Great expectations: inhibitory learning and change processes in exposure therapy for PTSD

Kooistra, M.J.

Chapter 3

Distress variability during exposure therapy and its relationship with PTSD symptom decline

Published as:
Kooistra, M. J., Hoeboer, C. M., Oprel, D. A. C., Schoorl, M., Van Der Does, W., Van Minnen, A., & De Kleine, R. A. (2024). Distress variability during exposure therapy and its relationship with PTSD symptom decline. *Journal of Behavior Therapy and Experimental Psychiatry, 85, 101983.* https://doi.org/10.1016/j.jbtep.2024.101983

Abstract

Background and objectives: Inhibitory Learning Theory (ILT) framework implies that in-session distress variability may promote extinction learning and thereby enhance exposure therapy efficacy. Thus far, research has mainly focused on in-session distress reduction. The aim of the current study was to assess whether in-session distress variability predicts next session PTSD symptom decline in PTSD patients receiving prolonged exposure (PE).
Methods: Eighty-six patients with PTSD received 14 to 16 sessions of PE. Using dynamic panel models, we assessed the temporal relation (i.e., within-persons) between in-session distress variability and PTSD symptom decline. Moreover, we assessed the averaged relation (i.e., between-persons) between in-session distress variability and PTSD symptom decline.
Results: Temporal analyses showed that in-session distress variability did not precede PTSD symptom improvement. Averaged analyses showed that distress variability was related to PTSD symptom improvement.
Limitation: The operationalization of distress variability appeared to deviate from its theoretical conceptualization.
Conclusions: In absence of distress reduction, distress variability can vary. However, our findings suggest that in-session distress variability does not drive symptom reduction during PE. In contrast, averaged over participants, distress variability was related to symptom improvement, suggesting that those with a more variable distress pattern across sessions show better treatment response. More empirical work is needed to shed light on the effect of distress variability during exposure sessions on treatment outcome and to offer grounds for clinical recommendations.

Keywords: PTSD, Prolonged Exposure, Inhibitory Learning, Distress variability, Change Mechanisms

Introduction

Prolonged exposure (PE) is an effective treatment for posttraumatic stress disorder (PTSD) but a substantial amount of patients do not improve sufficiently (Larsen et al., 2019; McLean et al., 2022). Understanding the underlying mechanisms of successful PE can help to optimize treatment outcomes of PTSD. Fear extinction is thought to be one of the most important mechanisms of action during PE and Inhibitory learning theory (ILT; Craske et al., 2008, 2014, 2022) posits that fear reduction through extinction is accomplished by the formation of inhibitory associations that compete with the original fear-eliciting associations. One of the strategies that have been proposed in ILT to strengthen the inhibitory associations, is to increase variability – in stimuli used during exposure as well as in contexts wherein exposure takes place. Following ILT, it has also been suggested that varying levels of distress during exposure sessions promote treatment effectiveness (Craske et al., 2022; Knowles & Olatunji, 2019). However, somewhat contradictory, the ILT framework also de-emphasizes the role of (in-session) distress patterns. How distress variability relates to exposure effectiveness requires further examination.

Variability of distress levels has been proposed to be linked to beneficial treatment outcomes in different ways (Culver et al., 2012). First, it might allow the inhibitory association to be coupled with a variety of internal states during exposure. This variety of internal states is thought to increase the retrievability of the inhibitory association outside the exposure context, thereby strengthening the inhibitory association and making it more stable. Some suggest that variation in internal states may be viewed as context variability (Culver et al., 2012), which has been associated with a reduced return of fear after extinction in human laboratory studies (Bandarian-Balooch et al., 2015; Dunsmoor et al., 2014). Second, increased variability of in-session distress levels may provide multiple opportunities for patients to disconfirm aversive expected outcomes during exposure. The violation of expected outcomes has been shown to be crucial for learning (Rescorla & Wagner, 1972) and has been associated with enhanced extinction learning in laboratory settings (Brown et al., 2017; Gromer et al., 2022) and (sub-clinical) exposure outcomes (Deacon et al., 2013).

Greater variability of in-session distress levels has been related to better treatment outcomes in adults with public speaking – and contamination anxiety (Culver et al., 2012; Kircanski et al., 2012) and in clinically anxious children, diagnosed with e.g., obsessive compulsive disorder and generalized anxiety disorder (Kircanski & Peris, 2015; Waters et al., 2015). Conversely, a study that manipulated variability in exposure intensity in a sample of adults with an unacceptable obsessional thought did not find that in-session distress variability predicted treatment outcomes (Jacoby

et al., 2019). Another study also found no effect of fear variability on exposure therapy treatment outcome in a sample of pediatric OCD patients (Benito et al., 2018). These inconsistencies warrant a closer examination of the proposed effect of in-session distress variability on treatment outcomes, especially in clinical samples such as PTSD patients, which have not been studied before.

Prior to ILT, rather than in-session distress variability, in-session distress reduction was presumed to be an important index of change during PE. In-session distress reduction (often referred to as within-session habituation or WSH) was thought to provide new information that is incompatible with the pre-existing information about the feared stimuli, thereby weakening links between stimuli and fear responses (Foa & Kozak, 1986). However, since its conceptualization, multiple studies have not found that in-session distress reduction was related to long-term outcomes of PE (see for a review Asnaani et al., 2016). In line with ILT, some studies suggest that in-session distress reduction is not an important target to achieve during PE (Brown, Zandberg, et al., 2019; Sripada & Rauch, 2015).

In-session distress reduction is generally operationalized as the difference between the peak of distress in the session and the distress at the end of the session. However, distress levels may drop earlier in the exposure session, enabling corrective learning to occur, and may increase again as the session continues. For example, after an initial decrease in distress when recounting the traumatic event, the patient might subsequently focus on other distressing parts of the memory (Foa et al., 2007), causing an increase in distress. As such, corrective learning could have occurred, but this would not be reflected in in-session distress reduction from the peak to the end of the session. Possibly, variability of distress levels better captures the process of corrective learning during exposure therapy.

It should also be noted that most studies examining the relationship between in-session distress reduction and symptom decline only looked at between-person effects, and not at within-person effects. That is, these studies have examined whether those with more in-session distress reduction on average have lower PTSD symptoms following treatment. However, within-person effects are thought to be especially relevant for indices of change (Falkenström et al., 2020; Kazdin, 2007), as these effects provide more convincing evidence that a change in the proposed mechanisms leads to a subsequent change in the targeted outcome. Indeed, we recently showed that in-session distress reduction was predictive of next-session reduction in PTSD symptoms when testing the within-person effects (Hoeboer et al., 2022). In other words, in-session distress reduction preceded PTSD symptom improvement and appeared to be an indicator of change during PE. It should still be clarified how patterns of in-session distress present themselves, how they relate to each other, and to treatment outcome.

The aim of the current study was twofold. First, we aimed to provide descriptive information on levels of in-session distress variability and distress reduction throughout PE. Little is known about the session-to-session distress patterns during PE, as averaged effects over treatment are generally reported. We evaluated descriptive statistics (e.g., frequencies, mean, etc.) of distress variability and distress reduction indices across all sessions and examined the relationship between these two variables. Second, we aimed to assess whether in-session distress variability predicted change in PTSD symptoms, through temporal (within-person) analyses. Following ILT, we expected that in-session variability of levels of distress predicted next session change in PTSD symptoms. To enhance comparability with previous findings on this relationship, we carried out additional analyses to assess the between-person effect of distress variability on change in PTSD symptoms.

Methods

Design

The IMPACT study is a randomized controlled trial comparing the effectiveness of standard PE, intensified PE (iPE), and phase-based treatment (PBT), in which PE was preceded by Skills Training in Affective and Interpersonal Regulation (STAIR). For more information on the design and other outcomes of the IMPACT study, we refer to previously published papers (Oprel et al., 2018, 2021). The IMPACT study has been approved by the Medical Ethical Committee of Leiden University Medical Center (NL57984.058.16).

Participants

Participants were diagnosed with posttraumatic stress disorder following childhood abuse (CA-PTSD) established with the Clinician Administered PTSD Scale (CAPS-5; Boeschoten et al., 2018) with at least moderate severity of PTSD symptoms (CAPS-5 score ≥26) and at least one specific memory of the traumatic event. The index trauma had to be related to sexual abuse and/or physical abuse that occurred before the age of 18 and was committed by a primary caretaker or an authority figure. Participants had to be between ages 18 and 65 and proficient in the Dutch language. Exclusion criteria were: (1) involvement in a compensation case or legal procedures concerning admission or stay in The Netherlands, (2) pregnancy, (3) severe non-suicidal self-injury (NSSI) which required hospitalization during the past three months, (4) severe suicidal behavior: a suicide attempt during the past three months or acute suicidal ideations with serious intent to die with a specific plan for suicide and preparatory acts, (5) severe disorder in the use of alcohol or drugs in last three months, (6) cognitive impairment (estimated IQ < 70), (7) changes in

psychotropic medication in the two months before inclusion, and (8) engagement in any current psychological treatment.

The IMPACT study sample consisted of 149 participants. For the current study, we excluded participants in the PBT condition (n = 50) as PE in this condition was preceded by emotion regulation skills training. This skills training may influence the proposed working mechanism of the subsequent PE and may thereby affect the outcomes of the analysis. Furthermore, we excluded participants who did not complete at least 2 treatment sessions (n = 13), as this was a requirement for the temporal analyses. Our final sample, therefore, consisted of 86 participants, most of whom were female (79%), and with ages ranging from 20 to 60 years (M = 36.8, SD = 11.5). Power analyses for sample size justification were conducted for the parent trial (Oprel et al., 2021). We previously did a study using a similar statistical method and comparable estimated effect sizes (medium effects; Cohen's d = 0.5) and found to have sufficient power to detect significant effects using a sample of N = 86 (Hoeboer et al., 2022).

In total, our dataset included 1069 therapy sessions. The first session (n = 86) included psychoeducation and treatment planning. From the second session onwards, sessions included imaginal exposure (n = 983). On average, participants completed 12.4 sessions (SD = 3.6), with a minimum of 3 and a maximum of 16.

Procedure

Eligibility for the study was assessed during a baseline assessment. After this baseline assessment, patients were randomly allocated to one of the three treatment conditions with a 1:1:1 ratio (PE, iPE, and PBT). Patients in the PE condition received a maximum of 16 weekly, 90-min sessions. Patients in the iPE condition received a maximum of 14 90-min sessions, starting with 12 sessions over 4 weeks and followed by two sessions after one and two months. Exposure in the iPE condition was delivered by two therapists due to practical considerations. The treatment manual for both conditions was based on the protocol described by Foa et al. (2007). In both conditions, the first session consisted of psycho-education and a case conceptualization. The other sessions consisted of 60-min imaginal and in-vivo exposure. Between sessions, patients were instructed to perform homework assignments (e.g., listening to audiotape recordings of the imaginal exposure or performing in-vivo exercises).

All therapists were trained in the protocol and had to pass an exam with pilot patients to ensure competency in the exposure protocol. To further ensure the therapist's adherence to the treatment protocol, they received weekly group supervision (supervised by RAdK and AvM). The therapists had at least a master's degree in psychology and on average 10 years of experience in mental health

services. A random selection of PE sessions (approximately 10% of the total PE sessions) was rated by independent observers for treatment adherence based on the Dutch translation of the original adherence rater checklist scale. Protocol adherence was high (*M* session elements completed = 90%, *SD* = 18%).

The majority of patients completed 14 sessions (*n* = 55, 64%), but only a few participants completed session 15 (*n* = 18, 21%) and session 16 (*n* = 15, 17%), mostly because the iPE condition contained 14 sessions. Due to insufficient observations in these sessions (*n* = 33), they were omitted from our temporal analyses.

Measures

PTSD symptoms

Self-reported PTSD symptoms, the primary outcome of this study, were measured with the weekly version of the PTSD checklist for DSM-5 (PCL-5; Blevins et al., 2015). The PCL-5 consists of 20 self-report items that are rated on a 5-point Likert scale ranging from 0 (not at all) to 4 (extremely). The total PCL-5 score ranges from 0 to 80, with higher scores indicating higher symptom severity. The PCL-5 is considered to have good psychometric properties with high Cronbach's α (.94) in previous studies (Blevins et al., 2015). The PCL-5 was administered at the start of each session and was completed with reference to the index trauma. It should be noted that we did not have PCL-5 scores for each session in the iPE condition as the PCL-5 was administered once a week. For the iPE condition, this meant that the PCL-5 was administered every three sessions (as participants received three sessions per week). In total, we had PCL-5 data for 782 sessions.

In-session distress indices

From the second PE session onwards, participants rated subjective units of distress (SUD) for every 10 minutes of the in total 60-minute exposure (7 in-session time points) on a scale from 0 (*no distress*) to 100 (*maximum distress*). In line with previous work (Culver et al., 2012; Jacoby et al., 2019; Kircanski et al., 2012; Waters et al., 2015) in-session variability of distress was calculated for each session by taking the standard deviation of SUDs within a specific session (hereafter referred to as VAR-SD). Also in line with previous work (Badour et al., 2017; Harned et al., 2015; Hendriks et al., 2018; Hoeboer et al., 2022), in-session distress reduction (hereafter referred to as WSH) was calculated for each session by subtracting the SUD end (i.e., the last reported SUD score within a session) from the SUD peak (i.e., the highest reported SUD score within a session). Contrary to previous work, we did not average these outcomes across all sessions, as we used the data per session for our within-subjects analyses. For our post hoc between-subject analyses, we did average these outcomes across all sessions. In total, there were 933 sessions with SUD data.

Statistical analyses

The data analysis plan of this study was pre-registered at the open science framework (OSF; osf.io/n26xm). The first aim of this study was to provide descriptive information on distress variability and its relationship with distress reduction. We calculated the within-person correlation between VAR-SD and WSH across the 933 sessions for which SUDs data was available. We also calculated frequencies of the amount of distress variability and distress reduction. We created separate groups for the level of distress variability and distress reduction (low, medium, and high) based on the 33rd and 66th percentiles. This provided categories for within session distress variability and distress reduction, allowing us to plot distress patterns of (1) high variability and high reduction, (2) high variability and low reduction, (3) low variability and high reduction, and (4) low variability and low reduction. Sessions within the medium groups were not plotted, as these were hard to interpret and not very insightful.

For our second aim, to assess whether in-session variability in SUD levels predicted next session PTSD symptom reduction, we used dynamic panel models based on maximum likelihood estimation (Allison et al., 2017). This allowed us to assess within-person effects. Models were fitted using the Lavaan and dpm package (Rosseel, 2012) in Rstudio (version 2022.12.0). First, we assessed the effect of in-session distress variability on PTSD outcome. We used PCL-5 scores as dependent variable with the auto-regressive effect of the PCL-5 scores and the cross-lagged effect of distress variability (VAR-SD) per session as independent variable. To illustrate with an example, PCL-5 scores of session 4 were predicted by PCL-5 scores of session 3 and VAR-SD scores of session 3. As the iPE condition had fewer PCL-5 measurement points (session 1, 4, 7, 10, 12, 13, and 14), PCL-5 scores of session 4 were in this condition predicted by PCL-5 scores of session 1 and VAR-SD scores of session 3. To test temporality, we ran the reversed model with PCL-5 as the independent variable and VAR-SD as dependent variable. We also ran additional analyses to test the effect of condition on the relationship between VAR-SD and PTSD symptoms, as (1) the delivery format of exposure therapy in the two conditions might affect the outcomes of our analysis and (2) the PCL-5 was not administered in every session in the iPE condition which may have affected the autoregressive effect. We initially planned to evaluate whether distress variability added predictive value over in-session distress reduction when both variables were entered in the model, but this model did not converge.

Post hoc (i.e., after preregistration), we conducted additional analyses to assess the between-person effect of distress variability across sessions ('averaged distress variability') on PTSD symptom improvement. This was tested using a dynamic panel model, with PCL-5 scores as dependent variable, the autoregressive effect of the

PCL-5 score and the averaged distress variability (fixed effect) as independent variables.

Results

Means and standard deviation of PCL-5 scores, within-session distress variability, and distress reduction can be found in Table 1. Over treatment, PCL-5 scores decreased from 54.15 (*SD* = 12.65) in session 1 to 31.22 (*SD* = 23.10) in session 14. On average, distress variability was larger in early sessions (session 1; *M* = 16.71) compared to later sessions (session 14; *M* = 9.38) and varied between participants (range *SD*s of VAR-SD across sessions is 8.42–10.03). SUD-WSH also decreased over the course of treatment and varied between participants (see Table 1).

Table 1. Descriptive information per session of PTSD symptomatology and within-session distress variability and – reduction

	PCL-5			SUD: VAR-SD			SUD: WSH		
Session	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
1	85	54.29	12.72						
2	43	55.93	12.59	86	16.71	8.54	86	25.48	23.81
3	44	54.25	15.74	85	16.38	10.03	85	24.94	26.53
4	83	50.61	15.32	83	14.47	8.97	83	22.35	21.00
5	42	46.95	17.87	79	14.55	9.89	79	23.99	21.72
6	40	46.10	18.42	73	13.36	9.41	73	21.63	20.73
7	73	42.93	18.83	73	15.16	8.57	73	20.41	18.76
8	35	38.03	21.71	69	14.26	8.84	69	18.96	17.63
9	33	34.94	21.17	64	13.78	8.92	64	18.91	19.51
10	66	36.50	20.42	64	13.62	8.70	64	21.20	21.67
11	28	34.07	23.94	63	13.39	8.42	63	19.53	17.89
12	62	32.08	20.07	60	11.85	8.44	60	19.67	21.30
13	62	30.35	20.95	56	11.09	8.68	56	15.02	16.58
14	54	31.22	23.10	46	9.38	8.81	46	15.07	19.22

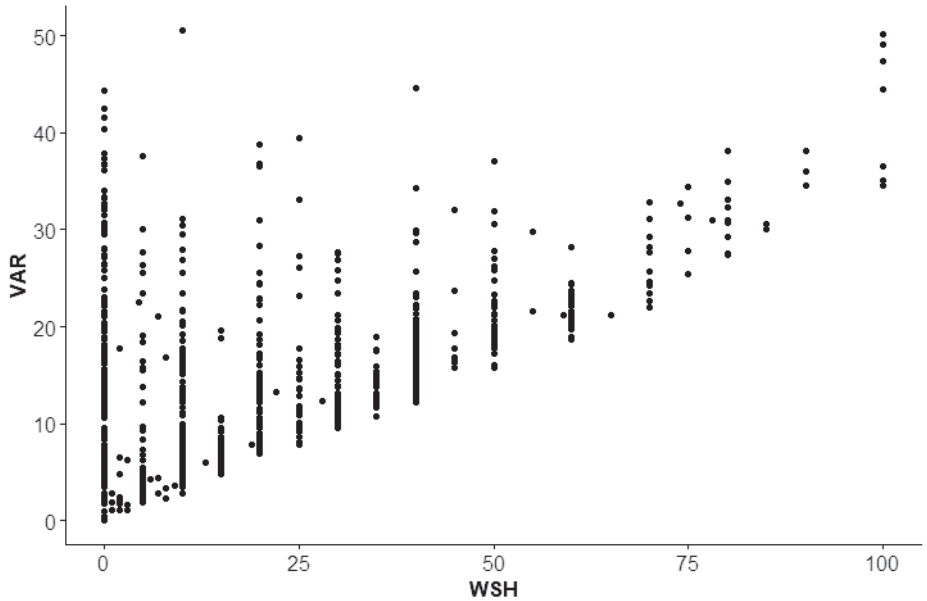
Note. PCL-5 = PTSD checklist for DSM-5; SUD = subjective units of distress; VAR-SD = variability operationalized as the standard deviation of SUDs per session (Kircanski et al., 2012); WSH = within-session habituation/distress reduction.

Association between within-session distress reduction an variability

As expected, the within-person correlation showed that within-session distress variability (VAR-SD) and within-session distress reduction (WSH) were significantly and positively correlated ($r = 0.53$, $p < 0.001$). However, an absence of distress reduction (i.e., SUD-WSH = 0) could co-occur with a wide range of in-session distress variability (i.e., VAR-SD varies between 0 and 45; see Figure 1).

To gain more insight in divergent or convergent distress variability and – reduction patterns, we created three groups for distress variability and – reduction (i.e., low, medium and high) based on their approximate 33rd percentiles (see Table 2). A pattern of high distress variability and high distress reduction was most common (*n* = 175, 18.8%), followed closely by a pattern of low distress variability and low distress reduction (*n* = 160, 17.1%). Approximately a tenth of the sessions (*n* = 83, 8.9%) show a pattern of high distress variability and low distress reduction.

Figure 1. Relation between in-session distress variability and reduction



Note. VAR = distress variability (original operationalization, Kircanski et al., 2012); WSH = within-session habituation/distress reduction.

Surprisingly, we found through visual inspection (see Figure 2) that the operationalization of distress variability by calculating the standard deviation of distress levels in a session (as was done in previous studies; Culver et al., 2012;

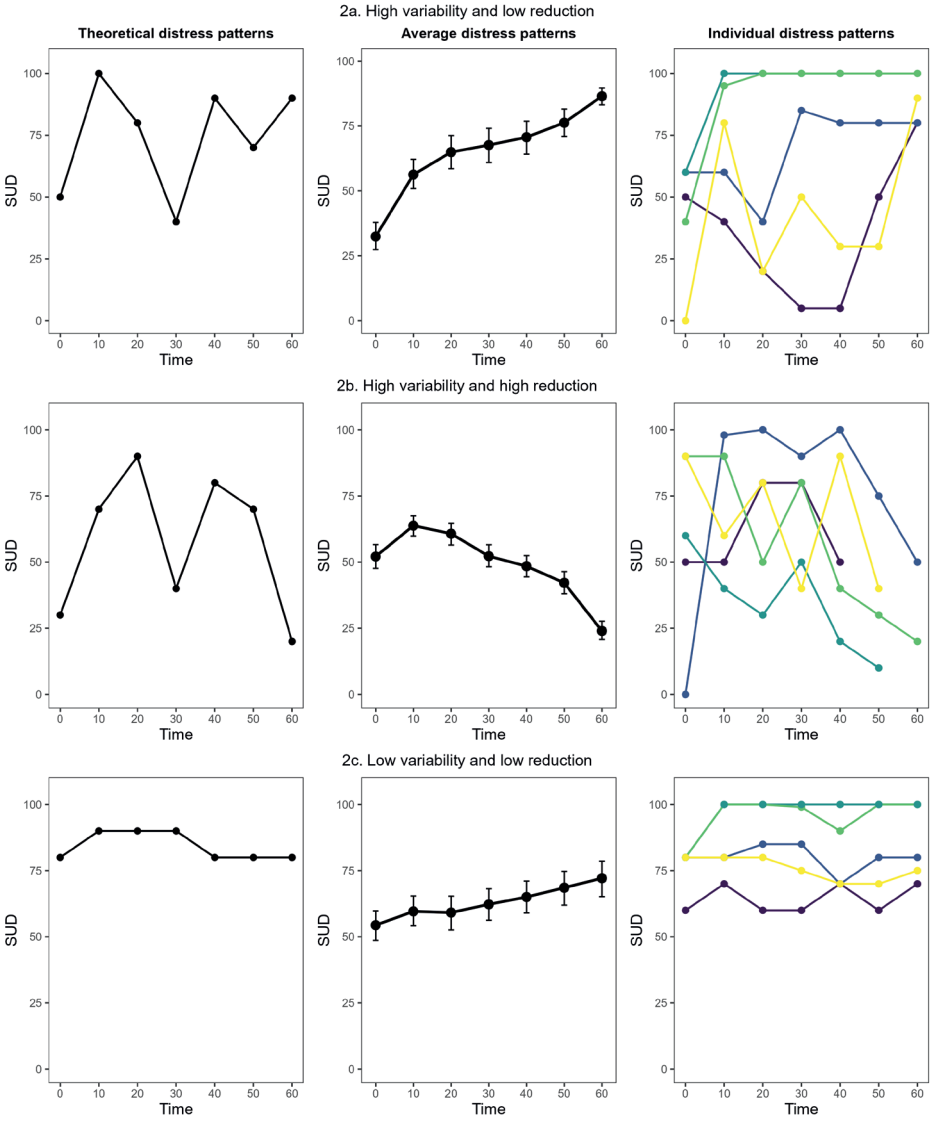
Jacoby et al., 2019; Kircanski et al., 2012; Waters et al., 2015) did not always capture the ups-and-downs pattern, which is crucial for the theoretical concept of distress variability (as it suggests opportunities for corrective learning). For instance, the top right graph of Figure 2 shows a session (green line) that is assigned to high variability, but this session merely shows an increase in distress levels. Merely increasing distress levels within the relatively high variability group was not a rare occasion (see Appendix A, Figure A1). Figure 2 shows what we theoretically expect a pattern within a category to look like, the average of distress patterns within a category, and distress patterns of five randomly selected sessions. Especially in the 'high variability and low reduction' category, the plots indicate that VAR-SD does not fully align with theory. To further develop the field, we came up with alternative metrics for distress variability (post hoc), which are discussed in Appendix B.

Table 2. Three group crosstabulation of distress variability and distress reduction

		VAR-SD			
		Low	Medium	High	Total
		n (%)	*n* (%)	*n* (%)	*n* (%)
WSH	Low	160 (17.1)	59 (1.6)	83 (8.9)	302 (32.4)
	Medium	149 (16.0)	127 (13.6)	53 (5.7)	329 (35.3)
	High	0 (0)	127 (13.6)	175 (18.8)	302 (32.4)
	Total	309 (33.1)	313 (33.5)	311 (33.3)	933 (100)

Note. VAR-SD = distress variability; WSH = within-session habituation/distress reduction. In this table, *n* refers to the number of sessions.

Figure 2. Examples of SUD trajectories



Note. Plots on the left depict the pattern that is theoretically expected within a category (based on hypothetical data). Plots in the middle show the averaged distress pattern with the 95% Confidence Interval. Plots on the right show distress patterns of five randomly selected sessions.

Temporal analyses

Outcomes of the temporal analyses can be seen in Table 3. Within-session distress variability (VAR-SD) was not significantly related to lower PTSD symptoms in the next session, b = -0.05, SE = 0.05, z = -0.95, p = .345. We found no evidence for differences between the two treatment conditions (PE vs. iPE)[1], b = 0.08, SE = 0.06, z = 1.36, p = .172. The reversed effect was also not significant, i.e., PTSD symptoms were not significantly related to less within-session distress variability (VAR-SD) in the next session, b = 0.05, SE = 0.03, z = 1.61, p = .108. We refer to Appendix B for the temporal analyses with the alternative distress variability metrics.

Table 3. Temporal analyses of PTSD symptoms and distress variability

Temporal effect	Estimate	SE	*z*-value	*p*-value
Predicting PCL-5 from VAR-SD				
Lagged VAR-SD	-0.05	0.05	-0.95	.345
Autoregressive PCL-5	0.70	0.05	15.55	<.001
Predicting VAR-SD from PCL-5				
Lagged PCL-5	0.05	0.03	1.61	.108
Autoregressive VAR-SD	0.24	0.04	5.55	<.001

Note. PTSD = posttraumatic stress disorder; PCL-5 = PTSD checklist for DSM-5; VAR-SD = in-session distress variability following original operationalization (Kircanski et al., 2012).

Between-person analyses

Correcting for the autoregressive effect of PTSD symptoms, averaged distress variability (VAR-SD) was a significant predictor of a decrease in PTSD symptoms over the course of treatment, b = 0.12, SE = 0.03, z = 3.71, p < .001.

Discussion

In this study we aimed to provide descriptive information on in-session distress variability and distress reduction during PE, and to assess whether in-session distress variability predicted next session PTSD symptom improvement. We found

1 Based on a reviewer's suggestion we ran a sensitivity analysis on the PE condition only (n = 44). In this model, within-session distress variability (VAR-SD) was not significantly related to lower PTSD symptoms in the next session, b = -0.09, SE = 0.06, z = -1.41, p = .159. Nor did we find evidence for the reversed effect: PTSD symptoms were not significantly related to less within-session distress variability (VAR-SD) in the next session, b < 0.01, SE = 0.04, z = 0.06, p = .954.

that sessions that contained no in-session distress reduction could have a wide range of in-session distress variability. That is, some PE sessions show distress variability but not reduction. In-session levels of distress variability were not predictive of next session PTSD symptom decline. Our post hoc analyses showed that the average level of distress variability over the sessions was related to better treatment outcomes (i.e., more PTSD symptom improvement) during treatment.

The positive relation between distress variability and distress reduction was not surprising, as reduction is a form of variability. More to our surprise, when looking at distress patterns that had a relatively high score on distress variability in our dataset, these also contained sessions marked by a steady in- or decrease in distress. It seems that participants with such distress patterns did not have multiple opportunities to disconfirm aversive expected outcomes regarding the infinite duration or intensity of distress (i.e., they did not show more 'ups' and 'downs'), yet scored high on this variability index. This operationalization of variability (calculating the standard deviation of in-session distress levels) does not fully align with our theoretical starting point that variability would better capture the process of corrective learning during exposure therapy than distress reduction (i.e., WSH). Up until now, studies on distress variability in adult samples have been carried out in clinical analogue populations (Culver et al., 2012; Jacoby et al., 2019; Kircanski et al., 2012). Suggestively, distress patterns may be different in clinical samples, where lack of in-session ups-and-downs may be more common. We explored alternative operationalizations of within-session distress variability that could better capture the up-and-down pattern (see Appendix B). Arriving at a suitable metric that captures distress variability proved to be difficult and distress variability and distress reduction are not entirely separable. Our alternative metrics further the field, but future research (across different samples) is needed.

Our hypothesis that in-session distress variability predicted next-session PTSD symptom improvement was not confirmed, neither with the original conceptualization of distress variability nor with our alternative conceptualizations. This is in line with two earlier studies (Benito et al., 2018; Jacoby et al., 2019), but not with several others (Culver et al., 2012; Kircanski et al., 2012; Waters et al., 2015). One difference between the studies that found an effect and ours is that these studies assessed averaged distress variability. However, especially temporal (or, within-person) effects are important when assessing change mechanisms, as they are more indicative of change processes than averaged effects (Falkenström et al., 2020; Kazdin, 2007). In our post hoc analysis, we found that averaged (or, between-person) distress variability was related to better treatment outcome. Our findings indicate that, on average, patients who show more distress variability show more improvement in PTSD symptoms. However, a patient who showed more variability

in one session did not show more next-session improvement. Therefore, in-session distress variability may reflect who responds well to treatment in general, rather than it being a mechanism of change during PE.

Taken together with our previous study, which showed that more in-session distress reduction (i.e., WSH) was related to more next-session PTSD symptom reduction (Hoeboer et al., 2022), the results of this study suggest that reduction in distress is a better predictor of PTSD symptom reduction than distress variability. This is in line with the original propositions of 'emotional processing theory' (Foa & Kozak, 1986), but not with several other earlier studies (see for a review Cooper, Clifton, et al., 2017). Partly based on these earlier findings, striving for in-session distress reduction has been de-emphasized in ILT (Craske et al., 2008, 2014). Moreover, ILT proposes to sometimes use strategies designed to maintain heightened in-session distress levels (Craske et al., 2014). Clinically, our current findings suggest that exposure interventions for PTSD do not need to be tailored to increase in-session distress variability.

Crucially, although increasing distress variability has been linked to the inhibitory learning approach, this approach also posits that we should move away from distress levels as a yardstick for successful exposure sessions and rather focus on the reduced credibility of the expected aversive outcome. Distress variability is difficult to control and operationalize in clinical settings. Multiple factors can cause distress variability, such as exposure length and stimulus. In contrast, several studies conducted in pre-clinical samples have suggested that a focus on the disconfirmation of expected outcomes enhances extinction learning and could thereby optimize exposure therapy outcomes (Brown et al., 2017; Deacon et al., 2013). However, other recent studies that aimed to translate these findings to more clinical populations have reported mixed results (Buchholz et al., 2022; de Kleine et al., 2017; Elsner et al., 2022; Krause et al., 2022). Therefore, the processes of expectancy violation and corrective learning during exposure therapy also warrants further investigation.

The current study also has some limitations. Variability has mostly been assessed in clinical analogue studies, while the current study used a clinical PTSD sample (with multiple traumas). Although the use of this sample increased ecological validity, it decreased the controllability of the variables under investigation. Importantly, extinction theory is about extinction processes on a single stimulus (e.g., a loud tone; CS) with an outcome specifically related to this CS (e.g., startle reflex; CR). In the current work, the exposure stimulus (trauma memory) could vary between sessions and the outcome index (PCL-5) assessed weekly PTSD symptoms, which may or may not have a direct link to the CS targeted in the prior exposure session. More studies with a single stimulus and an outcome directly linked to that stimulus

are needed to gain further insight into the effect of distress variability in clinical samples. Second, distress variability was measured in-session with 10-min intervals. Possibly, this is too infrequent to capture changes in distress. Related to this, the metrics we used to capture distress variability might not reflect the theoretical conceptualization of allowing for more corrective learning. The cause of distress change (i.e., reduction due to escape behavior or learning, see also Benito et al., 2018) is also unknown in the current study and future studies may benefit from function-based coding of distress (for instance via video-ratings, Alpert et al., 2021; Alpert, Hayes, et al., 2023 for examples). Third, we used the most commonly used peak-end operationalization of WSH

(Cooper, Clifton, et al., 2017; Foa & McLean, 2016; Hendriks et al., 2018; Nacasch et al., 2015). However, this metric may not fully capture the process of within-session distress reduction as this calculation misses instances of reduction that appear earlier in the session (Benito et al., 2018). As with distress variability, the operationalization of in-session distress needs refining and more granular indices of exposure processes are needed to ground adaptations in therapeutic procedures. Fourth, we conducted the averaged analyses post hoc and these outcomes should be interpreted with caution, and be replicated in studies with a priori hypotheses. Additionally, we did not have PCL-5 data for each session in the iPE condition, resulting in less power and lower precision in the assessment of change processes in this condition. Future work may more easily overcome these challenges by administering abbreviated measures, such as the 4- or 8-item PCL-5 (Price et al., 2016), although the decreased amount of variance due to fewer items may decrease power again (i.e., smaller differences are harder to detect). Finally, the current study is a re-analysis of data from the IMPACT study and the findings of the current study should be replicated in other samples. Notwithstanding these limitations, the current study is the first to assess the effect of distress variability on exposure treatment outcome in a clinical sample of PTSD patients.

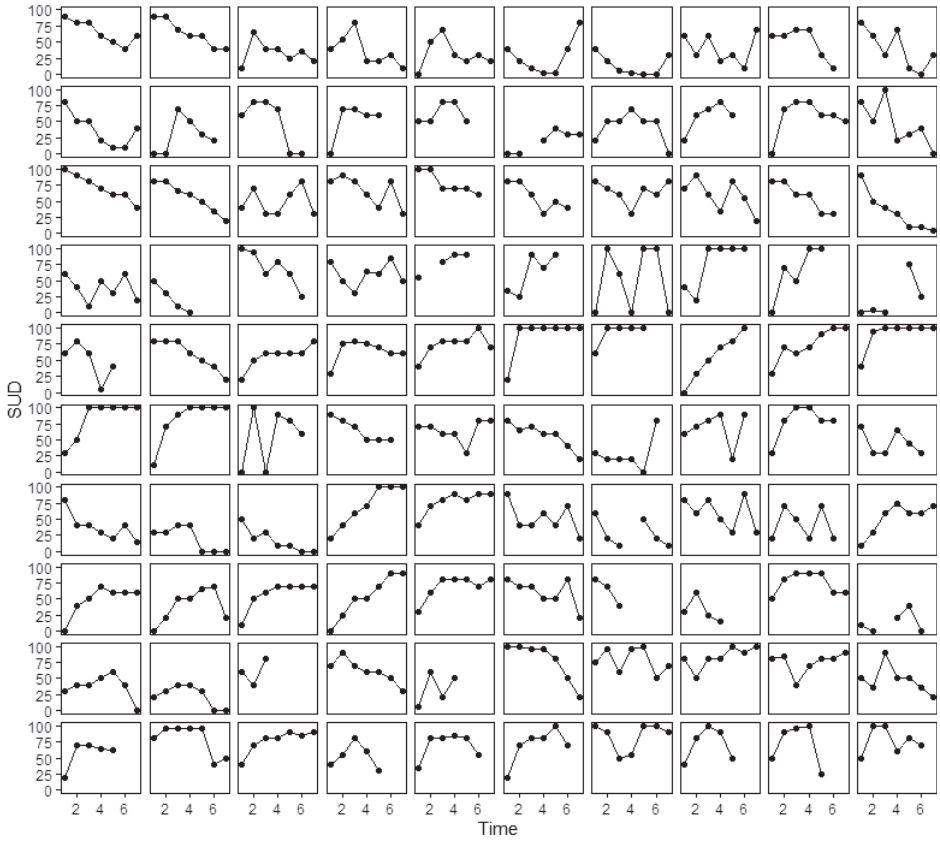
The results of this study showed that in-session distress variability during PE did not predict subsequent PTSD symptom decline. Averaged levels of distress variability were related to a greater decline in PTSD symptoms. On a conceptual level, the operationalization of distress variability requires further investigation. We argue that distress variability is not a predictor of symptom decline during exposure therapy for PTSD and that more research should be conducted on the temporal relationship between in-session distress patterns and symptom change to ground clinical recommendations on exposure procedures.

Appendix A

Visual inspection of high distress variability patterns

We plotted the distress patterns of 100 random sessions that had relatively high distress variability to show that VAR-SD (as was done in previous studies; Culver et al., 2012; Jacoby et al., 2019; Kircanski et al., 2012; Waters et al., 2015) does not always capture the ups-and-downs-pattern and that patterns with only 'ups' in this group are not a rare occasion.

Figure A1. Individual session patterns of distress with high distress variability, random selection (n = 100)



Appendix B

Alternative metrics for in-session distress variability

Introduction

Based on visual inspection of in-session distress patterns (also see Figure 2 [ms.] and Figure A1), we found that the operationalization of distress-variability via the standard deviation of SUD scores per session (Culver et al., 2012; Jacoby et al., 2019; Kircanski et al., 2012; Waters et al., 2015) was not in line with how distress variability is theoretically thought to improve treatment outcomes. Namely, some sessions showed a pattern where there were no (multiple) opportunities to disconfirm expected aversive outcomes, but were still relatively high in variability. Moreover, distress variability and reduction overlap, both theoretically and mathematically. Sessions that have a large distress reduction by definition have a large variability if one uses the SUD-SD metric. Although some overlap between the variability and reduction constructs may remain, they should be clearly distinguishable. We therefore explored other operationalizations of distress variability. We explore two alternative metrics: 1) the average of absolute difference scores between subsequent time points within a session (VAR-abs), i.e., the sum of ups-and-downs in a session , and 2) the residual sum of squares of the linear regression across SUD timepoints (VAR-RSS), i.e., the variation of SUD scores controlled for the general pattern of SUDs decline (or increase) within a session. First, we assessed whether these alternative operationalizations better captured the theoretical concept of distress variability, by visual inspection and their relationship with WSH. Second, we assessed whether distress variability using these alternative metrics was predictive of PTSD symptom improvement.

Methods

Measures

During each exposure session, subjective units of distress (SUDs) was rated by participants every 10 minutes. We created two alternative metrics for within-session distress variability. The first alternative metric for in-session distress variability was calculated by taking the absolute difference between each subsequent time point in a session and then by averaging these differences into one score (VAR-abs). The second alternative metric was calculated by fitting a linear regression across the SUDs for each session and taking the residual sum of squares (RSS) of this model

as distress variability (VAR-RSS), thereby creating a metric that is not contaminated by the general time effect of the in-session distress pattern.

Statistical analyses

We re-ran our main analyses (post-hoc, after pre-registration), but now with the alternative metrics for distress variability. We used dynamic panel models based on maximum likelihood estimation (Allison et al., 2017). This allowed us to assess within-person effects. Models were fitted using the Lavaan and dpm package (Rosseel, 2012) in Rstudio (version 2022.12.0). We used PCL-5 scores as dependent variable with the auto-regressive effect of the PCL-5 scores the cross-lagged effect of distress variability per session as independent variable, using VAR-abs in one analysis and using VAR-RSS in another analysis. Additionally, we ran the reversed models, with the alternative distress variability metrics as the outcome variable, and cross-lagged PCL-5 and auto-regressive distress variability as predictors.

Results

Descriptives

See Table B1 and B2 for the correlations between the in-session distress metrics. VAR-SD (the original operationalization of distress variability; Culver et al., 2012; Kircanski et al., 2012) showed strong, significant correlations with VAR-abs and VAR-RSS. The novel metrics also exhibit a decrease in their correlations with WSH; with VAR-RSS demonstrating a slightly more pronounced difference than VAR-abs.

Table B1. Within-person correlations of in-session distress metrics

	VAR-SD	VAR-abs	VAR-RSS	WSH
VAR-SD	.	.	.	.
VAR-abs	.76**	.	.	.
VAR-RSS	.66***	.68***	.	.
WSH	.53***	.46***	.44***	.

Note. VAR-SD = in-session distress variability following original operationalization; VAR-abs = first alternative metric for distress variability; VAR-RSS = second alternative metric for distress variability; WSH = within-session habituation/distress reduction.

Temporal analyses

See table B2 for outcomes of the temporal analyses with the alternative VAR metrics. VAR-abs was not significantly related to lower PTSD symptoms in the next session, *b* = -0.03, *SE* = 0.02, *z* = -1.38, *p* = .166. The reversed effect was also not significant, i.e., PTSD symptoms were not significantly related to lower VAR-abs in the next session, *b* = 0.13, *SE* = 0.07, *z* = 1.88, *p* = .060. VAR-RSS was not significantly related to lower PTSD symptoms in the next session, *b* < -0.01, *SE* < 0.01, *z* = -1.82, *p* = .06. The reversed effect was also not significant, i.e., PTSD symptoms were not significantly related to lower VAR-RSS in the next session, *b* = 2.94, *SE* = 2.92, *z* = 1.01, *p* = .314.

Between-person analyses

Correcting for the autoregressive effect of PTSD symptoms, the averaged VAR-abs over sessions was a significant predictor of a decrease in PTSD symptoms over the course of treatment, *b* = -0.12, *SE* = 0.03, *z* = -3.60, *p* <.001, as well as the averaged VAR-RSS, *b* < -0.01, *SE* < 0.01, *z* = -3.53, *p* <.001.

Table B2. Temporal analyses of alternative distress variability metrics

Temporal effect	Estimate	SE	*z*-value	*p*-value
Predicting PCL-5 from VAR-abs				
Lagged VAR-abs	-0.03	0.02	-1.38	.166
Autoregressive PCL-5	0.68	0.04	15.84	<.001
Predicting VAR-abs from PCL-5				
Lagged PCL-5	0.13	0.07	1.88	.060
Autoregressive VAR-abs	0.14	0.04	3.51	<.001
Predicting PCL-5 from VAR-RSS				
Lagged VAR-RSS	<-0.01	<0.01	-1.82	.060
Autoregressive PCL-5	1.04	0.01	72.17	<.001
Predicting VAR-RSS from PCL-5				
Lagged PCL-5	2.94	2.92	1.01	.314
Autoregressive VAR-RSS	0.39	0.3	11.98	<.001

Note. PCL-5 = PTSD checklist for DSM-5; VAR-abs = first alternative metric for distress variability; VAR-RSS = second alternative metric for distress variability.

Brief discussion

In this appendix, we have explored alternative metrics that aim to capture in-session distress variability. The high within-person correlation between the different distress variability metrics support the notion that they are reflecting a similar construct,

indicating that the new metrics may be valid alternatives. When using the alternative metrics, the correlation with WSH decreased somewhat - albeit to a limited extent. It is possible that distress variability and distress reduction may not be further disentangled. The outcomes of the temporal analyses with these alternative metrics offer no new interpretations. We refer to our main manuscript for a more in-depth discussion on distress variability.