



## Negative effects during multicomponent group-based treatment: A multisite study

Martina Pourová, Tomáš Řiháček, Luboš Chvála, Zbyněk Vybíral & Jan R. Boehnke

To cite this article: Martina Pourová, Tomáš Řiháček, Luboš Chvála, Zbyněk Vybíral & Jan R. Boehnke (2023) Negative effects during multicomponent group-based treatment: A multisite study, *Psychotherapy Research*, 33:3, 282-297, DOI: [10.1080/10503307.2022.2095237](https://doi.org/10.1080/10503307.2022.2095237)

To link to this article: <https://doi.org/10.1080/10503307.2022.2095237>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 01 Jul 2022.



Submit your article to this journal [↗](#)



Article views: 680



View related articles [↗](#)



View Crossmark data [↗](#)



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

RESEARCH ARTICLE

## Negative effects during multicomponent group-based treatment: A multisite study

MARTINA POUROVÁ <sup>A</sup>, TOMÁŠ ŘIHÁČEK <sup>A</sup>, LUBOŠ CHVÁLA <sup>A</sup>,  
ZBYNĚK VYBÍRAL <sup>A</sup>, & JAN R. BOEHNKE <sup>B</sup>

<sup>a</sup>Department of Psychology, Faculty of Social Studies, Masaryk University, Brno, Czech Republic & <sup>b</sup>School of Health Sciences, University of Dundee, Dundee, Scotland

(Received 19 January 2022; revised 26 May 2022; accepted 15 June 2022)

### ABSTRACT

**Objective** Negative effects (NEs) in group treatments remain an under-researched area. This study aimed to explore the prevalence of various types of NEs in a multicomponent group-based treatment and to determine their predictors. **Method:** A total of 330 patients participating in a multicomponent group-based treatment were recruited across seven clinical sites. At the end of treatment, the Negative Effects Questionnaire (NEQ) was used to measure NEs. Item-level descriptive analysis was conducted to explore the prevalence of various types of NEs, and structural equation modeling was used to determine predictors of these NEs. **Results:** The most frequently reported type of NEs was the worsening of symptoms, and the single most frequently reported item was the resurfacing of unpleasant memories. Predictors of NEs included the overall distress level, alexithymia, attachment avoidance, low working alliance, problem activation, and worse outcomes; psychological mindedness was a protective factor. **Conclusion:** Patients who experience higher levels of distress at the beginning of treatment, who perceive the group working alliance as problematic, and who experience high in-session emotional arousal related to their problem seem to be especially prone to reporting NEs. Furthermore, the findings do not support the assumption that NEs are a prerequisite for therapeutic change.

**Trial registration:** [ISRCTN.org identifier: ISRCTN13532466](https://www.isrctn.com/ISRCTN13532466).

**Keywords:** negative effects; predictors; group therapy; multisite study; structural equation modeling

**Clinical or methodological significance of this article:** Negative effects are a highly prevalent phenomenon in multicomponent group-based treatment. Our results suggest that especially those patients who experience higher levels of distress at the beginning of treatment, who perceive the working alliance as problematic, and who experience high in-session emotional arousal related to their problem are prone to reporting NEs. Since NEs were connected with worse outcomes in our study, practitioners should pay special attention to patients who are more likely to have these experiences.

Negative effects (NEs) in psychotherapy have received increasing attention in the last decade, and various types of NEs have been recognized in the literature including deterioration, nonresponse, emergence of novel symptoms, stigmatization, and dependency on therapy (Ladwig et al., 2014; Linden, 2013; Marmarosh, 2021; Rozental et al., 2014). In our study, we focus on those negative effects that are (a) self-reported by patients, (b)

experienced at some point during treatment but not necessarily persisting after the end of treatment, and (c) attributed by the patient, at least in part, to the treatment. This definition corresponds to the concepts of adverse treatment reactions (Schermuly-Haupt et al., 2018) and treatment-emergent reactions (Linden, 2013). Both Schermuly-Haupt et al. (2018) and Linden (2013) further distinguished between NEs produced by a correctly conducted

\*Correspondence concerning this article should be addressed to Tomáš Řiháček, Department of Psychology, Faculty of Social Studies, Masaryk University, Joštova 10, 602 00 Brno, Czech Republic. Email: [rihacek@fss.muni.cz](mailto:rihacek@fss.muni.cz)

treatment and those caused by malpractice. Although the latter differentiation is conceptually important, it is not always possible – from a patient’s perspective – to assess if the treatment was conducted properly. Therefore, in our study, we did not differentiate between these two subcategories.

The nature of group-based treatments differs in some respects from treatments conducted in the individual setting (Holmes & Kivlighan Jr, 2000): On the one hand, group therapy offers less opportunities for problem clarification, problem solution, gaining insight, and developing emotional awareness, compared to individual therapy; and on the other hand, it offers more opportunities for the development of new relationships and sources of social support, possibilities of identification with other members of the group, receiving feedback and validation from others, and experiencing altruism during the treatment process. Given these specific factors and a higher complexity of social interactions in group therapy (Burlingame et al., 2018; Johansson & Werbart, 2009; Strauss, 2021) NEs in group-based settings deserve special attention.

Empirical literature on negative effects in group therapy is scarce. Schneibel et al. (2017) have found that 60-65% of all patients experienced deterioration of their mood state or unwanted treatment reactions (Schneibel et al., 2017). In Linden et al.’s (2020) naturalistic trial study, almost everyone experienced burdens or side effects in group CBT therapy. Two qualitative studies revealed that NEs commonly encountered by patients in group therapy include difficulties in verbalizing emotions, being exposed to other members’ past experiences, managing own overwhelming feelings, feeling exposed, and inter-subjective misinterpretation (Akerman & Geraghty, 2016; Rankanen, 2014). Although these findings demonstrate the extent of the problem, they did not provide deeper understanding of the nature of these NEs.

Several measurement tools have been developed in the last decade to measure NEs. Apart from the Unwanted to Adverse Treatment Reaction (UE-ATR; Linden, 2013) checklist completed by the therapist, there are several patient-rated measures, including the Experiences of Therapy Questionnaire (ETQ; Parker et al., 2013); the Assessment of Negative Effects of Psychotherapy (INEP; Ladwig et al., 2014); and the Negative Effect Questionnaire (NEQ; Rozental et al., 2016). We used the NEQ in our study since it was the only one of the measures available in Czech (Chvála et al., 2020). The NEQ contains six domains of NEs, including worsening of symptoms, perceived quality of psychotherapy, dependency on therapy/therapist, stigma, hopelessness, and a sense of failure.

In studies that used the NEQ, the most frequently experienced NEQ factor was the worsening of symptoms factor (Hoffmann et al., 2021; Maroti et al., 2021; Rozental et al., 2016; Schaeuffele et al., 2020; van de Leur et al., 2020), which includes, for instance, experiencing increased stress, worries, anxiety, or dejection. Within this factor, the most frequently reported items included the resurfacing of unpleasant memories, with a prevalence of 28.6% to 38.46% (Maroti et al., 2021; Rozental et al., 2016; Schaeuffele et al., 2020); experiencing more stress, reported by 26.0% to 37.7% (Rozental et al., 2016; Schaeuffele et al., 2020); and experiencing more anxiety, reported by 36.3% to 37.2% (Hoffmann et al., 2021; Rozental et al., 2016). However, only Schaeuffele et al.’s (2020) study was conducted in the context of group-based multimodal treatment; the remaining studies were conducted in the context of internet-based individual therapy. Overall, Rozental et al. (2019) found that 50% of patients reported some degree of adverse experience during internet-based therapy on one or more dimensions measured by the NEQ.

### Predictors of Negative Effects

To better understand the emergence of negative treatment reactions, we need to examine their predictors. To our knowledge, no empirical study has investigated the predictors of negative effects thus far; therefore, we adopted an exploratory approach. Our study was based on data from an uncontrolled pre–post study on the effectiveness of psychotherapy and mechanisms of change (Pourová et al., 2022; Řiháček et al., 2022). For this study on NEs, we selected potential predictors from variables that were available in that dataset and were expected to predict negative effects based on existing evidence. The first group of predictors included patient *demographic variables*, i.e., age, gender, and education. Patients’ demographic variables have previously been investigated, with qualitative research suggesting that not paying attention to demographic identity issues, related to faith, gender, or race, together with a lack of understanding from the therapist can lead to NEs (Curran et al., 2019).

Patients’ *pretreatment characteristics* comprise the second group of predictors, including diagnosis, initial symptom severity, and several personal characteristics. The severity of patients’ condition, in general, and personality disorders, specifically, are believed to contribute to increased reporting of NEs in psychotherapy (Ladwig et al., 2014). Furthermore, NEs can be exacerbated by patients’ interpersonal problems (Grawe, 2004; Roback, 2000) that hinder

the development of the working alliance. These problems can be related to attachment issues and adverse childhood experiences (Lorenc et al., 2020; Macdonald et al., 2016). NEs can also be related to conditions that hinder the processing of negative emotions, such as alexithymia (Nunes da Silva et al., 2018; Ogrodniczuk et al., 2011) and dissociation (Spitzer et al., 2007; Zoet et al., 2018). In contrast, psychological mindedness may function as a protective factor that allows patients to explore their experience with insight and curiosity (Nyklíček & Denollet, 2009) and thus potentially turn NEs into benefits.

The third group included variables related to the *therapeutic process*. The working alliance is known to be related to group psychotherapy outcomes (Robak et al., 2013), and low alliance ratings and alliance ruptures (Allredge et al., 2021) are thus potential predictors of NEs. Furthermore, group cohesion is an important indicator of the overall group climate (Ogrodniczuk et al., 2006; Wongpakaran et al., 2013), and low group cohesion may jeopardize therapeutic outcomes (Burlingame et al., 2018; Joyce et al., 2007). Finally, problem actuation, a variable that refers to the patient's actual emotional experience of their problem in a therapy session, is hypothesized to be positively related to change (Mander et al., 2013). However, it did not predict outcomes in Mander et al.'s study, and the experience of the problem may, in fact, be perceived as a NE by patients.

The fourth group of variables were outcome variables (i.e., depression, anxiety, and well-being). In Schaeuffele et al.'s (2020) study, participants reported that negative effects affected their well-being "slightly to moderately". Therefore, it is reasonable to expect that negative treatment experiences reactions would be connected with worse outcomes or deterioration. On the other hand, there is an assumption that patients need to come into direct contact with painful feelings and thoughts to overcome their problems (e.g., Gassmann & Grawe, 2006), which is shared across many therapeutic schools and has been recognized as one of the common factors of therapeutic change (e.g., Lampropoulos, 2001). Some authors also emphasized that patients need to endure treatment despite disliking it to be able to benefit from it (Barnes et al., 2013). This leaves the question of the relationship between NEs and outcomes open.

### Aim of the Study

Empirical literature on NEs during psychotherapy is still scarce and no study has focused on NEs in the context of multicomponent group-based treatment. Furthermore, although some studies investigated

predictors of drop-out, deterioration, and non-improvement, no study has empirically examined predictors of NEs as defined in our study. Therefore, the aim of this study was (a) to explore the prevalence of various types of NEs in multicomponent group-based treatment and (b) to determine which pretreatment, process, and outcome variables were related to these experiences. The study was based on a secondary analysis of data on a multi-site sample of patients in multicomponent group-based treatment (see Pourová et al., 2022; Řiháček et al., 2022, for reports on the primary analyses). Although we collected evidence supporting the inclusion of the abovementioned predictors in our analysis, we did not find enough evidence to hypothesize which variables would predict which kind of NEs specifically. Therefore, we proceeded in an exploratory manner.

## Method

### Patients

This study included 330 patients which was 74% of those who agreed to participate in the study and 45% of the total number of patients accepted for treatment (see Supplement 1 for a flowchart of patient enrollment). Patients were recruited at seven clinical sites; 73.6% were women, and their ages ranged from 18 to 74 years old ( $M = 40.3$  years old,  $SD = 10.8$  years). Most patients were classified under an IDC-10 F4x diagnosis (69.4%). Six percent had multiple diagnoses, most often a combination of a personality disorder and an F3x or F4x diagnosis. See Table 1 for the sample description. Most of these patients (76.7%) participated in outpatient programs, while the remaining patients (23.3%) participated in inpatient programs.

### Treatment and Therapists

The treatments were held in seven clinical centers in the Czech Republic. It was a multicomponent treatment based on face-to-face group psychotherapy with one or two therapists. The group therapy was supplemented by therapy community meetings, relaxation training, thematic group education, ergotherapy, expressive therapy (such as art therapy, music therapy, bibliotherapy, drama therapy), physiotherapy, cognitive training, and consultations with social workers. Typically, patients received five sessions of this treatment per week, with each session lasting 90 min. Treatment length varied between four and 12 weeks across sites, and the mean individual treatment dose (i.e., days of treatment) was  $M = 32.81$  days ( $SD = 11.7$ ).

Table 1. Demographic characteristics of patients (N = 330).

Gender	n (%)	Nationality	n (%)
Women	243 (73.6%)	Czech	317 (96.1%)
Men	86 (26.1%)	Slovak	6 (1.8%)
Missing	1 (0.3%)	Other	5 (1.5%)
Age		Missing	2 (0.6%)
Mean (SD)	40.3 (10.6)	Education	
Missing	2 (0.6%)	Primary	11 (3.3%)
Household		Secondary	169 (51.2%)
In	177 (53.6%)	Tertiary	148 (44.8%)
partnership		Missing	2 (0.6%)
Alone	64 (19.4%)	Psychiatric	
With parents	32 (9.7%)	diagnosis	
Other	56 (17.0%)	F0x	6 (1.8%)
Missing	1 (0.3%)	F1x	5 (1.5%)
Marital status		F2x	3 (0.9%)
Single	155 (47.0%)	F3x	59 (17.9%)
Married	113 (34.2%)	F4x	229 (69.4%)
Divorced	60 (18.2%)	F5x	6 (1.8%)
Widowed	1 (0.3%)	F6x	48 (14.5%)
Missing	1 (0.3%)	F7x	0 (0.0%)
Occupation		F8x	0 (0.0%)
Employed	145 (43.9%)	F9x	0 (0.0%)
Entrepreneur	26 (7.9%)		
Unemployed	48 (14.5%)		
Maternity	7 (2.1%)		
leave			
Student	19 (5.8%)		
Retired	5 (1.5%)		
Disability	32 (9.7%)		
pension			
Other	14 (4.2%)		
Missing	34 (10.3%)		

The group therapy was led by 25 therapists (16 women). Their age ranged from 25 to 59 years ( $M = 44.13$ ,  $SD = 10.29$ ), and their length of practice varied between 1 and 25 years ( $M = 12.21$ ,  $SD = 7.30$ ). Psychotherapists' self-classified theoretical orientations included psychoanalysis and psychoanalytic psychotherapy ( $n = 9$ ), psychodynamic psychotherapy ( $n = 6$ ), gestalt therapy ( $n = 4$ ), person-centered approach ( $n = 3$ ), Daseinsanalysis ( $n = 1$ ), and integrative psychotherapy ( $n = 2$ ).

**Measures**

**Negative effects questionnaire (NEQ).** We used the Czech version of the NEQ (Chvála et al., 2020). The NEQ is a self-report questionnaire that consists of 32 items, each describing one type of NE or effect of psychological treatment. Patients scored each item on three scales. First, they indicated whether they experienced the effect (“yes” or “no”). Second, if their answer was “yes”, they scored the severity of the effect (i.e., “how negatively it affected me”) on a 5-point Likert-type scale ranging from

“not at all” to “extremely”. Third, they were asked to attribute the effect to “the treatment I receive” and/or “other circumstances”. Rozentel et al. (2016) reported a six-factor structure of the NEQ, including worsening of symptoms, quality of the psychotherapy and the therapeutic relationship, patients' dependency on psychotherapy, stigma, hopelessness, and a sense of failure. We tested the model using confirmatory factor analysis and found an acceptable fit,  $\chi^2(451) = 793.851$ ,  $p < 0.001$ , SRMR = 0.079, RMSEA (robust) = 0.062 [0.055; 0.069], TLI (robust) = 0.821. The TLI was suboptimal, but since the RMSEA of the null model used in its calculation was 0.153, this could not be interpreted as a lack of fit (Kenny, 2020). Therefore, we proceeded with this model in our analysis. In terms of reliability, Cronbach's  $\alpha$  was .85 for symptoms ( $k = 10$  items), .84 for quality ( $k = 11$ ), .50 for dependency ( $k = 2$ ), .69 for stigma ( $k = 2$ ), .73 for hopelessness ( $k = 4$ ), and .75 for failure ( $k = 3$ ).

**Pretreatment measures**

**Brief dissociative experience scale (DES-B).**

The DES-B (Dalenberg & Carlson, 2010) is a self-report questionnaire measuring dissociation. It consists of eight items, each rated on a five-point Likert scale from “not at all” to “more than once a day”. The baseline Cronbach's  $\alpha$  was .72 ( $n = 325$ ).

**Psychological treatment inventory-alexithymia scale (PTI-AS).**

The PTI-AS (Gori, Giannini, Palmieri, Salvini, & Schuldberg, 2012) is a self-report questionnaire measuring alexithymia. It consists of five items, each rated on a five-point Likert scale from “not at all” to “a great deal”. The baseline Cronbach's  $\alpha$  was .85 ( $n = 328$ ).

**Experiences in close relationship-relationship structure (ECR-RS).**

The ECR-RS (Fraley et al., 2011) is a self-report questionnaire measuring global attachment style. It consists of nine items, each rated on a seven-point Likert scale from “strongly disagree” to “strongly agree”. The scale has two factors: global avoidance and global anxiety. The baseline Cronbach's  $\alpha$  for global avoidance was  $\alpha = .35$  ( $n = 326$ ) and for global anxiety was  $\alpha = .87$  ( $n = 326$ ).

**Adverse childhood experiences scale (ACEs).**

The ACEs (Felitti et al., 1998) is a self-report questionnaire that measures various types of adverse childhood experiences in a patient's early life history. It consists of 14 items, each rated “yes” or “no”, and the total score is computed as a sum of



the “yes” scores. The baseline Cronbach’s  $\alpha$  was = .80 ( $n = 326$ ).

**Balanced index of psychological mindedness (BIPM).** The BIPM (Nykliček & Denollet, 2009) is a 14-item self-report measure of psychological mindedness. Each item is rated on a five-point Likert-type scale from “not true” to “very much true”. The scale consists of two subscales, namely, insight (i.e., a lack of insight into the internal phenomena) and interest (i.e., interest in attending to these phenomena, typically feelings). Their Cronbach’s  $\alpha$  coefficients at baseline were  $\alpha = .73$  ( $n = 324$ ) and  $\alpha = .81$  ( $n = 323$ ), respectively.

**Demographic questionnaire.** The demographic questionnaire contained questions about patients’ age, gender, education, household, nationality, occupation, and marital status.

#### Process measures

**Group cohesiveness scale (GCS).** The GCS (Klocek et al., 2020; Wongpakaran et al., 2013) is a self-report questionnaire for measuring group cohesion. It consists of seven items, each rated from “strongly disagree” to “strongly agree”. The Cronbach’s  $\alpha$  was .87 ( $n = 306$ ) at the first measurement.

**Group session rating scale (GSRS).** The GSRS (Quirk et al., 2013) is a self-report questionnaire for measuring the working alliance in group psychotherapy. It consists of four visual analog scales, each rated from 0 to 100. The total score is computed as a sum of all items. The Cronbach’s  $\alpha$  was .83 ( $n = 308$ ) at the first measurement.

**Scale for the multiperspective assessment of general change mechanisms in psychotherapy (SACiP).** The SACiP (Mander et al., 2013) is a self-report measure developed to assess general psychotherapeutic mechanisms, including problem actuation, resource activation, clarification of meaning, mastery, emotional bond, and agreement on collaboration. In this study, only the problem actuation subscale was used. It consists of three items, each rated on a five-point Likert-type scale ranging from “doesn’t fit at all” to “fits exactly”. Two modifications of the scale were made. First, patients were asked to rate problem actuation with regard to the last week instead of the last session only. Second, the items were reworded to refer to group psychotherapy sessions. The Cronbach’s  $\alpha$  for the problem actuation subscale was .66 ( $n = 308$ ) at the first measurement.

#### Outcomes measures

**Patient health questionnaire-9 (PHQ-9).** The PHQ-9 (Dášová et al., 2016; Kroenke et al., 2001) is a self-report questionnaire for screening the severity of depressive symptoms over the last two weeks. It consists of nine items, each rated on a four-point Likert scale from “not at all” to “nearly every day”. The baseline Cronbach’s  $\alpha$  was .81 ( $n = 328$ ).

**Generalized anxiety disorder screener (GAD-7).** The GAD-7 (Löwe et al., 2008) is a self-report questionnaire for screening anxiety symptoms over the last two weeks. It consists of seven items, each rated from “not at all” to “nearly every day”. The baseline Cronbach’s  $\alpha$  was .85 ( $n = 325$ ).

**Well-being index (WHO-5).** The WHO-5 (Bech et al., 2003) is a self-report questionnaire for assessing psychological well-being in terms of hedonia. It consists of five items, each rated from “all the time” to “at no time”. The baseline Cronbach’s  $\alpha$  was .86 ( $n = 327$ ).

#### Procedure

The study was approved by the Research Ethics Committee of Masaryk University (ref. no. EKV-2017-029-R1). All measures for which a Czech version was not available (i.e., DES-B, PTI-AS, ACEs, BIPM, GCS, and SACiP) were translated into Czech from the English version. Five native Czech speakers (a psychology student, two psychologists, and two laypeople) created five independent Czech translations. A group of three people (the two psychologists and the psychology student) then discussed all the translations and consolidated them into a single version. This version was then back-translated into English by a bilingual, native English speaker and compared to the original English version. Finally, the Czech version was field-tested with five respondents to check the comprehensibility of the items.

The data were collected between 2018 and 2019. All patients who participated in the treatment were invited to participate in the research. All patients who agreed and provided informed consent were administered (1) a pretreatment battery that included the demographic questionnaire, the pretreatment measures (i.e., DES-B, PTI-AS, ECRS, ACEs, and BIPM), and the outcome measures (i.e., PHQ-9, GAD-7, WHO-5), (2) a weekly assessment battery that included the process measures (i.e., GCS, GSRS, and SACiP), and (3) a posttreatment battery that included the outcome measures

(i.e., PHQ-9, GAD-7, WHO-5) and the NEQ. The batteries also included other measures not analyzed in this study (Řiháček, 2018). The data were collected in paper-and-pencil form.

### Statistical Analysis

Data analysis was conducted using R software version 4.0.5 (R Core Team, 2021). Before the analysis, we preprocessed the NEQ variables. Since we were interested only in those NEs that were – at least in part – attributed to therapy, we recoded the NEQ items accordingly. Specifically, a severity score of zero was assigned to items where respondents (1) did not rate the effect severity and indicated that they did not experience the given effect at all or (2) experienced the given effect but attributed it fully to extratherapeutic factors. A similar procedure was used in the original study (Rozenal et al., 2016).

The first aim of the study was to explore the prevalence of various types of NEs. We conducted an item-level descriptive analysis, treating the severity scores as ordinal variables. The results are reported in a graph that shows the percentage of patients who experienced the effect at each level of the scale, as well as the overall proportion of patients who experienced the effect (i.e., they scored the effect as “slightly” or higher). We also computed 95% confidence intervals for the proportions, assuming underlying normal distributions. Furthermore, we computed the mean severity rating for each item for those observations in which the effect was rated as “present”. The severity ratings thus represent the adversity of the effect in case it was experienced, without taking the frequency of its occurrence into account.

The second aim was to identify the predictors of NEs. This analysis was conducted within the structural equation modeling framework using the “lavaan” package in R (Rosseel, 2012). First, we tested the six-factor measurement model for the NEQ using the robust maximum likelihood (MLR) estimator and the full information maximum likelihood (FIML) method to account for missing data. In factors comprised of only two items, the factor loadings were constrained to the same value to ensure factor identification. Second, the NEQ factors were included as dependent variables, with the other variables included as predictors, including the set of demographic variables (age, gender, and education), the patient’s baseline psychopathology (diagnosis, baseline outcome level), other patients’ baseline characteristics (DES-B, PTI-AS, ECR-RS, ACEs, and BIPM), process variables (GSRS, GCS, and problem actuation), therapeutic change, and

the length of treatment (i.e., days of treatment attended by a patient). Education was dichotomized (tertiary vs. lower), and diagnosis was converted into a set of dummy variables (only F3x, F4x, and F6x were included since other diagnoses were represented only marginally). Scores for the continuous predictors were computed as averages of the scales’ items. If a patient answered less than 80% of the scale’s items, their response was considered missing data. Since the three outcome measures (i.e., PHQ-9, GAD-7, and WHO-5) were highly correlated both in terms of baseline values (absolute values between  $r = .59$  and  $.72$ ) and pre–post change scores (absolute values between  $r = .57$  and  $.68$ ), we created one composite outcome variable as a sum of the standardized scores of these three outcome variables. The composite outcome variable thus represented an overall distress level (WHO-5 scores were reversed for that purpose). The outcome change score was computed by subtracting the post-treatment score from the pretreatment score (positive values thus represented improvement). Furthermore, since the process variables were measured weekly, we included patients’ individual means, as well as their standard deviations, as predictors. This was motivated by our expectation that both the overall level of a process variable (represented by the mean) and its instability (represented by the standard deviation) may be related to NEs. The NEQ item loadings were fixed to the values obtained from the measurement model alone. This was done to ensure that the measurement model was not altered by the predictors. To evaluate the model’s overall predictive power, we compared the full structural model to a model in which all regression paths were fixed to zero (subsequently referred to as the baseline model).

Finally, we conducted a bias analysis to explore how patients who answered the NEQ differed from those who did not. We compared these two groups in terms of their baseline characteristics, as well as the therapeutic process. For continuous variables, we reported the standardized mean difference (Cohen’s  $d$ ) and tested the statistical significance using a  $t$  test. For dichotomous variables, we reported  $\phi$  and used the  $\chi^2$  test to test the difference in proportions.

## Results

### Sample and Missing Data

A total of  $N = 330$  patients completed the treatment and the posttreatment assessment. Of these patients, 317 answered all NEQ items, 10 patients missed one item, two missed two items and one missed eight

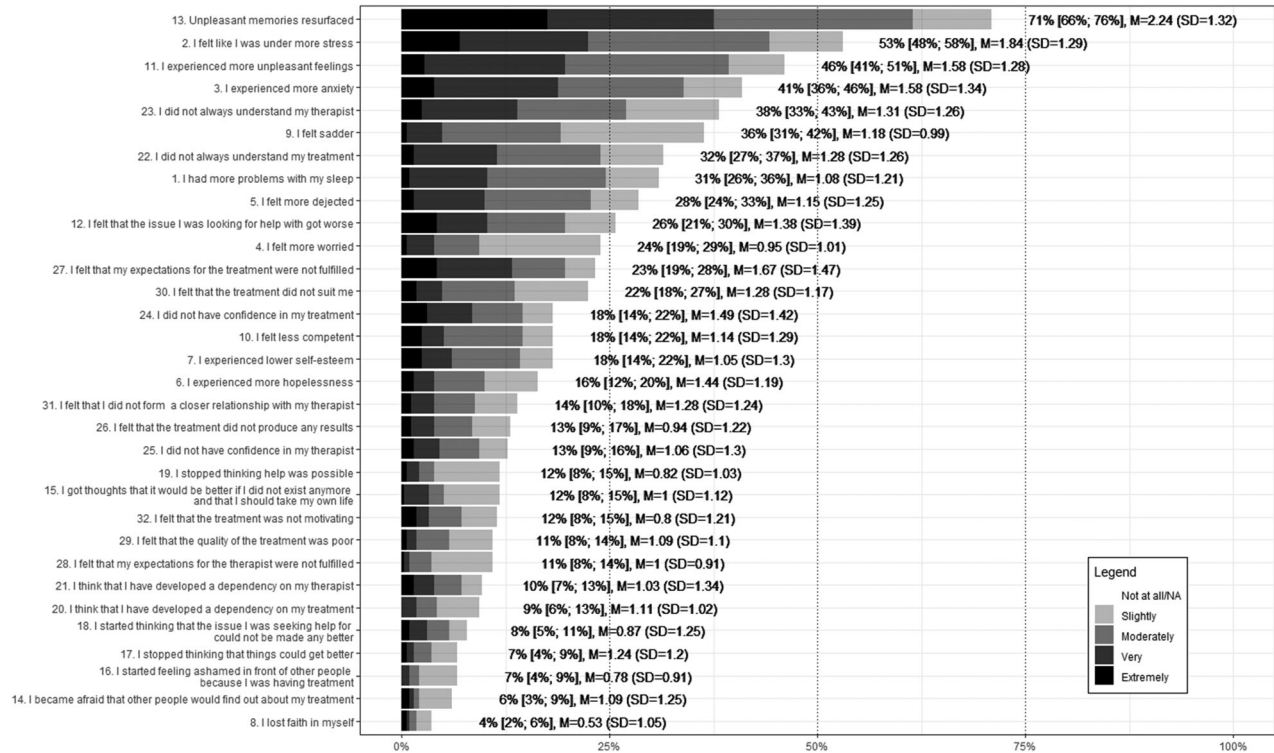


Figure 1. Prevalence and severity of negative effects in group treatment (items drawn from the NEQ). Note: For each item, the overall percentage of patients who reported the effect is presented, as well as the 95% confidence interval (in square brackets), the mean intensity and standard deviation.

items. All missingness patterns were unique, suggesting that the missingness was not item-specific. The received dose of psychotherapy ranged from 2 to 12 weeks ( $Mdn = 7$ ).

### Prevalence of Negative Effects

Figure 1 shows the proportion of patients who reported various types of NEs irrespective of the severity of the experience (see Supplement 4 for proportions per site). The most prevalent type of NE was the resurfacing of unpleasant memories (Item 13), which was reported by 71% of the sample. Two other types of NEs were reported by approximately half of the sample, namely, experiencing more unpleasant feelings (Item 11, 54%) and experiencing more stress (Item 2, 46%). Several other types of NEs were reported by more than a quarter of the sample: experiencing more anxiety (Item 3, 41%), feeling sadder (Item 9, 38%), not always understanding the treatment (Item 22, 36%), feeling more dejection (Item 5, 32%), having more problems sleeping (Item 1, 31%), having more worries (Item 4, 28%), and experiencing more hopelessness (Item 6, 26%). The remaining 22 items were reported by less than 25% of the sample.

Overall, the most prevalent types of NEs belonged to the Symptoms subscale, except for suicidal ideation (Item 15), which was reported by only 12% of the sample. On the other hand, Items 14 and 16, which belonged to the Stigma subscale, were reported by very few patients (8% and 4%, respectively). Somewhat higher prevalence was reported for Items 10, 7, and 8 belonging to the Failure subscale (18%, 13%, and 10%); Items 19, 18, and 17 belonging to the Hopelessness subscale (18%, 18%, and 13%); and Items 20 and 21 belonging to the Dependency subscale (16% and 9%, respectively). Items belonging to the Quality of Psychotherapy subscale received more variable ratings; the most prevalent items included not understanding one's treatment (Item 22, 36%), not understanding one's therapist (Item 23, 24%), and feeling that one's expectations for the treatment were not fulfilled (Item 27, 22%).

### Severity of Negative Effects

The severity of NEs was calculated only for patients who reported the given type of effect and, therefore, did not take the prevalence into account (see Figure 1; Supplement 4 for breakdown per site). Interestingly, the severity ratings resembled the prevalence



to a large degree, with the most prevalent NE (the resurfacing of unpleasant memories, Item 13) also being perceived as the most severe ( $M = 2.24$ ) by the patients. Other items rated high in severity included experiencing more unpleasant feelings (Item 11,  $M = 1.84$ ), experiencing more stress (Item 2,  $M = 1.58$ ), and experiencing more anxiety (Item 3,  $M = 1.58$ , all belonging to the Symptoms subscale).

**Prediction of Negative Effects**

The descriptive statistics of the predictors are reported in Table 2 (see Supplement 5 for unstandardized coefficients), and the results of the prediction analysis are presented in Table 3. Twenty-seven cases were removed due to missing values of predictors. Overall, the fit of the structural model was acceptable,  $\chi^2(1209) = 1824.495$ ,  $p < 0.001$ ,  $BIC = 21945.241$ ,  $SRMR = 0.079$ ,  $RMSEA(robust) = 0.045 [0.041; 0.050]$ ,  $TLI(robust) = 0.769$ . The TLI was suboptimal, but since the RMSEA of the null model was 0.097, this could not be interpreted as a lack of fit (Kenny, 2020). Factor loadings and correlations of the NEQ measurement model are presented in Supplement 2. The fit of the full structural model

outperformed the fit of the baseline model (see Supplement 3).

**Demographic variables.** The only demographic variable that predicted NEs was age: younger patients more often reported symptom deterioration ( $\beta = -.16$ ) and stigma ( $\beta = -.13$ ) during treatment.

**Diagnosis and baseline severity.** Taking the diagnosis and baseline severity of patients into account, we found that a diagnosis of affective disorder predicted perceived low quality of therapy ( $\beta = .15$ ) and worsening of symptoms ( $\beta = .14$ ). A diagnosis of a neurotic, stress-related, or somatoform disorder also predicted perceived low quality of therapy ( $\beta = .16$ ) but was also related to lower hopelessness ( $\beta = -.15$ ). A diagnosis of personality disorder was not a significant predictor of any type of NEs. Furthermore, the higher the baseline distress level was, the lower the perceived quality of therapy ( $\beta = .22$ ) and the higher hopelessness ( $\beta = .19$ ).

**Other patients' baseline characteristics.** Only a few effects were found for other baseline characteristics. The interest dimension of psychological mindedness was related to lower scores for Stigma ( $\beta = -.13$ ), Failure ( $\beta = -.12$ ), and Symptoms ( $\beta$

Table 2. Descriptive statistics and bias analysis.

	NEQ completers ( $N = 330$ )	NEQ noncompleters ( $N = 114$ )	Difference <sup>a</sup>
Age	$M = 40.25 (SD = 10.75)$	$M = 35.99 (11.93)$	$d = 0.39^{***}$
Gender	74%	78%	$\phi = 0.03$
Education: tertiary	45%	31%	$\phi = 0.12^*$
F3x	18%	22%	$\phi = 0.04$
F4x	69%	73%	$\phi = 0.03$
F6x	15%	18%	$\phi = 0.03$
Depression (baseline)	$M = 1.58 (SD = 0.62)$	$M = 1.70 (SD = 0.73)$	$d = -0.18$
Anxiety (baseline)	$M = 1.66 (SD = 0.68)$	$M = 1.76 (SD = 0.74)$	$d = -0.14$
Well-being (baseline)	$M = 1.40 (SD = 0.86)$	$M = 1.42 (SD = 1.00)$	$d = -0.02$
Dissociation	$M = 1.21 (SD = 0.74)$	$M = 1.32 (SD = 0.75)$	$d = -0.15$
Alexithymia	$M = 3.29 (SD = 0.97)$	$M = 3.34 (SD = 1.12)$	$d = -0.05$
Attachment avoidance	$M = 4.09 (SD = 1.19)$	$M = 4.05 (SD = 1.41)$	$d = 0.03$
Attachment anxiety	$M = 4.18 (SD = 1.85)$	$M = 4.33 (SD = 1.84)$	$d = -0.08$
Adverse childhood experiences	$M = 4.22 (SD = 3.14)$	$M = 4.45 (SD = 3.00)$	$d = -0.07$
Psychological mindedness (insight)	$M = 2.21 (SD = 0.77)$	$M = 2.25 (SD = 0.78)$	$d = -0.05$
Psychological mindedness (interest)	$M = 2.31 (SD = 0.79)$	$M = 2.28 (SD = 0.83)$	$d = 0.04$
Working alliance – mean	$M = 293.21 (SD = 67.82)$	$M = 278.11 (SD = 73.63)$	$d = 0.22^*$
Working alliance – SD	$M = 43.72 (SD = 27.16)$	$M = 51.14 (SD = 30.41)$	$d = -0.26^*$
Group cohesion – mean	$M = 3.86 (SD = 0.60)$	$M = 3.66 (SD = 0.68)$	$d = 0.32^{**}$
Group cohesion – SD	$M = 0.42 (SD = 0.23)$	$M = 0.45 (SD = 0.28)$	$d = -0.12$
Problem actuation – mean	$M = 2.84 (SD = 0.58)$	$M = 2.68 (SD = 0.76)$	$d = 0.25^*$
Problem actuation – SD	$M = 0.55 (SD = 0.27)$	$M = 0.58 (SD = 0.38)$	$d = -0.10$
Treatment dose (days)	$M = 32.81 (SD = 11.7)$	$M = 27.44 (SD = 16.5)$	$d = 0.41^{***}$

Note: NEQ = Negative Effects Questionnaire; F3x, F4x, and F6x = diagnoses according to the ICD-10; M = mean; SD = standard deviation. \*  $p < .05$ , \*\*  $p < .01$ , and \*\*\*  $p < .001$ .

<sup>a</sup>For continuous variables, Cohen's  $d$  and the significance of a t test are reported. For dichotomous variables, phi and the significance of a chi-square test are reported.

Table 3. The structural equation model simultaneously predicting negative effects (latent factors according to NEQ subscales, all coefficients standardized, N = 303).

Predictors	NEQ factors					
	Symptoms	Quality	Dependency	Stigma	Hopelessness	Failure
<i>Demographic variables</i>						
Age	-.16 [-.25; -.06]**	-.08 [-.18; .02]	-.15 [-.3; .00]	-.13 [-.24; -.01]*	-.06 [-.17; .05]	.00 [-.13; .13]
Gender	.08 [-.01; .16]	-.02 [-.13; .08]	.06 [-.06; .18]	.08 [-.06; .22]	.08 [-.03; .19]	.07 [-.03; .18]
Education (tertiary)	-.01 [-.10; .08]	.07 [-.02; .17]	.08 [-.07; .22]	.12 [-.02; .26]	-.04 [-.15; .07]	.00 [-.11; .11]
<i>Diagnosis and baseline severity</i>						
F3x	.14 [.02; .26]*	.15 [.03; .26]*	.08 [-.07; .23]	.15 [-.01; .31]	-.04 [-.20; .11]	.10 [-.09; .30]
F4x	.07 [-.06; .20]	.16 [.01; .30]*	.15 [-.07; .36]	.07 [-.05; .19]	-.15 [-.30; .00]*	.03 [-.17; .24]
F6x	.04 [-.08; .15]	.12 [-.05; .30]	.01 [-.20; .21]	-.02 [-.15; .10]	-.07 [-.21; .07]	.04 [-.12; .21]
Overall distress (baseline) <sup>a</sup>	.12 [-.01; .25]	.22 [.08; .36]**	.11 [-.11; .34]	.08 [-.08; .24]	.19 [.05; .34]**	.13 [-.04; .30]
<i>Other patients' baseline characteristics</i>						
Dissociation	-.02 [-.11; .07]	-.02 [-.13; .09]	.00 [-.15; .15]	.06 [-.05; .18]	-.02 [-.16; .11]	.02 [-.13; .17]
Alexithymia	.08 [-.03; .18]	.04 [-.08; .16]	.12 [-.06; .30]	-.07 [-.21; .06]	.00 [-.14; .14]	.13 [.00; .25]*
Attachment avoidance	.05 [-.04; .14]	-.14 [-.26; -.03]*	.04 [-.12; .19]	-.12 [-.33; .09]	-.12 [-.26; .02]	-.05 [-.19; .09]
Attachment anxiety	-.02 [-.11; .07]	-.07 [-.21; .07]	.01 [-.19; .21]	.15 [.00; .29]	.09 [-.01; .20]	-.03 [-.14; .08]
Adverse childhood exp.	.08 [-.01; .17]	.01 [-.09; .11]	.10 [-.03; .23]	-.01 [-.14; .11]	.01 [-.10; .12]	.03 [-.09; .16]
Psych. mindedness (insight)	.01 [-.08; .10]	-.11 [-.22; .00]*	.08 [-.07; .22]	.00 [-.14; .13]	-.08 [-.21; .04]	.03 [-.09; .14]
Psych. mindedness (interest)	-.10 [-.19; -.01]*	-.03 [-.12; .06]	-.06 [-.19; .07]	-.13 [-.26; -.01]*	-.08 [-.19; .02]	-.12 [-.25; .00]*
<i>Psychotherapy process variables</i>						
Working alliance (M)	-.23 [-.36; -.10]***	-.26 [-.42; -.09]**	-.13 [-.32; .06]	-.01 [-.17; .15]	-.24 [-.43; -.04]*	-.17 [-.38; .03]
Working alliance (SD)	.02 [-.10; .15]	-.13 [-.28; .01]	.08 [-.08; .24]	-.03 [-.12; .06]	-.15 [-.30; .00]*	.06 [-.10; .22]
Group cohesion (M)	.02 [-.10; .13]	-.11 [-.25; .03]	.15 [-.02; .32]	-.05 [-.22; .11]	-.01 [-.15; .13]	.00 [-.17; .17]
Group cohesion (SD)	-.05 [-.15; .05]	-.01 [-.10; .07]	.05 [-.10; .20]	.07 [-.04; .18]	.05 [-.06; .16]	.00 [-.11; .12]
Problem actuation (M)	.30 [.21; .39]***	.05 [-.06; .17]	.11 [-.05; .27]	.00 [-.11; .11]	.15 [.01; .29]*	.20 [.08; .32]**
Problem actuation (SD)	.17 [.08; .26]***	.13 [.02; .23]*	-.05 [-.19; .08]	.09 [-.03; .22]	.00 [-.11; .12]	.10 [-.01; .21]
<i>Therapeutic change</i>						
Overall improvement <sup>a</sup>	-.30 [-.40; -.19]***	-.22 [-.33; -.11]***	-.12 [-.27; .03]	.05 [-.10; .20]	-.23 [-.36; -.10]***	-.32 [-.45; -.19]***
Length of treatment <sup>b</sup>	-.02 [-.14; .11]	.10 [-.06; .27]	.09 [-.15; .33]	-.13 [-.33; .08]	.09 [-.09; .27]	.12 [-.04; .28]
<i>Correlations among NEQ factors</i>						
Symptoms	–	.52	.19	.27	.62	.77
Quality		–	.04	.26	.67	.49
Dependency			–	.10	.08	-.00
Stigma				–	.31	.27
Hopelessness					–	.76

Note: NEQ = Negative Effects Questionnaire; F3x, F4x, and F6x = diagnoses according to the ICD-10 (the diagnoses were dummy-coded as present/absent to reflect the possibility of multiple diagnoses per patient). Values represent standardized regression coefficients and their 95% confidence intervals. Positive coefficients for gender indicate a higher prevalence in men. Dummy variables representing the site effects (treated as fixed effects) were omitted from the table; their coefficients ranged from -.20 to -.09 for symptoms, -.16 to .01 for quality, -.12 to .21 for dependency, -.11 to .00 for stigma, -.21 to -.03 for hopelessness, and -.13 to .01 for failure. \*  $p < .05$ , \*\*  $p < .01$ , and \*\*\*  $p < .001$ .

<sup>a</sup>A composite variable computed as the sum of sample-standardized PHQ-9, GAD-7, and (reversed) WHO-5 scores.

<sup>b</sup>Days of treatment attended by a patient.

= -.10), while the insight dimension of psychological mindedness was related to better perceived quality of treatment ( $\beta = -.11$ ). The avoidance dimension of attachment was also related to better perceived quality of therapy ( $\beta = -.14$ ), but the anxiety dimension did not predict any type of NE significantly. Higher scores in alexithymia were related to a higher sense of failure in therapy ( $\beta = .13$ ). None of the effects for dissociation and adverse childhood experiences reached statistical significance.

**Psychotherapy process variables.** While group cohesion was unrelated to negative effects, the mean value of the working alliance was negatively related to complaints about treatment quality ( $\beta = -.26$ ), hopelessness ( $\beta = -.24$ ), and worsening of symptoms ( $\beta = -.23$ ). Intraindividual variability in the perception of the working alliance also predicted lower hopelessness ( $\beta = -.15$ ). Moreover, problem actuation was related to several types of NEs: the mean value of problem actuation was positively related to worsening symptoms ( $\beta = .30$ ), a sense of failure in therapy

( $\beta = .20$ ), and a sense of hopelessness ( $\beta = .15$ ). Moreover, the more variable a patient's ratings of problem actuation were, the more worsening of symptoms ( $\beta = .17$ ) and the lower perceived quality of treatment ( $\beta = -.13$ ) they reported.

**Therapeutic change.** Finally, the less patients improved (of the more they deteriorated) during the treatment, the more they tended to experience a sense of failure ( $\beta = -.32$ ), worsening of symptoms ( $\beta = -.30$ ), a higher sense of hopelessness ( $\beta = -.23$ ), and lower perceived quality of treatment ( $\beta = -.22$ ).

### Bias Analysis

Patients who responded to the NEQ at posttreatment ( $N = 330$ ) differed from the rest of the intent-to-treat sample in several aspects ( $N = 114$ ; see Table 2). As expected, the responders received higher treatment doses. Furthermore, they tended to be older, slightly more educated, and slightly more depressed and anxious at baseline than nonresponders. On average, they evaluated the quality of the working alliance, group cohesion, and problem actuation as higher and more stable compared to nonresponders.

### Discussion

This study aimed to determine the prevalence of various kinds of NEs and to explore how these effects related to pretreatment, process, and outcome variables. Consistent with the literature, patients in our study most frequently reported worsening of symptoms (Hoffmann et al., 2021; Maroti et al., 2021; Rozental et al., 2016; Schaeuffele et al., 2020; van de Leur et al., 2020), including unpleasant memories and feelings, stress, anxiety, sadness, dejection, and sleep problems (see Figure 1). The prevalence of these symptomatic complaints suggests that a temporary worsening of symptoms may be considered an expected side effect of dealing with traumatic or adverse issues or perhaps even an indispensable part of therapeutic change (Barnes et al., 2013; Gassmann & Grawe, 2006). However, our data do not support this assumption since all but one NEQ subscale correlated with worse outcomes. Furthermore, the Symptom subscale yielded medium to high correlations with treatment dissatisfaction, hopelessness, and a sense of failure, all of which were also related to worse outcomes. Thus, our findings suggest that symptomatic deterioration during treatment, on average, should be considered a detrimental effect that decreases patients' likelihood of successful outcomes and may lead to demoralization (Kissane & Clark, 2002;

Shimokawa et al., 2010; Strauss et al., 2015). This is in line with Brakemeier et al.'s (2015) finding that subjective deterioration during treatment was associated with lower likelihood of remission at discharge. Nevertheless, the cross-sectional nature of our findings does not allow us to make causal claims, and a longitudinal design is necessary to test these preliminary conclusions.

From the clinical perspective, various types of NEs probably differ in their gravity. Arguably, the most prevalent item, the resurfacing of unpleasant memories, could be seen as less detrimental than, for instance, a sense of one's own failure or suicidal ideation. Many therapeutic approaches, including experiential approaches or those based on exposure, require patients to face unpleasant memories or experiences almost by definition. Therefore, we explored how the resurfacing of unpleasant memories was related to outcome at the item level. We found a weak correlation ( $r_s = .18$ ) with worse outcomes that vanished after controlling for the baseline distress level. Therefore, the resurfacing of unpleasant memories, on average, did not relate to treatment outcome. Nevertheless, in a qualitative study, Grafanaki and McLeod (2002) suggested that patients' reaction to being confronted with unpleasant memories or experiences may depend on patients' readiness to deal with them, and further qualitative work exploring the importance and impact of negative NEs from the perspective of the patients seems necessary.

Patients in our sample tended to report NEs more often than those in Strauss et al.'s (2021) study that primarily included patients in individual outpatient settings. For instance, while 71% of patients in our sample reported the resurfacing of unpleasant memories, only 40% did so in a predominantly individual psychotherapy sample (Strauss et al., 2021), and a similar pattern was observed for most NEQ items. An especially concerning finding is the prevalence of suicidal thoughts attributed to treatment (12% in our sample). Although the presence of suicidal ideation in psychotherapy samples is not surprising (e.g., 31% of patients had an unfavorable suicidal ideation trajectory in Alexopoulos et al., 2021), the fact that suicidal ideation was presumably caused or aggravated by the treatment itself calls for further attention. The prevalence of this phenomenon in our sample was ten times higher than in Strauss et al.'s (2021) sample, suggesting that group therapy may be more likely to generate negative effects compared to individual treatments. For instance, one patient's sharing of distressing life stories and troubles with the group can uncover forgotten unpleasant memories in the rest of the group (Bernard et al., 2008). However, the differences in prevalence may be

caused by other characteristics of the samples, such as the type or duration of treatment and differences in the study design. Furthermore, Supplement 4 shows that there was considerable variation in prevalence and reported severity across sites, suggesting that the occurrence of NEs (or the willingness to share them via a questionnaire) may vary based on the site setting. However, confidence intervals of these estimates were large due to small subsample sizes and, therefore, any conclusions about the variability must be made with caution.

The demographic characteristics were not related to NEs, except for a relationship between lower age and higher symptom deterioration during treatment. The diagnosis and baseline severity predicted these effects somewhat better. Unlike in Ladwig et al.'s (2014) study, personality disorders (F6x) did not significantly predict any category of NEs in our sample. Affective disorders (F3x) were related to higher worsening of symptoms, which is in line with existing findings on symptom deterioration during the treatment of depression (Brakemeier et al., 2015). Surprisingly, stress-related and somatoform disorders (F4x) seemed to protect patients from a sense of helplessness. Since this group of disorders is typically based on dysfunctional attitudes that are easier to influence in therapy, these patients may more easily regain a sense of control over their lives compared to those in other diagnostic groups (Rodziński et al., 2019). Both F3x- and F4x-type diagnoses predicted lower perceived treatment quality.

Other patients' baseline characteristics that were related to NEs included alexithymia, attachment avoidance, and psychological mindedness. Alexithymia predicted a sense of patient failure in therapy. This finding is in line with studies that described a connection between alexithymia and negative indicators of mental health (Kauhanen et al., 1996; Vanheule et al., 2007). Attachment avoidance predicted higher reported satisfaction with treatment. This is consistent with Diener and Monroe's (2011) meta-analysis on the relationship between adult attachment style and therapeutic alliance in individual psychotherapy. This may mean either that patients with higher levels of avoidance in relationships are less critical of therapists' failures or less ready to express their critique.

Psychological mindedness serves as a protective factor that can help patients cope with psychologically demanding situations (Kronström et al., 2009). The insight dimension of psychological mindedness predicted lower dissatisfaction with treatment quality in our study, which may mean that patients with higher levels of insight are better able to understand treatment procedures or tolerate mild levels of patient-treatment or patient-therapist misfit. The

interest dimension of psychological mindedness protected patients from experiences of worsening symptoms, stigma, and failure. Patients' interest in their own experience or the psychological meaning of symptoms can facilitate the therapeutic process, increase their ability to benefit from the treatment, or, again, increase their tolerance for therapy situations that do not meet their expectations (Beitel et al., 2004; Grant, 2001).

The association between the working alliance and NEs resonates with the fact that the working alliance is a robust predictor of outcomes in both individual (Flückiger et al., 2018; Horvath et al., 2011) and group (Joyce et al., 2007; Marziali et al., 1997) psychotherapy. While the working alliance predicted worsening of symptoms and the perceived quality of psychotherapy, it was unrelated to a sense of stigmatization. Although group cohesion was related to outcomes in Marziali et al.'s (1997) study, it did not predict any of the NEQ subscales in the current study.

As expected, problem actuation predicted several domains of NEs. Coupled with the finding that problem actuation is unrelated to outcome (Mander et al., 2013; Řiháček et al., 2022), this suggests that the role of problem actuation is detrimental rather than helpful. This is contrary to the theoretical assumption of Grawe's integrative model that patients need to come into direct contact with painful feelings and thoughts to overcome their problems (Gassmann & Grawe, 2006). Given the wide trans-theoretical consensus regarding problem actuation/confrontation as a psychotherapy change principle, it would be premature to reject it solely on the basis of a single study. However, a more nuanced perspective is needed to understand the role of this principle in facilitating change. For instance, it may only work when accompanied by other relational conditions, such as a warm interpersonal style of the therapist (Nissen-Lie et al., 2010), as well as patient-related conditions, such as resource activation (Gassmann & Grawe, 2006) or readiness for change and tolerance for unpleasant emotions. To understand for whom and under which conditions problem actuation/confrontation facilitates changes, we must study it from a longitudinal perspective and in interaction with other variables.

### Limitations

Our study had several limitations. First, our sample included only those patients who completed the treatment and responded to the posttreatment measures. Therefore, patients who were less motivated to complete the treatment or less committed



to participate in the research project were underrepresented. The analysis revealed that the patients in the sample tended to be older, more depressed and anxious, and generally more satisfied with the treatment compared to those who did not complete the treatment or did not answer the NEQ. These characteristics may have biased the results (Rozenal et al., 2018), although we cannot easily determine in which direction. On the one hand, the NEQ completers were, on average, older and rated the working alliance as better. Since these characteristics were connected with less NEs, this would lead to an underestimation of the prevalence. On the other hand, the NEQ completers had, on average, a higher baseline level of distress. Since this variable was related to higher occurrence of NEs, this would lead to an overestimation. Furthermore, this study was conducted in the context of daily group treatment. These patients may suffer from higher levels of distress than patients in less intensive treatments. Therefore, more research is needed to assess NEs in other populations (e.g., patients suffering from less severe problems, patients with personality disorders, or patients with psychotic disorders).

Second, high sample heterogeneity in terms of the auxiliary treatment components (e.g., art therapy), length of the treatment, therapists' theoretical orientation, and therapists' experience can be seen both as a strength and a limitation of the study. Resulting from a multi-site naturalistic study, our findings probably better reflect the reality of the everyday therapeutic practice than findings from more controlled studies. To address potential bias introduced by this heterogeneity, we (a) analyzed the prevalence and severity of NEs both for the total sample and per site (see Supplement 4) and (b) controlled for site effect in the structural equation model.

Third, the Cronbach's alpha was suboptimal for some measures (ECR-RS global avoidance and NEQ dependency). Both these measures are very brief and, therefore, lower reliabilities were expected. Nevertheless, they were lower than in other studies (see, e.g., Moreira et al., 2015, for the ECR-RS and Rozenal et al., 2016, for the NEQ). While the predictive value of questionnaire-based measures is not directly related to their reliability (Smits et al., 2018), it could have limited our ability to detect relationships for these two measures.

Fourth, the types of NEs reported by the patients were limited to those represented in the NEQ items. The NEQ does not focus on group-specific NEs, such as hassles between members, highly critical interpersonal feedback about one's personal shortcomings ("feedback overload", Roback, 2000). Furthermore, the attribution of the effect to the treatment or other circumstances was made

solely by the patients. However, their judgment may be biased by factors such as limited insight or social desirability and must be interpreted with caution. Although the patients were reassured that neither their psychotherapists nor the clinic staff would see their responses, the mere fact that the data were collected at the clinic could have influenced patients' willingness to share their reactions.

Fifth, the Symptoms subscale of the NEQ is conceptually akin to the outcome measures. Although the framing instructions differed for the NEQ (asking patients whether they had experienced a negative effect at any time during their treatment) and for the outcome measures (asking patients how they had been doing over the last couple of days), it is possible that patients' overall evaluation of the treatment experience influence both types of measures. Arguably, it may be difficult from the patients' perspective to disentangle these nuanced aspects of experience. Thus, from this point of view, it would be desirable to measure NEs prospectively. Furthermore, the NEQ was a part of the post-treatment battery of measures and patients' responses in the NEQ could thus be influenced by answering the outcome measures at the same time. Apart from serving as potentially useful real-time feedback for psychotherapists, this would allow us to better disentangle NEs that occur as part of the therapeutic process from those of the overall outcome.

Sixth, the study was based on NEs as retrospectively reported by patients. Research in other areas has shown that people can form incorrect memories about past events and as a consequence evaluate people and circumstances associated with such "false memories" for example more negatively (Muschalla & Schönborn, 2021). Our findings should thus be interpreted keeping in mind that we investigated patients' subjective and retrospective evaluations. But since unpleasant memories are a common NE (e.g., Rozenal et al., 2016), this also points to an important area for further research and study design development to investigate the nature of such memories and their development more closely in the context of negative effects.

## Conclusion

We found that the most frequently reported negative effect in multicomponent group-based treatment was the worsening of symptoms. In contrast to some theoretical considerations, our data did not support the assumption that NEs are a prerequisite for therapeutic change. In contrast, they were related to worse patient outcomes. While we failed to find any

relationship between demographic characteristics and NEs, the diagnosis and baseline symptom severity of patients better predicted these effects. Other patient-related predictors included alexithymia and attachment avoidance; we also found psychological mindedness served as a protective factor. Although group cohesion did not yield statistically significant effects, the working alliance and patient problem actuation did. As the occurrence of NEs seems to be connected with worse outcomes, practitioners should pay special attention to patients who are more likely to report these effects. Our results suggest that especially those patients who experience higher levels of distress at the beginning of treatment, who perceive the working alliance with their therapists as problematic, and who experience high in-session emotional arousal related to their problem are more likely to report NEs. As psychotherapists, we may need to develop higher sensitivity to signs of patients' discomfort and encourage patients to share their negative treatment experiences.

### Disclosure Statement

No potential conflict of interest was reported by the author(s).

### Funding

Czech Science Foundation Grant GA18-08512S and Czech Ministry of Education, Youth, and Sports Grant MUNI/A/1216/2020 funded this work.

The data and the R script will be made available upon reasonable request for three years after the end of the project. Afterwards, it will be made available at <https://osf.io/dfrma/>

### ORCID

Martina Pourová  <http://orcid.org/0000-0003-1769-9439>

Tomáš Řiháček  <http://orcid.org/0000-0001-5893-9289>

Luboš Chvála  <http://orcid.org/0000-0002-4479-5699>

Zbyněk Vybíral  <http://orcid.org/0000-0003-2239-4859>

Jan R. Boehnke  <http://orcid.org/0000-0003-0249-1870>

### References

Akerman, G., & Geraghty, K. A. (2016). An exploration of clients' experiences of group therapy. *Therapeutic Communities: The*

- International Journal of Therapeutic Communities*, 37(2), 101–108. <https://doi.org/10.1108/TC-12-2015-0026>
- Alexopoulos, G. S., Raue, P. J., Banerjee, S., Mauer, E., Marino, P., Soliman, M., Kanellopoulos, D., Solomonov, N., Adeagbo, A., Sirey, J. A., Hull, T. D., Kiosses, D. N., & Areán, P. A. (2021). Modifiable predictors of suicidal ideation during psychotherapy for late-life major depression. A machine learning approach. *Translational Psychiatry*, 11(1), 1–8. <https://doi.org/10.1038/s41398-021-01656-5>
- Allredge, C. T., Burlingame, G. M., Yang, C., & Rosendahl, J. (2021). Alliance in group therapy: A meta-analysis. *Group Dynamics: Theory, Research, and Practice*, 25(1), 13–28. <https://doi.org/10.1037/gdn0000135>
- Barnes, M., Sherlock, S., Thomas, L., Kessler, D., Kuyken, W., Owen-Smith, A., Lewis, G., Wiles, N., & Turner, K. (2013). No pain, no gain: Depressed clients' experiences of cognitive behavioural therapy. *British Journal of Clinical Psychology*, 52(4), 347–364. <https://doi.org/10.1111/bjc.12021>
- Bech, P., Olsen, L., Kjoller, M., & Rasmussen, N. (2003). Measuring well-being rather than the absence of distress symptoms: A comparison of the SF-36 mental health subscale and the WHO-five well-being scale. *International Journal of Methods in Psychiatric Research*, 12(2), 85–91. <https://doi.org/10.1002/mpr.145>
- Beitel, M., Ferrer, E., & Cecero, J. J. (2004). Psychological mindedness and cognitive style. *Journal of Clinical Psychology*, 60(6), 567–582. <https://doi.org/10.1002/jclp.10258>
- Bernard, H., Burlingame, G., Flores, P., Greene, L., Joyce, A., Kobos, J. C., Leszcz, M., Macnair Semands, R. R., Piper, W. E., Slocum Mceneaney, A. M., & Feirman, D. (2008). Clinical practice guidelines for group psychotherapy. *International Journal of Group Psychotherapy*, 58(4), 455–542. <https://doi.org/10.1521/ijgp.2008.58.4.455>
- Brakemeier, E. L., Radtke, M., Engel, V., Zimmermann, J., Tuschen-Caffier, B., Hautzinger, M., Schramm, E., Berger, M., & Normann, C. (2015). Overcoming treatment resistance in chronic depression: A pilot study on outcome and feasibility of the cognitive behavioral analysis system of psychotherapy as an inpatient treatment program. *Psychotherapy and Psychosomatics*, 84(1), 51–56. <https://doi.org/10.1159/000369586>
- Burlingame, G. M., McClendon, D. T., & Yang, C. (2018). Cohesion in group therapy: A meta-analysis. *Psychotherapy*, 55(4), 384–398. <https://doi.org/10.1037/pst0000173>
- Chvála, L., Řiháček, T., Polakovská, L., Vybíral, Z., & Rozental, A. (2020). Adaptation of the negative effects questionnaire into the Czech context. *Psychotherapy Research*, 30(4), 510–519. <https://doi.org/10.1080/10503307.2019.1622053>
- Curran, J., Parry, G. D., Hardy, G. E., Darling, J., Mason, A. M., & Chambers, E. (2019). How does therapy harm? A model of adverse process using task analysis in the meta-synthesis of service users' experience. *Frontiers in Psychology*, 10, 347. <https://doi.org/10.3389/fpsyg.2019.00347>
- Dalenberg, C., & Carlson, E. (2010). *New versions of the Dissociative experiences scale: The DES-R (revised) and the DES-B (brief)*. Paper presented at annual meeting of the International Society for traumatic stress studies, November, Montreal, Quebec.
- Daňšová, P., Masopustová, Z., Hanáčková, V., Kicková, K., & Korábová, I. (2016). Metoda patient health questionnaire-9: Česká verze [The patient health questionnaire-9: The Czech version]. *Ceskoslovenska Psychologie*, 60(5), 468–481.
- Diener, M. J., & Monroe, J. M. (2011). The relationship between adult attachment style and therapeutic alliance in individual psychotherapy: A meta-analytic review. *Psychotherapy*, 48(3), 237–248. <https://doi.org/10.1037/a0022425>

- Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V. K., Koss, M. P., & Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. *American Journal of Preventive Medicine*, 14(4), 245–258. [https://doi.org/10.1016/S0749-3797\(98\)00017-8](https://doi.org/10.1016/S0749-3797(98)00017-8)
- Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, 55(4), 316–340. <https://doi.org/10.1037/pst0000172>
- Fraleigh, R. C., Heffernan, M. E., Vicary, A. M., & Brumbaugh, C. C. (2011). The experiences in close relationships—relationship structures questionnaire: A method for assessing attachment orientations across relationships. *Psychological Assessment*, 23(3), 615–625. <https://doi.org/10.1037/a0022898>
- Gassmann, D., & Grawe, K. (2006). General change mechanisms: The relation between problem activation and resource activation in successful and unsuccessful therapeutic interactions. *Clinical Psychology & Psychotherapy*, 13(1), 1–11. <https://doi.org/10.1002/cpp.442>
- Gori, A., Giannini, M., Palmieri, G., Salvini, R., & Schulberg, D. (2012). A new measure for alexithymia: Psychometric properties of the Psychological Treatment Inventory Alexithymia Scale (PTI-AS). *Psychology*, 03(03), 231–236. <http://doi.org/10.4236/psych.2012.33032>
- Grafanaki, S., & McLeod, J. (2002). Experiential congruence: Qualitative analysis of client and counsellor narrative accounts of significant events in time-limited person-centred therapy. *Counselling and Psychotherapy Research*, 2(1), 20–32. <https://doi.org/10.1080/14733140212331384958>
- Grant, A. M. (2001). Rethinking psychological mindedness: Metacognition, self-reflection, and insight. *Behaviour Change*, 18(1), 8–17. <https://doi.org/10.1375/bech.18.1.8>
- Grawe, K. (2004). *Psychological therapy*. Hogrefe & Huber.
- Hoffmann, D., Rask, C. U., Hedman-Lagerlöf, E., Jensen, J. S., & Frostholm, L. (2021). Efficacy of internet-delivered acceptance and commitment therapy for severe health anxiety: Results from a randomized, controlled trial. *Psychological Medicine*, 51(15), 2685–2695. <https://doi.org/10.1017/S0033291720001312>
- Holmes, S. E., & Kivlighan Jr, D. M. (2000). Comparison of therapeutic factors in group and individual treatment processes. *Journal of Counseling Psychology*, 47(4), 478–484. <https://doi.org/10.1037/0022-0167.47.4.478>
- Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy*, 48(1), 9–16. <https://doi.org/10.1037/a0022186>
- Johansson, L., & Werbart, A. (2009). Patients' views of therapeutic action in psychoanalytic group psychotherapy. *Group Analysis*, 42(2), 120–142. <https://doi.org/10.1177/0533316409104361>
- Joyce, A. S., Piper, W. E., & Ogradniczuk, J. S. (2007). Therapeutic alliance and cohesion variables as predictors of outcome in short-term group psychotherapy. *International Journal of Group Psychotherapy*, 57(3), 269–296. <https://doi.org/10.1521/ijgp.2007.57.3.269>
- Kauhanen, J., Kaplan, G. A., Cohen, R. D., Julkunen, J., & Salonen, J. T. (1996). Alexithymia and risk of death in middle-aged men. *Journal of Psychosomatic Research*, 41(6), 541–549. [https://doi.org/10.1016/S0022-3999\(96\)00226-7](https://doi.org/10.1016/S0022-3999(96)00226-7)
- Kenny, D. A. (2020). *Measuring model fit*. <http://davidakenny.net/cm/fit.htm>.
- Kissane, D. W., & Clark, D. M. (2002). Demoralization: Its phenomenology and importance. *Australian & New Zealand Journal of Psychiatry*, 36(6), 733–742. <https://doi.org/10.1046/j.1440-1614.2002.01086.x>
- Klocek, A., Řiháček, T., & Cigler, H. (2020). Psychometric evaluation of the Czech version of group cohesiveness scale (GCS) in a clinical sample: A Two-dimensional model. *Frontiers in Psychology*, 11, 3404. <https://doi.org/10.3389/fpsyg.2020.595651>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9. *Journal of General Internal Medicine*, 16(9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Kronström, K., Salminen, J. K., Hietala, J., Kajander, J., Vahlberg, T., Markkula, J., Rasi-Hakala, H., & Karlsson, H. (2009). Does defense style or psychological mindedness predict treatment response in major depression? *Depression and Anxiety*, 26(7), 689–695. <https://doi.org/10.1002/da.20585>
- Ladwig, I., Rief, W., & Nestoriuc, Y. (2014). Welche risiken und nebenwirkungen hat psychotherapie? - entwicklung des inventars zur erfassung negativer effekte von psychotherapie (INEP). *Verhaltenstherapie*, 24(4), 252–263. <https://doi.org/10.1159/000367928>
- Lampropoulos, G. K. (2001). Common processes of change in psychotherapy and seven other social interactions. *British Journal of Guidance & Counselling*, 29(1), 21–33. <https://doi.org/10.1080/03069880020019356>
- Linden, M. (2013). How to define, find and classify side effects in psychotherapy: From unwanted events to adverse treatment reactions. *Clinical Psychology & Psychotherapy*, 20(4), 286–296. <https://doi.org/10.1002/cpp.1765>
- Linden, M., Muschalla, B., & Walter, M. (2020). Gender and side effects of group cognitive behavior psychotherapy. *Archives of Psychiatry and Mental Health*, 4(1), 014–018. <https://doi.org/10.29328/journal.apmh.1001012>
- Lorenc, T., Lester, S., Sutcliffe, K., Stansfield, C., & Thomas, J. (2020). Interventions to support people exposed to adverse childhood experiences: Systematic review of systematic reviews. *BMC Public Health*, 20(1), 1–10. <https://doi.org/10.1186/s12889-020-08789-0>
- Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. *Medical Care*, 46(3), 266–274. <https://www.jstor.org/stable/40221654> <https://doi.org/10.1097/MLR.0b013e318160d093>
- Macdonald, G., Livingstone, N., Hanratty, J., McCartan, C., Cotmore, R., Cary, M., Glaser, D., Byford, S., Welton, N. J., Bosqui, T., Bowes, L., Audrey, S., Mezey, G., Fisher, H., Riches, W., & Churchill, R. (2016). The effectiveness, acceptability and cost-effectiveness of psychosocial interventions for maltreated children and adolescents: An evidence synthesis. *Health Technology Assessment*, 20(69), 1–508. <https://doi.org/10.3310/hta20690>
- Mander, J. V., Wittorf, A., Schlarb, A., Hautzinger, M., Zipfel, S., & Sammet, I. (2013). Change mechanisms in psychotherapy: Multiperspective assessment and relation to outcome. *Psychotherapy Research*, 23(1), 105–116. <https://doi.org/10.1080/10503307.2012.744111>
- Marmarosh, C. L. (2021). Ruptures and repairs in group psychotherapy: Introduction to the special issue. *Group Dynamics: Theory, Research, and Practice*, 25(1), 1–12. <https://doi.org/10.1037/gdn0000150>
- Maroti, D., Ek, J., Widlund, R. M., Schubiner, H., Lumley, M. A., Lillengren, P., Bileviciute-Ljungar, I., Ljótsson, B., & Johansson, R. (2021). Internet-Administered emotional awareness and expression therapy for somatic symptom disorder With centralized symptoms: A preliminary efficacy trial. *Frontiers in Psychiatry*, 12, 104. <https://doi.org/10.3389/fpsyg.2021.620359>
- Marziali, E., Munroe-Blum, H., & McCleary, L. (1997). The contribution of group cohesion and group alliance to the outcome



- of group psychotherapy. *International Journal of Group Psychotherapy*, 47(4), 475–497. <https://doi.org/10.1080/00207284.1997.11490846>
- Moreira, H., Martins, T., Gouveia, M. J., & Canavarro, M. C. (2015). Assessing adult attachment across different contexts: Validation of the Portuguese version of the experiences in close relationships–relationship structures questionnaire. *Journal of Personality Assessment*, 97(1), 22–30. <https://doi.org/10.1080/00223891.2014.950377>
- Muschalla, B., & Schönborn, F. (2021). Induction of false beliefs and false memories in laboratory studies—A systematic review. *Clinical Psychology & Psychotherapy*, 28(5), 1194–1209. <https://doi.org/10.1002/cpp.2567>
- Nissen-Lie, H. A., Monsen, J. T., & Rønnestad, M. H. (2010). Therapist predictors of early patient-rated working alliance: A multilevel approach. *Psychotherapy Research*, 20(6), 627–646. <https://doi.org/10.1080/10503307.2010.497633>
- Nunes da Silva, A., Vasco, A., & Watson, J. (2018). Alexithymia and therapeutic alliance: A multiple case study comparing good and poor outcome cases. *Research in Psychotherapy: Psychopathology, Process and Outcome*, 21(2), 83–97. <https://doi.org/10.4081/ripppo.2018.313>
- Nykliček, I., & Denollet, J. (2009). Development and evaluation of the balanced index of psychological mindedness (BIPM). *Psychological Assessment*, 21(1), 32–44. <https://doi.org/10.1037/a0014418>
- Ogrodniczuk, J. S., Piper, W. E., & Joyce, A. S. (2006). Treatment compliance in different types of group psychotherapy. *Journal of Nervous & Mental Disease*, 194(4), 287–293. <https://doi.org/10.1097/01.nmd.0000207366.49820.85>
- Ogrodniczuk, J. S., Piper, W. E., & Joyce, A. S. (2011). Effect of alexithymia on the process and outcome of psychotherapy: A programmatic review. *Psychiatry Research*, 190(1), 43–48. <https://doi.org/10.1016/j.psychres.2010.04.026>
- Parker, G., Fletcher, K., Berk, M., & Paterson, A. (2013). Development of a measure quantifying adverse psychotherapeutic ingredients: The experiences of therapy questionnaire (ETQ). *Psychiatry Research*, 206(2-3), 293–301. <https://doi.org/10.1016/j.psychres.2012.11.026>
- Pourová, M., Řiháček, T., Boehnke, J. R., Šimek, J., Saic, M., Kabát, J., & Šilhán, P. (2022). Effectiveness of a multicomponent group-based treatment in patients with medically unexplained physical symptoms: A multisite naturalistic study. Manuscript submitted for publication.
- Quirk, K., Miller, S., Duncan, B., & Owen, J. (2013). Group session rating scale: Preliminary psychometrics in substance abuse group interventions. *Counselling and Psychotherapy Research*, 13(3), 194–200. doi:10.1080/14733145.2012.744425
- Ranankan, M. (2014). Clients' positive and negative experiences of experiential art therapy group process. *The Arts in Psychotherapy*, 41(2), 193–204. <https://doi.org/10.1016/j.aip.2014.02.006>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>. <https://www.R-project.org/>
- Řiháček, T. (2018). Effectiveness of psychotherapy in patients with medically unexplained physical symptoms. *ISRCTN*. <https://doi.org/10.1186/ISRCTN13532466>
- Řiháček, T., Čeveliček, M., Boehnke, J. R., Pourová, M., & Roubal, J. (2022). Mechanisms of change in multicomponent group-based treatment for patients suffering from medically unexplained physical symptoms. *Psychotherapy Research*. <https://doi.org/10.1080/10503307.2022.2061874>
- Roback, H. B. (2000). Adverse outcomes in group psychotherapy: Risk factors, prevention, and research directions. *The Journal of Psychotherapy Practice and Research*, 9(3), 113–122.
- Robak, R. W., Kangos, K. A., Chiffrieller, S. H., & Griffin, P. W. (2013). The working alliance in group counseling: An exploratory study. *Psychological Reports*, 113(2), 591–604. <https://doi.org/10.2466/21.02.PR0.113x19z7>
- Rodziński, P., Ostachowska, A., Cyranka, K., Rutkowski, K., Sobański, J. A., Citkowska-Kisielewska, A., Dembińska, E., & Mielimaka, M. (2019). Neurotic personality traits and associated dysfunctional attitudes as factors predisposing patients to suicidal ideations at the end of intensive psychotherapy. *Psychiatria Polska*, 53(4), 883–899. <https://doi.org/10.12740/PP/89795>
- Rosseel, Y. (2012). lavaan: An Rpackage for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Rozental, A., Andersson, G., Boettcher, J., Ebert, D. D., Cuijpers, P., Knaevelsrud, C., Ljótsson, B., Kaldo, V., Titov, N., & Carlbring, P. (2014). Consensus statement on defining and measuring negative effects of internet interventions. *Internet Interventions*, 1(1), 12–19. <https://doi.org/10.1016/j.invent.2014.02.001>
- Rozental, A., Andersson, G., & Carlbring, P. (2019). In the absence of effects: An individual patient data meta-analysis of non-response and its predictors in internet-based cognitive behavior therapy. *Frontiers in Psychology*, 10(Article 589), 1–15. <https://doi.org/10.3389/fpsyg.2019.00589>
- Rozental, A., Castonguay, L., Dimidjian, S., Lambert, M., Shafran, R., Andersson, G., & Carlbring, P. (2018). Negative effects in psychotherapy: Commentary and recommendations for future research and clinical practice. *BjPsych Open*, 4(4), 307–312. <https://doi.org/10.1192/bjo.2018.42>
- Rozental, A., Kottorp, A., Boettcher, J., Andersson, G., & Carlbring, P. (2016). Negative effects of psychological treatments: An exploratory factor analysis of the Negative Effects Questionnaire for monitoring and reporting adverse and unwanted events. *PLoS ONE*, 11(6), e0157503. <https://doi.org/10.1371/journal.pone.0157503>
- Schaeuffele, C., Homeyer, S. L., Perea, L., Scharf, L., Schulz, A., Knaevelsrud, C., Renneberg, B., & Boettcher, J. (2020). The unified protocol as an internet-based intervention for emotional disorders: Randomized controlled trial. *PsyArXiv*, 1–26. <https://doi.org/10.31234/osf.io/528tw>
- Schermuly-Haupt, M.-L., Linden, M., & Rush, A. J. (2018). Unwanted events and side effects in cognitive behavior therapy. *Cognitive Therapy and Research*, 42(3), 219–229. <https://doi.org/10.1007/s10608-018-9904-y>
- Schneibel, R., Wilbertz, G., Scholz, C., Becker, M., Brakemeier, E. L., Bschor, T., Zobel, I., & Schmoll, D. (2017). Adverse events of group psychotherapy in the in-patient setting - results of a naturalistic trial. *Acta Psychiatrica Scandinavica*, 136(3), 247–258. <https://doi.org/10.1111/acps.12747>
- Shimokawa, K., Lambert, M., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, 78(3), 298–311. <https://doi.org/10.1037/a0019247>
- Smits, N., van der Ark, L. A., & Conijn, J. M. (2018). Measurement versus prediction in the construction of patient-reported outcome questionnaires: Can we have our cake and eat it? *Quality of Life Research*, 27(7), 1673–1682. <https://doi.org/10.1007/s11136-017-1720-4>
- Spitzer, C., Barnow, S., Freyberger, H. J., & Joergen Grabe, H. (2007). Dissociation predicts symptom-related treatment outcome in short-term inpatient psychotherapy. *Australian & New Zealand Journal of Psychiatry*, 41(8), 682–687. <https://doi.org/10.1080/00048670701449146>
- Strauss, B. (2021). “You can’t make an omelet without breaking eggs”: Studies on side effects and adverse events in group



- psychotherapy. *International Journal of Group Psychotherapy*, 71 (3), 472–480. <https://doi.org/10.1080/00207284.2021.1890089>
- Strauss, B., Gawlytta, R., Schleu, A., & Frenzl, D. (2021). Negative effects of psychotherapy: Estimating the prevalence in a random national sample. *BjPsych Open*, 7(6), e186. <https://doi.org/10.1192/bjo.2021.1025>
- Strauss, B. M., Lutz, W., Steffanowski, A., Wittmann, W. W., Boehnke, J. R., Rubel, J., Scheidt, C. E., Caspar, F., Vogel, H., Altmann, U., Steyer, R., Zimmermann, A., Bruckmayer, E., von Heymann, F., Kramer, D., & Kirchmann, H. (2015). Benefits and challenges in practice-oriented psychotherapy research in Germany: The TK and the QS-PSY-BAY projects of quality assurance in outpatient psychotherapy. *Psychotherapy Research*, 25(1), 32–51. <https://doi.org/10.1080/10503307.2013.856046>
- van de Leur, J. C., Buhrman, M., Åhs, F., Rozental, A., & Jansen, G. B. (2020). Standardized multimodal intervention for stress-induced exhaustion disorder: An open trial in a clinical setting. *BMC Psychiatry*, 20(1), 1–14. <https://doi.org/10.1186/s12888-020-02907-3>
- Vanheule, S., Desmet, M., Meganck, R., & Bogaerts, S. (2007). Alexithymia and interpersonal problems. *Journal of Clinical Psychology*, 63(1), 109–117. <https://doi.org/10.1002/jclp.20324>
- Wongpakaran, T., Wongpakaran, N., Intachote-Sakamoto, R., & Boripuntakul, T. (2013). The group cohesiveness scale (GCS) for psychiatric inpatients. *Perspectives in Psychiatric Care*, 49(1), 58–64. <https://doi.org/10.1111/j.1744-6163.2012.00342.x>
- Zoet, H. A., Wagenmans, A., van Minnen, A., & de Jongh, A. (2018). Presence of the dissociative subtype of PTSD does not moderate the outcome of intensive trauma-focused treatment for PTSD. *European Journal of Psychotraumatology*, 9(1), 1468707. <https://doi.org/10.1111/j.1744-6163.2012.00342.x>