cambridge.org/neu

Original Article

Cite this article: Migó M, Chou T, Widge AS, Peters AT, Ellard K, Dougherty DD, and Deckersbach T. (2023) Neural correlates of learning accommodation and consolidation in generalised anxiety disorder. *Acta Neuropsychiatrica* **35**:218–225. doi: 10.1017/neu.2022.16

Accepted: 5 December 2021 First published online: 27 May 2022

Keywords:

Anxiety; Generalized anxiety disorder; Learning; Hippocampus; Basal ganglia

Author for correspondence: Darin D. Dougherty, Email: ddougherty@partners.org

Neural correlates of learning accommodation and consolidation in generalised anxiety disorder

Marta Migó¹, Tina Chou¹, Alik S. Widge^{1,3}, Amy T. Peters¹, Kristen Ellard¹, Darin D. Dougherty¹ and Thilo Deckersbach^{1,2}

¹Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; ²University of Applied Sciences, Diploma Hochschule, Germany and ³Department of Psychiatry, University of Minnesota, Minneapolis, MN, USA

Abstract

Objective. Anxiety can interfere with attention and working memory, which are components that affect learning. Statistical models have been designed to study learning, such as the Bayesian Learning Model, which takes into account prior possibilities and behaviours to determine how much of a new behaviour is determined by learning instead of chance. However, the neurobiological basis underlying how anxiety interferes with learning is not yet known. Accordingly, we aimed to use neuroimaging techniques and apply a Bayesian Learning Model to study learning in individuals with generalised anxiety disorder (GAD). Methods. Participants were 25 controls and 14 individuals with GAD and comorbid disorders. During fMRI, participants completed a shape-button association learning and reversal task. Using a flexible factorial analysis in SPM, activation in the dorsolateral prefrontal cortex, basal ganglia, and hippocampus was compared between groups during first reversal. Beta values from the peak of these regions were extracted for all learning conditions and submitted to repeated measures analyses in SPSS. Results. Individuals with GAD showed less activation in the basal ganglia and the hippocampus only in the first reversal compared with controls. This difference was not present in the initial learning and second reversal. Conclusion. Given that the basal ganglia is associated with initial learning, and the hippocampus with transfer of knowledge from short- to long-term memory, our results suggest that GAD may engage these regions to a lesser extent during early accommodation or consolidation of learning, but have no longer term effects in brain activation patterns during subsequent learning.

Significant outcomes

- Most importantly, our paper uses Bayesian modelling approaches to track learning in a clinical population.
- We provide evidence that GAD patients may engage the basal ganglia and the hippocampus to a lesser extent during early learning accommodation or consolidation, but show no difference in neural activation in subsequent learning.

Limitations

- The most important limitations of our study were our small sample size, and our clinically heterogeneous patient group, who had a number of comorbidities.
- The use of this heterogeneous group strengthens the generalizability of our findings, since the majority of GAD patients have comorbid disorder, but it also means that our findings are not necessarily specific to only having GAD.

© The Author(s), 2022. Published by Cambridge University Press on behalf of Scandinavian College of Neuropsychopharmacology.



Introduction

Generalized anxiety disorder (GAD) is an anxiety disorder characterised by excessive, uncontrollable, and often irrational worry about events or activities (American Psychiatric Association, 2013). Trait anxiety, the primary component of GAD, has been theorised, in the Attentional Control Theory (Eysenck *et al.*, 2007; Derakshan and Eysenck, 2009), to interfere with inhibition, shifting, and updating processes of working memory, which predicts that individuals with anxiety will perform worse on cognitively demanding tasks. For instance, individuals with high math anxiety demonstrate reduced working memory span, and in turn, longer reaction times and more performance errors when completing mental additions



concurrently with a memory load task (Ashcraft and Kirk, 2001). Additionally, high-trait anxious individuals show slower target identification when completing a response–conflict task under low attentional demands and when threat-related stimuli are absent (Bishop, 2009). Together, anxiety is linked to impairments in attention and memory (Ashcraft and Kirk, 2001; Eysenck *et al.*, 2007; Bishop, 2009), which are components that can affect learning.

Notably, a study exploring probabilistic learning in individuals with GAD who were given either positive or negative affective reinforcements showed that those with GAD learn the correct probabilistic choices at a slower rate over time and to a lesser degree than control participants regardless of reinforcement type (LaFreniere and Newman, 2019). GAD is also associated with underperformance during a n-back learning task, regardless of working memory load, under threat (Vytal *et al.*, 2016). Moreover, decreased reward learning under stress is associated with depression severity in GAD (Morris and Rottenberg, 2015). However, all prior studies included a stressor or affective component. Hence, there is a need to evaluate the purely cognitive learning component and to investigate whether non-affective learning is impaired in GAD in no-threat environments.

Real-time learning has rarely been explored in GAD. However, exploring real-time, trial-by-trial learning could reveal more nuanced impairments in GAD and account for time-sensitive abnormalities associated with GAD, such as deficits in learning or cognitive flexibility during specific trials or time periods. Bayes Theorem is a commonly used statistical approach, which aims to take into account the probability of some event occurring, given some a priori feature, and can then be used to model the chances of an event happening as a result of learning, given its odds of happening by chance (Ahn, et al., 2017). This approach facilitates a more precise understanding of the rate of inference and accuracy when learning characteristics of new environments (Wilson et al., 2007). It has proven to be a useful tool to quantify learning, allowing for the trial-by-trial study of learning in psychiatric patients (Ahn et al., 2017; Widge et al., 2017). However, this statistical model has only been previously applied to studying the degree of learning in psychiatric populations with anxiety-related and mood disorders under conditions of threat or monetary reward/punishment (Harlé et al., 2017; Aylward et al., 2019; Piray et al., 2019). While prior studies utilised Bayesian Learning Theorem, results appeared inconclusive, as Aylward et al. (2019) and Harlé et al. (2017) concluded that patients with mood and anxiety disorders and participants with higher trait-anxiety were quicker to update their behaviour in response to negative or threatening outcomes, while Piray et al., concluded that trait anxiety was accompanied by disruption of optimal learning in threatening situations. This enhanced or hindered learning process can be further broken down into GAD patients' ability to flexibly update their learned knowledge or associations (learning accommodation) and their ability to transfer short-term learning to longterm learning (learning consolidation). Learning accommodation takes place when a new piece of knowledge substitutes an old, previously learned fact. Learning consolidation takes place when a piece of knowledge becomes more quickly accessible in one's mind, and more automatic. Together, it is not yet known how GAD and its neural underpinnings interfere with non-affective learning, learning flexibility, and knowledge consolidation in a nonthreatening environment, in real time.

The purpose of this study is to investigate neural correlates of real-time non-affective learning accommodation and

consolidation in individuals with GAD and healthy controls (HC) in a stress-free environment. Consequently, we examine Bayesian learning coefficients as measures of initial learning, learning consolidation, and learning accommodation and use them as regressors to average brain activation time course in a functional magnetic resonance imaging (fMRI) associative learning paradigm in a population with GAD and comorbid mood, anxiety, and impulse-control disorders.

In previous research, the basal ganglia has been identified as an area related to initial rapid learning and learning accommodation (Graybiel, 1995; Packard and Knowlton, 2002). Additionally, the hippocampus has been identified as an area related to the transfer of knowledge (Jarrard, 1993; Myers et al., 2003). Furthermore, the dorsolateral prefrontal cortex (dlPFC) has been associated with working memory (Barbey et al., 2013), and learning accommodation (Xue et al., 2013; Bartolo and Averbeck, 2020). Reduced activation of the hippocampus (Bannerman et al., 2004; Engin and Treit, 2007), the basal ganglia (Wu et al., 1991; Marchand, 2010), and the dlPFC has also been associated with anxiety (Bishop, 2009; Balderston et al., 2017). Consequently, we hypothesise that the basal ganglia will show decreased activation in the GAD group in those blocks requiring initial learning and memory accommodation, rather than those requiring memory consolidation. Conversely, we hypothesise that individuals with GAD will display decreased activation of the hippocampus in those blocks requiring memory consolidation, rather than those blocks requiring memory accommodation or initial learning. Finally, we hypothesise that individuals with GAD will display decreased activation of the dlPFC in all blocks requiring working memory and learning accommodation.

Material and methods

Participants

Study participants (ages 19-70) were 25 healthy participants and 14 individuals with GAD recruited at Massachusetts General Hospital. See Table 1 for participant demographics. All participants were right handed, had corrected to normal vision, and provided written informed consent prior to participation in accordance with the guidelines of the Mass General Brigham Human Research Office/Institutional Review Board. All healthy controls reported no history of significant medical, psychiatric, or neurological illness, no history of substance abuse within the past 3 months, were not currently taking antidepressants, mood stabilisers or benzodiazepines, all had normal or corrected-tonormal vision and were all right-handed. All psychiatric patients had been diagnosed with GAD with the Mini International Neuropsychiatric Interview (Sheehan et al., 1998). Individuals also had comorbid mood (bipolar disorder [n=4], major depressive disorder [n=7]), impulse control (obsessive-compulsive disorder [n=2]), and other anxiety disorders (social anxiety disorder [n=4], panic disorder [n=4], agoraphobia [n = 2]).

Associative learning fMRI task paradigm

During fMRI, participants completed a shape-button association learning and reversal task, modelled after (Cools *et al.*, 2002). In this task, participants were shown shapes in the centre of a computer screen and were asked to choose one of three keys during each trial to learn which key each of the six shapes was associated with (Fig. 1). When the button pressed was correct, the

Table 1. Participants' demographics

Measure		HC (N)	HC (%)	GAD (N)	GAD (%)	<i>p</i> -value
Sex	Female	13	52	7	50	0.1
	Male	12	48	7	50	
Race	White	21	84	14	100	0.29
	Black	2	8	0	0	
	Asian	2	8	0	0	
Ethnicity	Latino/Hispanic	2	8	0	0	0.16
	Non-Latino/Hispanic	23	92	14	100	
		НС (М)	HC (SD)	GAD (M)	GAD (SD)	<i>p</i> -value
Age		32.8	8.4	30.5	12.2	0.5

Note. Non-parametric measures tested for significance using Chi-squared tests. Parametric measures tested for significance using t-tests.



Figure 1. Stimuli: Above are the six stimuli that were presented to participants, which had to be associated with three different buttons during the task.

shape turned green, and when it was incorrect, the shape turned red (Fig. 2). A blue circle was presented interspersed with the shapes, which turned green upon the selection of any button. Instructions were written on the screen during the training phase, and participants were given the opportunity to ask clarifying questions. The experiment had three phases; an initial learning block lasting 82 trials and two reversal blocks (first reversal and second reversal) lasting 100 trials. In the learning stage, three of the six shapes were shown at greater, pseudo-randomized frequency initially so that the participant could learn more effectively, followed by a decrease in the frequency of presentation of those shapes and an increase in the frequency of the other three shapes for the second half of the learning phase. In the first reversal block, the first half of shape-button associations were pseudo-randomized, leaving the second half of the shape-button associations unaltered. In the second reversal block, the second half of the shapes were pseudo-randomized, while the first half remained associated to the same button as the previous phase. These blocks were designed to trigger two specific learning processes: accommodation and consolidation. To correctly encode the changes in button association, participants needed to update (or accommodate) their previous learning. However, to correctly recall the previous associations that had stayed constant, participants needed to consolidate their previous learning. Participants had 1.4 seconds to respond to each shape.

MRI scanning

All MRI scans were completed at the Athinoula A. Martinos Center for Biomedical Imaging. MRI data were collected using a 3.0-T whole-body scanner (Trio-System), equipped for echo planar imaging (Siemens Medical Systems, Iselin NJ) with a 32-channel head coil. Foam cushions were used to restrict head movements. Task images were displayed using a rear projection system and MATLAB stimulus presentation software (Psychtoolbox, MATLAB. (2010).version 7.10.0 (R2010a). Natick, Massachusetts: The MathWorks Inc.). The structural sequences involved a high-resolution, four-multiecho, T1-weighted, magnetisation-prepared, gradient-echo image (TR = 2510 ms, TE = 1.64 ms, flip angle = 7° , voxel size = $1.0 \times 1.0 \times 1.0$ mm; Van Der Kouwe et al. (2008)). Functional images were acquired using a multiband SMS-3 T2*-weighted echo-planar-imaging (EPI) sequence sensitive to blood-oxygen-level dependent (BOLD) contrast (TR = 2200 ms, TE = 30 ms, flip angle = 75° , voxel size = $2.0 \times 2.0 \times 2.0$ mm).

Analysis

Learning model

The latent learning state of each participant during each trial was estimated using Bayesian inference applied independently for each shape. For this behavioural analysis, the prior likelihood was **Figure 2.** Task Paradigm: Once the stimulus appeared on the screen, participants had to press a button (out of three possible ones). If the figure turned green, the correct button had been pressed, meaning that the association was correct. If the figure turned red, the participant had to try a new button next time the same stimulus was presented again. Here, in the Initial Learning block (top), the presented stimulus was associated with button number 3. However, in one of the reversal blocks (bottom), the presented stimulus changed its association to button number 2. Participants were then expected to press button number 2 when the stimulus above was shown, instead of button number 3.

Figure 3. The prior distribution is shown on the left and the resulting posterior distributions are shown for either a correct or an incorrect response are shown on the right. In this example, the participant correctly recalled the shape-button association for 75% of associations after the first correct response for that association. So, the prior distribution has a mean of 1/n where n is the number of key choices available to the participant and a standard deviation such that if the participant chooses correctly, it looks like the distribution in the top right. which has a maximum at 0.75. If the participant chose wrong, their latent learning state distribution would have, in this example and in most cases, a maximum at 0.



Learning

1 2 3

1 2 3

assumed to 1/n where n is the number of key choices available which reflects the assumption that the participant is naive to the shape-button associations at the start of every block had an equal likelihood of choosing each button, and therefore had an equal likelihood of selecting the correct association on the first trial. Starting with the prior distribution on the first trial, this likelihood distribution was then updated, independently for each shape using Bayes rule, and the probability with the maximum value was taken as an estimate of the latent learning state x_i during trial *i*. The standard deviation of the prior distribution was chosen based on the likelihood that, after choosing the correct association, participant correctly remembered the correct association on the next trial where that shape was presented. The value of the standard deviation was chosen such that after an update was applied using Bayes rule, the maximum likelihood probability was equal to the probability that the participant correctly recalled the shape-button association after the first correct response. After reversals, the mean of the prior distribution was assumed to be $1/x_j$

where xj is the latent learning state of the trial immediately before the reversal. This reflects that participants are unlikely to respond correctly the first trial after the reversal if they have learned the shape-button association well. The model was chosen this way because we had strong prior knowledge about the likelihood of answering an unknown association correctly and the likelihood of recalling an association. This analysis yielded an estimate of the latent learning state x_i for each trial *i* (Yousefi *et al.*, 2019). Trials with no response were excluded, and participants who had not responded to more than 20% (seven in total: five controls, two patients) of the stimuli were also excluded. See Fig. 3 for reference.

fMRI

Analyses were conducted using SPM8 (Wellcome Department of Cognitive Neurology, London, UK). Raw data were presubtracted using phase and magnitude fieldmap images, slice time corrected, realigned and unwarped, co-registered, convolved into threedimensional space established by the Montreal Neurological



Figure 4. Individual data points and group differences in the basal ganglia during the first reversal block (right BA 48; x = 18, y = -14, z = 24). p = 0.03.



Figure 5. Individual data points and group differences in the hippocampus during the first reversal block (right BA 54; x = -36, y = -32, z = -10). p = 0.03.

Institute, segmented, normalised, and smoothed. For motion correction, during preprocessing, the realignment step generated a set of motion-related parameters and tracked participants' motion in the x, y, and z directions, as well as in pitch, roll, and yaw. This set of six parameters was included as covariates in the first-level individual-subject analysis to correct for individual variation in head motion. In first levels, Bayesian learning coefficients were input as parametric modulators to individual functional data (Zorowitz *et al.*, 2019). Trials with no response were excluded.

Statistical analysis

In SPM, we conducted a GLM random effects flexible factorial analysis to identify group differences in brain activation during the initial learning, first reversal, and second reversal blocks of the task. More specifically, we created contrasts of interest: GAD > HC and

HC > GAD initial learning block, GAD > HC and HC > GAD first reversal block, GAD > HC and HC > GAD second reversal block. We then created anatomical masks of our *a priori* regions using the Wake Forest University Pick Atlas (Maldjian *et al.*, 2003), the bilateral dlPFC, hippocampus, and basal ganglia. We next identified significant clusters in these *a priori* regions for our contrasts of interest. Finally, we applied AFNI's 3DClustSim's (Forman *et al.*, 1995; Cox, 1996) correction to identify the minimum cluster sizes in our *a priori* regions which would survive multiple comparisons at $\alpha < 0.05$ (voxel correction size of 134 voxels in the basal ganglia, 123 voxels in the hippocampus, and 189 voxels in the dlPFC). In order to better visualise the significant group differences, we extracted beta weight values from spherical ROIs (radius = 2 mm) (http://marsbar.sourceforge.net) centred on the coordinates of the peak voxel of each significant cluster.

Results

Behavioural data

The average response accuracy, whether a subject correctly picked the shape-button association, was 79.7% (SD = 13.2%) in the HC group and 83.8% (SD = 8.2%) in the GAD group. However, subjects learned shapes heterogeneously, meaning that some subjects learned some button-shape associations correctly right away but took more time to or failed to learn other associations until the following reversal. Nonetheless, there were no differences between the groups in response accuracy in the Initial Learning condition (t(37) = 1.37, p = 0.18), the first-reversal condition (t(37) = 1.40, p = 0.17), the second reversal condition (t(37) = 0.04, p = 0.97), or across all conditions (t(37) = 1.05, p = 0.30).

The average response time was 0.91 seconds (SD = 0.18) in the HC group and 0.87 seconds (SD = 0.08) in the GAD group. The groups did not differ in response time in the Initial Learning condition (t(37) = 0.95, p = 0.35), the first reversal condition (t(37) = 1.19, p = 0.24), the second reversal condition (t(37) = 0.10, p = 0.92), or across all conditions (t(37) = 0.73, p = 0.47).

fMRI results

Basal ganglia

The control group (M = 0.62, SD = 1.47) showed greater activation in the basal ganglia (right Brodmann Area 48; x = 18, y = -14, z = 24, Z-score = 2.36) compared to the GAD group (M = -0.61, SD = 1.94) during the first reversal block (Figs. 4 and 6). There were no significant differences, between the groups, in the initial learning or second reversal blocks.

Hippocampus

The control group (M = 0.42, SD = 0.77) showed greater activation in the hippocampus (right Brodmann Area 54; x = -36, y = -32, z = -10, Z-score = 2.62) compared to the GAD group (M = -0.48, SD = 1.80) during the first reversal block (Figs. 5 and 7). There were no significant differences, between the groups, in the initial learning or second reversal blocks.

dlPFC

There were no significant differences between the groups, during any of the blocks, in the dlPFC. We have included post-hoc whole brain analysis results in Supplemental Table 1.



Figure 6. Basal ganglia activation.

Figure 7. Hippocampus activation.

Discussion

Using a simple fMRI button-association paradigm, the present study aimed to compare neural activation associated with degree of learning in a psychiatric population with GAD and healthy controls. Results revealed a significant hypoactivation of both the hippocampus and the basal ganglia in the GAD group compared to the HC group, in the first reversal block.

Behavioral results from the learning task showed that subjects from both groups understood the task and learned at a reasonable speed and to a reasonable degree of accuracy. There was no difference between the groups in degree of learning, response accuracy, or in response time, implying that overall learning was unaffected in the GAD group. These findings could suggest that GAD does not negatively impact associative learning, learning accommodation, or learning consolidation in stress-free environments, or that our sample was not large enough to capture behavioural group differences.

Findings in the basal ganglia revealed a significant hypoactivation in the GAD group in comparison to the HC group only for the first reversal bock. This difference was not present in the Initial Learning block and appeared to resolve by the second reversal, implying that anxiety may only affect brain activation in early stages of learning accommodation and re-consolidation. It is notable that this difference occurred during early set shifting, which is emblematic of the perseveration characteristic in GAD. However, HCs did not show a constant activation of the basal ganglia throughout the task, as was the case with the hippocampus. In this case, HCs increased activation of the basal ganglia during the first reversal block, meaning that they engaged this area in response to having to accommodate new button-image associations and having to re-consolidate previous associations for the first time. However, this increased activation dissipated by the second reversal. This pattern of findings is in line with previously reported evidence that the basal ganglia can be differentially activated during various forms of learning and memory tasks in healthy individuals such as working memory tasks (e.g., McNab and Klingberg, 2008; Moore et al., 2013), and positive and negative association learning (e.g., Seger, 2006), and habit learning (e.g.,

Packard and Knowlton, 2002). This is demonstrated by the fact that during first reversal, which requires intact associative learning and working memory, the healthy individuals showed increased basal ganglia and hippocampal activation (Figs. 4 and 5). However, while controls presented an increased activation of the basal ganglia in an initial learning consolidation and accommodation, GAD was associated with basal ganglia *hypoactivation* during learning accommodation/consolidation. Our finding supports previous research associating basal ganglia hypoactivation with GAD (Wu *et al.*, 1991), as well as previous studies showing basal ganglia activation in relation to learning in controls (Packard and Knowlton, 2002).

Similarly, results found in the hippocampus partially aligned with our original hypothesis. While HCs showed similar hippocampus activation throughout the task, GAD patients showed hypoactivation in this area in the block where initially learned associations changed for the first time. However, activation abnormalities appeared to resolve by the second reversal block, instead of remaining hypoactivated as we expected. This could be due to individuals with GAD engaging this ROI to a lesser extent while first adjusting or accommodating new memories or requiring more time to consolidate memories. However, given that our behavioural data does not show any processing speed deficits, these differences in basal ganglia and hippocampal activation could also reveal compensatory neural mechanisms during the early stages of memory consolidation or accommodation. While some previous research has dissociated hippocampal activation patterns in relation to anxiety and memory (Bannerman et al., 2004; McHugh et al., 2004; Bertoglio, et al., 2006), our finding suggests that anxiety and memory could functionally overlap in the hippocampus. Our result gives support to previous research associating hippocampal hypoactivation with anxiety (Mah, et al., 2016), as a result of stress-induced damage in the brain, which in its turn could have altered the activation of hippocampal regions more typically associated with memory.

As briefly mentioned above, the functional brain differences between the groups in the first reversal block were not observed in the behavioural data. This could have been caused by compensatory mechanisms that the GAD group may have engaged in, unlike the HC group. Indeed, according to the processing-efficiency hypothesis, individuals with GAD may require greater activation of certain brain regions supporting cognitive control (such as our ROIs), in order to maintain equivalent performance to healthy controls (Eysenck *et al.*, 2007). Another theory of cognitive control proposed that this reduced cognitive efficiency may be the result of changes in the temporal dynamics of these brain regions' recruitment (Fales *et al.*, 2008), which we also observed in our findings.

Finally, no significant results between the groups were found in the dlPFC, which did not align with our hypothesis. However, previous studies exploring the role of the dlPFC in learning in anxiety and anxiety-related disorders did not explore nonthreatening associative learning, nor did they explore reversal learning (Wheelock *et al.*, 2014; Balderston *et al.*, 2017). Hence, our null findings suggest either that a GAD diagnosis does not alter dlPFC activation during associative learning or associative reversal learning, or that our study was affected by a number of the limitations described below, which could have reduced our power to detect a difference in dlPFC activation between the clinical and control groups.

It must be noted that our results' interpretations are speculative, as our findings were constrained by a number of limitations. Most importantly, our study was composed of a small sample size, which reduced our statistical power and overall predictive validity. Additionally, while every individual in our psychiatric group had been diagnosed with GAD, all patients had also been diagnosed with comorbid mood, anxiety-related, and impulse-control disorders, increasing variability within the GAD group. Future research should test our findings using a larger and less heterogeneous psychiatric group in order to investigate specific associations with anxiety in particular.

Our findings suggest that GAD may affect regional brain activation in initial but not later stages of learning accommodation and consolidation. Particularly, the hippocampus and the basal ganglia may be areas of interest in the study of associative learning consolidation and accommodation in GAD in non-threatening environments, and treatments targeting these areas may have the potential to reduce anxiety symptomatology linked to learning difficulties.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/neu.2022.16.

Author contributions. Marta Migó substantially contributed to the conception, design, and analysis of data and drafted and revised the content. Tina Chou substantially contributed to the conception, design, and analysis of data, and drafted and revised the content. Alik Widge substantially contributed to the conception and design of this project and revised important intellectual content. Amy Peters substantially contributed to the conception and design of this project and revised important intellectual content. Kristen Ellard substantially contributed to the conception and design of this project and revised important intellectual content. Kristen Ellard substantially contributed to the conception and design of this project and revised important intellectual content. Darin Dougherty substantially contributed to the conception and design of this project and revised important intellectual content. Thilo Deckersbach substantially contributed to the conception and design of this project and revised important intellectual content. All authors approved of the final version of this paper.

Financial support. National Institute of Neurological Disorders and Stroke (NINDS), Grant/ Award Number: T32 NS100663-01; NIMH, Grant/Award Number: K23 MH074895- 01; Defense Advanced Research Projects Agency (DARPA), Grant/Award Number: W911NF-14-2-0045; NIH, Grant/ Award Number: R01 MH045573-24 and 1R01AT008563-01A1; Medtronic.

Competing interest. Migó has no conflict of interest to report.

Dr Chou's research is currently funded by NIH T32 NS100663 and the Tiny Blue Dot Foundation.

Dr Widge has pending patent applications related to electrographic markers and brain stimulation methods to ameliorate mental illness. Dr Widge has received device donations and consulting income from Medtronic and consulting income from Livanova, Dandelion Science and Circuit Therapeutics.

Dr Peters' research is currently funded by NIMH K23MH122676 and the Dauten Family Center for Bipolar Treatment Innovation.

Dr Ellard's research is currently funded by NIMH K23 MH120348.

Dr Dougherty's research has been funded by NIMH, NARSAD, IOCDF, Medtronic, Cyberonics, Eli Lilly, McNeil, and Northstar. He has received honoraria, consultation fees and/or royalties from Medtronic, Wyeth, Jazz Pharmaceuticals Bristol Myers Squibb, Brand Ideas, and Reed Elsevier.

Dr Deckersbach's research has been funded by NIMH, NARSAD, TSA, and OCF. He has received honoraria, consultation fees and/or royalties from Medacorp, MGH Psychiatry Academy, BrainCells Inc., Systems Research and Applications Corporation, Boston University, the Catalan Agency for Health Technology Assessment and Research, the National Association of Social Workers Massachusetts, and Oxford University Press. He has also participated in research funded by NIH, NIA, Janssen Pharmaceuticals, The Forest Research Institute, Shire Development Inc., Medtronic, Cyberonics, and Northstar.

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

References

- Ahn WY, Haines N and Zhang L (2017) Revealing neurocomputational mechanisms of reinforcement learning and decision-making with the HBayesDM Package. *Computational Psychiatry* 1(0), 24–57. DOI: 10.1162/ CPSY_a_00002.
- American Psychiatric Association (2013) Diagnostic and Statistical Manual of Mental Disorders, 5th edn. American Psychiatric Association.
- Ashcraft MH and Kirk EP (2001) The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General* 130(2), 224–237. DOI: 10.1037/0096-3445.130.2.224.
- Aylward J, Valton V, Ahn WY, Bond RL, Dayan P, Roiser JP and Robinson OJ (2019) Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nature Human Behaviour* 3(10), 1116–1123. DOI: 10.1038/s41562-019-0628-0.
- Balderston NL, Vytal KE, O'Connell K, Torrisi S, Letkiewicz A, Ernst M and Grillon C (2017) Anxiety patients show reduced working memory related DIPFC activation during safety and threat: research article: anxiety patients show reduced DIPFC activity. *Depression and Anxiety* 34(1), 25–36. DOI: 10.1002/da.22518.
- Bannerman DM, Rawlins JNP, McHugh SB, Deacon RMJ, Yee BK, Bast T, Zhang WN, Pothuizen HHJ and Feldon J (2004) Regional dissociations within the hippocampus—memory and anxiety. *Neuroscience & Biobehavioral Reviews* 28(3), 273–283. DOI: 10.1016/j.neubiorev.2004.03. 004.
- Barbey AK, Koenigs M and Grafman J (2013) Dorsolateral prefrontal contributions to human working memory. *Cortex* 49(5), 1195–1205. DOI: 10.1016/j.cortex.2012.05.022.
- Bartolo R and Averbeck BB (2020) Prefrontal cortex predicts state switches during reversal learning. *Neuron* **106**(6), 1044–1054. DOI: 10.1016/j.neuron. 2020.03.024.
- Bertoglio LJ, Joca SRL and Guimarães FS (2006) Further evidence that anxiety and memory are regionally dissociated within the hippocampus. *Behavioural Brain Research* 175(1), 183–188. DOI: 10.1016/j.bbr.2006.08.021.
- Bishop SJ (2009) Trait anxiety and impoverished prefrontal control of attention. *Nature Neuroscience* 12(1), 92–98. DOI: 10.1038/nn.2242.
- Cools R, Clark L, Owen AM and Robbins TW (2002) Defining the Neural Mechanisms of Probabilistic Reversal Learning Using Event-Related Functional Magnetic Resonance Imaging. *The Journal of Neuroscience* 22(11), 4563–4567. DOI: 10.1523/JNEUROSCI.22-11-04563.2002.

- Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research* 29(3), 162–173. DOI: 10.1006/cbmr.1996.0014.
- Derakshan N and Eysenck MW (2009) Anxiety, processing efficiency, and cognitive performance: new developments from attentional control theory. *European Psychologist* 14(2), 168–176. DOI: 10.1027/1016-9040.14.2.168.
- Engin E and Treit D (2007) The role of hippocampus in anxiety: intracerebral infusion studies. *Behavioural Pharmacology* 18(5-6), 365–374. DOI: 10.1097/ FBP.0b013e3282de7929.
- Eysenck MW, Derakshan N, Santos R and Calvo MG (2007) Anxiety and cognitive performance: attentional control theory. *Emotion* 7(2), 336–353. DOI: 10.1037/1528-3542.7.2.336.
- Fales CL, Barch DM, Burgess GC, Schaefer A, Mennin DS, Gray JR and Braver TS (2008) Anxiety and cognitive efficiency: differential modulation of transient and sustained neural activity during a working memory task. *Cognitive, Affective, & Behavioral Neuroscience* 8(3), 239–253. DOI: 10.3758/ CABN.8.3.239.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA and Noll DC (1995) Improved assessment of significant activation in Functional Magnetic Resonance Imaging (FMRI): use of a cluster-size threshold. *Magnetic Resonance in Medicine* 33(5), 636–647. DOI: 10.1002/ mrm.1910330508.
- Graybiel AM (1995) Building action repertoires: memory and learning functions of the basal ganglia. *Current Opinion in Neurobiology* 5(6), 733–741. DOI: 10.1016/0959-4388(95)80100-6.
- Harlé KM, Guo D, Zhang S, Paulus MP and Yu AJ (2017) "Anhedonia and anxiety underlying depressive symptomatology have distinct effects on reward-based decision-making" edited by E. Vasilaki. *PLoS One* 12, e0186473. DOI: 10.1371/journal.pone.0186473.
- Jarrard LE (1993) On the role of the hippocampus in learning and memory in the rat. *Behavioral and Neural Biology* **60**(1), 9–26. DOI: 10.1016/0163-1047(93)90664-4.
- LaFreniere LS and Newman MG (2019) Probabilistic learning by positive and negative reinforcement in generalized anxiety disorder. *Clinical Psychological Science* 7(3), 502–515. DOI: 10.1177/2167702618809366.
- Mah L, Szabuniewicz C and Fiocco AJ (2016) Can anxiety damage the brain? Current Opinion in Psychiatry 29(1), 56–63. DOI: 10.1097/YCO. 000000000000223.
- Maldjian JA, Laurienti PJ, Kraft RA and Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of FMRI data sets. *NeuroImage* 19(3), 1233–1239. DOI: 10.1016/S1053-8119(03)00169-1.
- Marchand WR (2010) Cortico-basal ganglia circuitry: a review of key research and implications for functional connectivity studies of mood and anxiety disorders. Brain Structure and Function 215(2), 73–96. DOI: 10.1007/ s00429-010-0280-y.
- McHugh SB, Deacon RMJ, Rawlins JNP and Bannerman DM (2004) Amygdala and ventral hippocampus contribute differentially to mechanisms of fear and anxiety. *Behavioral Neuroscience* **118**(1), 63–78. DOI: 10.1037/ 0735-7044.118.1.63.
- McNab F and Klingberg T (2008) Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience* 11(1), 103–107. DOI: 10.1038/nn2024.
- Moore AB, Li Z, Tyner CE, Hu X and Crosson B (2013) Bilateral basal ganglia activity in verbal working memory. *Brain and Language* **125**(3), 316–323. DOI: 10.1016/j.bandl.2012.05.003.

- Morris BH and Rottenberg J (2015) Heightened reward learning under stress in generalized anxiety disorder: a predictor of depression resistance? *Journal of Abnormal Psychology* **124**(1), 115–127. DOI: 10.1037/a0036934.
- Myers CE, Shohamy D, Gluck MA, Grossman S, Kluger A, Ferris S, Golomb J, Schnirman G and Schwartz R (2003) Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *Journal of Cognitive Neuroscience* 15(2), 185–193. DOI: 10.1162/089892903321208123.
- Packard MG and Knowlton BJ (2002) Learning and memory functions of the basal ganglia. Annual Review of Neuroscience 25(1), 563–593. DOI: 10.1146/ annurev.neuro.25.112701.142937.
- Piray P, Ly V, Roelofs K, Cools R and Toni I (2019) Emotionally aversive cues suppress neural systems underlying optimal learning in socially anxious individuals. *The Journal of Neuroscience* **39**(8), 1445–1456. DOI: 10.1523/ JNEUROSCI.1394-18.2018.
- Seger CA (2006) The basal ganglia in human learning. *The Neuroscientist* **12**(4), 285–290. DOI: 10.1177/1073858405285632.
- Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R and Dunbar GC (1998) The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry* 59, 22–33; quiz 34-57.
- Van Der Kouwe AJW, Benner T, Salat DH and Fischl B (2008) Brain morphometry with multiecho MPRAGE. NeuroImage, 40(2), 559–569. DOI: 10.1016/j.neuroimage.2007.12.025.
- Vytal KE, Arkin NE, Overstreet C, Lieberman L and Grillon C (2016) Inducedanxiety differentially disrupts working memory in generalized anxiety disorder. BMC Psychiatry 16(1), 62. DOI: 10.1186/s12888-016-0748-2.
- Wheelock MD, Sreenivasan KR, Wood KH, Hoef LWV, Deshpande G and Knight DC (2014) Threat-related learning relies on distinct dorsal prefrontal cortex network connectivity. *NeuroImage* 102, 904–912. DOI: 10.1016/j. neuroimage.2014.08.005.
- Widge AS, Ellard KK, Paulk AC, Basu I, Yousefi A, Zorowitz S, Gilmour A, Afzal A, Deckersbach T, Cash SS, Kramer MA, Eden UT, Dougherty DD, Eskandar EN (2017) Treating refractory mental illness with closed-loop brain stimulation: progress towards a patient-specific transdiagnostic approach. *Experimental Neurology* 287(Pt 4), 461–472. DOI: 10.1016/j. expneurol.2016.07.021.
- Wilson A, Fern A, Ray S and Tadepalli P (2007) Multi-Task Reinforcement Learning: A Hierarchical Bayesian Approach. Proceedings of the 24th international conference on Machine learning - ICML '07, Corvalis, Oregon: ACM Press, 1015–22.
- Wu JC, Buchsbaum MS, Hershey TG, Hazlett E, Sicotte N and Johnson JC (1991) PET in generalized anxiety disorder. *Biological Psychiatry* 29(12), 1181–1199. DOI: 10.1016/0006-3223(91)90326-H.
- Xue G, Xue F, Droutman V, Lu Z-L, Bechara A and Read S (2013) "Common neural mechanisms underlying reversal learning by reward and punishment" edited by S. A. Simon. *PLoS One* 8, e82169. DOI: 10.1371/journal.pone. 0082169.
- Yousefi A, Paulk AC, Basu I, Mirsky JL, Dougherty DD, Eskandar EN, Eden UT and Widge A-S (2019) COMPASS: an open-source, general-purpose software toolkit for computational psychiatry. *Frontiers in Neuroscience* 12, 957. DOI: 10.3389/fnins.2018.00957.
- Zorowitz S, Rockhill AP, Ellard KK, Link KE, Herrington T, Pizzagalli DA, Widge AS, Deckersbach T and Dougherty DD (2019) The neural basis of approach-avoidance conflict: a model based analysis. *Eneuro* 6(4). DOI: 10.1523/ENEURO.0115-19.2019.