



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Post-traumatic stress disorder as a disorder of prediction

**Citation for published version:**

Seriès, P 2019, 'Post-traumatic stress disorder as a disorder of prediction', *Nature Neuroscience*, vol. 22, no. 3, pp. 334-336. <https://doi.org/10.1038/s41593-019-0345-z>

**Digital Object Identifier (DOI):**

[10.1038/s41593-019-0345-z](https://doi.org/10.1038/s41593-019-0345-z)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Nature Neuroscience

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# COMPUTATIONAL PSYCHIATRY

## PTSD as a Disorder of Prediction

**Disproportionate reactions to unexpected stimuli as well as greater attention to perceived threat are cardinal symptoms of PTSD. Computational Psychiatry helps explain how these responses develop and result from abnormalities in learning and prediction during and after traumatic events.**

**Peggy Seriès. University of Edinburgh. pseries@inf.ed.ac.uk**

Following a terrifying event, such as military combat or rape, 5-30% of individuals (1) will develop Post Traumatic Stress Disorder (PTSD). For them, the intense fear they have experienced leaves a debilitating trace that will interfere with their future life. PTSD symptoms include flashbacks, nightmares, hyper-arousal and severe anxiety, as well as uncontrollable thoughts about the event and behavioral strategies to avoid environments that may trigger the symptoms.

Why does PTSD develop for a fraction but not all individuals submitted to similar experiences? Is there a biological vulnerability for the disorder, or a biological signature of its consequences that could be used as a diagnostic marker and guide the development of new therapies? Recent studies by Homan et al (2) and Brown et al (3) in combat-exposed veterans show how computational psychiatry can help answering those questions. As with other mental disorders (4), the key might be to model PTSD as resulting from (subconscious) inferential biases and impaired belief updating.

Common theories propose that PTSD results from abnormalities in learning during and after the traumatic event (5). Fear conditioning could explain why neutral stimuli (people, place, sounds etc.) that have been associated with the traumatic event acquire the capacity to trigger and maintain anxiety long after the trauma itself. Why this association doesn't weaken over time has been attributed to either the fact that it was abnormally strong in the first place, or - more likely - to deficits in extinction processes, i.e. a failure for the association to weaken when the same cues are encountered without leading to the traumatic event. This could be a result of patients' avoidance strategies: individuals with PTSD avoid encountering such cues again and thus may never experience them as being safe. Other theories assume on the contrary that PTSD is related to basic deficits in acquiring associations between specific cues and the traumatic event. This would

result in associating the trauma with the environment as a whole, resulting in heightened contextual anxiety and/or overgeneralization of fear to all cues resembling the initial cues. In environments not related to the traumatic event, PTSD patients have also robustly been found to exhibit reduced habituation of responses to repeatedly presented novel, intense or fear-relevant stimuli, as well as greater sensitization of fear-related autonomic responses. Despite the popularity of those theories, the specific components of anomalous learning in PTSD remain unclear.

Computational modeling is ideally placed to help formalize and quantitatively test hypotheses regarding such potential abnormalities. In the laboratory, we can explore how individuals learn to predict the association between different cues and threats (such as electric shocks) and their flexibility in using, updating or forgetting those predictions. Computational modeling can then reveal inter-individual differences in internal learning and evaluation processes that are otherwise inaccessible to raw data analysis (4).

Homan et al. (2) used a fear-conditioning task with a group of combat-exposed veterans presenting a wide range of PTSD symptoms (Figure 1). Participants had to passively learn the pairing between two face images and mild electric shocks. Face A was paired with an electric shock in one third of the trials, while Face B was never paired with the shock. The acquisition phase was immediately followed by a reversal phase. After reversal, face B is now likely to lead to the shock, while face A is no longer paired with the shock. To assess conditioning, the authors measured skin conductance response (SCR). Interestingly, PTSD severity had no effect on the acquisition of the conditioned response before or after the reversal: all participants seemed to learn equally well. However, a modeling approach uncovered subtle differences.

Homan et al. used a basic reinforcement learning (Rescorla- Wagner) model and a Pearce-Hall hybrid model to fit the SCR data. Both types of reinforcement models compute a “value” for each face cue, iteratively updated at each trial, based on the discrepancy between the expected and obtained outcome, i.e. the prediction error. However, the hybrid model replaces the constant learning rate of the Rescorla-Wagner model by a dynamic “associability” parameter, which reflects attention allocation to cues that has been previously accompanied by surprise. Associability dynamically modulates value learning by accelerating it for cues whose predictions are poor (large prediction errors) and decelerating it when predictions become reliable.

In line with previous studies (3, 6), Homan et al. found that the hybrid model accounted for the SCR data better than the basic model. Moreover, after fitting the model to individual participants’ data, they found that PTSD severity was associated with one particular model quantity: the prediction error weight, which can be seen as a learning rate for associability. In line with Brown et al. (3), they found that highly symptomatic combat veterans were more influenced by prediction errors, weighing them more strongly as they adjusted trial by-trial attention to cues.

Using model-based fMRI, they went one step further and asked about the neural correlates of such differences: where and how strongly is the computation of value, prediction errors and associability reflected in the neural activity. One of the main structures implicated in PTSD is the amygdala, considered as the threat processing center and locus of associative learning (1). The amygdala has been found to be smaller in size and hyperactive in PTSD. Other structures are also involved in providing context and meaning to the traumatic events: in particular the PFC and the hippocampus. PTSD patients typically show reduced activation of the PFC and hippocampus, which is thought to correspond with reduced top-down inhibitory control of the amygdala, possibly explaining the hyper-responsivity of the amygdala to fearful stimuli (1).

Homan et al. found that neural activity in the amygdala was associated with the computation of value for the face images. PTSD was associated with lower neural tracking of value in the amygdala and the striatum, in addition to smaller amygdala volumes. Moreover, and departing from (3), the authors found lower tracking of associability (and less so of prediction error) in the striatum, hippocampus and dACC in individuals with higher PTSD severity. They suggest that the higher weight assigned to prediction errors might be a compensatory adjustment for the decreased neural tracking of associability.

Computational psychiatry of PTSD is in its infancy, and quantifying individual differences in internal learning and evaluation processes is an important first step. By framing PTSD in a predictive coding framework, these recent findings may provide new keys to understanding the disorder: the increased weight given to surprising outcomes might explain disproportionate reactions to unexpected stimuli or events, as well as heightened orienting and attentional biases toward negative information (3). It could also explain the aberrant learning and synaptic plasticity long postulated to be at the core of PTSD (1), that aversive outcomes could be experienced as less predictable and less avoidable, and the documented aversion to ambiguity in aversive environments in PTSD (7).

The next steps will be to clarify how those results compare with previous findings (5) and whether they extend (or not) to other paradigms such as instrumental (3) and reward (8) learning. It will be crucial to verify that these individual differences correspond to vulnerabilities for the disorder (as opposed to its consequences) and how they relate to the different dimensions of PTSD symptoms (re-experiencing, avoidance, hyper-arousal). It will be also important to show that they are specific to PTSD, as opposed to depression (3) or other anxiety disorders (9), which have also been found to relate to learning rates of associative learning in dynamic aversive environments (10). Ultimately, computational studies will need to focus on developing models that can integrate

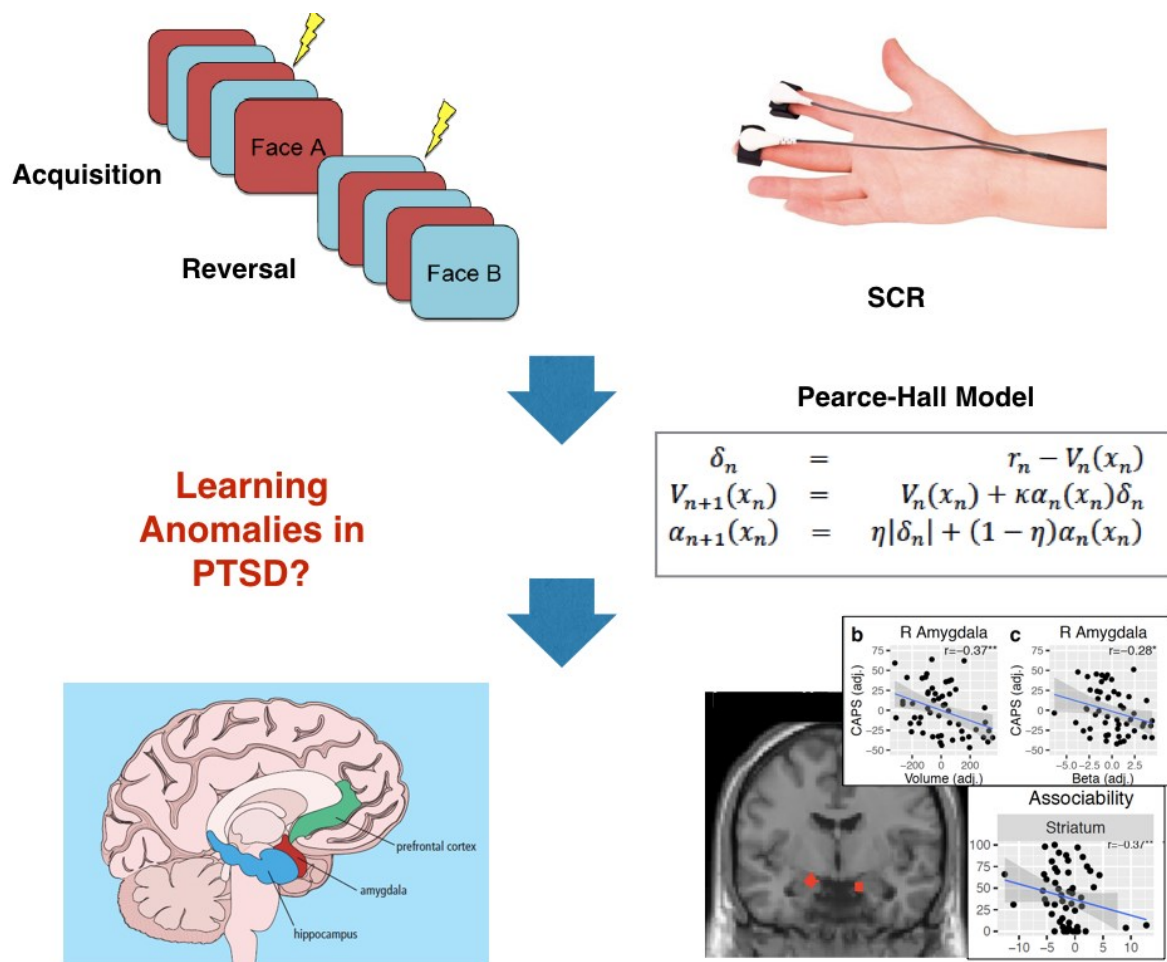
theories of abnormal learning, during and after traumatic events, with the explanation of some or all symptom clusters into a single framework.

Importantly, such learning mechanisms are also at the core of the therapies that have shown to be effective in PTSD (11): prolonged exposure, cognitive processing therapy and trauma-focused CBT. These treatments all try to counteract avoidance strategies and to directly address - and update - the associations (memories, feelings, thoughts) made during the traumatic events. Despite the relative success of these techniques, the mechanisms behind both their strengths and their weaknesses are inadequately understood and it has been suggested that up to 33% of people with PTSD are resistant to treatment (12). We need to understand how those therapies work when they do, possibly by identifying the relationship between individual learning differences (such as increased attention to surprising outcomes) and treatment success. Ultimately we will need to design new therapies informed by a better understanding of the role of inference and learning in the genesis and maintenance of psychological distress (13).

## References

1. Mahan, A. L. & Ressler, K. J. Fear conditioning, synaptic plasticity and the amygdala: implications for posttraumatic stress disorder. *Trends Neurosci.* 35(1): 24-35 (2012).
2. Homan, P., Levy, I., Feltham, E., Gordon, C., Hu, J., et al. Neural computation of threat in the aftermath of combat trauma. *Nat. Neurosci.* <https://dx.doi.org/xxx> (2019).
3. Brown, V. M., Zhu, L., Wang, J. & Chiu, P. H. Associability-modulated loss learning is increased in post-traumatic stress disorder. *eLife*; 7:e30150 (2018).
4. Huys, Q. J., Maia, T. V. & Frank, M. J. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neurosci.* 19: 404–413 (2016).
5. Lissek, S. & van Meurs, B. Learning models of PTSD: Theoretical accounts and psychobiological evidence. *International Journal of Psychophysiology* 98(3), 594–605 (2015).
6. Li, J, Schiller, D., Schoenbaum, G., Phelps, E. A. & Daw, N. D. Differential roles of human striatum and amygdala in associative learning. *Nature Neurosci.* 14:1250–1252. (2011).
7. Ruderman, L., Ehrlich, D., Roy, A., Pietrzak, R. H., Harpaz-Rotem, I. & Levy, I. Posttraumatic symptoms and aversion to ambiguous losses in combat veterans. *Depression and Anxiety* 33(7), 606–613. (2016)
8. Myers, C. E. Learning to Obtain Reward, but Not Avoid Punishment, Is Affected by Presence of PTSD Symptoms in Male Veterans: Empirical Data and Computational Model. *PloS One.* (2013).

9. Raymond, J. G., Steele, J. D. & Seriès P., Modeling Trait Anxiety: From Computational Processes to Personality, *Front Psychiatry*. 8: 1. (2017).
10. Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X. & Bishop, S. J. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat Neurosci*. 18(4): 590-6. (2015).
11. Watkins, L. E., Sprang, K. R. & Rothbaum, B. O. Treating PTSD: A Review of Evidence-Based Psychotherapy Interventions. *Frontiers in Behavioral Neuroscience*. 12: 258 (2018).
12. Green, B. Post-traumatic stress disorder: new directions in pharmacotherapy. *Advances in Psychiatric Treatment*. 19:3, pp. 181-19 (2013).
13. Moutoussis, M., Shahar, N., Hauser T. U. & Dolan, R. J. Computation in Psychotherapy, or How Computational Psychiatry Can Aid Learning-Based Psychological Therapies *Computational Psychiatry*. 2:50-73. (2018).



**Figure 1. A Computational Psychiatry approach to investigating possible learning anomalies in PTSD.** (A) Homan et al. (2) recorded skin conductance responses (SCR) during a fear-conditioning paradigm in combat-exposed veterans. Face A was first paired with a mild electric shock. After reversal, Face B was paired with the shock while Face A was no longer associated with it. (B) They then modelled the SCR data using various reinforcement models which compute a “value”  $V$  for each face cue  $x$ , iteratively updated at each trial  $n$ , based on the discrepancy between the expected ( $V_n(x_n)$ ) and obtained outcome ( $r_n$ ), i.e. the prediction error  $\delta_n$ . The best fitting model was found to be a hybrid Pearce-Hall model, which includes an associability variable ( $\alpha_n$ ), which reflects attention allocation to cues that has been previously accompanied by surprise. They determined the best-fitting parameters of the model for each individual and found that PTSD severity was associated with increased prediction error weight ( $\eta$ ). (C) The model-based time-series was then convolved with the hemodynamic response function and then regressed against fMRI data with a focus on regions known to be involved in PTSD (the amygdala, the striatum, the hippocampus and the dACC). They found that the neural computations that were shaped by these altered prediction error weights contributed to the symptoms of PTSD: aversive value encoding in the amygdala and striatum, and associability computations in the striatum, dACC, and hippocampus. They also found that the right amygdala computations contributed to the

symptomatology above and beyond the effects of smaller amygdala volumes, suggesting additive effects of right amygdala volume and function.