



The social cognitive appraisal of trustworthiness in individuals with dimensional levels of post-traumatic stress symptoms: a translational study

Tanya Saraiya^{a,b}, Dominic Fareri^a, Teresa López-Castro^b, Denise Hien^c, Eric Fertuck^b and Robert Melara^b

^aDerner School of Psychology, Adelphi University, Garden City, NY, USA; ^bPsychology Department, The City College of New York, CUNY, New York, NY, USA; ^cCenter of Alcohol Studies, Graduate School of Applied and Professional Psychology, Rutgers University, Piscataway, NJ, USA

ABSTRACT

Background: Misappraisals in evaluating the trustworthiness of others may be one mechanism contributing to the interpersonal difficulties individuals with posttraumatic stress disorder (PTSD) face.

Objective: This study used a translational experimental design to examine the behavioural and neural correlates underlying the appraisal of facial stimuli morphed on dimensions of trustworthiness across three groups: individuals with high posttraumatic stress symptoms (HPTS), low posttraumatic stress symptoms (LPTS), and healthy controls (HC).

Methods: Participants (N = 70) rated how trustworthy to untrustworthy they perceived three facial morphs (trustworthy, neutral, and untrustworthy) while undergoing electroencephalography (EEG).

Results: Behavioural results showed that the HPTS group rated the untrustworthy morph as more untrustworthy than the HC group ($\beta = 0.20$, $SE = .07$, 95% CI [0.06, 0.33], $z = 2.88$, $p = .004$). The HPTS group also showed no variation in response time across morphs ($\chi^2(2) = 0.92$, $p = 0.63$), while the LPTS and HC groups did ($\chi^2(2) = 9.60$, $p = .008$; $\chi^2(2) = 23.62$, $p < .001$). EEG data revealed significant group by morph interactions at the N170 latency and the Vertex Positive Potential (VPP): the HPTS and LPTS identified the untrustworthy morph faster than the HCs, but diverged to the degree to which they encoded each facial morph.

Conclusions: Taken together our results suggest that HPTS individuals demonstrate an early attentional avoidance of faces morphed on dimensions of trustworthiness. This early, pre-conscious, avoidance may be one mechanism contributing to the miscalculations individuals with PTSD make in interpersonal situations.

El Juicio socio-cognitivo de la confiabilidad en individuos con distintos niveles dimensionales de síntomas de estrés postraumático: un estudio traslacional

Antecedentes: Los errores de apreciación al evaluar la confiabilidad de los demás pueden ser un mecanismo que contribuye a las dificultades interpersonales que enfrentan las personas con trastorno de estrés postraumático (TEPT).

Objetivo: Este estudio utilizó un diseño experimental traslacional para examinar los correlatos conductuales y neurales subyacentes a la evaluación de los estímulos faciales transformados en dimensiones de confiabilidad en tres grupos: individuos con síntomas de estrés postraumático alto (SEPA), síntomas de estrés postraumático bajo (SEPBA) y controles sanos (CS)

Métodos: Los participantes (N = 70) calificaron cuán confiables a no confiables percibieron tres transformaciones faciales (confiables, neutrales y no confiables) mientras se sometían a electroencefalografía (EEG).

Resultados: Los resultados de comportamiento mostraron que el grupo SEPA calificó la transformación no confiable como más confiable que el grupo CS ($\beta = 0.20$, $SE = .07$, IC 95% [0.06, 0.33], $z = 2.88$, $p = .004$). El grupo SEPA tampoco mostró variación en el tiempo de respuesta entre las transformaciones. (Common.EditSubmissionSteps.Transform.EquationText (2) = 0.92, $p = 0.63$), mientras que los grupos SEPBA y CS sí lo hicieron (Common.EditSubmissionSteps.Transform.EquationText (2) = 9.60, $p = .008$; Common.EditSubmissionSteps.Transform.EquationText (2) = 23.62, $p < .001$). Los datos del EEG revelaron interacciones significativas de grupo por transformación en la latencia N170 y el potencial positivo de vértice (PPV): el SEPA y el SEPBA identificaron la transformación no confiable más rápido que los CS, pero divergieron en el grado en que codificaron cada transformación facial.

Conclusiones: Tomados en conjunto, nuestros resultados sugieren que las personas con SEPA demuestran una evitación temprana de las caras transformadas en dimensiones de confiabilidad. Esta evitación temprana y pre-consciente puede ser un mecanismo que contribuye a los errores de cálculo que las personas con TEPT hacen en situaciones interpersonales.

ARTICLE HISTORY

Received 25 July 2019
Revised 24 September 2019
Accepted 9 October 2019

KEYWORDS

Trauma; social cognition; EEG; VPP; N170; trustworthiness

PALABRAS CLAVES

trauma; cognición social; EEG; VPP; N170; confiabilidad

关键词

创伤; 社会认知; EEG; VPP; N170; 可信度

HIGHLIGHTS

- Individuals with higher PTSD symptoms appraise untrustworthy faces as more untrustworthy than participants with lower PTSD symptoms and healthy controls.
- Neural data suggests that participants with higher PTSD symptoms encode trustworthy and untrustworthy faces into early preconscious memory less than participants with lower PTSD symptoms.
- Overall, those with higher PTSD symptoms appear to exhibit an attentional avoidance of trustworthy and untrustworthy faces, which leads to misappraisals in the assessment of trustworthiness.

不同创伤后应激症状水平个体对于可信度的社会认知评估:一项转化研究

背景:对他人可信度的误判可能是一种导致创伤后应激障碍 (PTSD) 人群人际交往困难的机制。

目标:本研究采用转化实验设计,考查了在高创伤后应激症状组 (HPTS), 低创伤后应激症状组 (LPTS) 和健康对照组 (HC) 这三组人中,涉及评估可信度维度变形的人脸刺激的行为和神经相关性。

方法:70名参与者在扫描脑电图 (EEG) 时评估他们对三种变形人脸 (可信的, 中立的和不可信的) 的信任程度。

结果:行为结果表明, HPTS组对不可信变形人脸的评价比HC组更高 ($\beta = 0.20$, $SE = .07$, 95% CI [0.06, 0.33], $z = 2.88$, $p = .004$)。同时HPTS组对不同人脸的反应时无差异 ($X^2(2) = 0.92$, $p = 0.63$), 而LPTS和HC组则有此差异 ($X^2(2) = 9.60$, $p = .008$; $X^2(2) = 23.62$, $p < .001$)。脑电数据揭示了在N170潜伏期和顶正电位 (VPP) 中组别与变形人脸间显著的交互作用:HPTS和LPTS组识别不可信人脸的速度比HC组快,但在编码各变形人脸的程度上有分歧。

结论:综上所述,我们的结果表明, HPTS组个体表现出对于可信度维度变形人脸的早期注意回避。这种早期的,前意识的回避可能是导致PTSD患者人际交往中做出误判的一种机制。

Social support is one of the largest buffers to developing PTSD after traumatic exposure (Dworkin, Ullman, Stappenbeck, Brill, & Kaysen, 2018; Meis et al., 2019; Ozer, Best, Lipsey, & Weiss, 2003). It reduces PTSD symptoms on a daily basis (Dworkin et al., 2018) and facilitates the formation of therapeutic alliances, which further promote recovery from PTSD (Keller, Zoellner, & Feeny, 2010). However, developing social bonds is strikingly difficult for an individual with PTSD. PTSD is marked by difficulties in interpersonal functioning (e.g. Campbell & Renshaw, 2018; Cloitre, Miranda, Stovall-McClough, & Han, 2005; Maercker & Horn, 2013; Stevens & Jovanovic, 2018) that lead to the long-term erosion of social support (King, Taft, King, Hammond, & Stone, 2006), loneliness, and isolation, further diminishing quality of life (Holt-Lunstad, Smith, Baker, Harris, & Stephenson, 2015). Moreover, the cardinal symptoms of PTSD – re-experiencing, numbing, and hyperarousal – place a significant burden on the ability to engage socially with others. The PTSD sufferer may withdraw and avoid others in their attempts to manage the distress and disturbance such symptoms yield. Thus, while relationships can enhance PTSD prognosis, they are difficult for an individual with PTSD to develop and maintain. This paradox calls for necessary intervention, first attained by understanding how individuals with PTSD perceive their social environments. This study assesses one social cognitive factor implicated in relationship formation – the appraisal of trustworthiness – through an experimental laboratory study with electroencephalography (EEG) among participants with dimensional levels of PTSD symptoms: high posttraumatic stress symptoms (HPTS), low posttraumatic stress symptoms (LPTS), and healthy controls with no trauma exposure (HC).

Social cognitive and information processing theories (e.g. Frith, 2008; Stevens & Jovanovic, 2018) identify the appraisal of trust as an early attentional process that has downstream effects on

behaviour. Three sequential components have been delineated: subjective, antecedent and behavioural. First, individuals assess (a) *subjective trust*, the appraisal of facial features which may indicate trustworthiness. This is followed by (b) *antecedent trust*, the psychological factors about the individual which imply trustworthiness, which then leads to (c) *behavioural trust*, where the perceiver engages in actions that denote feelings of trust in the other (Bellucci, Chernyak, Goodyear, Eickhoff, & Krueger, 2017; Suzuki, Misaki, Krueger, & Bodurka, 2015). In PTSD, this may present as an individual meeting a stranger at a party and inaccurately appraising their smile or gaze as trustworthy (*subjective trust*), learning about the personality of the stranger – for instance that he or she engages in risky behaviours – and still continuing to appraise the stranger as trustworthy (*antecedent trust*), followed by acting upon this formulated appraisal of trustworthiness by choosing to share something about themselves with this stranger (*behavioural trust*). To date, within trauma and PTSD literatures, behavioural trust has been the primary focus of one research group. Across a number of studies employing economic trust games and model-based fMRI, evidence has suggested that adolescent girls with histories of assault-related trauma are more likely to anticipate negative outcomes, become more desensitized to negative social outcomes, show difficulty discriminating between positive and negative social cues, and fail to assimilate new social information into their decision making (Cisler et al., 2015; Lenow, Cisler, & Bush, 2018; Lenow, Scott Steele, Smitherman, Kilts, & Cisler, 2014). Moreover, a follow-up analysis showed that these adolescents were more likely to misidentify the most trustworthy face in a trust game relative to controls (Lenow et al., 2018), alluding to deficits in subjective trust, although this was not the main aim of the authors' investigation.

The present investigation adds to this literature by specifically examining differences at the level of subjective trust. One advantageous way to elucidate brain-behavioural mechanisms underlying differences in the assessment of faces is to adopt a social cognitive affective neuroscience (SCAN) framework that combines behavioural laboratory tasks with EEG recordings (Lanius, Bluhm, & Frewen, 2011). EEG can capture early nonconscious cognitive, attentional, and perceptual processes when participants appraise faces (Read & Innis, 2017). Recordings can elucidate what neural mechanisms precede behaviour, and how neural activation changes over time when appraising a face. Of extant SCAN studies, most have focused on assessing the neural processing of angry, threatening, or neutral faces. Such studies have identified mixed findings in the positive amplitudes of P50, P200, and P300 event-related potentials (ERPs; for a review see, Karl, Malta, & Maercker, 2006; Lobo et al., 2015), a family of ERPs associated with arousal and the filtering of relevant and irrelevant auditory and visual stimuli in the environment (Karl et al., 2006). Solely one study, to date, has examined the appraisal of trustworthy facial stimuli (i.e. subjective trust) among individuals with PTSD using a SCAN-based framework, albeit without EEG recording. Fertuck et al. (2016) examined how facial stimuli morphed on dimensions of trustworthiness were perceived by individuals with PTSD, trauma-exposed healthy controls, and healthy controls with no trauma exposure. Contrary to the authors' hypotheses and Lenow et al.'s (2018) findings, individuals with PTSD rated faces as *more* trustworthy than trauma-exposed healthy controls, identifying misappraisals of subjective trust. The authors posited that such a bias may be a mechanism underlying the high rates of revictimization found among individuals with PTSD (Messman-Moore, Walsh, & DiLillo, 2010), where higher appraisals of trustworthiness may lead to entry into precarious situations or relationships.

Although there is limited evidence on ERPs associated with trustworthy appraisals among trauma-exposed participants, other ERP literature on trustworthy and happy facial stimuli has shown three components--the N170, the vertex positive potential (VPP), and the negative slow wave (NSW)--to respectively represent early, middle, and late components in the appraisal of trustworthiness. Specifically, the N170 has been associated with the early detection and encoding of the emotional valence of faces (Bentin, Allison, Puce, Perez, & McCarthy, 1996), where more negative amplitudes suggest faster, automatic processing (Chu, Bryant, Gatt, & Harris, 2016; Dzhelyova, Perrett, & Jentsch, 2012; Felmingham, Bryant, &

Gordon, 2003). The VPP, a related positive, middle ERP component at 200 milliseconds (ms), has also been associated with facial encoding (Klimova, Bryant, Williams, & Louise Felmingham, 2013; MacNamara, Post, Kennedy, Rabinak, & Phan, 2013), and literature has suggested that the N170 and VPP are related components that originate in the occipito-temporal cortex (Joyce & Rossion, 2005). Lastly, the NSW is a novel ERP from 800–1000 ms over frontocentral EEG sites, where positive amplitudes have been associated with negative images (Tso, Chiu, King-Casas, & Deldin, 2011) and negative amplitudes have been associated with trustworthy faces (Jessen & Grossmann, 2019), suggesting novelty detection. By examining these three components, we aimed to characterize differences in the early, middle, and late appraisal of subjective trust among individuals with trauma exposure and posttraumatic stress symptoms.

This study aimed to understand how individuals with dimensional levels of traumatic stress symptoms perceived facial stimuli morphed on dimensions of trustworthiness while undergoing EEG acquisition in order to elucidate the brain-behavioural mechanisms mediating the social cognitive appraisal of trustworthiness. A dimensional sample was chosen in order to probe the differences between trauma-exposed individuals who developed full PTSD by the time of study assessment and those who have coped in such a way that prevent such diagnostic progression, and thus present with milder symptoms of PTSD at the time of the study. Our first aim was to test the differences in the subjective appraisal (i.e. ratings) and reaction times to facial stimuli morphed on dimensions of trustworthiness across three groups of participants: high posttraumatic stress symptoms (HPTS), low posttraumatic stress symptoms (LPTS), and healthy controls (HC). We hypothesized that the HPTS group would rate all faces as more trustworthy than the LPTS group and that there would be no group differences in reaction time (cf. Fertuck et al., 2016). Our second aim was to test the group by facial stimuli morph interaction at the N170, VPP, and NSW. We hypothesized that the HPTS group would show a smaller N170 amplitude in comparison to the LPTS and HC groups across all morphs, suggesting poorer facial detection (Chu et al., 2016; Dzhelyova et al., 2012); a smaller amplitude for morphs at the extremes (i.e. untrustworthy and trustworthy morphs) at the VPP, suggesting poorer facial encoding (MacNamara et al., 2013); and show the most negative NSW in comparison to the LPTS and HC groups for the most trustworthy morph, suggesting novelty detection (Jessen & Grossmann, 2019). Through these aims, this study will delineate the differences in social appraisal of trustworthiness among trauma-exposed groups.

1. Method

Full experiment procedures consisted of an eligibility screening (online or by phone), an online survey, and the laboratory behavioural task with EEG recording.

1.1. Participants

Participants were recruited from May 2016 to November 2017 through online and paper flyers posted on The City College of New York's SONA-system and community-based sites (i.e. Craigslist and flyers at Columbia University). Recruitment procedures were tailored for each recruitment site. For the SONA-system, students signed up for a phone screening with a research assistant assessing for general eligibility criteria, which included participants being: (1) between the ages of 18–35; (2) fluent in English; (3) physically healthy to permit sitting at a computer; and (4) self-reporting normal or corrected vision. General exclusion criteria for all participants included: (1) self-reported history or diagnosis of severe and persistent mental illness (e.g. schizophrenia, psychosis); (2) self-reported use of psychotropic medications; (3) past or current self-reported diagnosis of a neurological syndrome (i.e. seizure disorder, brain trauma, and/or tumour disorders); and (4) hairstyles (i.e. braids, weaves, dreadlocks, etc.) that would obstruct EEG recordings of brainwave activity.

Exclusion of participants with hairstyles that interfered with EEG recording is an unfortunate limitation of this study. Our EEG equipment consisted of a BioSemi Active-Two recording system in a 160-electrode montage cap; such a system requires contact between electrodes and a participant's head scalp in order to have the conduction to obtain an EEG recording. There are some hairstyles that obstruct the connection between our electrode cap and scalp, thereby preventing electrodes the ability to capture EEG waves. For this reason, if participants reported having hairstyles that would interfere with EEG recordings, they were given the option to complete the study without EEG or reschedule at a later date if they planned on changing their hair. Some participants returned with different hairstyles, and thus we completed EEG, but some, naturally, did not as well. While we were still able to maintain a racially and ethnically diverse sample of participants (see, Results), this remains a significant limitation of our study and these methods that should be considered in the interpretation and generalizability of results.

Two types of flyers were posted for community-based sites, those that recruited individuals with

'trauma exposure' and those that recruited individuals who were 'healthy adults.' Interested participants contacted the study's email account, where they received a survey to assess screening. The survey asked general eligibility questions (identical to those that SONA-students received) in addition to specific eligibility criteria respective to each type of flyer. For both groups, eligible participants were automatically forwarded to the first part of the study, an online survey. Ineligible participants across all recruitment platforms were thanked for their time. All procedures were approved by the Institutional Review Boards at the City College of New York and Adelphi University.

A total of 252 participants met eligibility criteria and completed the online survey. Of those participants, 111 completed the experimental task. Removal of incomplete data and administration of an incorrect behavioural task ($n = 10$), as well as participants who completed the behavioural task without EEG acquisition ($n = 31$), yielded a total of $N = 70$ participants who completed the survey and experimental task with EEG acquisition. Participants from the City College of New York received \$40 and two study credits for completion of the entire study. Participants from the community received \$50 for completion of the entire study.

1.1.1. Study groups

Participants were categorized into three groups based on their answers to two questionnaires, the Life Events Checklist (LEC-5; Weathers, Blake, et al., 2013) and the PTSD Checklist-5 (PCL-5; Weathers, Litz, et al., 2013) (see, Measures). The HPTS group included individuals who endorsed exposure to at least one traumatic event on the LEC-5 and self-reported a sum total of PTSD symptoms greater than 32 on the PCL-5, a value denoted in the psychometric literature as a cut-off for PTSD (Weathers, Litz, et al., 2013). The LPTS group included individuals who endorsed exposure to at least one traumatic event on the LEC-5 and a sum score of PTSD symptoms ranging from 1–32 on the PCL-5. The HC group included individuals who reported either (a) no exposure to a traumatic life event on the LEC-5 or (b) exposure to one traumatic event on the LEC-5, but with an asymptomatic presentation, defined as a sum score of 0 on the PCL-5. Two participants endorsed exposure to a single traumatic event – for one participant a serious accident at work and for the other participant assault with a weapon. Both of these participants endorsed zero PTSD symptoms on the PCL-5. Behavioural analyses were run with and without these participants in the HC group and showed no statistically meaningful differences. For this reason,

both participants were included in the HC group. Groups were also matched on age, race, sex, ethnicity, occupation, and education level.

1.2. Measures

1.2.1. Demographic characteristics

Age, race/ethnicity, education, and years of education were assessed with basic demographic questions. Participants self-identified their race and ethnicity and were categorized in the following way: White, Black or African American, Native American, Asian, or Other. Participants identified their ethnicity as Hispanic or non-Hispanic.

1.2.2. Trauma exposure

The LEC-5 assessed exposure to traumatic events as defined by the DSM-5 (American Psychiatric Association, 2013). The LEC-5 assessed for directly experiencing or witnessing traumatic events, including interpersonal violence, crime-related trauma, natural disaster, and serious illness (Weathers, Blake, et al., 2013). Responses were dichotomized (yes/no) and summed to provide a total number of traumatic exposures for each participant. In addition, three categories of trauma exposure were created based off of the endorsement of LEC items: accidental traumas (e.g. natural disaster, fire/explosion, transportation accident, serious accident at work or home, and exposure to a toxic substance), intentional traumas (e.g. physical assault, sexual assault, other unwanted sexual experience, combat, and captivity), and other traumas (e.g. life threatening illness or injury, severe human suffering, sudden violent death, sudden accidental death, serious injury or harm you caused to someone else, and any other stressful experience). Previous research has found the LEC-5 to have high inter-rater agreement with a Cohen's $\kappa = .61$ among a sample of college students (Gray, Litz, Hsu, & Lombardo, 2004).

1.2.3. PTSD symptoms

Posttraumatic stress symptoms were measured by the PCL-5, a 20-item self-report symptom measure (Weathers, Litz, et al., 2013). Participants were asked to think of their most stressful event and then respond to the frequency with which they experienced a list of PTSD symptoms within the past month. Likert responses ranged from 0 = Not at All to 4 = Extremely. Sum scores (range: 0–80) were used to assess overall degree of PTSD severity. The internal consistency of the PCL-5 was high, Cronbach's $\alpha = 0.96$.

1.2.4. Internalizing symptoms

Given the overlap in depression and anxiety symptoms with PTSD symptoms, internalizing symptoms were included as covariates in order to account for any variance in models that were not attributed to solely

PTSD symptoms. Two measures were used to assess psychological distress and anxiety symptoms. Psychological distress was assessed by the 53-item Brief Symptom Inventory (BSI; Derogatis, 2001). Participants were required to endorse how frequently they were distressed by items, such as 'feeling blue'. Responses ranged from 0–4. Sum scores created the Global Severity Index (range 0–212) where higher scores indicated greater psychological distress. The internal consistency for the BSI was high, Cronbach's $\alpha = 0.98$. The State-Trait Anxiety Inventory for Adults (STAI; Spielberger, 1983) was used to assess in the moment and trait-level anxiety symptoms. Responses ranged from 1–4; summed scores range from 20–80 for each respective subscale. The internal consistency for the state scale and trait scale were high, Cronbach's $\alpha = 0.97$ and Cronbach's $\alpha = 0.93$, respectively.

1.2.5. Experimental procedure

Facial stimuli were taken from Todorov and colleagues (Oosterhof & Todorov, 2009; Todorov, Baron, & Oosterhof, 2008; <http://tlab.princeton.edu/databases/trustworthinessfaces/>), matching the stimuli used in Fertuck et al. (2016). Trustworthy facial stimuli consisted of four Caucasian, computerized, male, black-and-white facial avatars. Trustworthy faces were parametrically morphed within respective face identities into variations of emotional intensity. For the purposes of this study, three morphs were chosen: 0% (untrustworthy), 50% (neutral), and 100% (trustworthy). All facial stimuli were presented in grey scale and the periphery of faces was blurred in order to occlude non-facial features (Figure 1). It should be noted that the facial stimuli we used are limited, capturing one gender and racial group, thereby excluding the important interaction of different genders, races, and ethnicities in the appraisal of trustworthiness. While we used these stimuli in order to compare our findings with Fertuck et al. (2016), future work, as mentioned in the discussion, should incorporate diverse facial stimuli mirroring participant demographics (cf. Liddell et al., 2019). In this study, there were 15 practice trials followed by 96 actual trials. Facial stimuli were presented in the centre of the screen; participants were tasked with assessing facial trustworthiness via a keyboard response: 1 = very trustworthy, 2 = somewhat trustworthy, 3 = 50/50 (i.e. neutral), 4 = somewhat untrustworthy, and 5 = very untrustworthy. Trials were separated by an inter-trial interval of 800–1200 ms to prevent adaptation or anticipation of stimulus presentation. Participants were verbally instructed to respond as quickly as they could to the face. The task was presented in Presentation (Neurobehavioral Systems, Berkeley, CA, USA) on a Dimension 5150 Dell desktop computer and was presented to participants on a 17-inch Dell Model P1130 RGB computer monitor with



Figure 1. Facial trustworthy morphs. Morphs from the left to right are: trustworthy (100% trustworthy), neutral morph (50% trustworthy, 50% untrustworthy), and the untrustworthy morph (0% trustworthy) morph (Oosterhof & Todorov, 2009; Todorov et al., 2008).

a refresh rate of 75 Hz. Participants were seated at approximately 60 cm from the monitor.

1.2.6. Psychophysiological recordings and processing

EEG was recorded using a BioSemi Active-Two recording system in a high-density, 160-electrode montage to amplify spatial resolution of neural activity. Participants completed all computer tasks while undergoing EEG acquisition in a dimly lit, sound-attenuated chamber. Blinks and other eye movements were monitored by electrooculogram (EOG) from two electrode montages, one on the infra- and supra-orbital ridges of the right eye (VEOG), the other on the outer canthi of each eye (HEOG). Trials containing mastoid activity exceeding $100 \mu\text{V}$ were rejected. Data were preprocessed using the FieldTrip toolbox in MATLAB (Oostenveld, Fries, Maris, & Schoffelen, 2011). Trials contaminated by blinks, eye movements, or other movement artefacts were defined as z-values on the VEOG, HEOG, and lowermost scalp channels exceeding 4.5 in a frequency band between 1 and 140 Hz; artefact trials were automatically removed using FieldTrip. Stimulus-locked waveforms (sweep time = 1000 ms), averaged separately for each morph (trustworthy, neutral, untrustworthy), were referenced to linked mastoids band-pass filtered between .1 and 30 Hz and corrected 200 ms pre-stimulus baseline.

Three ERPs related to facial stimuli processing and/or PTSD were examined: (1) the N170 amplitude, defined as the largest negative amplitude between at 170–212 ms after stimulus onset, and the N170 latency, over nine sites corresponding with the fusiform face area: E30, E31, E32, E17, E18, E19, E14, E15, E16 (Batty & Taylor, 2003; Blau, Maurer, Tottenham, & McCandliss, 2007; Righart & de Gelder, 2008); (2) the Vertex Positive Potential (VPP), defined as the largest positive amplitude between 133–216 ms after stimulus onset over nine centrally located sites: C25,

C24, D2, D3, D4, D15, C1, C27, D14 (MacNamara et al., 2013); (3) and the Negative Slow Wave (NSW), defined as the average amplitude 600–1000 ms after stimulus onset over nine frontal sites: D6, D7, D8, D9, D10, D11, D20, D21, D22 (Jessen & Grossmann, 2019; Tso et al., 2011).

1.2.7. Data analysis

Behavioural analyses tested the equivalence of groups on demographic (e.g. age, race, gender, and years of education) and clinical variables (e.g. psychological distress, trait anxiety, and state anxiety) with chi-square tests for categorical variables and Tukey's Honestly Significant Difference test for continuous variables. Reaction times that were ± 3 SDs above or below the group's specific mean were classified as outliers and removed from analyses; 151 observations were removed. Reaction time was not normally distributed (Shapiro-Wilk test: $p < .05$) and was transformed into a natural logarithmic (\log_e or \ln) scale. Ratings showed a normal distribution.

Two multilevel mixed effect linear regressions were completed on both outcome variables of interest: \log_e and ratings. A multilevel mixed effect linear regression was chosen in order to account for the hierarchical nature of the data where trials were nested into participants who were nested into groups. In addition, due to the removal of reaction times and ratings that were outliers, there were unequal numbers of trial observations across participants. A multilevel mixed effect linear regression was able to account for such variability in the number of trials across participants. Models nested trials (i.e. time) into participants, the random effect. An iteration of models was run from the unconditional model with solely the outcome variable, thereby depicting the overall variance of the model, to final models that included fixed effect predictors of group (HPTS, LPTS, HC), morph (0%, 50%, 100%), the interaction of group by morph, in addition to trial as a random

slope to account for within-task performance variability. Analyses were completed with and without covariates (e.g. gender, race, ethnicity, distress, state anxiety, and trait anxiety). Final reported models excluded covariates due to an absence of change in findings when covariates were included in the model. Moreover, to maximize statistical parsimony, we reported models with less predictors.

A series of model diagnostics were performed, including deviance, log likelihood, Akaike information criteria (AIC), Bayesian information criteria (BIC), and the likelihood ratio test. Models were fit with maximum likelihood estimation and an independent covariance structure. The reference group for both models was HCs and the neutral level of the morph. Recoded models had the trustworthy morph and LPTS group as the referent group. Analyses were completed in Stata IC 15.0 (StataCorp LLC, 2017), where fixed effects are presented as standardized regression coefficients (β) for each factor of groups, morphs, and group by morph interactions. Additional analyses were completed to assess joint effects for the main effects and the interaction, pairwise comparisons at each level of morph and group, and partial interactions. These analyses are reported as χ^2 .

Time-locked ERP data were assessed by mixed model analyses of variance (ANOVAs) on respective amplitudes and latencies using Statistica® software.

Group was included as a between-subjects factor and morph as within-subject factors. To guard against violations of the sphericity assumption with repeated-measures data, all main effects and interactions reported as significant were reliable after Greenhouse-Geisser correction (Greenhouse & Geisser, 1959).

2. Results

2.1. Demographic and clinical characteristics

Table 1 presents the demographic and clinical variables of interest for study participants by group. There were no significant group differences in demographic variables, although the difference between the three groups in racial composition was approaching significance (Monte Carlo two-sided significance test, $p = .061$, 99% CI [.055, .067]). There were expected differences in clinical variables. The mean number of traumatic events differed significantly among all three groups (HC: $M = .11$, $SD = .32$, LPTS: $M = 3$; $SD = 2.12$; HPTS: $M = 4.89$, $SD = 2.42$). In regards to trauma type, HC showed significantly lower accidental traumas ($M = .06$, $SD = .24$) than the LPTS ($M = 1.52$, $SD = .92$) and HPTS ($M = 1.63$, $SD = 1.04$). Moreover, the HPTS showed a greater number of intentional traumas ($M = 1.93$, $SD = 1.36$) than the HC ($M = .56$, $SD = .92$) and LPTS ($M = .06$, $SD = .24$). For other

Table 1. Demographic and clinical characteristics of study population (N = 70).

| Variables | HC | LPTS | HPTS |
|---------------------------|----------------------------|----------------------------|----------------------------|
| | (n = 18) | (n = 25) | (n = 27) |
| | M (SD) or n (%) | | |
| Demographics | | | |
| Age | 21.72 (4.98) | 23.24 (6.11) | 24.52 (5.21) |
| Gender (% female) | 8 (44.4) | 16 (64) | 17 (63) |
| Annual Household Income | | | |
| Less than \$25,000 | 8 (44.4) | 15 (60) | 16 (59.3) |
| \$25,000-\$49,999 | 5 (27.8) | 5 (20) | 6 (22.2) |
| \$50,000 and above | 5 (27.8) | 5 (20) | 5 (18.5) |
| Education (years) | 13.78 (2.07) | 13.48 (1.50) | 13.44 (1.70) |
| Race ¹ | | | |
| White | 6 (33.3) | 9 (36) | 7 (25.9) |
| Black/African American | 3 (16.7) | 4 (16) | 14 (51.9) |
| Asian | 7 (38.9) | 9 (36) | 6 (22.2) |
| Other | 2 (11.1) | 3 (12) | 0 (0) |
| Ethnicity | | | |
| Hispanic | 7 (38.9) | 11 (44) | 8 (29.6) |
| Non-Hispanic | 11 (61.1) | 14 (56) | 19 (70.4) |
| Clinical variables | | | |
| Traumatic Events | .11 (.32) _a | 3 (2.12) _b | 4.89 (2.42) _c |
| Accidental Traumas | .06 (.24) _a | 1.52 (.92) _b | 1.63 (1.04) _b |
| Intentional Traumas | .06 (.24) _a | .56 (.92) _a | 1.93 (1.36) _b |
| Other Traumas | 0 (0) _a | .92 (.95) _b | 1.33 (1.04) _b |
| PTSD Symptoms | 2.56 (4.00) _a | 12.56 (10.17) _b | 45.07 (11.01) _c |
| Psychological Distress | 16.67 (18.37) _a | 25.56 (20.60) _a | 81.67 (41.36) _b |
| Trait Anxiety | 38.22 (11.46) _a | 41 (11.32) _a | 54.70 (12.52) _b |
| State Anxiety | 32.39 (11.26) _a | 35.96 (11.96) _a | 50.56 (15.93) _b |

Note: Different subscripts denote significant group differences ($p < .05$) via pairwise comparisons. For example, cells that share a subscript show no statistical difference from one another whereas cells with different subscripts denote significant differences between groups. Pairwise comparisons were completed with either Chi-Square for categorical variables or Tukey's Honestly Significant Difference test for continuous variables. HC = healthy controls, LPTS = low posttraumatic stress symptoms, HPTS = high posttraumatic stress symptoms, and PTSD = posttraumatic stress disorder.

¹A Monte Carlo simulation was completed due to small cell counts with a 10,000 sampled table with a starting seed of 2,000,000. Results showed that a two-sided Monte Carlo significant test, $p = .061$, 99% confidence interval [.055, .067].

traumas, the HC showed no endorsement ($M = 0$, $SD = 0$) in comparison to the LPTS ($M = .92$, $SD = .95$) and HPTS ($M = 1.33$, $SD = 1.04$). PTSD symptoms also significantly differed among all three groups, where HPTS showed the highest ($M = 45.07$, $SD = 11.01$), followed by LPTS ($M = 12.56$, $SD = 10.17$) and HC ($M = 2.56$, $SD = 4.00$). Psychological distress, state anxiety, and trait anxiety significantly differed between the HPTS group (distress: $M = 81.67$, $SD = 41.36$; state anxiety: $M = 50.56$, $SD = 15.93$; trait anxiety: $M = 54.70$, $SD = 12.52$) and both the LPTS (distress: $M = 25.56$, $SD = 20.60$; state anxiety: $M = 35.96$, $SD = 11.96$, trait anxiety: $M = 41$, $SD = 11.32$) and HC groups (distress: $M = 16.67$, $SD = 18.37$; state anxiety: $M = 32.39$, $SD = 11.26$; trait anxiety: $M = 38.22$, $SD = 11.46$). There were no significant differences in psychological distress and state and trait anxiety between the LPTS and HC groups.

2.2. Behavioural results

2.2.1. Reaction time

The first set of multilevel models tested for differences in the outcome of natural logarithmic (ln) reaction time (Table 2). Likelihood tests demonstrated that model 2 fit the data better than model 1, $\chi^2(1) = 201.98$, $p < .0001$. Model 3 did not fit the data better than model 2, $\chi^2(2) = 2.59$, $p = 0.27$, but model 4 fit the data better than model 3, $\chi^2(6) = 34.06$, $p < .0001$. Model 4 was also superior to model 2, $\chi^2(8) = 36.65$, $p < .0001$ and had the lowest AIC.

Model 4 (column 5) showed a significant fixed effect for morph, $\chi^2(2) = 10.49$, $p = 0.005$ and group \times morph interaction, $\chi^2(4) = 27.27$, $p < 0.001$. There was no significant main effect for group, $\chi^2(2) = 2.61$, $p = 0.27$. Reaction times significantly differed for the trustworthy morph ($\beta = 0.06$, $SE = .03$, 95% CI [.01, .11], $z = 2.27$, $p = 0.02$), the neutral morph, ($\beta = -.06$, $SE = .02$, 95% CI [-.11, -.02], $z = -2.84$, $p = .004$), and the untrustworthy morph ($\beta = 0.13$, $SE = .03$, 95% CI [.08, .18], $z = 4.86$, $p < .001$). Interactions showed that the HPTS group responded faster to the untrustworthy morph ($\beta = -0.14$, $SE = .03$, 95% CI [-.21, -.08], $z = -4.15$, $p < .001$) and the trustworthy morph ($\beta = -0.08$, $SE = .03$, 95% CI [-.15, -.01], $z = -2.36$, $p = .01$) in comparison to the HC group. Moreover, pairwise comparisons showed that the HPTS group responded faster to the neutral morph (Contrast = $-.20$, $SE = .10$, 95% CI [-.40, -.001], $z = -1.98$, $p = .05$) and the trustworthy morph (Contrast = $-.20$, $SE = .10$, 95% CI [-.40, -.002], $z = -1.98$, $p = .05$) than the LPTS group. Further, tests of simple effects showed that LPTS ($\chi^2(2) = 9.60$, $p = .008$) and HC ($\chi^2(2) = 23.62$, $p < .001$) significantly differed in their reaction time for each level of morph, whereas HPTS did not ($\chi^2(2) = 0.92$, $p = 0.63$, Figure 2).

2.2.2. Ratings

Table 3 presents the multilevel mixed linear model for participants' ratings of trustworthy facial stimuli.

Table 2. Two-way multilevel mixed effect linear regression on natural logarithmic reaction time ($N = 70$).

| Variables | Model 1 | Model 2 | Model 3 | Model 4 |
|---------------------------|----------------|----------------|----------------|-----------------|
| | β (SE) | | | |
| Intercept (SE) | 7.27 (0.05)*** | 7.35 (0.04)*** | 7.31 (0.09)*** | 7.25 (0.09)*** |
| Fixed Effects | | | | |
| Group | | | | |
| HC | | | (ref.) | (ref.) |
| LPTS | | | 0.14 (0.11) | 0.18 (0.11) |
| HPTS | | | -0.02 (0.11) | 0.06 (0.11) |
| Morph | | | | |
| Neutral | | | | (ref.) |
| Trustworthy | | | | 0.06 (0.03)* |
| Untrustworthy | | | | 0.13 (0.03)*** |
| Group*Morph | | | | |
| LPTS*trust | | | | 0.00 (0.04) |
| LPTS*untrustworthy | | | | -0.12 (0.04)** |
| HPTS*trust | | | | -0.08 (0.03)* |
| HPTS*untrustworthy | | | | -0.14 (0.03)*** |
| Random Effects | | | | |
| Variance (intercept) | 0.15 (0.03) | 0.13 (0.02) | 0.13 (0.02) | 0.13 (0.02) |
| Variance (trial) | | 0.00 | 0.00 | 0.00 |
| Error Variance (residual) | 0.21 (0.00) | 0.20 (0.00) | 0.20 (0.00) | 0.20 (0.00) |
| Fit Statistics | | | | |
| Log Likelihood | -4464.75 | -4363.76 | -4362.47 | -4345.44 |
| Deviance | 8931.50 | 8723.52 | 8724.93 | 8690.88 |
| ICC | 0.41 | 0.39 | 0.38 | 0.38 |
| AIC | 8935.50 | 8735.52 | 8736.94 | 8714.88 |
| BIC | 8955.91 | 8762.73 | 8777.74 | 8796.49 |

Note: Reference group is healthy controls. Reference morph is the neutral morph. Recoded models for reference groups are reported in text. HC = healthy controls, LPTS = low posttraumatic stress symptoms, HPTS = high posttraumatic stress symptoms, ICC = intraclass correlation coefficient, AIC = Akaike information criterion, BIC = Bayesian information criterion. Significance values are: * $p < .05$. ** $p < .01$. *** $p < .001$.

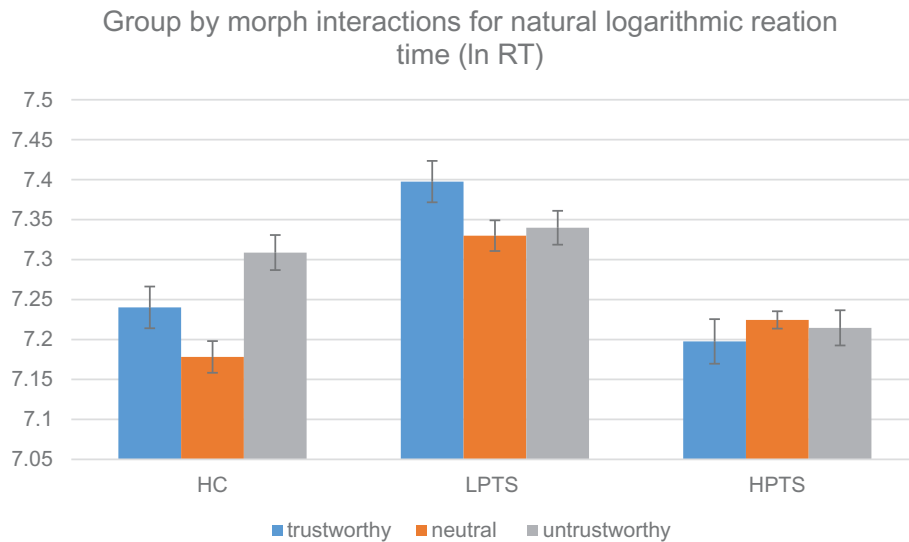


Figure 2. Behavioural differences in the natural logarithmic reaction times by group and morph. HC = healthy controls, LPTS = low posttraumatic stress symptoms, and HPTS = high posttraumatic stress symptoms.

Likelihood ratio tests showed that model 2 fit the data better than model 1, $\chi^2(1) = 85.03, p < .0001$, and again, that model 3 did not fit the data better than model 2, $\chi^2(2) = 0.76, p = 0.68$. However, model 4 remained to fit the model better than model 3, $\chi^2(6) = 973.58, p < .0001$, and model 2, $\chi^2(8) = 974.34, p < .0001$. Model 4 also showed the lowest AIC and BIC values.

There was a significant fixed effect for morph, $\chi^2(2) = 976.47, p < .001$. All three groups rated all three morphs significantly differently (trustworthy morph:

$\beta = 0.30, SE = .05, 95\% CI [0.20, 0.41], z = 13.96, p < .001$; neutral morph: $\beta = -0.29, SE = .04, 95\% CI [-0.38, -0.20], z = -6.48, p < .001$; untrustworthy morph: $\beta = 0.73, SE = .05, 95\% CI [0.63, 0.84], z = 13.96, p < .001$). Neutral morphs were rated as the most trustworthy, followed by trustworthy morphs, and then untrustworthy morphs. We also observed a group by morph interaction ($\chi^2(4) = 9.95, p = .04$). The group x morph interaction (Figure 3) was driven by the HPTS group where HPTS rated the untrustworthy morph as more

Table 3. Two-way multilevel mixed effect linear regression on ratings (N = 70).

| Variables | Model 1 | Model 2 | Model 3 | Model 4 |
|----------------------|----------------|----------------|----------------|----------------|
| | β (SE) | | | |
| Intercept | 3.16 (0.06)*** | 3.05 (0.07)*** | 3.00 (0.14)*** | 2.65 (0.14)*** |
| Fixed Effects | | | | |
| Group | | | | |
| HC | | | (ref.) | (ref.) |
| LPTS | | | 0.01 (0.18) | -0.03 (0.18) |
| HPTS | | | 0.13 (0.18) | 0.03 (0.18) |
| Morph | | | | |
| neutral | | | | (ref.) |
| trustworthy | | | | 0.30 (0.05)*** |
| untrustworthy | | | | 0.73 (0.05)*** |
| Group*Morph | | | | |
| LPTS*trust | | | | -0.02 (0.07) |
| LPTS*untrustworthy | | | | 0.11 (0.07) |
| HPTS*trust | | | | 0.04 (0.07) |
| HPTS*untrustworthy | | | | 0.20 (0.07)** |
| Random Effects | | | | |
| Variance (intercept) | 0.23 (0.04) | 0.31 (0.06) | 0.30 (0.06) | 0.30 (0.06) |
| Variance (trial) | | 0.00 | 0.00 | 0.00 |
| Variance (residual) | 0.93 (0.02) | 0.90 (0.02) | 0.90 (0.02) | 0.78 (0.01) |
| Fit Statistics | | | | |
| Log Likelihood | -9289.60 | -9247.08 | -9246.7 | -8759.91 |
| Deviance | 18579.20 | 18494.16 | 18493.4 | 17519.82 |
| ICC | 0.20 | 0.25 | 0.25 | 0.28 |
| AIC | 18585.2 | 18502.16 | 18505.4 | 17543.82 |
| BIC | 18605.6 | 18529.37 | 18546.2 | 17625.43 |

Note: Reference group is healthy controls. Reference group for morphs is the neutral morph. Recoded models for reference groups are reported in text. HC = healthy controls, LPTS = low posttraumatic stress symptoms, HPTS = high posttraumatic stress symptoms, ICC = intraclass correlation coefficient, AIC = Akaike information criterion, BIC = Bayesian information criterion. Significance values are: * $p < .05$. ** $p < .01$. *** $p < .001$.

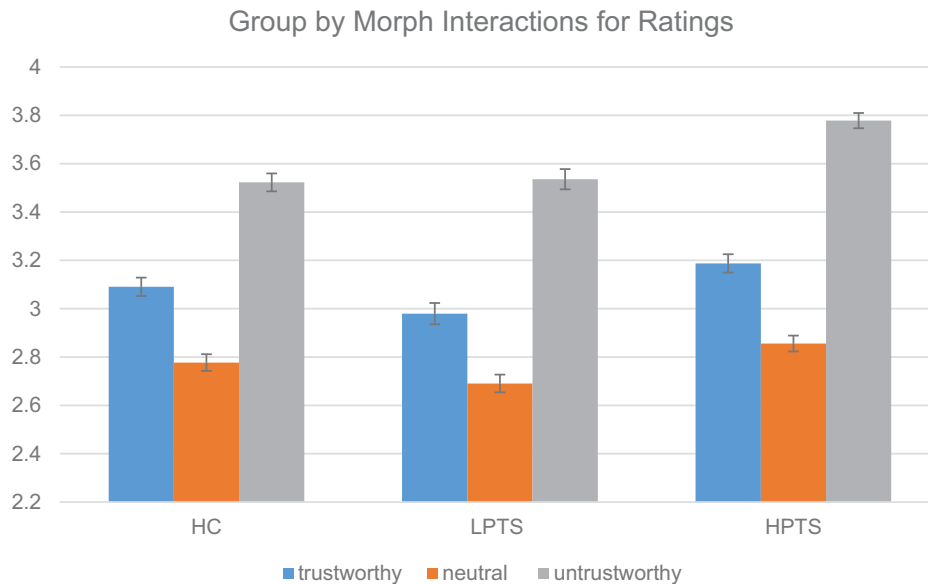


Figure 3. Behavioural differences in trustworthy ratings by group and morph. The rating scale is: 1 = very trustworthy, 2 = somewhat trustworthy, 3 = 50/50, 4 = somewhat untrustworthy, and 5 = very untrustworthy. HC = healthy controls, LPTS = low posttraumatic stress symptoms, and HPTS = high posttraumatic stress symptoms.

untrustworthy than HCs ($\beta = 0.20$, $SE = .07$, 95% CI [0.06, 0.33], $z = 2.88$, $p = .004$) but not the LPTS group. There was no significant main effect for group, $\chi^2(2) = 0.62$, $p = 0.73$.

2.3. Event-related potentials

2.3.1. N170

There was no main effect of group at the N170 amplitude, $F(2,132) = .11$, $p = ns$, $MS_e = 32.55$, Cohen's $f = .008$, but there was a significant group by morph interaction at the N170 latency, $F(4,132) = 3.23$, $p = .01$, $MS_e = .0004$, Cohen's $f = .20$. Post-hoc tests revealed that the HC group showed a significantly later N170 latency (186.66 ± 2.71 ms) than the LPTS (179.78 ± 1.56 ms, $p = .01$) and HPTS (177.89 ± 1.56 ms, $p = .006$) groups in response to the untrustworthy morph.

2.3.2. Vertex positive potential (VPP)

There was a significant group x morph interaction at the VPP, $F(4,132) = 2.41$, $p = .05$, $MS_e = 64.09$, Cohen's $f = .71$ (Figure 4). Post-hoc tests revealed a significant difference for the neutral morph: the LPTS group (1.03 ± 1.24 μV) showed a smaller amplitude than the HPTS group (2.68 ± 0.99 μV , $p = .04$). In addition, to assess brain-behaviour relationships, correlations between task averages were calculated. There was a negative association between behavioural reaction time and the amplitude at the VPP, $r(8) = -0.54$.

2.3.3. Negative slow wave (NSW)

There was a significant main effect of morph for the NSW,

$F(2,132) = 3.29$, $p = .04$, $MS_e = 87.31$, Cohen's $f = .16$, where the untrustworthy morph was associated with a more positive NSW amplitude (1.56 ± 0.53 μV) in comparison to the neutral (0.34 ± 0.67 μV) and trustworthy morphs (0.54 ± 0.61 μV). Further, there was a positive association between the behavioural ratings of facial stimuli and the NSW, $r(8) = 0.34$.

3. Discussion

This study used a SCAN-based translational design in order to understand the behavioural and neural correlates of the social cognitive appraisal of trustworthiness among trauma-exposed individuals with dimensional PTSD symptoms. Contrary to our hypotheses, individuals with high posttraumatic stress symptoms rated untrustworthy faces as *more untrustworthy* than healthy controls. In addition, individuals with higher PTSD symptoms responded to all faces in the same amount of time whereas healthy controls and individuals with lower PTSD symptoms showed variability in their reaction times. Analysis of ERP data suggested that individuals with both high and low posttraumatic stress symptoms recognized the untrustworthy face faster than the healthy control group, but differed in how they encoded different faces. Specifically, individuals with higher PTSD symptoms appeared to encode the neutral face more than individuals with lower levels of traumatic stress symptoms. Together, these results suggest that subjective trust among individuals with higher levels of posttraumatic stress symptoms consists of an early detection of untrustworthy faces, but a subsequent poor encoding of the extreme faces (e.g.

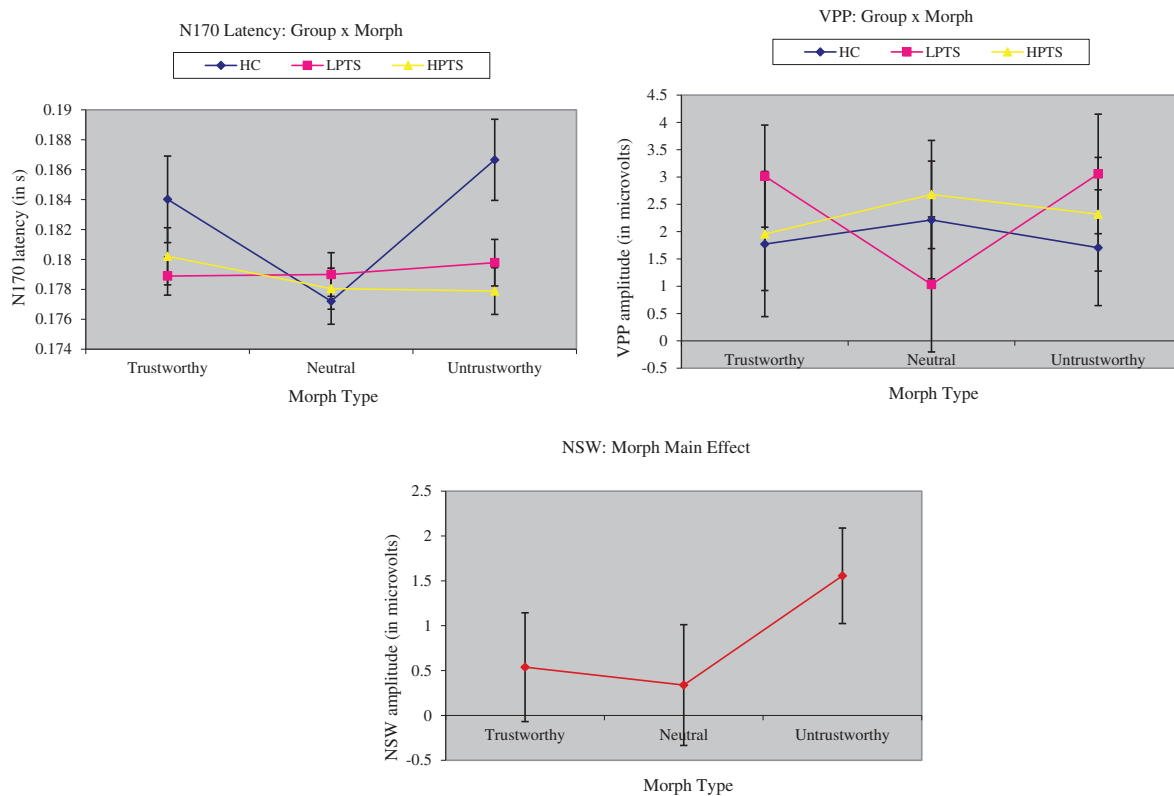


Figure 4. Event-related potentials at the N170 latency, VPP, and NSW. The first graph shows group by morph differences at the N170 latency. The second graph shows group by morph differences at the VPP, and the third graph shows group differences at the NSW. VPP = vertex positive potential, NSW = negative slow wave, HC = healthy controls, LPTS = low posttraumatic stress symptoms, and HPTS = high posttraumatic stress symptoms.

untrustworthy and trustworthy). This could be characterized as an early hypervigilance followed by an attentional avoidance during early preconscious processing of faces for individuals with high levels of posttraumatic stress symptoms.

The first aim of this study was to assess group differences in the appraisal and reaction time to faces on dimensions of trustworthiness. We found differences in the appraisal of untrustworthy faces, where individuals with higher PTSD symptoms perceived untrustworthy faces as *more untrustworthy* than healthy controls. In addition, we found differences in reaction time; individuals with higher PTSD symptoms responded in the same amount of time to all faces whereas healthy controls and individuals with lower PTSD symptoms showed longer response times that varied depending on the face presented. Comparatively, Fertuck et al. (2016) found no group differences in reaction time and found that individuals with PTSD to rate faces as more trustworthy than healthy controls with trauma exposure. The divergence between our study's findings and Fertuck et al.'s (2016) can be interpreted in a few ways. First, Fertuck et al. (2016) used a sample that met full criteria for PTSD (i.e. through a clinician administered assessment) and compared them with trauma-exposed healthy controls. Although individuals in our study with high PTSD symptoms met the cut-off criteria on

the PCL-5 for PTSD (Weathers, Litz, et al., 2013), they were not diagnosed with full or subthreshold PTSD by clinicians, and thus may be characteristic of a less severe PTSD group. If this were true, then our findings could suggest a progression within PTSD where individuals with high levels of posttraumatic stress symptoms, but not necessarily PTSD, show an untrustworthy bias (i.e. perceiving untrustworthy faces as more untrustworthy), followed by individuals with PTSD rating all faces as more trustworthy. Alternatively, both of these studies could highlight how individuals with high posttraumatic stress symptoms or PTSD exhibit an overall poor appraisal of trustworthy faces with the likelihood to exhibit bias in *either* direction. The tendency to show bias in either direction is characteristic of classic approach and avoidance responses widely documented in experimental tasks with threatening facial stimuli (Cisler & Koster, 2010; Clausen et al., 2016; Fani, Bradley-Davino, Ressler, & McClure-Tone, 2011; Hayes, VanElzaker, & Shin, 2012; Pine et al., 2005; Pollak & Kistler, 2002; Pollak & Tolley-Schell, 2003), where deficits in appraising threat yield overly approaching or avoidant behaviours. In our sample, given that individuals with higher PTSD symptoms took the same amount of time to rate all morphs and were *faster* than the other two groups, the higher PTSD group may show an overall avoidance or conversely,

a hypervigilance that leads them to respond quickly to faces on the computer screen. In comparison, Fertuck et al. (2016) may have captured a PTSD sample that exhibited approach behaviour, indicated by higher trustworthiness appraisals.

The second aim of this study was to examine early, middle, and late neural components associated with the appraisal of trustworthy faces in order to characterize subjective trust. The earliest temporal component, the N170, is face-specific ERP associated with early nonconscious recognition of faces in the external environment. Contrary to our hypotheses and literature using happy, angry, trustworthy, and untrustworthy facial stimuli (Chu et al., 2016; Dzhelyova et al., 2012), we found no differences in the N170 amplitude, but we did find differences in the *latency* of the N170, suggesting differences in the moment in time when faces were recognized by participants (D'Hondt et al., 2017; Zhao, Meng, An, & Wang, 2019; Zheng et al., 2016). Both trauma-exposed groups recognized the untrustworthy face faster than healthy controls, plausibly highlighting a hypervigilance for untrustworthy faces. However, both trauma-exposed groups differed in the degree to which they encoded facial morphs at the VPP. Individuals with higher levels of traumatic stress symptoms and healthy controls showed larger VPP amplitudes to the neutral face than individuals with lower levels of traumatic stress symptoms. Larger VPP amplitudes have been associated with superior encoding and attention to structural differences in facial features (Klimova et al., 2013; MacNamara et al., 2013; Wheatley, Weinberg, Looser, Moran, & Hajcak, 2011). Thus, healthy controls and individuals with high PTSD symptoms appear to have encoded the neutral face more than the trustworthy and untrustworthy faces. In contrast, individuals with lower levels of traumatic stress symptoms appear to have encoded the trustworthy and untrustworthy faces more than the neutral face. Taken together with behavioural results, these data may suggest that individuals with lower levels of posttraumatic stress symptoms may engage in compensatory nonconscious strategies when appraising trustworthy faces. They appear to spend more time looking at neutral morphs and then, perhaps through realizing these faces are ambiguous, encode these structural facial features less. In contrast, participants with higher PTSD symptoms mimic healthy controls, encoding the neutral face more and the trustworthy and untrustworthy faces less, a finding corresponding to other studies (Klimova et al., 2013; MacNamara et al., 2013).

Lastly, one unexpected finding in this study that should be mentioned was the rating of neutral faces as more trustworthy than the trustworthy faces across all three groups. Although Todorov et al. (2008) normed the facial stimuli used in this study on a small sample of undergraduates at Princeton University, the differences between our sample and theirs may have led to differences in the appraisal of these faces. Our sample consists of a highly diverse group, both racially (68.6% Non-White) and socio-economically (55.7% earning less than \$25,000/year) residing in an urban area. In comparison, although limited demographic information is provided in Todorov et al.'s (2008) study, we speculate whether these undergraduates were predominantly Caucasian and affluent. Participants in the present study may have perceived normed, neutral male, Caucasian faces as more trustworthy than normed trustworthy faces with more emotional expression. The interaction of perceiver (i.e. participant) and presented facial stimulus is an important limitation to this work, and arguably, many experimental laboratory tasks. Future work should consider diverse facial stimuli, varying on dimensions of age, race, and gender.

4. Limitations

Several limitations should be noted when interpreting the findings of this study. First, our data are cross-sectional, limiting our ability to test for causal mechanisms underlying these processes. Our results may be correlated with the symptom profiles of each group, a preceding risk factor, or a consequence of trauma exposure and traumatic stress symptoms. Moreover, we examined trustworthiness as an early appraisal process, not looking at the downstream effects this appraisal may have on cognition and behaviour. Although this was our aim, the dynamic process of trustworthy appraisals, learning, and their consequence on behaviour is not captured in our design. We also refrained from interpreting the correlations between task averages and ERP amplitudes given the exploratory nature of this study. Additional research could compare their findings with this study's to facilitate in interpretation. Future research would also benefit from including a memory task to corroborate the inferences we have made from ERP results. As aforementioned, the faces in this study were not diverse or representative of the population; a diverse set of facial stimuli could highlight unique interactions with the racial, ethnic, and sex breakdown of our participants. Future studies should utilize other stimuli (cf. Liddell et al., 2019) that may then reveal the interactions between the demographics of the participant and demographics of the facial stimulus. Furthermore, this study did not assess

whether type of trauma exposure may have influenced the results. Future work could distinguish whether certain types of trauma exposure may show differences in trustworthiness appraisals and additionally, whether behavioural and neural results show particular associations with PTSD clusters or symptom severity.

5. Implications & conclusion

This is the first study to our knowledge to capture the neural mechanisms underlying subjective trust among a sample of individuals with trauma and PTSD. Our findings have important clinical implications. First, individuals with high PTSD symptoms seem to have demonstrated an early attentional avoidance when appraising trustworthy faces, which may have led them to rate untrustworthy faces as more untrustworthy in comparison to healthy controls. These miscalculations are most likely exacerbated by the high posttraumatic stress symptom group's poor preconscious encoding of trustworthy and untrustworthy faces, despite their hypervigilant detection of untrustworthy faces when they first appeared on the computer screen. Comparatively, the low posttraumatic stress symptom group showed possible compensatory responses which may be associated with the prevention of future PTSD development. Although they quickly detected untrustworthy faces akin to the high PTSD group, ERP findings suggested continued encoding of these faces. This could suggest that individuals with lower levels of PTSD symptoms appraise trustworthy and untrustworthy faces more accurately by spending more time looking and subsequently encoding more salient social signals (i.e. (un)trustworthy faces) relative to less salient signals (i.e. neutral faces). The healthy controls, in comparison, appear to have taken a moderate amount of time to respond to faces and encoded faces to a lesser degree than individuals with lower PTSD symptoms. It is plausible that healthy controls may not need to attenuate attentional processes due to a lack of trauma exposure. Indeed, the similar neural results among the high PTSD group and the healthy control group may suggest that the high PTSD group has not 'learned' how to modify attentional encoding of faces to compensate for trauma exposure and high traumatic stress symptoms.

Clinically, this study suggests that individuals with PTSD have difficulties in early preconscious appraisals which could contribute to problems in social functioning. This may explain why an individual with PTSD may re-enter abusive relationships, isolate from others, or experience difficulty in social relationship formation. Clinicians may also notice the

employment of attentional and behavioural avoidance strategies to regulate affect when seeing other trustworthy or untrustworthy faces.

In conclusion, PTSD has been traditionally modelled as a disorder stemming from disruption in fear learning and threat reactivity (Liddell et al., 2019) despite considerable debate on the limitation of such models given the heterogeneity in PTSD presentations (Suvak & Barrett, 2011). This study provides preliminary evidence for additional conceptualizations of PTSD as a disorder with disruptions in the social cognition, namely in the appraisal of trustworthiness at the behavioural, attentional, and neural level. Deficits in the appraisal of trustworthy faces could translate to difficulties in assessing the trustworthiness of individuals in real life, potentially thwarting the ability to affiliate with meaningful social bonds facilitating recovery.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Teresa López-Castro  <http://orcid.org/0000-0003-2521-6329>

Eric Fertuck  <http://orcid.org/0000-0003-3785-6630>

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). Washington, DC: American Psychiatric Publishing.
- Batty, M., & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research*, 17(3), 613–620.
- Bellucci, G., Chernyak, S. V., Goodyear, K., Eickhoff, S. B., & Krueger, F. (2017). Neural signatures of trust in reciprocity: A coordinate-based meta-analysis. *Human Brain Mapping*, 38(3), 1233–1248.
- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, 8(6), 551–565.
- Blau, V. C., Maurer, U., Tottenham, N., & McCandliss, B. D. (2007). The face-specific N170 component is modulated by emotional facial expression. *Behavioral and Brain Functions*, 3(7), 1–13.
- Campbell, S. B., & Renshaw, K. D. (2018). Posttraumatic stress disorder and relationship functioning: A comprehensive review and organizational framework. *Clinical Psychology Review*, 65, 152–162.
- Chu, D. A., Bryant, R. A., Gatt, J. M., & Harris, A. W. F. (2016). Failure to differentiate between threat-related and positive emotion cues in healthy adults with childhood interpersonal or adult trauma. *Journal of Psychiatric Research*, 78, 31–41.
- Cisler, J. M., Bush, K., Scott Steele, J., Lenow, J. K., Smitherman, S., & Kilts, C. D. (2015). Brain and behavioral evidence for altered social learning mechanisms

- among women with assault-related posttraumatic stress disorder. *Journal of Psychiatric Research*, 63, 75–83.
- Cisler, J. M., & Koster, E. H. W. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical Psychology Review*, 30(2), 203–216.
- Clausen, A. N., Youngren, W., Sisante, J.-F. V., Billinger, S. A., Taylor, C., & Aupperle, R. L. (2016). Combat PTSD and implicit behavioral tendencies for positive affective stimuli: A brief report. *Frontiers in Psychology*, 7. doi:10.3389/fpsyg.2016.00758
- Cloitre, M., Miranda, R., Stovall-McClough, K. C., & Han, H. (2005). Beyond PTSD: Emotion regulation and interpersonal problems as predictors of functional impairment in survivors of childhood abuse. *Behavior Therapy*, 36(2), 119–124.
- D'Hondt, F., Lassonde, M., Thebault-Dagher, F., Bernier, A., Gravel, J., Vannasing, P., & Beauchamp, M. H. (2017). Electrophysiological correlates of emotional face processing after mild traumatic brain injury in preschool children. *Cognitive, Affective, & Behavioral Neuroscience*, 17(1), 124–142.
- Derogatis, L. R. (2001). *Brief symptom inventory (BSI)-18. Administration, scoring and procedures manual*. Minneapolis: NCS Pearson, Inc.
- Dworkin, E. R., Ullman, S. E., Stappenbeck, C., Brill, C. D., & Kaysen, D. (2018). Proximal relationships between social support and PTSD symptom severity: A daily diary study of sexual assault survivors. *Depression and Anxiety*, 35(1), 43–49.
- Dzhelyova, M., Perrett, D. I., & Jentsch, I. (2012). Temporal dynamics of trustworthiness perception. *Brain Research*, 1435, 81–90.
- Fani, N., Bradley-Davino, B., Ressler, K. J., & McClure-Tone, E. B. (2011). Attention bias in adult survivors of childhood maltreatment with and without posttraumatic stress disorder. *Cognitive Therapy and Research*, 35(1), 57–67.
- Felmingham, K. L., Bryant, R. A., & Gordon, E. (2003). Processing angry and neutral faces in post-traumatic stress disorder: An event-related potential study. *Neuroreport*, 14(5), 777–780.
- Fertuck, E. A., Tsoi, F., Grinband, J., Ruglass, L., Melara, R., & Hien, D. A. (2016). Facial trustworthiness perception bias elevated in individuals with PTSD compared to trauma exposed controls. *Psychiatry Research*, 237, 43–48.
- Frith, C. D. (2008). Social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1499), 2033–2039.
- Gray, M. J., Litz, B. T., Hsu, J. L., & Lombardo, T. W. (2004). Psychometric Properties of the Life Events Checklist. *Assessment*, 11(4), 330–341.
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24(2), 95–112.
- Hayes, J. P., VanElzakker, M. B., & Shin, L. M. (2012). Emotion and cognition interactions in PTSD: A review of neurocognitive and neuroimaging studies. *Frontiers in Integrative Neuroscience*, 6. doi:10.3389/fnint.2012.00089
- Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality: A meta-analytic review. *Perspectives on Psychological Science*, 10, 227–237.
- Jessen, S., & Grossmann, T. (2019). Neural evidence for the subliminal processing of facial trustworthiness in infancy. *Neuropsychologia*, 126, 46–53.
- Joyce, C., & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: The effect of reference electrode site. *Clinical Neurophysiology*, 116(11), 2613–2631.
- Karl, A., Malta, L. S., & Maercker, A. (2006). Meta-analytic review of event-related potential studies in post-traumatic stress disorder. *Biological Psychology*, 71(2), 123–147.
- Keller, S. M., Zoellner, L. A., & Feeny, N. C. (2010). Understanding factors associated with early therapeutic alliance in PTSD treatment: Adherence, childhood sexual abuse history, and social support. *Journal of Consulting and Clinical Psychology*, 78(6), 974–979.
- King, D. W., Taft, C., King, L. A., Hammond, C., & Stone, E. R. (2006). Directionality of the association between social support and posttraumatic stress disorder: A longitudinal investigation. *Journal of Applied Social Psychology*, 36(12), 2980–2992.
- Klimova, A., Bryant, R. A., Williams, L. M., & Louise Felmingham, K. (2013). Dysregulation in cortical reactivity to emotional faces in PTSD patients with high dissociation symptoms. *European Journal of Psychotraumatology*, 4(1), 20430.
- Lanius, R. A., Bluhm, R. L., & Frewen, P. A. (2011). How understanding the neurobiology of complex post-traumatic stress disorder can inform clinical practice: A social cognitive and affective neuroscience approach. *Acta Psychiatrica Scandinavica*, 124(5), 331–348.
- Lenow, J., Cisler, J., & Bush, K. (2018). Altered trust learning mechanisms among female adolescent victims of interpersonal violence. *Journal of Interpersonal Violence*, 33(1), 159–179.
- Lenow, J. K., Scott Steele, J., Smitherman, S., Kilts, C. D., & Cisler, J. M. (2014). Attenuated behavioral and brain responses to trust violations among assaulted adolescent girls. *Psychiatry Research: Neuroimaging*, 223(1), 1–8.
- Liddell, B. J., Cheung, J., Outhred, T., Das, P., Malhi, G. S., Felmingham, K. L., ... Bryant, R. A. (2019). Neural correlates of posttraumatic stress disorder symptoms, trauma exposure, and postmigration stress in response to fear faces in resettled refugees. *Clinical Psychological Science*, 1–15. doi:10.1177/2167702619841047
- Lobo, I., Portugal, L. C., Figueira, I., Volchan, E., David, I., Garcia Pereira, M., & de Oliveira, L. (2015). EEG correlates of the severity of posttraumatic stress symptoms: A systematic review of the dimensional PTSD literature. *Journal of Affective Disorders*, 183, 210–220.
- MacNamara, A., Post, D., Kennedy, A. E., Rabinak, C. A., & Phan, K. L. (2013). Electrocortical processing of social signals of threat in combat-related post-traumatic stress disorder. *Biological Psychology*, 94(2), 441–449.
- Maercker, A., & Horn, A. B. (2013). A socio-interpersonal perspective on PTSD: The case for environments and interpersonal processes. *Clinical Psychology & Psychotherapy*, 20(6), 465–481.
- Meis, L. A., Noorbaloochi, S., Hagel Campbell, E. M., Erbes, C. R., Polusny, M. A., Velasquez, T. L., ... Spont, M. R. (2019). Sticking it out in trauma-focused treatment for PTSD: It takes a village. *Journal of Consulting and Clinical Psychology*, 87(3), 246–256.
- Messman-Moore, T. L., Walsh, K. L., & DiLillo, D. (2010). Emotion dysregulation and risky sexual behavior in revictimization. *Child Abuse & Neglect*, 34(12), 967–976.
- Oostenfeld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011, 156869.

- Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion, 9*(1), 128–133.
- Ozer, E. J., Best, S. R., Lipsey, T. L., & Weiss, D. S. (2003). Predictors of posttraumatic stress disorder and symptoms in adults: A meta-analysis. *Psychological Bulletin, 129*, 52–73.
- Pine, D. S., Mogg, K., Bradley, B. P., Montgomery, L., Monk, C. S., McClure, E., ... Kaufman, J. (2005). Attention bias to threat in maltreated children: Implications for vulnerability to stress-related psychopathology. *The American Journal of Psychiatry, 162*(2), 291–296.
- Pollak, S. D., & Kistler, D. J. (2002). Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proceedings of the National Academy of Sciences, 99*(13), 9072–9076.
- Pollak, S. D., & Tolley-Schell, S. A. (2003). Selective attention to facial emotion in physically abused children. *Journal of Abnormal Psychology, 112*(3), 323–338.
- Read, G. L., & Innis, I. J. (2017). Electroencephalography (Eeg). In Eds. (J. Matthes, C. S. Davis, & R. F. Potter), *The international encyclopedia of communication research methods*. doi:10.1002/9781118901731.iecrm0080
- Righart, R., & de Gelder, B. (2008). Rapid influence of emotional scenes on encoding of facial expressions: An ERP study. *Social Cognitive and Affective Neuroscience, 3*(3), 270–278.
- Spielberger, C. D. (1983). *Manual for the state-trait anxiety inventory STAI (form Y)*. Palo Alto, CA: Consulting Psychologists Press.
- StataCorp. (2017). *Stata statistical software: Release 15*. College Station, TX: StataCorp LLC.
- Stevens, J. S., & Jovanovic, T. (2018). Role of social cognition in post-traumatic stress disorder: A review and meta-analysis. *Genes, Brain and Behavior, e12518*. doi:10.1111/gbb.12518
- Suvak, M. K., & Barrett, L. F. (2011). Considering PTSD from the perspective of brain processes: A psychological construction approach. *Journal of Traumatic Stress, 24*(1), 3–24.
- Suzuki, H., Misaki, M., Krueger, F., & Bodurka, J. (2015). Neural responses to truth telling and risk propensity under asymmetric information. *PloS One, 10*(9), e0137014.
- Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model based approach. *Social Cognitive and Affective Neuroscience, 3*(2), 119–127.
- Tso, I. F., Chiu, P. H., King-Casas, B. R., & Deldin, P. J. (2011). Alterations in affective processing of attack images following September 11, 2001. *Journal of Traumatic Stress, 24*(5), 538–545.
- Weathers, F. W., Blake, D. D., Schnurr, P. P., Kaloupek, D. G., Marx, B. P., & Keane, T. M. (2013). *The life events checklist for DSM-5 (LEC-5)*. Retrieved from Scale available from the National Center for PTSD: www.ptsd.va.gov
- 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)*. Retrieved from Scale available from the National Center for PTSD: www.ptsd.va.gov
- Wheatley, T., Weinberg, A., Looser, C., Moran, T., & Hajcak, G. (2011). Mind perception: Real but not artificial faces sustain neural activity beyond the N170/VPP. *PloS One, 6*(3), e17960.
- Zhao, J., Meng, Q., An, L., & Wang, Y. (2019). An event-related potential comparison of facial expression processing between cartoon and real faces. *PloS One, 14*(1), e0198868.
- Zheng, Y., Li, H., Ning, Y., Ren, J., Wu, Z., Huang, R., ... She, S. (2016). Sluggishness of early-stage face processing (N170) is correlated with negative and general psychiatric symptoms in schizophrenia. *Frontiers in Human Neuroscience, 10*. doi:10.3389/fnhum.2016.00615