The Therapy Process Questionnaire - Factor analysis and psychometric properties of a multidimensional self-rating scale for high-frequency monitoring of psychotherapeutic processes

Günter Schiepek\(^1,2,3\) | Barbara Stöger-Schmidinger\(^2\) | Helmut Kronberger\(^2\) | Wolfgang Aichhorn\(^1,2\) | Leonhard Kratzer\(^4\) | Peter Heinz\(^4\) | Kathrin Viol\(^1,2\) | Anna Lichtwarck-Aschoff\(^5\) | Helmut Schöller\(^1,2\)

Abstract

Many outcome measures and session-related questionnaires in psychotherapy are designed for weekly or biweekly administration. Yet, today, technical developments allow for higher frequency assessments to monitor human change dynamics more closely by daily assessments. For this purpose, the Therapy Process Questionnaire (TPQ) was developed, with a specific focus on inpatient psychotherapy. In this article, we present an explorative and confirmative factor analysis of the TPQ on the basis of the time series data of 150 patients collected during their hospital stay (mean time series length: 69.1 measurement points). A seven-factor solution was identified, which explains 68.7% of variance and associates 43 items onto the factors, which are “well-being and positive emotions,” “relationship with fellow patients,” “therapeutic relationship and clinical setting,” “emotional and problem intensity,” “insight/confidence/therapeutic progress,” “motivation for change,” and “mindfulness/self-care.” The internal consistency (Cronbach’s \(\alpha\)), the inter-item correlations of the subscales, and the discriminative power of the items are excellent. The TPQ can be applied in practice and research for creating time series with equidistant measurement points and time series lengths, which are appropriate for the application of nonlinear analysis methods. Especially in clinical practice, it is important to identify precursors of phase transitions, changing synchronization patterns, and critical or instable periods of a process, which now is possible by internet- or app-based applications of this multidimensional questionnaire.

KEYWORDS

factor analysis, nonlinear dynamics, process monitoring, psychotherapy feedback, Therapy Process Questionnaire
1 | INTRODUCTION

Outcome monitoring and feedback on therapeutic progress has become popular and has been adopted by many mental health providers all over the world (e.g., Delgadillo et al., 2017; Gibbons et al., 2015; Schiepek et al., 2016; Wampold, 2015). Lambert (2007, 2010) describes it as an important feature of good clinical practice and asks for an integration of monitoring procedures into routines of mental health care. Studies and meta-analyses report on effects of psychotherapy feedback in outcome optimization (“routine outcome monitoring”) at different degrees (mostly small to moderate effect sizes), particularly in patients at risk for deterioration (Brattland et al., 2018; de Jong, van Sluis, Nugter, Heiser, & Spinhoven, 2012; de Jong, Segaar, Ingenhoven, van Busschbach, & Timman, 2018; de Jong, Barkham, Wolpert, & the INSPIRE Research Network, in prep.; Gibbons et al., 2015; Kendrick et al., 2016; Knap, Kösters, Schöfer, Becker, & Puschner, 2009; Lambert, 2017; Lambert, Whipple, & Kleinstäuber, 2018; Shimokawa, Lambert, & Smart, 2010). Although there is a wide agreement on the usefulness of process and outcome monitoring (Wampold, 2015), there are different ways of applying it to clinical practice. The question of how it should be implemented and realized has been answered from different viewpoints.

Diverging perspectives have been previously offered to address the question if human change processes should be conceptualized as linear and input-dependent or as nonlinear, complex, and dynamic. Within the linear approach, treatment courses seem to be more or less continuous, characterized by straight lines or moderately damped trajectories (dose-effect curves), and predictable by standard tracks using the reference of patients with the same diagnosis or the same symptom severity at the beginning (Finch, Lambert, & Schaalje, 2001). In consequence, measures at low frequencies (session by session, weekly, or bi-weekly) seem to be sufficient. Within the nonlinear dynamic systems paradigm (Haken & Schiepek, 2010: Gelo & Salvatore, 2016; Schiepek, Heinzel, Karch, Plöderl, & Strunk, 2016; Schiepek & Pincus, under review), trajectories of change may have complex shapes, are usually characterized by discontinuous pattern transitions, which become evident as sudden gains or losses (Hayes, Laurenceau, Feldman, Strauss, & Cardaciotto, 2007; Lutz et al., 2013; Stiles et al., 2003) or rupture-repair sequences in the working alliance (Gumz, Brähler, Geyer, & Erices, 2012; Stiles et al., 2004), demonstrate critical instabilities and other precursors of discontinuous transitions (Schiepek, Tominschek, & Heinzel, 2014), and are chaotic, that is, not predictable in the long run (Schiepek et al., 2017). In consequence, this creates specific demands on sampling rates and time series length.

Most feedback procedures in psychotherapy are working within the linear low-frequency approach and focus on treatment outcomes. Widely used measures are the Outcome Questionnaire (Lambert et al., 2004), the World Health Organization Wellbeing Index (Bech, Gudex, & Johansen, 1996), the Mental Health subscales of the Medical Outcomes Questionnaire (Ware, Snow, Kosinski, & Gandek, 1993), the Health of the Nation Outcome Scales (Wing et al., 1998), or the Depression Anxiety Stress Scales (Lovibond & Lovibond, 1995; Nilges & Essau, 2015). Besides the outcome, process-mediating aspects and known therapeutic mechanisms and mediators also should be respected, for example, client factors (resources, motivation to change, or engagement), working alliance, emotions, self-relatedness, expectancies, self-esteem, or self-efficacy (Duncan, Miller, Wampold, & Hubble, 2010; Norcross & Lambert, 2011; Schiepek, Aichhorn, et al., 2016). There is a diversity of questionnaires focusing on such ingredients of the therapeutic work. Usually, these questionnaires are designed as post-session reports, for example, the Therapy Session Report (Orlinsky & Howard, 1967), the Vanderbilt Psychotherapy Process Scales (Smith, Hilsenroth, Baity, & Knowles, 2003), the Session Evaluation Questionnaire (Stiles, 1980; Stiles et al., 1994), the Session Impact Scale (Elliott & Wexler, 1994), or the Berne Post-Session Report (Flückiger, Regli, Zwahlen, Hostetteler, & Caspar, 2010) and its advanced development, the Scale for the Multiperspective Assessment of General Change Mechanisms in Psychotherapy (SACiP; Mander et al., 2013). The multidimensional Questionnaire for the Evaluation of Psychotherapeutic Processes (Lutz & Böhne, 2008) is applied at some few time points during the process with the ratings referring to a period of 1 week. Other session-related scales focus uniquely on the therapeutic alliance, such as the Working Alliance Inventory (Horvath & Greenberg, 1986), the Helping Alliance Questionnaire (Alexander & Luborsky, 1986), the California Psychotherapy Alliance Scales (Gaston, 1991), or the Combined Alliance Scale (Hatcher & Shannon, 2005) and its short version, the Scale of the Therapeutic Alliance-Revised (STA-R; Brockmann et al., 2011). A measure that was introduced by Mander et al. (2014) combines the items of the SACiP and the STA-R to a further instrument, the Individual Therapy Process Questionnaire. There are also scales that focus on specific constructs or specific types of interventions, such as the Applied Mindfulness Process Scale (Li, Black, & Garland, 2015).

Usually, outcome measures do not cover mechanisms of change and usually grasp longer periods of time (e.g., one, two, or more...
weeks). Questionnaires on secondary outcomes (e.g., on the quality of life or interpersonal relations) also focus on specific constructs and grasp longer periods of time. Other process questionnaires are explicitly focused on the experiences of a patient within sessions (usually all alliance questionnaires) or refer to inter-session periods or are linked to specific interventions. In opposition to all possible alternatives, the unique feature of the Therapy Process Questionnaire (TPQ) is its aptitude for high-frequency (daily) monitoring and its capacity to capture different change-related constructs.

Compared with the sampling rate of sessions, which usually take place at irregular time distances, daily measurements appear to be an achievable alternative. Days are more frequent and also equidistant and by this fulfill the criteria of (a) a complete recording of therapies (not only as an irregular event sampling), (b) frequent and (c) continuous measurements, and (d) considering practicalities of data collection. A further point is a therapeutic one: daily self-assessments—especially when combined with electronic diaries—activate the reasoning on one’s own thoughts, feelings, and behaviour (mindfulness and metacognition) and by this support self-regulation and auto-catalyse developmental processes. This is a mechanism of change that is reported by many patients and therapists and is also documented in case studies (e.g., Kratzer et al., 2018; Schiepek, Eckert, Aas, Wallot, & Wallot, 2015; Schiepek, Stöger-Schmidinger, Aichhorn, Schöller, & Aas, 2016; Schiepek, Aichhorn, Schöller, & Kronberger, 2018) and in a controlled study on the effects of process monitoring on chronically addicted inpatients (Patzig & Schiepek, 2015). More systematic studies are needed on the psychological effects of continuous self-reports and on the synergistic effects between monitoring other mechanisms of change.

Figure 1 illustrates how the dynamics of a time series (daily ratings of self-esteem from a patient with borderline personality disorder) is distorted and the information on the dynamic pattern is lost if measurement points are successively omitted. The rapid cycling of self-esteem characterizing the first weeks of a treatment vanishes if ratings are only made on every second (Figure 1b) or fourth day (Figure 1c), weekly (Figure 1d), or at mixed weekly and fortnightly intervals, the most common periodicity of therapy sessions (Figure 1e,f). Corresponding to the loss of information, the dynamics of the presented time series appear more and more linear with the shape of the curve depending on the chosen measurement points. Following the Nyquist theorem, sampling rates have to be at least twice as high as the eigenvalues of the systems under investigation (Shannon, 1949).

The nonlinear high-frequency approach to psychotherapy takes seriously what in many theoretical and empirical writings was highlighted: The dynamic and non-ergodic nature of human cognitions, emotions, and behaviour needs a switch from linear group statistics to a paradigm of understanding and measuring complex systems by using nonstationary time series analysis and even individual methods, combining idiographic with nomothetic science (Haken & Schiepek, 2010; Molenaar, 2004, 2007; Molenaar, Sinclair, Rovine, Ram, & Corneal, 2009; Piccirillo, Beck, & Rodebaugh, 2019; Schiepek et al., 2015; Schiepek, Heinzl, et al., 2016; Schiepek, Stöger-Schmidinger, et al., 2016; Tzur-Bitan, Meiran, & Shahar, 2010; Tzur-Bitan, Meiran, Steinberg, & Shahar, 2012). Chaotic dynamics produced by the functioning of nonlinear complex systems indeed are not predictable but realize specific dynamic features, which are important for understanding human change. One is the sensitive dependency of the dynamics on minimal input, which creates adaptability and flexibility of the systems to the environment (e.g., therapeutic interventions), another one is the realization of attractors (more or less stable patterns of the dynamics), and a third one is transitions between attractors depending on the control parameters of a system (Strunk & Schiepek, 2006). Within this context, clusters of outcomes may be predicted by any characteristics of psychotherapy, but the concrete trajectories of change will be unpredictable by principal reasons. Instead of including more and more predictors in linear regression models, the consequence will be to create short-term predictions (early warning systems) of critical events or changing dynamics in practice and to create nonlinear system models of involved variables in theory. A sampling rate of daily measures has proven to allow for insight in nonlinear patterns and its transitions (see Figure 1).

About two decades ago, we introduced the TPQ (Haken & Schiepek, 2010; Nischk, Grothe, & Schiepek, 2000) as a method for day-by-day assessments of psychotherapeutic processes. Since about 2 years, an optimized version of the TPQ is available for the routine assessment of psychotherapeutic change dynamics. The optimization was based on many years of clinical experience in inpatient settings.

\[ \text{FIGURE 1} \quad \text{Distortion of the dynamics of a time series by omitting measurement points. The time series represents the self-esteem of a patient with borderline personality disorder diagnosis during her hospital stay.} \]

(a) The original time series with daily measures (opaque in b–f). (b) Every second day is omitted as missing day. Fluctuations of the first weeks of the time series vanish, if ratings are only made on every fourth day (c) or weekly with some variation around an exact 7-day rhythm (d). A major loss of information and possible source of therapeutic misjudgment occurs with the common practice of occasional weekly and fortnightly measurement intervals (e,f).
Despite of the 47 items of the questionnaire, daily ratings never caused big problems. The compliance rates were amazingly high (up to 80%; Schiepek, Aichhorn, et al., 2016), which is due to the facts that (a) the self-reports could be done by electronic devices via the Synergistic Navigation System (SNS), (b) the monitoring was defined as part of the hospital and treatment routine, and—perhaps the most important—(c) the feedbacks on the resulting data (visualized as graphs and diagrams) and analysis results were integrated into the psychotherapeutic process. Patients and therapists refer on these results for planning and evaluating the process (continuous cooperative process control) and would not miss this kind of treatment support. Here, we present a factor analysis and statistical item characteristics of this questionnaire.

2 | METHODS AND MATERIALS

2.1 | Subjects

The patients of this study were treated at two psychotherapy centres, the Department of In-patient Psychotherapy at the University Hospital of Psychiatry, Psychotherapy and Psychosomatics (Paracelsus Medical University Salzburg, Austria) and the Department of Psychotraumatology at the Clinic St. Irmingard (Prien am Chiemsee, Germany). The time series of 150 patients (120 females and 30 males) were included in the analysis on the basis of a criterion of less than 20% missing data in a time series. The mean number of missing data in this sample was 2.5 days (=measurement points; standard deviation [SD]: 4.0), which corresponds to a compliance rate of 96.4%. The mean time series length was 69.1 days (SD: 22.6; see Table 1). The inclusion criterion of patients with less than 20% missing data (by this we excluded about 20% of the patients treated during the reference period of about 2 years) is due to the necessity of having time series with high variability (missing data produce straight lines in the process) to get a realistic picture of the dynamics and to get valid inter-item correlations. This procedure is justified because this is not a compliance study, which should identify realistic compliance and missing rates (compare Schiepek, Aichhorn, et al., 2016), but a factor analysis, which requires a sufficient density of time series data.

For 108 out of 150 patients, ICD-based Symptom Rating (ISR)-based assessments at the beginning of the hospital stay (pretreatment) were available (Tritt, 2015; Tritt et al., 2008), which was used for a

| TABLE 1 Patient characteristics of the sample included in the factor and item analysis of the Therapy Process Questionnaire |
|-------------------------------------------------|----------------|----------------|
| N                                              | 150            | 105            |
| m/f                                            | 30/120         | 27/78          |
| Age, AM (SD)                                   | 41.3 (11.8)    | 41.71 (11.48)  |
| Time series length (days), AM (SD)             | 69.1 (22.6)    | 74.04 (22.45)  |
| Missing days, AM (SD)                          | 2.5 (4.0)      | 2.3 (4.1)      |
| Compliance rate, AM% (SD%)                     | 96.4 (5.7)     | 96.9 (5.4)     |
| First-order diagnoses                          | Total N(%)     | UHPS N(%)      |
| Bipolar affective disorder (F31.1, F31.3, F31.6, F31.7) | 5 (3.3) | 5 (4.8) |
| Depressive episode/recurrent depressive disorder (F32.1, F32.2, F32.3, F33.0, F33.1, F33.2, F33.3, F33.4) | 40 (26.7) | 40 (38.1) |
| Phobic anxiety disorders/other anxiety disorders (F40.1, F41.0, F41.2) | 5 (3.3) | 5 (4.8) |
| Obsessive-compulsive disorder (F42.0, F42.1, F42.2) | 8 (5.3) | 8 (7.6) |
| Reaction to severe stress/adjustment disorders (F43.0, F43.1, F43.2) | 66 (44.0) | 44 (41.9) | 22 (48.9) |
| Dissociative (conversion) disorders (F44.8, F44.81, F44.9) | 10 (6.6) | 0 | 10 (22.2) |
| Somatoform disorders (F54.0, F45.40, F45.41) | 3 (2.0) | 0 | 3 (6.7) |
| Disorders of adult personality and behaviour (F60.30, F60.31, F60.6, F60.8, F61.0) | 8 (5.3) | 0 | 8 (17.8) |
| Other disorders (F10.21, F23.8, F25.1, F48.1, F91.0) | 5 (3.3) | 3 (2.9) | 2 (4.4) |

Abbreviations: AM, arithmetic mean; SD, standard deviation; SIP, patients treated at the Department for Psychotraumatology, Clinic St. Irmingard, Prien; UHPS, subsample of patients who were treated at the Department of Inpatient Psychotherapy, University Hospital of Psychiatry, Psychotherapy and Psychosomatics, Salzburg.
construct validation of the subscales of the TPQ (not all patients filled in the ISR pretreatment and posttreatment). The ISR is a first-order criterion outcome measure, assessing symptom severity corresponding to the criteria of the diagnostic F-clusters of the ICD-10. It is not a process monitoring system, but we can expect—and this would be at least a partial construct validation—that some of the factors of the TPQ may be positively or negatively linked to symptom severity; for example, the factors (see below) "well-being and positive emotions" (WPE), "motivation for change" (MOT), or "mindfulness/self-care" (MSC) are supposed to be negatively correlated with symptom severity, whereas "emotional and problem intensity" (EPI) may be positively correlated with symptom severity.

The diagnoses of the patients (following the nomenclature of the ICD-10) are shown in Table 1. The two main diagnostic clusters of the sample are "depressive episode/recurrent depressive disorder" (26.7%) and "reaction to severe stress/adjustment disorders" (44.0%). The mean numbers of co-diagnoses are 0.80 (bipolar affective disorders), 1.13 (depressive episode/recurrent depressive disorder), 0.60 (phobic anxiety disorders/other anxiety disorders), 1.13 (obsessive-compulsive disorder [OCD]), 2.73 (reaction to severe stress/adjustment disorders), 2.17 (dissociative [conversion] disorders), 1.67 (somatoform disorders), 2.88 (disorders of adult personality and behaviour), and 2.00 (other disorders).

Ethical approval for the application of the TPQ by using an internet-based device for patient monitoring and the usage of the retrieved data was given by the ethical committee of the Salzburg County Governance. Daily self-ratings by using the TPQ were based on a written informed consent of all patients. All procedures were in accordance with the Helsinki Declaration as revised in 2013.

2.2 Questionnaires and monitoring technology

The TPQ was administered in a day-by-day routine during inpatient psychotherapy. For reasons of cross-validation, we used the ISR (Fischer, Titt, Klapp, & Fliege, 2009, 2011; Titt, 2015; Titt et al., 2008), which was applied at the beginning (pre) and at the end (post) of the treatment. The subscales of the ISR are "depression," "anxiety," "OCD," "somatoform disorder," "eating disorder," and an additional scale with problems not related to the other subscales. The total score of the ISR averages all subscales by a weight of 1, the additional scale by a weight of 2. Both questionnaires, the TPQ and the ISR, were administered by an internet- or app-based (this is up to free choice) monitoring system, the SNS, which was developed for the assessment of processes and outcomes in naturalistic settings (Schiepek, Aichhorn, et al., 2016; Schiepek, Aichhorn, & Schöller, 2018; Schiepek, Stöger-Schmidinger, et al., 2016).

2.3 Construction and history of the TPQ

The items of the TPQ were created by a literature research on psychological factors contributing to the engagement of patients in their change process ("process involvement"), a construct close to that of "self-relatedness," which is a core concept of the Generic Model of Psychotherapy (Orlinsky, 2009; Orlinsky & Howard, 1986; Orlinsky, Ronnestad, & Willutzki, 2004). The questionnaire was factor analysed on the basis of data from a process–outcome study with inpatients treated in routine practice (Honermann & Schiepek, 2006; N = 94 cases, seven factors with 59.6% explained variance). A second factor analysis was based on another sample of inpatients (Schiepek, Aichhorn, & Strunk, 2012; N = 149 cases, five factors with 74.5% explained variance). Beginning from 2007, the TPQ was presented and filled in by an electronic, internet-based device, the SNS (Schiepek et al., 2015; Schiepek, Aichhorn, et al., 2016). After some years of experience in different inpatient settings, the items were revised by the team of the Department of Inpatient Psychotherapy of the Christian Doppler Medical Center Salzburg. On the basis of clinical experience, some items were omitted, others were newly introduced. Items were omitted because of unclear meaning (e.g., "Today, I felt compassion"—to whom, self or others? or "Today I felt confused.") New items were introduced because mindfulness, self-care, and body-related experiences seemed important to most of the therapists in the team. The hitherto combined 7-point Likert scales and visual analogue scales were replaced by visual analogue scales only for all 47 items. This revised TPQ was used from spring 2016 up to now. The data used for the exploratory factor analysis and confirmatory factor analysis (CFA) were collected during this period.

2.4 Statistics

The exploratory factor analysis was realized by the methods of principal component analysis (PCA) and principal axis factoring using SPSS 24.0.0.1 (64 Bit version). The adequacy of the data for factor analysis was tested by the Kaiser–Meyer–Olkin (KMO) score and the Bartlett’s test of sphericity. The KMO score is a measure of the proportion of variance among the included variables that might be common variance; scores between 0.8 and 1 indicate the appropriateness of the sample for factor analysis. The Bartlett’s test checks if the correlation matrix of the included items deviates from the unity matrix, that is, if the correlations are unequal to 0.

The CFA (Brown, 2015) was realized in R (Studio Version 1.1.463; © 2009–2018, RStudio Inc with R-version 3.5.0 [2018-04-23], and R packages “lavaan”[version 0.6–3] and “psych” [version 1.8.10]) using the χ² fit criterion, the Tucker–Lewis index (TLI), the comparative fit index (CFI), and the root mean square error of approximation (RMSEA). The TLI (Tucker & Lewis, 1973) analyses the discrepancy between the chi-squared value of the hypothesized model and the null model. This nonnormed fit index resolves some problems of the negative bias of the normed fit index. Its range is between 0 and 1, with a cut-off for values >.95 indicating a good model fit (Hu & Bentler, 1999).
The CFI analyses the model fit by examining the discrepancy between the data and the hypothesized model while adjusting for the issues of sample size inherent in the chi-squared test of model fit and the normed fit index (Bentler, 1990). CFI values range from 0 to 1, with larger values indicating better fit. CFI values >.95 can be accepted as an indicator of good fit (Bentler & Bonett, 1980; Hu & Bentler, 1999).

The RMSEA avoids issues of sample size by analysing the discrepancy between the hypothesized model, with optimally chosen parameter estimates, and the population covariance matrix (Hooper, Coughlan, & Mullen, 2008). It ranges from 0 to 1, with smaller values indicating better model fit. A value of .06 or less is indicative of acceptable model fit. The standardized root mean square residual (SRMR) indicates the square root of the difference between the residuals of the sample covariance matrix and the hypothesized covariance model (Kline, 2005). Values of the SRMR range from 0 to 1, with models obtaining values less than .08 being deemed acceptable (Hu & Bentler, 1999).

The item analysis and psychometric characteristics of the TPQ includes Cronbach’s α, a measure of the internal consistency of a scale, the mean intercorrelation between the items of a subscale, and the mean discriminative power, which is the mean correlation of each item of a subscale with the total score of the subscale. The calculation was realized by SPSS 24.0.0.1 (64 Bit version, subprogram “reliability analysis”). An overall calculation of Cronbach’s α was realized for the factors (subscapes) on the basis of the complete set of linked multiple time series of all items over 10.442 measurement points. A further calculation of Cronbach’s α values was realized for each time point. Given the fact that the mean time series length of our sample is 69.1 (the next integer value is 70, see Table 1), we calculated the values of each measurement point from 1 to 70. This procedure ensures that more than 50% of all subjects (85 out of 150) were included in the calculation. In Table 2, the mean Cronbach’s α of the measurement points from 1 to 70 and the minimum and maximum values out of 70 calculations are reported. The analysis procedure was realized as proposed by Cronbach (1951) using R version 3.5.3 (R Core Team, 2019) and the ltm package (Rizopoulos, 2006).

### TABLE 2  Characteristics of the factors (subscapes) of the TPQ

| Factor                               | EV   | Var%  | VarC% | Cα   | Ca(70) (min–max) | IIC  | MDP  | n   |
|--------------------------------------|------|-------|-------|------|-----------------|------|------|-----|
| Well-being and positive emotions (WPE)| 6.81 | 15.84 | 15.84 | 0.944| 0.922 (0.881–0.945)| 0.711| 0.816| 7   |
| Relationship with fellow patients (RFP)| 6.13 | 14.27 | 30.11 | –   | –               | 0.552| 0.552| 2   |
| Therapeutic alliance and clinical setting (TAS) | 4.81 | 11.19 | 41.30 | 0.944| 0.919 (0.855–0.946)| 0.740| 0.832| 6   |
| Emotional and problem intensity (EPI)  | 4.75 | 11.05 | 52.35 | 0.919| 0.904 (0.826–0.935)| 0.595| 0.717| 9   |
| Insight/confidence/therapeutic progress (ICP)| 3.47 | 8.08  | 60.43 | 0.943| 0.931 (0.854–0.953)| 0.625| 0.766| 10  |
| Motivation for change (MOT)           | 2.47 | 5.75  | 66.17 | 0.928| 0.891 (0.821–0.921)| 0.721| 0.812| 5   |
| Mindfulness/self-care (MSC)           | 1.08 | 2.51  | 68.68 | 0.924| 0.866 (0.789–0.891)| 0.753| 0.824| 4   |

Abbreviations: Ca, Cronbach’s α based on the complete set of multiple time series of all items over 10.442 measurement points; Ca(70), mean of all values of Cronbach’s α calculated for the measurement points 1 to 70; EV, eigenvalues of the extracted factors (stop criterion: eigenvalue >1); IIC, mean inter-item correlations of the factors; MDP, mean discriminant power of the items of each factor; [min–max], smallest and largest value of Cronbach’s c calculated for the measurement points 1 to 70 (the mean time series length of 70 still includes 85 out of 150 patients); n, number of items of each factor; TPQ, Therapy Process Questionnaire; Var%, explained variance of the factors; VarC%, cumulated explained variance of the factors.

### 3  RESULTS

#### 3.1  Factor analysis

The database of the factor analysis was a multiple time series of the 47 items of the TPQ. The time series of all 150 subjects were linked together, resulting in an artificial time series of 10.442 measurement points. The inter-item correlation matrix of all items was calculated on this multiple time series. A KMO score of 0.968 and a highly significant (χ² = 426.042.12, degrees of freedom [df] = 903, p < .001) Bartlett’s test of sphericity confirmed the adequacy of the data for factor analysis, that is, the correlations in the correlation matrix are unequal to 0.

#### 3.1.1  Exploratory factor analysis

The factor analysis was realized by the methods of PCA and principal axis factoring, with similar results for the number of extracted factors and the assignment of the items to the factors with an assignment criterion of 0.5 for the rotated factor matrix. As the factor loadings obtained by varimax and oblimin direct (delta = 0) rotation yielded similar results, we report on the solutions of the PCA method with varimax rotation. The exploratory factor analysis (PCA) revealed a seven-factor solution, with the criterion of an eigenvalue >1. The seven-factor solution accounts for 68.68% of the variance of the inter-item correlations. Table 2 shows the eigenvalues of the extracted factors, the variance, which is explained by each factor, and the cumulated variance of all factors. Table 3 presents the items of the TPQ and their association to the factors, the factor loadings of all items on the corresponding factor, the arithmetic mean and the standard deviation of the factor’s and item’s time series, and the communalities of the items (h²). All items realize high and distinct factor loadings, indicating exclusive item–factor associations of each item to one specific factor.

Items with lower factor loadings (cut-off criterion: <.5) or insufficient distinctness of the loadings between different factors were eliminated from the list of items (Table 4). This concerns the items “Today, I was...”
### TABLE 3 Arithmetic mean, standard deviation, factor loadings, and communalities ($h^2$) of the items corresponding to the factors of the TPQ

| Item (VAS description)                                                                 | AM  | SD  | WPE | RFP | TAS | EPI | ICP | MOT | MSC | $h^2$ |
|---------------------------------------------------------------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------|
| **Well-being and positive emotions (WPE)**                                               |     |     |     |     |     |     |     |     |     |       |
| 1. Today I felt joy (not at all–very much)                                             | 41.99 | 28.22 | 0.748 | 0.102 | 0.060 | -0.285 | 0.280 | 0.139 | 0.167 | 0.774 |
| 2. Today I experienced moments of happiness and light-heartedness (not at all–very much) | 39.48 | 28.21 | 0.746 | 0.083 | 0.042 | -0.311 | 0.280 | 0.109 | 0.147 | 0.768 |
| 3. Today I felt energized (not at all–very much)                                        | 40.73 | 26.65 | 0.670 | 0.069 | 0.064 | -0.328 | 0.337 | 0.154 | 0.123 | 0.770 |
| 4. Today my self-esteem was (very low–very high)                                        | 38.70 | 26.74 | 0.651 | 0.124 | 0.041 | -0.346 | 0.348 | 0.174 | 0.250 | 0.703 |
| 5. Today I felt comfortable in my body (not at all–very comfortable)                   | 36.53 | 27.71 | 0.645 | 0.113 | 0.018 | -0.352 | 0.367 | 0.108 | 0.175 | 0.721 |
| 6. Today I was satisfied with myself (not at all–very much)                            | 43.12 | 27.92 | 0.621 | 0.080 | 0.065 | -0.373 | 0.376 | 0.221 | 0.262 | 0.788 |
| 7. Today I felt valued (not at all–very much)                                           | 49.26 | 29.20 | 0.534 | 0.111 | 0.053 | -0.144 | 0.305 | 0.225 | 0.215 | 0.501 |
| **Relationship with fellow patients (RFP)**                                             |     |     |     |     |     |     |     |     |     |       |
| 8. Today I experienced myself in contact with fellow patients (not at all tense–very much tense) | 67.35 | 25.92 | -0.145 | -0.568 | -0.034 | 0.358 | -0.038 | -0.032 | 0.002 | 0.664 |
| 9. I feel at ease with other patients (not at all–very much)                          | 35.94 | 29.30 | 0.177 | 0.703 | 0.214 | -0.218 | 0.163 | 0.122 | 0.086 | 0.479 |
| **Therapeutic alliance and clinical setting (TAS)**                                    |     |     |     |     |     |     |     |     |     |       |
| 10. I perceive the work with my therapist(s) as helpful (not at all–very much)        | 84.28 | 20.11 | 0.080 | 0.033 | 0.905 | -0.047 | 0.124 | 0.123 | 0.053 | 0.860 |
| 11. I feel understood by my therapist(s) (not at all–very much)                       | 84.72 | 19.49 | 0.083 | 0.071 | 0.889 | -0.044 | 0.080 | 0.113 | 0.041 | 0.824 |
| 12. I feel supported by my therapist(s) in my personal development (not at all–very much) | 83.90 | 20.63 | 0.091 | -0.014 | 0.885 | -0.072 | 0.161 | 0.061 | 0.065 | 0.827 |
| 13. I feel safe and supported in the clinic (not at all–very much)                    | 82.48 | 21.25 | 0.018 | 0.121 | 0.792 | -0.100 | 0.121 | 0.129 | 0.039 | 0.675 |
| 14. I can be open and honest with my therapist(s) (not at all–very much)              | 87.87 | 17.03 | -0.052 | 0.018 | 0.784 | -0.086 | 0.169 | 0.189 | 0.018 | 0.627 |
| 15. This clinic is the right place for me and my personal development (not at all–very much) | 82.65 | 21.62 | 0.030 | 0.020 | 0.778 | -0.053 | 0.067 | 0.090 | 0.050 | 0.685 |
| **Emotional and problem intensity (EPI)**                                              |     |     |     |     |     |     |     |     |     |       |
| 16. Today I felt anxious (not at all–very much)                                        | 47.95 | 32.30 | -0.138 | -0.056 | -0.007 | 0.814 | -0.071 | -0.072 | -0.098 | 0.700 |
| 17. Today I felt tense and restless (not at all–very much)                            | 58.01 | 29.94 | -0.269 | 0.009 | -0.061 | 0.792 | -0.143 | -0.059 | 0.008 | 0.667 |
| 18. Today my problems/complaints were ... (absent–very intense)                       | 53.93 | 27.17 | -0.171 | -0.034 | -0.006 | 0.792 | -0.081 | -0.027 | -0.061 | 0.723 |
| 19. Today I felt impaired by my complaints in my daily routine (not at all–very much) | 51.80 | 28.79 | -0.256 | -0.001 | -0.067 | 0.775 | -0.136 | -0.092 | 0.017 | 0.690 |
| 20. Today I felt sad (not at all–very much)                                           | 48.61 | 33.52 | -0.129 | -0.036 | -0.106 | 0.743 | -0.069 | -0.013 | -0.061 | 0.587 |
| 21. Today I felt guilty (not at all–very much)                                         | 41.18 | 34.13 | -0.090 | -0.196 | 0.037 | 0.676 | 0.009 | -0.127 | -0.180 | 0.575 |
| 22. Today I felt helpless with regard to my problems (not at all–very much)          | 43.28 | 30.18 | -0.255 | -0.001 | -0.158 | 0.662 | -0.151 | -0.125 | -0.092 | 0.544 |
| 23. Today I felt shame (not at all–very much)                                         | 39.92 | 33.50 | -0.137 | -0.261 | -0.027 | 0.633 | -0.015 | -0.013 | -0.216 | 0.524 |
| 24. Today I felt angry (not at all–very much)                                         | 37.30 | 33.40 | -0.017 | -0.111 | -0.115 | 0.581 | 0.004 | -0.048 | -0.060 | 0.359 |
| **Insight/confidence/therapeutic progress (ICP)**                                      |     |     |     |     |     |     |     |     |     |       |
| 25. Today I became aware of relations that were not clear to me before (not at all–very much) | 42.22 | 28.73 | 0.153 | 0.047 | 0.135 | -0.050 | 0.818 | 0.150 | 0.136 | 0.757 |
| 26. I now understand myself and my problems better (not at all–very much)            | 45.13 | 27.22 | 0.208 | 0.027 | 0.166 | -0.151 | 0.792 | 0.168 | 0.209 | 0.786 |
| 27. Today I came closer to a solution for my problems (not at all–very much)          | 40.72 | 27.66 | 0.298 | 0.071 | 0.131 | -0.169 | 0.792 | 0.213 | 0.110 | 0.817 |
| 28. Today I had new insights about how to better deal with my life circumstances (not at all–very much) | 39.25 | 27.39 | 0.305 | 0.065 | 0.108 | -0.155 | 0.784 | 0.171 | 0.115 | 0.782 |

(Continues)
Today I avoided situations which were related to my problems 38.02 28.65
Today I felt determined to tackle my problems (not at all–very much) 61.02 27.48
Today I was committed to accomplish my goals (not at all–very much) 57.41 27.24
Today I had my goals clearly in mind (not at all–very much) 54.19 28.42
Today my interest in the topics of therapy was very low (not at all–very much) 39.70 27.95
Today I was able to manage my emotions 47.35 25.87
Today I felt confident to approach burdensome issues in my life (not at all–very much) 46.85 28.76
Today I could take a positive view on myself and my progress (not at all–very much) 46.97 27.95
Today I felt confident that I will resolve my issues (not at all–very much) 50.61 27.27

Motivation for change (MOT)

| Item (VAS description) | AM  | SD  | WPE | RFP | TAS | EPI | ICP | MOT | MSC | $h^2$ |
|------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| Today I avoided situations which were related to my problems | 38.02 | 28.65 | -0.053 | -0.243 | 0.036 | 0.280 | 0.045 | -0.267 | 0.068 |
| Today I was able to manage my emotions | 47.35 | 25.87 | 0.470 | 0.124 | 0.056 | -0.409 | 0.373 | 0.232 | 0.359 |

Note. Factor loadings indicate the item–factor association.

Abbreviations: AM, arithmetic mean of all factors and items, with reference to a scale range from 0 to 100; SD, standard deviation; TPQ, Therapy Process Questionnaire; VAS, visual analogue scales.

TABLE 4 Excluded items

| Item (VAS description) | AM  | SD  | WPE | RFP | TAS | EPI | ICP | MOT | MSC | $h^2$ |
|------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| Today I was able to cope with interpersonal conflicts | 50.66 | 28.12 | 0.332 | 0.270 | 0.073 | -0.181 | 0.229 | 0.259 | 0.215 |
| Today I experienced a conflict with the therapist(s) or other members of the team of professionals | 7.72 | 15.82 | 0.113 | -0.213 | -0.181 | 0.200 | 0.149 | -0.067 | 0.038 |
| Today I avoided situations which were related to my problems | 38.02 | 28.65 | -0.053 | -0.243 | 0.036 | 0.280 | 0.045 | -0.267 | 0.068 |
| Today I was able to manage my emotions | 47.35 | 25.87 | 0.470 | 0.124 | 0.056 | -0.409 | 0.373 | 0.232 | 0.359 |

Note. Items were excluded from the list of respected items because of low and insufficiently differentiating factor loadings (cut-off: <.5). The table shows the factor loadings and communalities ($h^2$) of the four excluded items.

Abbreviations: AM, arithmetic mean, with reference to a scale range from 0 to 100; EPI, emotional and problem intensity; ICP, insight/confidence/therapeutic progress; MOT, motivation for change; MSC, mindfulness/self-care; RFP, relationship with fellow patients; SD, standard deviation; TAS, therapeutic alliance and clinical setting; WPE, well-being and positive emotions.

able to cope with interpersonal conflicts." "Today I experienced a conflict with the therapist(s) or other members of the team of professionals," and "Today I avoided situations which were related to my problems." One item was excluded not only because it differentiated insufficiently, but also because it was reported by the patients as unclear and difficult to understand ("Today I was able to manage my emotions"). After these four items were omitted, the TPQ now comprises 43 items.
TABLE 5  | Chi-square and fit indices of the CFA

| Measure | Value |
|---------|-------|
| $\chi^2$ | 22,995.72 (p < .001) |
| TLI     | 0.984 |
| CFI     | 0.985 |
| RMSEA   | 0.049 |
| SRMR    | 0.051 |

Note. Cut-off scores for good model fit: TLI, >.95; CFI, >.95; RMSEA, <.05; SRMR, <.06.

Abbreviations: CFA, confirmatory factor analysis; CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; TLI, Tucker-Lewis Index.

3.1.2  | Confirmatory factor analysis

All fit indices of the CFA (Brown, 2015) confirmed the exploratory factor analysis, as can be seen in Table 5. The $\chi^2$ of the CFA fit is 22,995.72 (p < .001). All criteria (relative fit indices) of our CFA (the TLI, CFI, RMSEA, and SRMR) exceed respectively fall below the thresholds of good model fits (Table 5; compare Baumgartner & Homburg, 1996; Hu & Bentler, 1999).

3.1.3  | Intercorrelations

The intercorrelations between the extracted factors confirm relations that can be expected from psychotherapy research and practical experience (Table 6). The EPI factor is negatively correlated with all other factors, with the most negative correlation ($r = -.57$) realized with the WPE factor. The highest positive correlations are given between WPE and MSC ($r = .72$), between ICP and MOT ($r = .69$), between ICP and WPE ($r = .66$), and between ICP and MSC ($r = .64$). The WPE-RFP correlation is almost twice as high ($r = .40$) as the WPE-TAS correlation ($r = .21$). The negative correlation between EPI and RFP is twice as high as that between EPI and TAS ($r = -.44$ vs. $r = -.18$), but MOT is more highly correlated with TAS ($r = .43$) than with RFP ($r = .27$; for an interpretation of the results, see Section 4).

3.2  | Item analysis and psychometric data

For each subscale of the TPQ, some important psychometric characteristics were calculated: Cronbach's $\alpha$, a measure of the internal consistency of a scale, the mean intercorrelations between items, and the mean discriminative power, which is the mean correlation of each item of a subscale with the total score of the subscale. Whereas the internal consistency and the discriminative power should be high, the item intercorrelations should be in a mean range, because the items of a subscale should not identically represent the same feature of an "object." Table 2 presents the results of the item analysis and subscale characteristics.

Concerning the calculation of Cronbach's $\alpha$, it should be respected that the procedure is based on some restrictive assumptions, that is, unidimensionality, uncorrelated errors, and essentially tau-equivalence (Rammstedt & Beierlein, 2014). On the basis of two items as in the RFP subscale, these assumptions cannot be tested, and, in consequence, no values are reported for RFP in Table 2. Beside the calculation of Cronbach's $\alpha$ based on the complete set of multiple time series of all items over 10.442 measurement points, another calculation was realized for each time point. Because the mean of the time series length of our sample is 70, we calculated the values for 1 to 70 measurement points. In Table 2, the mean Cronbach's $\alpha$ of 70 time series points and the minimum and maximum values out of the 70 calculations are reported.

3.3  | Construct validity

The partial construct validity of the factors (subscales) of the TPQ is demonstrated by the intercorrelations with the subscales of the ISR. Because of the fact that the TPQ is not a measure of symptom severity or of other outcome criteria, the correlations between the subscales of the two questionnaires will only reveal a partial construct validity. However, EPI or, reversely, WPE, may directly correspond to symptom severity, whereas other psychological constructs represented by the subscales of the TPQ may intensify or weaken symptom severity, in other words, may be moderators of

TABLE 6  | Factor intercorrelations

|                   | WPE | RFP | TAS | EPI | ICP | MOT | MSC |
|-------------------|-----|-----|-----|-----|-----|-----|-----|
| Well-being and positive emotions (WPE) | 0.40 | 0.21 | -0.57 | 0.66 | 0.56 | 0.72 |
| Relationship with fellow patients (RFP) | 0.22 | -0.44 | 0.26 | 0.27 | 0.30 |
| Therapeutic alliance and clinical setting (TAS) | -0.18 | 0.31 | 0.43 | 0.26 |
| Emotional and problem intensity (EPI) | -0.28 | -0.30 | -0.44 |
| Insight/confidence/therapeutic progress (ICP) | 0.69 | 0.64 |
| Motivation for change (MOT) | 0.56 |
| Mindfulness/self-care (MSC) | |

Note. The correlations are based on the z-transformed time series of the factors. These time series are produced by including all patients of the sample, which resulted in an artificial time series of 10,442 measurement points. All correlations are significant at $p < .001$. The applied family-wise error correction for multiple comparisons did not change the significance levels of the correlations.
psychopathology. The correlations were calculated between the ISR at intake to the hospital and the mean of the first seven values of the time series of the TPQ factors. Given daily measures, seven measurement points correspond to 1 week, which ensures that we take a more stable level of the dynamics than by taking a single day. A single measure always can represent by chance a very high or a very low occurrence, which would distort the result. As shown in Table 7, the subscales depression, anxiety, eating disorder, the additional scale, and the total score of the ISR were significantly negative correlated with WPE and positively correlated with EPI. ICP, MOT, and MSC were negatively correlated with depression, anxiety, eating disorder, the additional scale, and the total score of the ISR. EPI was positively correlated with the OCD and the eating disorder subscale, and the MSC factor correlated negatively with the somatoform disorder subscale of the ISR. It should be noted that symptoms related to OCD, somatoform, and eating disorders were less important for most of the patients of our sample than depression- and anxiety-related symptoms (compare Table 1). Although we outlined no specific hypotheses on the correlations between the RFP and the TAS factors with the subscales of the ISR, these correlations are shown in Table 7. There are no significant correlations of RFP with the ISR subscales but significant negative correlations of the TAS with the depression subscale and the additional subscale of the ISR. More depressed patients may have a poorer relationship with professionals at the beginning of the hospital stay but not with fellow patients.

For each factor of the TPQ, the time series can be calculated by averaging the respective items of the factors or by calculating z-transformed dynamics on the averaged items. Figure 2 shows the evolution of the factors representing the therapeutic process of a

**TABLE 7** Correlations between the factors of the TPQ and the subscales of the ISR

| Factors | WPE | RFP | TAS | EPI | ICP | MOT | MSC |
|---------|-----|-----|-----|-----|-----|-----|-----|
| Dep     | -0.61** | -0.06 | -0.28** | 0.53** | -0.49** | -0.42** | -0.49** |
| Anx     | -0.39** | 0.03 | -0.06 | -0.37** | -0.19 | -0.21 | -0.29** |
| OCD     | -0.16 | 0.09 | 0.06 | 0.32** | -0.15 | 0.03 | -0.11 |
| Som     | -0.17 | 0.04 | 0.09 | 0.14 | -0.04 | -0.07 | -0.11 |
| Eat     | -0.24** | 0.00 | 0.04 | 0.20 | -0.09 | 0.08 | -0.2 |
| Add     | -0.55** | 0.00 | -0.25 | 0.49** | -0.35 | -0.30 | -0.41** |
| Total   | -0.51** | -0.02 | -0.10 | 0.50** | -0.29** | -0.21 | -0.38** |

**Note.** Intercorrelations between the arithmetic mean of the factor values at the first seven measurement points (beginning of the treatment) and the subscales of the ICD-10 based Symptom Rating (ISR), taken at the first days of hospital stay. Subscales: Dep (depression), Anx (anxiety disorder), OCD (obsessive-compulsive disorder), Som (somatoform disorder), Eat (eating disorder), Add (additional scale), Total (total score).

**Abbreviations:** EPI, emotional and problem intensity; ICP, insight/confidence/therapeutic progress; MOT, motivation for change; MSC, mindfulness/self-care; RFP, relationship with fellow patients; TAS, therapeutic alliance and clinical setting; TPQ, Therapy Process Questionnaire; WPE, well-being and positive emotions.

**FIGURE 2** Time series of the seven factors of the TPQ, illustrated by a single case (104 measurement points). The time series are calculated by averaging the items associated to each factor and then applying a z-transformation, which shows the time series in units of standard deviations on the y-axis around the mean of zero. This is the way how factor dynamics is presented in the SNS. Usually, the z-transformed factor dynamics shows the shape of a process more pronounced than time series resulting only from averaging items.
patient diagnosed with a complex post-traumatic stress disorder. The synchronized order transition during this psychotherapy after about one third of the process can be seen in different factors, for example, EPI, ICP, MOT, and MSC. Other subscales (e.g., TAS) reveal more stable dynamics after an instable period at the beginning, representing stable boundary conditions of the transient destabilization during the self-organizing process.

4 | DISCUSSION

The factor analysis of the TPQ revealed a seven-factor solution including 43 items distributed on the factors (subscales). Four items were excluded for statistical and clinical reasons. The factors address a broad variety of change mechanisms, which are discussed in the scope of common factors research on psychotherapy (“contextual model,” Wampold & Imel, 2015).

One stable and frequently replicated result from this research concerns the importance of the therapeutic alliance as a contributor to change (e.g., Flückiger, del Re, Wampold, & Horvath, 2018; Norcross & Lambert, 2011; Wampold, 2015), which is represented by the factor TAS of the TPQ and, also, though investigated to a minor degree, the RFP. In a clinical context, fellow patients take the role of a social network supporting the patient and his development (e.g., Jörgensen, Römö, & Rundmo, 2009). Although it may seem to represent an artificial setting, the RFP at a hospital ward is a contributor and mediator of change because—among other reasons—intensive social interactions enable social learning (Adler & Stead, 2015). In addition to this, the experience of social relations may be a secondary outcome criterion. Other contributors are positive emotions and well-being (compare the results from positive psychology, for example, Seligman & Csikszentmihalyi, 2000; Sin & Lyubomirsky, 2009), as well as the activation of resources (e.g., Grawe, 2004), represented by the factor WPE. This factor may be seen as a supporting variable for the personal development of a patient and at the same time as an important secondary outcome criterion (Wampold et al., 2017). WPE reinforces change and may be an indicator of successful development. In a theoretical model of therapeutic change (Schiepek et al., 2017), experiences of success and positive emotions (both are covered by the WPE factor) are related to motivation for change, insight, and problem reduction, as well as to the long-term evolution of personality traits (e.g., reduced hopelessness or increased self-efficacy, mindfulness, and competences in emotion regulation; Schöller et al., 2018).

Insight (e.g., Castonguay & Hill, 2007; Grawe, 2004), trust in one’s personal development (including experiences of self-efficacy; e.g., Maddux, 2013), and experienced success reinforcing the therapeutic work (e.g., Catty, 2004) and acting as an agent of self-efficacy are well-known mechanisms of change (represented by the factor ICP). Of course, MOT is an intensely investigated factor contributing to therapeutic progress (e.g., Grawe, 2004; Ryan, Lynch, Vansteenkiste, & Deci, 2011). During the last two decades, mindfulness and positive self-regard (represented by the factor MSC) were recognized as important therapeutic mechanisms (Bateman & Fonagy, 2013, 2015) and created therapy approaches of its own (e.g., Mindfulness-based Stress Reduction; e.g., Grossman, Niemann, Schmidt, & Walach, 2004). Independent of any specific therapeutic approach, mindfulness and self-regard contribute to personal development and stress-reduction in a more general sense. Finally, symptom severity and problem intensity represent a primary outcome criterion, which, in many cases, for example, in all mood and anxiety disorders, is related to worrying and stressful emotions. Both components, symptom severity and negative emotions, are integrated in the factor EPI. In addition to this, many psychotherapeutic approaches emphasize the role of worrying emotions for personal development, for example, during confrontation with anxiety-provoking or trauma-related personal issues, avoidance reduction, or taking in mind suppressed needs, phantasies, and burdensome reminders. Emotion-focused psychotherapy (Greenberg, 2002) specifically focuses on the experience of emotions as a core mechanism of therapeutic change. The theoretical constructs as represented by the factors of the TPQ reflect the importance of interpersonal experiences for personal development (TAS and RFP), as stressed by the contextual model of psychotherapy and by the frequently replicated findings on the contribution of the therapeutic alliance to the outcome (e.g., Flückiger et al., 2018). The TAS and the RFP factors explain a substantial part of the variance of the TPQ (see Table 2) and are significantly correlated to other factors, as EPI and MOT (see Table 6). Five from seven factors are related to the patient and refer to numerous findings, which underline that the patient plays a substantial (if not the main) role in producing the process and creating the outcome of psychotherapy (Boholt & Tallman, 2010; Duncan, Miller, & Sparks, 2004; Orlinsky et al., 2004; Prochaska, Norcross, & DiClemente, 1994).

Some interesting insight into the mechanisms of change is given by the interfactor correlations. The high positive correlations between ICP and MOT, WPE, and MSC underline the important role of insight and confidence in the progress for all other ingredients and contributors to symptom reduction, positive emotions, and well-being—in other words, to therapeutic success and good outcome. The closer connectedness of WPE to the RFP, compared with the relationship with the professional therapists, underlines that in clinical settings (inpatient treatment), the relationship with the fellow patients is important and even may be more important and supporting than the alliance with the professional therapists (e.g., Jörgensen et al., 2009). A cautious interpretation of the higher MOT-TAS correlation ($r = .43$) compared with the MOT-REP correlation ($r = .27$) may be that motivation for change is more supported by professionals than by fellow patients, whereas fellow patients play a greater role for positive feelings and well-being (TAS-EPI: $r = -.18$ vs. RFP-EPI: $r = -.44$; TAS-WPE: $r = .21$ vs. RFP-WPE: $r = .40$).

The factor structure of the TPQ corresponds to the theoretical constructs (variables) of a mathematical model, which explains some essential features of change dynamics in psychotherapy (Schiepek et al., 2017; Schöller et al., 2018). The variables of this model are “therapeutic success,” “motivation for change,” “dysphoric emotions,” “symptom severity,” and “insight.” The factors of the TPQ as presented here are not identical to these variables but very similar. It should be noted that the therapeutic alliance factor is also represented in the
model, not as a variable but as one out of four parameters that modulate the shape of the nonlinear functions interconnecting the variables. The discrete iterations produced by a set of five coupled nonlinear difference equations—each equation representing the dependency of a variable on other variables or on itself—may be interpreted as discrete day-by-day measures of a psychotherapeutic process. In consequence, the validation of the model can be realized by time series as produced by the TPQ (Schöller et al., 2019). The combination of data-driven computer simulations and time series data available at "real time" opens the way for further developments in artificial intelligence and process control in psychotherapy.

Equidistant high-frequency measures as realized by the TPQ allow for the application of time series analysis methods, which can be used for the identification of nonstationarities and pattern transitions in human change processes. One of these methods is dynamic complexity, which combines the amplitude, the frequency, and the distribution of the measurement values over the available scaling range (Haken & Schiepek, 2010; Schiepek & Strunk, 2010). It is calculated within a gliding window, which runs over the complete time series. Dynamic complexity may be characteristic for specific dynamics (e.g., the emotional instability of borderline personality disorder) or may be a precursor of phase transitions (critical instability). Another feature of psychopathological dynamics is increased or decreased synchronization of emotions and cognitions, which can be seen in colour-coded inter-item correlation matrices calculated in running windows over multiple time series. Locally increased synchronization also may be a precursor of phase transitions (Haken & Schiepek, 2010; Scheffer et al., 2009). Recurrence plots portray recurrent patterns and transients within a time series in a time × time diagram (Eckmann, Oliffson Kamphorst, & Ruelle, 1987; Webber & Zbilut, 1994).

The complexity of trajectories of change, which is mirrored by these analysis methods, is reflected by the fact that the trajectories do not move on standard tracks (Schiepek et al., under review). Compared with standard tracks, which are produced by averaging time series of patients with the same diagnosis or similar symptom severity at intake, patterns of change as assessed by the TPQ are much more complex, unpredictable, and chaotic. Deviations from standard tracks do not predict poor outcome (Schiepek et al., under review); on the contrary, they result from order transitions and critical instabilities, which are necessary conditions for success (Haken & Schiepek, 2010).

4.1 Strengths

The data of our study was taken from naturalistic inpatient settings. Patients with a range of diagnoses, which is representative for inpatient treatment, filled in the TPQ during the complete period of their hospital stay. The compliance rate was high, resulting in a very low percentage of missing data. The factor structure, which could be extracted from the covariance matrix of all items, was clear and easy to interpret. The cumulated explained variance of the factors is comparatively high, and the item, as well as subscale characteristics (internal consistency, inter-item correlations of the subscales, and discriminative power of the items), may be declared as excellent. The CFA clearly verified the exploratory factor analysis. The items and the factors of the TPQ correspond to broadly accepted and validated constructs on mechanisms of change in psychotherapy.

Other than outcome measures, which usually focus on symptom severity and, by this, are counterproductive for daily assessments because patients continuously would be minded on problems instead of personal development, many items of the TPQ refer to resources, competencies, or experiences, which could drive progress. This makes the continuous high-frequency application of the TPQ an auto-catalysing therapeutic intervention of its own, which should be tested in further studies.

4.2 Limitations

The factor analysis was based on the time series data of 150 patients. Of course, the sample size could have been larger, and, in consequence, a replication study based on a bigger sample of not only inpatients but also outpatients should be realized. This is intended and realistic because of the fact that the TPQ is applied in different hospitals, treatment centres, and therapeutic practices. A disadvantage may be seen in the unequal assignment of the items to the factors, with a range from two items in the RFP factor and 10 items in the ICP factor. However, this is an empirical and statistical result, which has to be recognized and cannot be manipulated. In the strictest sense, the results potentially do not apply to the English translation we present in the manuscript but to the German version of the TPQ. Consequently, future psychometric investigations of the English version of the TPQ are of importance and should be realized by an independent replication and validation study.

A 43-item self-report for daily assessment from one perspective may be too labour-intensive; from another perspective, it would be not intensive enough to match the expectations of "high-frequency" monitoring. However, years of experience in different clinical settings proved that patients not only accept this investment of time but also appreciate it as a quiet moment of self-reflection, mentalization, and focusing their own personal development. On the other hand, higher frequencies of self-reports (e.g., by event sampling) would interrupt the day, focus the attention too much on the task of self-rating, and finally produce more missing data in the case of losing one day of feedback. Given the claim of a full assessment of a psychotherapeutic process, daily measures have revealed to be a good compromise.

4.3 Perspectives

The TPQ is not restricted to specific disorders or treatment approaches. None of the items picks up specific symptoms of specific diagnoses. On the contrary, the questionnaire addresses different psychological constructs and general mechanisms of change. In consequence, it is not only a transdiagnostic measure but also independent
from specific therapeutic concepts, techniques, or psychotherapeutic confessions. As experience shows, therapists and treatment centres working with completely different methods can use the TPQ-based process monitoring.

The use of the TPQ in outpatient settings is possible if the items representing the RFP (evidently, fellow patients do not exist in outpatient psychotherapy) and representing the therapeutic alliance will be omitted. Outpatient psychotherapy usually has a rhythm of weekly or bi-weekly sessions, and it is useless to ask for the quality of the professional relationship each day. All other items can be applied to the monitoring of outpatient psychotherapy.

The TPQ standardizes the sampling rate because it imposes self-assessments at a daily rhythm. This is the reason why the TPQ can be used in very different inpatient, day-treatment, or outpatient settings for routine monitoring and, by this, allows for a direct comparison of very different treatment cases and processes. Its usability not only in studies but in everyday routine practice creates big data sets, which can be analysed for features of the dynamics and their relation to the outcome. Process features such as sudden gains or losses, other kinds of dynamic or order transitions, crisis-repair sequences, critical instabilities, changing synchronization patterns, or early drop-outs can be assessed and even defined in a general way. The definition, assessment, and analysis of such phenomena requires a standardized high-frequency sampling rate.

The option of a differentiated access to dynamic patterns not only allows for new target concepts in diagnostics, such as the diagnostics of stability or instability of the process, of dynamic connectivity and transient synchronization patterns, or of precursors of critical transitions, but also for more general definitions of outcome. Traditionally, we focus on pre–post measures of primary or secondary outcome criteria, which can now be complemented by criteria of changed dynamic patterns, for example, reduced pathological oversynchronization of cognitions and emotions (e.g., Schiepek, Stöger-Schmidinger, et al., 2016), increased flexibility and adaptability of psychological processes ( Kashdan & Rottenberg, 2010 ), or modified emotional dynamics, for example, from emotional instability to stability ( e.g., in borderline personality disorder), or the other way round, from emotional rigidity to flexibility ( e.g., in major depressive disorder, alexithymia, or addictive disorders).

As we noted, the TPQ was developed for applications in routine real-time monitoring by electronic devices, using internet and app-based systems (e.g., the SNS). By this, it can be seen as an important step towards the digitalization of psychotherapy. In psychotherapy 4.0, computer-assisted methods open new ways for treatment control, just in time applications of interventions, combining web-based and face-to-face treatments, or integrated care linking treatments over different settings (e.g., inpatient–outpatient). Certainly, psychotherapy 4.0 will not be the solution for everything, but combined with new developments in artificial intelligence, it will contribute to the development of useful concepts and help to bridge challenging gaps, for example, between personalized treatment and big data, idiographic and nomothetic approaches, or process control and the personal responsibility of patients.

ACKNOWLEDGEMENTS
We are grateful to David Pincus, PhD, (Chapman University) for a detailed check of the wording of the items that were translated from German to English language.

AUTHORS’ CONTRIBUTIONS
G.S. developed the original version of the TPQ, realized the revision together with H.K. and B.S.-S., wrote the paper, and supervised the study and the statistical procedures. H.S. and L.K. realized the statistical analysis. K.V. supervised and checked the statistical procedures. A. L.-A. and G.S. translated the items into English language. L.K., P.H., W. A., B.S.-S., and H.K., together with their teams, were responsible for the data collection in different treatment centres.

ORCID
Kathrin Vial https://orcid.org/0000-0003-3482-801X
Anna Lichtwarck-Aschoff https://orcid.org/0000-0002-4365-1538

REFERENCES
Adler, N. E., & Stead, W. W. (2015). Patients in context—EHR capture of social and behavioral determinants of health. New England Journal of Medicine, 372, 698–701. https://doi.org/10.1056/NEJMp1413945
Alexander, L. B., & Luborsky, L. (1986). The Penn Helping Alliance Scales. In L. S. Greenberg, & W. M. Pinsof (Eds.), The psychotherapeutic process: A research handbook (pp. 325–366). New York, NY: Guilford Press.
Bateman, A., & Fonagy, P. (2013). Mentalization-based treatment. Psychoanalytic Inquiry, 33, 595–613. https://doi.org/10.1080/07351690.2013.835170
Bateman, A., & Fonagy, P. (2015). Borderline personality disorder and mood disorders: Mentalizing as a framework for integrated treatment. Journal of Clinical Psychology, 71, 792–804. https://doi.org/10.1002/jclp.22206
Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modelling in marketing and consumer research: A review. International Journal of Research in Marketing, 13, 139–161. https://doi.org/10.1016/0167-8116(95)00038-0
Bech, P., Gudex, C., & Johansen, K. S. (1996). The WHO Well-being Index: Validation in diabetes. Psychotherapy and Psychosomatics, 65, 183–190. https://doi.org/10.1159/000289073
Bentler, P. M. (1990). Comparative fit indexes in structural models. Psychological Bulletin, 107(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238
Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88, 588–606. https://doi.org/10.1037/0033-2909.88.3.588
Bohart, A. C., & Tallman, K. (2010). Clients: The neglected common factor in psychotherapy. In B. Duncan, S. Miller, B. Wampold, & M. Hubble (Eds.), 2010The heart and soul of change (2nd. ed.) (pp. 83–111). Washington DC: American Psychological Association.
Brottland, H., Koksvik, J. M., Burkeland, O., Grøwe, R. W., Klöckner, C., Linaker, O. M., ... Iversen, V. C. (2018). The effects of routine outcome monitoring (ROM) on therapy outcomes in the course of an implementation process: A randomized clinical trial. Journal of Counseling Psychology, 65(5), 641–652. https://doi.org/10.1037/cou0000286
Brockmann, J., Kirsch, H., Hatcher, R., Andreas, S., Benz, S., & Sammet, I. (2011). Dimensionen der therapeutischen Beziehung aus Patienten
introduction to complexity in human behavior and cognition]. München: Elsevier.

Tritt, K. (2015). ICD-10-Symptom-Rating (ISR) - das Handbuch zum Fragebogen (Handbook of the ICD-10-based Symptom-Rating [ISR]). Berlin: Neobooks.

Tritt, K., von Heymann, F., Zacharias, I., Söllner, W., & Löw, T. (2008). Entwicklung des Fragebogens "ICD-10-Symptom-Rating" (ISR) [Development of the Questionnaire "ICD 10 Symptom Rating" (ISR)]. Zeitschrift für Psychosomatische Medizin Und Psychotherapie [Journal for Psychosomatic Medicine and Psychotherapy], 54, 409–418. https://doi.org/10.13109/zptm.2008.54.4.409

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. Psychometrika, 38, 1–10. https://doi.org/10.1007/BF02291170

Tzur-Bitan, D., Meiran, N., & Shahar, G. (2010). The importance of modeling comorbidity using an intra-individual, time-series approach. Behavioral and Brain Sciences, 33(2–3), 172–173. https://doi.org/10.1017/s0140525x10000816

Tzur-Bitan, D., Meiran, N., Steinberg, D. M., & Shahar, G. (2012). Is the looming maladaptive cognitive style a central mechanism in the (Generalized) Anxiety–(Major) Depression comorbidity: An intra-individual time series study. International Journal of Cognitive Therapy, 5, 170–185. https://doi.org/10.1521/iict.2012.5.2.170

Wampold, B. (2015). Routine outcome monitoring: Coming of age with the usual developmental challenges. Psychotherapy, 52, 458–462. https://doi.org/10.1037/pst0000037

Wampold, B. E., Flückiger, C., del Re, A. C., Yulish, N. E., Frost, N. D., Pace, B. T., ... Hilsenroth, M. J. (2017). In pursuit of truth: A critical examination of meta-analyses of cognitive behavior therapy. Psychotherapy Research, 27, 14–32. https://doi.org/10.1080/10503307.2016.1249433

Wampold, B. E., & Imel, Z. E. (2015). The great psychotherapy debate: The evidence for what makes psychotherapy work (2nd ed.). New York: Routledge.

Ware, J. E., Snow, K. K., Kosinski, M., & Gandek, B. (1993). SF-36 Health Survey: Manual and interpretation guide. Boston: The Health Institute, New England Medical Centre.

Webber, C. L., & Zbilut, J. P. (1994). Dynamical assessment of physiological systems and states using recurrence plot strategies. Journal of Applied Physiology, 76, 965–973. https://doi.org/10.1152/jappl.1994.76.2.965

Wing, J., Beevor, A. S., Curtis, R. H., Park, S. B. G., Hadden, S., & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS): Research and development. British Journal of Psychiatry, 172, 11–18. https://doi.org/10.1192/bjp.172.1.11

How to cite this article: Schiepek G, Stöger-Schmidinger B, Kronberger H, et al. The Therapy Process Questionnaire - factor analysis and psychometric properties of a multidimensional self-rating scale for high-frequency monitoring of psychotherapeutic processes. Clin Psychol Psychother. 2019;26:586–602. https://doi.org/10.1002/cpp.2384