Behaviour Research and Therapy 51 (2013) 274–283

Contents lists available at SciVerse ScienceDirect

Behaviour Research and Therapy

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

# Imagery Rescripting as treatment for complicated PTSD in refugees: A multiple baseline case series study

Arnoud Arntz [a,b,*], Dana Sofi [a,c], Gerard van Breukelen [a]

[a] Maastricht University, The Netherlands
[b] Netherlands Institute for Advanced Study in the Humanities and Social Sciences, The Netherlands
[c] Osperon, den Bosch, The Netherlands

ARTICLE INFO

*Article history:*
Received 10 October 2012
Received in revised form
13 February 2013
Accepted 20 February 2013

*Keywords:*
PTSD
Refugees
Imagery Rescripting
CBT
Clinical trial

ABSTRACT

This study tested the effectiveness of Imagery Rescripting (ImRs) for complicated war-related PTSD in refugees. Ten adult patients in long-term supportive care with a primary diagnosis of war-related PTSD and Posttraumatic Symptom Scale (PSS) score > 20 participated. A concurrent multiple baseline design was used with baseline varying from 6 to 10 weeks, with weekly supportive sessions. After baseline, a 5-week exploration phase followed with weekly sessions during which traumas were explored, without trauma-focused treatment. Then 10 weekly ImRs sessions were given followed by 5-week follow-up without treatment. Participants were randomly assigned to baseline length, and filled out the PSS and the BDI on a weekly basis. Data were analyzed with mixed regression. Results revealed significant linear trends during ImRs (reductions of PSS and BDI scores), but not during the other conditions. The scores during follow-up were stable and significantly lower compared to baseline, with very high effect sizes (Cohen's $d = 2.87$ (PSS) and 1.29 (BDI)). One patient did clearly not respond positively, and revealed that his actual problem was his sexual identity that he couldn't accept. There were no dropouts. In conclusion, results indicate that ImRs is a highly acceptable and effective treatment for this difficult group of patients.



During the last decade there has been an increasing interest in Imagery Rescripting (ImRs) as a treatment or treatment ingredient for a variety of disorders, including PTSD and other anxiety disorders, depression, eating disorders, sleep problems and personality disorders. In ImRs, the patient imagines the (start of a) traumatic (or otherwise negative) experience, and then imagines an intervention that changes the course of events so that a more satisfying outcome is achieved. In original applications often the full trauma was imagined, before rescripting started. For instance, Arntz, Tiesema, and Kindt (2007; also Kindt, Buck, Arntz, & Soeter, 2007) added ImRs to Imaginal Exposure (IE), assuming that it would be ineffective to avoid exposure to the complete trauma memory. Although this study found the combination of IE and ImRs to be better tolerable (significantly less dropout) and more effective in non-fear emotions like shame, guilt, anger, and anger control than IE alone, attempts to apply the technique to highly complex cases necessitated changes in the application of ImRs. Often, these patients refused to relive the full trauma, or dissociated, or ran away. We therefore tried out to start rescripting already during events preceding the actual trauma, so that the patient imagined to be rescued from the trauma and did not have to imagine all the horrible details and feelings of helplessness, shame and guilt associated with the trauma proper. As a side effect, one ImRs often takes no more than 10–15 min; in contrast to the minimal 60 min that was used in the early Arntz et al. (2007) study, and 2–4 ImRs exercises can be done during one session. Clinical observations indicated good effects and high acceptability of the new procedure. Interestingly, this new procedure matches well with new insights from fundamental research, that stress the importance that the event triggering retrieval of the memory should contain new (hence, unexpected) information to bring about a reconsolidation of the memory in a different form (Finnie & Nader, 2012). The new procedure is now described in protocols (Arntz, 2011; 2012; Arntz & van Genderen, 2009), but has not been tested as treatment for complex PTSD.

One form of complicated PTSD is war-related PTSD in refugees. These patients are often considered to be very fragile, and usually have many current social problems that contribute to psychological dysfunction, such as unemployment; loss of social status (e.g., education not recognized in the host country); social

* Corresponding author. Clinical Psychological Science, Maastricht University, PO Box 616, The Netherlands.
*E-mail address:* Arnoud.Arntz@maastrichtuniversity.nl (A. Arntz).

0005-7967/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.brat.2013.02.009

isolation; boredom; uncertainty about their legal status (or having experienced that for a prolonged time); complicated bereavement; having no partner; and fears about their relatives that didn't flee. Also contributing to complexity is that their traumas are usually multiple and horrible, and that these refugees might have been both victims and perpetrators.

Only a few studies specifically tested PTSD treatments in refugees (see for a review Crumlish & O'Rourke, 2010). Without proper treatment, long-term natural course perspectives for refugees who suffer from the aftermaths of traumas are poor, especially when problems don't recede in the first years of arrival in the host country (Vaage et al., 2010). Various treatments for PTSD have been tested in refugees, but the methodological quality of the trials is, according to a systematic review, generally low (Crumlish & O'Rourke, 2010). Nevertheless, the authors concluded that there is evidence for effectiveness of narrative exposure therapy and cognitive-behavioral therapy. Effect sizes vary considerably, as do proportions with remitted PTSD. From the Crumlish and O'Rourke (2010) review, we calculated a mean remission rate between 40 and 50% (depending on completers vs ITT analyzes), with a range of 0–71% (ITT: 0–59%) for the most effective treatments, CBT and narrative exposure therapy.

The aim of the present study was to assess the effectiveness of ImRs as a treatment for war-related PTSD in refugees. As an initial test, we used a concurrent multiple baseline design with 10 patients with PTSD that reported levels of PTSD and depression on a weekly basis. By randomizing patients over five conditions of baseline length, we separated time from the consecutive conditions, so that the effects of treatment could be distinguished from that of time per se. We compared 10 sessions of ImRs to baseline. We also tested the effects of a 5 week exploration period, thus assessing whether expressing understanding and empathy for the traumas explains treatment effects. A 5 week follow-up period was used to assess the stability of the effects. Treatments were provided by the second author, a junior therapist, after one day of training in ImRs. This means that the current study is also important in view of implementation possibilities.

The reasons for choosing the concurrent multiple baseline design include the following. Although usually conceived as the gold standard, the randomized controlled trial (RCT) has limitations in its practicality, external validity, and costs (Hawkins, Sanson-Fisher, Shakeshaft, D'Este, & Green, 2007; Onghena, 2005; Onghena & Edgington, 2005). Like an RCT, a concurrent multiple baseline design can demonstrate that a change has occurred, that the change is the result of the intervention – and not of time, and that the change is significant (Hawkins et al., 2007; Onghena, 2005; Onghena & Edgington, 2005). Practical advantages over RCTs are that the design requires fewer participants (also an ethical advantage), and that participants act as their own controls – increasing power. An initial evaluation of a treatment is often done in an open trial. Compared to the open trial, the concurrent multiple baseline design has many advantages, as it is a true experiment (by the experimental manipulation of time when treatment starts), so that more causal inferences can be drawn than from an open trial, that offers little possibilities to control for time effects and attention. By adding an exploration phase in our design, and by the fact that during the baseline phase the usual supportive treatment is given, we controlled for nonspecific factors like attention and talking about traumas, further increasing the experimental control over testing the effects of ImRs. However, it should be realized that multiple baseline designs are not suitable for direct comparison of two or more active treatments that have strong lasting effects. For such studies, between-group designs are needed (i.e., RCTs). Moreover, multiple baseline designs are more suited for problems that are stabilized (so that there is no large time effect during baseline) than for problems with natural recovery. As it was not our aim in this phase of research to compare ImRs to another potentially powerful treatment, and we wanted to test ImRs in patients with chronic PTSD despite being in usual care, the concurrent multiple baseline design was a good option.

## Methods

### *Participants*

Participants were 10 patients from the mental health care institute Osperon in den Bosch, the Netherlands. This institute is specialized in the supportive care for refugees, but felt a necessity to offer trauma-focused treatments for patients suffering from PTSD and needing processing of their traumatic experiences. Inclusion criteria were: (1) primary diagnosis of PTSD as assessed with the SCDI-1, resulting from war-related traumas; (2) a score of >20 on the PSS at screening, (3) age 18–65; (4) ability to communicate with therapist with or without interpreter; (5) willingness to participate in the study (signed informed consent). Exclusion criteria were: (1) life-time psychosis (though psychotic features along depression were allowed) or bipolar disorder type 1; (2) IQ < 80; (3) acute suicide risk; (4) substance dependence; (5) start of new medication within 3 months before start of the study (medication used for longer periods could be continued; stopping medication during the study was allowed). No other evidence based treatment of PTSD was allowed during the study. To increase external validity we didn't require that participants could speak, write or read Dutch. Assessments and treatments were done in the language of their preference, if necessary with help of professional translators. The 10 participants were recruited from 22 patients screened for participation, of whom 4 were excluded because of a PSS score <20, and 8 because another diagnosis was primary. Fig. 1 presents the patient flow. Table 1 gives an overview of the characteristics of the participants. The study protocol was approved by the ethical committee of the Faculty of Psychology and Neurosciences of Maastricht University.

### *Procedure*

Patients in regular supportive care but not improving enough from that care with a clinical diagnosis of PTSD due to war-related traumas were approached by the second author. A full SCID-1 was applied to assess Axis-1 disorders. In case of a suspicion of low IQ, a full intelligence test was applied (turned out to be not necessary). Potential participants were fully informed about the study both verbally and by written information, and gave written consent if they agreed to participate, which they all did. After 10 participants were included, they were randomized to baseline length (2 participants per length) by the following procedure, executed by an independent person in the presence of institute's psychiatrist and the second author. Each name was written on a separate piece of paper and folded up so that the name was invisible. The 10 papers were put in a bin and stirred by the independent person. Next the independent person drew two papers, and the participants named on the papers were assigned to the longest baseline. The next two drawn participants were assigned to the second longest baseline, etc., until there were two left who were assigned to the shortest baseline condition. The patient was not informed about the allocated baseline length (treatment as usual continued), and the therapist just started with the exploration phase when indicated by the outcome of the randomization.

### *Instruments*

The Dutch SCID-1 was used to assess axis-1 disorders (van Groenestijn, Akkerhuis, Kupka, Schneider, & Nolen, 1999).

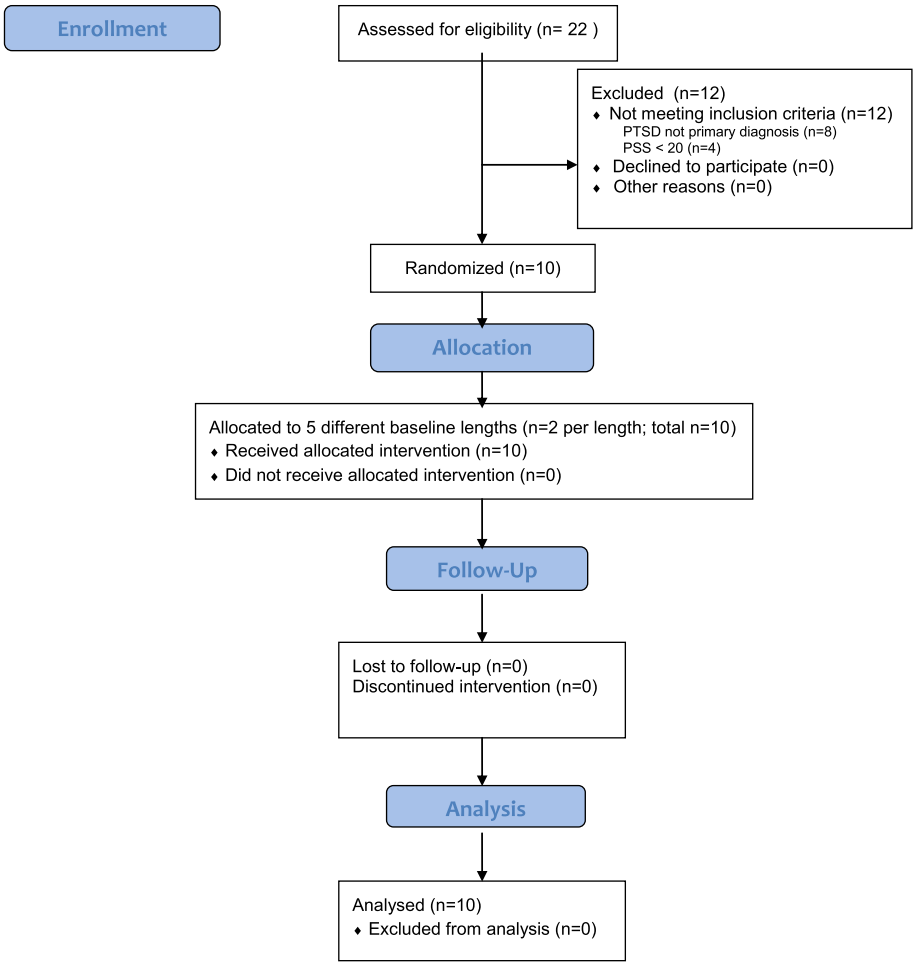


**Fig. 1.** CONSORT flow diagram.

Lobbestael, Leurgans, and Arntz (2011) report interrater agreement in the fair to excellent range between raters trained in our institute. The mean value of Cohen's Kappa between raters for Axis-I disorders was .71, for PTSD .77. The Posttraumatic Symptom Scale (PSS; Foa, Riggs, Dancu, & Rothbaum, 1993) assesses the frequency of PTSD symptoms according to the DSM-IV during the last week (range 0–51). It has excellent internal consistency (e.g., Cronbach alpha = .88 in a Dutch sample, Arntz et al., 2007), and adequate validity (Engelhard, Arntz, & van den Hout, 2007; Foa et al., 1993). A cut-off equal to or greater than 15 on the PSS score has been proposed as highly sensitive for the diagnosis of PTSD (Wohlfarth, van den Brink, Winkel, & ter Smitten, 2003), while another study found a cutoff score of 14 (Coffey, Gudmundsdottir, Beck, Palyo, & Miller, 2006). A score of <14 was thus used as criterion for remission from PTSD. We used the interview form of the PSS which has similar reliability as the self-report form (Cronbach alpha = .86) and excellent validity (Foa & Tolin, 2000). The PSS was the primary outcome measure. The Beck Depression Inventory version 2 (BDI, Beck, Steer, & Brown, 1996), range 0–61, was used to assess depressive symptoms, and was the secondary outcome measure. The BDI has excellent internal consistency (Cronbach alpha = .91; Beck, Steer, Ball, & Ranieri, 1996) and is often used as outcome instrument in treatment research. Norms are 0–13: minimal depression; 14–19: mild depression; 20–28: moderate depression; and 29–63: severe depression. The BDI was also taken as an interview, given the reading problems of some participants. For both instruments, we used officially translated versions in languages suitable to the specific participant.

### *Design*

A concurrent multiple baseline design (also: stepped wedge or staggered baseline design) was used, with a baseline varying in length from 6 to 10 weeks, with 2 participants randomly allocated to each of the 5 lengths. As the recruitment of all participants precedes randomization, and all participants started with the baseline phase at the same time, the multiple baseline design is called "concurrent" (Carr, 2005). This design is deemed to be superior to nonconcurrent designs, as it has higher power of verification of the effects of the experimental manipulation because it controls for historical effects (Carr, 2005). The variation in baseline length offers the possibility to differentiate between time effects and experimental effects of the treatment (and the treatment-control, exploration). After baseline, a 5 week exploration phase started – which was used as a control for the effects of attention to the participants' stories of trauma. A 10 week treatment phase followed, during which ImRs was applied. A 5 week follow-up period followed immediately to assess the midterm effects of ImRs. A 3-months follow-up assessment was added to explore longer term effects, results will be reported separately together with a qualitative analysis of patients' views. No systematic way to perform power analysis for this type of design is known to us. As an indication, the study would have 80% power to

**Table 1**
Demographic data of participants ($N = 10$).

| Variable | | Mean (SD)/Number (%) |
|---|---|---|
| Age (in yrs) | Range 25–56 | 39.9 (12.24) |
| Gender | Female | 2 (20%) |
| | Male | 8 (80%) |
| Country of origin | Iraq | 3 (30%) |
| | Turkey | 2 (20%) |
| | Kosovo | 1 (10%) |
| | Afghanistan | 1 (10%) |
| | Guinea | 1 (10%) |
| | Somalia | 1 (10%) |
| | Ghana | 1 (10%) |
| Religion | Islam | 6 (60%) |
| | Christian | 3 (30%) |
| | None | 1 (10%) |
| Marital status | Married/living together | 4 (40%) |
| | Single | 5 (50%) |
| | LAT | 1 (10%) |
| Vocational status | Unemployed | 4 (40%) |
| | Employed | 6 (60%) |
| Educational level (highest completed) | Primary school not completed | 1 (10%) |
| | Primary school | 4 (40%) |
| | High school | 1 (10%) |
| | University | 4 (40%) |
| Language of treatment | Kurdish | 3 (30%) |
| | Arabic | 2 (20%) |
| | Dutch | 2 (20%) |
| | French | 1 (10%) |
| | Servo-Croatian | 1 (10%) |
| | Farsi | 1 (10%) |
| Duration of PTSD (yrs) | Range 2–27 | 13.1 (8.1) |
| Traumas (targeted in ImRs) | Witnessing relatives/ intimates/friends killed | 8 (80%) |
| | Threatened to be killed | 6 (60%) |
| | Sexual Abuse | 5 (50%) |
| | Torture & other physical abuse | 4 (40%) |
| | Life-threatening injuries not by direct violence | 1 (10%) |
| Age of first trauma | Range: 7–27 (yrs) | 18.3 (7.8) |
| Secondary axis-1 diagnoses | None | 4 (40%) |
| | MDD | 5 (50%) [4 chronic] |
| | Anxiety Disorder | 1 (10%) |
| | Somatization Disorder | 1 (10%) |
| Treatment duration at institute (yrs) | Range 1.5–5 | 2.9 (1.19) |
| Medication at baseline | No medication | 0 (0%) |
| | SSRI's | 10 (100%) |
| | Neuroleptics | 8 (80%) |
| | Benzodiazepines | 6 (60%) |
| | Mood stabilizers | 5 (50%) |
| | Sum | 2.90 (1.10) |
| Medication at follow-up | No medication | 4 (40%) |
| | SSRI's | 6 (60%) |
| | Neuroleptics | 4 (40%) |
| | Benzodiazepines | 0 (0%) |
| | Mood stabilizers | 0 (0%) |
| | Sum | 1.00 (.94) |

detect a change of Cohen's $d = 1$ or higher at alpha = .05, two-tailed, if the paired $t$-test of the pre to post change were used to evaluate the treatment effect.

### *Treatments and therapist*

All participants were in regular supportive mental health care, which was continued during the baseline phase. During the exploration phase therapist and patient explored the traumas that caused the PTSD, their context, and their meaning for the patient. They made a decision which traumas were to be addressed in the ImRs phase, and in which order. No hierarchy was made, but in case of extended traumas these were divided in subtraumas that could be targeted separately with ImRs. The order of (sub)traumas to be addressed was chosen by the patient. During the ImRs phase, the therapist explained the technique in detail, and started with its application. Generally, the therapist started with intervening in the image, taking care that the intervention was imagined before the actual trauma took place, but late enough to have an arousing expectation triggered in the patient. If the patient felt strong enough, (s)he could take the lead in the rescripting, assisted by the therapists or others if needed (see Arntz & Weertman, 1999). Generally, two to three ImRs exercises were done per session. Examples of imagined rescripting include: imagining members of patient's tribe successfully defending patient's family against attack by another tribe; imagining revenge by killing the perpetrator and getting money as compensation; imagining giving the killed a decent funeral; imagining to be very strong and successfully defend against a rapist. During the follow-up period, no treatment was provided. The weekly assessments were taken by the therapist. The treatments were provided by one junior therapist (the second author) after a one-day training by the first author. Occasional telephone supervision was offered, but turned out to be needed only once.

### *Statistical analysis*

Mixed regression was used to assess the differences between the exploration, treatment (ImRs) and follow-up phases on the one hand, and baseline on the other hand, in average scores and linear change. The fixed model part consisted of 1) a general linear time effect starting with time = 0 when the first assessment was taken for an individual, and 2) dummy indicators for the exploration, treatment, and follow-up phases (thus contrasting each to baseline), and 3) three centered time-within-condition covariates, one for every phase except baseline to assess time by phase interaction, that is, changes in the time effect across phases (cf. Vlaeyen, De Jong, Geilen, Heuts, & Van Breukelen, 2001). The random model part consisted of a random intercept to capture between-subject outcome variation, plus ARMA11 for the within-subject covariance structure. Random slopes to allow interindividual variation in time and condition effects led to reduced fit of the model or convergence problems, and were therefore not included. The analytic strategy was to first test for a general time effect, next to assess the full model with all predictors entered, and then to delete in backward fashion the time by phase interactions that were N.S. If the main time effect was N.S., it was deleted at the last step. Cohen's $d$ was calculated as effect size of the change of a phase with respect to baseline: $d$ = the mean outcome difference between baseline and current phase, derived from the fixed part of the mixed regression divided by the sd of the residual outcome variance (the patient-specific outcome mean per phase has as variance random intercept (between subject variance) + (residual (within-subject) variance/number of measurements per phase); the square root of this subject-specific variance is the denominator for $d$).

## Results

### *PSS*

Fig. 2 shows the individual PSS scores of the 10 participants during the different phases. Visual inspection suggests decreases in PSS during ImRs in 8 of 10 participants, and lower PSS scores during FU than during baseline in all but one participant (nr. 3). Mixed regression revealed a highly significant linear effect of time, $t(18.52) = -4.80$, $p < .001$. When all predictors were entered, the time-within-exploration and time-within-follow-up effects appeared to be nonsignificant, $p$'s $> .60$. After stepwise deleting, the

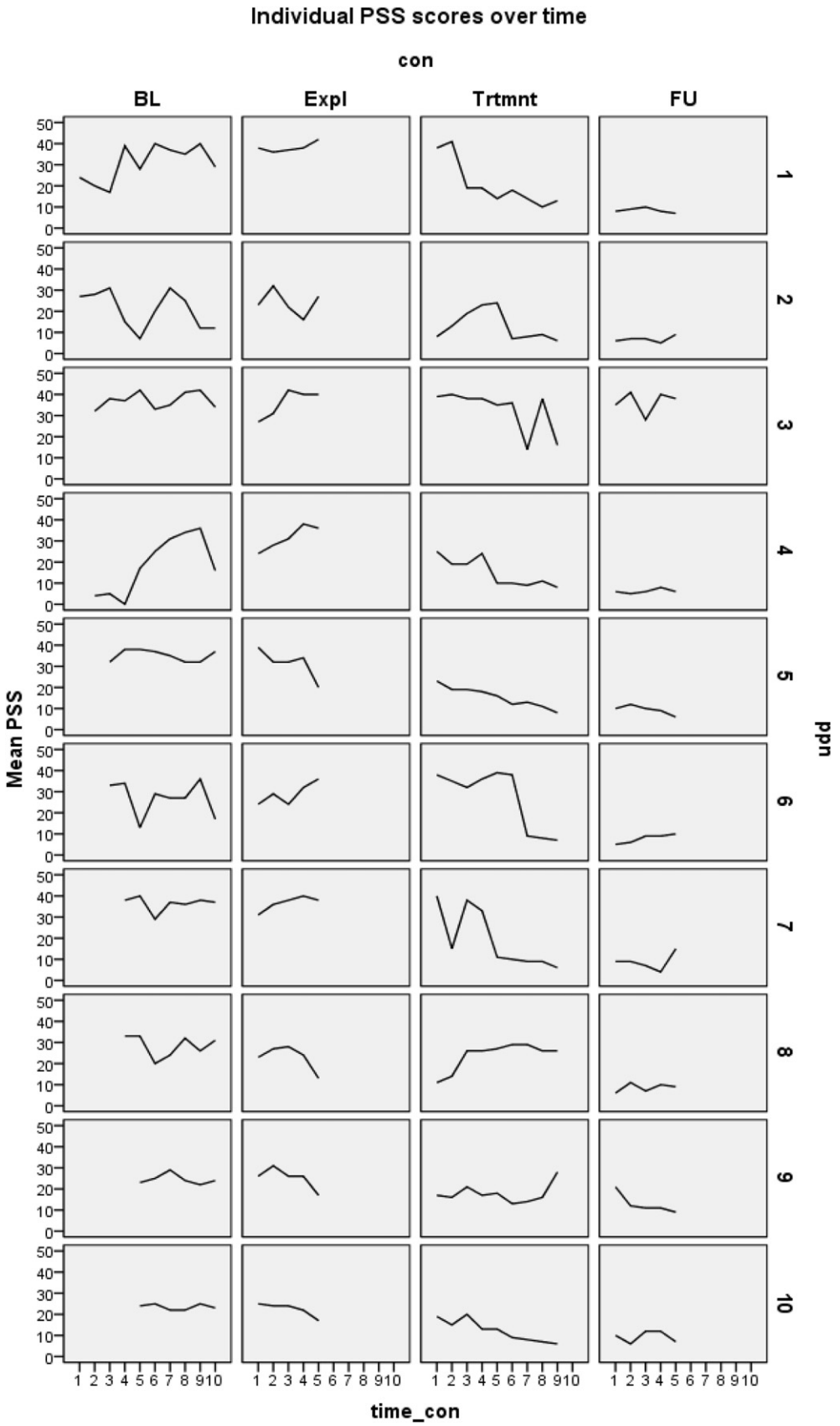


**Fig. 2.** Individual PSS scores. Ppn = participant's number; time_con = week number within condition; BL = baseline; Expl = exploration; Trtmnt = ImRs treatment; FU = follow-up.

main effect of time appeared to be nonsignificant, $p = .57$, and was therefore also deleted. Table 2 presents the final results of the mixed regression analysis. Exploration had no effect on the PSS ratings, but the main effect of treatment (i.e., the change halfway treatment compared to baseline) was significant, as was the main effect of follow-up (as compared to baseline). The time-within-treatment effect was significant, showing a steep decrease of PTSD symptoms. Effect sizes of treatment vs baseline, and follow-up vs baseline were very high (Table 2). Fig. 3 depicts both the observed means and the predicted means from the analysis. Inspection of the residuals indicated outliers during follow-up caused by high scores by participant 3 (to be discussed in Discussion). The analyzes were repeated without participant 3, but yielded similar results with higher effect sizes.

*Remission from PTSD*

Using the cut-off criterion of 14 on the PSS, 9 of 10 patients scored during follow-up on average <14 and can thus be classified as remitted. The non-remitted participant (nr.3) had clearly not improved (36.4).

*BDI*

Fig. 4 shows the individual BDI scores of the 10 participants during the different phases. Visual inspection suggests decreases in BDI during ImRs in all but two participants, and lower BDI scores during FU in 8 out of 10 participants. Mixed regression revealed a highly significant main linear effect of time, $t(14.43) = -4.29$, $p = .001$. When all predictors were entered, the time-within-exploration, and time-within-follow-up effects appeared to be nonsignificant, $p$'s $> .33$. After stepwise deleting, the main effect of time appeared to be nonsignificant, $p = .54$, and was also deleted. Table 2 presents the final results of the mixed regression analysis. Exploration had no effect on the BDI ratings, but the main effect of treatment was significant, as was the main effect of follow-up. The slope during treatment was significant, showing a steep decrease of depressive symptoms. Effect sizes of treatment vs baseline, and follow-up vs baseline were very high (Table 2). Fig. 5 depicts both the observed means and the predicted means from the analysis. Inspection of the residuals indicated outliers during follow-up caused by high scores by participant 3 (to be discussed in Discussion). The analyzes were repeated without participant 3 but yielded similar results with higher effect sizes.

**Table 2**
Results of mixed regression analyzes.

| | Parameter | $\beta$ | Std. error | df | $t$ | $p$ | Effect size Cohen's $d$[a] |
|---|---|---|---|---|---|---|---|
| **PSS** | Intercept | 28.034 | 2.191 | 17.151 | 12.796 | <.001 | |
| | Exploration | 1.614 | 1.807 | 135.833 | .893 | .373 | −.284 |
| | Intervention | −8.608 | 2.020 | 37.418 | −4.261 | <.001 | 1.574 |
| | Follow-up | −16.416 | 2.457 | 22.844 | −6.680 | <.001 | 2.866 |
| | Time within Intervention | −1.733 | .389 | 99.810 | −4.460 | <.001 | |
| **BDI** | Intercept | 28.274 | 3.168 | 11.818 | 8.924 | <.001 | |
| | Exploration | −.070 | 1.785 | 97.880 | −.039 | .969 | .075 |
| | Intervention | −8.164 | 1.983 | 26.141 | −4.117 | <.001 | .891 |
| | Follow-up | −11.989 | 2.341 | 16.784 | −5.121 | <.001 | 1.286 |
| | Time within Intervention | −1.248 | .3869 | 94.935 | −3.226 | .002 | |

Note. Predictors were coded as follows: Dummy coding for Exploration (1, 0); Intervention (1, 0), and Follow-up (1, 0) so that Baseline was the reference category; Time-within-Condition: 0 for measurements outside the condition, centered time (with week as unit) for measurements within condition (e.g., −2,−1,0,1,2 for a 5-week condition); main Time effect: (0,1,2,3,…) with one unit per week starting with 0 at the first measurement during baseline per participant.
For PSS, intercept variance (BS) was 22.239, error variance (WS) 65.005, ARMA11 rho = .749, and ARMA11 phi = .581.
For BDI, intercept variance (BS) was 75.943, error variance (WS) 67.395, ARMA11 rho = .813, and ARMA11 phi = .419.
[a] Effect size Cohen's $d$ = effect/S.D., with S.D. derived from variance of the random part as explained in Methods, and effects equal to the beta's of the Mixed Regression model. Positive values denote improvement. Effect size of Intervention phase represents effect halfway the intervention (not at the end of the intervention).

*Change in depression status*

Averaged baseline BDI-scores indicated that 1 participant had mild depression, 6 moderate, and 3 severe. During 5-weeks follow-up there were 6 participants in the minimal depression category, 3 mild, and 1 severe (nr.3), based on their average BDI over the whole 5-week follow-up period.

*Additional findings*

*Medication use* reduced from an average of 2.90 (sd 1.10) types of medication at baseline to 1.00 (sd .94) at the end of the 5-week follow-up period, $t(9) = 10.59$, $p < .001$, Cohen's $d = 3.35$. At baseline all patients used medication, at follow-up four were medication free. All reduced medication use, and 7 of the ten patients even with two types of medication. At follow-up no benzodiazepines and mood stabilizers were used anymore (Table 1).

*Adverse effects.* No adverse effects were observed.

## Discussion

We tested ImRs as a treatment for complicated PTSD in refugees using a multiple baseline case series design, and found evidence for strong effects of this technique on PTSD-symptoms as assessed with the PSS, and on the secondary outcome of depressive symptoms, assessed with the BDI. Mixed regression analyzes revealed that there was no evidence for significant time effects within exploration and follow-up phases, whereas the linear time effect during ImRs was strong, indicating that ImRs had a positive impact on symptoms already during treatment. The general time effect disappeared after conditions were entered in the model, indicating that it is highly unlikely that effects are to be attributed to a time effect. Nine of 10 participants remitted from PTSD, a very high proportion, also when compared to the proportions reported in other studies in refugees (on average 40–50%, with a maximum of 71% remission, Crumlish & O'Rourke, 2010). Although some of this improvement might in principle be due to regression to the mean, the results of the mixed model indicate improvement in the treatment phase rather than a gradual improvement over all measurements.

The cases were highly complicated and in long-term care in the institute specialized in mental health care for refugees. Complications were not only related to the horrible traumas they had experienced, but also included factors outside psychopathology, but influencing it, like lack of integration in the Netherlands, having dealt with lengthy procedures (and uncertainty) to get a permit to stay in the Netherlands, being unemployed, having no partner, not being able to speak Dutch, etc. All used medication, 9 (90%) multiple, 8 (80%) used neuroleptics and five (50%) mood stabilizers, also indicating that they were severely distressed, and that their therapists felt that they needed stabilization with these medications. Despite such indicators of problematic functioning, responses to treatment were generally very positive – to the surprise of the center's team. Treatment also led to a strong reduction of medication use, with no patients using benzodiazepines or mood stabilizers anymore after treatment.

One case (nr. 3) was a statistical outlier during the FU phase. The participant did not respond positively to ImRs, at least not on the

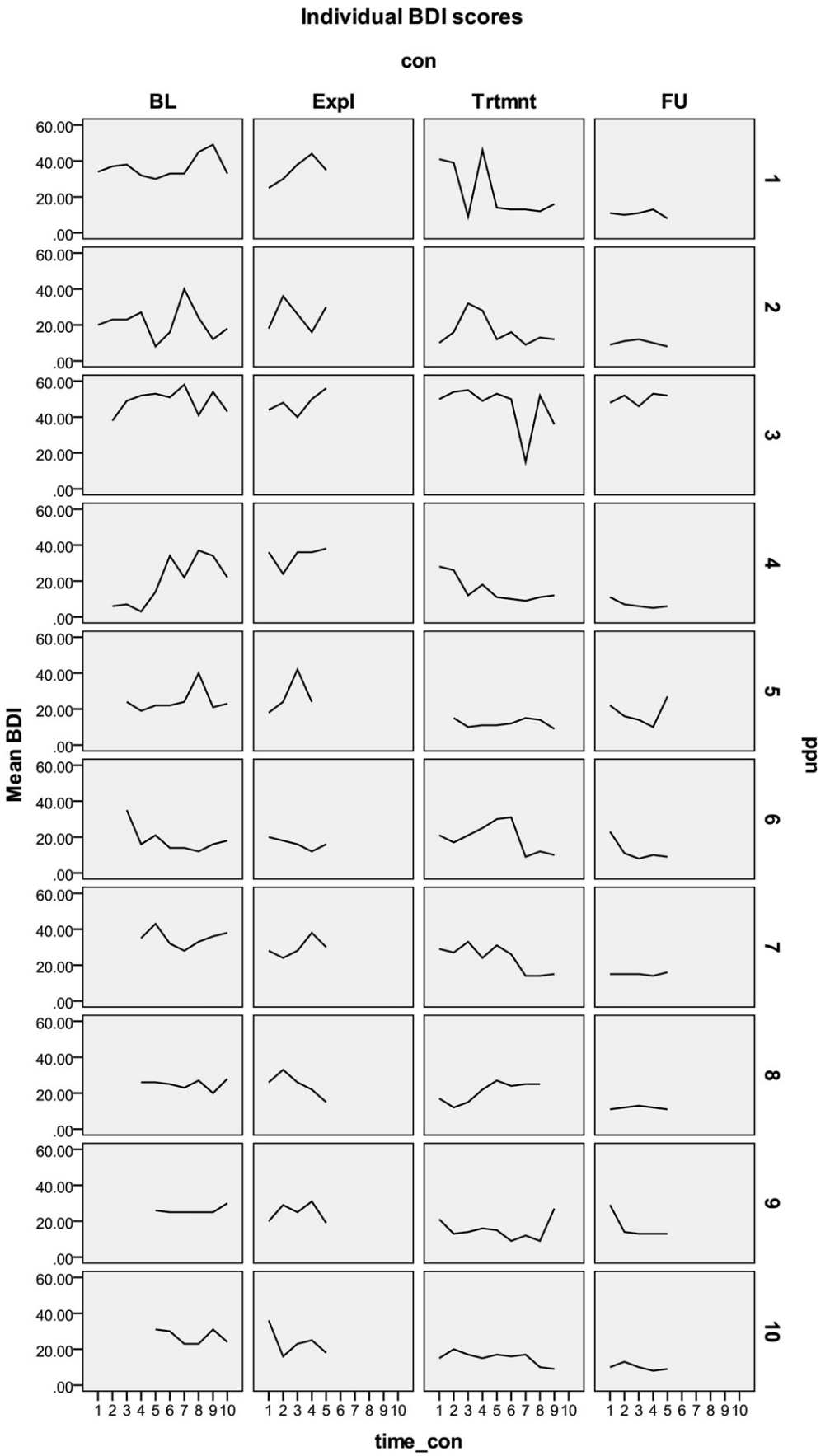


**Fig. 3.** Individual BDI scores. Ppn = participant's number; time_con = week number within condition; BL = baseline; Expl = exploration; Trtmnt = ImRs treatment; FU = follow-up.

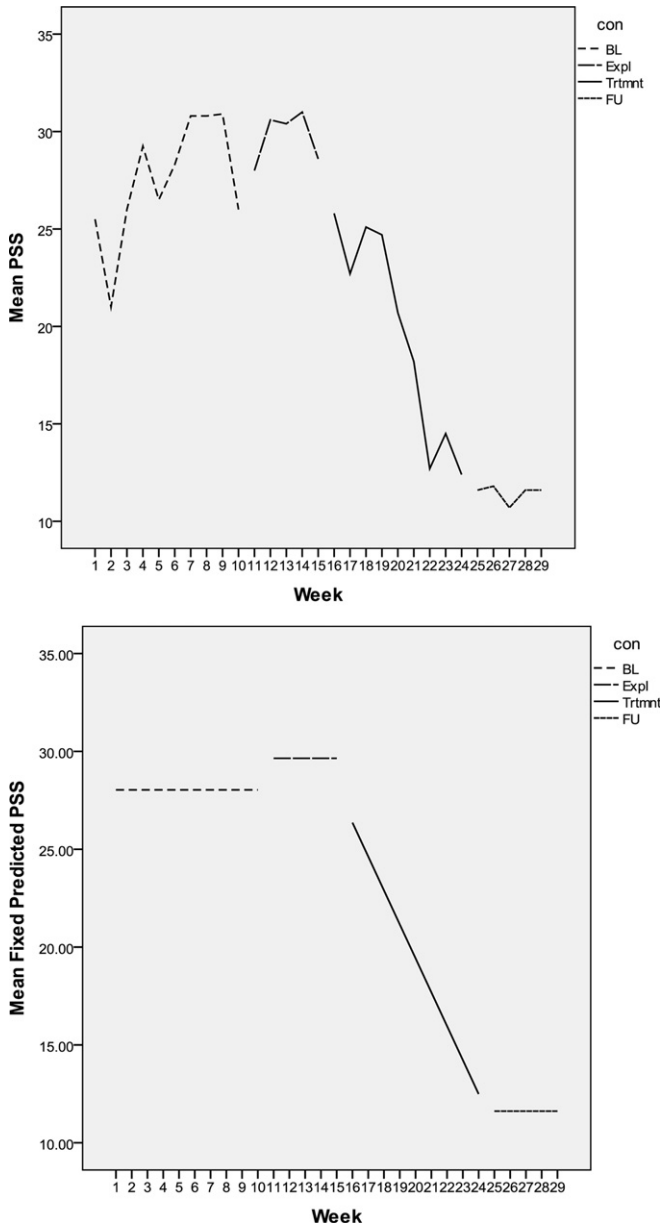


**Fig. 4.** Mean PSS scores: observed (upper panel) and predicted by the mixed regression model (lower panel). BL = baseline; Expl = exploration; Trtmnt = treatment; FU = follow-up; week = time index starting with 1 at the first assessment in participants 1 and 2. Note that for week 1–4, data are only available for a subset of the sample due to differences in baseline length.

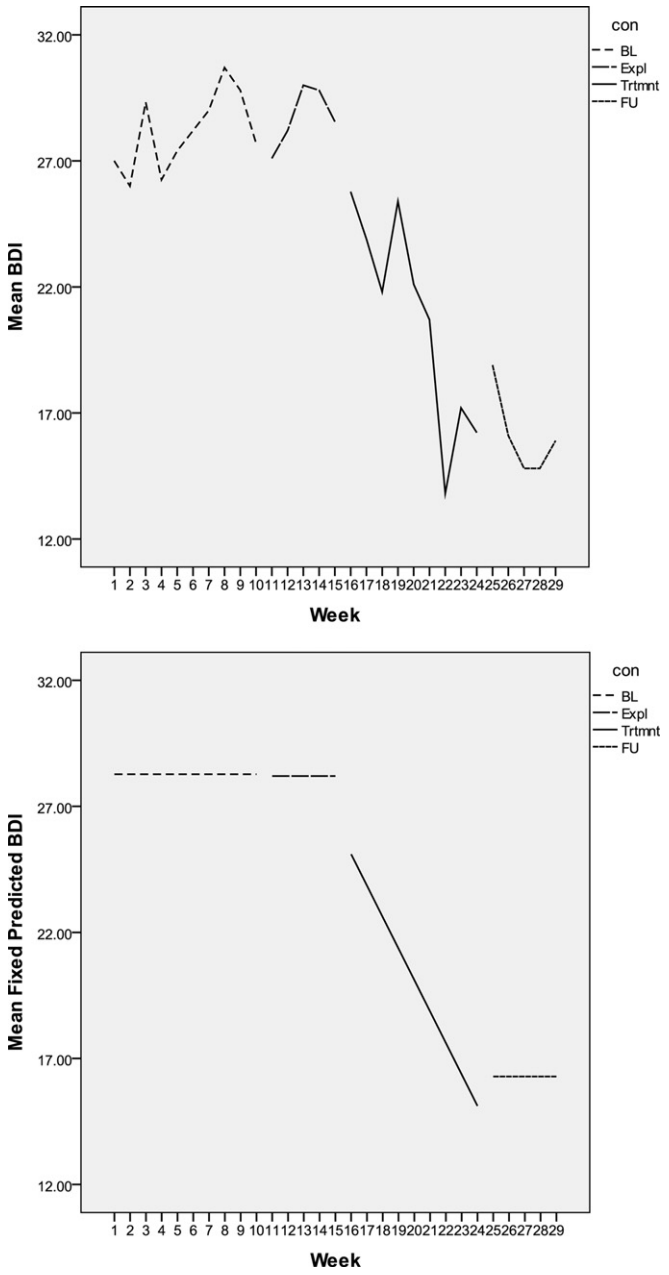


**Fig. 5.** Mean BDI scores: observed (upper panel) and predicted by the mixed regression model (lower panel). BL = baseline; Expl = exploration; Trtmnt = treatment; FU = follow-up; week = time index starting with 1 at the first assessment in participants 1 and 2. Note that for week 1–4, data are only available for a subset of the sample due to differences in baseline length.

level of PTSD and depression symptoms. The participant revealed at the end of treatment that his real problem was his struggle with his sexual identity, which he felt was unacceptable for his religious belief, and that he felt that processing traumatic memories would not help him with this. Thus, despite that ImRs failed to reduce PSS and BDI scores in this participant, the positive effect of treatment was that it helped the participant to understand that his sexual identity was his core problem and that he had to share this with his therapist.

Among the limitations of the present study the following deserve to be mentioned. First, all patients were treated by one and the same therapist so that it cannot be excluded that part of the effects are associated with this specific person. Second, the questionnaires were taken by the therapist, which might have influenced the ratings by the patient. An independent person blind to the condition participants are in would have been a protection against possible bias. Financial limitations prohibited us to do this, but researchers planning future studies should certainly consider this. Third, although the patients were recruited from a larger sample of potentially suitable patients, and explicit in and exclusion criteria were used, as with any open trial and case series study, it cannot be excluded that there was selective sampling. Fourth, although we controlled for time and for nonspecific therapeutic factors like attention, support, talking to a therapist, and empathy, not all possible alternative mechanisms that can explain the effects of ImRs could be controlled – an RCT would be necessary for that, e.g., to compare effects of ImRs to those of exposure alone (note

however that such a study has been done in non-refugee PTSD patients with an older form of ImRs, Arntz et al., 2007). Fifth, all patients were already in treatment at the institute and were more or less enough 'stabilized' to start trauma processing. No conclusions can be drawn about the effects of ImRs for similar complex patients in earlier phases of treatment. In other words, the amount of 'stabilization' needed for such patients before they can start ImRs is unknown, and obviously an important question. Taken together, the positive results call for a large scale RCTs addressing the issues related to the limitations mentioned.

A possible criticism on the study is its relatively small sample size. A sample size of $N = 10$ seems very small compared to RCTs that tend to get larger and larger, driven by power considerations. It should be stressed however that the very same power considerations led us to choose for this sample size. Statisticians have estimated that samples as small as $N = 4$ can suffice to demonstrate treatment effects with sufficient power using multiple baseline designs (Onghena, 2005), but given the lack of practical tools for power calculations for this type of design, we used a simple paired *t*-test power calculation to base our sample size on.

Another criticism might be that therapist and patient were not blind to treatment. This is however almost never the case in trials investigating psychological treatments, as the treatment is accomplished by active execution of procedures by therapist and patient. This makes a double blind approach like in placebo-controlled pharmacological trials usually impossible. However, the patients were blind as to length of baseline, as they were not informed about the length of the baseline and the therapist announced the Exploration phase only in the session where this phase started.

The technique of ImRs was trained in only 1 day to the therapists of the institute where the study took place. Local non-expert peer-supervision was provided, and only one expert telephone supervision session was needed. Acceptance by patients from different cultural background, and with different religious beliefs, was excellent. This indicates that ImRs can be quite easily disseminated and effectively implemented. ImRs is part of several treatment protocols, e.g., the cognitive therapy protocol for PTSD by Ehlers and Clark (2000); Ehlers et al. (2003), Ehlers, Clark, Hackmann, McManus, & Fennell (2005), the cognitive therapy protocol for Social Phobia by Clark et al. (2006), Wild, Hackmann, and Clark, (2007, 2008), and the schema therapy protocols for various personality disorders (e.g., Arntz, 2011; Arntz & van Genderen, 2009; Arntz & Jacob, 2012; Young, Klosko, & Weishaar, 2003). Experimental studies have indicated that ImRs is an effective technique for a range of disorders (see Arntz, 2012; for a review; see also Holmes, Arntz, & Smucker, 2007; Nilsson, Lundh, & Viborg, 2012). However, only a few studies have tested to what degree ImRs can offer an effective and complete treatment (see Nilsson et al., 2012; for a similar comment), but these studies indicate that ImRs is a viable alternative for more complex treatment packages.

The strong effects on depression are not surprising given the fact that treatment focusing on PTSD usually reduces concomitant depression. However, the very strong effects on depressive symptoms might at least partially be explained by the fact that ImRs is also a very effective treatment for depression (Brewin et al., 2009; Wheatley et al., 2007). The Brewin et al. (2009) study focused on treatment resistant chronic depression. Likewise, the present sample suffered from chronic problems despite years of supportive and pharmacological treatment.

Theorizing about and research into mechanisms underlying ImRs has just begun. One interesting possibility is that ImRs, at least when practiced in the way used in this study, does not rely on the mechanisms known to underlie extinction, that is the formation of a new memory trace that inhibits the original fear memory, but on a change of the meaning of the original memory – thus to a reconsolidation of the emotional memory in another form than when it was retrieved. Studies into such explanations have only just started but are of obvious importance as the possibility that ImRs works thru a different mechanism than extinction is clinically relevant: it might lead to quicker generalization and increased resistance for return of fear and other emotional problems (Arntz, 2012; Dibbets, Poort & Arntz, 2011; Hagenaars & Arntz, 2012). Moreover, this explanation of how ImRs works supports the use of early rescripting (i.e., start rescripting before the most traumatic part of the event is happening) as the unexpectedness of the different outcome than the trauma might be essential for reconsolidation of the trauma memory with a different emotional meaning (Finnie & Nader, 2012).

An alternative explanation is offered by Brewin's theory about retrieval competition (Brewin, 2006). This theory is more akin to the current understanding of extinction, in that the original memory representation is not changed by treatment, but that treatment offers a new, more functional representation that has to compete with the original representation. Applied to ImRs, the rescripting would form an alternative representation of the trauma and its meaning, which competes with the original dysfunctional representation, and hopefully "wins" the competition most of the time.

It is also of interest to discuss ImRs in the context of Ehlers' theory on PTDS (Ehlers & Clark, 2000; Ehlers et al., 2005, 2002). According to this theory, stimuli preceding the actual trauma can become 'warning signals' and are essential in PTSD. They are prominent in intrusions (Ehlers et al., 2002), and give the individual the sense of imminent threat – as if the trauma is going to happen again. These 'warning signals' might thus be excellent starting points to start rescripting – as they are experienced as predictors of the trauma (and underlie reliving experiences), and an unexpected change in the sequence of events, preventing the trauma and meeting the needs of the person, might thus be more effective than fully reliving the whole trauma.

To summarize, the present study yielded first evidence for the effectiveness of ImRs as a treatment for complicated PTSD. The strong effects support our clinical observation that starting rescripting earlier in the image, so that the actual trauma is prevented and safety is brought into the image, improves the effectiveness of the technique, and helps to apply it with even the most horrible trauma's, as it is not necessary to imagine and relive the whole trauma in detail. More research is necessary to test whether others can replicate the effects, if possible with assessments done by an independent assessor blind for condition. In addition to further research into clinical effectiveness, future studies should address the possible underlying mechanisms of ImRs.

## Declaration of interests

The first author published on Imagery Rescripting in scientific articles and books, and occasionally gives workshops on the technique. Financial profits are transferred to Maastricht University, not to his private account. No other interests declared.

## References

Arntz, A. (2011). Imagery rescripting for personality disorders. *Cognitive and Behavioral Practice, 18*, 466–481.

Arntz, A. (2012). Imagery rescripting as a therapeutic technique: review of clinical trials, basic studies, and research agenda. *Journal of Experimental Psychopathology, 3*, 198–208. http://dx.doi.org/10.5127/jep.024211.

Arntz, A., & van Genderen, H. (2009). *Schema therapy for borderline personality disorder*. Oxford: Wiley-Blackwell.

Arntz, A., & Jacob, G. (2012). *Schema therapy in practice: An introductory guide to the schema mode approach*. Chichester: Wiley.

Arntz, A., Tiesema, M., & Kindt, M. (2007). Treatment of PTSD: a comparison of imaginal exposure with and without imagery rescripting. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 345–370. http://dx.doi.org/10.1016/j.jbtep.2007.10.006.

Arntz, A., & Weertman, A. (1999). Treatment of childhood memories; theory and practice. *Behaviour Research and Therapy, 37*, 715–740. http://dx.doi.org/10.1016/S0005-7967(98)00173-9.

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. (1996). Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment, 67*, 588–597. http://dx.doi.org/10.1207/s15327752jpa6703_13.

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

Brewin, C. R. (2006). Understanding cognitive behaviour therapy: a retrieval competition account. *Behaviour Research and Therapy, 44*, 765–784. http://dx.doi.org/10.1016/j.brat.2006.02.005.

Brewin, C. R., Wheatley, J., Patel, T., Fearon, P., Hackmann, A., Wells, A., et al. (2009). Imagery rescripting as a brief stand-alone treatment for depressed patients with intrusive memories. *Behaviour Research and Therapy, 47*, 569–576. http://dx.doi.org/10.1016/j.brat.2009.03.008.

Carr, J. E. (2005). Recommendations for reporting multiple-baseline designs across participants. *Behavioral Interventions, 20*, 219–224. http://dx.doi.org/10.1002/bin.191.

Clark, D. M., Ehlers, A., Hackmann, A., McManus, F., Fennell, M. J. V., Grey, N., et al. (2006). Cognitive therapy vs. exposure and applied relaxation in social phobia: a randomized controlled trial. *Journal of Consulting and Clinical Psychology, 74*, 568–578. http://dx.doi.org/10.1037/0022-006X.74.3.568.

Coffey, S. F., Gudmundsdottir, B., Beck, G., Palyo, S. A., & Miller, L. (2006). Screening for PTSD in motor vehicle accident survivors using the PSS-SR and IES. *Journal of Traumatic Stress, 19*, 119–128.

Crumlish, N., & O'Rourke, K. (2010). Systematic review of treatments for post-traumatic stress disorder among refugees and asylum-seekers. *Journal of Nervous and Mental Disease, 198*, 237–251.

Dibbets, P., Poort, H., & Arntz, A. (2011). Adding imagery rescripting during extinction leads to less ABA renewal. *Journal of Behavior Therapy and Experimental Psychiatry, 43*, 614–624.

Ehlers, A., & Clark, D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy, 38*, 319–345. http://dx.doi.org/10.1016/S0005-7967(99)00123-0.

Ehlers, A., Clark, D. M., Hackmann, A., McManus, F., & Fennell, M. (2005). Cognitive therapy for posttraumatic stress disorder: development and evaluation. *Behaviour Research and Therapy, 43*, 413–431. http://dx.doi.org/10.1016/j.brat.2004.03.006.

Ehlers, A., Clark, D. M., Hackmann, A., McManus, F., Fennell, M., Herbert, C., et al. (2003). A randomized controlled trial of cognitive therapy, a self-help booklet, and repeated assessments as early interventions for posttraumatic stress disorder. *Archives of General Psychiatry, 60*, 1024–1032. http://dx.doi.org/10.1001/archpsyc.60.10.1024.

Ehlers, A., Hackmann, A., Steil, R., Clohessy, S., Wenninger, K., & Winter, H. (2002). The nature of intrusive memories after trauma: the warning signal hypothesis. *Behaviour Research and Therapy, 40*, 995–1002.

Engelhard, I., Arntz, A., & van den Hout, M. A. (2007). Low specificity of post-traumatic stress disorder symptoms: a comparison of patients with post-traumatic stress disorder, patients with other anxiety disorders, and individuals without psychopathology. *British Journal of Clinical Psychology, 46*, 449–456.

Finnie, P. S. B., & Nader, K. (2012). The role of metaplasticity mechanisms in regulating memory destabilization and reconsolidation. *Neuroscience and Biobehavioral Reviews, 36*, 1667–1707.

Foa, E. B., & Tolin, D. F. (2000). Comparison of the PTSD symptom scale–interview version and the clinician-administered PTSD scale. *Journal of Traumatic Stress, 13*, 181–191.

Foa, E. B., Riggs, D. S., Dancu, C. V., & Rothbaum, B. O. (1993). Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *Journal of Traumatic Stress, 6*, 459–473.

van Groenestijn, M. A. C., Akkerhuis, G. W., Kupka, R. W., Schneider, N., & Nolen, W. A. (1999). *Gestructureerd klinisch interview voor de vaststelling van DSM-IV as-I stoornissen (SCID-I)* [*Structured clinical interview for DSM-IV axis I disorders (SCID-I)*]. Lisse, The Netherlands: Swets & Zeitlinger.

Hagenaars, M. A., & Arntz, A. (2012). Reduced intrusion development after post-trauma imagery rescripting; an experimental study. *Journal of Behavior Therapy and Experimental Psychiatry, 44*, 808–814. http://dx.doi.org/10.1016/j.jbtep.2011.09.005.

Hawkins, N. G., Sanson-Fisher, R. W., Shakeshaft, A., D'Este, C., & Green, L. W. (2007). The multiple baseline design for evaluating population-based research. *American Journal of Preventive Medicine, 33*, 162–168.

Holmes, E. A., Arntz, A., & Smucker, M. R. (2007). Imagery rescripting in cognitive behaviour therapy: images, treatment techniques and outcomes. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 297–305. http://dx.doi.org/10.1016/j.jbtep.2007.10.007.

Kindt, M., Buck, N., Arntz, A., & Soeter, M. (2007). Perceptual and conceptual processing as predictors of treatment outcome in PTSD. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 491–506. http://dx.doi.org/10.1016/j.jbtep.2007.10.002.

Lobbestael, J., Leurgans, M., & Arntz, A. (2011). Inter-rater reliability of the structured clinical interview for DSM-IV axis I disorders (SCID I) and axis II disorders (SCID II). *Clinical Psychology and Psychotherapy, 18*, 75–79. http://dx.doi.org/10.1002/cpp.693.

Nilsson, J.-E., Lundh, L.-G., & Viborg, G. (2012). Imagery rescripting of early memories in social anxiety disorder: an experimental study. *Behaviour Research and Therapy, 50*, 387–392. http://dx.doi.org/10.1016/j.brat.2012.03.004.

Onghena, P. (2005). Single-case designs. In David C. Howell (Series Ed.) & B. S. Everitt (Vol. Ed.), *Encyclopedia of statistics in behavioral science*, Vol. 4, (pp. 1850–1854). Chichester: Wiley.

Onghena, P., & Edgington, E. S. (2005). Customization of pain treatments single-case design and analysis. *Clinical Journal of Pain, 21*, 56–68.

Vaage, A. B., Thomsen, P. H., Silove, D., Wentzel-Larsen, T., Van Ta, T., & Hauff, E. (2010). Long-term mental health of Vietnamese refugees in the aftermath of trauma. *British Journal of Psychiatry, 196*, 122–125. http://dx.doi.org/10.1192/bjp.bp.108.059139.

Vlaeyen, J., De Jong, J., Geilen, M., Heuts, P., & Van Breukelen, G. (2001). Graded exposure in vivo in the treatment of pain-related fear: a replicated single-case experimental design in four patients with chronic low back pain. *Behaviour Research and Therapy, 39*, 151–166.

Wheatley, J., Brewin, C. R., Patel, T., Hackmann, E. A., Wells, A., Fisher, P., et al. (2007). "I'll believe it when I can see it": imagery rescripting of intrusive sensory memories in depression. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 371–385. http://dx.doi.org/10.1016/j.jbtep.2007.08.005.

Wild, J., Hackmann, A., & Clark, D. M. (2007). When the present visits the past: updating traumatic memories in social phobia. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 386–401. http://dx.doi.org/10.1016/j.jbtep.2007.07.003.

Wild, J., Hackmann, A., & Clark, D. M. (2008). Rescripting early memories linked to negative images in social phobia: a pilot study. *Behavior Therapy, 39*, 47–56.

Wohlfarth, T. D., van den Brink, W., Winkel, F. W., & ter Smitten, M. (2003). Screening for posttraumatic stress disorder: an evaluation of two self-report scales among crime victims. *Psychological Assessment, 15*, 101–109.

Young, J. E., Klosko, J., & Weishaar, M. E. (2003). *Schema therapy: A practitioner's guide*. New York: Guilford.