Journal of Behavior Therapy and Experimental Psychiatry 64 (2019) 54–63

Contents lists available at ScienceDirect

Journal of Behavior Therapy and Experimental Psychiatry

journal homepage: www.elsevier.com/locate/jbtep

Is the hierarchy necessary? Gradual versus variable exposure intensity in the treatment of unacceptable obsessional thoughts

Ryan J. Jacoby[a,b,*], Jonathan S. Abramowitz[a], Shannon M. Blakey[a], Lillian Reuman[a]

[a] University of North Carolina at Chapel Hill, Davie Hall, Campus Box 3270, Chapel Hill, NC, 27599, United States
[b] Massachusetts General Hospital / Harvard Medical School, 185 Cambridge St, Boston, MA, 02114, United States

ARTICLE INFO

Keywords:
Exposure therapy
Obsessive compulsive disorder
Fear
Habituation
Psychophysiology
Adulthood (18 yrs & older)

ABSTRACT

Background and objectives: Although research suggests that introducing varying levels of fear during exposure enhances outcomes for some anxiety-related problems, this has not been examined in the context of obsessions. The current preliminary study tested the hypothesis that introducing variability in exposure intensity would improve long-term outcomes relative to traditional gradual (hierarchical) exposure

Methods: Adults ($N = 40$) with a moderately distressing unacceptable obsessional thought were randomly assigned in parallel to four twice-weekly sessions of: (a) gradual exposure (EXP-G; $n = 19$) emphasizing hierarchical exposure completion, or (b) variable exposure (EXP-V; $n = 21$) emphasizing variability in exposure intensity

Results: There were no significant differences in pre to post changes between groups using self-report, interview, or behavioral outcomes (as evaluated by an independent assessor blind to treatment condition). Group comparisons at 3-month follow-up did not reach statistical significance but were moderate in magnitude. Specifically, as measured by clinical interview (the Yale-Brown Obsessive-Compulsive Scale; primary outcome) and self-report, individuals in the EXP-G group maintained gains at 3-month follow-up, while the EXP-V group continued to improve. Treatment expectancies and satisfaction were comparable for both groups. Five participants withdrew from the EXP-G condition, and none withdrew from the EXP-V condition. In contrast to previous studies, variability in subjective and physiological fear during exposure did not predict outcomes

Limitations: The study employed an analogue sample with moderate unacceptable obsessions, and results should be replicated in clinical samples

Conclusions: Variable exposure warrants future study to understand the mechanisms, moderators, and implications of this novel approach

1. Introduction

Cognitive behavioral therapy (CBT), including exposure and response prevention (ERP), is the most efficacious psychological treatment for obsessive compulsive disorder (OCD), with 69% of patients experiencing clinically significant improvement (Eddy, Dutra, Bradley, & Westen, 2004). Exposure can be conducted via encounters with external fear cues or via imaginal exposure to obsessional thought/images that provoke anxiety. Response prevention entails refraining from rituals that serve as an escape from obsessional fear. Traditionally, ERP is conducted using a gradual approach in which stimuli are rank-ordered hierarchically according to a patient's predicted subjective units of distress (SUDS). Treatment typically begins with exposures of lesser intensity that are repeated until habituation (i.e., reduction) of fear occurs before progressing gradually until habituation to the most anxiety-provoking exposures is accomplished.

This method of implementing exposure is derived from emotional processing theory (EPT; Foa, Huppert, & Cahill, 2006), which proposes that habituation within and between sessions is a critical indicator that corrective learning is taking place. Despite EPT's empirical support, however, a substantial percentage of patients (14–31%) are non-responders to ERP (e.g., Norberg, Calamari, Cohen, & Riemann, 2008), and of those who respond, 50–60% experience partial symptom relapse (e.g., Eisen et al., 2013). These attenuated outcomes, and the fact that "return of fear" can occur, suggest that enhancing the long-term efficacy of ERP remains an important clinical need.

Findings from basic research on extinction learning indicate that hierarchy-driven gradual exposure may inadvertently *attenuate* long-

* Corresponding author. Massachusetts General Hospital, Department of Psychiatry, 185 Cambridge Street, Suite 2000, Boston, MA 02114, United States.
E-mail addresses: rjjacoby@mgh.harvard.edu (R.J. Jacoby), jabramowitz@unc.edu (J.S. Abramowitz), sblakey@unc.edu (S.M. Blakey), reuman@unc.edu (L. Reuman).

https://doi.org/10.1016/j.jbtep.2019.02.008
Received 25 April 2018; Received in revised form 24 January 2019; Accepted 20 February 2019
Available online 26 February 2019
0005-7916/

term outcomes by failing to extinguish fear of anxiety itself (Craske et al., 2008; Jacoby & Abramowitz, 2016). Specifically, gradual exposure may ironically nurture anticipatory anxiety of items toward the top of the hierarchy by communicating that patients are not "ready" for intense anxiety until they have first confronted lower/"safer" levels (Abramowitz & Arch, 2014). Moreover, patients sometimes rely on habituation as an escape from anxiety (e.g., "I will do this exposure because I know my anxiety will come down") and thus use exposure to *control* anxiety rather than learning that it is safe and manageable (Craske, Liao, Brown, & Vervliet, 2012). The prioritization of habituation during exposure might also result in the misinterpretation of normal surges of anxiety following treatment as signals of danger or failure, ultimately contributing to relapse (Craske et al., 2008). Accordingly, although gradual exposure may facilitate performance *during* exposure, it might make one vulnerable to return of fear in the *long-run*.

Research on fear extinction suggests that a promising alternative for optimizing long-term ERP outcomes is to foster "desirable difficulties" during exposure. One way to do this is to vary the intensity of exposures, rather than proceeding hierarchically (e.g., Craske et al., 2008). Variation is a "difficulty" because it introduces challenges *during* exposure and slows the rate at which fear declines in the short-term. Yet this is "desirable" as it results in more durable long-term learning and may minimize return of fear (Hermans, Craske, Mineka, & Lovibond, 2006).[1] First, variability is beneficial because it maximizes opportunities for surprise; that is, when there is greater discrepancy between fear-based predictions (i.e., expectancies; e.g., "I can't handle uncertainty and anxiety") and what actually occurs (e.g., anxiety and uncertainty are safe and manageable) this results in more stable extinction learning (Rescorla & Wagner, 1972). Second, variability in exposure maximizes retrieval cues for extinction, as the experience of fluctuating fear during exposure allows for a broader variety of emotions and internal stimuli to become associated with extinction (Bouton, 2000). Third, variable exposure requires patients to engage in higher order learning (i.e., synthesizing vs. simply remembering), which also maximizes long-term extinction (Bjork & Bjork, 2006). Finally, because encounters with obsessional stimuli present unexpectedly in the real world, variable exposure might best prepare patients to manage anxiety following therapy (and inoculate them against return of fear).

Two studies suggest that varying exposure intensity enhances long-term outcomes. First, individuals with spider phobia (Rowe & Craske, 1998) who received "varied-stimulus" exposure (i.e., using four spiders of varying "shape, color, hairiness, quickness, and size," p.723) experienced less habituation during exposure trials, but also displayed less return of fear when presented with a spider at follow-up relative to the "constant-stimulus" group (in which all exposures were conducted with the same spider). Similarly, among individuals with fear of heights (Lang & Craske, 2000), participants in the variable group (who looked over balconies of a tall building in random order; e.g., 5th, 2nd, 10th floor, etc.) reported less anxiety at follow-up (despite higher peak fear during exposure) relative to the gradual exposure group (who systematically progressed up the floors; i.e., 1st, 2nd, 3rd floor, etc.). Two additional studies attempting to experimentally vary fear during exposure failed to replicate these results in individuals fearful of public speaking (Culver, Stoyanova, & Craske, 2012) and contamination (Kircanski et al., 2012). However, in both studies variability in fear responding during exposure was a predictor of superior outcomes irrespective of treatment condition.

Thus, while some findings suggest that learning to tolerate varying levels of fear during exposure enhances long-term outcomes for some anxiety-related problems, replication and extension is needed to clarify these mixed results and investigate their applications beyond more circumscribed fears. In the present study, we examined this question through a preliminary application to exposure therapy for unacceptable obsessional thoughts (i.e., regarding sex, immorality, religiosity, or harm), the presentation of OCD most associated with attenuated outcomes (e.g., Williams et al., 2014). Considering such obsessions often intrude abruptly without identifiable external triggers and are associated with elevated intolerance of uncertainty (e.g., Jacoby, Fabricant, Leonard, Riemann, & Abramowitz, 2013), individuals with these obsessions may especially benefit from interventions that maximize surprise, variability, and uncertainty during exposure.

Accordingly, the current preliminary study compared the processes and short- and long-term outcomes of gradual (EXP-G) and variable (EXP-V) exposure for unacceptable obsessional thoughts. The two interventions were intentionally brief and specific (i.e., targeting one obsessional thought) to maximize internal validity. First, we hypothesized that EXP-V would demonstrate superior outcome relative to EXP-G at 3-month follow-up, but not immediately post-treatment (given that the benefits of variable exposure emerged at later assessments; Lang & Craske, 2000). Second, we predicted that within-session fear variability, but not habituation, would predict changes in obsessional symptoms at post-treatment and follow-up across conditions. Finally, we examined the feasibility of conducting variable exposure with this sample, predicting that EXP-V would demonstrate comparable treatment credibility and satisfaction to EXP-G.

2. Method

2.1. Participants

Participants were 40 adults with an unacceptable obsessional thought pertaining to sex, immorality, religion, or harm (See Table 1) who responded to study advertisements near a university in the southeastern United States from September 2015–May 2016. We used an analogue sample of individuals with at least moderate obsession-related distress, but did not require a diagnosis of OCD given that: (a) 80–99% of the population experiences unwanted thoughts similar in content to clinical obsessions (e.g., Radomsky et al., 2014), and (b) "clinical" and "non-clinical" obsessions are associated with the same developmental and maintenance factors that can be targeted in exposure-based treatments (Abramowitz et al., 2014). A phone screen confirmed the presence of the obsession and ruled out unwanted thoughts related to traumatic events, body image, or generalized worry. The distress item from the Anxiety and Related Disorders Interview Schedule identified thoughts that provoked at least moderate distress (i.e., score ≥ 4; Brown & Barlow, 2014) from 0 (*none*) to 8 (*extreme*).

Participants taking psychiatric medications ($n = 13$) were on a stable dosage for at least 30 days before beginning the study. Exclusion criteria were: (a) previous exposure-based CBT; (b) current suicidal ideation (score > 1 on the suicidality item of the Beck Depression Inventory [BDI-II]; Beck, Steer, & Brown, 1996), current substance use disorder, mania or psychosis (Mini-International Neuropsychiatric Interview [MINI]; Sheehan et al., 1998); (c) current anxiolytic or stimulant medication; (d) heart, respiratory, or neurological condition; or (e) pregnancy. See Fig. 1 for a CONSORT diagram of study enrollment.

2.2. Measures

The **Yale-Brown Obsessive-Compulsive Scale** (Y-BOCS; Goodman, Price, Rasmussen, & Mazure, 1989) is an interview measure of OCD symptoms that includes a symptom checklist and a 10-item severity scale assessing obsessions (items 1–5) and compulsions (6–10) along the following parameters: time, interference, distress, resistance, and control (total scores 0–40). The measure has good reliability, validity, and sensitivity to change following treatment; $\alpha = 0.88$ in the present sample.

The **Dimensional Obsessive-Compulsive Scale–Unacceptable Thoughts** (DOCS-UT; Abramowitz et al., 2010) is a self-report measure

[1] Of note, these concepts are consistent with inhibitory learning models of exposure therapy (Craske et al., 2008).

Table 1
Socio-demographic and baseline clinical characteristics of the sample by EXP group.

	Total sample $N = 40$	EXP-G $n = 19$	EXP-V $n = 21$	Test for difference
Recruitment, % PSYC 101 (*n*)	47.5% (19)	36.8% (7)	57.1% (12)	$\chi^2(1) = 1.64$, $p = .20$, $\varphi = .21$
Age (years), *M* (*SD*)	24.85 (8.68)	26.58 (9.38)	23.29 (7.89)	$t(38) = 1.21$, $p = .24$, $d = .38$
Gender, % female (*n*)	70.0% (28)	63.2% (12)	76.2% (16)	$\chi^2(2) = 0.92$, $p = .63$, $\varphi = .15$
Race, % (*n*) [a]				$\chi^2(1) = 0.01$, $p = .94$, $\varphi = .01$
African American or Black	12.5% (5)	10.5% (2)	14.3% (3)	
White or Caucasian	62.5% (25)	63.2% (12)	61.9% (13)	
Asian	15.0% (6)	10.5% (2)	19.0% (4)	
Biracial or Multiracial	10.0% (4)	15.8 (3)	4.8% (1)	
Ethnicity, % Hispanic or Latino/a (*n*)	7.5% (3)	10.5 (2)	4.8 (1)	
Years of Education, *M* (*SD*)	15.35 (3.41)	16.32 (3.64)	14.48 (3.01)	t (38) = 1.75, $p = .09$, $\varphi = .55$
Y-BOCS clinical cut-off	77.5% (31)	68.4% (13)	85.7% (18)	$\chi^2(1) = 1.71$, $p = .19$, $\varphi = .21$
Primary obsessional thoughts				
Violence, aggression, or harm	67.5% (27)			
Immorality	17.5% (7)			
Sex	7.5% (3)			
Religion	2.5% (1)			
Other	5% (2)			
Compulsions and neutralizing strategies				
Suppression	87.5% (35)			
Distraction	87.5% (35)			
Reassurance-seeking (i.e., from self or others)	75.0% (30)			
Analyzing	75.0% (30)			
Neutralizing (or "canceling out")	45.0% (18)			
Other (e.g., praying, mental reviewing, deep breathing)	25.0% (10)			
Avoiding the thought and things that trigger it	50% (20)			

Note. EXP-G = Gradual Exposure group; EXP-V = Variable Exposure group; Y-BOCS = Yale-Brown Obsessive Compulsive Scale.
[a] Chi-square for participant race compares percentages of white vs. non-white participants due to small sample size of certain racial and ethnic groups (*ns* < 5).

of clinical severity of the unacceptable thoughts (UT) OCD symptom dimension. Respondents rate symptoms on time occupied, avoidance behaviors, distress, functional interference, and difficulty disregarding obsessions and resisting compulsive urges (total scores 0–20). The DOCS has good reliability and converges well with other OCD symptom measures; $\alpha = 0.86$ in the present sample.

We adapted the **Behavioral Approach Test** (BAT) developed by Steketee, Chambless, Tran, and Worden (1996) as an *in vivo* measure of participants' responses to their target obsession. Participants were instructed to complete as many of the following tasks as they could: (a) State the thought aloud, (b) Write the thought on a piece of paper, (c) Write the thought on your hand in ink, (d) Say "I will never know for certain whether one day [the thought] might come true," and (e) Say "I hope and pray [the thought] comes true." Participants informed the experimenter if there were any tasks they did not wish to complete because they were too difficult. Participants also reported after each step if they performed a neutralizing ritual from a list of options (e.g., distraction), and the number of steps participants completed without ritualizing was calculated (0–5).

Fear level was monitored using the **Subjective Units of Distress Scale** (SUDS; Wolpe, 1973) from 0 (*no distress/fear/anxiety*) to 100 (*extreme distress/fear/anxiety*). Participants provided SUDS ratings after completing each BAT step, as well as every minute during exposures. From these ratings, we calculated average SUDS levels across the five BAT tasks; $\alpha = 0.93$ in the present sample. **Skin Conductance Level** (SCL; via two electrodermal activity [EDA] silver-silver chloride [Ag/AgCl] adhesive electrodes [11 mm diameter contact] filled with electrolyte gel [0.5% chloride salt] placed on the medial segments of the index and middle fingers of the participant's non-dominant hand) and **Heart Rate** (HR; via three electrocardiogram [ECG] silver-silver chloride [Ag/AgCl] adhesive electrodes [11 mm diameter contact] filled with electrolyte gel [7% chloride salt] placed on the participant's torso in the standard Right Arm - Left Arm - Left Leg [RA-LA-LL] configuration) were measured using the BioPac MP150 as objective physiological measures of fearful responding (i.e., arousal) during the BAT and exposure sessions.

The **Credibility/Expectancy Questionnaire** (CEQ; Devilly & Borkovec, 2000) is a 6-item measure with two subscales (3 items each): (1) treatment rationale credibility (CEQ-C) is cognitively based (e.g., "How logical does the therapy offered to you seem?"), and (2) treatment expectancy (CEQ-E) is emotionally based (e.g., "How much do you really *feel* that therapy will help you to reduce your symptoms?"). The CEQ demonstrates adequate internal consistency and test-retest reliability (Devilly & Borkovec, 2000); $\alpha = 0.73$-0.79 in the present sample.

The **Client Satisfaction Questionnaire** (CSQ; Nguyen, Attkisson, & Stegner, 1983) is a widely used 8-item measure of treatment satisfaction (total scores 8–32). The CSQ demonstrates high internal consistency, as well as criterion and construct validity (Nguyen et al., 1983). It has been used in studies examining patient satisfaction with exposure therapy (Tolin, Diefenbach, & Gilliam, 2011); $\alpha = 0.92$ in the present sample.

2.3. Procedure

Overview. Participants attended four twice-weekly intervention sessions and two follow-up visits: (a) baseline assessment (PRE, 30 min) followed by psychoeducation and exposure list generation (Session 1, 30 min); (b) three exposure sessions (Sessions 2–4, 45–60 min each), with the last exposure session followed by post-assessment (POST, 30 min); and (c) two follow-up assessments 1- and 3-months after post (1MFU and 3MFU, 30 min each). On the day of appointments, participants were asked to refrain from using the following due to their effects on physiological measures: (a) alcohol, recreational drugs, anti-anxiety, sleep, or stimulant medications, and (b) caffeinated or tobacco products two hours prior to the session.

Session 1. Following informed consent, an independent evaluator (IE) blind to experimental condition administered the MINI and BDI-II to determine eligibility, as well as the Y-BOCS, DOCS-UT, and BAT. Participants were then randomly assigned in parallel (1:1; via a computerized random number generator with a concealed allocation sequence) to either gradual (EXP-G; $n = 19$) or variable (EXP-V; $n = 21$) exposure. Next, a doctoral-level experimenter provided psychoeducation about the cognitive-behavioral conceptualization of obsessions and the rationale for EXP from one of two perspectives. In EXP-G, the rationale emphasized the goal of fear reduction via habituation (e.g., 50% reduction in SUDS). In EXP-V, the rationale emphasized exposure to

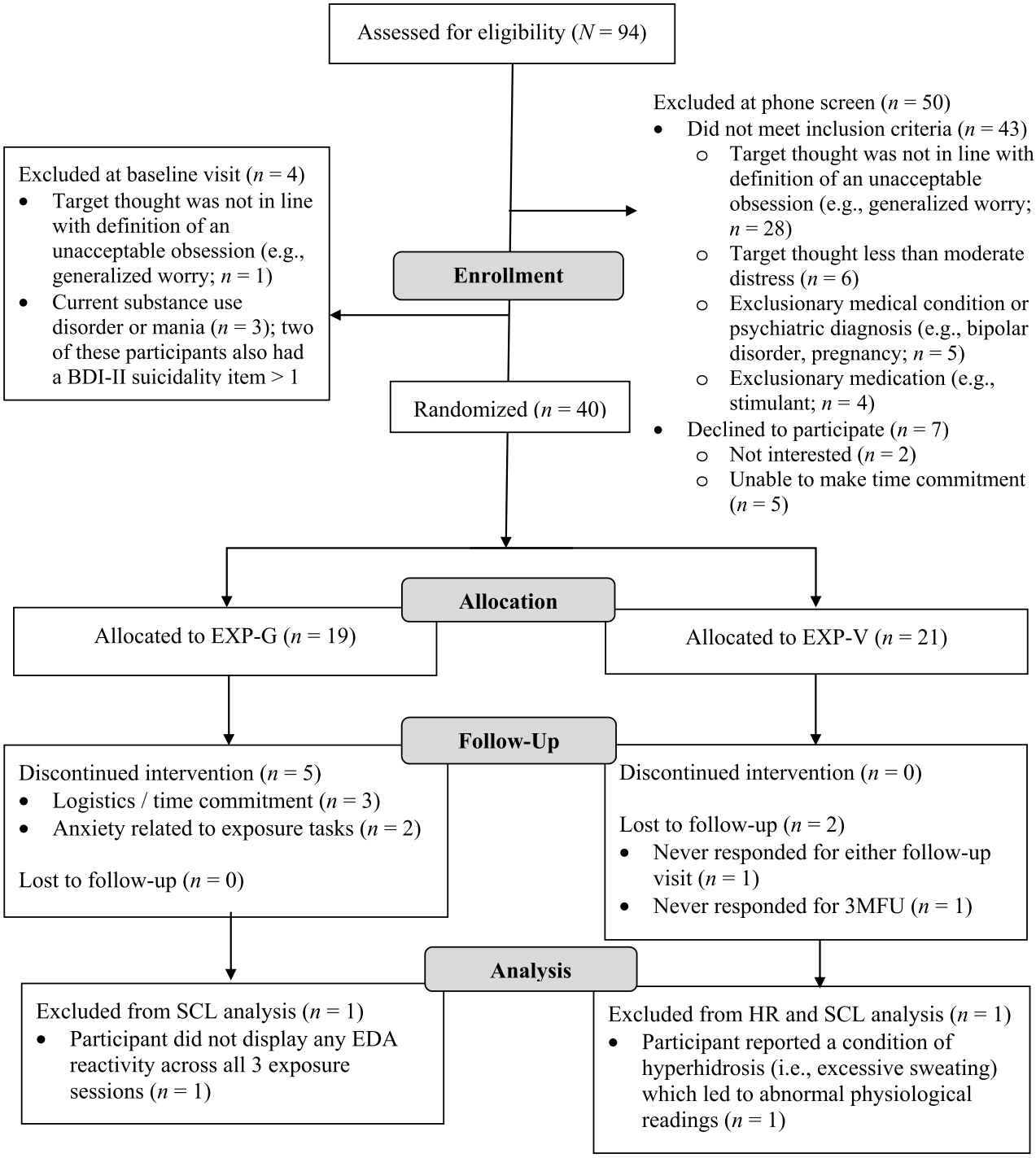

Fig. 1. CONSORT flow diagram.

varying levels of fear to practice fear tolerance and maximize surprise. Both rationales emphasized the importance of refraining from maladaptive safety behaviors (e.g., distraction, thought suppression) which interfere with the learning process. The experimenter and participant collaboratively generated a list of nine exposure stimuli, categorized by the intensity of distress they were anticipated to provoke (three each mild, moderate, and high intensity).

Sessions 2–4. After attaching the EDA/ECG electrodes, participants read a neutral reading comprehension passage (i.e., about science discoveries; 8th grade reading level) for three minutes while physiological data (ECG, GSR) were collected to determine baseline physiological responding. SUDS were also recorded every minute. Sessions then included three experimenter-guided exposures lasting eight minutes each (*SD* = 1 min). In EXP-G, exposure proceeded hierarchically from mildly, to moderately, to highly intense stimuli so that intensity gradually built between (but not within) sessions. In EXP-V, the exposure stimuli were chosen at random so that mild, moderate, and high intensity exposures could occur in any order. This aimed to maximize uncertainty, variability in exposure intensity, and variability in corresponding physiological arousal. In both conditions, the experimenter helped the participant refrain from anxiety-reducing rituals by periodically inquiring if they were using any of the previously identified safety behaviors to remove or control the thought. Session 4 also included a discussion of relapse prevention.

Post and follow-up visits. The POST, 1MFU, and 3MFU assessment visits were conducted by an IE who remained blind to EXP condition and included the administration of the previously described measures. After the final visit, participants were debriefed and provided with referrals.

2.4. Data analytic strategy

Power. With a sample of 40 participants we have 80% power to detect significant group differences at follow-up between gradual vs. variable exposure at the .05 level if the true effect size is 0.91 or greater, which is comparable to previous studies reporting effects of 0.94 (e.g., Lang & Craske, 2000).

Data Reduction. The following calculations were made consistent with previous research (e.g., Kircanski et al., 2012) to summarize the fear indices: (a) **Baseline Fear**: average SUDS, SCL, and HR during the 3 min baseline (before the first EXP task was described); (b) **Mean Fear**: average SUDS, SCL, and HR, (c) **Within-session habituation** (**WSH**): peak fear level (i.e., highest SUDS, SCL and HR for each EXP session) minus final fear level (i.e., final SUDS rating and mean SCL and HR for the final minute), and (d) **Fear variability**: intra-individual standard deviation (**ISD**) of fear (as measured by SUDS, SCL, and HR). Each of these values was calculated within session and then averaged across the 3 EXP sessions.

Multi-level Models. Analyses were intention-to-treat and involved all participants who were randomly assigned. Due to the nested structure of the data, we used multi-level modeling (MLM) with compound symmetric error structures. Model predictors included: (a) experimental condition (Level 2; EXP-G vs. EXP-V; dummy coded 0 and 1 respectively, such that a main effect of condition suggests an effect of EXP-V above and beyond EXP-G), (b) EXP indices of habituation and fear variability (Level 2; WSH and ISD averaged across the 3 EXP sessions), and (c) session number (Level 1; nominally coded; i.e., making no assumptions about the time spacing between sessions). Outcome measures included: (a) interview (i.e., Y-BOCS), (b) self-report (i.e., DOCS-UT), and (c) behavioral (i.e., mean SUDS; number of BAT steps completed without ritualizing) indices.

3. Results

3.1. Baseline characteristics and group comparisons

Clinical characteristics of unacceptable thoughts. The most common obsessional themes (See Table 1) concerned violence, aggression, or harm (e.g., image of stabbing a loved one), followed by thoughts of immoral behavior (e.g., urge to yell derogatory comments). The majority of the sample was above the clinical cut-off of 16 on the Y-BOCS at PRE (Goodman et al., 1989). Compulsions and neutralizing strategies reported also appear in Table 1.

Examination of pre-treatment clinical measures. As seen in Table 1, randomization produced two comparable groups without any statistically significant differences on our sociodemographic variables (i.e., differences were small in magnitude). Group means for study measures at all assessment points appear in Table 2. PRE Y-BOCS and DOCS-UT scores indicated moderate severity of obsessions and compulsions. On average, participants experienced mild levels of depression and moderate distress during the BAT. Independent samples *t*-tests revealed no significant group differences at PRE on any measure (all $ps > .20$, $ds < -0.49$), thus ensuring that randomization successfully produced two equitable groups.

Group comparisons of treatment credibility/expectancy, satisfaction, and drop-out. Group means for the CEQ and CSQ also appear in Table 2. There were no differences on participant-rated credibility of the two exposure rationales, $t(37) = 0.06$, $p = .95$, $d = 0.02$ [95% CI: −0.61, 0.65], or their expectancies of improvement, $t(37) = 0.63$, $p = .53$, $d = 0.20$ [95% CI: −0.43, 0.83]. Participants thought the treatment rationales were "somewhat" to "very" logical and expected to see 50% reduction in symptoms. Similarly, there were no differences in how satisfied participants were with the intervention they received, $t(33) = -0.79$, $p = .43$, $d = -0.27$ [95% CI: −0.95, 0.41]. Satisfaction levels approached those of individuals receiving ERP in a clinical setting (*Ms* = 29.00–29.92; Tolin et al., 2011).

Five participants (13.3%) dropped out prematurely from the study and all were in the EXP-G condition. Specifically, 1 dropped after the first session (i.e., before beginning any exposures) and the other 3 dropped after Session 2 (i.e., between the first and second exposure sessions). Each stated a logistical reason for dropping (e.g., time commitment) and 2 participants mentioned that the tasks were "anxiety-provoking" as a reason for discontinuing.

Exposure Process Characteristics. There were no group differences in the duration of exposure exercises: EXP-G: 8.31 min (*SD* = 1.21) vs. EXP-V: 7.98 min (*SD* = 0.70); $t(37) = 1.04$, $p = .31$, $d = 0.34$, 95% CI [-0.30, 0.97]. As a manipulation check of variability, we calculated the number of times exposure intensity changed for the two groups. For participants in the EXP-G group, exposure intensity increased twice (from mild to moderate and from moderate to high intensity). For participants in the EXP-V group, exposure intensity changed 6.05 times on average (*SD* = 1.28; range = 4–8), which was significantly higher; $t(20) = -14.45$, $p < .001$, $d = -4.30$, 95% CI [-5.34, −3.09].

3.2. Multi-level models

Level 1 and Level 2 residuals were normally distributed and homoscedastic and there were no problematic trends in the data. Effect size calculations include: (a) Standardized Mean Gain (ESsg) for paired samples to compute PRE-POST and PRE-3MFU effect sizes, and (b) Cohen's *d* calculations to compute between groups effect sizes at POST and 3MFU (i.e., for EXP-V compared to EXP-G). Figs. 2–5 graphically depict changes in outcomes over time.[2]

Overall Exposure Outcomes. As an initial test of EXP efficacy, we examined whether EXP conditions combined produced significant immediate (PRE-POST) and longer-term (POST-3MFU) reductions in obsessional symptoms (using planned comparisons to test whether changes were significantly different from zero). Simple contrasts indicated that there were significant changes from PRE-POST, with large effect sizes for all four outcome measures in the expected directions (See Table 3).[3] Additionally, there continued to be significant decreases from POST-3MFU for Y-BOCS and DOCS-UT (but not for BAT mean SUDS levels or BAT steps completed without ritualizing).

Hypothesis 1. To test our first hypothesis, we used planned comparisons to examine whether experimental condition (EXP-G vs. EXP-V) predicted changes in each of the four outcome measures from PRE-POST and POST-3MFU (i.e., two-way cross-level EXP condition × time interactions). First, simple contrasts were not significant when comparing the difference between the EXP-G and EXP-V changes from PRE-POST for the four outcome measures (See Table 3). From POST-3MFU, the simple contrast approached significance ($p = .056$) when comparing the difference between the EXP-G and EXP-V changes in Y-BOCS score (model estimated difference in changes = 4.44 points).

[2] MLM analyses include both 1MFU and 3MFU data in the models. For the sake of parsimony, interpretation of results focuses on the long-term (i.e., 3-month) follow-up outcomes. There were no significant changes in scores from POST to 1MFU except that DOCS-UT scores significantly decreased for the EXP-V group only ($p = .02$).

[3] Significant decreases in Y-BOCS, DOCS-UT, and BAT mean SUDS levels, and significant increases in BAT number of steps completed without ritualizing (indicating improvements in treatment).

Table 2
Descriptive statistics for clinical measures by EXP group and time point.

	EXP-G Group				EXP-V Group			
	PRE n = 19	POST n = 14	1MFU n = 13	3MFU n = 14	PRE n = 21	POST n = 21	1MFU n = 20	3MFU n = 19
Y-BOCS	18.53 (6.47)	10.71 (6.46)	12.92 (5.56)	11.50 (6.96)	18.29 (3.27)	11.38 (5.78)	9.00 (5.73)	7.79 (5.06)
DOCS-UT	7.21 (3.26)	4.93 (3.50)	4.15 (2.23)	4.00 (2.45)	7.19 (2.38)	3.71 (2.41)	2.45 (2.42)	2.42 (2.59)
BAT								
SUDS	39.55 (18.98)	14.90 (11.86)	14.44 (13.83)	12.07 (11.43)	48.12 (21.42)	21.12 (16.36)	20.38 (18.32)	16.09 (13.68)
Steps completed without ritualizing	2.53 (1.54)	3.71 (1.64)	3.33 (1.61)	4.00 (1.24)	2.33 (1.43)	3.81 (1.21)	3.70 (1.63)	3.26 (1.56)
CEQ								
Credibility	18.37 (3.80)				18.30 (3.54)			
Expectancy[a]	0.29 (3.35)				−0.28 (2.16)			
CSQ		25.07 (4.01)				26.29 (4.70)		

Note. EXP-G = Gradual Exposure group; EXP-V = Variable Exposure group; Y-BOCS = Yale-Brown Obsessive Compulsive Scale; DOCS-UT = Dimensional Obsessive-Compulsive Scale–Unacceptable Thoughts; BAT = Behavioral Approach Test; SUDS = Subjective Units of Distress Scale; CEQ = Credibility/Expectancy Questionnaire; CSQ = Client Satisfaction Questionnaire.

[a] CEQ-E scores are calculated by standardizing the items to z-scores before summing to create an expectancy score.

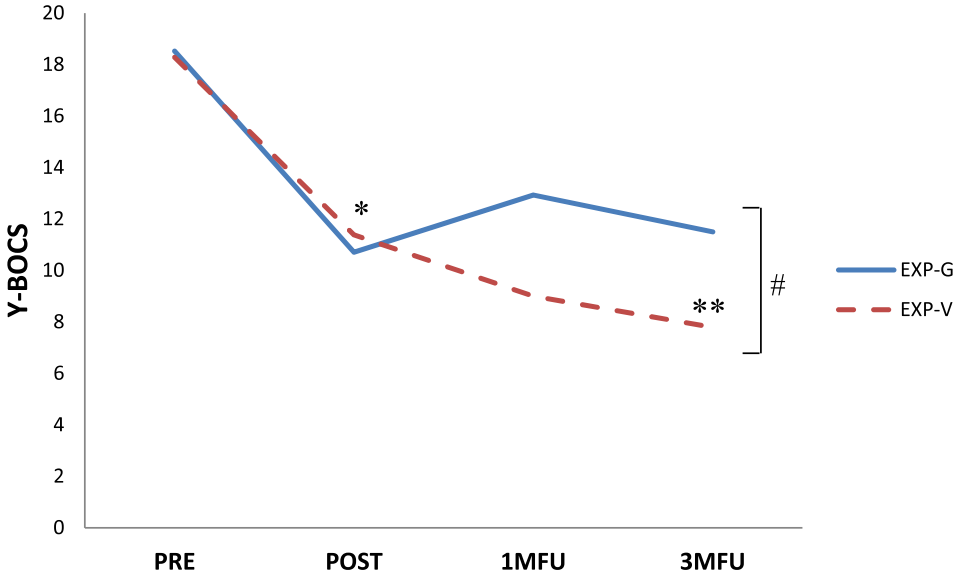

Fig. 2. Changes in Y-BOCS scores by EXP group.

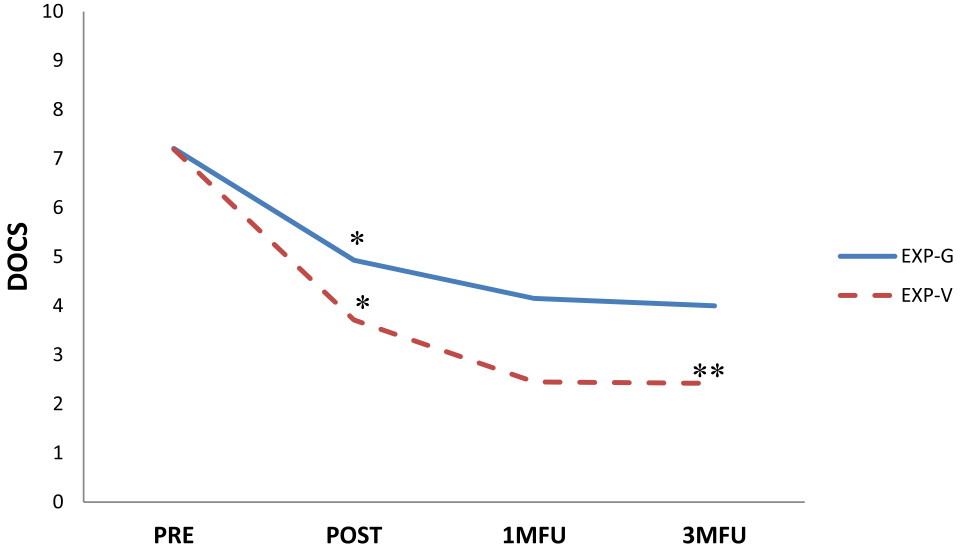

Fig. 3. Changes in DOCS-UT scores by EXP group.

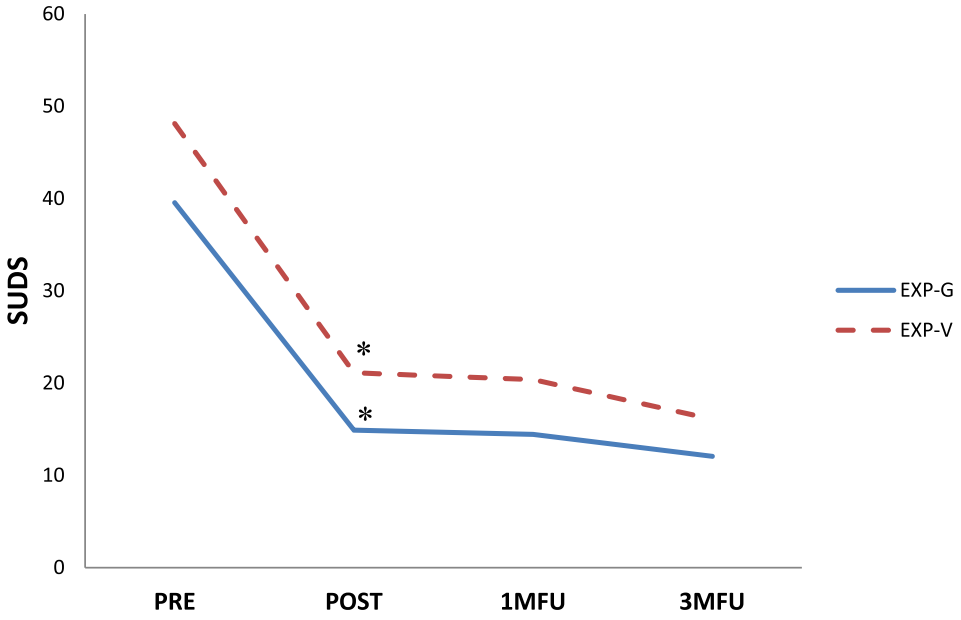

Fig. 4. Changes in BAT Mean SUDS by EXP group.

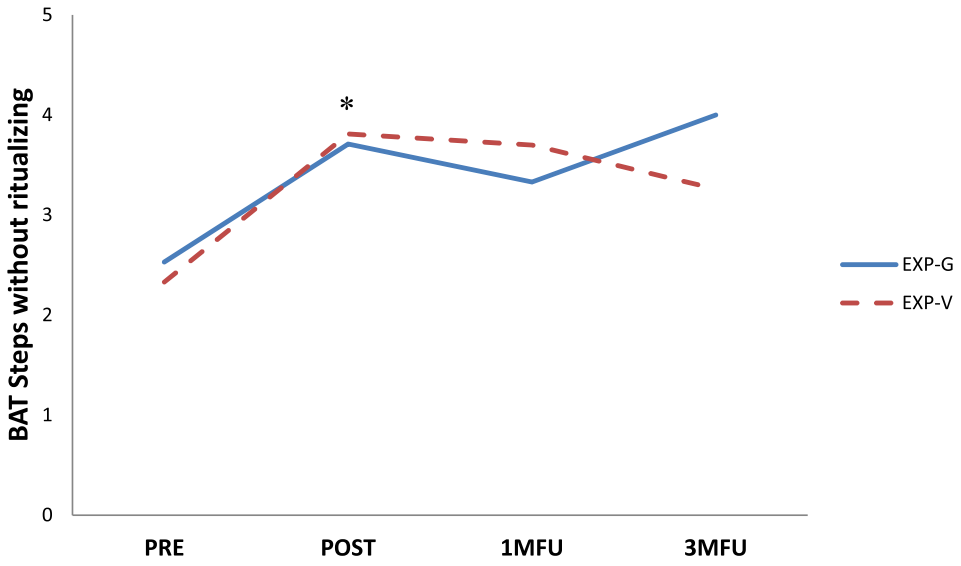

Fig. 5. Changes in BAT Number of Steps Completed without Ritualizing by EXP group.

Specifically, although Y-BOCS scores did not change significantly from POST-3MFU for the EXP-G group, $t(93) = 0.45$, $p = .66$, ESsg = −0.12, 95% CI [-0.80, 0.57], scores did decrease significantly for the EXP-V group, $t(93) = -2.47$, $p = .015$, ESsg = 0.69, 95% CI [0.12, 1.27]. Similarly, although DOCS-UT scores did not change significantly from POST-3MFU for the EXP-G group, $t(93) = -1.44$, $p = .15$, ESsg = 0.30, 95% CI [-0.23, 0.83], scores did decrease significantly for the EXP-V group, $t(93) = -2.32$, $p = .02$, ESsg = 0.51, 95% CI [0.10, 0.93]. Simple contrasts were not significant, however, when comparing the difference between EXP-G and EXP-V changes from POST-3MFU for the DOCS-UT or the final two outcomes (i.e., BAT mean SUDS and BAT steps completed without ritualizing).

Clinically significant change. We used the Jacobson and Truax method (1991) to determine the number of participants in each group that (a) achieved end-state functioning within the non-clinical distribution of Y-BOCS scores (Frost, Steketee, Krause, & Trepanier, 1995); i.e., below the cutoff score of 13.5 using Method c (Jacobson & Truax, 1991, p. 13), and (b) demonstrated reliable change of 6.33 points (p. 14). All dropouts were scored as unimproved and included in a full intent-to-treat analysis.[4] The number of patients in EXP-G vs. EXP-V that attained clinically significant and reliable change was not significantly different at post-treatment, but significantly more participants in the EXP-V group met this criterion with a moderate effect size at follow-up (See Table 4).

Hypothesis 2. To test our second hypothesis, we fit models including: (a) session number, (b) pre-treatment scores and baseline fear measures (i.e., SUDS, SCL, or HR during the 3-min baseline) as

[4] If participants who dropped from treatment or were lost to follow-up were not included in analyses ($n = 33$), at posttreatment, 6 (42.9%) patients in the EXP-G condition and 8 (42.1%) in the EXP-V condition attained both clinically significant and reliable change; these proportions were not significantly different, $\chi^2(1) = 0.002$, $p = .97$, $\varphi = -0.01$. At follow-up, this status was attained by 7 (50%) patients in the EXP-G condition and 15 (78.9%) in the EXP-V condition; the proportions were not significantly different, $\chi^2(1) = 3.04$, $p = .08$, $\varphi = 0.30$.

Table 3
Multi-level model planned comparisons and effect sizes.

	Both EXP Groups Combined				EXP-G vs. EXP-V			
	PRE-POST *F*-statistic	PRE-POST ESsg [95% CI]	POST-3MFU *F*-statistic	PRE-3MFU ESsg [95% CI]	PRE-POST *t*-statistic	POST *d* [95% CI]	POST-3MFU *t*-statistic	3MFU *d* [95% CI]
Y-BOCS	22.97**	1.35 [0.87, 1.83]	3.16*	1.63 [1.01, 2.25]	−0.48	−0.11 [−0.78, 0.57]	1.94	0.63 [−0.10, 1.32]
DOCS-UT	30.34**	1.11 [0.71, 1.51]	3.74*	1.52 [1.04, 2.00]	1.10	0.42 [−0.27, 1.10]	0.40	0.62 [−0.10, 1.31]
BAT SUDS	40.95**	1.26 [0.82, 1.71]	1.31	1.56 [0.98, 2.15]	0.89	−0.42 [−1.10, 0.27]	0.47	−0.31 [−1.01, 0.40]
BAT Steps completed without ritualizing	16.76**	−0.91 [−1.27, −0.54]	1.58	−0.75 [−1.14, −0.35]	−0.75	−0.07 [−0.75, 0.61]	1.62	0.51 [−0.22, 1.22]

Note: *$p < .05$; **$p < .001$. EXP-G = Gradual Exposure group; EXP-V = Variable Exposure group; Y-BOCS = Yale-Brown Obsessive Compulsive Scale; DOCS-UT = Dimensional Obsessive-Compulsive Scale–Unacceptable Thoughts; BAT = Behavioral Approach Test; SUDS = Subjective Units of Distress Scale. Effect size calculations include: (a) Standardized Mean Gain (ESsg) for paired samples to compute PRE-POST and PRE-3MFU effect sizes, and (b) Cohen's *d* calculations to compute between groups effect sizes at POST and 3MFU (i.e., for EXP-V compared to EXP-G).

Table 4
Clinically significant and reliable change using primary outcome measure (Y-BOCS).

	EXP-G	EXP-V	Test for difference
Post-treatment	31.6% (6)	47.6% (10)	$\chi^2(1) = 1.07$, $p = .301$, $\varphi = .17$
3-month follow-up	36.8% (7)	81.0% (17)	$\chi^2(1) = 8.09$, $p = .004$, $\varphi = .45$

Note: EXP-G = Gradual Exposure group; EXP-V = Variable Exposure group; Y-BOCS = Yale-Brown Obsessive Compulsive Scale.

Table 5
Subjective and psychophysiological variability and within-session habituation as predictors of outcome.

	t	*p*	*d*
DV: Y-BOCS			
Model 1: Covariates: PRE Y-BOCS, baseline SUDS			
ISD SUDS	1.16	0.25	0.24
WSH SUDS	−1.31	0.19	−0.27
Model 2: Covariates: PRE Y-BOCS, baseline SCL			
ISD SCL	−1.33	0.19	−0.28
WSH SCL	−0.61	0.54	−0.13
Model 3: Covariates: PRE Y-BOCS, baseline HR			
ISD HR	−0.40	0.69	−0.08
WSH HR	−0.25	0.80	−0.05
DV: DOCS-UT			
Model 1: Covariates: PRE DOCS-UT, baseline SUDS			
ISD SUDS	0.47	0.64	0.10
WSH SUDS	−1.10	0.27	−0.23
Model 2: Covariates: PRE DOCS-UT, baseline SCL			
ISD SCL	−1.08	0.28	−0.23
WSH SCL	0.28	0.78	0.06
Model 3: Covariates: PRE DOCS-UT, baseline HR			
ISD HR	−0.18	0.86	−0.03
WSH HR	0.82	0.42	0.17
DV: BAT SUDS			
Model 1: Covariates: PRE BAT SUDS, baseline SUDS			
ISD SUDS	2.27	0.03*	0.47
WSH SUDS	−1.46	0.15	−0.30
Model 2: Covariates: PRE BAT SUDS, baseline SCL			
ISD SCL	−0.12	0.91	−0.03
WSH SCL	−0.44	0.66	−0.09
Model 3: Covariates: PRE BAT SUDS, baseline HR			
ISD HR	0.16	0.88	0.03
WSH HR	−0.63	0.53	−0.13
DV: BAT steps completed without ritualizing			
Model 1: Covariates: PRE BAT steps completed without ritualizing, baseline SUDS			
ISD SUDS	−1.30	0.20	−0.27
WSH SUDS	0.59	0.56	0.12
Model 2: Covariates: PRE BAT steps completed without ritualizing, baseline SCL			
ISD SCL	3.18	0.002*	0.68
WSH SCL	−0.29	0.78	−0.06
Model 3: Covariates: PRE BAT steps completed without ritualizing, baseline HR			
ISD HR	0.21	0.83	0.04
WSH HR	1.56	0.12	0.33

Note: *$p < .05$; DV = Dependent Variable; Y-BOCS = Yale-Brown Obsessive Compulsive Scale; ISD = Intra-individual Standard Deviation; WSH = Within-session Habituation; DOCS-UT = Dimensional Obsessive-Compulsive Scale–Unacceptable Thoughts; BAT = Behavioral Approach Test; SUDS = Subjective Units of Distress Scale; SCL = Skin Conductance Level; HR = Heart Rate.

covariates, and (c) fear variability (ISD) and within-session habituation (WSH) as time-invariant predictors (i.e. averaged across EXP sessions). A separate model was tested for each of the three fear measurements (i.e., SUDS, SCL, and HR) for each of the four outcome variables.

As displayed in Table 5, after controlling for PRE Y-BOCS scores and baseline SUDS, ISD and WSH SUDS did not significantly-predict Y-BOCS scores. Similarly, after controlling for PRE Y-BOCS scores and baseline SCL, ISD and WSH SCL did not significantly-predict Y-BOCS scores.

Finally, after controlling for PRE Y-BOCS scores and baseline HR, ISD and WSH HR did not significantly-predict Y-BOCS scores. Null results of models with the other three outcomes (i.e., DOCS-UT, BAT mean SUDS, BAT steps completed without ritualizing) parallel those of the Y-BOCS, and thus numeric results are summarized in Table 5.[5]

4. Discussion

In the present study, both gradual and variable exposure led to significant reductions in obsessional symptoms with large pre-to post-treatment effect sizes that were maintained (or further increased) at 3-month follow-up. These results are consistent with meta-analytic findings demonstrating the efficacy of ERP (Olatunji, Davis, Powers, & Smits, 2013). Although the percentage of participants who achieved clinically significant change was low, the intervention was designed to be brief and specific (i.e., 4 sessions) versus the typical 16-session dose of ERP.

For our first hypothesis, we found no statistically significant differences in the degree of improvement between the two conditions. These results are in line with null findings from past studies comparing gradual versus variable exposure for contamination fears (Kircanski et al., 2012). Thus, exposure therapy might be highly effective regardless of the order in which stimuli are presented, and ceiling effects limit the detection of differential outcomes. Of note, however, non-significant trends favoring variable exposure at 3-month follow-up (as hypothesized) were observed for interview and self-report measures of obsessional symptoms. Moreover, effect sizes comparing EXP-G vs. EXP-V differences at 3-month follow-up (while not statistically significant) were moderate in magnitude. Additionally, significantly more participants in EXP-V achieved clinically significant and reliable change at 3-month follow-up relative to EXP-G. Thus, these preliminary findings indicate that future study of how to enhance exposure with variability may be warranted, as discussed further below.

For our second hypothesis, results were consistent with previous findings that within-session habituation is not associated with clinical outcomes (e.g., Baker et al., 2010; Meuret, Seidel, Rosenfield, Hofmann, & Rosenfield, 2012) and bolster recommendations against relying solely on fear reduction as an indicator of exposure outcome. Contrary to theories that variability in subjective and physiological levels of fear may optimize extinction learning, findings from previous studies (e.g., Kircanski et al., 2012), and our second hypothesis, however, no measure of fear variability during exposure consistently predicted outcome (and effect sizes were small in magnitude).

Finally, participants in the two conditions expressed no differences in the credibility of the treatment rationales (i.e., fear reduction vs. tolerance), or in satisfaction with the intervention they received. Previous studies (Arch, Twohig, Deacon, Landy, & Bluett, 2015) have also demonstrated that participants perceive empirically-derived rationales for exposure to be similarly credible. These findings, therefore, provide preliminary evidence contrary to clinician concerns that variable exposures would seem unappealing or even intolerable, at least in comparable samples with moderate distress.

Although this study had several strengths such as a longitudinal design and multi-method assessment, some caveats deserve discussion. First, unlike more circumscribed fears/phobias in which identifying low, medium, and high intensity exposure items is more clear, obsessions are thematically heterogeneous and thus there are many ways that variability can be induced. While this is a strength for studies seeking to maximize variability during ERP, this also is a challenge since participants may not always anticipate which exposures will generate low, moderate, and high fear levels.

[5] As can be seen, only two predictors emerged as significant after controlling for the aforementioned variables in the model: (1) mean ISD SUDS was a significant predictor of BAT mean SUDS, and (2) mean ISD SCL was a significant predictor of BAT mean number of steps completed without ritualizing.

Second, although analogue studies are useful for testing the application of fundamental scientific principles (Vervliet, Craske, & Hermans, 2013), the extent to which our findings generalize to clinical populations is unknown. Third, the current study did not assign homework exercises between sessions (as is traditionally done in clinical settings), given that homework compliance is predictive of treatment outcome (e.g., Abramowitz, Franklin, Zoellner, & DiBernardo, 2002), and we did not want to introduce a confounding factor that might have better accounted for the results. However, future investigations could examine whether encouraging participants to incorporate variability into their exposure practice between sessions in different environmental contexts would demonstrate more robust effects.

Fourth, participants in the current study had not previously received CBT for anxiety (to simplify interpretation, as a null result could emerge either due to an ineffective treatment or due to the difficulty of treating refractory patients). However, it would be worth examining methods for optimizing exposure therapy among individuals that have previously failed standard CBT. Fifth, while we were sufficiently powered to detect large effects, more modest effects could not have been detected with our sample size. Thus, future research with larger samples would be warranted to determine whether variable exposure might optimize treatment for OCD (e.g., 64 per group to detect moderate effects 0.50 or greater with 80% power). Finally, research should extend the assessment period to determine whether outcomes might continue to diverge at a later time point.

A question for future investigation is whether gradual vs. variable exposure influences beliefs about one's self-efficacy to approach feared stimuli. All five participants who withdrew from this study received gradual exposure, two cited high levels of anxiety as a reason for discontinuing, and four dropped out between the low and moderate intensity sessions. Research on whether gradual exposure perpetuates “fear of fear” could use ecological momentary assessment to elucidate the relation between pre-session anticipatory anxiety and session attendance. Additionally, there may be important moderators of outcome not considered in the current study. For example, individuals with high levels of intolerance of uncertainty might particularly benefit from interventions that maximize variability and uncertainty during exposure. Consistent with precision medicine initiatives, research on such moderators could help therapists “prescribe” gradual vs. variable exposure depending on patients' pre-treatment levels.

Finally, the current study indicates that clinicians may have more flexibility in the order of exposure tasks during ERP (versus needing to follow a hierarchy). One potential adaptation of the methods in the present study to clinical settings would be for patient and therapist to agree on a subset of exposure items to be randomly selected if the patient were not yet willing to confront all items on the list. Furthermore, exposures could be selected based on life interference (i.e., values), which is consistent with Acceptance and Commitment Therapy (Twohig et al., 2018). Conducting exposures in this manner allows patient values to drive treatment, introduces desirable difficulties, and does not overrely on habituation of anxiety.

Disclosure statement

The authors declare no conflict of interest.

Clinical trials registration

This study was registered on ClinicalTrials.gov [number NCT03416504]. Since this was a preliminary pilot study with an analogue sample, we did not pre-register this study.

Funding source declaration

This work was supported by the Association for Behavioral and Cognitive Therapies (ABCT), the American Psychological Association (APA), and the University of North Carolina at Chapel Hill. These

funding sources were not involved in the original study design; in the collection, analysis or interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

Protection of human subjects

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Acknowledgements

We would like to extend our sincerest thanks to: Siobhan Hsu, Dana Mansfield, Mackenzie Brown, and Nina Notman for their assistance with data collection; Drs. Donald Baucom, David Penn, Stacey Daughters, and Daniel Bauer for their valuable input on the study design and an earlier version of this manuscript; Christopher Wiesen for his statistical consultation; and the participants who volunteered for this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbtep.2019.02.008.

References

Abramowitz, J. S., & Arch, J. J. (2014). Strategies for improving long-term outcomes in cognitive behavioral therapy for obsessive-compulsive disorder: Insights from learning theory. *Cognitive and Behavioral Practice, 21*, 20–31. https://doi.org/10.1016/j.cbpra.2013.06.004.

Abramowitz, J. S., Deacon, B. J., Olatunji, B. O., Wheaton, M. G., Berman, N. C., Losardo, D., ... Hale, L. R. (2010). Assessment of obsessive-compulsive symptom dimensions: Development and evaluation of the Dimensional Obsessive-Compulsive Scale. *Psychological Assessment, 22*(1), 180–198.

Abramowitz, J. S., Fabricant, L. E., Taylor, S., Deacon, B. J., McKay, D., & Storch, E. A. (2014). The utility of analogue studies for understanding obsessions and compulsions. *Clinical Psychology Review, 34*(3), 206–217. https://doi.org/10.1016/j.cpr.2014.01.004.

Abramowitz, J. S., Franklin, M. E., Zoellner, L. A., & DiBernardo, C. L. (2002). Treatment compliance and outcome in obsessive-compulsive disorder. *Behavior Modification, 26*(4), 447–463. https://doi.org/10.1177/0145445502026004001.

Arch, J. J., Twohig, M. P., Deacon, B. J., Landy, L. N., & Bluett, E. J. (2015). The credibility of exposure therapy: Does the theoretical rationale matter? *Behaviour Research and Therapy, 72*, 81–92. https://doi.org/10.1016/j.brat.2015.05.008.

Baker, A., Mystkowski, J., Culver, N., Yi, R., Mortazavi, A., & Craske, M. G. (2010). Does habituation matter? Emotional processing theory and exposure therapy for acrophobia. *Behaviour Research and Therapy, 48*(11), 1139–1143. https://doi.org/10.1016/j.brat.2010.07.009.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Bjork, R. A., & Bjork, E. L. (2006). Optimizing treatment and instruction: Implications of a new theory of disuse. In L.-G. Nilsson, & N. Ohta (Eds.). *Memory and society: Psychological perspectives* (pp. 116–140). New York, NY US: Psychology Press.

Bouton, M. E. (2000). A learning theory perspective on lapse, relapse, and the maintenance of behavior change. *Health Psychology, 19*(1), 57–63. https://doi.org/10.1037/0278-6133.19.Suppl1.57.

Brown, T. A., & Barlow, D. H. (2014). *Anxiety and Related Disorders Interview Schedule for DSM-5* (1st ed.). Oxford ; New York: Oxford University Press.

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy, 46*(1), 5–27. https://doi.org/10.1016/j.brat.2007.10.003.

Craske, M. G., Liao, B., Brown, L., & Vervliet, B. (2012). Role of inhibition in exposure therapy. *Journal of Experimental Psychopathology, 3*(3), 322–345. https://doi.org/10.5127/jep.026511.

Culver, N. C., Stoyanova, M., & Craske, M. G. (2012). Emotional variability and sustained arousal during exposure. *Journal of Behavior Therapy and Experimental Psychiatry, 43*(2), 787–793. https://doi.org/10.1016/j.jbtep.2011.10.009.

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the Credibility/Expectancy Questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry, 31*(2), 73–86.

Eddy, K. T., Dutra, L., Bradley, R., & Westen, D. (2004). A multidimensional meta-analysis of psychotherapy and pharmacotherapy for obsessive-compulsive disorder. *Clinical Psychology Review, 24*(8), 1011–1030.

Eisen, J. L., Sibrava, N. J., Boisseau, C. L., Mancebo, M. C., Stout, R. L., Pinto, A., et al. (2013). Five-year course of obsessive-compulsive disorder: Predictors of remission and relapse. *Journal of Clinical Psychiatry, 74*(3), 233–239. https://doi.org/10.4088/JCP.12m07657.

Foa, E. B., Huppert, J. D., & Cahill, S. P. (2006). Emotional Processing Theory: An update. In B. O. Rothbaum (Ed.). *Pathological anxiety: Emotional processing in etiology and treatment* (pp. 3–24). New York, NY, US: Guilford Press.

Frost, R. O., Steketee, G., Krause, M. S., & Trepanier, K. L. (1995). The relationship of the Yale-Brown Obsessive Compulsive Scale (YBOCS) to other measures of obsessive compulsive symptoms in a nonclinical population. *Journal of Personality Assessment, 65*(1), 158–168. https://doi.org/10.1207/s15327752jpa6501_12.

Goodman, W. K., Price, L. H., Rasmussen, S. A., & Mazure, C. (1989). The Yale-Brown Obsessive Compulsive Scale: II. Validity. *Archives of General Psychiatry, 46*(11), 1012–1016.

Hermans, D., Craske, M. G., Mineka, S., & Lovibond, P. F. (2006). Extinction in human fear conditioning. *Biological Psychiatry, 60*(4), 361–368. https://doi.org/10.1016/j.biopsych.2005.10.006.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19.

Jacoby, R. J., & Abramowitz, J. S. (2016). Inhibitory learning approaches to exposure therapy: A critical review and translation to obsessive-compulsive disorder. *Clinical Psychology Review, 49*, 28–40. https://doi.org/10.1016/j.cpr.2016.07.001.

Jacoby, R. J., Fabricant, L. E., Leonard, R. C., Riemann, B. C., & Abramowitz, J. S. (2013). Just to be certain: Confirming the factor structure of the Intolerance of Uncertainty Scale in patients with obsessive-compulsive disorder. *Journal of Anxiety Disorders, 27*(5), 535–542. https://doi.org/10.1016/j.janxdis.2013.07.008.

Kircanski, K., Mortazavi, A., Castriotta, N., Baker, A. S., Mystkowski, J. L., Yi, R., et al. (2012). Challenges to the traditional exposure paradigm: Variability in exposure therapy for contamination fears. *Journal of Behavior Therapy and Experimental Psychiatry, 43*(2), 745–751. https://doi.org/10.1016/j.jbtep.2011.10.010.

Lang, A. J., & Craske, M. G. (2000). Manipulations of exposure-based therapy to reduce return of fear: A replication. *Behaviour Research and Therapy, 38*(1), 1–12. https://doi.org/10.1016/S0005-7967(99)00031-5.

Meuret, A. E., Seidel, A., Rosenfield, B., Hofmann, S. G., & Rosenfield, D. (2012). Does fear reactivity during exposure predict panic symptom reduction? *Journal of Consulting and Clinical Psychology, 80*(5), 773–785. https://doi.org/10.1037/a0028032.

Nguyen, T. D., Attkisson, C. C., & Stegner, B. L. (1983). Assessment of patient satisfaction: development and refinement of a service evaluation questionnaire. *Evaluation and Program Planning, 6*(3–4), 299–313.

Norberg, M. M., Calamari, J. E., Cohen, R. J., & Riemann, B. C. (2008). Quality of life in obsessive-compulsive disorder: An evaluation of impairment and a preliminary analysis of the ameliorating effects of treatment. *Depression and Anxiety, 25*(3), 248–259.

Olatunji, B. O., Davis, M. L., Powers, M. B., & Smits, J. A. J. (2013). Cognitive-behavioral therapy for obsessive-compulsive disorder: A meta-analysis of treatment outcome and moderators. *Journal of Psychiatric Research, 47*(1), 33–41. https://doi.org/10.1016/j.jpsychires.2012.08.020.

Radomsky, A. S., Alcolado, G. M., Abramowitz, J. S., Alonso, P., Belloch, A., Bouvard, M., ... Wong, W. (2014). Part 1—You can run but you can't hide: Intrusive thoughts on six continents. *Journal of Obsessive-Compulsive and Related Disorders, 3*(3), 269–279. https://doi.org/10.1016/j.jocrd.2013.09.002.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, & W. F. Prokasy (Eds.). *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Rowe, M. K., & Craske, M. G. (1998). Effects of varied-stimulus exposure training on fear reduction and return of fear. *Behaviour Research and Therapy, 36*(7–8), 719–734.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., ... Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 59*, 22–33.

Steketee, G., Chambless, D. L., Tran, G. Q., & Worden, H. (1996). Behavioral Avoidance Test for obsessive compulsive disorder. *Behaviour Research and Therapy, 34*(1), 73–83.

Tolin, D. F., Diefenbach, G. J., & Gilliam, C. M. (2011). Stepped care versus standard cognitive-behavioral therapy for obsessive-compulsive disorder: A preliminary study of efficacy and costs. *Depression and Anxiety, 28*(4), 314–323. https://doi.org/10.1002/da.20804.

Twohig, M. P., Abramowitz, J. S., Smith, B. M., Fabricant, L. E., Jacoby, R. J., Morrison, K. L., ... Ledermann, T. (2018). Adding acceptance and commitment therapy to exposure and response prevention for obsessive-compulsive disorder: A randomized controlled trial. *Behaviour Research and Therapy, 108*, 1–9. https://doi.org/10.1016/j.brat.2018.06.005.

Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear extinction and relapse: state of the art. *Annual Review of Clinical Psychology, 9*, 215–248. https://doi.org/10.1146/annurev-clinpsy-050212-185542.

Williams, M. T., Farris, S. G., Turkheimer, E. N., Franklin, M. E., Simpson, H. B., Liebowitz, M., et al. (2014). The impact of symptom dimensions on outcome for exposure and ritual prevention therapy in obsessive-compulsive disorder. *Journal of Anxiety Disorders, 28*(6), 553–558. https://doi.org/10.1016/j.janxdis.2014.06.001.

Wolpe, J. (1973). *The practice of behavior therapy* (2nd ed.). Oxford England: Pergamon.