Crisis-repair-sequences - considerations on the assessment and classification of breaches in the therapeutic relationship

Authors: Antje Gumz*¹, Elmar Brähler², Denise Kästner¹, Rainer Erices³

Email addresses: antje.gumz@web.de; elmar.braehler@medizin.uni-leipzig.de; denise-kaestner@gmx.de; rainer@erices.de

Institutional addresses:

¹ Department of Psychotherapy and Psychosomatic Medicine, University of Leipzig, Germany

² Department of Medical Psychology and Medical Sociology, University of Leipzig, Germany

³ Institut für Geschichte und Ethik der Medizin, Friedrich-Alexander-Universität Erlangen-Nürnberg

* Corresponding author

Antje.Gumz@web.de

Department of Psychotherapy and Psychosomatic Medicine, University of Leipzig, Semmelweisstraße 10, 04103 Leipzig, Germany, Tel: +49 163 5026408
Abstract

Background: Recent research indicates that temporary deteriorations of variables monitored continuously in the course of the therapeutic relationship are important characteristics of psychotherapeutic change. These so-called “rupture-repair” episodes were assessed and put in relation to the outcome of therapy using different mathematical methods. Presently, there is no standardized procedure to identify such episodes in the therapeutic relationship.

Methods: The study deals with the criteria for identifying rupture-repair episodes that have been established so far. It proposes modifications of these criteria which prospectively could make it possible to identify rupture-repair episodes more precisely and consistently. The authors developed an alternative criterion. This crisis-repair criterion is able to include rupture patterns which had not been considered before, as well as characterizing the length of the rupture-repair episodes. As a model application, the different criteria were used with continuous assessments of the therapeutic interaction in psychodynamic therapy courses (10 courses averaging 31.5 sessions, 1 course over 200 sessions).

Results: The analysis revealed that the number of the identified rupture-repair episodes differed depending on the criterion that was used. With the proposed criterion, considerably more episodes were identified. Moreover, the authors showed that there are different patterns of rupture-repair sequences in therapy processes. The most frequent pattern was the simple V-shape. The second most common pattern was a decline over more than one session with a sudden repair. The longest downward trend comprised a period of six sessions.

Summary: The findings of the study give insight into basic mechanisms of change within the therapeutic relationship. The implications of these mechanisms for future research aims were discussed. The authors propose to use the new developed criterion to standardize the identification of crises within the psychotherapeutic process.
Background

In recent years, more and more authors have focused on discontinuous changes in psychotherapy processes. Especially in terms of symptom courses, there is considerable evidence to suggest that change within psychotherapeutic processes occurs in a nonlinear, discontinuous manner. Different types of discontinuous change are described in the literature. These include abrupt improvements (mostly referred to as “sudden gains” or “rapid early responses”) and deteriorations (“sudden losses”) from one session to another (for review see [1;2]). Another type of discontinuous change is V-shaped, representing an abrupt deterioration in the trajectory of measured variables followed by a subsequent return to the previous or a higher level. Hayes et al. [3] observed this type of discontinuity in addition to “sudden gains” in the progression of depressive symptom trajectories.

Numerous studies have provided evidence linking the quality of the therapeutic relationship to the outcomes of all treatment modalities [4]. The studies relate to different components of the therapeutic relationship such as the working alliance, the transference–countertransference configuration, or the real relationship [5;6]. Far fewer authors have examined the developments of the therapeutic relationship over the entire therapeutic process focusing on characteristic features of change. Accordingly, discontinuous changes in the trajectories of the therapeutic relationship across the whole process have rarely been analyzed until now. Recent research indicates that temporary deteriorations of variables monitored continuously in the course of the therapeutic relationship are important characteristics of psychotherapeutic change. These so-called “rupture-repair” episodes were assessed and put in relation to the outcome of therapy using different mathematical methods. Presently, there is no standardized procedure to identify such episodes in the therapeutic relationship.

The existing studies regarding this topic [7;8;2;9] established a relation to the concept of ruptures in the therapeutic alliance. Safran et al. [10;4] discussed at length the importance
of investigating alliance ruptures. They suggested that alliance ruptures inevitably occur in treatment courses and that these breaches can provide an opportunity for therapeutic change. Stiles et al. [7] provided “the first nomothetic (statistical) demonstration of this theoretically expected effect” (p. 89) in longitudinal data. They showed that patients whose course of alliance was characterized by rupture-repair sequences made larger gains in treatment compared to other patients. Similarly, Strauss et al. [8] ascertained that most of the patients who reported rupture-repair episodes also reported symptom reductions of 50% or more in all outcome measures.

Prior to this, researchers had investigated the course of the therapeutic relationship on the basis of aggregated data by applying cluster analysis procedures. Different change profiles of alliance ratings (linear und stable pattern in the majority of cases) and their connection to the therapeutic outcome are described in the literature [11;12]. Kivlighan and Shaughnessy [13] found three patterns of alliance development in the course of counseling processes: the stable, linear and quadratic patterns. The quadratic pattern correlated with a positive outcome. Stiles et al. [7] wanted to replicate the latter findings on the alliance curves of 79 depressed patients (psychodynamic interpersonal and cognitive-behavioral short-term therapies, 2 groups - 8 or 16 sessions). They revealed that V-shaped strong deteriorations with subsequent improvements to the previous or a higher level (alliance rupture-repair episodes) proved crucial in their sample, rather than the quadratic pattern. In 2006, Strauss et al. [8] observed alliance rupture-repair episodes in 30 patients with obsessive-compulsive and avoidant personality disorder in cognitive-behavioral courses of therapy.

Gumz et al. [2;9] examined interaction variables over a single course of 125 sessions and over six courses of 29 to 35 sessions of psychodynamic therapy. The studies were based on the assumption that there is a temporal association between high instability of the examined variables and discontinuous change [14]. (It is possible that there is a relationship between
discontinuous changes and the theory of self-organizing systems. Several authors modeled psychotherapy as a dynamic process with stable and critically unstable episodes and focused on the existence of more or less pronounced fluctuations [e.g. 15; 16; 17]). Within both the longer and the shorter processes, the authors identified several sequences of stable and unstable episodes. They registered a temporal connection between abrupt improvements to a higher mean score of interaction variables and local instability maxima and noticed that the unstable episodes in each case were characterized by one or more pronounced negative slumps (V-shape). Persistent discontinuous improvements in interaction perceptions thus were preceded by temporary abrupt strong deteriorations.

We want to mention that we only refer to those studies which identified the ruptures a) with the aid of self-report assessments, b) indirectly, on the basis of the differences in the value of variables of the therapeutic relationship and c) on a macro level between consecutive sessions (discontinuities in therapeutic processes). There are other authors who focused on the micro level, directly, in which patients or therapists assessed whether alliance ruptures and their resolutions occurred within a single session [18; 19], indirectly, via instruments measuring aspects of the therapeutic relationship within a single session [10], or external observers rated the in-session ruptures and resolutions [10; 20].

Aims of the present study

In the context of our process research, we closely explored every detail of individual process characteristics in the course of psychodynamic processes. One or more temporary and more or less pronounced deteriorations in the experienced therapeutic interaction played a role in the majority of the processes. The above-mentioned authors who examined the V-shaped discontinuities in the course of variables of the therapeutic alliance [7;8] used different criteria to identify these rupture-repair sequences. We found differences in the mathematical construction of these two criteria which made it difficult to decide which of them is better
suited for selecting the most relevant rupture-repair episodes in the analyzed curves. We noticed that the criteria do not fully capture specific patterns of change. We observed that the rupture-repair episodes can be characterized with respect to their length and precise curve shape. In addition, when attempting to apply the criteria we were confronted with a deficiency in the criteria definition. From these observations and preliminary considerations we derived the objectives of the present study:

(a) We explain and compare the methods previously used for measuring rupture-repair episodes in longitudinal data [7;8].

(b) We consider modifications of these original criteria in order to standardize the procedures.

(c) We suggest an alternative criterion. Using this criterion it is possible to include rupture patterns which had not been identified before as well as characterize the length of the rupture-repair episodes.

(d) We demonstrate the differences between the criteria and the effect of the modifications using several courses of individual psychodynamic therapy as a model.

(e) We classify different patterns of rupture-repair sequences that may occur in therapy processes. We examine the frequency at which these patterns occur and the length of the episodes.

(f) We give recommendations on how to use these criteria.

Methods

Methods previously used for measuring V-shaped changes

*The criterion developed by Stiles et al.*

Stiles et al. calculated the following four parameters for each individual course using regression analysis: Intercept at the time of midtreatment (midtreatment intercept), the linear trend (slope), the quadratic trend (curve), and the variability (root-mean-square-error, RMSE,
which can be understood as the deviation of the raw values from the fitted curve). In order to extract typical change profiles, they computed a cluster analysis using these four parameters. They defined a rupture as “an unusually low score” and determined that there is a rupture-repair episode if an observed value falls at least two times the RMSE below the value predicted by the fitted curve. This general definition was specified by four criteria: 1.) Low scores in the first or the last session are not considered a rupture. 2.) Patients who showed a negative linear trend curve were excluded from the analysis due to the reason that ruptures in the course of a generally deteriorating alliance are never considered to be repaired. 3.) The value of a rupture session has to be less than the value of the previous session in order not to mark ruptures in generally increasing curves. 4.) The value must fall below a pre-specified limit to be defined as a rupture. This specification was necessary for the following reason: Relatively stable courses show a low intraindividual variability (a low RMSE). Thus, a comparatively small deviation from the fitted curve is sufficient to be defined as a rupture. The restriction avoids identifying ruptures in stable courses with permanently high values.

**The criterion developed by Strauss et al.**

The criterion developed by Strauss et al. includes specifications for the rupture and for the repair. Strauss et al. calculated the standard deviation for all individual courses and averaged these scores. They identified a rupture-repair episode if they could find a decrease and a subsequent increase in the curve which were both at least equal to the averaged standard deviation. They added that this may not be followed by an unrepaired decrease of the same magnitude. Strauss et al. did not explicitly label this latter specification as an exclusion criterion; nonetheless, it implies that repaired ruptures should be ignored entirely if there is an unrepaired rupture afterwards.

**Modifications and supplementations of the criteria for measuring the V-shaped changes**
Four modifications are presented in the following sections. We applied these modifications to both criteria. For this, we adapted the criterion developed by Stiles et al. to the criterion developed by Strauss et al. and vice versa. The effect of the modifications is demonstrated in the section entitled “Model application of the criteria”.

**Modification 1 - Threshold for identifying ruptures**

We adapted the threshold for identifying ruptures of the Stiles et al. measure (two times the RMSE) to the Strauss et al. measure (one standard deviation), meaning that we identified ruptures which are only one time the RMSE below the predict.

Similarly, we adapted the threshold of the Strauss et al. measure to the Stiles et al. measures; i.e, we examined the data with two standard deviations in addition to the originally proposed constant of one standard deviation.

**Modification 2 - Interindividual and intraindividual variability**

Stiles et al. considered the intraindividual variability only. As a result, minor fluctuations occurring in relatively stable courses might be overrated and misclassified as rupture-repair episodes. Stiles et al. avoided this problem by establishing the additional condition that the value must be under a predetermined limit to be defined as a rupture. However, it remained unclear how they determined the cut-off. It seemed somewhat arbitrary and limits the application to the questionnaire used by Stiles et al. (Agnew Relationship Measure). Moreover, this procedure implies that slight downward trends in stable courses featuring permanently low values would also be marked as ruptures. To solve this problem, we combined the intraindividual with the interindividual variability: We computed the mean of all individual root-mean-square-errors and compared it with the interindividual RMSE. In low variability profiles (intraindividual < interindividual RMSE) we applied the interindividual root-mean-square-error. In high variability profiles (intraindividual > interindividual RMSE) we retained the intraindividual RMSE. The latter is important to avoid
overestimating the number of ruptures in courses which can be better described as permanently fluctuating.

Strauss et al. considered the interindividual variability only. As a result, intraindividually less significant declines in highly fluctuating courses could be marked as ruptures. We applied the intraindividual variability in cases in which the intraindividual variability was higher than the interindividual variability in order not to overestimate the number of ruptures in highly fluctuating courses.

(Within the long term course of 200 therapy sessions, we used the intraindividual value solely (SD 41.46) in order not to merge the values of therapies with different lengths. The interindividual averaged SD value of the subsample with 29-35 therapy sessions was 38.69.)

**Modification 3 - Exclusion criteria**

We ignored the second exclusion criterion of Stiles et al. and included patients with a negative linear trend. Contrary to the specifications of Strauss et al., we further included the rupture-repair episodes of those courses excluded with the original criterion due to an unrepaired rupture at the end of treatment.

**Modification 4 - Adjusting the repair value to the rupture value**

In the original Strauss et al. criterion, decreases extending over two standard deviation scores or more are considered to be repaired if there is a subsequent increase of just one standard deviation score. This may result in an inaccurate evaluation of the endpoint of an episode, with the misclassification of an insufficient increase as a full repair. Therefore, we determined that decreases which reach or exceed the value of two standard deviations should only be considered to be repaired if the following increase reached the value of two standard deviations as well.

**Supplements**
We found the following deficiency in the definition of the criteria: Stiles et al. as well as Strauss et al. provided no information as to how to deal with cases in which two or more consecutive decreases fall below the threshold. We decided to classify these cases as one rupture-repair episode only and to aggregate all session-to-session differences after the first rupture mark.

**The criterion developed by our research group (crisis-repair criterion)**

We developed an alternative criterion which is able to include rupture patterns which had not been considered before as well as the length of the rupture-repair episodes. It is based on summing up the differences of session-to-session values until the point where the direction changes is reached. (Constant values are not regarded as a direction change. Sessions with constant values are classified as belonging to the rupture, instead of belonging to the repair. This specification may be of importance when determining the length of a rupture or repair.) We determined that a rupture must be identified as taking place if the summation value of a decline reaches at least one standard deviation. With this procedure, it is possible to take into account ruptures that develop in the form of gradual downward trends. Moreover, this procedure allows us to determine the beginning and the end of a decline more exactly. We defined a rupture as repaired if there is a positive difference between the start value and a subsequent value or a difference value that lies within one standard deviation. As described in the "Modification 2" paragraphs we compared the intraindividual with the interindivial standard deviation; in each case, we favored the stricter of the two values. We referred to the suggested alternative criterion as the “crisis-repair criterion” to imply that we do not only include leaps (ruptures), but also take gradual downward trends into account.

**Model application of the criteria**
In the following section we show the differences of the results between the original criteria. We also demonstrate the effect of the modifications and the results which were assessed by the criterion developed by our research group. The calculations on the basis of the described sample are intended to illustrate the problem. As we do not aim to draw clinical conclusions from the calculations, we keep the sample description short.

**Sample, measuring instrument and data collection**

The analysis is based on ten courses of depth psychology-based psychotherapy (a primarily conflict-centered approach with limitation of the treatment goal and restriction of regressive processes) with 29 to 35 sessions and one 200-session extract from an analytic long-term therapy course. The subsample with 29-35 therapy sessions (average 31.5) included 10 female patients, aged 20 to 40 years. Patients filled in the SCID-II patient questionnaire [21]. They were diagnosed according to ICD-10-GM definitions [22] after the clinical interview. All of them were diagnosed with depression and personality disorder (four of them with narcissistic personality disorder, three with dependent personality disorder, two with avoidant personality disorder and one with obsessive-compulsive personality disorder). The therapies were conducted by 6 female and 2 male therapists (aged 33 to 45 years; between 4 and 17 years of practical professional experience). The patient of the long-term therapy was male, 38 years old at the beginning of therapy, and was diagnosed with panic disorder, agoraphobia, depression, and dependent personality disorder. He was treated by a female therapist (aged 35 years, 9 years of practical professional experience).

Exclusion criteria included a bipolar or psychotic disorder, suicidal tendencies, and the use of psychotropic medication. Each of the 50-minute sessions took place with the patient and therapist seated opposite one another. Patients’ conflicts were addressed, with attention paid to transference and countertransference reactions; therapists incorporated
countertransference phenomena into their interventions and, among other things, also focused on the here and now of the therapeutic relationship.

Patients were extensively informed prior to commencing therapy and provided written informed consent concerning the use of their data for research purposes. Procedures for this study were approved by the University of Leipzig Ethics Commission.

We used the Intrex questionnaire [23;24] for continuous assessment of the therapeutic interaction. The Intrex was developed on the basis of the “Structural Analysis of Social Behavior” model [25] which is one of the most influential approaches for the categorization of interpersonal interactions. SASB has been subject to considerable empirical examination and wide application [26]. It is a circumplex model which contends that variables that measure interpersonal relations are arranged around a circle in a two-dimensional space. A circle is defined by a horizontal dimension of affiliation and a vertical dimension of interdependence. In the model there are three circles, one for each of three foci of action: The transitive focus captures behavior with which an actor attempts to influence an interaction partner, the intransitive focus describes the reactive behavior of the actor, and the level of the introject reflects the way in which a person interacts with him/herself.

Patients completed the short form of the questionnaire immediately after each therapy session. In each case, they rated the relationship behavior for the transitive and intransitive focus and in two different directions (“How did I behave towards my therapist in today’s session?” and “How did my therapist behave towards me in today’s session?”). Eight items exist for each focus. These items assess the interaction on a 10-point interval scale ranging from 0 (“never/not at all”) to 100 (“always/completely”). (To get an impression of the items patients rated, the English version of the Intrex short form is provided in the Appendix). We combined the Intrex scores of items 2 to 4 and 6 to 8 for each of the 2 directions and foci to form the weighted affiliation index (see Appendix). This was proposed by Pincus et al., who
showed that a weighted sum of the clusters has attractive distribution characteristics as well as good validity [27]. A high affiliation index signifies an experience of affectionate interaction between patient and therapist.

Results

Original criteria of Stiles et al. and Strauss et al.

Table 1 contains the number of ruptures measured with the original criteria of Stiles et al. and Strauss et al. and the number of courses with at least one rupture-repair episode. The total of identified ruptures is considerably different. Using the criterion of Stiles et al., we found 12 ruptures in 7 courses, compared to 23 ruptures in 7 courses with the criterion of Strauss et al. Accordingly, we observed a discrepancy in the course of the long-term therapy (4 vs. 27 ruptures).

Three courses either had to be excluded or did not contain a rupture using either the Stiles et al. or the Strauss et al. criteria. Two of them were the same. (In the Stiles et al. procedure two patients had to be excluded because of an overall negative linear trend. Within one course no rupture-repair episode could be identified. This course was highly fluctuating, resulting in a high RMSE and hence a threshold below which no value fell. In the Strauss et al. procedure one case had to be excluded due to an unrepaired rupture in the last session. The other two cases were relatively stable and did not contain a rupture-repair episode.)

Table 1

<table>
<thead>
<tr>
<th>Subsample with 29-35 therapy sessions (n = 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total ruptures/ RREs/ CREs</td>
</tr>
<tr>
<td>Stiles</td>
</tr>
<tr>
<td>12</td>
</tr>
</tbody>
</table>
Courses with at least one rupture/ RRE/ CRE  | 7  | 7  | 9  | 7  
Courses with no rupture/ RRE/ CRE  | 1  | 2  | 1  | 3  
Excluded courses  | 2  | 1  | -  | -  

|  | Single case with 200 therapy sessions (n = 1) |  |
|-----------------|---------------------------------------------|
| Total ruptures/ RREs/ CREs | 4  | 27  | 34  | 9  |

|  | Complete sample (n = 11) |  |
|-----------------|---------------------------|
| Total ruptures/ RREs/ CREs | 16  | 50  | 69  | 24 |
| Courses with at least one rupture/ RRE/ CRE  | 8  | 8  | 10  | 8  
| Courses with no rupture/ RRE/ CRE  | 1  | 2  | 1  | 3  
| Excluded courses  | 2  | 1  | -  | -  |

Note. RRE = Rupture-repair episode; CRE = Crisis-repair episode.

aStiles = original criterion by Stiles et al. [7], two RMSEs; bStrauss = original criterion by Strauss et al. [8], one SD; cCrisis-repair criterion one SD = original version; dCrisis-repair criterion two SDs = modified version.

### Effects of the modifications

Table 2 and Table 3 show the results of the comparison of the different modifications.

**Modification 1 - Threshold for identifying ruptures**

Reducing the threshold of the Stiles et al. criterion to one root-mean-square-error below the fitted curve, we identified 26 rupture-repair episodes within 8 of the shorter courses (instead of 12 ruptures in 7 courses with the original criterion). Within the long-term course we identified 26 ruptures (instead of 4 with the original criterion).

When elevating the Strauss et al. criterion to 2 standard deviation scores, we detected 9 ruptures within 6 courses (compared to 23 in 7 courses, using the original criterion). Within the long-term course we identified 6 ruptures (instead of 27 with the original criterion).

As expected, the total of ruptures counted with the modified Stiles et al. and the original Strauss et al. criterion converged when the thresholds were adapted.

**Modification 2 - Interindividual and intraindividual variability**
When applying the combination of intraindividual and interindividual variability, we detected 9 ruptures within 6 courses following the Stiles et al. procedure (compared to 12 ruptures within 7 courses with the original criterion). Using the Strauss et al. criterion, the total number of ruptures remained the same.

Modification 3 - Exclusion criteria

After ignoring the exclusion criteria, the total number of ruptures assessed by the Stiles et al. criterion increased from 12 to 14. In each of the two previously excluded courses, one rupture could be detected. When looking at the Strauss et al. criterion, the number changed considerably (31 with the modified vs. 23 episodes with the original criterion). This pronounced difference was due to the fact that eight ruptures could be detected in the previously excluded course that was highly fluctuating, showing a high intraindividual variability. (If this course had not been excluded, the application of Modification 2 would have shown a more pronounced difference.)

Modification 4 - Adjusting the repair value to the rupture value

As a result of the application of this modification, the number of rupture-repair episodes dropped by three and one more patient was excluded in the shorter courses (20 ruptures in 6 courses, compared to 23 in 7 courses with the original criterion). Within the long-term course we identified 21 ruptures (instead of 27 with the original criterion).

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Comparison of the Stiles et al. criterion with its modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Original criterion</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Subsample with 29-35 therapy sessions (n = 10)</td>
<td></td>
</tr>
<tr>
<td>Total ruptures</td>
<td>12</td>
</tr>
<tr>
<td>Courses with at least one rupture</td>
<td>7</td>
</tr>
</tbody>
</table>
Courses without rupture  
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>0</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
</table>
Excluded courses  
|   | 2 | 2 | 2 | - |

Single case with 200 therapy sessions (n = 1)

|   | 4 | 26 | - | - |

\(^a\)Modification to one RMSE threshold for identifying ruptures. \(^b\)Including the interindividual in addition to the interindividual variability favoring the individually stricter value (applied to the subsample with 29-35 therapy sessions only). \(^c\)Inclusion of cases excluded with the original criterion because of a negative linear trend.

### Table 3

#### Comparison of the Strauss et al. criterion with its modifications

<table>
<thead>
<tr>
<th></th>
<th>Original criterion</th>
<th>Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1(^a)</td>
</tr>
<tr>
<td>Subsample with 29-35 therapy sessions (n = 10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total RREs</td>
<td>23</td>
<td>9</td>
</tr>
<tr>
<td>Courses with at least one RRE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Courses without RRE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Excluded courses</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Single case with 200 therapy sessions (n = 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total RREs</td>
<td>27</td>
<td>6</td>
</tr>
</tbody>
</table>

Note: RRE = Rupture-repair episode.

\(^a\)Modification to one-SD threshold for identifying ruptures. \(^b\)Including the intraindividual SD in addition to the interindividual SD favoring the individually stricter value (applied to the subsample with 29-35 therapy sessions only). \(^c\)Inclusion of cases excluded with the original criterion because of an unrepaired rupture at the end of treatment. \(^d\)Adjusting the repair value to the rupture value.

### Findings on the proposed crisis-repair criterion

Comparing the results of the crisis-repair criterion with the original criteria of Stiles et al. and Strauss et al., there is a considerable difference in the number of ruptures (Table 1). We detected a total of 35 rupture-repair episodes within 9 of the shorter courses and 34 episodes within the long-term therapy when using a threshold of 1 standard deviation score (compared to 23 within 7 of the shorter courses and 27 within the long-term course using the Strauss et al. criterion). Applying the stricter value of 2 standard deviation scores, we identified 15...
episodes in 7 shorter courses and 9 in the long-term course (compared to 12 within 7 shorter courses and 4 within the long-term course using the Stiles et al. criterion).

**Classification of patterns of crisis-repair-sequences, frequencies and length**

Based on theoretical considerations, we characterized five typical crisis-repair patterns that occur in the progress of therapy. We primarily distinguished the subtypes on the basis of whether deteriorations or repairs of the therapeutic relationship comprise one session or more than one session.

We distinguished the following subtypes:

- **Pattern 1 (“jump in - jump out”, V-shape):** Deteriorations developing in leaps from one session to the next comprising one session only that are repaired in the following session.

- **Pattern 2 (“jump in - slide out”):** Deteriorations developing in leaps from one session to the next comprising one session only that are repaired gradually, meaning that the repair comprises more than one session.

- **Pattern 3 (“jump in - jump out”):** Deteriorations developing gradually over more than one session and repaired in the form of leaps from one session to the next.

- **Pattern 4 (“slide in - slide out”):** Deteriorations developing gradually over more than one session and repaired gradually, meaning that the repair comprises more than one session.

- **Pattern 5 (“complex pattern”):** Sequences in which two or more ruptures are separated by a change of direction.

The classified subtypes of crisis-repair episodes are depicted schematically in Figure 1.

*Figure 1. Classification of theoretical patterns of crisis-repair sequences. Schematic picture of the five patterns and the frequencies of their occurrence. Pattern 1 (“jump in - jump out”, V-form): Deteriorations developing in leaps from one session to the next, comprising one session only and being repaired in the following session. Pattern 2 (“jump in - slide out”): Deteriorations developing in leaps from one session to the next, comprising one session only and being repaired gradually in the following session.*
session only. Gradual repair, comprising more than one session (at least two). Pattern 3 ("slide in - jump out"): Deteriorations developing gradually over more than one session, being repaired in the following session. Pattern 4 ("slide in - slide out"): Deteriorations developing gradually over more than one session, being repaired gradually (repair comprises more than one session). Pattern 5 ("complex pattern"): Sequences in which two or more ruptures are separated by a change of direction.

We assigned each episode calculated with the crisis-repair criterion to one of the theoretical types and ascertained the frequency distribution of the types. Table 4 and Figure 1 contain the frequency of occurrence of the subtypes of crisis-repair episodes separately identified for the shorter courses, the long-term course of therapy and the complete sample.

In those cases in which deterioration occurred over more than one session ("slide in", patterns 3 and 4) we further specified whether the deterioration arose from gradual downward trends only ("slide in without a jump", patterns 3a and 4a) or in combination with one or more sudden sharp declines ("slide in including one or more jumps", patterns 3b and 4b). We did not subdivide cases with a repair extending over more than one sessions (patterns 2 and 4) as these cases occurred very rarely.

The most frequent pattern was the simple V-shape ("jump in - jump out", pattern 1, 28 ruptures, 40.6%). The second most common pattern was a decline over more than one session including a jump with a sudden repair ("slide in including one or more jumps - jump out", pattern 3b, 22 cases, 31.9%). These two patterns together account for nearly three quarters of the crisis-repair episodes in our sample. Patterns with a sliding repair occurred significantly less frequently ("slide out", patterns 2 and 4, 8 cases, 11.6%). The complex pattern was even rarer (3 cases, 4.3%).

Sudden declines and increases in the value of the therapeutic relationship variables (leaps) are a central element of therapeutic progress. 57 of a total of 66 declines (86.4%)
included at least one jump. Three quarters of the declines extending over more than one session included at least one jump.

Our criterion enables us to measure the beginning and the length of downward and upward trends exactly. This makes it possible to determine how many sessions the crises and repairs lasted. Within the complete sample, 30 deteriorations of the therapeutic relationship comprised 1 session only (45.4%), 25 cases (37.9%) extended over 2 sessions, 9 declines extended over 3 sessions, 1 extended over 4, and 1 extended over 6 sessions. The repair segments are shorter altogether: Within the complete sample, 58 repairs comprised 1 session only (87.9 %), 6 extended over 2 sessions (9.1 %), 1 extended over 3 sessions, and 1 extended over 4 sessions. (Complex patterns were not counted for these calculations.)

The results may briefly be summarized as follows: Most of the deteriorations of the therapeutic relationship comprised one or two sessions. Declines extending over more than three sessions occurred infrequently. The longest downward trend comprised a period of six sessions. Almost all repairs occurred within one session. The length of the entire crisis-repair episode for the complete sample averaged 2.9 sessions.

**Table 4**

Frequency distribution of the theoretical patterns of crisis-repair-episodes (CRE)

<table>
<thead>
<tr>
<th>Patterns</th>
<th>Subsample with 29-35 therapy sessions (n = 10)</th>
<th>Single case with 200 therapy sessions (n = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total CREs</td>
<td>1a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>Total CREs</td>
<td>34.3%</td>
<td>5.7</td>
</tr>
<tr>
<td>Courses with at least one CRE</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>%</td>
<td>47.1</td>
<td>0</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>Complete sample (n = 11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total CREs</td>
<td>28</td>
<td>2</td>
</tr>
<tr>
<td>%</td>
<td>40.6</td>
<td>2.9</td>
</tr>
</tbody>
</table>

*Note.* CRE = Crisis-repair episode.

Pattern 1 = jump in - jump out; Pattern 2 = jump in - slide out; Pattern 3 = slide in - jump out; Pattern 4 = slide in - slide out; Pattern 5 = slide in without a jump (gradual downward trend only); Pattern 6 = slide in including a jump; Pattern 7 = slide in - slide out; Pattern 8 = slide in without a jump (gradual downward trend only); Pattern 9 = slide in including a jump; Complex pattern = Two or more ruptures are intermitted through a change of direction.

**Discussion**

Temporary deteriorations within the therapeutic relationship are a distinctive feature of change over the course of the processes, which was also noted in our sample. The number of identified crises varies significantly, depending on the criterion used for measurement. Stiles et al. [7] pointed out that their rupture–repair criterion was developed *ad hoc* and should be regarded as preliminary; they indicated that application to other samples remained necessary. Strauss et al. [8] remarked that their method of quantifying alliance ruptures “is most similar to that of Stiles et al. (2004)” (p. 343). However, our testing yielded differing results, depending on the criterion that was used. Thus, the comparability of the results of studies which used different criteria seems to be limited in terms of the frequency of the ruptures and the ascertained number of courses with at least one rupture. Applying our modifications, the discrepancies in the ascertained number of rupture-repair episodes became smaller. This refers to all of our proposed modifications. Therefore, we recommend using a combination of interindividual and intraindividual variability as described in the “Methods” section. In doing so, we avoid overrating minor fluctuations occurring in relatively stable courses or declines occurring in highly fluctuating courses which are intraindividually less significant. We propose to use the stricter threshold of two standard deviations for identifying ruptures to
avoid misclassifying normal fluctuations as rupture-repair episodes. The exclusion criteria should be disregarded for standardization.

One additional note about the exclusion criteria suggested by Stiles et al. and Strauss et al. should be added: Stiles et al. noted that ruptures occurring in the course of a generally deteriorating alliance can never be regarded as fully repaired and that patients showing a negative linear trend curve should be excluded from the analysis for this reason. Strauss et al. ignored all repaired ruptures in cases with an unrepaired rupture at the end of treatment and removed these courses from the analysis completely. These exclusion criteria which call for removing potentially unsuccessful courses from the analysis should be reconsidered, especially in connection with outcome studies, as this approach seems to contain a circular argument. Also, the theoretical content of the exclusion criteria can be challenged. Rupture-repair sequences are not regarded as a separate mechanism of change when using these exclusion criteria. In theory, a rupture without repair at the end of therapy might be in connection with the termination of therapy. Notwithstanding, conflicts may have been treated in the context of previous episodes.

As mentioned in the “Methods” section, the criterion definition by Strauss et al. proved to be ambiguous for certain cases. It was not clear how we should assess those cases in which a “jump in” whose value exceeded one standard deviation difference occurs twice in succession (“double jump in”). It was equally unclear at what point these cases should be classified as repaired. Do they have to be repaired twice or does the sum of the two consecutive session ruptures needs to be repaired? Or does only the second rupture session need to be taken into account? In the latter case the extent of the rupture would be ignored or the extent of the repair would be overestimated. These specifications should be clearly defined to improve the comparability of prospective studies. We propose to count two consecutive jumps as one rupture only in order not to distort the number of identified
ruptures. We also recommend adjusting the repair value to the rupture value as, in our view, the crisis within the therapeutic relationship can only be seen as overcome if the previous level is reached again.

In comparison with the criteria of Stiles et al. and Strauss et al., more ruptures were identified with the crisis-repair criterion proposed by us. Using our criterion, it becomes possible to identify crises developing in small steps over a number of sessions; in this way, we also captured gradual downward trends. Furthermore, we involved all of the modifications at once. We differentiated five crisis-repair subtypes. The most frequent pattern was the simple V-shape ("jump in - jump out", 40.6%). The second most common pattern was a decline over more than one session including a jump with a sudden repair ("slide in including one or more jumps - jump out", 31.9%). It was striking that the episodes in more than half the cases showed more complex progress than a simple high-low-high pattern. It appeared that both the crises and the repairs may extend over several sessions. The length of the crises in the complete sample averaged 1.8 and the length of the repairs 1.2 sessions. The longest downward trend comprised a period of 6 sessions. In this light, we can assume that crisis-repair episodes exist at different temporal levels in the process. Like other authors proved, ruptures can be found within a single session that are repaired before the end of the same session [4;18;19;20]. Besides this, there are ruptures that are repaired in the following session (which results in a high-low-high pattern) and there are ruptures that extend over several sessions. Stiles et al. remarked that their criteria "were crude, justified by making use of ratings that were gathered only once for each session". They further stated that "moment-by-moment ratings of the session process are potentially much more sensitive, albeit more laborious" (p. 91). Alternatively, we can assume that the length of the episodes is related to specific characteristics of the therapeutic process. For example, it can be hypothesized that there is a relation to the extent of the entanglements of the therapists with a pathological
hostile interaction pattern of the patient. Prolonged crises may potentially occur in more disturbed patients with severe interpersonal problems. Less severe interpersonal problems may lead to more subtle crises that may be resolved within the same or at least within the next session. It is also conceivable that the length of the episodes is negatively related to therapeutic qualities such as the professional experience. As a result of the present observations, it would be interesting for future analyses to examine the potential influence of the mentioned variables on the length of the crisis-repair episodes and the effect of the length of the episodes on the outcome. For this purpose it is necessary to define in which session a crisis begins and in which session it should be considered repaired. Moreover, it seems to be useful to exactly define an episode to analyze the sessions in detail in order to reveal the mechanisms behind the crisis-repairs in the course of the therapeutic relationship. To create a basis for this kind of objective, it makes sense to examine the criteria that determine what a rupture or what a rupture-repair sequence is in depth.

At least one jump was seen in 86.4% of the declines. Thus, discontinuous sudden declines and increases in the trajectories of the therapeutic relationship are a central element of therapeutic progress. However, it also must be stated that we found gradual downward trends (isolated or in combination with a jump) in more than half the cases. Strictly speaking, it must be decided beforehand whether leaps (discontinuities) or crises of varying length can be assumed to be the crucial mechanism of change. If we theoretically view the leap to be crucial, the criterion of Strauss et al. seems to be advisable. If we assume a crisis of varying length (including gradual trends) to be crucial, the criterion proposed by us seems to be advisable. If we theoretically focus on the importance of extreme lows, the criterion of Stiles et al. appears to be useful. In a narrow sense, we can not clearly define whether it is the leap only which is related to the better outcome or if a gradual slide into a crisis has an influence as well. It must be examined whether gradual trends are also found in other samples.
We recommend use of the criterion proposed by us. The advantage of using this criterion is that ruptures with different patterns and lengths are included. This means we regard all crises, both gradual and sudden deteriorations, in the experienced therapeutic relationship as important mechanisms of change without limiting our scope to single discontinuities, i.e. sudden deteriorations. This provides a wider range for different assumptions.

When looking at the repairs, we can state that the gradual trend played a less important role. We found an isolated jump in 84% of the cases which lets us conclude that a breach in the therapeutic relationship can be dissolved completely immediately after focusing in most cases.

We would like to point out that it is useful to ascertain variables continuously throughout the processes. The results may be distorted if only selected sessions are examined as jumps between the measurements can occur. A part of the ruptures found in the study of Strauss et al. may have been developed gradually over several sessions or the number of ruptures may be even higher than documented as this study was conducted with larger intervals between the measurements.

It should be mentioned that the number of identified ruptures might depend on the measuring instrument. Stiles et al. used the Agnew Relationship Measure [28], while Strauss et al. used the California Psychotherapy Alliance Scale [29]. It is possible that certain different aspects of the therapeutic relationship are detected by different instruments. It might be that the Intrex [23;24] is a particularly sensitive instrument for showing fluctuations in the therapeutic relationship as it measures patterns of the therapeutic relationship (the interaction between therapist and patient) directly and in a relatively differentiated manner.

In our present state of knowledge, many questions still remain open. Further analyses on this topic are required. Empirical studies of the therapeutic relationship and of the factors
influencing that relationship are essential to clinical practice as change always emerges from the context of the patient-therapist relationship independently of the applied therapeutic technique. The findings of the study give an insight into basic mechanisms of change within the therapeutic relationship. Explorative elaborations of patterns of change, based on the detailed examination of single cases, can contribute to the development of realistic hypotheses and create a sound basis for future research. After all, the methodological concerns we address might be adaptable to different research topic areas where the analysis of fluctuations in a variable of interest over time is relevant.

Conclusions

The number of identified rupture-repair episodes varies significantly depending on the measuring criterion that is used. Therefore, we recommend standardizing the procedure to identify such episodes occurring in the course of the therapeutic relationship.

We propose the use of the criterion developed by us. The advantage of using this criterion is that ruptures with different patterns and lengths are included. That means we regard all crises, both gradual and sudden deteriorations in the therapeutic relationship, as important mechanisms of change as opposed to focusing on single discontinuities, i.e. sudden deteriorations. This leaves more scope for assumptions. With the criterion proposed by us we were able to identify considerably more rupture episodes compared to the use of criteria described previously. When applying the original criteria by Stiles et al. [7] and Strauss et al. [8] the modifications proposed by us should be used (stricter threshold of two standard deviations for identifying ruptures, combination of interindividual and intraindividual variability, disregarding the exclusion criteria). We also recommend adjusting the repair value to the rupture value and counting two consecutive jumps as one rupture only when using the Strauss criterion.
We classified different theoretical patterns of rupture-repair sequences in therapy processes. The most frequent pattern was the simple V-shape (40.6%). The second most common pattern was a decline over more than one session including a jump with a sudden repair (31.9%). The longest downward trend comprised a period of six sessions. We conclude that crises extend over different temporal levels in the process.

When looking at the repairs, we found an isolated jump in 84% of the cases. We conclude that a breach in the therapeutic relationship can be dissolved completely immediately after focusing, in most cases.

**Competing interests**

There were no competing interests, and no personal relationships, academic competition, or intellectual commitments that might bias the work or interfere with objective judgment.

**Authors’ contribution**

AG conceived of the study, participated in its coordination and drafted the manuscript. DK participated in the design of the study and performed the statistical analysis. EB and RE contributed to the discussions about the topic. RE drafted the manuscript.

**References**


2. X: 2010 (authors).


10. Safran JD, Muran JC: **The Resolution of Ruptures in the Therapeutic Alliance.**  


Appendix

English version of the Intrex short form. Intrex scores of items 2 to 4 and 6 to 8 which formed the weighted affiliation index [23]:

"How did my therapist behave towards me in today’s session?"

transitive

2 She/ he lets me speak freely, and warmly tries to understand me even if we disagree.

3 She/ he happily, gently, very lovingly approaches me, and warmly invites me to be as close as I would like.

4 With much kindness, she/ he teaches, protects, and takes care of me.

6 She/ he puts me down, blames me, punishes me.

7 Without thought about what might happen, she/ he wildly, hatefully, destructively attacks me.

8 Without giving it a thought, she/ he uncaringly ignores, neglects, abandons me.
intransitive

2 She/ he clearly and comfortably expresses his own thoughts and feelings to me.
3 She/ he relaxes, freely plays, and enjoys being with me as often as possible.
4 She/ he learns from me, relies upon me, accepts what I offer.
6 With much sulking and fuming, she/ he scurries to do what I want.
7 With much fear and hate, she/ he tries to hide from or get away from me.
8 She/ he walls himself off from me and doesn't react much.

“How did I behave towards my therapist in today’s session?”

transitive.

2 I let her/ him speak freely, and warmly try to understand her/ him even if we disagree.
3 I happily, gently, very lovingly approach her/ him, and warmly invite her/ him to be as close as she/ he would like.
4 With much kindness, I teach, protect, and take care of her/ him.
6 I put her/ him down, blame her/ him, punish her/ him.
7 Without thought about what might happen, I wildly, hatefully, destructively attack her/ him.
8 Without giving it a thought, I uncaringly ignore, neglect, abandon her/ him.

intransitive

2 I clearly and comfortably express my own thoughts and feelings to her/ him.
3 I relax, freely play, and enjoy being with her/ him as often as possible.
4 I learn from her/ him rely upon her/ him, accept what she/ he offers.
6 With much sulking and fuming, I scurry to do what she/ he wants.
7 With much fear and hate, I try to hide from or get away from her/ him.
8 I wall myself off from her/ him and don't react much.
The two items which are directly on the affiliation axis within the circumplex model (Items 3 and 7) contribute more strongly to the weighted affiliation index. Weighted affiliation values range from -280 to 280. We did not distinguish according to focus or direction in our analysis. The value is thus based on 24 items.

\[
A = \frac{0.5 \cdot \text{item 2} + 1 \cdot \text{item 3} + 0.5 \cdot \text{item 4} - 0.5 \cdot \text{item 5} + 0.5 \cdot \text{item 6} - 1 \cdot \text{item 7} + 0.5 \cdot \text{item 8}}{\text{number of incorporated items}}
\]

We disregarded the control index since this index is, in our view, more strongly dependent on individual psychopathology and less clearly interpretable in terms of therapeutic progress. [26] have suggested that hostility can be seen as a central component of psychopathology. In almost all investigations, negative values on the affiliation axis have been found to be significantly related to symptoms and diagnoses. According to the authors, the significance of the control axis for the definition of psychopathology is, in contrast, less well-established.
Figure 1

- **Pattern 1**: Measured value = 41%
- **Pattern 2**: Measured value = 3%
- **Pattern 3**: Measured value = 44%
- **Pattern 4**: Measured value = 9%
- **Complex pattern**: Measured value = 4%