Grawe, K. (2000). Temporal interaction of process variables in psychotherapy. *Psychotherapy Research*, *10*, 296-308. doi:10.1093/ptr/10.3.296

Tschacher, W., Ramseyer, F., & Grawe, K. (2007). Der Ordnungseffekt im Psychotherapieprozess [The order effect in the psychotherapy process]. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *36*, 18-25. doi:10.1026/1616-3443.36.1.18

Williams, D. M. (2010). Outcome expectancy and self-efficacy: Theoretical implications of an unresolved contradiction. *Personality and Social Psychology Review*, *14*, 417-425. doi:10.1177/1088868310368802

Wittchen, H. U., Wunderlich, U., Gruschwitz, S., & Zaudig, M. (1997). *SCID: Clinical interview for DSM-IV* (German version). Göttingen, Germany: Hogrefe.

Figure W2

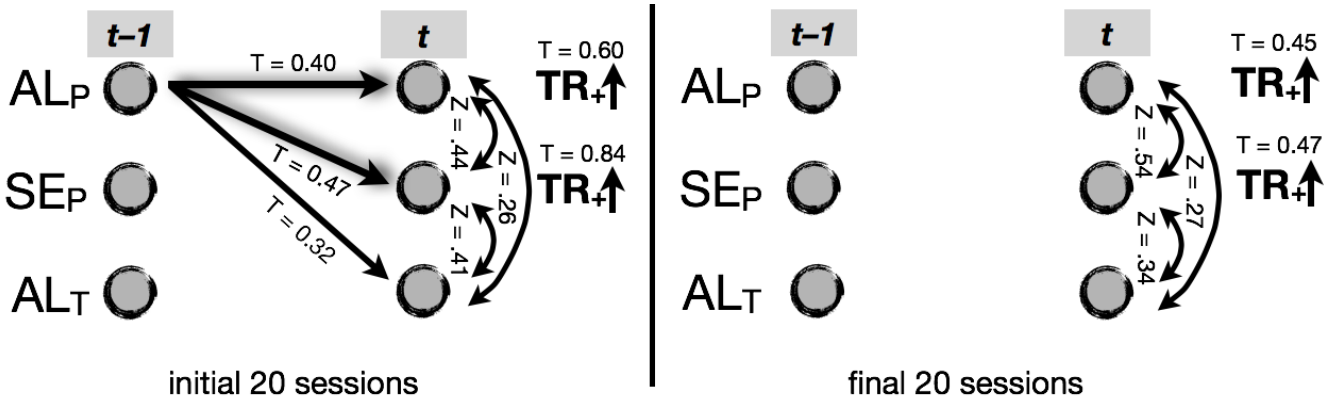

Figure W2. Prototypical mechanisms of change for initial (first 20 sessions) and final phases (last 20 sessions) of therapy.
VAR parameters that differ significantly between phases are depicted with bold/shadowed arrows.
AL_P = alliance patient; SE_P = self-efficacy patient; AL_T = alliance therapist.
TR+/- = linear trend of variable; *Z* = correlation of residuals.

Table W3

Table W3
Results of Multiple Regression Analysis of Outcome by VAR Parameters and Linear Trends (n = 54).

	Multiple Regression				
Outcome	R^2	*F*	*p*	Significant Predictors	
CTI Effect Size (Cohen's *d*)	37.3	2.05	.040	TR_SE_P	(β = .42; *p* = .032)
				CL_T–>CL_T	(β = .40; *p* = .037)
				MA_T–>MA_T	(β = .36; *p* = .024)
				AL_T–>AL_P	(β = .56; *p* = .019)
				CL_T–>SE_P	(β = .61; *p* = .001)
				MA_T–>AL_P	(β = .80; *p* = .005)
				MA_T–>CL_T	(β = .46; *p* = .024)

Note. R^2 = adjusted R^2
AL_P = alliance patient; SE_P = self-efficacy patient; AL_T = alliance therapist;
CL_T = clarification interventions therapist; MA_T = mastery interventions therapist;
TR_ = linear trend of variable
β = Standardized beta

Ramseyer, F., Kupper, Z., Caspar, F., Znoj, H., & Tschacher, W. (2014). Time Series Panel Analysis (TSPA) – Multivariate modeling of temporal associations in psychotherapy process. *Journal of Consulting and Clinical Psychology*.

Table W4. Code for SAS

```
/*********************/
/* SAS-code for TSPA */
/*********************/

/* raw-data has to be in long-format and must contain the following information: */
/* PAT_ID SESSION VARIABLE_1 VARIABLE_2 VARIABLE_3 VARIABLE_4 VARIABLE_5 */
/* dataset has to be sorted by session */

/** data-steps  **/
LIBNAME IN "C:\Users\name\Documents\My SAS Files\9.3\directory1";
/* directory containing file with raw-data */
LIBNAME SAVE "C:\Users\name\Documents\My SAS Files\9.3\directory2";
/* directory to export output files (VAR_PAR, AIC_PAR) */

data A;
set in.RAWDATA;
/* read rawdata in SAS-format (.sas7bdat) */

/** VAR-model **/
proc varmax outest=VAR_PAR outstat=AIC_PAR;
/* estimates are written to VAR_PAR , fit-statistics are written to AIC_PAR */
model variable_1 variable_2 variable_3 variable_4 variable_5 /
p=1 trend=linear;
/* maximal lag (p=) is set to 1, linear trend is used */
causal group1=(variable_2 variable_3 variable_4 variable_5)
group2=(variable_1);
/* tests whether group2 granger-causes the variables in group1 */
OUTPUT OUT=FOREC_DATA;
/* forecast data and residuals are written to FOREC_DATA */
by PAT_ID;
/* analyses are performed on individual patients (=Pat_ID) */
run;

/* save VAR parameters & STD in aggregated dataset -> calculate T-values */
data save.VAR_PAR;
set VAR_PAR;
run;

/* save fit-statistics in aggregated dataset */
data save.AIC_PAR;
set AIC_PAR;
run;

/* correlation of residuals */
PROC CORR DATA=FOREC_DATA OUTP=RESID_CORR;
VAR RES1 RES2 RES3 RES4 RES5;  /* 5 factors */
BY PAT_ID;
run;
```

Table W5. Code for *R*

```
# R-code for TSPA analysis using the library 'vars'
library(vars) # load library for VAR-models and time-series analysis

# 1. read raw data (with text headers; tab-separated)
example <- read.csv("~/Documents/JCCP/dataset.txt", h=T, sep="\t")

# 2. subsetting the data to process only one PAT at a time
# automatic sub-processing may be done using the library 'plyr'
exp_pat <- example[ which(example$patnr=="12345"), ]
# select patient with ID number 12345

# 3. convert dataset to matrix and select relevant variables for VAR
varF <- exp_pat[ , c("AL_P", "SE_P", "AL_T", "CL_T", "MA_T")]
# select rows , and variables (c) to be estimated by VAR

# 4. remove missing values (delete missings)
varF_ok <- varF[complete.cases(varF), ]

# 5. use information criteria for the selection of the best lag
VARselect(varF_ok, lag.max=5, type = c("both"), season = NULL, exogen = NULL)
# type = both -> estimates both the constant (default) and
the linear trend of variables

# 6. do the VAR-analysis (idiographic level / tspa_i)
tspa_i <- VAR(varF_ok, p = 1, type = c("both"), ic = c("AIC"))
summary(tspa_i)

# 7. causality analysis (performed for each variable separately)
causality(tspa_i, cause = "AL_P", vcov.=NULL, boot=FALSE)$Granger
# similar code for other variables

# 8. extraction of residuals
sync_assoc <- residuals(tspa_i)
```