Received: 28 April 2023 | Revised: 25 September 2023 | Accepted: 7 October 2023

DOI: 10.1002/hipo.23583

RAPID COMMUNICATION

WILEY

Perceived plausibility modulates hippocampal activity in episodic counterfactual thinking

Kaylee Miceli[1,2] | Ricardo Morales-Torres[2,3] | Ari Khoudary[4] | Leonard Faul[2] | Natasha Parikh[5] | Felipe De Brigard[1,2,3]

[1]Department of Philosophy, Duke University, Durham, North Carolina, USA

[2]Center for Cognitive Neuroscience, Levine Science Research Center, Duke University, Durham, North Carolina, USA

[3]Department of Psychology and Neuroscience, Duke University, Durham, North Carolina, USA

[4]Department of Cognitive Sciences, University of California, Irvine, Irvine, California, USA

[5]Department of Psychology and Neuroscience, University of North Carolina, Chapel Hill, Chapel Hill, North Carolina, USA

Correspondence
Felipe De Brigard, Center for Cognitive Neuroscience, Levine Science Research Center, Duke University, 308 Research Drive, Room C03E, Durham, NC 27708, USA.
Email: felipe.debrigard@duke.edu

Abstract

Episodic counterfactual thinking (ECT) consists of imagining alternative outcomes to past personal events. Previous research has shown that ECT shares common neural substrates with episodic future thinking (EFT): our ability to imagine possible future events. Both ECT and EFT have been shown to critically depend on the hippocampus, and past research has explored hippocampal engagement as a function of the perceived plausibility of an imagined future event. However, the extent to which the hippocampus is modulated by perceived plausibility during ECT is unknown. In this study, we combine two functional magnetic resonance imaging datasets to investigate whether perceived plausibility modulates hippocampal activity during ECT. Our results indicate that plausibility parametrically modulates hippocampal activity during ECT, and that such modulation is confined to the left anterior portion of the hippocampus. Moreover, our results indicate that this modulation is positive, such that increased activity in the left anterior hippocampus is associated with higher ratings of ECT plausibility. We suggest that neither effort nor difficulty alone can account for these results, and instead suggest possible alternatives to explain the role of the hippocampus during the construction of plausible and implausible ECT.

KEYWORDS

episodic counterfactual thinking, episodic future thinking, hippocampus, parametric modulation, perceived plausibility

In the past two decades, a substantial amount of neuropsychological and neuroimaging research has revealed that our capacity to imagine possible future personal events (episodic future thinking [EFT]) shares common neural structures with our capacity to imagine possible ways past personal events could have occurred but did not (episodic counterfactual thinking [ECT]; Schacter et al., 2015). The neural structures common to these two kinds of episodic hypothetical thoughts largely overlap with core regions of the brain's default mode network, including the hippocampus (De Brigard, Addis, et al., 2013). Indeed, evidence suggests that individuals with schizophrenia, who show marked hippocampal volumetric reduction, present parallel deficits in EFT and ECT (Hooker et al., 2000). Likewise, individuals with hippocampal amnesia also show parallel difficulties in EFT and ECT (Mullally & Maguire, 2014). And more recently, it has been shown that patients in relapsing–remitting stages of multiple sclerosis show deficits in EFT and ECT predicted by a lack of white matter integrity in the hippocampal portion of the cingulum (Ayala et al., 2022).

Despite these commonalities, both EFT and ECT differ along several dimensions (De Brigard & Parikh, 2019). One such dimension, for instance, is temporal. Although both EFT and ECT involve hypothetical events, the former involves thoughts about possible *future* episodes while the latter involves thoughts about possible *past* events. Exploring differences along this temporal dimension for both EFT and ECT has been the subject of a number of behavioral and neuroimaging

Kaylee Miceli and Ricardo Morales-Torres share equal co-authorship.

 |

studies (for reviews, see De Brigard & Parikh, 2019; Schacter et al., 2015). Less explored, however, is the modal dimension: the degree to which an imagined hypothetical event is perceived as more or less plausible. Although some studies have explored this dimension behaviorally (e.g., De Brigard, Addis, et al., 2013; Szpunar & Schacter, 2013, De Brigard et al., 2013), much less is known about its neural effects. Moreover, prior research on the effect of plausibility on hippocampal activity during EFT has yielded mixed results. Some studies have shown increased activity for less plausible EFT (Weiler et al., 2010), while others have shown that less plausible EFT are associated with reduced hippocampal activity (Roberts et al., 2017). However, the extent to which perceived plausibility during ECT modulates hippocampal activity is unknown.

Thus, the purpose of this study was to investigate the impact of perceived plausibility on hippocampal activity during ECT. We analyzed imaging data from two functional magnetic resonance imaging (fMRI) experiments that involved a total of 53 participants (Experiment 1: $N = 22$, 14 female, $M_{age} = 22.68$; $SD = 2.97$; Parikh et al., 2018; Experiment 2: $N = 31$, 22 female, $M_{age} = 21.41$; $SD = 2.88$; Khoudary et al., 2022) who were asked to engage in ECT and rate the plausibility of the imagined alternative events. Both experiments involved two sessions. In session one, which occurred outside the scanner, participants used a set of common event cues to generate a list of autobiographical events (180 in Experiment 1; 80 in Experiment 2). Session two, which occurred approximately 1 week later, took place in the scanner. Participants were cued with self-generated titles for their autobiographical memories and were asked to mentally generate a counterfactual alternative for each memory. They constructed and elaborated upon the counterfactual for a few seconds (12 s in Experiment 1; 8 s in Experiment 2), and then rated it along several dimensions. Critically, both experiments included a plausibility rating (1–7 scale in Experiment 1 and 1–4 scale in Experiment 2, with higher scores indicating greater plausibility). Of note, Experiment 1 contained both episodic and semantic counterfactuals. While this study modeled both the episodic and semantic counterfactual conditions in first-level models, we only analyzed trials belonging to the episodic counterfactual conditions for the parametric modulation assessment. Experiment 2, by contrast, involved an “internal” versus “external” manipulation, whereby participants were asked to engage in ECT by either imagining an alternative way they could have acted (internal) or an alternative way in which something in the context of the event could have occurred differently (external). Yet, since both conditions involved ECT, all trials in this experiment were analyzed for parametric modulation. After all, our main interest is to explore the role of hippocampal activity during ECT in general, regardless of the content of the mental mutation. For details on the preprocessing of fMRI data, read the Supporting Information.

Functional MRI data were modeled via univariate analyses using FMRIB software library (FSL) (Jenkinson et al., 2012). For each subject and run, we used a general linear model that included the onset and duration of each counterfactual simulation and the subsequent rating screens. We also included a regressor to model button presses associated with simulation construction and rating responses in order to account for any motor-related activity. Each regressor was convolved with a double-gamma hemodynamic response function. Crucially, to examine whether perceived plausibility modulated hippocampal activity, we included a parametric regressor whose intensity value was modulated by the within-run *z*-scored plausibility ratings. The unmodulated plausibility regressor, rating screen regressors, and their temporal derivatives were also included as regressors in modeling each individual functional run. Subsequently, we combined functional runs within each subject, and extracted the hippocampal estimates using the bilateral hippocampal region of interest (ROI) from the Harvard-Oxford subcortical atlas (Figure 2a). Additionally, we extracted hippocampal subareas estimates using both the anterior and posterior hippocampal masks from the Brainnetome Atlas, which utilizes functional and structural connectivity data to generate these parcellations (Fan et al., 2016).

Subjects with *z*-scores greater than 2.9 or smaller than −2.9 on the ROI estimates were excluded, as this was the most extreme value across experiments, meaning that no subjects were removed from Experiment 1, and only one from Experiment 2. Statistical analyses were performed in R (Version 4.0.2); linear mixed-effects models were fit with the lme4 package (Bates, 2015), significance values for the fitted coefficients were estimated with the lmerTest package (Kuznetsova et al., 2017), and visualizations were made using ggplot2 (Wickham, 2011). The linear mixed-effects model included the hippocampal hemisphere as a fixed factor and nested participants within-experiment as random factors, which allowed us to have a separate intercept for each experiment. We then conducted a Student's *t*-test to test the activity of the left and right hippocampus separately for each experiment, contrasting the region of interest (ROI) values against 0.

Ratings of plausibility are plotted in Figure 1. For Experiment 1, the mean rating of plausibility for ECT was 4.5 ($SD = 0.91$, $SEM = 0.16$), whereas for Experiment 2, the mean rating was 2.71 ($SD = 0.39$, $SEM = 0.08$). The mixed-effects model in which we grouped the effect of the parametric modulation analysis of both fMRI experiments revealed that the right hippocampal ROI was not significantly different from 0 (modeled as the intercept; $b = 0.61$, $t(1.06) = 0.9$, $p = .52$), whereas the left hippocampal ROI showed a positive effect ($b = 0.62$, $t(51) = 2.39$, $p = .02$). To further probe this result, we used *t*-tests to examine hemispheric effects independently in each study. As shown in Figure 2, The right hippocampal ROI was not significantly different from 0 in Experiment 1 ($t(21) = 1.51$, $p = .15$) or Experiment 2 ($t(29) = 1.23$, $p = .227$). By contrast, the left hippocampal ROI was significantly greater than 0 in both Experiment 1 ($t(21) = 2.12$, $p = .046$) and Experiment 2 ($t(29) = 2.23$, $p = .034$).

We also separately examined if plausibility modulated the anterior or posterior subareas of the hippocampus. In Experiment 1 ($t(21) = 2.38$, $p = .027$) and Experiment 2 ($t(29) = 2.49$, $p = .021$), only the anterior subareas of the left hippocampus was positively parametrically modulated by plausibility. However, no significant modulation was observed in the posterior subareas of the left hippocampus (Experiment 1: $t(21) = 1.57$, $p = .13$; Experiment 2: $t(29) = 1.94$, $p = .06$). Likewise, no significant modulation was observed for either the anterior or the posterior subareas of the right hippocampus (all *p*-values >.12, see Supporting Information for details). In summary, we found that perceived plausibility of an imagined counterfactual event

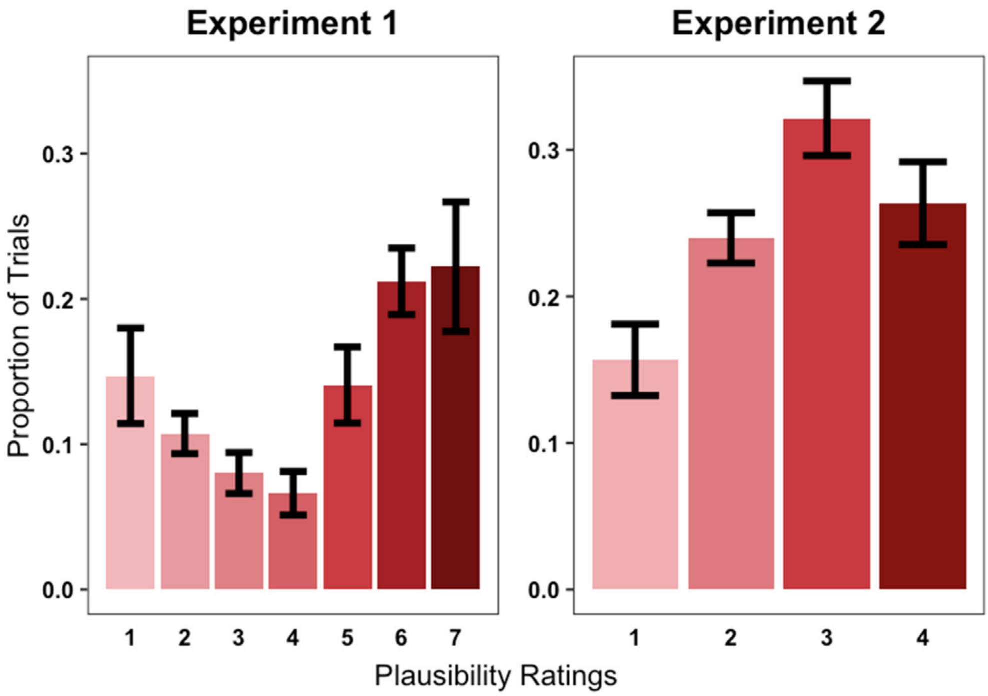

FIGURE 1 Proportion of trials for each plausibility rating averaged across subjects. Error bars represent standard error or the mean.

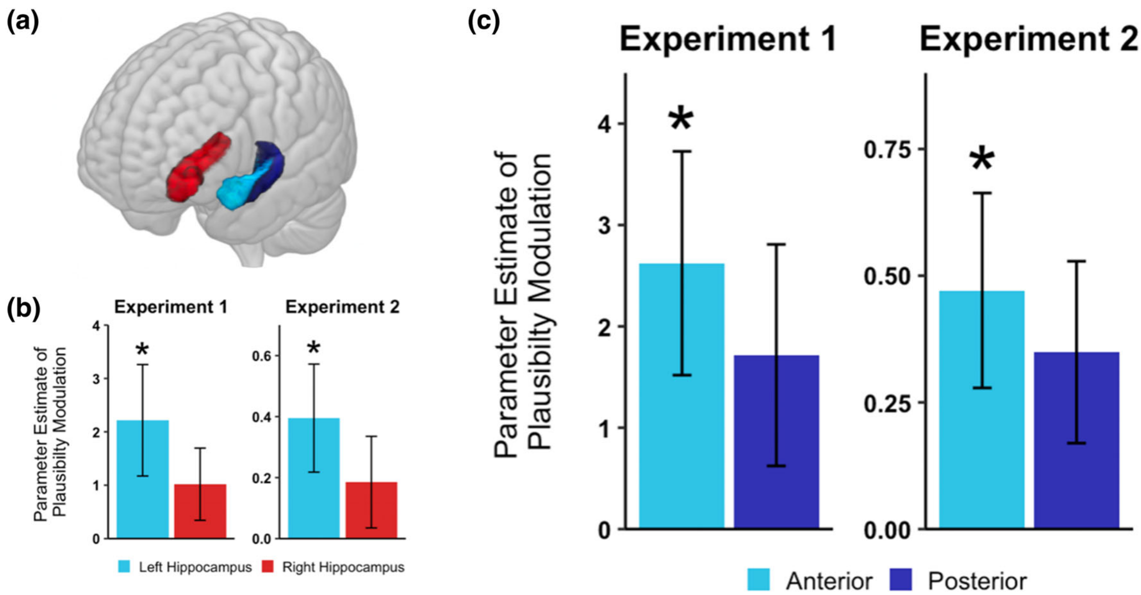

FIGURE 2 (a) Bilateral hippocampal ROI mask used to extract parameter estimates. The color red represents the mask used to extract the average values of both the left and right hippocampal hemispheres (only the right hemisphere mask is displayed). Cyan and blue represent the masks for extracting values from the anterior and posterior subareas of the hippocampus, respectively (only the left hemisphere masks are shown). (b) Average value for the hemispheric parameter estimates from the plausibility parametric modulation analysis, as a function of hippocampal ROI and experiment. Only the left hippocampus exhibited a significant increase in activation with greater plausibility ratings. (c) Average value of the parameter estimates for the left anterior and left posterior hippocampal subareas. For both experiments, the increase in activation with greater plausibility was restricted to the anterior subareas. Error bars represent the standard error of the mean. *$p < .05$.

parametrically modulated activity in the anterior subareas of the left hippocampus during the generation of an episodic counterfactual simulation, with neural activity increasing as a function of increased perceived plausibility, but we observed no modulatory effect in the right hippocampus.

In order to confirm the left lateralization of the parametric modulation of hippocampal activity as a function of plausibility, we followed the approach of van Mulukom et al. (2013) and computed a laterality index (LI; Wilke & Lidzba, 2007) for the anterior subareas of the hippocampus. This index measures the asymmetry of brain activation by calculating a ratio based on the number of voxels that surpass a bootstrapped threshold in each hemisphere. LI values range from −1 (indicating right-lateralized activity) to +1 (indicating left-lateralized activity). Consistent with the above results, in both Experiment 1 (LI = 0.58) and Experiment 2 (LI = 0.42) we found evidence of left lateralization of the parametric modulation of plausibility in the hippocampus.

Taken together, our findings indicate that as ratings of perceived plausibility of an episodic counterfactual episode increase, so does activity in the anterior subareas of the left hippocampus. These results contribute to a growing body of literature characterizing neural differences between types of episodic hypothetical thoughts. Specifically, we address an empirical gap concerning how plausibility modulates hippocampal activity during ECT. Because plausibility is a key dimension along which types of hypothetical thoughts can differ, it may be helpful to compare our findings with related results about the effect of perceived plausibility on hippocampal activity during EFT (an important comparison class of hypothetical thoughts that we did not empirically investigate in the present work). In one study, Weiler et al. (2010) asked participants to engage in EFT with general cues, and they found that right anterior hippocampal activity was negatively modulated by perceived plausibility, so that the more implausible the EFT, the more it recruited the right anterior portion of the hippocampus. They interpreted this finding as suggesting that implausible EFT demands more binding of unconnected episodic details than those perceived as plausible, and this increase in demands drives the increase in hippocampal activity. This explanation, however, is not available for our results, as they show a positive modulation instead. If the parametric increase in hippocampal activity as a function of perceived plausibility were to reflect increased binding demands during episodic simulation, then it should have been observed during ECT too, as presumably an implausible ECT requires the binding of episodic details that are more distantly associated with the original episodic memory than a plausible ECT (De Brigard et al., 2021; De Brigard & Parikh, 2019).

Interestingly, in a more recent study by Roberts et al. (2017), when participants engaged in EFT involving components that were coherent or congruent with each other, and thus were judged as more plausible, activity in the hippocampus was greater relative to when they engaged in less plausible EFT involving incongruent components. The inconsistency between these two findings (Roberts et al., 2017; Weiler et al., 2010) suggests that perceived plausibility during episodic hypothetical thinking—which both EFT and ECT belong to (De Brigard, 2014)—is likely more than just a matter of detail or effort. Extant evidence shows, for instance, that although difficulty affects perceived plausibility during ECT, it is not identical to it (Byrne, 2016). Moreover, the results from Khoudary et al. (2022) show that the association between perceived plausibility and difficulty during simulation is only moderate ($R_c^2 = 0.3$, see Supporting Information for details), corroborating that there is quite a bit of variance that is not shared by these two dimensions. Likewise, extant behavioral and neural evidence also suggest that, although related, the amount of detail experienced during the episodic simulation does not fully correlate with perceived plausibility, as it is possible to imagine detailed implausible hypothetical events (De Brigard et al., 2021). This claim is supported by data from Experiment 1, in which the level of detail of an ECT shows only a weak association with its plausibility ($R_c^2 = 0.16$, see Supporting Information for details).

Fully appreciating the role of plausibility in modulating hippocampal activity during episodic hypothetical simulations perhaps requires researchers to move away from simple accounts in terms of difficulty or amount of detail. Instead, we think that a useful strategy is to relate research in episodic simulation to extant computational models of plausibility in hypothetical thought, which suggest that plausibility relates to norm-violation (Phillips et al., 2019) as well as conceptual-coherence with prior knowledge (Connell & Keane, 2006). Since the hippocampus has been associated with both statistical and categorical learning (Schapiro et al., 2017), we think that a fruitful avenue for future research would be to explore the role of prior-knowledge and conceptual-coherence in the generation of both ECT and EFT.

Our results also suggest that there may be lateralized effects of plausibility in hippocampal activity during episodic hypothetical thinking. Although studies on EFT have shown preferential modulation of the right, as opposed to left, hippocampus during EFT generation (Addis et al., 2011; Addis & Schacter, 2012; Martin et al., 2011), it is unclear if this lateralization is due to the temporal direction of the episodic simulation, or if there is something specific to the content of the hypothetical thoughts that were generated. Further studies directly contrasting EFT and ECT may be able to shed light on this issue.

Finally, it is worth mentioning a few limitations of this study. First, although differences in detail and emotion likely cannot account for all of the variance in judgments of plausibility, these are important dimensions that we could not include as regressors in this study, but that likely should be included in future research. Second, the semantics of modal terms such as "plausible," "likely," "necessary," and "possible" are notoriously complex (Portner, 2009). As such, it may be possible that people interpret the notion of "plausibility" differently depending on whether they are thinking about EFT or ECT. Extant evidence suggests that people have no difficulty in judging counterfactual thoughts as more or less plausible (e.g., De Brigard et al., 2021), but further research should explore the hypothesis that these terms could be used systematically differently depending on the kind of hypothetical simulation. Finally, it may be possible that the differential contribution of the hippocampus during ECT and EFT is in turn influenced by functional connections with other brain regions. Future studies exploring this question should consider using connectivity analysis, and other whole-brain approaches, to further understand the role of the hippocampus during these two kinds of episodic hypothetical simulations.

ACKNOWLEDGMENT

We thank the members of the IMC Lab for their help while writing this paper.

DATA AVAILABILITY STATEMENT

Data and scripts for this study can be found at https://github.com/IMC-Lab/HippModCFT. Data, code, and CSV files for the original Experiment 2 (Khoudary et al., 2022) are available at https://github.com/IMC-Lab/conCFT. Unthresholded *z*-statistics for all of the univariate analyses of that study are available at https://identifiers.org/neurovault.collection:12098. For Experiment 1 (Parikh et al., 2018), data can be made available upon request.

ORCID

Kaylee Miceli https://orcid.org/0000-0002-6474-0812
Felipe De Brigard https://orcid.org/0000-0003-0169-1360

REFERENCES

Addis, D. R., Cheng, T., Roberts, R, P., & Schacter, D. L. (2011). Hippocampal contributions to the episodic simulation of specific and general future events. *Hippocampus*, *21*(10), 1045–1052.

Addis, D. R., & Schacter, D. L. (2012). The hippocampus and imagining the future: Where do we stand? *Frontiers in Human Neuroscience*, *5*, 173.

Ayala, O. D., Banta, D., Hovhannisyan, M., Duarte, L., Lozano, A., García, J. R., … De Brigard, F. (2022). Episodic past, future, and counterfactual thinking in relapsing-remitting multiple sclerosis. *NeuroImage: Clinical*, *34*, 103033.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Byrne, R. M. (2016). Counterfactual thought. *Annual Review of Psychology*, *67*, 135–157.

Connell, L., & Keane, M. T. (2006). A model of plausibility. *Cognitive Science*, *30*(1), 95–120. https://doi.org/10.1207/s15516709cog0000_53

De Brigard, F. (2014). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese*, *191*(2), 155–185.

De Brigard, F., Addis, D. R., Ford, J. H., Schacter, D. L., & Giovanello, K. S. (2013). Remembering what could have happened: Neural correlates of episodic counterfactual thinking. *Neuropsychologia*, *51*(12), 2401–2414.

De Brigard, F., Henne, P., & Stanley, M. L. (2021). Perceived similarity of imagined possible worlds affects judgments of counterfactual plausibility. *Cognition*, *209*, 104574.

De Brigard, F., & Parikh, N. (2019). Episodic counterfactual thinking. *Current Directions in Psychological Science*, *28*(1), 59–66.

De Brigard, F., Szpunar, K. K., & Schacter, D. L. (2013). Coming to grips with the past: Effect of repeated simulation on the perceived plausibility of episodic counterfactual thoughts. *Psychological Science*, *24*(7), 1329–1334.

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., Fox, P. T., Eickhoff, S. B., Yu, C., & Jiang, T. (2016). The human brainnetome atlas: A new brain atlas based on connectional architecture. *Cerebral Cortex*, *26*(8), 3508–3526. https://doi.org/10.1093/cercor/bhw157

Hooker, C., Roese, N. J., & Park, S. (2000). Impoverished counterfactual thinking is associated with schizophrenia. *Psychiatry*, *63*(4), 326–335.

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790.

Khoudary, A., O'Neill, K., Faul, L., Murray, S., Smallman, R., & De Brigard, F. (2022). Neural differences between internal and external episodic counterfactual thoughts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1866), 20210337.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26.

Martin, V. C., Schacter, D. L., Corballis, M. C., & Addis, D. R. (2011). A role for the hippocampus in encoding simulations of future events. *Proceedings of the National Academy of Sciences*, *108*(33), 13858–13863.

Mullally, S. L., & Maguire, E. A. (2014). Memory, imagination, and predicting the future: A common brain mechanism? *The Neuroscientist*, *20*(3), 220–234.

Parikh, N., Ruzic, L., Stewart, G. W., Spreng, R. N., & De Brigard, F. (2018). What if? Neural activity underlying semantic and episodic counterfactual thinking. *NeuroImage*, *178*, 332–345.

Phillips, J., Morris, A., & Cushman, F. (2019). How we know what not to think. *Trends in Cognitive Sciences*, *23*(12), 1026–1040.

Portner, P. (2009). *Modality*. OUP Oxford.

Roberts, R. P., Wiebels, K., Sumner, R. L., van Mulukom, V., Grady, C. L., Schacter, D. L., & Addis, D. R. (2017). An fMRI investigation of the relationship between future imagination and cognitive flexibility. *Neuropsychologia*, *95*, 156–172.

Schacter, D. L., Benoit, R. G., De Brigard, F., & Szpunar, K. K. (2015). Episodic future thinking and episodic counterfactual thinking: Intersections between memory and decisions. *Neurobiology of Learning and Memory*, *117*, 14–21.

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1711), 20160049.

Szpunar, K. K., & Schacter, D. L. (2013). Get real: Effects of repeated simulation and emotion on the perceived plausibility of future experiences. *Journal of Experimental Psychology: General*, *142*(2), 323–327. https://doi.org/10.1037/a0028877

van Mulukom, V., Schacter, D. L., Corballis, M. C., & Addis, D. R. (2013). Re-imagining the future: Repetition decreases hippocampal involvement in future simulation. *PLoS One*, *8*(7), e69596.

Weiler, J. A., Suchan, B., & Daum, I. (2010). Foreseeing the future: Occurrence probability of imagined future events modulates hippocampal activation. *Hippocampus*, *20*(6), 685–690.

Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*, *3*(2), 180–185.

Wilke, M., & Lidzba, K. (2007). LI-tool: A new toolbox to assess lateralization in functional MR-data. *Journal of Neuroscience Methods*, *163*(1), 128–136.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Miceli, K., Morales-Torres, R., Khoudary, A., Faul, L., Parikh, N., & De Brigard, F. (2024). Perceived plausibility modulates hippocampal activity in episodic counterfactual thinking. *Hippocampus*, *34*(1), 2–6. https://doi.org/10.1002/hipo.23583