Clinical Psychology & Psychotherapy

WILEY

RESEARCH ARTICLE

A Scale for Measuring Positive and Negative Experiences of Psychotherapy (PNEP): First Psychometric Findings of a New Instrument for Monitoring Clients' Experiences

Brechje Dandachi-FitzGerald | Sanne T. L. Houben | Nick J. Broers | Harald Merckelbach

Universiteit Maastricht Faculty of Psychology and Neuroscience, Maastricht, Netherlands

Correspondence: Brechje Dandachi-FitzGerald (b.fitzgerald@maastrichtuniversity.nl)

Received: 7 December 2023 | **Revised:** 3 May 2024 | **Accepted:** 13 May 2024

Funding: This study is supported by Stichting Wetenschapsbevordering Klinisch psychologen en Klinisch neuropsychologen (20-001WF).

Keywords: adverse events | negative experiences | positive experiences | psychotherapy outcome | side effects

ABSTRACT

Background: Clients' adverse experiences during psychotherapy are rarely monitored in clinical practice or research trials. One obstacle here is the lack of a measure to gauge both positive and negative experiences during psychotherapy. We developed and evaluated a new instrument for measuring such experiences.

Method: The Positive and Negative Experiences of Psychotherapy (PNEP) questionnaire was developed based on pilot data, a literature review, and two existing scales for measuring primarily adverse experiences during psychotherapy. Mental healthcare clients ($N = 200$) anonymously completed and evaluated the PNEP. Subsequently, a sample of professionals ($N = 34$) who underwent psychotherapy in the context of their training filled in the PNEP twice, with a 2-week interval in between.

Results: The positive and negative experiences subscales of the PNEP were found to possess excellent internal consistencies ($\alpha s \geq 0.90$). The PNEP test–retest reliability was 0.93 for the positive experiences subscale and 0.78 for the negative experiences subscale. For the positive subscale, four factors were extracted: symptom reduction and positive well-being, high quality of therapy and therapeutic relation, personal growth and acceptance and interpersonal functioning. For the negative subscale, exploratory factor analysis suggested a three-factor solution: escalation of symptoms and emotional distress, low quality of therapy and therapeutic relation and (self-)stigmatization and dependency. Participants related a median of 13 positive and six negative therapy experiences to their most recent treatment. The most frequently endorsed negative experiences were having more negative thoughts and memories, feeling emotionally overwhelmed and an increase in stress due to the therapy. A minority of participants (10.5%) reported no negative treatment experiences.

Conclusions: In the field of psychotherapy, the evaluation of risks and benefits is crucial for assessing safety and effectiveness. The PNEP could be a promising instrument for achieving this objective, although further research is needed to replicate and expand upon the current findings.

1 | Introduction

Psychological therapies are effective in reducing, for example, severe depression and anxiety symptoms (e.g. Wakefield et al. 2021). Still, some patients deteriorate rather than improve during or after psychotherapy (Barlow 2010; Moos 2005). For example, close to 5% of almost 15000 patients who had recently been in psychotherapy reported lasting bad effects of their treatment (Crawford et al. 2016). Negative reactions to or during psychotherapy may involve not only symptom

Summary

- Clients may report negative treatment experiences alongside positive ones, even when they perceive the treatment as beneficial.
- The most commonly reported negative treatment experiences are increasingly experiencing negative thoughts and memories, feeling overwhelmed by emotions, experiencing heightened stress and/or tensions and feeling vulnerable or unprotected.
- To comprehensively monitor treatment progress and evaluate treatment outcomes, it is essential to measure both positive and negative treatment experiences.
- Preliminary psychometric results indicate that the Positive and Negative Experiences of Psychotherapy (PNEP) scale exhibits properties that make it a potentially suitable tool for assessing and appraising both favourable and unfavourable treatment encounters.
- Future studies should further examine the psychometric properties of the PNEP and explore the role of negative experiences in improving overall treatment outcomes in research and clinical settings.

exacerbation (e.g. suicidal ideation after a therapy session, increased anxiety during exposure), relapse and drop-out but also destabilization of social life (e.g. strains in family relations; Dandachi-FitzGerald, Otgaar, and Merckelbach 2023). To develop clinical practice guidelines for psychotherapy, it is imperative that professionals have access to accurate information regarding the balance between benefits and potential drawbacks associated with these therapeutic approaches (Halfond, Wright, and Bufka 2021). However, there is a paucity of data on the prevalence and intensity of negative experiences in psychotherapy trials. The field struggles with 'the identification and detection of harms associated with psychotherapy', and this challenge may well be linked to the absence of a classification scheme for negative experiences during psychotherapy. In an attempt to address this, Linden (2013) proposed a taxonomy. Briefly, it starts with unwanted events, defined as 'all events that occur parallel to treatment in relation to the patient' (Linden 2013, 287). Subsequent steps involve determining whether this unwanted event is (1) caused by the therapy or other factors (e.g. increased mood problems due to being fired at the job because of reorganization while in psychotherapy) and if so (2) whether it is a side effect (e.g. heightened anxiety during exposure therapy) or an iatrogenic effect (e.g. false memories from suggestive trauma recovery techniques). This taxonomy assumes that causality can be established. However, relying solely on patient reports makes it challenging, if not impossible, to distinguish between the different categories of unwanted events. With this in mind, we prefer using the term 'experience' over 'effect', acknowledging that patients' perceptions of positive and negative experiences, along with their attributions, can offer valuable insights.

Self-report instruments measuring negative experiences during psychotherapy do not usually belong to the standard test battery in psychological therapy trials (Cuijpers 2021). For instance, in 2010, only 21% of the published trials ($N = 132$) included any monitoring of harms reported by patients (Jonsson et al. 2014). Although there has been improvement, with a recent systematic review noting that 60% of published preregistered psychotherapy trials explicitly reported harmful events, considerable heterogeneity in conceptualization, monitoring and reporting persists, impeding the accurate assessment of risks and benefits in psychotherapy interventions (Klatte et al. 2023). Not surprisingly, the Lancet Psychiatry Commission (Holmes et al. 2018, 257) concluded, 'Historically, psychological therapy trials have been poor at both monitoring hypothesized side-effects and deterioration, and reporting serious adverse events'. Relatedly, instruments assessing negative therapy effects are rarely included in routine outcome measurements (Kendrick et al. 2016).

Various patient-rated instruments exist for assessing experiences with psychotherapy. Examples include the Experiences with Therapy Questionnaire (ETQ; Parker et al. 2013), the Inventory of Negative Effects of Psychotherapy (INEP; Ladwig, Rief, and Nestoriuc 2014) and the Side Effects of Psychotherapy Scale (SEPS; Moritz et al. 2015). The two English-language patient-rated instruments with the most well-documented psychometric data are the Negative Effects Questionnaire (NEQ; Rozental et al. 2016) and the Positive and Negative Effects of Psychotherapy Scale (PANEPS; Moritz et al. 2019; Peth et al. 2018), which will be further considered here. The 32 items of the NEQ were selected from a larger pool of 60 items generated by a consensus meeting of experts (Rozental et al. 2014) and through qualitative assessments of patients' negative treatment experiences during online psychotherapy trials (Boettcher et al. 2014; Rozental et al. 2015) as well as a literature review. Given the background and item content of the NEQ, there is no a priori reason to presume that this scale would be unsuitable for assessing negative experiences in face-to-face psychotherapy. Each NEQ item refers to a negative experience and respondents are asked whether it occurred to them (yes/no). If answered affirmatively, respondents are asked to rate the impact of the negative experience on a 5-point Likert scale (0 = *not at all*, 1 = *slightly*, 2 = *moderately*, 3 = *very*, 4 = *extremely*) and to relate the negative experience to a cause (i.e. the therapy per se or other circumstances). Thus, negative therapy experiences are conceptualized as adverse experiences that the client or patient ascribes to the therapy. Rozental et al. (2015) conducted an exploratory factor analysis (EFA) performed on data collected in a mixed clinical sample ($N = 653$) and found a six-factor solution (with factors labelled *Symptoms*, *Quality*, *Dependency*, *Stigma*, *Hopelessness* and *Failure*) accounting for 57.6% of the variance. The authors also found the full scale of the NEQ to possess an excellent internal consistency (Cronbach's $\alpha = 0.95$), with internal consistencies for the six factors ranging from acceptable to excellent (Cronbach's α range = 0.72–0.93).

In contrast to the NEQ, the PANEPS (Moritz et al. 2019) covers both negative and positive experiences, though the latter is relatively under-represented, compromising less than 25% of items. The 43 statements of the PANEPS are rated on a 4-point scale (1 = *true*, 2 = *rather true*, 3 = *rather not true*, 4 = *not true*). The PANEPS, a shortened and revised version of the Side Effects of Psychotherapy Scale (SEPS; Moritz et al. 2015), was developed based on discussions with clinicians and patients suffering from obsessive-compulsive disorder (OCD). Additional sources of input were two already existing German assessments tools [i.e. Inventory of Negative Effects of Psychotherapy (INEP;

Ladwig, Rief, and Nestoriuc 2014) and the Unwanted Event-Adverse Treatment Reaction (UE-ATR) checklist for therapists (Linden 2013)]. The psychometric properties of the PANEPS were examined in a sample of 135 patients with a current or previous depressive episode (Moritz et al. 2019). A principal component analysis identified a four-factor solution, labelled *Positive effects*, *Side effects*, *Unethical conduct* and *Malpractice*, accounting for 46.3% of the variance. The internal consistencies of these factors ranged between acceptable and excellent reliability (i.e. Cronbach's $\alpha = 0.72–0.92$).

Several recent studies have employed scales such as the NEQ, PANEPS or similar tools (i.e. the INEP or SEPS) to systematically explore negative therapy experiences. A recurrent finding is that a substantial proportion of patients (i.e. 22%–93%), report having encountered at least one adverse aspect of treatment (e.g. Gerke et al. 2020; Holsting et al. 2017; Moritz et al. 2019; Rheker et al. 2017; Rozental et al. 2019; Strauss et al. 2021). The reported negative experiences include feeling down after therapy, dependency on the therapist or therapy, fear of disclosure, post-session exhaustion, heightened stress and anxiety and resurfacing negative memories. The prevalence of such experiences varies widely from study to study, probably due to different sample characteristics and clinical settings. Even at the lower end of this range, it is evident that there is a need for scales that thoroughly address adverse experiences in psychotherapy trials and clinical practice. This is essential for assessing treatment safety and tolerability.

Notwithstanding the good psychometric qualities of the NEQ and PANEPS, their content primarily alludes to harmful effects. This may induce negative priming, even up to the point of nocebo effects, by creating expectations about the occurrence of aversive side effects (e.g. Faasse and Petrie 2013; Herzog et al. 2019). What the field needs, then, is an instrument that includes items that address both positive and negative experiences. With this in mind, the current research aims to contribute to the development of an instrument enabling researchers and clinicians to measure patients' experiences during psychotherapy in a balanced and unbiased way (see also Paveltchuk et al. 2022). Drawing inspiration from NEQ and PANEPS, and after thorough review of pertinent literature, we developed and evaluated a novel instrument aimed at comprehensively gauging positive and negative experiences associated with psychotherapy.

2 | Method

We conducted three studies to develop and evaluate the Positive and Negative Experiences of Psychotherapy (PNEP) questionnaire. The first study focused on developing the PNEP based on panel evaluation ($N = 17$), a literature review and two existing scales for adverse experiences during psychotherapy.

In the second study, mental healthcare clients ($N = 200$) anonymously completed and evaluated the PNEP. This study aimed to investigate order effects by presenting two versions of the PNEP randomly: (1) positive experiences before negative experiences and (2) negative experiences before positive experiences. This study also examined the internal consistency and factor structure of the two subscales and gathered user feedback on aspects such as item clarity, PNEP length and informational value. Additionally, we looked at the total positive and negative treatment experiences and explored the relationship of two therapy characteristics with these total scores: ongoing or completed therapy and treatment duration.

Finally, the third study aimed to assess test–retest reliability with a 2-week interval in professionals ($N = 34$) who completed 50 psychotherapy sessions during training.

3 | Study 1 (Pilot): Panel Evaluation

Participants were recruited from the (Dutch) Association of Experts by Experience and the professional network of the researchers. Experts by experience refers to peer providers in mental health facilities (e.g. Farkas and Boevink 2018). In the Netherlands, an expert by experience is someone who has personal experience with psychopathology, often severe mental health conditions, and has completed an official education programme lasting 1 or 2 years to become a certified expert by experience (Van Bakel et al. 2013). In total, 20 experts by experience and (former) mental health patients were invited to evaluate NEQ and PANEPS items in terms of qualities such as comprehensiveness of topics and readability (Berghs 2020). We used the Dutch version of the NEQ. As for the PANEPS, its English version was translated to Dutch using back-and-forth translation. The data of three participants were excluded because they failed to respond ($n = 1$) or dropped out of the study due to medical reasons ($n = 2$). The final sample consisted of 17 participants (nine women). Participants had a mean age of 36.9 years ($SD = 14.3$) and a mean experience with psychotherapy of 13.6 years ($SD = 12.7$), and a majority (65%) had a relatively high educational level. The group was diverse regarding the psychopathology for which they had received psychotherapy, with trauma-related disorders, anxiety and mood disorders and personality disorders reported most often. The most frequently mentioned forms of psychotherapy were talking therapy (not further specified), cognitive behavioural therapy, eye movement desensitization and reprocessing and schema therapy. Participants completed in a counterbalanced order the NEQ and the PANEPS, followed by an evaluation questionnaire containing both quantitative and qualitative items. User-friendliness was rated on a Likert scale from 1 (*very poor*) to 10 (*excellent*). Language difficulty was evaluated on a 5-point Likert scale (1 = *very difficult to understand*; 5 = *very easy to understand*). Follow-up open-end questions asked participants to elaborate on the strengths and weaknesses of both questionnaires and comprehensiveness of topics. In addition, participants were queried on the relevance of instruments measuring negative effects. Upon returning the completed questionnaires, a structured telephone interview took place to allow for clarification (e.g. in cases where handwriting was difficult to read) and to ensure that any additional feedback points were captured. However, the telephone interview did not yield new information beyond what participants had already provided in the questionnaires. The data were stored in a privacy folder within the research data management programme. Participants received a gift voucher of 22.50 euros for their participation. The pilot study was approved by the standing ethical committee of the Faculty of Psychology and Neuroscience of Maastricht University [Master_207_10_04_2019].

4 | Results: Study 1

Both questionnaires were deemed to be user-friendly [$M_{NEQ} = 7.65$ $SD_{NEQ} = 1.37$ and $M_{PANEPS} = 7.29$, $SD_{PANEPS} = 1.72$], and their items were easy to understand [$M_{NEQ} = 4.29$, $SD_{NEQ} = 0.69$ and $M_{PANEPS} = 4.12$, $SD_{PANEPS} = 0.70$]. For the NEQ, the most frequently reported strength was the answer format with follow-up questions ($n = 6$; 35.3%). Also, the option to elaborate in an open-end question about other negative effects was mentioned as a strength by five participants (29.4%). For the PANEPS, eight participants (47.1%) liked its clear questions and layout, and five participants (29.4%) appreciated the broad scope of topics it covered, including boundary violations. Also, four participants (23.5%) judged the positive items to be a strength of the PANEPS, although they also indicated that positive and negative items could be better balanced. Four participants (23.5%) were critical about the answer format of the PANEPS. Specifically, they found it challenging to distinguish between 'not true' and 'not applicable' categories, felt the categories were too coarse (e.g. desiring an additional category with *a bit* true or untrue) and missed an option to explain an answer or to list additional experiences.

In a broad sense, the advantages of the NEQ were mentioned as disadvantages of the PANEPS and vice versa. Crucially, participants perceived both questionnaires as not representative of their personal experiences with psychotherapy. Although negative experiences were said to occur on a regular basis, the panel stressed that positive experiences during psychotherapy should also be queried to obtain a more balanced view.

4.1 | Item Design

Based on these pilot data and the core domains of negative experiences identified in Herzog et al. (2019), a new questionnaire was developed: the PNEP. The main impetus for developing a new instrument was twofold: (1) We wanted an instrument that includes all core domains as identified in Herzog et al. (2019); (2) we wanted an instrument that assesses both negative and positive experiences in a balanced manner and allows for querying the patient's perspective regarding attribution of causality. Permission to adapt and use parts from both the NEQ and the PANEPS for developing the PNEP was generously granted by their respective authors (S. Moritz and J. Peth, personal communication, April 12, 2019; A. Rozental, personal communication, April 24, 2020). In this process, careful consideration was given to (1) presenting items covering positive and negative psychotherapy experiences; (2) including an open-end question for both positive and negative experiences; (3) leveraging NEQ and PANEPS items for addressing negative experiences; and (4) adopting an answer format that allows to differentiate between occurrence and causal attribution of a (negative) experience.

Positive experiences items were drawn from the PANEPS and a literature review focusing on studies examining patients' perceptions of positive, beneficial or helpful aspects of therapy and positive outcomes of psychotherapy (e.g. Binder, Holgersen, and Nielsen 2010; Fava and Guidi 2020; Hoyer et al. 2020; Timulak and Keogh 2017). All in all, the positive experiences subscale of the PNEP comprised 33 items, encompassing seven domains: (1) positive well-being; (2) acceptance (self—by others); (c) progress (personal growth, self-insight/knowledge); (4) high quality of therapy; (5) high quality of the therapeutic relationship; (6) autonomy (resilience/coping skills); and (7) positive developments/improvement in daily life (work, family, relationships).

The negative experiences items were mainly drawn and adapted from the NEQ and the PANEPS. We included a few additional items to tap into experiencing new symptoms (e.g. 'New, unpleasant memories surfaced that were unknown to me before I started therapy'), feelings that are connected to being in therapy (i.e. feeling emotionally overwhelmed and feeling vulnerable or unprotected) and therapy expectations (e.g. 'the therapy took longer than I expected or was told'). In total, the negative experiences subscale included 36 items that span eight domains: (1) symptoms (new/worsening of); (2) stigma (self—by others); (3) lack of treatment response/demoralization/loss of hope; (4) low quality of therapy; (5) low quality of the therapeutic relationship; (6) dependency on the therapist; (7) negative developments and stress in daily life (work, family, relationships); and (8) therapist's transgressive behaviour/misconduct.

Each PNEP item first provides respondents with the description of an experience (e.g. 'I felt more stress'), of which they have to indicate whether they experienced this. Upon answering affirmatively, respondents are prompted to answer three follow-up questions. The first pertains to the personal impact of the experience (3-point Likert scale: *a bit, quite a lot, a lot*). The second addresses the duration of the experience (4-point Likert scale: *shortly*, *multiple days/weeks*, *multiple months* or *permanent*). The third question is about the probable cause of the experience (therapy vs. other circumstances). A positive or negative experience is considered a therapy-related experience, at least according to respondents, when (1) they affirmatively answer the item and (2) attribute the experience to therapy.[1]

Both the positive and negative subscales conclude with an 'other' item providing room for explaining the experience in an open text box, followed by open questions as described below. Also, at the end of the PNEP, respondents are asked to rate to what extent they benefitted from their most recent therapy using a Visual Analogue Scale that ranges from −100 (*severely deteriorated*) to +100 (*strongly improved*). The PNEP items are listed in Data S1.

5 | Study 2: Psychometric and Qualitative Evaluation PNEP

5.1 | Data Collection

Recruitment took place through advertisements on Psychosenet.nl, an online platform for mental health problems (e.g. psychotic spectrum disorders, mood disorders and trauma-related psychopathology), and the Dutch Association of Experts by Experience website. To further increase the diversity of the sample, the advertisement was also sent to JADOS, a supported living organization for persons with autism spectrum disorder, and posted on social media (e.g. LinkedIn). The study was advertised as an opportunity for participants to evaluate a new measure for assessing positive and negative treatment experiences.

The advertisement contained a link to the online platform Qualtrics. Participants first read information about the survey's aim and how long it would take to complete it (approximately 20–30 min). The introductory text also explained that there was no compensation for participation, that participants had the right to withdraw from the survey at any time and that data collection was anonymous. Participants were not informed about the random allocation to two versions of the PNEP (see below). After providing consent, participants answered demographical questions that addressed age, gender and educational level. Next, questions about the diagnoses and psychotherapy (history) were given. Following this, the PNEP items were presented. Participants randomly received one of the two versions, that is, first negative followed by positive items or vice versa. Following this, they were presented with the PNEP evaluation questionnaire, designed to evaluate the clarity and length of the PNEP, identify potentially omitted topics and gather suggestions for further improving the questionnaire (see Table 1 for the questionnaire items and answer categories). Upon completion of the study, participants were fully debriefed on the randomization of PNEP versions. Ethical approval was obtained from the standing Ethical Review Committee of the Faculty of Psychology and Neuroscience of Maastricht University [ERCPN-220_43_03_2020_A1].

5.2 | Participants

The link to the study was opened 371 times. Incomplete questionnaires were excluded ($n = 171$), of which 86 records (50.3% of 171) were of participants who did answer the demographical items. Non-completers did not differ from completers with regard to age[2] [$\chi^2(4) = 7.17$, $p = 0.13$], gender[3] [$\chi^2(1) = 0.01$, $p = 0.93$] and educational level [$\chi^2(3) = 3.75$, $p = 0.29$]. The final sample consisted of 200 participants. The majority were women, relatively highly educated and between 25 and 54 years old. As can be seen in Table 2, the sample was heterogeneous in terms of psychopathology, type of psychotherapy method and treatment duration. According to the participants, a median of two diagnoses ($IQR = 1$–3) were the targets of the most recent psychotherapy. Approximately half of the participants reported having undergone one type of treatment, whereas the other half said to have received two or more psychotherapy methods.

5.3 | Data Analysis

Data were analysed with SPSS version 26. Using a 2 (order: negative–positive vs. positive–negative) × 2 (valence: negative vs. positive items) ANOVA with repeated measures on the last factor, we examined whether the order of administration had an effect on the number of reported negative and positive therapy experiences. Next, descriptive statistics were used to summarize the demographics of the sample, frequency of negative and positive therapy experiences and questions about the usefulness of the questionnaire. For exploratory purposes, we conducted independent samples *t*-tests to assess the influence of two psychotherapy characteristics—specifically, whether the therapy was ongoing or concluded and whether the therapy duration was less than 1 year or 1 year or longer—on the mean total scores of positive and negative treatment experiences.

TABLE 1 | Evaluation of the Positive and Negative Experiences of Psychotherapy (PNEP) questionnaire (Study 2, $N = 200$).

Evaluative statements	*n*	%
The items are clear		
Fully agree	43	21.5
Agree	76	38.0
Neutral	22	11.0
Somewhat disagree	54	27.0
Fully disagree	5	2.5
The length of the questionnaire is doable		
Fully agree	52	26.0
Agree	82	41.0
Neutral	25	12.5
Somewhat disagree	36	18.0
Fully disagree	5	2.5
I have sometimes answered 'no' to a question, just to avoid getting follow-up questions		
No	139	69.5
Yes, sometimes	54	27.0
Yes, regularly (> 5 times)	7	3.5
I missed topics in the questionnaire		
Yes, namely [open text]	45	22.5
No	92	46.0
I do not know	63	31.5
This questionnaire could be informative/important to monitor progress during therapy		
Fully agree	51	25.5
Agree	72	36.0
Neutral	53	26.5
Somewhat disagree	12	6.0
Fully disagree	12	6.0
This questionnaire could be informative/important to evaluate the therapy outcome		
Fully agree	50	25.0
Agree	80	40.0
Neutral	52	26.0
Somewhat disagree	12	6.0
Fully disagree	6	3.0

Using Pearson product–moment correlations, we examined relationships between key variables. The factor structure of the positive and negative subscales of the PNEP was examined with EFA using principal axis factoring. Assuming factors to be correlated, we used oblique rotations (direct oblimin) with delta set at zero and the number of iterations set at 25. Bartlett's test of sphericity was

TABLE 2 | Sample characteristics ($N = 200$).

	n	%
Gender		
Women	152	76.0
Men	42	21.0
Other	6	3.0
Age category (years)		
18–24	32	16.0
25–34	48	24.0
35–44	36	18.0
45–54	49	24.5
55–64	24	12.0
65–74	11	5.5
Education		
Low	5	2.5
Medium	63	31.5
High	132	66.0
Diagnosis that was focus of psychotherapy[a]		
Trauma/posttraumatic stress disorder	65	32.5
Mood disorders	61	30.5
Personality disorders	44	22.0
Autism spectrum disorders	34	17.0
Anxiety disorders	28	14.0
Psychotic spectrum disorders	20	10.0
Eating disorders	9	4.5
ADHD	6	3.0
Other	17	8.5
Treatment method during psychotherapy[a]		
Talking therapy (not otherwise specified)	108	54.0
Cognitive behavioural therapy	68	34.0
EMDR	56	28.0
Schema therapy	31	15.5
Group therapy	24	12.0
Relationship/family therapy	9	4.5
Emotion-focused therapy	15	7.5
Mentalization-based therapy	6	3.0
Interpersonal therapy	6	3.0
Other[b]	37	18.5
I do not know	4	2.0
End of psychotherapy		
Still ongoing	90	45.0
This year (2020)	18	9.0
Last year	31	15.5
Between 2 and 5 years ago	33	16.5
> 5 years ago	27	13.5
Psychotherapy duration		
< 2 months	12	6.0
Between 2 and 6 months	34	17.0
Between 6 and 12 months	36	18.0
Between 1 and 2 years	48	24.0
Between 2 and 5 years	50	25.0
> 5 years	20	10.0

[a]Multiple answers possible;
[b]Most open answer responses in the other category ($n = 18$) alluded to a (more) specific treatment form (e.g. exposure, DGT, ACT, narrative exposure therapy). In total, 23 of the 37 participants (62.2%) who endorsed the other category, also ticked one of the predetermined categories. Fourteen participants (7% of the total sample) exclusively chose the other option.

performed to evaluate the suitability of the EFA and the correlation between items. The Kaiser–Meyer–Olkin (KMO) test was performed to measure the sampling adequacy of the data. A number of methods were used to ensure the validity of the factor solution. The factor solution was based on the Kaiser criterion (i.e. eigenvalues > 1), the scree test and potential substantive considerations. For a sample of 200 participants, the scree plot is a relatively reliable source for factor reduction, and so we deemed it appropriate to use it. Cronbach's α was calculated to determine the internal consistencies of the final factors (Cortina 1993).

6 | Results: Study 2

6.1 | Effect of Order

In total, 104 participants received first the negative items, followed by the positive items. For 96 participants this order was reversed. The main effect of valence was significant, with participants generally reporting more positive than negative experiences [$F(1,198) = 37.43$, $p < 0.001$, partial $\eta 2 < 0.16$]. The main effect of order remained statistically non-significant: $F(1,198) = 0.48$, $p = 0.49$. Importantly, there was no significant interaction effect between order of administration and total number of reported negative and positive experiences [$F(1,198) = 0.01$, $p = 0.91$]. Therefore, data of both order versions of the PNEP were collapsed for subsequent analyses.

6.2 | Descriptive Statistics

Cronbach's α was 0.94 for positive items and 0.90 for negative items (see also below). A median of 13 positive therapy experiences

(*IQR* = 5–20) and six negative therapy experiences (*IQR* = 3–11) were reported. As for the positive therapy experiences, feeling accepted by the therapist, feeling that the therapy was conducted in a good way and feeling understood and supported by the therapist were most often mentioned (see Table 3). Therapy experiences related to the respondent's development were also regularly reported, including better coping with problems and increased self-acceptance (49.5%; *n* = 99). In total, 13 respondents (6.5%) did not report any positive therapy experiences.

As for negative therapy experiences, distressing thoughts and memories, feeling emotionally overwhelmed and an increase in stress were most often mentioned (see Table 3). Developing suicidal thoughts was reported by 27 (13.5%) respondents. Six respondents (3.0%) indicated that they had experienced unacceptable behaviour in the form of verbal abuse (e.g. shouting, name-calling), coercion or mockery. Two respondents (1.0%) reported sexually transgressive behaviour. Twenty-one respondents (10.5%) did not report any negative therapy experiences. The negative and positive therapy experiences varied in intensity and duration. A detailed overview of the response frequencies can be found in Data S1.

The mean total scores of positive and negative treatment experiences did not differ between the group for which psychotherapy was still ongoing ($n = 90$; $M_{pnep\_positive} = 13.73$, $SD = 7.70$; $M_{pnep\_negative} = 7.72$, $SD = 6.86$) and the group for which psychotherapy was concluded ($n = 110$; $M_{pnep\_positive} = 12.92$, $SD = 9.86$, $M_{pnep\_negative} = 7.57$, $SD = 6.46$); $t(198) = 0.64$, $p = 0.523$ and $t(198) = 1.58$, $p = 0.874$, respectively. However, there was a difference with a modest effect size in the mean total positive treatment experiences between the group that received therapy for less than 1 year ($n = 82$, $M_{pnep\_positive} = 11.39$, $SD = 8.37$) and those who received therapy for 1 year or longer ($n = 118$; $M_{pnep\_positive} = 14.60$, $SD = 9.11$); $t(198) = -2.53$, $p = 0.012$, $d = 0.367$. The difference in mean total negative treatment experiences between those who received psychotherapy for less than 1 year ($M_{pnep\_negative} = 6.63$, $SD = 5.56$) and those that received psychotherapy for 1 year or longer ($M_{pnep\_negative} = 8.34$, $SD = 7.21$) fell short of significance; $t(198) = -1.80$, $p = 0.073$, $d = 0.264$.

6.3 | Correlations Between Therapy Experiences and Therapy Evaluation

There was an inverse relationship between number of positive and negative experiences ($r_s = -0.37$, $p < 0.001$). The median rating on the item 'Overall, how well has psychotherapy helped you?' (VAS from −100 to + 100) was 55.5 (*IQR* = 8.5–83.3). Reporting more negative therapy experiences was

TABLE 3 | Most frequent positive and negative psychotherapy experiences across Study 2 and Study 3 samples.

	PNEP item	Study 2[a] (*N* = 200)			Study 3[b] (*N* = 34)		
		n	**%**	**95% CI**	***n***	**%**	**95% CI**
Positive psychotherapy experiences							
24.	I felt that the therapist accepted me	137	68.5	[61.6, 74.9]	32	94.1	[80.3, 99.3]
18.	In my opinion the therapy was well executed	136	68.0	[61.1, 74.4]	32	94.1	[80.3, 99.3]
23.	The therapist understood and supported me	134	67.0	[60.0. 73.5]	33	97.1	[84.7, 99.9]
22.	I felt a match with my therapist	133	66.5	[59.5, 73.0]	32	94.1	[80.3, 99.3]
20.	The therapist had informed me well about the therapy	127	63.5	[56.4, 70.2]	26	76.5	[62.3, 90.8]
15.	I understood myself better	123	61.5	[54.4, 68.3]	32	94.1	[80.3, 99.3]
8.	I learnt to accept myself better	98	49.0	[42.1, 56.0]	32	94.1	[80.3, 99.3]
Negative psychotherapy experiences							
3.	I suffered more from negative thoughts and memories	110	55.0	[47.8, 62.0]	10	29.4	[15.1, 47.5]
6.	I was overwhelmed by emotions	99	49.5	[42.4, 56.6]	12	35.3	[19.8, 53.5]
1.	I suffered more from stress and/or tensions	94	47.0	[39.9, 54.2]	5	14.7	[5.0, 31.1]
7.	I felt vulnerable or unprotected	75	37.5	[30.8, 44.6]	14	41.2	[24.7, 59.3]
9.	The symptoms for which I had sought treatment increased	65	32.5	[26.1, 39.5]	2	5.9	[−2.0, 13.8]
10.	I started to suffer from other, new symptoms that I did not had before the therapy	24	12.0	[7.5, 16.5]	5	14.7	[5.0, 31.1]
31.	The relationship with my family came under pressure	25	12.5	[7.9, 17.1]	5	14.7	[5.0, 31.1]

Abbreviations: CI = confidence interval, PNEP = Positive and Negative Experiences of Psychotherapy.
[a]Data from 200 (former) patients and experiential experts, who filled in the PNEP on an online platform.
[b]Data from T1 in a test–retest set-up in which 34 professionals filled in the PNEP for their mandatory personal therapy.

associated with a less favourable assessment of therapy outcome ($r_s = -0.44$, $p < 0.001$), whereas reporting more positive therapy experiences was associated with a higher score on this outcome measure ($r_s = 0.77$, $p < 0.001$). The subgroup of participants who clearly indicated improvement in therapy (i.e. VAS > VAS_{median} 55.5; $n = 99$) reported a median of five negative therapy effects ($IQR = 2–7$).

6.4 | Principal Axis Factoring

To obtain an acceptable case-to-variables ratio, separate EFA were performed for the positive (33 items) and negative (36 items) experiences subscales.

For the positive subscale, the Kaiser–Meyer–Olkin was 0.92, and the Bartlett's test of sphericity was significant, indicating that the data were appropriate for EFA. Seven factors met Kaiser's criterion and in combination explained 60.53% of the variance. The scree plot showed inflexions at a two-factor and a four-factor solution. Based on Kaiser's criterion and on interpretability of the factors, we decided to extract a four-factor solution including 31 items (see Table 4). The first three extracted factors met Kaiser's criterion (eigenvalues Factor 1 = 11.09, accounting for 33.6% of the variance; Factor 2 = 1.38 explaining an additional 4.2% of variance; and Factor 3 = 1.14, describing 3.4% of variance). The fourth factor had an eigenvalue of 0.94 and explained 2.9% of the variance. After rotation, Factor 1 relates to the experience of symptom reduction and positive well-being, Factor 2 covers the therapy and therapeutic relationship, and Factor 3 can best be conceptualized as indexing personal growth and acceptance. Finally, Factor 4 relates mostly to interpersonal functioning (with family and friends and in school or the workplace). The rotated solution yielded factors with good reliabilities for the first three factors (Cronbach's αs > 0.80). However, with a Cronbach's α of 0.60, the reliability of the fourth factor was questionable.

For the negative items, the Kaiser–Meyer–Olkin was 0.84, and the Bartlett's test of sphericity was significant, again suggesting that the data are suitable for EFA. Ten factors had eigenvalues over Kaiser's criterion and together explained 63.07% of the variance. Looking at the scree plot, which showed an inflexion at three factors explaining 37.8% of the variance, and the interpretability of the factors, we extracted a three-factor solution that encompassed 29 items (see Table 5). All three factors met Kaiser's criterion (eigenvalues Factor 1 = 8.48, accounting for 23.5% of the variance; Factor 2 = 3.01, explaining an additional 8.6% of variance; and Factor 3 = 2.04, describing 5.7% of variance). After rotation, Factor 1 relates to experiencing more symptoms and emotional distress; Factor 2 gauged the quality of the therapy and therapeutic relationship; and Factor 3 related to (self-) stigmatizing and dependency. The rotated solution yielded factors with good reliabilities for the first two factors (Cronbach's αs > 0.80) and with a Cronbach's α of 0.72 an acceptable reliability of the third factor.

6.5 | PNEP Evaluation

The majority of participants indicated that administration of the PNEP might yield important information during ($n = 123$, 61.5%) and at the end ($n = 130$, 65.0%) of psychotherapy (see Table 1). Although most participants ($n = 119$, 59.5%) agreed with the statement that the PNEP items were clear, a considerable minority ($n = 59$, 29.5%) somewhat or fully disagreed. Responses to open-end evaluation questions indicated participants found it challenging to attribute an experience to a cause. The most frequently mentioned comment was that the dichotomous choice between therapy and other circumstances was artificial and that there should also be an answer option referring to a situation in which both contributed to the experience. Participants sometimes said that the questionnaire was too extensive, particularly the three follow-up questions on severity, duration and probable cause. Consistent with this, 27% of the sample ($n = 54$) mentioned occasionally responding 'no' to an item to avoid getting the follow-up questions. Seven participants (3.5%) reported they had been doing this on a regular basis (> 5 times).

7 | Study 3: Test–Retest Reliability

Based on the recommendations of Study 2 participants, we slightly modified the PNEP. Most importantly, we added a third option (i.e. 'partly the therapy/partly other circumstances') to the items that asked for an attribution of experiences during psychotherapy. Further, we reduced the length by substituting the two follow-up questions on impact and duration with one question on the intensity of the experience (5-point Likert scale: *not at all*, *slightly*, *moderately*, *very*, *extremely*). Lastly, we adjusted the answer format of the final item, evaluating overall therapy, from a Visual Analogue Scale to a 7-point Likert scale (1 = *severely deteriorated*; 7 = *strongly improved*) for enhanced interpretability and measurement reliability. The nine items (seven negative and two positive items) that did not load on any of the factors in the EFA (Study 2) were retained in the PNEP, given the response frequencies (e.g. 27% of the participants reported that new, unknown bad memories surfaced) or the important content of the item (e.g. items on verbally and sexually transgressive therapist behaviour). The modified version has been translated into English using back-and-forth translation and can be found in Data S2.

In a small-scale study, we assessed the test–retest reliability of the modified PNEP. Participants were recruited from either the final year or within the last 5 years of completing their postgraduate training programme in clinical psychology/psychotherapy at the training institute RINO Zuid, Eindhoven, The Netherlands. Only participants were included who had completed the mandatory personal therapy of minimal 50 sessions. This approach facilitated a targeted evaluation of the reliability of the PNEP in individuals undergoing a similar professional treatment with consistent duration while also contributing to the overall goal of obtaining a heterogeneous sample across three separate studies in line with the broad purpose design of the PNEP questionnaire. Participants were asked to complete the PNEP twice with a 2-week interval while reflecting on their experience with personal therapy. At T1, a short questionnaire with items pertaining to demographic and professional background was administered, followed by the PNEP. An item to check for inattentive responding was included in each subscale of the PNEP at T1 and T2 (i.e. 'If I read this item carefully, I will answer "yes" here' and 'my therapist is a human being'). The

TABLE 4 | Principal axis factoring for a four factor solution for the positive psychotherapy experiences using oblique rotation.

	PNEP item	Factor loading			
		1	2	3	4
Factor 1: Symptom reduction—positive well-being					
3.	I felt more comfortable in my own skin	**0.81**	0.06	−0.25	0.02
1.	I felt better	**0.61**	0.25	−0.13	−0.06
6.	I felt happy more often	**0.61**	−0.06	−0.07	0.28
2.	The symptoms for which I came to therapy have decreased	**0.56**	0.17	−0.25	0.02
4.	I could enjoy things more	**0.56**	−0.02	−0.18	0.17
14.	I got new hope	**0.45**	−0.05	−0.32	0.11
5.	I often times felt more calm and relaxed	**0.39**	0.02	−0.18	0.17
20.	The therapist has informed me well about the therapy	0.35	0.24	0.09	0.03
Factor 2: High quality therapy—therapeutic relation					
23.	The therapist understood and supported me	0.03	**0.80**	0.00	0.07
24.	I felt that the therapist accepted me	−0.04	**0.77**	−0.02	−0.03
22.	I felt a match with my therapist	0.02	**0.62**	−0.03	0.04
19.	I had a lot of trust in the therapy method	0.06	**0.41**	−0.32	0.11
18.	In my opinion the therapy was well executed	0.34	0.35	−0.11	0.06
Factor 3: Personal growth and acceptance					
29.	I have learned to take more responsibility for myself	−0.27	0.09	**−0.75**	0.09
13.	I learned to live life more from what I consider to be valuable and important	0.16	−0.05	**−0.59**	−0.03
28.	I gained more control over my life	0.16	0.04	**−0.53**	0.10
27.	I took better care of myself	0.16	−0.05	**−0.52**	0.17
9.	I gained more confidence in my own abilities	0.10	0.11	**−0.52**	0.14
16.	I could further develop myself as a person	0.10	0.12	**−0.50**	0.04
8.	I learned to accept myself more	0.27	0.08	**−0.49**	−0.08
25.	I learned to deal with problems better	0.21	0.19	**−0.47**	0.04
15.	I understood myself better	0.16	0.20	**−0.44**	0.06
26.	I learned where to go when I need help	0.03	0.24	**−0.40**	0.14
30.	I was able to perform my daily activities better (e.g. work, study, daytime activities)	0.13	0.04	−0.32	0.25
Factor 4: Interpersonal functioning					
12.	People close to me (family, partner, friends) were proud of me because I went to therapy	0.10	0.10	0.29	**0.64**
32.	Relationships with people close to me (e.g. partner, friends) improved	0.01	−0.03	−0.11	**0.56**
31.	The relationship with my family improved	0.01	−0.03	−0.01	**0.55**
10.	New, pleasant memories surfaced that were unknown to me before I started therapy	0.01	0.02	−0.03	**0.41**
33.	At work/study/daytime activities they were aware of my therapy/diagnosis, and this had a positive effect on me	−0.10	0.08	−0.12	**0.38**
11.	I was proud that I tried therapy	0.05	0.14	−0.12	0.35

(Continues)

TABLE 4 | (Continued)

	PNEP item	Factor loading			
		1	2	3	4
17.	I started to plan the future	0.19	−0.09	−0.28	0.30
Items not included in the factor solution					
7.	After sleep, I became fitter and more well rested	0.20	−0.07	−0.13	0.28
21.	The therapist has prepared me well for the period after ending therapy	0.25	0.06	−0.09	0.27

Note: $N = 200$. Factor loadings above 0.36 are in bold.
Abbreviation: PNEP = Positive and Negative Experiences of Psychotherapy.

study was approved by the standing Ethical Review Committee of the Faculty of Psychology and Neuroscience of Maastricht University [ERCPN-243_134_10_2021].

7.1 | Participants

The invitation to participate was emailed to 172 eligible participants. Out of these, 16 automatic replies were received, either because the e-mail address was no longer active ($n = 13$) or because the participant was on temporary leave (e.g. maternal leave; $n = 3$), leaving a sample of 156 eligible participants. Of these, 46 participants initiated the survey and 40 participants completed all items, yielding a response rate of 25.6%. After 2 weeks, 34 of the 40 participants (85.0%) completed the PNEP for the second time. None of the participants incorrectly answered the check items for inattentive responding. Below, we will focus on the subsample of 34 participants with complete records.

Participants were, on average, 41.5 years ($SD = 6.0$; range = 31–58). The vast majority were women (97.1%). Out of the participants, 13 (38.2%) were licensed psychotherapist, 14 (41.2%) were licensed clinical psychologist, and seven (20.6%) were healthcare psychologist in training to become clinical psychologist. Eleven participants (32.4%) had completed the personal therapy with one therapist, whereas 23 participants had two separate trajectories (each minimal 25 sessions). This latter subgroup was instructed to focus on one specific trajectory and, with that in mind, fill out the PNEP twice. Personal therapy was completed in the past year for five participants (14.7%), the past 1–2 years for nine participants (26.5%), the past 3 years for eleven participants (32.4%), and 4 years or longer ago for nine participants.

8 | Results: Study 3

The median time to fill in the modified PNEP the first time was 15.3 min ($IQR = 11.4–19.5$). Cronbach's αs for positive and negative experiences at T1 and T2 were all > 0.75. On average, participants reported a median of 20.5 ($IQR = 16.0–27.0$) positive experiences and 2.5 ($IQR = 1.0–5.3$) negative experiences at T1. At T2, medians for positive and negative experiences were 20.5 ($IQR = 16.0–27.0$) and 1.0 ($IQR = 0.0–4.0$), respectively. The test–retest stability (Pearson product–moment correlation) between positive experiences at T1and T2 was 0.93. For negative experiences, it was $r = 0.78$. For positive therapy experiences, the most frequently endorsed items were well in line with that found in Study 2 (see above and Table 3). Overall, the professional sample reported negative therapy experiences less frequently than the patient sample. However, the four most frequently endorsed negative therapy experiences were similar across both samples.

9 | Discussion

The study's key findings can be summarized as follows. First, although participants of the pilot study appreciated existing scales (i.e. NEQ; Rozental et al. 2016; PANEPS; Moritz et al. 2019; Peth et al. 2018), many argued that these scales were biased towards negative experiences. Together with the systematic review of Herzog et al. (2019) on instruments measuring negative therapy experiences, this provided an important impetus to develop the PNEP, which addresses both positive and negative experiences.

Second, the positive and negative subscales of the PNEP demonstrated adequate internal reliabilities, with Cronbach's αs > 0.75 in studies 2 and 3. Third, the test–retest reliability of the (slightly modified) PNEP (Study 3) was good, although higher for positive than for negative experiences, which is an interesting observation in itself (see below). Fourth, the EFA (Study 2) extracted factors aligned with a priori theoretical assumptions and replicated empirical findings of psychotherapy effects and outcomes (e.g. Deres et al. 2020; Flückiger et al. 2018; Norcross and Lambert 2018), lending credibility to their content.

All in all, the PNEP appears to be a psychometrically sound instrument that can be used to evaluate psychotherapy effects. Interestingly, the pattern of frequently endorsed positive and negative experiences was rather similar across former patients and professionals who had undergone training therapy and aligned well with the literature on therapy effects. That is, positive experiences such as feeling accepted, understood and supported by the therapist and experiencing that the therapy has been conducted well all relate to therapeutic alliance (i.e. clarity and consensus on tasks and goals in therapy and the client-therapist bond; Bordin 1994). Several studies have documented the impact of therapeutic alliance on psychotherapy outcome (e.g. Flückiger et al. 2018; Moos 2005; Norcross and Lambert 2018). Apart from symptom reduction, participants in Study 2 and Study 3 often endorsed aspects such as a better understanding of oneself, increased self-acceptance, personal growth and being better able to deal with problems. This fits well with qualitative studies on what patients' value in psychotherapy (Binder, Holgersen, and Nielsen 2010; Hoyer

TABLE 5 | Principal axis factoring for a three factor solution for the negative psychotherapy experiences using oblique rotation.

	PNEP item	Factor loading		
		1	2	3
Factor 1: Symptom escalation—emotional distress				
1.	I suffered more from stress and/or tensions	**0.72**	−0.11	−0.15
3.	I suffered more from negative thoughts and memories	**0.64**	0.03	−0.01
9.	The symptoms for which I had sought treatment, increased	**0.62**	−0.12	−0.07
6.	I was overwhelmed by emotions	**0.58**	−0.19	0.05
2.	I had more trouble sleeping	**0.51**	0.10	−0.03
17.	I felt hopeless	**0.51**	−0.20	0.08
5.	I was gloomier	**0.50**	−0.10	−0.03
8.	I got thoughts that it would be better if I did not exist anymore or that I should take my own life	**0.44**	−0.07	0.21
4.	I felt more anxious	**0.42**	−0.16	0.14
7.	I felt vulnerable or unprotected	**0.39**	−0.24	0.07
33.	I was unable to perform my daily activities	0.36	−0.09	0.20
Factor 2: Low-quality therapy—therapeutic relationship				
26.	I did not feel understood by the therapist	0.03	**−0.79**	−0.05
24.	I felt that the therapist had insufficient knowledge	0.07	**−0.75**	−0.08
28.	I did not feel a connection between me and the therapist	−0.02	**−0.75**	−0.12
27.	I felt a distance in the relation with the therapist	0.07	**−0.74**	−0.10
25.	I felt that the therapist was not taking me seriously	−0.07	**−0.67**	−0.003
20.	I lost trust in therapy	0.12	**−0.66**	−0.09
23.	I did not quite understand the purpose of the therapy	−0.03	**−0.53**	0.25
16.	I did not notice any improvement	0.08	**−0.52**	0.07
21.	The therapy was only focused on 'taking away the problem'; there was no positive goal that we were trying to achieve	−0.02	**−0.52**	0.19
18.	I lost trust in myself; my sense of self-worth decreased	0.30	**−0.48**	0.13
Factor 3: (Self-)stigmatization—dependency				
14.	I had the feeling that people thought I am crazy	−0.05	−0.10	**0.68**
30.	I did not dare to make a decision without first consulting my therapist	−0.01	0.08	**0.61**
15.	I am scared that people will find out I am (have been) in psychotherapy	0.06	0.03	**0.52**
22.	In my opinion I received an incorrect diagnosis	−0.01	−0.30	**0.45**
13.	People that are close to me (family, partner, friends) were embarrassed because I am (have been) in psychotherapy	−0.01	0.10	**0.45**
12.	I was ashamed that I am (have been) in psychotherapy	0.06	−0.01	**0.41**
29.	I felt too dependent on my therapist	0.11	−0.07	0.33
31.	The relationship with my family came under pressure	0.29	−0.06	0.31
Items not included in the factor solution				

(Continues)

TABLE 5 | (Continued)

	PNEP item	Factor loading		
		1	2	3
10.	I started to suffer from other, new symptoms that I did not have before the therapy	0.23	−0.13	0.14
11.	New, unpleasant memories surfaced that were unknown to me before I started therapy	0.18	0.12	0.20
19.	The therapy took longer than I had anticipated or was told it would	0.20	−0.04	0.23
32.	My relationship with my partner deteriorated	0.23	−0.10	0.15
34.	At work/study/daytime activities they were aware of my therapy/diagnosis and this had a negative effect on me	0.12	−0.17	0.25
35.	I experienced verbal abuse, for example yelling, name calling, mockery/compulsion from the therapist	0.09	−0.25	0.07
36.	In therapy I experienced sexual violence, for example inappropriate sexual comments, touches or sexual abuse by my therapist	−0.11	−0.21	0.25

Note: $N = 200$. Factor loadings above 0.36 are in bold.
Abbreviation: PNEP = Positive and Negative Experiences of Psychotherapy.

et al. 2020; Timulak and Keogh 2017). Considering the episodic and/or long-lasting nature of psychological problems (e.g. Gustavson et al. 2018), it might be equally important to define the success of therapy by factors such as learning to cope, accepting and valuing oneself and living in accordance with personal values, alongside the reduction in symptom levels at the end of psychotherapy.

As for negative experiences, having more negative thoughts and memories, feeling emotionally overwhelmed, feeling vulnerable and unprotected and an increase in stress due to therapy were the most frequently endorsed items across the two samples. The vast majority of participants in Study 2 ($n = 179$, 89.5%) indicated having experienced at least one negative therapy effect. This percentage is close to the upper bound of the 22%–93% range found across different studies (e.g. Gerke et al. 2020; Holsting et al. 2017; Moritz et al. 2019; Rheker et al. 2017; Rozental et al. 2019; Strauss et al. 2021). Note that this wide range across studies is likely to be the result of heterogeneity in patient samples (e.g. patients with OCD, depressive disorders, specific phobia, receiving inpatient vs. outpatient psychotherapy), differences in interventions and differences in assessment methods. The high frequency of negative experiences that we found could be due to the relatively extensive list of PNEP items that address such experiences. Additionally, our recruitment strategy may have contributed to the heightened percentage of negative experiences in that we targeted platforms such as Psychosenet.nl and the website of the Dutch Association of Experiential Experts. In doing so, we likely recruited individuals with severe mental health conditions, as evidenced by participants holding a median of two diagnoses and over half receiving treatment for more than 1 year.

Although the impact of negative therapy experiences on treatment outcome remains understudied, our findings in Study 2 reveal that even participants who considered their psychotherapy successful reported on average five negative therapy experiences. This suggests that certain negative therapy experiences may be unavoidable and perhaps even necessary for treatment progress (e.g. increased anxiety during exposure therapy). Be that as it may, we also found that reporting more negative therapy experiences was associated with fewer positive therapy experiences and, notably, a less favourable evaluation of the therapy result. Interestingly, the negative experiences subscale of the PNEP appeared to be less stable across time than the positive experiences subscale. It may well be the case that post-therapy, some patients develop a new interpretation of their problems along with a new reference point from which to assess psychotherapy (Anvari et al. 2022; Valkonen, Hänninen, and Lindfors 2011). Clearly, this issue warrants further research. In line with our findings, Verkooyen, Broers, and Dandachi-Fitzgerald (2024) utilized the PNEP during mid-treatment and found that, whereas all patients ($N = 80$) reported positive therapy experiences, 69% of them ($n = 55$) also reported negative experiences. The authors conducted a regression analysis and found that positive therapy experiences predicted treatment outcomes (i.e. OQ-45 score end therapy) after correction for baseline severity (i.e. OQ-45 at baseline) and demographics, whereas negative therapy experiences did not. However, consistent with our findings, reporting more negative therapy experiences was linked to fewer positive therapy experiences at mid- and end of treatment and overall lower patient evaluation of therapy success. Monitoring negative experiences during therapy is important, as shown in a randomized controlled trial of Muschalla et al. (2023). These authors demonstrated that systematically monitoring and discussing potential side effects during therapy sessions had a positive impact on the therapeutic alliance when compared to the control group. This improvement in the therapeutic alliance may, in turn, positively influence treatment outcomes. In line with this, most of our participants felt that administration of the PNEP during and/or at the end of

the therapy had added value. Thus, systematically evaluating both positive and negative experiences might improve the patient–therapist collaboration in therapy.

Importantly, our participants also provided valuable points for improvement of the questionnaire. Based on this feedback, we slightly amended the PNEP. We encourage researchers and clinicians to use and further refine the PNEP.

Our study is not without limitations. First and foremost, the PNEP and similar instruments are self-reports. We do not know how accurate people are in determining positive and negative experiences and relating them to causes. The social-psychological literature (Ross and Nisbett 2011) would suggest that in general people display all kinds of biases when probed for this type of evaluations. Therefore, a much-needed complementary strategy would be to survey important others (e.g. family members) as well as therapists along with clients themselves.

Relatedly and second, our approach is not well able to distinguish between those negative experiences that occur because of the therapy and negative experiences that occur during or after therapy. For example, when a patient experiences more bad thoughts and memories while in psychotherapy, this might be not only due to therapy factors (e.g. talking about the past in therapy) but also due to extra-therapy factors (e.g. a family gathering that stirs up bad memories). Although often overlooked, this caveat also holds for positive experiences such as symptom reduction (De Smet et al. 2021; Lilienfeld et al. 2014).

A third limitation is that we do not know how representative our patient sample in Study 2 is. Although the sample was quite diverse in terms of DSM-5 classifications, type and duration of psychotherapy, people with a low educational level, men and the elderly were underrepresented (Ten Have et al. 2022). Moreover, because participants were tasked with assessing aspects such as comprehensiveness and readability, there is a possibility that the relatively high educational level of our sample contributed to an overestimation of these questionnaire features. Meanwhile, we have no reason to believe that it was mainly patients with either a markedly negative or positive experience who responded to the invitation to participate in our study (e.g. only a small minority reported to have experienced no positive or negative therapeutic experiences at all). Also, our results do not suggest that age, gender or education impacted the completion rate among those who started with the survey. Generalizability might also be a problem for Study 3. Here, only about a quarter of eligible participants completed the study. We do not know how representative their experiences with personal therapy are of the 75% of eligible participants who did not participate in our study. However, the main focus of Study 3 was to examine the test–retest reliability, and for this purpose, our sample provided a good starting point.

Fourth, the Study 2 sample size was adequate, albeit relatively modest for an EFA (McCallum et al. 1999; Rouquette and Falissard 2011). We encourage other researchers to use the PNEP in larger samples and carry out more fine-grained psychometrics in order to consider whether items need to be dropped so that stable subscales emerge that lend themselves to straightforward interpretations. In addition, future studies with larger samples might examine the underlying dimensions more thoroughly, and particularly, structural equation modelling could help to clarify the dynamics of negative and positive experiences and how they impact therapy outcome. Ultimately, this would contribute to a better construct definition of negative therapy experiences, something that is clearly needed (Paveltchuk et al. 2022).

The PNEP might be a suitable instrument for measuring negative psychotherapy reactions in clinical practice. Systematically monitoring therapy progress in terms of symptoms and daily functioning with instruments, such as the OQ-45 (Lambert et al. 2004) or the Outcome Rating Scale and Session Rating Scale (Campbell and Hemsley 2009), enhances therapy success (de Jong et al. 2021). Despite being lengthier, the PNEP's comprehensive assessment of both positive and negative aspects may offer valuable insights for therapists regarding therapy safety and tolerability. Perhaps most importantly, unlike the other instruments, the PNEP invites clients to attribute positive and negative experiences and this information might provide a good starting point for a helpful dialogue between client and therapist. This thorough evaluation is especially important for addressing negative experiences or harmful effects of psychotherapy, which therapists often overlook or attribute to the patient's psychopathology (Hannan et al. 2005; Hatfield et al. 2010; Teachman, White, and Lilienfeld 2021; Werbart, Annevall, and Hillblom 2019). Monitoring adverse events in psychotherapy may help overcome this blind spot for therapists and provide professionals with feedback on when to intervene, thereby reducing the risk of therapy failure. More generally, we need research to test whether monitoring both beneficial and harmful effects of psychotherapy has a protective effect in itself (Muschalla et al. 2023; Paveltchuk et al. 2022). In any case, such monitoring with, for example, the PNEP would increase the informational value of randomized control trials (Ellett and Chadwick 2021; Guidi et al. 2018; Ioannidis et al. 2004), which is essential for the development of clinical practice guidelines (e.g. Halfond, Wright, and Bufka 2021).

10 | Conclusions

Psychotherapy should be evaluated for risks and benefits to determine safety and success (Halfond, Wright, and Bufka 2021). The initial findings reported here indicate that the PNEP could be a suitable tool to achieve this goal, although replication and extension of the current findings are needed. Knowledge about negative experiences with psychotherapy is crucial for several reasons. For example, to determine the treatment burden of therapies (Dobler et al. 2018; Heinig et al. 2022), as well as to identify patients who are at risk of (serious) adverse events, this information, along with the success rate of the therapy, is needed to enable patients to give informed consent (Blease, Lilienfeld, and Kelley 2016; Halfond, Wright, and Bufka 2021). In addition, insight into the negative effects of therapies can inspire improvement of current treatments or the development of alternatives.

Endnotes

[1]To illustrate, in the sample to be described below ($N = 200$), 111 participants (62.5% of the total sample) endorsed the item 'I learned to accept myself more', with 98 of them (49%) attributing this to therapy. As another example, 90 participants (45%) indicated they developed suicidal thoughts, with 24 of them (13.5%) attributing this to therapy.

[2]The analysis involved five age categories (18–24 years; 25–34 years; 35–44 years; 45–54 years; 55–64 years). The 65–74 years' group was excluded because of the small number of participants in this category ($n = 11$).

[3]The analysis involved men versus women, because there were only six participants who indicated 'other'.

References

Anvari, F., E. Efendić, J. Olsen, R. C. Arslan, M. Elson, and I. K. Schneider. 2022. "Bias in Self-Reports: An Initial Elevation Phenomenon." *Social Psychological and Personality Science* 14: 727–737. https://doi.org/10.1177/19485506221129160.

Barlow, D. H. 2010. "Negative Effects From Psychological Treatments." *American Psychologist* 65, no. 1: 13–20. https://doi.org/10.1037/a0015643.

Berghs, M. 2020. "Evaluation of Two Questionnaires to Measure Negative Effects of Psychotherapy." Unpublished Master's thesis, Maastricht University.

Binder, P. E., H. Holgersen, and G. H. S. Nielsen. 2010. "What Is a "Good Outcome" in Psychotherapy? A Qualitative Exploration of Former Patients' Point of View." *Psychotherapy Research* 20, no. 3: 285–294. https://doi.org/10.1080/10503300903376338.

Blease, C. R., S. O. Lilienfeld, and J. M. Kelley. 2016. "Evidence-Based Practice and Psychological Treatments: The Imperatives of Informed Consent." *Frontiers in Psychology* 7: 1170. https://doi.org/10.3389/fpsyg.2016.01170.

Boettcher, J., A. Rozental, G. Andersson, and P. Carlbring. 2014. "Side Effects in Internet-Based Interventions for Social Anxiety Disorder." *Internet Interventions* 1, no. 1: 3–11. https://doi.org/10.1016/j.invent.2014.02.002.

Bordin, E. S. 1994. "Theory and Research on the Therapeutic Working Alliance: New Directions." In *The Working Alliance: Theory, Research, and Practice*, edited by A. O. Horvath and L. S. Greenberg, 13–37. John Wiley & Sons.

Campbell, A., and S. Hemsley. 2009. "Outcome Rating Scale and Session Rating Scale in Psychological Practice: Clinical Utility of Ultra-Brief Measures." *Clinical Psychologist* 13, no. 1: 1–9. https://doi.org/10.1080/13284200802676391.

Cortina, J. M. 1993. "What Is Coefficient Alpha? An Examination of Theory and Applications." *Journal of Applied Psychology* 78, no. 1: 98–104. https://doi.org/10.1037/0021-9010.78.1.98.

Crawford, M. J., L. Thana, L. Farquharson, et al. 2016. "Patient Experience of Negative Effects of Psychological Treatment: Results of a National Survey." *The British Journal of Psychiatry* 208, no. 3: 260–265. https://doi.org/10.1192/bjp.bp.114.162628.

Cuijpers, P. 2021. "Research on Negative Effects of Psychotherapies: The Next Steps." *Clinical Psychology: Science and Practice* 28, no. 2: 142–144. https://doi.org/10.1037/cps0000011.

Dandachi-FitzGerald, B., H. Otgaar, and H. Merckelbach. 2023. "When Psychotherapy Fails." In *Toward a Science of Clinical Psychology: A Tribute to the Life and Works of Scott O. Lilienfeld*, 301–319. Springer International Publishing.

de Jong, K., J. M. Conijn, R. A. Gallagher, A. S. Reshetnikova, M. Heij, and M. C. Lutz. 2021. "Using Progress Feedback to Improve Outcomes and Reduce Drop-Out, Treatment Duration, and Deterioration: A Multilevel Meta-Analysis." *Clinical Psychology Review* 85: 102002. https://doi.org/10.1016/j.cpr.2021.102002.

De Smet, M. M., C. Von Below, E. Acke, A. Werbart, R. Meganck, and M. Desmet. 2021. "When 'Good Outcome' Does Not Correspond to 'Good Therapy': Reflections on Discrepancies Between Outcome Scores and Patients' Therapy Satisfaction." *European Journal of Psychotherapy & Counselling* 23, no. 2: 156–176. https://doi.org/10.1080/13642537.2021.1923049.

Deres, A. T., P.-C. Bürkner, B. Klauke, and U. Buhlmann. 2020. "The Role of Stigma During the Course of Inpatient Psychotherapeutic Treatment in a German Sample." *Clinical Psychology & Psychotherapy* 27, no. 2: 239–248. https://doi.org/10.1002/cpp.2423.

Dobler, C. C., N. Harb, C. A. Maguire, C. L. Amour, C. Coleman, and M. H. Murad. 2018. "Treatment Burden Should Be Included in Clinical Practice Guidelines." *British Medical Journal* 363: k4065. https://doi.org/10.1136/bmj.k4065.

Ellett, L., and P. Chadwick. 2021. "Recommendations for Monitoring and Reporting Harm in Mindfulness for Psychosis Research." *The British Journal of Psychiatry* 219, no. 6: 629–631. https://doi.org/10.1192/bjp.2021.98.

Faasse, K., and K. J. Petrie. 2013. "The Nocebo Effect: Patient Expectations and Medication Side Effects." *Postgraduate Medical Journal* 89, no. 1055: 540–546. https://doi.org/10.1136/postgradmedj-2012-131730.

Farkas, M., and W. Boevink. 2018. "Peer Delivered Services in Mental Health Care in 2018: Infancy or Adolescence?" *World Psychiatry* 17, no. 2: 222–224. https://doi.org/10.1002/wps.20530.

Fava, G. A., and J. Guidi. 2020. "The Pursuit of Euthymia." *World Psychiatry* 19, no. 1: 40–50. https://doi.org/10.1002/wps.20698.

Flückiger, C., A. C. Del Re, B. E. Wampold, and A. O. Horvath. 2018. "The Alliance in Adult Psychotherapy: A Meta-Analytic Synthesis." *Psychotherapy* 55, no. 4: 316–340. https://doi.org/10.1037/pst0000172.

Gerke, L., A. Meyrose, I. Ladwig, W. Rief, and Y. Nestoriuc. 2020. "Frequencies and Predictors of Negative Effects in Routine Inpatient and Outpatient Psychotherapy: Two Observational Studies." *Frontiers in Psychology* 11: 2144. https://doi.org/10.3389/fpsyg.2020.02144.

Guidi, J., E.-L. Brakemeir, C. L. H. Bockting, et al. 2018. "Methodological Recommendations for Trials of Psychological Interventions." *Psychotherapy and Psychosomatics* 87, no. 5: 276–284. https://doi.org/10.1159/000490574.

Gustavson, K., A. K. Knudsen, R. Nesvåg, G. P. Knudsen, S. E. Vollset, and T. Reichborn-Kjennerud. 2018. "Prevalence and Stability of Mental Disorders Among Young Adults: Findings From a Longitudinal Study." *BMC Psychiatry* 18: 65. https://doi.org/10.1186/s12888-018-1647-5.

Halfond, R. W., C. V. Wright, and L. F. Bufka. 2021. "The Role of Harms and Burdens in Clinical Practice Guidelines: Lessons Learned From the American Psychological Association's Guideline Development." *Clinical Psychology: Science and Practice* 28, no. 1: 19–28. https://doi.org/10.1111/cpsp.12343.

Hannan, C., M. J. Lambert, C. Harmon, et al. 2005. "A Lab Test and Algorithms for Identifying Clients at Risk for Treatment Failure." *Journal of Clinical Psychology* 61, no. 2: 155–163. https://doi.org/10.1002/jclp.20108.

Hatfield, D., L. McCullough, S. H. Frantz, and K. Krieger. 2010. "Do We Know When Our Clients Get Worse? An Investigation of Therapists' Ability to Detect Negative Client Change." *Clinical Psychology & Psychotherapy* 17: 25–32. https://doi.org/10.1002/cpp.656.

Heinig, I., S. Knappe, J. Hoyer, et al. 2022. "Effective–and Tolerable: Acceptance and Side Effects of Intensified Exposure for Anxiety Disorders." *Behavior Therapy* 54, no. 3: 427–443. https://doi.org/10.1016/j.beth.2022.11.001.

Herzog, P., S. Lauff, W. Rief, and E.-L. Brakemeir. 2019. "Assessing the Unwanted: A Systematic Review of Instruments Used to Assess Negative Effects of Psychotherapy." *Brain and Behavior: A Cognitive Neuroscience Perspective* 9: e01447. https://doi.org/10.1002/brb.1447.

Holmes, E. A., A. Ghaderi, C. J. Harmer, et al. 2018. "The Lancet Psychiatry Commission on Psychological Treatments Research in Tomorrow's Science." *The Lancet Psychiatry* 5, no. 3: 237–286. https://doi.org/10.1016/S2215-0366(17)30513-8.

Holsting, A. F., H. F. Pedersen, M. T. Rask, L. Frostholm, and A. Schröder. 2017. "Is Psychotherapy for Functional Somatic Syndromes Harmful? A Mixed Methods Study on Negative Effects." *Journal of Psychosomatic Research* 98: 113–121. https://doi.org/10.1016/j.jpsychores.2017.05.010.

Hoyer, J., J. Čolić, G. Grübler, and A. T. Gloster. 2020. "Valued Living Before and After CBT." *Journal of Contemporary Psychotherapy* 50: 37–45. https://doi.org/10.1007/s10879-019-09430-x.

Ioannidis, J. P. A., S. J. W. Evans, P. C. Gøtzsche, et al. 2004. "Better Reporting of Harms in Randomized Trials: An Extension of the CONSORT Statement." *Annals of Internal Medicine* 141, no. 10: 781–788. https://doi.org/10.7326/0003-4819-141-10-200411160-00009.

Jonsson, U., I. Alaie, T. Parlin, and F. K. Arnberg. 2014. "Reporting of Harms in Randomized Controlled Trials of Psychological Interventions for Mental and Behavioral Disorders: A Review of Current Practice." *Contemporary Clinical Trials* 38, no. 1: 1–8. https://doi.org/10.1016/j.cct.2014.02.005.

Kendrick, T., M. El-Gohary, B. Stuart, et al. 2016. "Routine Use of Patient Reported Outcome Measures (PROMs) for Improving Treatment of Common Mental Health Disorders in Adults." *Cochrane Database of Systematic Reviews* 2016: CD011119. https://doi.org/10.1002/14651858.CD011119.pub2.

Klatte, R., B. Strauss, C. Flückiger, and J. Rosendahl. 2023. "Adverse Events in Psychotherapy Randomized Controlled Trials: A Systematic Review." *Psychotherapy Research*: 1–16. https://doi.org/10.1080/10503307.2023.2286992.

Ladwig, I., W. Rief, and Y. Nestoriuc. 2014. "Welche Risiken und Nebenwirkungen hat Psychotherapie?-Entwicklung des Inventars zur Erfassung Negativer Effekte von Psychotherapie (INEP)." *Verhaltenstherapie* 24, no. 4: 252–263.

Lambert, M. J., J. J. Morton, D. R. Hatfield, et al. 2004. *Administration and Scoring Manual for the Outcome Questionnaire (OQ 45.2)*. 3rd ed. Wilmington, DE: American Professional Credentialling Services LLC.

Lilienfeld, S. O., R. A. Ritschel, S. J. Lynn, R. L. Cautin, and R. D. Latzman. 2014. "Why Ineffective Psychotherapies Appear to Work: A Taxonomy of Causes of Spurious Therapeutic Effectiveness." *Perspectives on Psychological Science* 9, no. 4: 355–387. https://doi.org/10.1177/1745691614535.

Linden, M. 2013. "How to Define, Find and Classify Side Effects in Psychotherapy: From Unwanted Events to Adverse Treatment Reactions." *Clinical Psychology & Psychotherapy* 20, no. 4: 286–296. https://doi.org/10.1002/cpp.1765.

McCallum, R. C., K. F. Widaman, S. Zhang, and S. Hong. 1999. "Sample Size in Factor Analysis." *Psychological Methods* 4, no. 1: 84–99. https://doi.org/10.1037/1082-989X.4.1.84.

Moos, R. H. 2005. "Iatrogenic Effects of Psychosocial Interventions for Substance Use Disorders: Prevalence, Predictors, Prevention." *Addiction* 100, no. 5: 595–604. https://doi.org/10.1111/j.1360-0443.2005.01073.x.

Moritz, S., M. Fieker, B. Hottenrott, et al. 2015. "No Pain, no Gain? Adverse Effects of Psychotherapy in Obsessive–Compulsive Disorder and its Relationship to Treatment Gains." *Journal of Obsessive-Compulsive and Related Disorders* 5: 61–66. https://doi.org/10.1016/j.jocrd.2015.02.002.

Moritz, S., Y. Nestoriuc, W. Rief, J. P. Klein, L. Jelinek, and J. Peth. 2019. "It Can't Hurt, Right? Adverse Effects of Psychotherapy in Patients With Depression." *European Archives of Psychiatry and Clinical Neuroscience* 269: 577–586. https://doi.org/10.1007/s00406-018-0931-1.

Muschalla, B., J. Müller, A. Grocholewski, and A. Linden. 2023. "Effects of Talking about Side Effects Versus Not Talking About Side Effects on the Therapeutic Alliance: A Controlled Clinical Trial." *Acta Psychiatrica Scandinavica* 148: 208–216. https://doi.org/10.1111/acps.13543.

Norcross, J. C., and M. J. Lambert. 2018. "Psychotherapy Relationships That Work III." *Psychotherapy* 55, no. 4: 303–315. https://doi.org/10.1037/pst0000193.

Parker, G., K. Fletcher, M. Berk, and A. Paterson. 2013. "Development of a Measure Quantifying Adverse Psychotherapeutic Ingredients: The Experiences of Therapy Questionnaire (ETQ)." *Psychiatry Research* 206, no. 2-3: 293–301. https://doi.org/10.1016/j.psychres.2012.11.026.

Paveltchuk, F., S. E. D. Q. Mourão, S. Keffer, R. T. da Costa, A. E. Nardi, and M. R. de Carvalho. 2022. "Negative Effects of Psychotherapies: A Systematic Review." *Counselling and Psychotherapy Research* 22, no. 2: 267–278. https://doi.org/10.1002/capr.12423.

Peth, J., L. Jelinek, Y. Nestoriuc, and S. Moritz. 2018. "Unerwünschte Effekte von Psychotherapie bei depressiven Patienten–Erste Anwendung der Positive and Negative Effects of Psychotherapy Scale (PANEPS)." *Psychotherapie, Psychosomatik, Medizinische Psychologie* 68: 391–398. https://doi.org/10.1055/s-0044-101952.

Rheker, J., S. Beisel, S. Kräling, and W. Rief. 2017. "Rate and Predictors of Negative Effects of Psychotherapy in Psychiatric and Psychosomatic Inpatients." *Psychiatry Research* 254: 143–150. https://doi.org/10.1016/j.psychres.2017.04.042.

Ross, L., and R. E. Nisbett. 2011. *The Person and the Situation: Perspectives of Social Psychology*. Pinter & Martin Publishers.

Rouquette, A., and B. Falissard. 2011. "Sample Size Requirements for the Internal Validation of Psychiatric Scales." *International Journal of Methods in Psychiatric Research* 20, no. 4: 235–249. https://doi.org/10.1002/mpr.352.

Rozental, A., G. Andersson, J. Boettcher, et al. 2014. "Consensus Statement on Defining and Measuring Negative Effects of Internet Interventions." *Internet Interventions* 1, no. 1: 12–19. https://doi.org/10.1016/j.invent.2014.02.001.

Rozental, A., J. Boettcher, G. Andersson, B. Schmidt, and P. Carlbring. 2015. "Negative Effects of Internet Interventions: A Qualitative Content Analysis of Patients' Experiences With Treatments Delivered Online." *Cognitive Behaviour Therapy* 44, no. 3: 223–236. https://doi.org/10.1080/16506073.2015.1008033.

Rozental, A., A. Kottorp, J. Boettcher, G. Andersson, and P. Carlbring. 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: e0157503. https://doi.org/10.1371/journal.pone.0157503.

Rozental, A., A. Kottorp, D. Forsström, et al. 2019. "The Negative Effects Questionnaire: Psychometric Properties of an Instrument for Assessing Negative Effects in Psychological Treatments." *Behavioural and Cognitive Psychotherapy* 47, no. 5: 559–572. https://doi.org/10.1017/S1352465819000018.

Strauss, B., R. Gawlytta, A. Schleu, and D. Frenzl. 2021. "Negative Effects of Psychotherapy: Estimating the Prevalence in a Random National Sample." *BJPsych Open* 7, no. 6: E186. https://doi.org/10.1192/bjo.2021.1025.

Teachman, B. A., B. A. White, and S. O. Lilienfeld. 2021. "Identifying Harmful Therapies: Setting the Research Agenda." *Clinical Psychology: Science and Practice* 28, no. 1: 101–106. https://doi.org/10.1037/cps0000002.

Ten Have, M., M. Tuithof, S. van Dorsselaer, F. Schouten, and R. de Graaf. 2022. "NEMESIS. De psychische gezondheid van de Nederlandse bevolking [NEMESIS. The Psychological Health of the Dutch Population]." Retrieved February 3, 2023. https://cijfers.trimbos.nl/nemesis/nemesis-rapport-home/.

Timulak, L., and D. Keogh. 2017. "The client's Perspective on (Experiences of) Psychotherapy: A Practice Friendly Review." *Journal of Clinical Psychology* 73, no. 11: 1556–1567. https://doi.org/10.1002/jclp.22532.

Valkonen, J., V. Hänninen, and O. Lindfors. 2011. "Outcomes of Psychotherapy From the Perspective of the Users." *Psychotherapy Research* 21, no. 2: 227–240. https://doi.org/10.1080/10503307.2010.548346.

Van Bakel, M., S. Van Rooijen, D. Boertien, J. Kamoschinski, S. Liefhebber, and M. Kluft. 2013. *Beroepscompetentie Ervaringsdeskundigheid. [Professional Competences of Experts-by-Experience]*. Epos Press.

Verkooyen, R., N. Broers, and B. Dandachi-Fitzgerald. 2024. "No Pain, no Gain Revisited: The Impact of Positive and Negative Psychotherapy Experiences on Treatment Outcome." *Frontiers in Psychology* 15. https://doi.org/10.3389/fpsyg.2024.1378456.

Wakefield, S., S. Kellett, M. Simmonds-Buckley, D. Stockton, A. Bradbury, and J. Delgadillo. 2021. "Improving Access to Psychological Therapies (IAPT) in the United Kingdom: A Systematic Review and Meta-Analysis of 10-Years of Practice-Based Evidence." *British Journal of Clinical Psychology* 60, no. 1: 1–37. https://doi.org/10.1111/bjc.12259.

Werbart, A., A. Annevall, and J. Hillblom. 2019. "Successful and Less Successful Psychotherapies Compared: Three Therapists and Their Six Contrasting Cases." *Frontiers in Psychology* 10: 816. https://doi.org/10.3389/fpsyg.2019.00816.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.