PNAS RESEARCH ARTICLE | NEUROSCIENCE

A neural signature of the vividness of prospective thought is modulated by temporal proximity during intertemporal decision making

Sangil Lee[a,1], Trishala Parthasarathi[b], Nicole Cooper[c], Gal Zauberman[d], Caryn Lerman[e], and Joseph W. Kable[a]

Edited by Daniel Schacter, Harvard University, Cambridge, MA; received August 19, 2022; accepted September 27, 2022

Why do people discount future rewards? Multiple theories in psychology argue that one reason is that future events are imagined less vividly than immediate events, thereby diminishing their perceived value. Here we provide neuroscientific evidence for this proposal. First, we construct a neural signature of the vividness of prospective thought, using an fMRI dataset where the vividness of imagined future events is orthogonal to their valence by design. Then, we apply this neural signature in two additional fMRI datasets, each using a different delay-discounting task, to show that neural measures of vividness decline as rewards are delayed farther into the future.

fMRI | construal level theory | imagination | prospection | delay discounting

Significance

People tend to devalue, or discount, outcomes in the future relative to those that are more immediate; future discounting can explain people's difficulty in choosing healthier food or saving money for retirement. Several psychological theories propose that one reason delayed outcomes are discounted is that they are perceived less vividly than more immediate ones. Here we build a brain decoder for the vividness of future thought and apply it to participants' brain activity when they make decisions to show that the farther the outcomes occur into the future, the less vividly they are processed. This approach showcases the use of brain decoders to index unobtrusively thoughts or feelings that are otherwise difficult to measure.

Many of the most important choices we make in our daily lives involve tradeoffs between the present and the future. Should I spend money now or save it for retirement? Can I forego the pleasure of eating this dessert now in order to reach my weight loss goal and improve my health? In such intertemporal decisions, humans tend to devalue, or discount, outcomes in the future: a phenomenon known as delay discounting. In the laboratory, this tendency can be measured by presenting participants with choices between a smaller monetary amount available immediately and a larger monetary amount available after a delay. Patience as measured by laboratory intertemporal choice tasks predicts other important aspects of life such as drug and alcohol abuse, educational attainment, and personal finances (1–7).

Why, however, are delayed outcomes fundamentally less desirable? Psychologists have long pondered this important question. Although multiple factors may contribute to discounting (8), several theories suggest that one important factor is that distant outcomes are imagined less vividly than proximal ones. Rick and Loewenstein (9) have pointed out that in many intertemporal decisions, delayed outcomes are intrinsically less tangible than sooner ones. For example, while a calorie-rich dessert yields immediately perceivable pleasure for the eater, the promise of better future health is less appreciable. Similarly, construal level theory proposes that even when future outcomes are not intrinsically less tangible, people tend to use a process of high-level construal when thinking about future events that leads to their being represented in a relatively more abstract way (10, 11). In contrast, when people consider sooner events, they use low-level construal and represent them in a relatively more concrete manner. Many studies have provided some support for the central claim that the same outcome is perceived less concretely when it occurs farther in the future rather than more immediately (for review, see ref. 11) and linked these to changes in representation of delay discounting (12, 13). A related literature has shown that vividly imagining future events reduces delay discounting, again suggesting that delayed outcomes are discounted because they are perceived less vividly (14–16) (for review, see ref. 15). However, an ideal test of whether future outcomes are perceived less vividly during intertemporal decision making would measure vividness online and nonobtrusively while people are making intertemporal decisions. Functional brain imaging (fMRI) has the potential to provide such a nonobtrusive online test.

While many fMRI studies have compared brain activity for sooner versus later outcomes (for review, see ref. 17), this comparison does not isolate activity attributed solely to the vividness of outcomes, as sooner outcomes are generally valued more highly than delayed ones; that is, brain activity selectively responding for sooner versus later outcomes may reflect valuation, not necessarily vividness. Indeed, several previous studies that have compared sooner and later outcomes have found increased activity in the medial prefrontal cortex (mPFC) and posterior cingulate cortex (PCC), two regions with well-established roles in valuation (18–22). Also, studies that have used episodic imagery of future events to reduce delay discounting cannot easily dissociate the neural

Author affiliations: [a]Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104; [b]Department of Neuroscience, University of Pennsylvania, Philadelphia, PA 19104; [c]Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA 19104; [d]Yale School of Management, Yale University, New Haven, CT 06511; and [e]University of Southern California Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA 90033

Author contributions: T.P., N.C., G.Z., C.L., and J.W.K. designed research; S.L., T.P., and N.C. performed research; S.L. contributed new reagents/analytic tools; S.L. analyzed data; and S.L., T.P., N.C., G.Z., C.L., and J.W.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: lsangil@sas.upenn.edu.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2214072119/-/DCSupplemental.

Published October 24, 2022.

effects that are due to episodic imagery enhancing the vividness of future outcomes versus directly increasing the valuation of future outcomes (14, 16).

In the current study, we use a whole-brain multivariate predictor approach that provides a principled way to perform reverse inference by considering the pattern of activity across multiple brain regions and formally testing the predictive power of this neural pattern (23). For the prediction algorithm, we use thresholded partial least squares (T-PLS) (24) to construct a whole-brain multivariate neural predictor of the vividness of imagined future events. To train this predictor, we use a prospective imagination dataset (22), in which the vividness (high versus low) and valence (positive versus negative) of imagined future events were orthogonal, such that the subsequent neural predictor was specific to vividness and not valence. We then applied the whole-brain neural predictor of vividness in two separate delay-discounting task datasets with different evaluation schemes (bidding versus choice) to test whether the temporal distance of monetary options in intertemporal decision making modulates the neural signature of vivid imagination.

Results

Our first goal was to assess whether it is possible to create a whole-brain neural predictor of imagination vividness using T-PLS (Fig. 1), that does not also predict imagination valence. We used an fMRI dataset of 24 participants imagining possible future events that had been categorized a priori as high versus low in vividness and positive versus negative in valence (ref. 22 and Fig. 2*A*). Using a nested cross-validation approach, we built a whole-brain predictor using 23 participants at a time and tested whether it can predict the vividness but not the valence of imagined future events. Across the 24-folds, the T-PLS whole-brain predictors used between 8 and 17 partial least square components and were mostly thresholded to be around 0.5% of the total voxels (Fig. 2*B*).

We found that T-PLS can create neural predictors of imagination vividness that, in out-of-sample prediction, successfully discriminate the trial categories of high versus low vividness [mean prediction area under the receiver operating characteristic curve (AUC) = 62.06%, *t* test against 50%, $t(23) = 5.21$, Bonferroni $P < 0.0001$, Cohen's $d = 1.06$], but not the trial categories of positive versus negative valence [mean prediction AUC = 47.30%, *t* test against 50%, $t(23) = -1.72$, Bonferroni $P = 0.39$; Fig. 2*C*]. We found similar results with the participants' ratings of the vividness and valence of imagined future events (Fig. 2*D*). Mean out-of-sample correlation between the neural vividness score and vividness ratings was $r = 0.15$ [$t(23) = 3.55$, Bonferroni $P = 0.0068$, Cohen's $d = 0.73$], while the correlation between the neural vividness score and valence ratings was $r = -0.06$ [$t(23) = -2.14$, Bonferroni $P = 0.17$]. Finally, we verified that neural vividness scores were not significantly correlated with any of the other 10 aspects of imagination rated for these stimuli (e.g., arousal, subjective temporal distance; *SI Appendix*, section I).

The whole-brain prediction map of vividness involved various regions of the brain, mostly in a bilateral fashion (Fig. 3 and Table 1). Positive coefficients (predictive of higher vividness) were found in bilateral hippocampus, bilateral central orbitofrontal cortex (OFC), and left middle occipital gyrus. Negative coefficients (predictive of lower vividness) were found in bilateral precuneus/PCC, left inferior temporal gyrus, and right postcentral gyrus.

We next applied this whole-brain predictor of vividness in two separate delay-discounting tasks, in order to test whether the neural signature of vivid future thinking is higher when considering sooner rewards and lower when considering later rewards. In both datasets, we found that neural vividness scores were negatively correlated with the delay until the receipt of the reward, such that farther delays were associated with lower vividness scores. Firstly, in an intertemporal bidding task, participants ($n = 39$) were presented with a fixed monetary outcome of $75 at different delays and asked to report the immediate amount they would feel to be equivalent to the delayed outcome (21). For each trial, we calculated neural vividness scores by applying the above-developed whole-brain predictor to the activity for that trial. We found that the trial-by-trial neural vividness scores were correlated negatively with delay [Fig. 4; mean $r = -0.055$, $t(38) = -3.08$, $P = 0.0038$], such that shorter delays (i.e., more proximal future) were associated with higher vividness

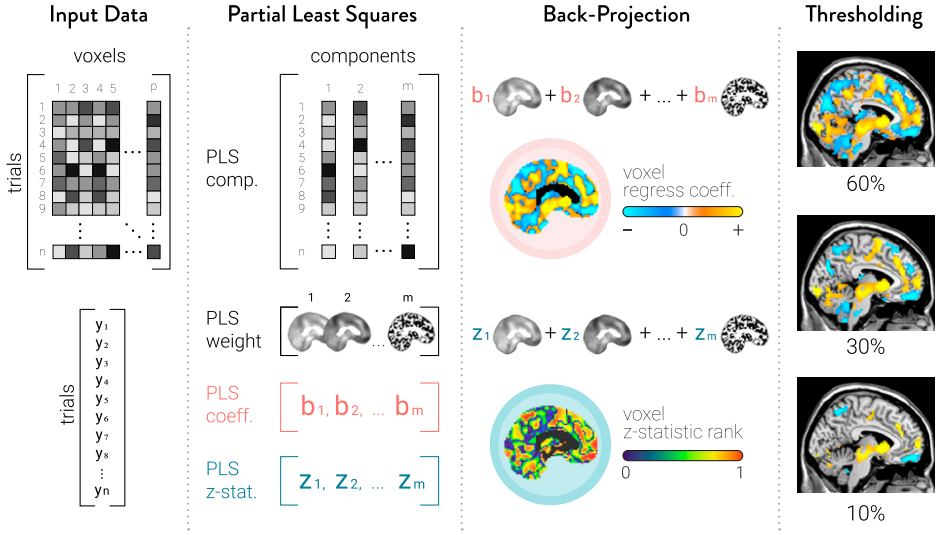

Fig. 1. T-PLS approach to building a whole-brain predictor, adapted with permission from ref. 24. From *Left* to *Right*, the T-PLS method is outlined. The first step performs partial least squares on the brain image data (X) and the dependent variable (Y) in order to extract components that maximally explain the covariance between X and Y. Each of these components are paired with weight maps that describe how each component is a weighted sum of the original voxels. They are also associated with regression coefficients and *t* statistics (approx. *z*-stat) from regressing the dependent variable onto the components. These regression coefficients and *z*-stats are multiplied with their respective weight maps to yield regression coefficients and *z*-stats in the original voxel space. Using the voxel-level *z*-stats, the whole-brain predictor is thresholded by removing less important voxels (i.e., voxels with smaller absolute *z*-stats).

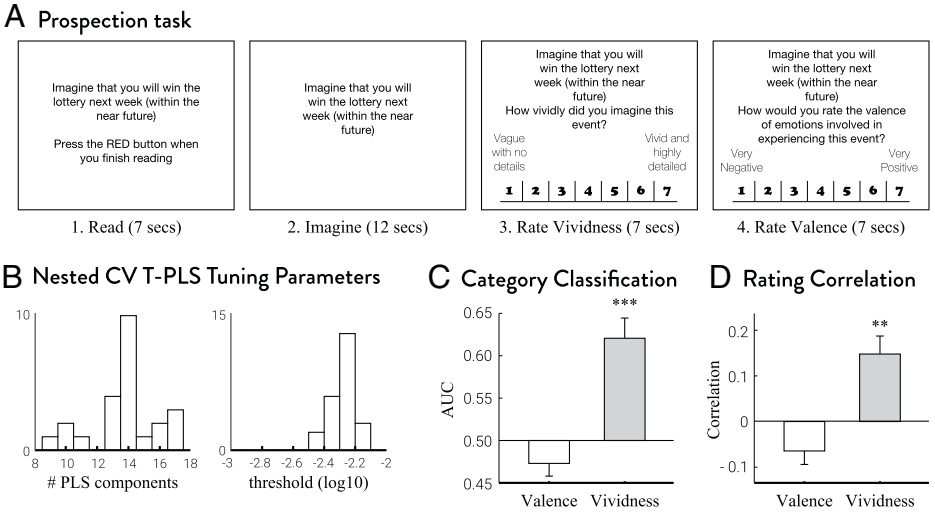

Fig. 2. Out-of-sample prediction of vividness and valence in the prospection dataset. (*A*) Schematic of the prospection task from ref. 22. (*B*) Distribution of T-PLS tuning parameters across the 24-fold cross-validation. (*C*) Classification performance on a priori trial categories of vividness (high versus low) and valence (positive versus negative) as measured by area under the receiver operating characteristic curve. (*D*) Correlation with vividness and valence ratings provided by participants. Error bars represent the SEM. Bonferroni-corrected $^{**}P < 0.01$, $^{***}P < 0.001$.

scores and longer delays (i.e., more distant future) with lower vividness scores.

We replicated this finding in a second dataset in which participants made discrete binary choices between immediate and delayed rewards. In this choice task, participants ($n = 166$) made choices between a fixed immediate reward of \$20 and a future reward that varied in amount (\$21 ∼ \$85) and delay (20 ∼180 d) across trials (25). Again, we found that the trial-by-trial neural vividness scores were correlated negatively with delay [Fig. 5; mean $r = -0.050$, $t(165) = -6.09$, Bonferroni $P < 0.0001$], such that shorter delays were associated with higher vividness scores. Furthermore, this association was specific to the delay to reward. The neural vividness scores were not significantly correlated with the delayed amount [mean $r = 0.016$, $t(165) = 2.12$, Bonferroni $P = 0.11$], and vividness was more strongly associated with delay than amount [paired t test, $t(165) = 3.10$, Bonferroni $P = 0.007$].

Lastly, we examined whether the correlation between neural vividness scores and delay varies systematically across time-on-task or people. Across time, we tested for habituation effects by comparing the mean correlation in the first two scanner runs against the mean correlation in the last two scanner runs. On average, the delay–vividness correlation in the earlier half of the experiment was higher than that in the later half of the experiment for both tasks (Fig. 6; bidding task: paired sign-rank $z = 2.22$, $P = 0.026$, Cohen's $d = 0.16$; choice task: paired sign-rank $z = 2.13$, $P = 0.033$, Cohen's $d = 0.25$). Across people, we tested whether the delay–vividness correlation was stronger for those with higher discount rates (Fig. 7). We found evidence for this relationship in the choice task across 166 people ($r = 0.2$, $P = 0.011$), but not in the bidding task across 39 people ($r = 0$, $P = 1$).

In parallel to our main analyses and results above, we also created a whole-brain predictor of valence that is orthogonal to vividness. While neural vividness scores were specifically associated with delay during intertemporal decision making, neural valence scores tracked the value of delayed rewards. In the bidding task, neural valence scores correlated with both the negative delay and the participants' bids; in the choice task, neural valence scores were correlated with both the negative delay and amount. We also tested whether the delay–valence correlation was stronger for those with higher discount rates; this was again true only for the choice task but not the bidding task (*SI Appendix*, section II).

Discussion

Multiple theories in psychology have suggested that delayed outcomes are discounted in value relative to immediate outcomes in part because more temporally distant options are perceived as less vivid and tangible than more temporally proximal ones (9–11, 14, 16). These theories have been supported by a range of various behavioral experiments (11, 12, 15, 26–29). Here we add converging neuroscientific evidence for these theories.

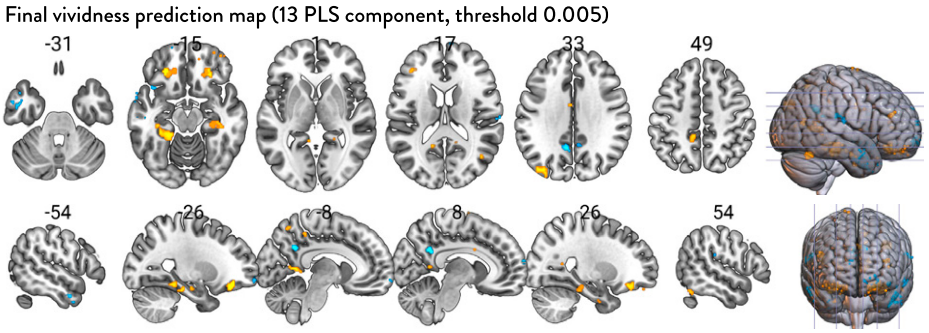

Fig. 3. Whole-brain predictor of the vividness of imagined future events. The warm colors indicate positive coefficients and cool colors indicate negative coefficients.

Table 1. Clusters in the whole-brain predictor of imagination vividness

Description	Size (number of voxels)	X	Y	Z
Positive				
Bilateral hippocampus	133 voxels (left)	−33	−34	−18
	87 voxels (right)	36	−32	−20
Bilateral central orbitofrontal cortex	80 voxels (left)	−23	35	−14
	32 voxels (right)	26	31	−14
Left middle occipital gyrus	65 voxels	−37	−80	35
Left calcarine	32 voxels	−5	−52	5
Right fusiform gyrus	17 voxels	52	−60	−24
Left precuneus/cingulate	12 voxels	−9	−42	49
Negative				
Bilateral precuneus/PCC	26 voxels (left)	−9	−54	31
	31 voxels (right)	8	−50	27
Left inferior temporal gyrus	24 voxels (cluster 1)	−39	11	−42
	24 voxels (cluster 2)	−47	−4	−30
Right postcentral gyrus	22 voxels	64	−20	15

Reported are clusters of voxels that have nonzero coefficients in the final predictor of the vividness of imagined future events, grouped by the sign of the coefficients and ordered by cluster size. From left to right, the region names, cluster size in voxels, and peak MNI coordinates are provided. Clusters that are 10 voxels or smaller have been omitted.

We used fMRI data during an imagination task to create a whole-brain, multivariate predictor specific to the vividness of prospective thought, independent of the valence of prospective thought. Then we show, in two separate delay-discounting datasets with markedly different task structure (one bidding task, one choice task), that the neural signature of vividness is modulated by the temporal distance to the delayed option under consideration. That is, the pattern of neural activity that predicts more vivid prospective thinking is stronger for more temporally proximal outcomes and weaker for more temporally distal ones. The neural signature of vividness was also more strongly modulated by the delay to reward than by the magnitude of reward. Furthermore, we find that this modulation of the neural signature of vividness by delay is stronger in earlier phases of the experiment, when participants may be more actively simulating delayed options, than in later phases, when participants may be more simply following response rules or habits. We also find some evidence that this modulation is stronger in higher discounters, as one might expect, given that delay has a stronger effect on behavior in higher discounters, though this relationship was only observed in the larger of the two datasets we examined. Taken together, these results show that, while people are making intertemporal decisions, an online, unobtrusive neural index of vivid thinking declines as the outcomes considered are delayed farther into the future.

Though the correlation between neural vividness scores and delay was significant, it was also small, at or below $r = 0.1$. We believe that the largest factor influencing the size of these correlations is the inherent noise in single-trial activity estimates in fMRI. Previous studies have simulated multivariate prediction analyses with fMRI, assuming typical signal-to-noise ratios, and found correlations between 0.1 and 0.3 depending on the signal-to-noise assumptions (30). In light of this, it is all the more remarkable that we observe significant correlations between neural prediction scores for vividness and a different variable (delay) in a completely different task (delay discounting). We have also shown that the size of this correlation is reduced by a habituation effect in both datasets and influenced by behaviorally relevant individual differences in the choice dataset. We also expect that the ability to decode mental activity from brain activity varies depending on the mental construct, the fidelity with which it can be elicited and measured, and aspects of brain organization (e.g., prediction is likely easier for motor activity, which is mapped on the cortical surface). Given these factors, we do not think that the correlations we observed were unexpectedly low, though the effect sizes we report should be interpreted in light of the various factors outlined above.

Our results complement previous tests of construal level theory using fMRI. These studies have shown that neural activity

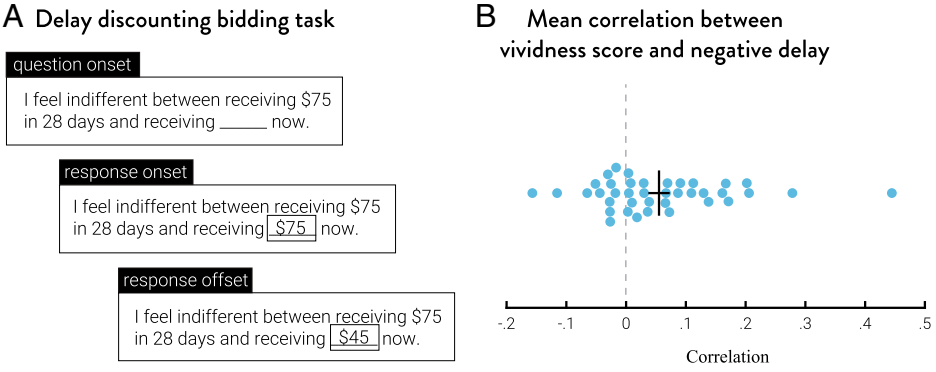

Fig. 4. Out-of-sample prediction of delay in an intertemporal bidding task. (*A*) Bidding task structure from intertemporal bidding dataset (21). Participants are first shown the delayed amount of $75 (fixed) and a variable delay and are asked to bid their immediate equivalent. (*B*) Per-subject correlation between trial-by-trial delay (sign flipped) and vividness prediction scores from the whole-brain predictor. The vertical bar represents the mean, and the horizontal bar represents the SEM ($n = 39$).

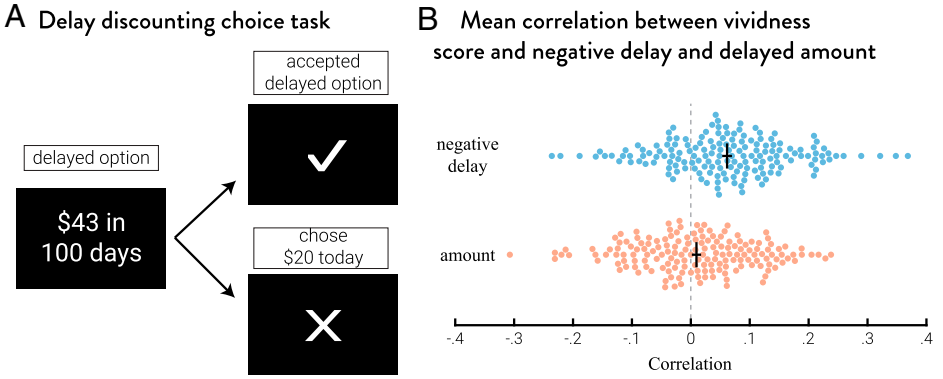

Fig. 5. Out-of-sample prediction of delay in an intertemporal choice task. (*A*) Choice task structure from intertemporal choice dataset (25). Participants are shown the delayed reward and are asked to either accept it or to reject it for $20 immediately. (*B*) Per-subject correlation between trial-by-trial delay (sign flipped) and vividness prediction scores in comparison to that between trial-by-trial amount and vividness prediction scores (delay has been sign flipped to facilitate this comparison). The vertical bar represents the mean, and the horizontal bar represents the SEM (n = 166).

associated with imagining near events, compared to distant events, overlaps with neural activity engaged by other forms of psychological proximity or by low- versus high-level construal (18, 31). Here we make two advances over these previous results. First, we distinguish between neural activity due to the vividness, versus the valence, of prospective thought. This is critical as several previous studies have found the strongest increases in activity for sooner, compared to later, events in the mPFC and PCC (18, 19), two regions that we have previously shown are associated with the valence of prospective thought (22). Second, we show that a neural index of vividness is modulated by the delay to the outcome during intertemporal decision making. This links reduced vividness directly to the discounting of future rewards, a process known to be associated with many important life outcomes (1–7). Finally, note that although our neural signature was constructed to decode the vividness rather concreteness of prospective thought (with the latter being the relevant construct in construal level theory (10, 11), vividness and concreteness ratings of the stimuli we used were highly correlated ($r \geq 0.80$) and therefore neural vividness scores were also correlated with concreteness (*SI Appendix*, section III).

The whole-brain prediction map for vividness is remarkably consistent with findings from other lines of research. Several previous studies have argued that the orbitofrontal cortex represents the features of potential outcomes during decision making (32–34), and that interactions with the hippocampus may be critical for generating these representations from memory (for review, see ref. 35). Furthermore, there is evidence that these regions play a role in valuing delayed rewards. Lesions to the OFC caused increased impatience (36), and reduced gray matter thickness in both the OFC (37) and the medial temporal lobe (38, 39) is associated with increased discounting. The hippocampus is more active when choosing delayed rewards under conditions with greater simulation demands (40), and the reduction in delay discounting by episodic future thinking is associated with increased connectivity between hippocampus and frontal brain regions (14, 16) and requires an intact medial temporal lobe (41). Correspondingly, we would expect that manipulating activity in these regions as people consider future outcomes, for example through noninvasive neurostimulation techniques, would alter the vividness with which those outcomes are imagined and the degree to which those outcomes are discounted.

To obtain the current results, we applied a modified partial least squares algorithm optimized to construct interpretable whole-brain predictors with minimal computation time (24). Though many different methods for constructing whole-brain predictors have been proposed (42–45), none have yet achieved widespread use in the field, presumably due to hurdles posed by their heavy computational demand and/or difficulty with interpretability. The T-PLS method we used here joins other methods

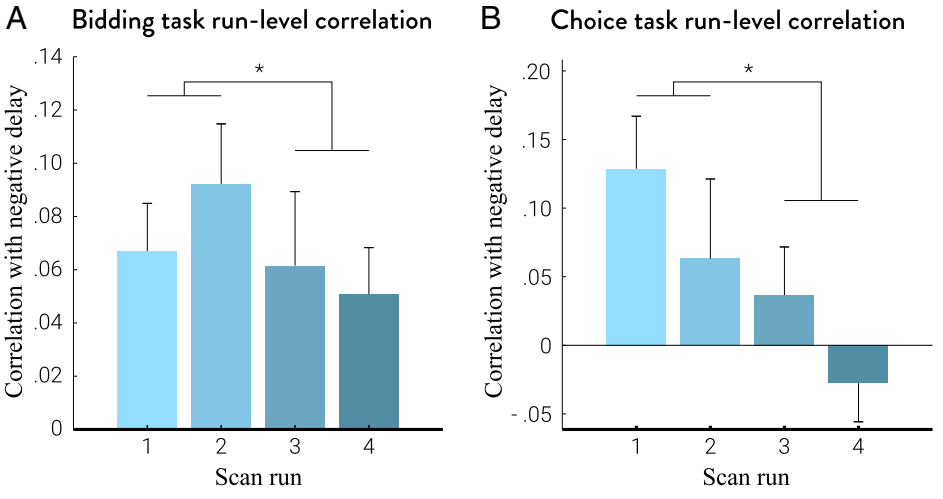

Fig. 6. Changes in vividness-delay correlation over time. Average correlation between negative delay and neural vividness score for each run, for the bidding task (*A*) and for the choice task (*B*). The mean of the first two runs' correlation is significantly higher than the mean of the last two runs' correlation. Error bars represent the SEM (Bonferroni-corrected *P < 0.05).

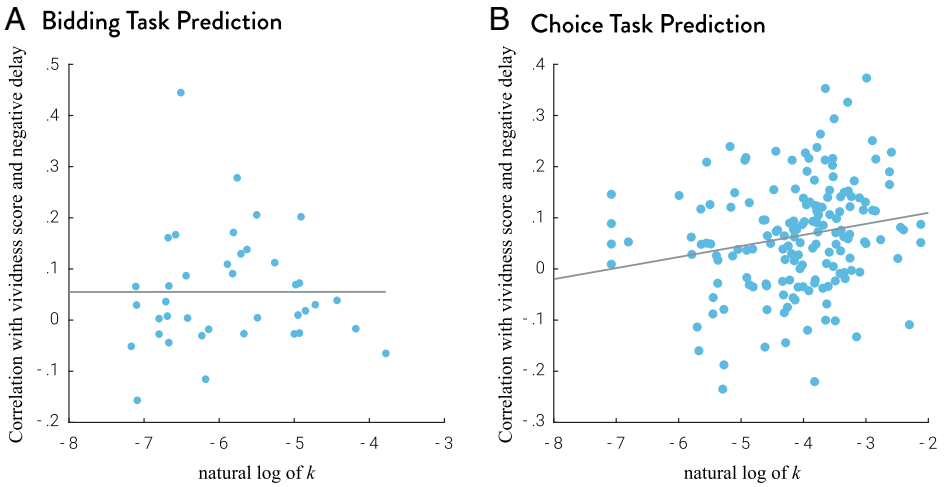

Fig. 7. Correlation between individual discount rate and vividness-delay correlation. Scatterplot between individual logk (abscissa) and individual vividness-delay correlation (ordinate) for the bidding task (*A*) and for the choice task (*B*).

in providing interpretable whole-brain predictors (42, 45), while significantly reducing the computational burden by using analytical computations. It is important to note that the end model of T-PLS, like many other whole-brain methods, is a linear model, but with regularization and variable selection. There are more sophisticated nonlinear models that have been used in smaller regions of the brain to decode more fine-grained patterns (46). In this paper, we illustrate what we think is the most promising and exciting potential use of a whole-brain predictor: decoding mental states online in order to test psychological hypotheses.

Materials and Methods

Prospection Dataset. We used a dataset from Lee et al. (22) to develop a whole-brain predictor of the vividness of imagined future events. The dataset was deidentified prior to use in our study and is available online at openneuro (doi:10.18112/openneuro.ds002835.v1.0.1). This study examined neural activity associated with the valence (positive versus negative) and vividness (high versus low) of imagined future events. Twenty-four participants underwent fMRI scanning while imagining 32 different future scenarios. In a 2 × 2 design (positive versus negative valence crossed with high versus low vividness), 8 different unique scenarios were selected for each condition based on pilot testing. Each scenario was repeated twice during the experiment. Participants completed four runs and imagined 16 scenarios per run. Each trial involved up to 5 s of participants reading the scenario cue, 12 s of imagination, and up to 14 s in which participants rated the vividness and valence of the imagined event on a 7-point Likert scale (7 s each). The trial duration was buffered such that the time the participants did not use in the cue and rating phases was appended to the inter-trial interval at the end of the trial to make the total duration of a single trial 34 s. The 32 scenarios used in this task are provided in *SI Appendix*, section IV.

Delay Discounting Datasets. We applied the neural predictor of vividness developed in the prospection dataset in two different delay-discounting datasets to test whether the neural signature of vividness is modulated by delay during intertemporal decisions. We use one bidding dataset and one choice dataset to evaluate the robustness of the results to different task structures. The first dataset we used was from Cooper et al. (21), which involved bidding on delayed rewards (deidentified dataset available online at openneuro; doi:10.18112/openneuro.ds002989.v1.0.0). A total of 40 participants were asked in each trial to indicate an immediate monetary amount that they would feel was equivalent to receiving $75 after a given delay, varying from 14 to 364 d. Each trial began with a screen of the form "I feel indifferent between receiving $75 in 28 days and receiving _____ now." After the prompt was shown for 3 to 5 s, participants were then allowed a maximum of 10 s to use a button pad to indicate their immediate equivalent amount within a range of $0 to $75. Each participant went through four scan runs, each of which involved 26 questions at different delays, ranging from 14 to 364 d. We removed one participant who bid $75 for all trials regardless of delay, as we were not sure whether the participant understood the task. An advantage of this dataset is that it presents participants with the exact same reward amount at varying delays, thereby allowing us to test whether the neural signature of imagination vividness is modulated by the delay. The flipside of this advantage is that only the delay, and not the amount, of the delayed reward is varied across trials. This limitation is addressed in the second dataset below.

The second dataset we used was from Kable et al. (25), which investigated the effects of cognitive training on neural activity during economic decision making (deidentified dataset available online at openneuro; doi:10.18112/openneuro.ds002843.v1.0.1). Here we use the data from the intertemporal choice task in the first, baseline, scanning session. A total of 166 participants completed four runs of the intertemporal choice task while being scanned. Each run consisted of 30 binary choices between a smaller immediate reward of $20 today that was held constant throughout the entire session and a larger delayed reward (e.g., $30 in 7 d) that varied in amount and delay from trial to trial. On each trial, the delayed option was shown on the screen; the immediate option was not displayed. Participants pressed the left/right buttons on a button pad to indicate whether they would like to accept the delayed option shown on the screen and forego the immediate reward of $20 or to reject the delayed option and take the immediate reward of $20. Participants had up to 4 s to respond, and after their response, a checkmark was shown on the screen if they accepted the delayed reward and an X was shown on the screen if they rejected it.

Image Acquisition. For all datasets, the images were collected with a Siemens 3T Trio scanner with a 32-channel head coil. High-resolution T1-weighted anatomical images were acquired using an MPRAGE sequence (T1 = 1,100 ms; 160 axial slices, 0.9375 × 0.9375 × 1.000 mm; 192 × 256 matrix). The prospection dataset's echo-planar imaging (EPI) sequence involved 44 axial slices with 181 volumes, the intertemporal bidding dataset's EPI sequence involved 45 axial slices with 150 to 152 volumes, and the intertemporal choice dataset's EPI sequence involved 53 axial slices with 104 volumes. The prospection dataset and the intertemporal choice dataset included B0 fieldmap images for distortion correction (Repetition Time (TR) = 1,000 ms, Echo Time (TE) = 2.69, and 5.27 ms for prospection dataset and TR = 1,270 ms, TE = 5, and 7.46 ms for the intertemporal choice dataset).

Image Preprocessing. All datasets were preprocessed via fMRIPrep (47). All BOLD runs were motion corrected, slice-time corrected, b0-map unwarped, registered, and resampled to a Montreal Neurological Institute (MNI) 2-mm template. fMRIPrep does not perform smoothing, so it was manually performed after estimating single trial activities (see below).

BOLD Deconvolution. We used beta-series regression (48) to estimate the BOLD activity associated with each trial in each of the three datasets. In the prospection dataset, we estimated the BOLD activity during the imagination period of 12 s. The regressors were time locked to imagination time onset with an event

duration of 12 s and convolved with a double gamma hemodynamic response function (HRF). In the two delay-discounting datasets, the regressors were time locked to the onset of the response period (when participants can input their response) with event duration of 0.1 s and convolved with a double gamma HRF function. In the intertemporal choice dataset only, the last trial of each run was excluded from analysis because the BOLD activity of the last trial was often not observed due to the termination of the scan. After the single-trial coefficients were estimated, all images were smoothed with a FWHM 8-mm Gaussian filter, which is the standard smoothing kernel for SPM software. Finally, we used the Automated Anatomical Labeling (AAL) atlas as a mask to retain only gray matter voxels.

Prediction Algorithm. To create a whole-brain predictor of the vividness of imagined future events, we used T-PLS (24). T-PLS is similar in approach to other methods for constructing whole-brain predictors that use principal components analysis (PCA) to reduce the dimensionality of the data followed by regression (42, 49, 50). The key advantage of T-PLS over PCA-based methods is that partial least squares (PLS) components maximally explain the covariance between the predictors and the outcome, whereas PCA components only explain the variance of the predictors. Thus, PLS yields data reduction that is more pertinent to prediction.

Building a whole-brain predictor of vividness is done in three steps (Fig. 1). First, we performed PLS to extract components that maximally explain the covariance between the single-trial images and the binary vividness trial categories (high versus low). These components consist of a map of weights for each voxel in the brain. PLS also automatically yields coefficients for each component that are equivalent to the regression coefficients one would obtain from regressing the dependent variable on the components. We also calculated the *t* statistics of each component as one would get from a regression model (here, given the large number of observations, we assume that the *t* statistics are approximations of *z* statistics). In the second step, we back project the PLS coefficients and *z* statistics into the original voxel space by multiplying them with the PLS weight maps. This yields coefficients for each of the brain voxels for easier interpretation. In the final step, we used the back-projected *z* statistics of each voxel to rank their variable importance and threshold the voxel coefficient map so that less important voxels are removed from the final predictor. This final predictor can be used to obtain a "vividness score" for each brain image by calculating the dot product between the predictor and the image.

The two tuning parameters–number of PLS components and the level of thresholding–are chosen based on cross-validation performance. In our paper, we use the "one-standard-error rule" to choose the most parsimonious model (least number of voxels) whose cross-validation performance is comparable with the best model (i.e., the tuning parameter combination that yields the highest cross-validation performance). This criterion is used ubiquitously in modern regressions and is often the default choice of tuning parameters in statistical packages (51–57).

Sensitivity and Specificity Analysis. To assess whether we can create a whole-brain predictor of vividness that is orthogonal to valence, we performed a nested 24-fold leave-one-person-out cross-validation within the prospection dataset. We trained the predictor on data from 23 participants and tested on the one left-out person. Within the 23 training participants' data, we employed an additional 23-fold leave-one-person-out cross-validation to find the optimal number of components and thresholding level. After the best parameters were found, the T-PLS model was fitted using all 23 participants and used to predict the left-out person's data. Specifically, we tested whether the T-PLS model can accurately classify the high versus low vividness trial categories but not positive versus negative trial categories and whether the classifier output correlates with vividness ratings but not with valence ratings. As a parallel to this analysis, we also repeated the same process to assess whether we can create a whole-brain predictor of valence that is orthogonal to vividness (*SI Appendix*, section II and Fig. S.1).

Constructing the Predictor of Vividness. Once we established that T-PLS can build a whole-brain predictor of vividness that is orthogonal to valence, we used the entire prospection dataset to build a whole-brain predictor of vividness to be used for prediction in other datasets. We used 24-fold leave-one-out cross-validation within the prospection dataset (but not nested cross-validation as in the sensitivity and specificity analysis) to choose one number of components and one thresholding value for the final predictor. Again, as a parallel to this analysis, we also repeated the same process to build a whole-brain predictor of valence to be used for prediction in other datasets (*SI Appendix*, section II and Fig. S.2).

To calculate an expression score for the neural signature of vividness during delay discounting, we calculated the dot product between the neural predictor of vividness and the brain image of estimated activity for each trial. These scores were then correlated with the delay until the receipt of the delayed reward (in days) in both datasets, and the delayed amount in the intertemporal choice dataset (since the delayed amount is constant in the intertemporal bidding dataset). The correlations were performed at the individual level, and each individual's correlation coefficient was used as a summary statistic to test whether there was a significant correlation at the group level. The parallel analysis using the neural signature of valence is provided in *SI Appendix*, section II and Fig. S.3.

Data, Materials, and Software Availability. Three previously published datasets were used for this work. The prospection dataset from Lee et al. (22, 58), delay discounting bidding dataset from Cooper et al. (21, 59), and delay discounting choice dataset from Kable et al. (25, 60). All three datasets are deidentified and available online at openneuro at their respective URLs.

ACKNOWLEDGMENTS. Portions of this work were developed from S.L.'s doctoral thesis. This study was supported by National Cancer Institute grants R01-CA-170297 (to C.L. and J.W.K.) and R35-CA-197461 (to C.L.), National Institute on Drug Abuse grant R01-DA-029149 (to J.W.K.), the National Institute on Aging grant P30 AG-012836-18, the Boettner Center for Pensions and Retirement Security, and NIH National Institute of Child Health and Development Population Research Infrastructure Program R24 HD-044964-9, all at the University of Pennsylvania.

1. K. N. Kirby, N. M. Petry, W. K. Bickel, Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *J. Exp. Psychol. Gen.* **128**, 78–87 (1999).
2. N. A. Shamosh, J. R. Gray, Delay discounting and intelligence: A meta-analysis. *Intelligence* **36**, 289–305 (2008).
3. S. M. Alessi, N. M. Petry, Pathological gambling severity is associated with impulsivity in a delay discounting procedure. *Behav. Processes* **64**, 345–354 (2003).
4. T. S. Schepis, A. McFetridge, T. M. Chaplin, R. Sinha, S. Krishnan-Sarin, A pilot examination of stress-related changes in impulsivity and risk taking as related to smoking status and cessation outcome in adolescents. *Nicotine Tob. Res.* **13**, 611–615 (2011).
5. O. Urminsky, G. Zauberman, "The psychology of intertemporal preferences" in *The Wiley Blackwell Handbook of Judgment and Decision Making*. G. Keren, G. Wu. Eds. (John Wiley & Sons, Ltd, 2015). pp. 141–181.
6. K. N. Kirby, G. C. Winston, M. Santiesteban, Impatience and grades: Delay-discount rates correlate negatively with college GPA. *Learn. Individ. Differ.* **15**, 213–222 (2005).
7. S. A. Atlas, E. J. Johnson, J. W. Payne, Time preferences and mortgage choice. *J. Mark. Res.* **54**, 415–429 (2017).
8. D. Soman *et al.*, The psychology of intertemporal discounting: Why are distant events valued differently from proximal ones? *Mark. Lett.* **16**, 347–360 (2005).
9. S. Rick, G. Loewenstein, Intangibility in intertemporal choice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **363**, 3813–3824 (2008).
10. Y. Trope, N. Liberman, Construal-level theory of psychological distance. *Psychol. Rev.* **117**, 440–463 (2010).
11. N. Liberman, Y. Trope, Traversing psychological distance. *Trends Cogn. Sci.* **18**, 364–369 (2014).
12. S. A. Malkoc, G. Zauberman, J. R. Bettman, Unstuck from the concrete: Carryover effects of abstract mindsets in intertemporal preferences. *Organ. Behav. Hum. Decis. Process.* **113**, 112–126 (2010).
13. S. A. Malkoc, G. Zauberman, Deferring versus expediting consumption: The effect of outcome concreteness on sensitivity to time horizon. *J. Mark. Res.* **43**, 618–627 (2006).
14. J. Peters, C. Büchel, Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-mediotemporal interactions. *Neuron* **66**, 138–148 (2010).
15. S. A. Rösch, D. F. Stramaccia, R. G. Benoit, Promoting farsighted decisions via episodic future thinking: A meta-analysis. *J. Exp. Psychol. Gen.* **151**, 1606–1635 (2022).
16. R. G. Benoit, S. J. Gilbert, P. W. Burgess, A neural mechanism mediating the impact of episodic prospection on farsighted decisions. *J. Neurosci.* **31**, 6771–6779 (2011).
17. R. M. K. Carter, J. R. Meyer, S. A. Huettel, Functional neuroimaging of intertemporal choice models: A review. *J. Neurosci. Psychol. Econ.* **3**, 27 (2010).
18. D. I. Tamir, J. P. Mitchell, The default network distinguishes construals of proximal versus distal events. *J. Cogn. Neurosci.* **23**, 2945–2955 (2011).
19. J. P. Mitchell, J. Schirmer, D. L. Ames, D. T. Gilbert, Medial prefrontal cortex predicts intertemporal choice. *J. Cogn. Neurosci.* **23**, 857–866 (2011).
20. O. Bartra, J. T. McGuire, J. W. Kable, The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* **76**, 412–427 (2013).
21. N. Cooper, J. W. Kable, B. K. Kim, G. Zauberman, Brain activity in valuation regions while thinking about the future predicts individual discount rates. *J. Neurosci.* **33**, 13150–13156 (2013).

22. S. Lee, T. Parthasarathi, J. W. Kable, The ventral and dorsal default mode networks are dissociably modulated by the vividness and valence of imagined events. *J. Neurosci.* **41**, 5243–5250 (2021).
23. R. A. Poldrack, Inferring mental states from neuroimaging data: From reverse inference to large-scale decoding. *Neuron* **72**, 692–697 (2011).
24. S. Lee, E. T. Bradlow, J. W. Kable, Fast construction of interpretable whole-brain decoders. *Cell Rep. Methods* **2**, 100227 (2022).
25. J. W. Kable *et al.*, No effect of commercial cognitive training on brain activity, choice behavior, or cognitive performance. *J. Neurosci.* **37**, 7390–7402 (2017).
26. R. Yi, A. Stuppy-Sullivan, A. Pickover, R. D. Landes, Impact of construal level manipulations on delay discounting. *PLoS One* **12**, e0177240 (2017).
27. C. Bischoff, J. Hansen, Influencing support of charitable objectives in the near and distant future: Delay discounting and the moderating influence of construal level. *Soc. Influ.* **11**, 217–229 (2016).
28. W. Mischel, N. Baker, Cognitive appraisals and transformations in delay behavior. *J. Pers. Soc. Psychol.* **31**, 254 (1975).
29. N. J. Kelley, B. J. Schmeichel, Thinking about death reduces delay discounting. *PLoS One* **10**, e0144228 (2015).
30. J. A. Mumford, B. O. Turner, F. G. Ashby, R. A. Poldrack, Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage* **59**, 2636–2643 (2012).
31. P. E. Stillman *et al.*, Neurological evidence for the role of construal level in future-directed thought. *Soc. Cogn. Affect. Neurosci.* **12**, 937–947 (2017).
32. K. A. Burke, T. M. Franz, D. N. Miller, G. Schoenbaum, The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. *Nature* **454**, 340–344 (2008).
33. J. D. Howard, J. A. Gottfried, P. N. Tobler, T. Kahnt, Identity-specific coding of future rewards in the human orbitofrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 5195–5200 (2015).
34. Y. K. Takahashi *et al.*, Neural estimates of imagined outcomes in the orbitofrontal cortex drive behavior and learning. *Neuron* **80**, 507–518 (2013).
35. D. Shohamy, N. D. Daw, Integrating memories to guide decisions. *Curr. Opin. Behav. Sci.* **5**, 85–90 (2015).
36. M. Sellitto, E. Ciaramelli, G. di Pellegrino, Myopic discounting of future rewards after medial orbitofrontal damage in humans. *J. Neurosci.* **30**, 16429–16436 (2010).
37. M. Pehlivanova *et al.*, Diminished cortical thickness is associated with impulsive choice in adolescence. *J. Neurosci.* **38**, 2471–2481 (2018).
38. K. M. Lempert *et al.*, Neural and behavioral correlates of episodic memory are associated with temporal discounting in older adults. *Neuropsychologia* **146**, 107549 (2020).
39. M. M. Owens *et al.*, Neuroanatomical foundations of delayed reward discounting decision making. *Neuroimage* **161**, 261–270 (2017).
40. M. Lebreton *et al.*, A critical role for the hippocampus in the valuation of imagined outcomes. *PLoS Biol.* **11**, e1001684 (2013).
41. D. J. Palombo, M. M. Keane, M. Verfaellie, The medial temporal lobes are critical for reward-based decision making under conditions that promote episodic future thinking. *Hippocampus* **25**, 345–353 (2015).
42. T. D. Wager *et al.*, An fMRI-based neurologic signature of physical pain. *N. Engl. J. Med.* **368**, 1388–1397 (2013).
43. A. Smith, B. D. Bernheim, C. Camerer, A. Rangel, Neural activity reveals preferences without choices. *Am. Econ. J. Microecon.* **6**, 1–36 (2014).
44. P. A. Kragel, K. S. LaBar, Multivariate neural biomarkers of emotional states are categorically distinct. *Soc. Cogn. Affect. Neurosci.*, **10**, 1437–1448 (2015).
45. L. Grosenick, B. Klingenberg, K. Katovich, B. Knutson, J. E. Taylor, Interpretable whole-brain prediction analysis with GraphNet. *Neuroimage* **72**, 304–321 (2013).
46. K. N. Kay, T. Naselaris, R. J. Prenger, J. L. Gallant, Identifying natural images from human brain activity. *Nature* **452**, 352–355 (2008).
47. O. Esteban *et al.*, fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nat. Methods* **16**, 111–116 (2019).
48. J. Rissman, A. Gazzaley, M. D'Esposito, Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage* **23**, 752–763 (2004).
49. L. J. Chang, P. J. Gianaros, S. B. Manuck, A. Krishnan, T. D. Wager, A sensitive and specific neural signature for picture-induced negative affect. *PLoS Biol.* **13**, e1002180 (2015).
50. T. D. Wager, L. Y. Atlas, L. A. Leotti, J. K. Rilling, Predicting individual differences in placebo analgesia: Contributions of brain activity during anticipation and pain experience. *J. Neurosci.* **31**, 439–452 (2011).
51. L. Breiman, J. Friedman, C. J. Stone, R. A. Olshen, *Classification and Regression Trees* (CRC Press, 1984).
52. D. Krstajic, L. J. Buturovic, D. E. Leahy, S. Thomas, Cross-validation pitfalls when selecting and assessing regression and classification models. *J. Cheminform.* **6**, 10 (2014).
53. J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1–22 (2010).
54. J. Qian, T. Hastie, J. Friedman, R. Tibshirani, N. Simon, Glmnet for Matlab, 2013 (2013). www.stanford.edu/~hastie/glmnet_matlab.
55. T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer Science & Business Media, 2009).
56. T. Hastie, R. Tibshirani, M. Wainwright, *Statistical Learning with Sparsity: The Lasso and Generalizations* (Chapman and Hall/CRC, 2019).
57. R. Tibshirani, G. Walther, T. Hastie, Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. Series B Stat. Methodol.* **63**, 411–423 (2001).
58. S. Lee, T. Parthasarathi, J. W. Kable, Prospection. OpenNeuro. doi:10.18112/openneuro.ds002835.v1.0.1. Accessed 1 July 2020.
59. N. Cooper, J. W. Kable, B. K. Kim, G. Zauberma, Delay discounting bidding. OpenNeuro. doi:10.18112/openneuro.ds002989.v1.0.0. Accessed 1 July 2020.
60. J. W. Kable *et al.*, Delay discounting choice. OpenNeuro. doi:10.18112/openneuro.ds002843.v1.0.1. Accessed 1 July 2020.