European Archives of Psychiatry and Clinical Neuroscience (2019) 269:577–586
https://doi.org/10.1007/s00406-018-0931-1

ORIGINAL PAPER

It can't hurt, right? Adverse effects of psychotherapy in patients with depression

Steffen Moritz[1] · Yvonne Nestoriuc[2] · Winfried Rief[3] · Jan Philipp Klein[4] · Lena Jelinek[1] · Judith Peth[1]

Received: 16 May 2018 / Accepted: 23 July 2018 / Published online: 7 August 2018

Abstract

Despite growing awareness of occasional adverse effects of psychological treatments, only a few instruments cover side effects and other unwanted effects of psychotherapy. For the present study, the Positive and Negative Effects of Psychotherapy Scale (PANEPS) was evaluated in a population of individuals with depression who had completed at least one course of face-to-face psychotherapy. A total of 135 individuals with a current or previous depressive episode as verified by a diagnostic interview filled out the online version of the PANEPS, which is designed to capture both positive and adverse events. Factor analysis yielded four dimensions: positive effects, side effects, malpractice, and unethical conduct. Internal consistency of the individual subscales was satisfactory to excellent (Cronbach's α: 0.72 and 0.92). Positive effects were reported by virtually all patients (95.6%). At the same time, approximately half of the sample noted at least one adverse event (52.6%). Among these, side effects (38.5%) and malpractice (26.7%) were significantly more prevalent than unethical conduct (8.1%). As expected, positive effects were negatively correlated with adverse events. Our results challenge the common clinical assumption that some degree of destabilization is necessary for symptom improvement. The survey was conducted anonymously, and the sample underwent diagnostic verification. The results indicate a need for improved treatment guidelines and mechanisms to monitor treatment.

Keywords Side effects · Adverse events · Psychotherapy · Depression · Malpractice unethical conduct

Introduction

Meta-analyses confirm that psychotherapy, especially cognitive behavioral therapy (CBT) and its variants, is effective in the treatment of psychological disorders. The evidence is especially large and consistent for anxiety disorders [1], obsessive–compulsive disorder (OCD) [2], and depression [1, 3] as well as borderline personality disorder [4, 5]. Even for disorders traditionally regarded as predominantly biologically determined and allegedly less amenable to psychological understanding and treatment, such as psychosis [6] and bipolar disorder [7], the evidence for the efficacy of psychotherapy is encouraging. Accordingly, national and international guidelines recommend (or even require) psychotherapy for the treatment of most psychiatric disorders (e.g., the NICE guidelines for United Kingdom or the DGPPN guidelines for Germany). For some disorders, particularly OCD, the effects of certain forms of psychotherapy seem to exceed those of psychopharmacotherapy [2]. For depression, the evidence is equivocal as to whether or not psychotherapy is more efficacious than psychopharmacological treatment [3, 8], but the usefulness of psychotherapy is beyond doubt and has also been demonstrated for short-term interventions [9].

The appreciation of psychotherapy as an effective standard treatment is in large part owing to the implementation of rigorous methodological standards in recent years, particularly the adaptation of CONSORT guidelines for nonpharmacological interventions [10]. High-quality studies are an indispensable prerequisite before a particular

✉ Steffen Moritz
moritz@uke.uni-hamburg.de

1 Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

2 Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

3 Department of Clinical Psychology and Psychotherapy, Philipps-University of Marburg, Marburg, Germany

4 Department of Psychiatry and Psychotherapy, Lübeck University, Lübeck, Germany

psychotherapeutic approach is recommended by clinical guidelines. Early research on psychotherapy usually did not pay much attention to the blinding of interviewers, sample size calculation, a priori definition of outcomes, and trial registration; neglecting any of these measures, however, may easily distort results due to biases such as post hoc "cherry-picking" or allegiance effects [11].

Adverse events (AE) of psychotherapy were only occasionally noted in the older literature [12]. It was not until the adoption of more rigorous research designs borrowed from pharmacological research that AE were routinely assessed [13, 14], although the criteria remain ill-defined. Often, the thresholds of symptom worsening (e.g., 20% worsening) that warrant thorough further examination by the principal investigator or an independent data safety monitoring board are arbitrary. Perhaps due to the complexity of psychotherapy, researchers continue to remain somewhat unclear as to what constitutes an AE and whether to attribute AE to psychotherapy or other causes. In recent years, a number of scales have been devised that tap AE in greater detail [15–17]. Following the nomenclature of pharmacological trials [18], AE are divided into side effects (SE), malpractice (MP), and unethical conduct (UC; see Fig. 1). SE are AE following proper application of an evidence-based treatment, while MP means that an evidence-based treatment is not applied properly or an inappropriate technique is adopted. Research shows that evidence-based and effective psychological treatments, for example, exposure, can lead to SE, even when applied appropriately [17, 19–21]. UC relates to a spectrum of abusive behaviors and other grossly inappropriate behaviors that at the extreme end of the spectrum may result in criminal prosecution (such as sexual abuse). The intimate situation of psychotherapy and the elevated vulnerability of the patient (e.g., proneness to emotional dependency) pose special risks. Although abuse has been acknowledged as a problem for decades [22, 23], estimated frequencies vary greatly across studies, and actual prevalence rates are still not known. The situation is similar for SE, where rates vary between 5.9% [24] and 93% [17], reflecting differences in populations (patients with different disorders, different therapists), settings (individual, group), reference points (one course of psychotherapy versus lifetime of experience with psychotherapy), and level of blindness of the assessment (e.g., interview, anonymous online assessment).

Growing knowledge about the SE of psychotherapy has created a dilemma for psychotherapists that has long been a dilemma for physicians. Informing patients about possible AE might exert a detrimental effect on outcome and increase the chances that an AE will in fact occur [25, 26]. However, failing to do so is clearly unprofessional and against the code of conduct as treatment failure may elicit fatalism (e.g., being chronically ill or treatment-unresponsive in general) in patients who are unprepared for such events. Thus, psychotherapists have to find a balance between imparting hope to the patient by stressing the potential of psychotherapeutic treatment, which is known to foster a better outcome

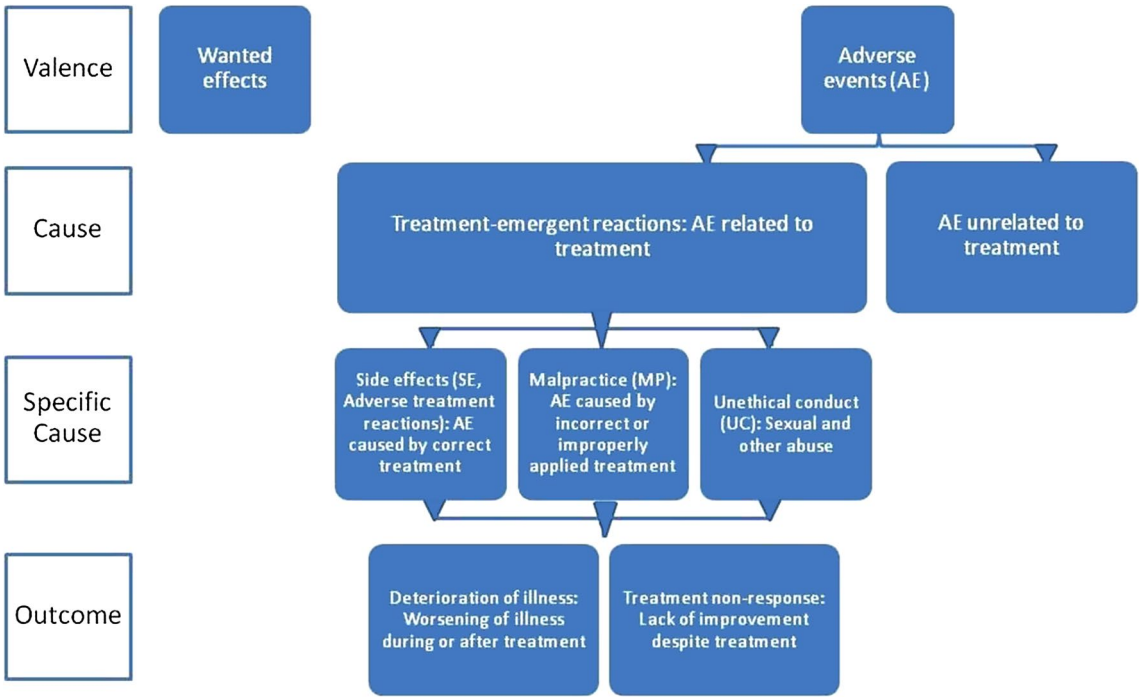

Fig. 1 Wanted (positive) effects and adverse events in psychotherapy

[27], and adequately informing the patient about the possible risks of the treatment.

The present study was concerned with AE in depression. We used the PANEPS, a revised and shortened version of the Side-Effects of Psychotherapy Scale (SEPS) [17], which has been tested in patients with OCD before but not in patients with depression. The scale taps both positive effects and AE to inform clinicians about possible associations and to verify/falsify anecdotal reports that some destabilization may be a prerequisite for improvement [28]. In a prior study, no evidence was found for this "no pain, no gain" hypothesis [17]. To the contrary, it was found that the more SE the patient reported, the poorer the outcome. Clearly, it remains to be tested whether an association determined in one population generalizes to others. We also implemented several changes to the scale. For example, we asked patients to base their responses on their most recent course of psychotherapy to reduce memory biases/forgetfulness. We also deleted some complex or unclear items. Moreover, we recruited a large sample to determine the factorial structure of the scale.

We administered the PANEPS items online to ensure the anonymity of the participants. Direct interviews or disclosure of personal information may result in an under-estimation of AE [16], particularly MP and UC, as patients may be afraid to disclose certain incidents during therapy for fear of legal consequences or of their identity becoming public if the authorities are informed. At the same time, we paid special attention to recruiting a well-characterized population by approaching patients who had undergone a clinical diagnostic interview and had consented to be recontacted for future scientific studies.

The study aimed to elucidate the psychometric properties of the PANEPS and the prevalence of AE in a well-characterized sample of patients with depression.

Methods

Study process and sample

The study was carried out from February to March 2015 as an anonymous online survey using Globalpark® (UniPark, version 10.5) and was in accordance with the updated Helsinki Declaration.

Patients with depressive disorder who had previously participated in an online intervention study [29, 30] were contacted by e-mail if they met the following inclusion criteria: a current or past depressive episode and completion of at least one course of face-to-face psychotherapy (an online intervention was not counted). Exclusion criteria were as follows: age under 18 or over 65 years, psychotic or manic symptoms, an abnormal response pattern during the interview/assessment (e.g., entering the same value throughout the questionnaires), or missing data on the PANEPS. During the course of this study, patients underwent a telephone screening that used the Mini International Neuropsychiatric Interview [31] to ascertain whether the patient met the inclusion criteria. The Hamilton Depression Rating Scale (HDRS) [32] was used for clinician assessment of depression severity. All patients displayed mild-to-moderate depressive symptoms (score: 5–14 points) on the Patient Health Questionnaire (PHQ-9) [33].

We contacted 209 persons in total, of whom 186 agreed to participate. A total of 135 participants fulfilled the study criteria and were considered for the final analyses (see Table 1 for a description of the sample).

The online survey recorded current depressive symptoms using the PHQ-9 [33]. This self-rating screening instrument taps the nine diagnostic criteria of depression in the DSM-IV (response options: 0 = not at all, 1 = several days, 2 = more than half the days, 3 = nearly every day). The total possible score ranges from 0 to 27. The psychometric properties of the scale are good [34].

In addition, we asked for further diagnostic information (e.g., comorbid disorders) and information pertaining to the most recently completed course of psychotherapy (e.g., type of treatment, number of sessions, premature termination). As compensation for their participation, we offered patients self-help material consisting of relaxation exercises upon completion of the survey.

The PANEPS is derived from the SEPS [17] and assesses both wanted and unwanted effects of psychotherapy. The item set ($N = 97$) of the SEPS was developed following discussions with patients and leaders of self-help groups and also took into consideration the content of existing scales, such as the Unwanted Events and Adverse Treatment Reaction Checklist for Psychotherapy (UE–ATR Checklist) [15] and the Inventory for the Assessment of Negative Effects of Psychotherapy (INEP) [16]. The SEPS/PANEPS follow the nomenclature proposed by Linden [15], which in turn is borrowed from the pharmacological literature. The SEPS subscales were eventually composed via an expert panel consisting of four psychologists; items were only allocated to a subscale by a unanimous decision. The PANEPS AE subscale is highly correlated with the INEP severity of negative effects ($r = 0.631$, $p < 0.001$) and misconduct subscales ($r = 0.744$, $p < 0.001$) speaking for its validity.

As some items were overly long or complicated, we recommended a revision and shortening of the SEPS in the pilot study [17]. Redundant items and items with a clear association with OCD were deleted as well as items with long and complicated content [17]. The PANEPS consists of 43 items. Responses had to be given on a four-point Likert scale (1 = true, 2 = rather true, 3 = rather not true, 4 = not true). An event/effect was coded to be present if participants endorsed the respective item ("true" or "rather true"). For some items,

Table 1 Characteristics of the sample ($N = 135$)

Variable	Category	*N*/*M* (*SD*)	%/Range
Age (in years)		46.85 (10.18)	20–67
Gender	Female	99	73.3
Education (in years)		11.82 (1.71)	8–18
PHQ-9 (total score)		8.42 (4.27)	0–21
PHQ-9 (total score > 10)		37	27.4
Number of comorbid diagnoses		0.83 (0.97)	0–4
Number of psychotherapies		2.10 (1.74)	1–15
Type of last treatment (according to patients)	Behavior therapy	44	32.5
	Cognitive behavioral therapy	14	10.4
	Psychodynamic psychotherapy	37	27.4
	Psychoanalysis	7	5.2
	Person-centered psychotherapy	21	15.6
	Interpersonal therapy	1	0.7
	Systemic therapy	2	1.5
	Unknown	9	6.7
Framework of last psychotherapeutic treatment	Outpatient	122	90.4
	Inpatient	12	8.9
	Day-patient care	1	0.7
Setting of last treatment	Single	112	83
	Group	2	1.5
	Both	21	15.6
Number of treatment sessions	1–5	6	4.4
	6–25	48	35.6
	26–120	67	49.6
	121–300	14	10.4
Regular termination of treatment	n/a	118	87.4
Cause of cancelation (multiple endorsements possible)	Not effective	5	3.7
	Therapeutic alliance	9	6.7
	Organizational reasons	3	2.2
	Other reasons	7	5.2
Gender of therapist	Female	93	68.9
Time since last treatment (in years)		2.70 (2.80)	0–13

the option "does not apply" could be chosen as well (e.g., if a patient was asked about effects of treatment on his or her partnership but had no partner at that time).

Results

Factor analysis of the PANEPS

To investigate the dimensional structure of the PANEPS, a principal component analysis was performed with a varimax rotation for all 43 items of the PANEPS. An assignment of individual items to a factor was made if (positive) factor loadings were at least 0.5. If one item loaded above this threshold on two subscales, the difference between the loadings on each of these subscales had to be greater than 0.1 to allow unequivocal allocation to a subscale. The principal component analysis resulted in the extraction of four components with an eigenvalue of greater than 1. In combination, these four factors explained 46.3% of the variance. The factors were labeled Positive Effects (PE, 5 items), Side Effects (SE, 7 items), Malpractice (MP, 11 items), and Unethical Conduct (UC, 6 items). A total of 29 items could be assigned to 1 of the 4 factors; 14 items showed no clear association to any of the factors and were accordingly excluded for subsequent analyses. The items of the PANEPS and their respective loadings on the four dimensions are displayed in Table 4 ("Appendix"). Cronbach's alpha was satisfactory to excellent (see Table 4 in "Appendix").

Frequency of positive effects (PE) and adverse events (AE)

The positive effects (PE) of psychotherapy were represented by five items. On average, three PE [($M = 3.33$, SD = 1.39; range: 0 ($n = 6$, 4.4%) to 5 ($n = 32$, 23.7%)] were reported (≥ 1 incident of positive effects: 95.6%). Four out of five patients ($n = 105$, 80.2%) agreed with the statement that they had more confidence in their own abilities after the therapy, and a similar number ($n = 101$, 77.1%) confirmed that they had learned to take responsibility for themselves. The PE scale mainly captured items relating to the psychological effects of psychotherapy. Three items that can be regarded as PE but were not part of the respective scale addressed psychosocial consequences: "My family is proud of me because I was in therapy" (33%), "The relationship with my partner has improved as a result of my therapy" (51.4), and "The therapy helped me to get a job" (44.3).

AE were represented by a total of 24 items (sum of the subscales MP, SE, and UC). On average, participants endorsed 2 AE [$M = 1.78$, $SD = 2.92$; range: 0 ($n = 64$, 47.4%) to 14 AE ($n = 1$)]. The frequency of AE is presented in Table 2. Of note, PE were affirmed more often than AE (although more AE items were posed), $t(134) = 4.988$, $p < .001$.

For AE, events of MP was endorsed most frequently [$M = 0.90$, SD = 1.91; range: 0 ($n = 99$, 73.3%) to 10 ($n = 1$)]. A total of 7.4% ($n = 10$) agreed to one item and 8.9% ($n = 12$) agreed to two to three items (≥ 1 incident of MP: 26.7%). The most common event of MP from the patients' perspective was the complaint that the therapy was only aimed at eliminating the problem without a positive goal orientation ($n = 23$, 17.4%), followed by positive responses to the statement "In my opinion, the applied therapeutic techniques were wrong" ($n = 22$, 16.9%).

SE were endorsed at a similar rate as MP [(pairwise comparison: $p = .430$; $M = 0.77$, SD = 1.28, range: 0 ($n = 83$, 61.5%) to 6 ($n = 2$, 1.5%)]; 19.3% ($n = 26$) endorsed one item and 8.9% ($n = 12$) endorsed two items (≥ 1 incident of SE: 38.5%). Fear of stigmatization was endorsed most frequently. One-fifth of the sample ($n = 27$, 20%) confirmed the statement "I am fearful that people I know will learn that I have been in therapy"; approximately every tenth participant ($n = 13$, 9.8%) noted interpersonal problems ("My relationships with my family and friends have deteriorated because of my therapy"). Of note, undergoing psychotherapy had a detrimental effect on insurance status in a subgroup of patients (16.5%), who endorsed the statement "My application for private insurance was rejected because I was in therapy that was paid for by my health insurance".

Unethical Conduct (UC) was rated the least often [$p < 0.001$ for all pairwise comparisons; $M = 0.11$, SD = 0.45; range: 0 ($n = 124$, 91.9%) to 4 items ($n = 1$)]. Yet, nine participants (6.7%) reported one incident of UC (≥ 1 incident of UE: 8.1%). Sexual abuse by the therapist was not reported by any of the respondents, but two patients (1.4%) endorsed the statement "During therapy, I had to endure sexual harassment by my therapist".

Table 2 Frequencies of adverse events (AE)

Number of AE	*n*	%	Cumulative %
0	64	47.4	47.4
1–3	51	38.8	86.2
4–8	12	8.9	95.1
9–12	5	3.6	98.7
> 12	3	2.2	100

Intercorrelations of PANEPS subscales and relationship to patient and therapist characteristics

A medium-to-strong correlation emerged between PE and MP and a small-to-medium but significant correlation emerged between PE and AE as well as SE and UC. The association between MP and UC was significant in the medium range (see Table 3).

Age was positively correlated with AE, particularly SE. Expectations (measured with a single item, 1 = very low, 7 = very high) regarding the most recent course of psychotherapy were not correlated with any of the PANEPS subscales (|rho| < 0.15, $p > 0.1$). Higher depressive symptoms were correlated with more PE and less AE, particularly SE.

We then conducted a series of t-tests that for exploratory purposes were not corrected for multiple comparisons. Completion of 13th grade ($n = 63$; below, $n = 71$) was associated with more AE, $t(133) = 2.281$, $p = .024$, particularly more UC $t(133) = 2.523$, $p = 0.013$, and MP, $t(133) = 2.796$, $p = 0.006$, but not more SE ($p = 0.582$). Gender of the patient (female: $n = 99$, male: $n = 36$) was not associated with any of the PANEPS subscales ($p > 0.7$), nor was gender of the therapist (female: $n = 93$, male: $n = 42$, $p > 0.1$), being in a partnership ($n = 79$, no: $n = 56$; $p > 0.2$), or whether or not there were diagnoses other than depression ($n = 64$; no additional diagnoses: $n = 71$; $p > 0.1$).

Subsequently, we compared patients who had undergone CBT ($n = 57$) versus psychodynamic therapies ($n = 43$). At trend level, the latter group showed less PE after treatment, $t(98) = 1.679$, $p = 0.096$. Patients in this group were significantly older, $t(98) = 2.044$, $p = 0.044$. When age was accounted for, the difference was nonsignificant ($p = 0.156$).

Psychological treatment by a physician ($n = 20$) was associated with more AE, $t(106) = 3.480$, $p = 0.001$, particularly more incidents of MP, in comparison to treatment provided by a psychologist ($n = 88$), $t(106) = 2.939$, $p = 0.005$. PE were

Table 3 Intercorrelation of PANEPS subscales and associations with demographic and psychometric information (Spearman's rho)

	Positive effects (PE)	Side effects (SE)	Malpractice (MP)	Unethical conduct (UC)	Adverse events total (AE)
Positive effects	–	−0.038	−0.454****	−0.064	−0.238**
Side effects	–	–	0.106	0.226**	0.670****
Malpractice	–	–	–	0.322****	699****
Unethical conduct	–	–	–	–	0.559****
Age	−0.034	0.250***	0.095	−0.056	0.180*
Expectations related to psychotherapy (1 = very low, 7 = very high)	−0.140	0.054	0.108	−0.017	0.089
PHQ-9 total	0.295***	−0.218*	−0.175*	−0.089	−0.212*

*$p < .05$, **$p < .01$, ***$p < .005$, ****$p < .001$

not different between groups (numerically, more PE were noted for psychologists), $t(106) = 1.709$, $p = 0.090$. Again, age was different between the two groups, $t(106) = 2.443$, $p = 0.016$, and when this was taken as a covariate, the difference for AE diminished to a trend ($p = 0.074$) and vanished for MP ($p = 0.143$).

Discussion

A total of 135 individuals with a current or previous depressive episode filled out the PANEPS, which both captures PE and AE. In line with prevalence rates, there was an excess of female participants [35]. Mirroring the situation in Germany [36], this was also true for therapists.

In line with meta-analyses demonstrating the efficacy of psychotherapy for depression [1, 3], virtually all patients in our study reported at least one PE of psychotherapy and—as in a forerunner study with OCD patients [17]—PE clearly outnumbered AE. Notable PE that were not part of the PE subscale were "The relationship with my partner has improved as a result of my therapy" (51.4%) and "The therapy helped me to get a job" (44.3%). While prevalence rates of PE were similar to the forerunner study (95.6% versus 95.3%), rates of AE were far lower (52.6% versus 92.9%), which is likely owing to different reference points (most recent course of psychotherapy versus lifetime experience with psychotherapy) and diagnosis (OCD versus depression). The rate, however, corresponds to recent estimates provided by Rheker and colleagues [37].

Corroborating findings from a prior trial [17], there was a negative correlation between PE and AE, which again brings into question the prominent clinical claim of "no pain, no gain". MP (e.g., administration of an evidence-based technique in a way that is not recommended by treatment manuals or guidelines) and SE (i.e., administration of an evidence-based technique in a way that is recommended by treatment manuals or guidelines but which leads to AE) were common (38.5% and 26.7%, respectively). Patients complained about (fears of) stigmatization and the treatment being aimed at eliminating the problem without a positive goal orientation, and some patients appraised the choice of therapeutic techniques as wrong. UC was reported the least often but was still high (8.1%); this included two reports of sexual harassment.

Older age was associated with more AE. Controlled designs are needed to test whether older people are more critical of treatment, more fearful of negative appraisal by others, make different attributions due to their greater life experience, or are more sensitive to SE of psychotherapy. Again, controlled designs are also needed to check whether AE are more frequent when physicians provide psychotherapy or when psychodynamic therapy is applied relative to CBT. Initial analyses tentatively supported these hypotheses, but the results were rendered nonsignificant when confounding variables were controlled for. The correlation between PE and symptom severity is counterintuitive at first but may point to a regression to the mean (i.e., more severely depressed patients benefit more from treatment).

The study had a number of limitations, some of which are inherent to psychotherapy research in general and are hard to rule out even when using the most scrupulous design. For example, no control group could be recruited. Neither self-reporting, as used in this study and which is prone to affective (e.g., mood-congruent memory) and cognitive biases (e.g., attributional biases), nor clinician-rated measurement of AE can provide a true picture about the prevalence of AE in view of potentially serious biases on both sides. As Linden pointed out [15], therapists are reluctant to disclose AE for fear that these events would be considered incidents of MP or misconduct for which they could be held legally accountable. Video recording therapeutic sessions may protect patients against misconduct and would likely raise psychotherapists' fidelity to guidelines relative to the naturalistic (unobserved) therapeutic context but might introduce other problems such as privacy. Therefore, studies

using such procedures would underestimate the true prevalence of AE. Patients, on the other hand, may not be able to distinguish between a true SE, MP, or other influences or sources (e.g., certain symptoms may be due to concomitant medication). However, some therapists are also poor at identifying AE [38]. In view of the detrimental effects of UC and incorrectly applied treatment on patients' well-being, we need to take up the challenge and address this topic despite the aforementioned and perhaps insolvable caveats. In line with Linden [15], therapists' sensitivity regarding this topic must be increased; therapists should routinely check for AE. Yet, we urgently need checklists similar to those available for medications regarding which SE are to be expected when using a particular therapeutic technique. Moreover, psychotherapists should undergo supervision for each patient and supervisors should ask about AE, whether or not they are related to treatment.

As stated, we cannot assume that patients would know whether a treatment technique is carried out *lege artis* and which treatments are evidence-based or not. This may be one reason why patients with more education in the forerunner study reported more SE than those with less education. For the present study, patients with more education reported more incidents of MP and UC. People with more education might know more about potential AE, which could enhance the probability of AE to emerge, and/or they might be better able to identify treatment-related AE as such (regardless of whether or not the incidence was indeed higher than in those with less education). These challenges to reliability are serious but should not lead to fatalism. While the saying "the customer is always right" may never fully apply to psychotherapeutic contexts, as a minimal commitment, clinicians should strive to maximize the satisfaction of users of the health care system, and the client's perspective must therefore be taken very seriously irrespective of its validity [39]. A number of factors speak to the truthfulness of the patients' responses in the current study: the study was anonymous; the psychometric characteristics of the scales were satisfactory to good; and, as mentioned before, the most severe AE, sexual abuse, was not endorsed, which might be expected if participants sought to sabotage the survey. Future studies should try to amalgamate clients' and therapists' perspectives on AE and also incorporate the views of supervisors—this could be aided by pre–post assessments. Although this and a prior study compared the efficacy of different therapeutic schools, we should now begin to look at single techniques to learn which techniques have the best versus the worst risk–benefit ratios. For example, in the forerunner study [17], new symptoms due to exposure and response prevention were affirmed by 42.2% of CBT patients who had undergone exposure. Such a fine-grained approach may help to develop "psychotherapeutic package inserts". Moreover, different types of treatment format (e.g., group, individual, online, hybrid/guided online intervention) should be tested for their profile of AE.

Compliance with ethical standards

Conflict of Interest None.

Appendix

See Table 4.

Table 4 Results from the factor analysis. Loadings, percentage endorsement (true, rather true), and allocation to factors

	n	% (rather) yes	MP	UC	SE	PE	Allocation
Subscale Positive Effects (PE; five items, items with bolded positive numbers belong to the scale); Cronbach's alpha: 72)							
1. Through this therapy, I was able to create new hope	132	83.4	**−0.550**	−0.218	−0.113	0.436	/
2. After therapy, I gained confidence in my own abilities	131	80.2	**−0.510**	−0.051	−0.057	**0.559**	PE
3. Through therapy, I have learned to take responsibility for myself	131	77.1	−0.171	0.040	0.020	**0.656**	PE
4. My performance has improved through therapy	131	71.8	−0.498	−0.197	−0.105	**0.522**	PE
5. My family is proud of me because I was in therapy	103	33.0	0.026	0.182	−0.205	0.327	/
6. I am proud that I tried therapy	130	70.0	0.006	0.182	−0.146	**0.592**	PE
7. The relationship with my partner has improved as a result of my therapy	107	51.4	−0.262	−0.171	−0.034	0.071	/
8. The therapist prepared me well for the time after the end of the therapy	132	62.9	**−0.572**	−0.087	−0.021	0.403	/
9. The therapy helped me to disengage from stressful relationships	116	50.9	−0.191	−0.126	0.185	**0.538**	PE
10. The therapy helped me to get a job	97	44.3	−0.319	0.115	0.040	0.356	/
Subscale Unethical Conduct (UE; six items, items with bolded numbers belong to the scale; Cronbach's alpha: 83)							
11. The therapist was intolerant of my sexual orientation	125	4.0	0.129	**0.708**	0.032	0.107	UC
12. The therapist tried to separate me from my partner	126	3.2	0.137	**0.764**	0.081	0.054	UC
13. My therapist insinuated that I was faking my symptoms	134	2.9	0.413	**0.567**	0.042	−0.170	UC
14. The therapist was intolerant of my religion	124	2.4	0.123	0.073	−0.244	−0.022	/
15. During therapy, I experienced sexual abuse by my therapist	134	0	−0.012	**0.888**	0.034	−0.063	UC
16. During therapy, I had to endure sexual harassment by my therapist	135	1.4	0.090	**0.668**	−0.061	0.009	UC
17. The therapist was prejudiced against foreigners	126	0	0.176	**0.712**	0.098	−0.037	UC
Subscale Malpractice (MP; 11 items, items with bolded numbers belong to the scale; Cronbach's alpha: 92)							
18. I experienced verbal abuse/mockery/impatience from the therapist	135	6.6	**0.623**	−0.004	0.091	0.021	MP
19. I had the feeling that the therapist was absent-minded during therapy and was not focused on me	134	9.7	**0.738**	0.291	0.029	−0.062	MP
20. I had the feeling that the therapist did not like me	135	4.5	**0.618**	0.291	−0.014	0.037	MP
21. The therapy was only aimed at "eliminating problems"; there was no positive goal-oriented work	132	17.4	**0.735**	0.165	0.011	−0.243	MP
22. I believe that the therapist's analysis of the cause of my problems was wrong	130	10.7	**0.768**	0.147	0.286	−0.006	MP
23. In my opinion, the applied therapeutic techniques were wrong	130	16.9	**0.712**	0.139	−0.083	−0.202	MP
24. The therapist had no understanding of my problems	133	8.3	**0.814**	0.054	0.109	−0.106	MP
25. In therapy, I learned behaviors that were very bad for me (e.g., saying "no" excessively, limiting myself)	133	4.5	**0.623**	0.396	0.305	−0.035	MP
26. In my opinion, I received a false diagnosis during therapy	131	4.6	**0.646**	−0.007	0.100	−0.111	MP
27. The therapist did not take me seriously	134	4.5	**0.814**	−0.026	−0.004	−0.068	MP
28. The treatment has been a financial strain	135	11.1	0.014	0.189	0.426	0.036	/
29. I feel that I have changed for the negative due to therapy	134	3.7	**0.544**	0.403	0.290	−0.103	MP
30. In the course of therapy, I made decisions that I regret in retrospect	135	6.7	0.335	0.418	0.215	−0.029	/
Subscale Side effects (SE; seven items, items with bolded numbers belong to the scale; Cronbach's alpha: 81)							
31. My application for private insurance was rejected because I was in therapy that was paid for by my health insurance	91	14.3	0.130	0.062	**0.511**	0.075	SE
32. I felt exhausted after the therapeutic sessions	134	51.5	0.446	0.021	0.328	0.029	/
33. After therapy ended, I felt worse because I missed the conversations with my therapist	134	29.1	0.039	0.006	0.394	−0.061	/
34. I am fearful that people I know will learn that I have been in therapy	135	20.0	−0.046	0.120	**0.601**	−0.360	SE
35. My relationships with my family and friends have deteriorated because of my therapy	133	9.8	0.364	−0.053	**0.634**	0.180	SE
36. The therapy has strained my relationship with my partner	117	15.4	0.089	0.400	0.334	0.296	/
37. Because of the exercises my therapist asked me to do, my free time was limited	127	7.1	0.068	0.012	0.498	0.096	/

Table 4 (continued)

	n	% (rather) yes	MP	UC	SE	PE	Allocation
38. It bothers me that my friends treat me differently after my therapy	128	4.7	0.320	−0.038	**0.729**	0.081	SE
39. Since therapy, I feel stigmatized as mentally ill	134	10.4	0.211	0.177	**0.652**	−0.178	SE
40. My family is ashamed of me because I was in therapy	117	16.2	0.068	0.210	**0.574**	−0.212	SE
41. I am ashamed because I was in therapy	135	8.9	−0.245	0.349	**0.521**	−0.293	SE
42. My promotion to tenure in a federal job* was not approved because I was in therapy that was paid for by my health insurance. * in German: *Verbeamtung*	88	1.1	0.099	**0.521**	0.346	0.079	/
43. Due to therapy, my work situation has worsened	126	9.5	0.455	0.219	0.438	−0.132	/

References

1. Cuijpers P, Sijbrandij M, Koole SL et al (2013) The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: a meta-analysis of direct comparisons. World Psychiatry 12:137–148. https://doi.org/10.1002/wps.20038
2. Skapinakis P, Caldwell D, Hollingworth W et al (2016) A systematic review of the clinical effectiveness and cost-effectiveness of pharmacological and psychological interventions for the management of obsessive–compulsive disorder in children/adolescents and adults. Health Technol Assess 20:1–392. https://doi.org/10.3310/hta20430
3. Cuijpers P, van Straten A, Schuurmans J et al (2010) Psychotherapy for chronic major depression and dysthymia: a meta-analysis. Clin Psychol Rev 30:51–62. https://doi.org/10.1016/j.cpr.2009.09.003
4. Kliem S, Kröger C, Kosfelder J (2010) Dialectical behavior therapy for borderline personality disorder: a meta-analysis using mixed-effects modeling. J Consult Clin Psychol 78:936–951. https://doi.org/10.1037/a0021015
5. Hofmann SG, Asnaani A, Vonk IJJ et al (2012) The efficacy of cognitive behavioral therapy: a review of meta-analyses. Cognitive Ther Res 36:427–440. https://doi.org/10.1007/s10608-012-9476-1
6. Mehl S, Werner D, Lincoln TM (2015) Does cognitive behavior therapy for psychosis (CBTp) show a sustainable effect on delusions? A meta-analysis. Front Psychol 6:1450. https://doi.org/10.3389/fpsyg.2015.01450
7. McMahon K, Herr NR, Zerubavel N et al (2016) Psychotherapeutic treatment of bipolar depression. Psychiatr Clin N Am 39:35–56. https://doi.org/10.1016/j.psc.2015.09.005
8. Karyotaki E, Smit Y, de Beurs DP et al (2016) The long-term efficacy of acute-phase psychotherapy for depression: a meta-analysis of randomized trials. Depression Anxiety 33:370–383. https://doi.org/10.1002/da.22491
9. Nieuwsma JA, Trivedi RB, McDuffie J et al (2012) Brief psychotherapy for depression: a systematic review and meta-analysis. Int J Psychiatry Med 43:129–151. https://doi.org/10.2190/PM.43.2.c
10. Boutron I, Moher D, Altman DG et al (2008) Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann Intern Med 148:295–309. https://doi.org/10.7326/0003-4819-148-4-200802190-00008
11. Leichsenring F, Abbass A, Hilsenroth MJ et al (2017) Biases in research: risk factors for non-replicability in psychotherapy and pharmacotherapy research. Psychol Med 47:1000–1011. https://doi.org/10.1017/S003329171600324X
12. Barlow DH (2010) Negative effects from psychological treatments: a perspective. Am Psychol 65:13–20. https://doi.org/10.1037/a0015643
13. Berk M, Parker G (2009) The elephant on the couch: side-effects of psychotherapy. Aust N Z J Psychiatry 43:787–794. https://doi.org/10.1080/00048670903107559
14. Linden M, Schermuly-Haupt ML (2014) Definition, assessment and rate of psychotherapy side effects. World Psychiatry 13:306–309. https://doi.org/10.1002/wps.20153
15. Linden M (2013) How to define, find and classify side effects in psychotherapy: from unwanted events to adverse treatment reactions. Clin Psychol Psychother 20:286–296. https://doi.org/10.1002/cpp.1765
16. Ladwig I, Rief W, Nestoriuc Y (2014) Welche Risiken und Nebenwirkungen hat Psychotherapie? – Entwicklung des Inventars zur Erfassung Negativer Effekte von Psychotherapie (INEP). Verhaltenstherapie. https://doi.org/10.1159/000367928
17. Moritz S, Fieker M, Hottenrott B et al (2015) No pain, no gain? Adverse effects of psychotherapy in obsessive-compulsive disorder and its relationship to treatment gains. J Obsess Compuls Relat Disord 5:61–66. https://doi.org/10.1016/j.jocrd.2015.02.002
18. Rief W, Barsky AJ, Glombiewski JA et al (2011) Assessing general side effects in clinical trials: reference data from the general population. Pharmacoepidemiol Drug Saf 20:405–415. https://doi.org/10.1002/pds.2067
19. Nestoriuc Y, Rief W (2012) Nebenwirkungen von Verhaltenstherapie. In: Linden M, Strauß B (eds) Risiken und Nebenwirkungen von Psychotherapie. MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin, pp 59–74
20. Olatunji BO, Deacon BJ, Abramovitz JS (2009) The cruelest cure? Ethical issues in the implementation of exposure based treatments. Cognit Behav Pract 16:172–180
21. Foa EB, Zoellner LA, Feeny NC et al (2002) Does imaginal exposure exacerbate PTSD symptoms? J Consult Clin Psychol 70:1022–1028. https://doi.org/10.1037/0022-006X.70.4.1022
22. Pope GG (1990) Abuse of psychotherapy: psychotherapist-patient intimacy. Psychother Psychosom 53:191–198. https://doi.org/10.1159/000288366
23. Borruso MT (1991) Sexual abuse by psychotherapists: the call for a uniform criminal statute. Am J Law Med 17:289–311
24. Crawford MJ, Thana L, Farquharson L et al (2016) Patient experience of negative effects of psychological treatment: results of a national survey. Br J Psychiatry 208:260–265. https://doi.org/10.1192/bjp.bp.114.162628
25. Planès S, Villier C, Mallaret M (2016) The nocebo effect of drugs. Pharmacol Res Perspect. https://doi.org/10.1002/prp2.208
26. Wells RE, Kaptchuk TJ (2012) To tell the truth, the whole truth, may do patients harm: the problem of the nocebo effect for informed consent. Am J Bioeth 12:22–29. https://doi.org/10.1080/15265161.2011.652798
27. Swift JK, Derthick AO (2013) Increasing hope by addressing clients' outcome expectations. Psychotherapy 50:284–287. https://doi.org/10.1037/a0031941

28. Gumz A, Kästner D, Geyer M et al (2010) Instability and discontinuous change in the experience of therapeutic interaction: an extended single-case study of psychodynamic therapy processes. Psychother Res J Soc Psychother Res 20:398–412. https://doi.org/10.1080/10503300903551021
29. Klein JP, Berger T, Schröder J et al (2013) The EVIDENT-trial: protocol and rationale of a multicenter randomized controlled trial testing the effectiveness of an online-based psychological intervention. BMC Psychiatry 13:239. https://doi.org/10.1186/1471-244X-13-239
30. Klein JP, Berger T, Schröder J et al (2016) Effects of a psychological internet intervention in the treatment of mild to moderate depressive symptoms: results of the evident study, a randomized controlled trial. Psychother Psychosom 85:218–228. https://doi.org/10.1159/000445355
31. Lecrubier Y, Sheehan DV, Weiller E et al (1997) The mini international neuropsychiatric interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. Eur Psychiatry 12:224–231. https://doi.org/10.1016/S0924-9338(97)83296-8
32. Hamilton M (1960) A rating scale for depression. J Neurol Neurosurg Psychiatry 23:56–62. https://doi.org/10.1136/jnnp.23.1.56
33. Kroenke K, Spitzer RL, Williams JBW (2001) The PHQ-9. J Gen Intern Med 16:605–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x
34. Gilbody S, Richards D, Brealey S, Hewitt C (2007) Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. J Gen Intern Med 22:1596–1602. https://doi.org/10.1007/s11606-007-0333-y
35. Albert PR (2015) Why is depression more prevalent in women? J Psychiatry Neurosci 40:219–221. https://doi.org/10.1503/jpn.150205
36. Nübling R, Schmidt J, Munz D (2010) Psychologische Psychotherapeuten in Baden-Württemberg – Prognose der Versorgung. Psychotherap J 9:46–52
37. Rheker J, Beisel S, Kräling S, Rief W (2017) Rate and predictors of negative effects of psychotherapy in psychiatric and psychosomatic inpatients. Psychiatry Res 254:143–150. https://doi.org/10.1016/j.psychres.2017.04.042
38. Hatfield D, McCullough L, Frantz SHB, Krieger K (2010) Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. Clin Psychol Psychother 17:25–32. https://doi.org/10.1002/cpp.656
39. Moritz S, Berna F, Jaeger S et al (2017) The customer is always right? Subjective target symptoms and treatment preferences in patients with psychosis. Eur Archiv Psychiatry Clin Neurosci 267:335–339. https://doi.org/10.1007/s00406-016-0694-5