Available online at www.sciencedirect.com

ScienceDirect

Behavior Therapy xxx (2022) xxx–xxx

Behavior Therapy

www.elsevier.com/locate/bt

Assessing Psychological Inflexibility Pertaining to Self in Patients With Posttraumatic Stress Disorder Using an Indirect Measure of (Nonassociative) Propositions

Petrus G.J. Janssen
Behavioural Science Institute, Radboud University
Psychotraumacentrum Zuid Nederland

Ciara McEnteggart
Perspectives Ireland

Dermot Barnes-Holmes
Ulster University

Elisa van Ee [1]
Behavioural Science Institute, Radboud University
Psychotraumacentrum Zuid Nederland

Jos Egger [1]
Centre of Excellence for Neuropsychiatry, Vincent van Gogh Institute for Psychiatry
Donders Institute for Brain, Cognition and Behaviour, Radboud University
Stevig Specialized and Forensic Care for People with Intellectual Disabilities, Dichterbij

Relational frame theory (RFT) is a modern behavioral account of human language and cognition, which focuses on relations or propositions, rather than associations, as core explanatory constructs. In an attempt to measure such propositions, RFT researchers have developed the implicit relational assessment procedure (IRAP). It has been argued that the size of an IRAP effect may provide a metric for psychological inflexibility. The current study aimed to determine whether psychological inflexibility, as measured by the self-focused Natural Language-IRAP (NL-IRAP), would be higher in a clinical sample of individuals with a diagnosis of PTSD ($N = 29$) when compared to a nonclinical sample. Subsequently, the study investigated whether the self-focused NL-IRAP could be used to predict the presence of a clinical diagnosis, using a ROC analysis. As predicted, higher levels of psychological inflexibility were observed for the clinical group. The self-focused NL-IRAP also correctly classified the presence of PTSD (AUC = 76%) with a sensitivity level of 79.3% and a specificity level of 59.2%. Overall, the use of the IRAP as a nonassociative clinical measure appears promising.

Keywords: relational frame theory; psychological inflexibility; propositional models; self; PTSD

This research was conducted with funding from the FWO. This work was supported by the Dutch Ministry of Defence.
Address correspondence to Petrus G.J. Janssen, Behavioral Science Institute, Radboud University, Postbus 9104, 6500 HE Nijmegen. e-mail: peter.janssen2@bsi.ru.nl.
[1] These authors contributed equally and share the senior authorship.

RELATIONAL FRAME THEORY (RFT) is a modern behavioral account of human language and cognition that focuses on relations or propositions, rather than associations, as core explanatory constructs. For RFT, the basis of human language is

Please cite this article as: Janssen, McEnteggart, Barnes-Holmes et al., Assessing Psychological Inflexibility Pertaining to Self in Patients With Posttraumatic Stress Disorder Using an Indirect Measure of (Nonassociative) Propositions, Behavior Therapy, https://doi.org/10.1016/j.beth.2022.05.002

derived relational responding (or relating), which accounts for the emergence of new learning that was not previously trained or directly prompted or instructed (Hayes et al., 2001). In an attempt to measure relating, RFT researchers developed a range of experimental preparations, including the implicit relational assessment procedure (IRAP; Barnes-Holmes et al., 2006; Kavanagh, Roelandt, et al., 2019), a computer-based task that requires participants to respond quickly and accurately in ways that are both consistent and inconsistent with their history of relational responding. The IRAP assumes that individuals respond more readily to history-consistent relations than history-inconsistent relations. The difference in response latency between history consistent and inconsistent responding is what generates the IRAP effect. Over the past decade, the IRAP has been used to examine relating in many clinical and nonclinical domains with robust effects (for a meta-analysis, see, for example, Vahey et al., 2015).

Many studies have attempted to employ the IRAP as a measure of implicit cognition, but one in which relatively rapid relational (or propositional) reasoning is targeted rather than "raw" associations in memory (for a detail discussion, see Hughes et al., 2011). Specifically, interpreting IRAP performances in purely associative terms, which is common using many other measures of implicit cognition, such as in the implicit association task (IAT; Greenwald et al., 1998), for example, has proven difficult because the IRAP requires that participants confirm or disconfirm specific propositions, rather than simply associating pairs of stimuli. Critically, the IRAP measures the speed and accuracy of the confirmatory and disconfirmatory responses and could therefore be seen as providing a "middle-ground" between *indirect* associative measures (e.g., the IAT) and *direct* self-report measures that ask participants to rate their agreement/disagreement with particular propositions. Other more recently developed methods, derived from the IRAP, have begun to emerge but they remain relatively small in number (e.g., De Houwer et al., 2015).

One of the potential benefits of using the IRAP in the clinical domain is that it appears to possess a reasonable level of predictive validity (Vahey et al., 2015) and it connects with the concept of psychological inflexibility that lies at the center of the theoretical model underlying acceptance and commitment therapy (ACT; Hayes et al., 2006). Specifically, it has been argued that the relative size of an IRAP effect, and the extent to which it is possible to manipulate it with laboratory and/or therapeutic interventions, may provide a metric for psychological inflexibility itself (Hussey & Barnes-Holmes, 2012; Kavanagh, Matthyssen, et al., 2019). For example, larger IRAP effects may be interpreted, in certain contexts, as evidence for higher levels of psychological inflexibility because participants find it difficult to reverse their response patterns across consistent versus inconsistent blocks of trials. For illustrative purposes, imagine an IRAP that was designed to target propositions concerning the self. One of the trial-types might ask a participant to respond to the proposition "I'm proud when I succeed in my exams." On consistent blocks of trials (i.e., consistent with common-sense expectations) participants would be required to confirm this statement as "True" (by pressing a key labelled "True"). On inconsistent blocks of trials, the opposite response pattern would be required (pressing a key labelled "False"). The assumption is that relatively large IRAP effects would indicate *low* flexibility because participants found it difficult to reverse their response patterns across consistent versus inconsistent blocks of trials (Murphy et al., 2019; O'Toole et al., 2009). In more concrete terms, a large IRAP effect in this context indicates that participants found it difficult to deny that they feel proud when they succeed.

Using a measure such as the IRAP to assess psychological inflexibility with regard to the self and how the self reacts to life events would certainly be relevant to recent arguments that psychological inflexibility may be a critical feature of a stable sense of self and psychological well-being. Or to put it another way, psychological inflexibility may be associated with psychological struggle (e.g., Barnes-Holmes et al., 2020; McEnteggart et al., 2017). The current study constitutes the first attempt to test this basic argument using both clinical and nonclinical samples. Specifically, the primary aim of the study was to measure the psychological inflexibility pertaining to the self using the IRAP with a clinical sample of individuals with a diagnosis of PTSD. This sample was employed because there is a growing body of research that supports a relationship between psychological trauma, the sense of self, and psychological suffering (see McEnteggart et al., 2017). Indeed, traumatized individuals often depict an instable and fragile sense of self (e.g., Berntsen & Rubin, 2006; Kashdan et al., 2006). The self-focused Natural Language-IRAP (NL-IRAP) in Kavanagh, Roelandt, et al. (2019) was employed in the current study. The self-focused NL-IRAP focused on self-based reactions to both positive and negative events and thus seemed directly relevant to the current population of individuals with

Please cite this article as: Janssen, McEnteggart, Barnes-Holmes et al., Assessing Psychological Inflexibility Pertaining to Self in Patients With Posttraumatic Stress Disorder Using an Indirect Measure of (Nonassociative) Propositions, Behavior Therapy, https://doi.org/10.1016/j.beth.2022.05.002

event-based traumatic histories. The key prediction was that larger IRAP effects will be observed for the clinical sample when compared to the nonclinical sample. Put simply, we predicted higher levels of psychological inflexibility in the clinical group. A second aim of the study was to determine if psychological inflexibility pertaining to the self, as measured by the self-focused NL-IRAP, could predict the presence of a clinical diagnosis, using a ROC analysis.

Method

PARTICIPANTS

Clinical Sample

Forty-eight participants with a clinical diagnosis of PTSD were recruited for the study, 10 females and 38 males. Participants ranged from 22–61 years old (M = 43.6) and were recruited via advertising and clinical referral from a psychotrauma centre in the Netherlands. Most participants experienced war-related combat trauma (70.8%), followed by childhood abuse (14.6%) and a combination of trauma types (e.g., physical or sexual abuse; 14.6%). All participants attended treatment for PTSD and were categorized as such based on the DSM-5 criteria using the Posttraumatic Stress Disorder Checklist (PCL-5; Weathers et al., 2013) and a clinical assessment. Each participant was paid an hourly rate of 10 euro.

Because participants sometimes failed to reach various performance criteria for the self-focused NL-IRAP (details provided subsequently), it was necessary to recruit more than the required 29 participants in order to yield an adequate dataset for analyses (see required sample size calculation below). Twelve of these participants were excluded from the study because they did not complete all stages of the self-focused NL-IRAP (relatively large attrition rates for IRAP studies employing clinical samples is not unusual because the task is perceived to be relatively challenging; for a meta-analysis, see Vahey et al., 2015). Therefore, 36 clinical participants successfully completed the study, 27 of these were male and 9 were female. No sex, age, PCL score or trauma distribution differences were found relative to the initial sample (p's > .05).

In the absence of a previous IRAP study employing the clinical sample recruited for the current research, we were guided by the results of the meta-analysis of IRAP effects in the clinical domain, indicating that a minimum of 29 is required to achieve a power of 0.8 for first order correlations (Vahey et al., 2015).

Nonclinical Sample

The data from a study employing the same IRAP and self-report measure (Study 2 in Kavanagh, Roelandt, et al., 2019) in a nonclinical university sample were used as a comparison sample. Participants were recruited through random convenience sampling. A total of 49 participants completed this study, 35 females and 14 males. Participants ranged from 18–49 years old (M = 24.5). Sample data for both the clinical and nonclinical groups are provided in Table 1.

MATERIALS AND APPARATUS

The study comprised the self-focused NL-IRAP (Barnes-Holmes et al., 2006) and the Community Assessment of Psychic Experiences (CAPE; Stefanis et al., 2002). The CAPE was employed because the role of the self has been implicated in psychotic-like experiences (Cicero et al., 2015; Savla et al., 2013) and has been used in numerous IRAP studies that have investigated the self in individuals who have experienced trauma and has been demonstrated to be a good predictor of psychological distress and IRAP performances (e.g., Kavanagh et al., 2018; Kavanagh, Matthyssen, et al., 2019; Kavanagh, Roelandt, et al., 2019; McEnteggart et al., 2016). The CAPE was also included in line with the methodology of Kavanagh, Roelandt, et al. (2019) and thus to assess whether this measure was able to predict group membership. Furthermore, this measure can readily be used in both clinical and nonclinical samples, thus offering a valid comparison measure for the two samples in the current study (e.g., Savla et al., 2013).

Self-Focused NL-IRAP

The self-focused NL-IRAP (Barnes-Holmes et al., 2006; Kavanagh, Roelandt, et al., 2019) required participants to respond to various statements about themselves (e.g., "My self-esteem increases if someone says I look good"). The self-focused NL-IRAP presented 16 statements, referring to an event (either positive or negative) *and* a positive or negative reaction to that event. The 16 statements were divided into four trial-types (see Figure 1). For example, consider the four statements: "My self-esteem increases if someone says I look good" (*Positive Event-Positive Reaction*); "I feel ugly if someone says I look good" (*Positive Event-Negative Reaction*); "I'm happy if a loved one dies" (*Negative Event-Positive Reaction*); and "If a loved one dies, I'm miserable" (*Negative Event-Negative Reaction*); see also Table 2. The response options "Yes" and "No" were presented at the bottom left- and right-hand

Please cite this article as: Janssen, McEnteggart, Barnes-Holmes et al., Assessing Psychological Inflexibility Pertaining to Self in Patients With Posttraumatic Stress Disorder Using an Indirect Measure of (Nonassociative) Propositions, Behavior Therapy, https://doi.org/10.1016/j.beth.2022.05.002

Table 1
Descriptive Statistics on all Variables for the Two Groups

Variable	Clinical Group ($n = 29$) %	Nonclinical Group ($n = 49$) %
Male	75	29
	M (SD)	*M (SD)*
Age	40.93 (8.08)	24.5 (5.14)
Self-focused NL-IRAP D_{IRAP}-scores		
Positive-Positive trial type	0.45 (0.33)	0.15 (0.27)
Positive-Negative trial type	0.39 (0.40)	0.15 (0.34)
Negative-Positive trial type	0.25 (0.41)	0.06 (0.32)
Negative-Negative trial type	0.43 (0.35)	0.25 (0.35)
CAPE		
Overall Frequency	1.84 (0.42)	1.72 (0.32)
Frequency of Positive Symptoms	1.41 (0.28)	1.43 (0.34)
Frequency of Negative Symptoms	2.24 (0.50)	1.94 (0.44)
Frequency of Depressive Symptoms	2.34 (0.61)	2.08 (0.52)
Overall Distress	1.75 (0.46)	2.15 (0.54)
Distress associated with Positive Symptoms	1.27 (0.29)	1.66 (0.44)
Distress associated with Negative Symptoms	2.06 (0.57)	2.07 (0.58)
Distress associated with Depressive Symptoms	2.43 (0.71)	2.54 (0.73)

Table 2
Natural Language Statements From the Self-Focused NL-IRAP

Trial-types	Stimuli
Positive Event – Positive Reaction	My self-esteem increases if someone says I look good. I feel liberated if my enemy dies. Winning the lottery makes me happy. I'm proud when I succeed in my exams.
Positive Event – Negative Reaction	I feel ugly if someone says I look good. I'm angry if my enemy dies. Winning the lottery disappoints me. It frustrates me if I succeed in my exams.
Negative Event – Positive Reaction	I'm happy if a loved one dies. Getting fines make me happy. Failing an exam is fantastic. I rejoice if someone I hate wins the lottery.
Negative Event – Negative Reaction	If a loved one dies, I'm miserable. Getting a fine makes me angry. Failing an exam is disappointing. It irritates me if someone I hate wins the lottery.

Note. Statements were presented to participants in Dutch. Trial type labels denote each of the two parts of the statement, but not necessarily the sequence in which they appeared in the statement.

corners on each trial. Further IRAP task and procedure details are described in more detail below.

CAPE

The CAPE (Stefanis et al., 2002) measures psychotic-like experiences in the general population. The scale consists of 42 items rated along three subscales: positive symptoms (20 items, e.g., "Do you ever feel as if there is a conspiracy against you?"), negative symptoms (14 items, e.g., "Do you ever feel that you experience few or no emotions at important events?") or depressive symptoms (eight items, e.g., "Do you ever feel sad?"). Each item is rated on two 4-point Likert scales from 0 (*never*) to 3 (*nearly always*) to indicate (1) the frequency of symptoms and (2) the

Please cite this article as: Janssen, McEnteggart, Barnes-Holmes et al., Assessing Psychological Inflexibility Pertaining to Self in Patients With Posttraumatic Stress Disorder Using an Indirect Measure of (Nonassociative) Propositions, Behavior Therapy, https://doi.org/10.1016/j.beth.2022.05.002

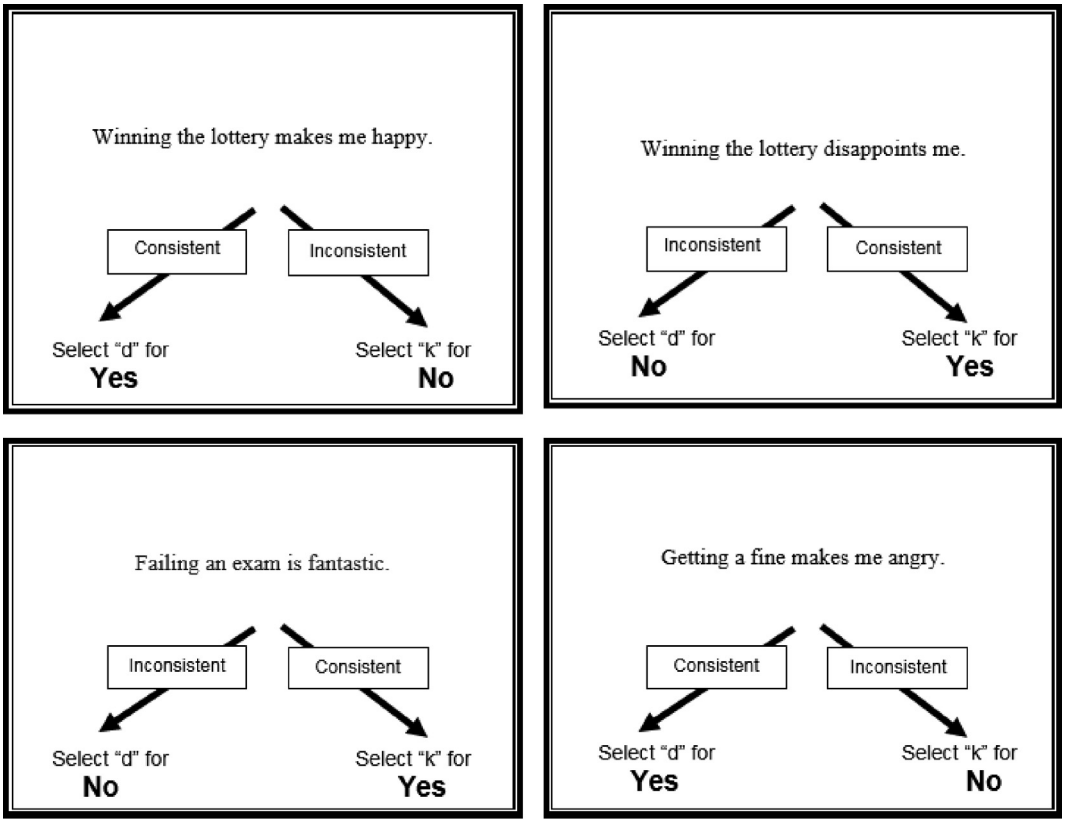

FIGURE 1 Examples of the Four Trial Types in the Self-Focused NL-IRAP: Positive Event-Positive Reaction, Positive Event-Negative Reaction, Negative Event-Positive Reaction, Negative Event-Negative Reaction. *Note.* The arrows and words *Consistent* and *Inconsistent* were not shown on-screen. Trial type labels denote each of the two parts of the statement, but not necessarily the sequence in which they appeared in the statement.

level of distress associated with each symptom. The CAPE provides overall frequency and distress scores of experiences, and total frequency and distress scores for each of the three subscales. In order to account for partial nonresponses, all scores are weighted for the number of valid answers per subscale (i.e., sum score divided by number of items completed). In all cases, higher scores indicate greater frequency or distress regarding symptoms, but there are no clinical cut-offs for this measure. The Dutch version was completed by participants. The scale has demonstrated adequate internal reliability (intraclass correlation coefficients) for the three subscales: positive = 0.63, negative = 0.64, and depression = 0.62 (Konings et al., 2006).

PROCEDURE

The current study was approved by the institutional review committee and by the Medical Ethical Committee of the University of Utrecht, the Netherlands (number 17/829). The clinical sample participated in a research room in a psychiatric facility. All participation was on an individual basis. As per the guidelines outlined in McEnteggart et al. (2017), the experimenter interacted with participants during all phases of the experiment. Participants were offered multiple opportunities to take breaks between the blocks of the self-focused NL-IRAP, which significantly extended the duration of the experiment. On average, these sessions (IRAP and CAPE administration) lasted between 1.5 and 2.5 hours (with regular breaks as requested) and all participation was completed in one session, in which the CAPE was completed in approximately 15 minutes. Informed consent was obtained from all participants. Each participant was exposed to the self-focused NL-IRAP and the CAPE, with the order of each counterbalanced across participants.

Self-Focused NL-IRAP

The IRAP consisted of a standardized procedure in line with Barnes-Holmes et al. (2006) and Kavanagh, Roelandt, et al. (2019). Additional information on the content and format of self-focused NL-IRAP can be found in Kavanagh, Roelandt, et al. (2019). The self-focused NL-

Please cite this article as: Janssen, McEnteggart, Barnes-Holmes et al., Assessing Psychological Inflexibility Pertaining to Self in Patients With Posttraumatic Stress Disorder Using an Indirect Measure of (Nonassociative) Propositions, Behavior Therapy, https://doi.org/10.1016/j.beth.2022.05.002

IRAP consisted of blocks of 32 trials (two presentations of the 16 statements) presented quasi-randomly. There were a maximum of eight pairs of practice blocks, followed by three pairs of test blocks. The content of the practice and test blocks were identical. It was particularly important in this IRAP to ensure that participants were responding to each of the statements *from their own perspective*. Hence, participants were instructed at the beginning of the self-focused NL-IRAP, as follows: "The program will present statements on the screen which refer to you. Please remember that when you see 'I' or 'me' on-screen, this refers to you (the participant)." On each trial, a self-related statement was presented in the middle of the screen (e.g., "I'm proud when I succeed in my exams"), with two response options ("Yes" and "No") at the bottom left and right of the screen. Participants were simply instructed to figure out, based on individual trial feedback, what the task involved. Participants responded on each trial using either the "d" key for the response option on the left or the "k" key for the response option on the right. The locations of the response options "Yes" and "No" alternated from trial to trial in a quasi-random order, such that they did not remain in the same left-right locations for more than three successive trials. The instruction "The previously correct and incorrect answers have been reversed" was presented between blocks of trials.

When participants selected the response option that was deemed correct within that block, an inter-trial interval of 400 ms was presented, after which the next trial occurred. When participants selected the response option that was deemed incorrect for that block, the stimuli remained on the screen and a red "X" appeared beneath the statement. The program proceeded to the 400 ms inter-trial interval (and next trial) only after the correct response option was selected. This pattern of trial presentations, with corrective feedback, continued until the entire block of 32 trials was presented. Trials were presented in a quasi-random order within each block, with the constraint that each statement was presented twice across the 32 trials. *Consistent* trial blocks required responding to the four trial-types that was deemed to be in accordance with positive events producing positive reactions and negative events producing negative reactions (i.e., *Positive Event-Positive Reaction/Yes, Positive Event-Negative Reaction/No, Negative Event-Positive Reaction/No,* and *Negative Event-Negative Event/Yes)*. *Inconsistent* trial blocks required responding to the four trial-types that was in accordance with positive events producing negative reactions and negative events producing positive reactions (i.e., *Positive Event-Positive Reaction/No, Positive Event-Negative Reaction/Yes, Negative Event-Positive Reaction/Yes,* and *Negative Event-Negative Reaction/No)*. The self-focused NL-IRAP always commenced with a consistent block of trials.

When participants completed each block of trials, the self-focused NL-IRAP program provided them with feedback on their performance during that block. The feedback consisted of a message informing them how accurately and how quickly they had responded overall during that block. The average speed of responding was calculated from stimulus onset to the first correct response across all 32 trials within the block. Participants were required to achieve a maximum median latency of no more than 8000 ms *on each trial-type*. The original Kavanagh, Roelandt, et al. (2019) study applied a median latency of 5000 ms for the nonclinical population, but with a clinical population, a longer response latency window was applied (see McEnteggart et al., 2017). It should be noted, however, that analyses of the latency data showed that the average latency for the clinical sample on the test blocks on the self-focused NL-IRAP was approximately 3000 ms.

Participants were also required to achieve a minimum accuracy of no less than 75%, at the trial-type level (i.e., no more than 2 errors were permitted per trial-type). If participants achieved both accuracy and latency criteria on any pair of practice blocks, they proceeded to the first pair of test blocks; if they failed on the eighth pair of practice blocks, participation in the experiment was terminated. A fixed set of six test blocks was presented with no accuracy or latency criteria required for participants to progress from one block to the next. However, percentage correct and median latency were again presented at the end of each block to encourage participants to maintain criterion-level responding from the practice blocks.

DATA-ANALYSIS

First, the self-focused IRAP and CAPE indices were examined using general linear model to probe potential differences between the clinical group relative to the nonclinical group. Second, it was examined whether scores on these indices could predict the presence of a clinical diagnosis. To investigate this, we conducted a Receiver Operator Characteristic (ROC) analysis in which the probability of a true positive, or a "hit" (i.e., sensitivity) is plotted against the probability of a false

Please cite this article as: Janssen, McEnteggart, Barnes-Holmes et al., Assessing Psychological Inflexibility Pertaining to Self in Patients With Posttraumatic Stress Disorder Using an Indirect Measure of (Nonassociative) Propositions, Behavior Therapy, https://doi.org/10.1016/j.beth.2022.05.002

positive or a "false alarm" (i.e., 1 minus specificity, see Fawcett, 2006). From this, the area under the curve (AUC) can be calculated, which is the statistical likelihood that a randomly chosen member of the "positive" group (in this case, the clinical group) will have a higher score than a randomly chosen member of the "negative" group (in this case, the nonclinical group). Therefore, a test with perfect ability to predict group membership would have an AUC of 100%, and a test with no ability to detect group membership would have an AUC of ~50%.

Results

SELF-FOCUSED NL-IRAP

Consistent with standard IRAP practice, mean response latencies for consistent and inconsistent blocks were initially divided according to trial-type and calculated for each participant (see Barnes-Holmes, Barnes-Holmes, et al., 2010). Based on the latency and accuracy criteria, four participants failed to complete the practice blocks (and did not proceed to the test blocks) on the self-focused NL-IRAP. Hence, all four datasets were excluded from further analyses. For the remaining participants, the same accuracy and latency criteria were applied in the test blocks, except that the criteria now applied across all six test blocks. This meant that no more than eight errors were permitted per trial-type across the six test blocks. Using these criteria, three participants failed to complete the self-focused NL-IRAP. All three datasets were excluded from further analyses, leaving the final number of datasets in the analyses at $N = 29$.

Consistent with the majority of published IRAP studies, D_{IRAP}-scores for the self-focused NL-IRAP were calculated for each of the four trial-types (see Table 1, Barnes-Holmes, Barnes-Holmes, et al., 2010), such that positive D_{IRAP}-scores during consistent blocks indicated responding "Yes" more quickly than "No" on *Positive Event-Positive Reaction* and on *Negative Event-Negative Reaction* trial-types, and responding "No" more quickly than "Yes" on *Positive Event-Negative Reaction* and on *Negative Event-Positive Reaction* trial-types. Negative D_{IRAP}-scores indicated the opposite pattern: "No" more quickly than "Yes" on *Positive Event-Positive Reaction* and on *Negative Event-Negative Reaction* trial-types, and responding "Yes" more quickly than "No" on *Positive Event-Negative Reaction* and on *Negative Event-Positive Reaction* trial-types. This scoring algorithm uses the same rationale as the D_{IAT} algorithm and is typically employed as it has proven to be effective at taking individual differences into account, such as age, motor skills and/or cognitive ability (e.g., Barnes-Holmes, Barnes-Holmes, et al., 2010; Greenwald et al., 2003; O'Toole & Barnes-Holmes, 2009; Vahey et al., 2015).

A 2 × 4 mixed repeated measures Analysis of Variance (ANOVA) was conducted for group and trial-type. There was a significant main effect for trial-type, $F(3, 74) = 6.876$, $p < .001$, $\eta^2 = .083$, and for group, $F(1, 76) = 14.285$, $p < .001$, $\eta^2 = .158$; the interaction was nonsignificant ($p > .05$). Post-hoc tests in the form of four independent *t*-tests indicated that three of the four trial-types differed significantly between the groups: *Positive Event-Positive Reaction*, $t(76) = -4.291$, $p < .001$, $d = 1.01$, 95% confidence interval (CI) [0.52, 1.49], *Positive Event-Negative Reaction*, $t(76) = -2.778$, $p = .007$, $d = 0.68$, CI [0.18, 1.12], and *Negative Event-Positive Reaction*, $t(76) = -2.374$, $p = .02$, $d = 0.56$, CI [0.09, 1.02].

ROC analyses for each of the four trial-types were conducted, and the D_{IRAP}-scores on the *Positive Event-Positive Reaction*, AUC = 0.76, $p < .001$, and *Positive Event-Negative Reaction* trial-types, AUC = 0.65, $p = 0.02$, were good predictors of the clinical group (see Figure 2). Using the better predictor of the two, a cut-off of .24 for the *Positive Event-Positive Reaction* D_{IRAP}-score yields a sensitivity level of 79.3% and a specificity level of 59.2% (see Figure 2). A cut-off of .10 for the *Positive Event-Negative Reaction* D_{IRAP}-score yields a sensitivity level of 79.3% and a specificity level of 55.1%.

CAPE

The CAPE weighted overall and subscale scores are provided in Table 1. A 2 × 2 multivariate ANOVA with overall frequency and distress was conducted for group. No group effect was found ($p > .05$). Additionally, ROC analyses of the CAPE data were also conducted and findings yielded non-significant results (all p's > .05).

Discussion

The current study was the first attempt to use the self-focused NL-IRAP as a measure of psychological inflexibility, with regard to self, in a clinically diagnosed population currently attending a specialized psychiatric facility for the diagnosis and treatment of PTSD. As predicted, larger IRAP effects were observed for the clinical sample when compared to the nonclinical sample, suggesting higher levels of psychological inflexibility in the clinical group. The second aim of the study was

Please cite this article as: Janssen, McEnteggart, Barnes-Holmes et al., Assessing Psychological Inflexibility Pertaining to Self in Patients With Posttraumatic Stress Disorder Using an Indirect Measure of (Nonassociative) Propositions, Behavior Therapy, https://doi.org/10.1016/j.beth.2022.05.002

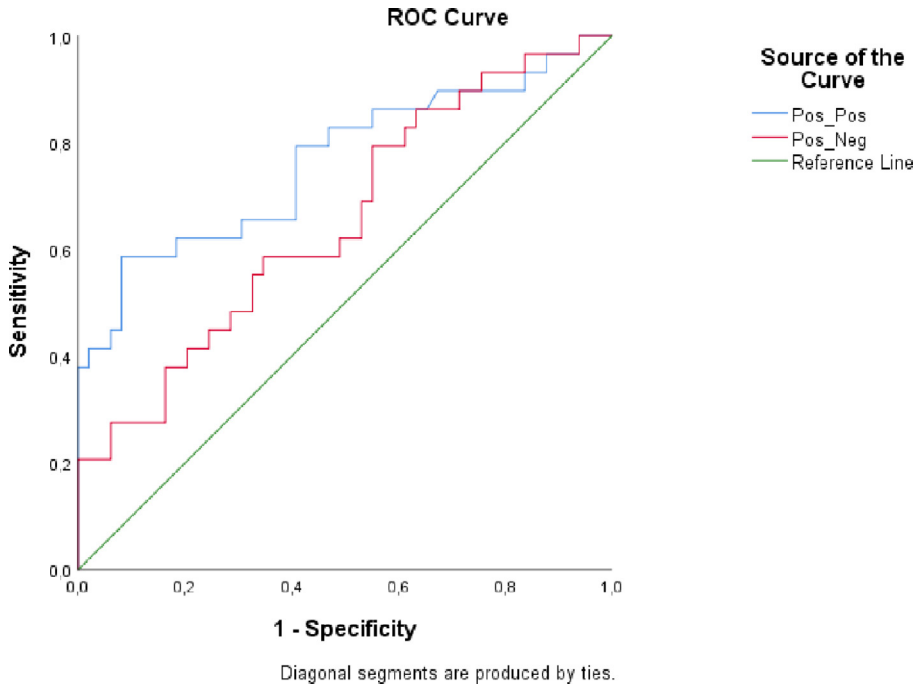

FIGURE 2 ROC Curve of the Positive Event-Positive Reaction and Positive Event-Negative Reaction trial-types.

to determine if psychological inflexibility, as measured by the self-focused NL-IRAP, could predict the presence of a clinical diagnosis, using a ROC analysis. Two trial-types (i.e., *Positive Event-Positive Reaction* and *Positive Event-Negative Reaction*) were significant predictors of such diagnosis. However, only the *Positive Event-Positive Reaction* trial-type correctly classified the group with an acceptable AUC of 76%, with a sensitivity level of 79.3% and a specificity level of 59.2%.

Interestingly, the CAPE failed to classify the clinical group from a nonclinical group at a statistically significant level. Whilst the CAPE was also included in line with the methodology of Kavanagh, Roelandt, et al. (2019) and thus to examine whether this measure was able to predict group membership, future studies could also employ alternative measures of psychological stress and well-being, including self-concept or trauma-related indices. On balance, the CAPE targeted internal psychological experiences, and the associated distress, and thus it was not a *direct* measure of how participants reacted to their own psychological content. In contrast, the self-focused NL-IRAP directly measured reaction times to propositions concerning (hypothetical) life events, which may have increased its utility in accurately differentiating the two groups. For example, perhaps some or indeed many of the participants in the clinical sample did not produce high scores on the CAPE, which would reduce its sensitivity for the whole group. A post-hoc analysis of the data did indeed indicate that over half of the participants failed to score above 1.74 (weighted score) on the CAPE frequency scale and thus there was considerable overlap between the distribution of scores between the clinical and nonclinical samples. Taken together, results indicated that the CAPE did not differentiate between the clinical and nonclinical samples, yet future research may also probe group differences using more direct trauma related indices.

In this context, it is interesting that measuring the ease with which participants react to their own psychological events (using the self-focused NL-IRAP) appeared to provide reasonable sensitivity in differentiating between clinical and nonclinical groups. The IRAP employed in the current study targeted propositions (with regard to self), but measured accuracy/latency rather than self-rating scales. This result highlights a possible advantage for propositional over associative models of human cognition in the clinical domain (see Smyth et al., 2008). At the very least, it suggests that it may be useful to conceptualize clinically relevant measures as lying on a continuum. For example, at one end of the continuum, an evaluative priming task might be considered as a task that aims to measure associations in memory, with a self-report scale at the other end aiming to measure propositions concerning the clinically relevant behaviors; the IRAP, which presents propositions but measures reaction time rather than ratings, could be seen as lying "mid-way" along the continuum between these two extremes.

Please cite this article as: Janssen, McEnteggart, Barnes-Holmes et al., Assessing Psychological Inflexibility Pertaining to Self in Patients With Posttraumatic Stress Disorder Using an Indirect Measure of (Nonassociative) Propositions, Behavior Therapy, https://doi.org/10.1016/j.beth.2022.05.002

Another interesting finding that emerged from the study was that the only trial-types that classified the clinical group from the nonclinical group were the trial-types pertaining to positive events. This finding suggests that the inflexibility involved in reactions to positive and negative events for the clinical group may not be due to a generic or global level of inflexibility, as measured by the self-focused NL-IRAP. For instance, it could be argued that the valence of (or sensitivity to) positive events may have been greater for the clinical group employed in the current study due to their direct experience of traumatic events. Intuitively, of course, one might have expected the difference between the clinical and nonclinical samples to have been focused on the *negative* events (given that they were diagnosed with PTSD), but it appears that this was not the case. On balance, perhaps a history of PTSD served to create a type of positive hyper-sensitivity for positive (over negative) events (Berntsen et al., 2011; Elman et al., 2018; Walker et al., 2003); in addition, it may be that negative reactions to negative life-events are so strong in both normative and clinical populations that their reactions to such events will fail to discriminate between the two groups. With that said, the difference between the two groups on the *Negative-Negative* trial-type were in the intuitively expected direction (i.e., the clinical group produced a larger IRAP effect), and thus perhaps a larger sample, yielding increased statistical power, would have produced a significant result.

One potential limitation in the study design was the possible presence of unmeasured factors that may have influenced IRAP performances, including comorbid condition(s) within the clinical sample (e.g., acquired brain injury, substance abuse). Future research could address this as well as the role of other factors of interest that may or may not affect IRAP performances, such as differences related to trauma history. Specifically, it may have been interesting to record and assess the nature and extent of each individual's PTSD-related traumatic experiences in order to investigate any potential differences in the IRAP effects. For instance, would individuals with fewer traumatic experiences demonstrate less sensitivity to positive events, or vice versa? Moreover, the nonclinical group was not assessed or screened on the presence of psychiatric diagnoses or the extent of clinical levels of psychopathology. The latter would be of interest, given that (sub-)clinical levels of mental health complaints are prevalent in the general and student populations (e.g., Ayuso-Mateos et al., 2010; Beiter et al., 2015; Remes et al., 2016). Indeed, future studies should incorporate assessments of psychiatric complaints, as particular conditions such as anxiety symptomatology may negatively affect the level of psychological flexibility (Tirch et al., 2012). Lastly, following Kavanagh, Roelandt, et al. (2019), the CAPE was employed as a measure of psychological distress. However, given that the CAPE involves an index for psychotic-like experiences and distress, future studies should additionally include other trauma-related indices to assess clinical levels associated with psychotrauma more directly.

Future research could further address the clinical applicability of the IRAP by optimizing and probing different task performance parameters, such as using less stringent accuracy and latency thresholds, thereby potentially reducing attrition rates—which are common and expected in both clinical and nonclinical samples (see, for example, Vahey et al., 2015). However, relatively strict performance criteria seems to be a prerequisite for the IRAP data to be valid (i.e., inducing rapid relational responding; Barnes-Holmes et al., 2006; Kavanagh, Roelandt, et al., 2019) and reliable (e.g., shorter latencies have shown to improve reliability properties; Barnes-Holmes, Murphy, et al., 2010; Drake et al., 2016; Golijani-Moghaddam et al., 2013). Additionally, McEnteggart et al. (2017) provides concrete recommendations for using the IRAP in order to circumvent the practical difficulties associated with the task and to further reduce attrition, particularly within a clinical population (e.g., such as initial exposure to the IRAP to familiarize with the procedure and providing additional task instructions). Moreover, future studies could examine whether the predictive properties of the IRAP as a clinical tool could be further increased by including IRAP content tailored to specific clinical populations of interest, such as using trauma related IRAP statements in PTSD samples. Nevertheless, whilst future research should further examine IRAP performances in traumatized individuals and other forms of psychopathology in clinical contexts, the current findings suggest that use of the IRAP as a nonassociative clinical measure appears promising.

Conflict of Interest Statement

The authors declare that there are no conflicts of interest.

References

Ayuso-Mateos, J. L., Nuevo, R., Verdes, E., Naidoo, N., & Chatterji, S. (2010). From depressive symptoms to depressive disorders: The relevance of thresholds. *The British Journal of Psychiatry, 196*(5), 365–371. https://doi.org/10.1192/bjp.bp.109.071191.

Please cite this article as: Janssen, McEnteggart, Barnes-Holmes et al., Assessing Psychological Inflexibility Pertaining to Self in Patients With Posttraumatic Stress Disorder Using an Indirect Measure of (Nonassociative) Propositions, Behavior Therapy, https://doi.org/10.1016/j.beth.2022.05.002

Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2020). Updating RFT (more field than frame) and its implications for process-based therapy. *The Psychological Record, 70*(4), 1–20. https://doi.org/10.1007/s40732-019-00372-3.

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist, 32*(7), 169–177. https://doi.org/10.1007/BF03395684.

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the implicit relational assessment procedure (IRAP) and the relational elaboration and coherence (REC) model. *The Psychological Record, 60*(3), 527–542. https://doi.org/10.1007/BF03395726.

Barnes-Holmes, D., Murphy, A., Barnes-Holmes, Y., & Stewart, I. (2010). The Implicit Relational Assessment Procedure: Exploring the impact of private versus public contexts and the response latency criterion on pro-white and anti-black stereotyping among white Irish individuals. *The Psychological Record, 60*(1), 57–79. https://doi.org/10.1007/BF03395694.

Beiter, R., Nash, R., McCrady, M., Rhoades, D., Linscomb, M., Clarahan, M., & Sammut, S. (2015). The prevalence and correlates of depression, anxiety, and stress in a sample of college students. *Journal of Affective Disorders, 173*, 90–96. https://doi.org/10.1016/j.jad.2014.10.054.

Berntsen, D., & Rubin, D. C. (2006). The centrality of event scale: A measure of integrating a trauma into one's identity and its relation to post-traumatic stress disorder symptoms. *Behaviour Research and Therapy, 44*(2), 219–231. https://doi.org/10.1016/j.brat.2005.01.009.

Berntsen, D., Rubin, D. C., & Siegler, I. C. (2011). Two versions of life: Emotionally negative and positive life events have different roles in the organization of life story and identity. *Emotion, 11*(5), 1190–1201. https://doi.org/10.1037/a0024940.

Cicero, D. C., Docherty, A. R., Becker, T. M., Martin, E. A., & Kerns, J. G. (2015). Aberrant salience, self-concept clarity, and interview-rated psychotic-like experiences. *Journal of Personality Disorders, 29*(1), 79–99. https://doi.org/10.1521/pedi_2014_28_150.

De Houwer, J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). The relational responding task: Toward a new implicit measure of beliefs. *Frontiers in Psychology, 6*, 319. https://doi.org/10.3389/fpsyg.2015.00319.

Drake, C. E., Seymour, K. H., & Habib, R. (2016). Testing the IRAP: Exploring the reliability and fakability of an idiographic approach to interpersonal attitudes. *The Psychological Record, 66*(1), 153–163. https://doi.org/10.1007/s40732-015-0160-1.

Elman, I., Upadhyay, J., Langleben, D. D., Albanese, M., Becerra, L., & Borsook, D. (2018). Reward and aversion processing in patients with post-traumatic stress disorder: Functional neuroimaging with visual and thermal stimuli. *Translational Psychiatry, 8*(1), 1–15. https://doi.org/10.1038/s41398-018-0292-6.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.

Golijani-Moghaddam, N., Hart, A., & Dawson, D. L. (2013). The implicit relational assessment procedure: Emerging reliability and validity data. *Journal of Contextual Behavioral Science, 2*(3–4), 105–119. https://doi.org/10.1016/j.jcbs.2013.05.002.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464–1480. https://doi.org/10.1037//0022-3514.74.6.1464.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. Plenum.

Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behaviour Research and Therapy, 44*(1), 1–25. https://doi.org/10.1016/j.brat.2005.06.006.

Hughes, S., Barnes-Holmes, D., & De Houwer, J. (2011). The dominance of associative theorizing in implicit attitude research: Propositional and behavioral alternatives. *The Psychological Record, 61*(3), 465–496. https://doi.org/10.1007/BF03395772.

Hussey, I., & Barnes-Holmes, D. (2012). The implicit relational assessment procedure as a measure of implicit depression and the role of psychological flexibility. *Cognitive and Behavioral Practice, 19*(4), 573–582. https://doi.org/10.1016/j.cbpra.2012.03.002.

Kashdan, T. B., Uswatte, G., Steger, M. F., & Julian, T. (2006). Fragile self-esteem and affective instability in posttraumatic stress disorder. *Behaviour Research and Therapy, 44*(11), 1609–1619. https://doi.org/10.1016/j.brat.2005.12.003.

Kavanagh, D., Barnes-Holmes, Y., Barnes-Holmes, D., McEnteggart, C., & Finn, M. (2018). Exploring differential trial-type effects and the impact of a read-aloud procedure on deictic relational responding on the IRAP. *The Psychological Record, 68*(2), 163–176. https://doi.org/10.1007/s40732-018-0276-1.

Kavanagh, D., Matthyssen, N., Barnes-Holmes, Y., Barnes-Holmes, D., McEnteggart, C., & Vastano, R. (2019). Exploring the use of pictures of self and other in the IRAP: Reflecting upon the emergence of differential trial type effects. *International Journal of Psychology and Psychological Therapy, 19*(3), 323–336. https://doi.org/10.1037/0022-3514.63.4.596.

Kavanagh, D., Roelandt, A., Raemdonck, L., Barnes-Holmes, Y., Barnes-Holmes, D., & McEnteggart, C. (2019). The on-going search for perspective-taking IRAPs: Exploring the potential of the natural language-IRAP. *The Psychological Record, 69*(2), 291–314. https://doi.org/10.1007/s40732-019-00333-w.

Konings, M., Bak, M., Hanssen, M., Van Os, J., & Krabbendam, L. (2006). Validity and reliability of the CAPE: A self-report instrument for the measurement of psychotic experiences in the general population. *Acta Psychiatrica Scandinavica, 114*(1), 55–61. https://doi.org/10.1111/j.1600-0447.2005.00741.x.

McEnteggart, C., Barnes-Holmes, Y., & Adekuoroye, F. (2016). The effects of a voice hearing simulation on implicit fear of voices. *Journal of Contextual Behavioral Science, 5*(3), 154–159. https://doi.org/10.1016/j.jcbs.2016.06.003.

McEnteggart, C., Barnes-Holmes, Y., Mellon, M., van Dongen, L., Hendriks, A., Barnes-Holmes, D., Egger, J., & Harte, C. (2017). Recommendations for using the IRAP with a medicated in-patient population with a diagnosis of

Please cite this article as: Janssen, McEnteggart, Barnes-Holmes et al., Assessing Psychological Inflexibility Pertaining to Self in Patients With Posttraumatic Stress Disorder Using an Indirect Measure of (Nonassociative) Propositions, Behavior Therapy, https://doi.org/10.1016/j.beth.2022.05.002

psychosis. *International Journal of Psychology & Psychological Therapy, 17*(3), 317–325.

Murphy, C., Lyons, K., Kelly, M., Barnes-Holmes, Y., & Barnes-Holmes, D. (2019). Using the Teacher IRAP (T-IRAP) interactive computerized programme to teach complex flexible relational responding with children with diagnosed autism spectrum disorder. *Behavior Analysis in Practice, 12*(1), 52–65. https://doi.org/10.1007/s40617-018-00302-9.

O'Toole, C., & Barnes-Holmes, D. (2009). Three chronometric indices of relational responding as predictors of performance on a brief intelligence test: The importance of relational flexibility. *The Psychological Record, 59*(1), 119–132. https://doi.org/10.1007/BF03395652.

O'Toole, C., Barnes-Holmes, D., Murphy, C., O'Connor, J., & Barnes-Holmes, Y. (2009). Relational flexibility and human intelligence: Extending the remit of Skinner's Verbal Behavior. *International Journal of Psychology and Psychological Therapy, 9*(1), 1–17.

Remes, O., Brayne, C., Van Der Linde, R., & Lafortune, L. (2016). A systematic review of reviews on the prevalence of anxiety disorders in adult populations. *Brain and Behavior, 6*(7), 1–33. https://doi.org/10.1002/brb3.497.

Savla, G. N., Vella, L., Armstrong, C. C., Penn, D. L., & Twamley, E. W. (2013). Deficits in domains of social cognition in schizophrenia: A meta-analysis of the empirical evidence. *Schizophrenia Bulletin, 39*(5), 979–992. https://doi.org/10.1093/schbul/sbs080.

Smyth, S., Barnes-Holmes, D., & Barnes-Holmes, Y. (2008). Acquired equivalence in human discrimination learning: The role of propositional knowledge. *Journal of Experimental Psychology: Animal Behavior Processes, 34*(1), 167–177. https://doi.org/10.1037/0097-7403.34.1.167.

Stefanis, N. C., Hanssen, M., Smirnis, N. K., Avramopoulos, D. A., Evdokimidis, I. K., Stefanis, C. N., Verdoux, H., & Van Os, J. (2002). Evidence that three dimensions of psychosis have a distribution in the general population. *Psychological Medicine, 32*(2), 347–358. https://doi.org/10.1017/S0033291701005141.

Tirch, D. D., Leahy, R. L., Silberstein, L. R., & Melwani, P. S. (2012). Emotional schemas, psychological flexibility, and anxiety: The role of flexible response patterns to anxious arousal. *International Journal of Cognitive Therapy, 5*(4), 380–391. https://doi.org/10.1521/ijct.2012.5.4.380.

Vahey, N. A., Nicholson, E., & Barnes-Holmes, D. (2015). A meta-analysis of criterion effects for the Implicit Relational Assessment Procedure (IRAP) in the clinical domain. *Journal of Behavior Therapy and Experimental Psychiatry, 48*, 59–65. https://doi.org/10.1016/j.jbtep.2015.01.004.

Walker, W. R., Skowronski, J. J., & Thompson, C. P. (2003). Life is pleasant—and memory helps to keep it that way! *Review of General Psychology, 7*(2), 203–210. https://doi.org/10.1037/1089-2680.7.2.203.

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). The PTSD Checklist for DSM-5 (pcl-5). *Scale available from the National Center for PTSD at* www.ptsd.va.gov, *10*(4).

Received: February 11, 2021
Accepted: May 5, 2022
Available online: xxxx

Please cite this article as: Janssen, McEnteggart, Barnes-Holmes et al., Assessing Psychological Inflexibility Pertaining to Self in Patients With Posttraumatic Stress Disorder Using an Indirect Measure of (Nonassociative) Propositions, Behavior Therapy, https://doi.org/10.1016/j.beth.2022.05.002