Statistical learning of successor representations is related to on-task replay

Lennart Wittkuhn[1,2,*], Lena M. Krippner[1,3] & Nicolas W. Schuck[1,2,*]

[1]Max Planck Research Group NeuroCode, Max Planck Institute for Human Development, Berlin, Germany
[2]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany
Lentzeallee 94, D–14195 Berlin, Germany
[3]Harding Center for Risk Literacy, University of Potsdam, Faculty of Health Sciences, Potsdam, Germany
Virchowstraße 2–4, D–14482 Potsdam, Germany
[*]Correspondence to
wittkuhn@mpib-berlin.mpg.de (ORCiD: 0000-0003-2966-6888)
schuck@mpib-berlin.mpg.de (ORCiD: 0000-0002-0150-8776)

Abstract

Humans automatically infer higher-order relationships between events in the environment from their statistical co-occurrence, often without conscious awareness. Neural replay of task representations, which has been described as sampling from a learned transition structure of the environment, is a candidate mechanism by which the brain could use or even learn such relational information in the service of adaptive behavior. Human participants viewed sequences of images that followed probabilistic transitions determined by ring-like graph structures. Behavioral modeling revealed that participants acquired multi-step transition knowledge through gradual updating of an internal successor representation (SR) model, although half of participants did not indicate any knowledge about the sequential task structure. To investigate neural replay, we analyzed dynamics of multivariate functional magnetic resonance imaging (fMRI) patterns during short pauses from the ongoing statistical learning task. Evidence for sequential replay consistent with the probabilistic task structure was found in occipito-temporal and sensorimotor cortices during short on-task intervals. These findings indicate that implicit learning of higher-order relationships establishes an internal SR-based map of the task, and is accompanied by cortical on-task replay.

Introduction

The representation of structural knowledge in the brain in form of a so-called *cognitive map* has been a topic of great interest. A common assumption is that a cognitive map provides the basis for flexible learning, inference, and generalization (Tolman, 1948; Wilson et al., 2014; Schuck et al., 2016; Behrens et al., 2018), and yet is based on individual experiences that provide structural information only indirectly (Schapiro et al., 2013; Garvert et al., 2017). The brain must therefore extract statistical regularities from continuous experiences, and then use these regularities as the starting point for the formation of abstract, map-like knowledge. A mechanism through which abstract knowledge could be used to generate flexible behavior is on-task replay (e.g., Sutton, 1991; Kurth-Nelson et al., 2016), the rapid reactivation of trajectories simulated from an internal cognitive map. In this paper, we investigated whether on-task replay of cognitive map-like knowledge occurs in the human brain while participants learn statistical regularities.

The extraction of statistical regularities from experience is known as *statistical learning* (Schapiro and Turk-Browne, 2015; Garvert et al., 2017; Sherman et al., 2020). Statistical learning is automatic and incidental, as it occurs without any instructions or premeditated intention to learn, and often leads to implicit knowledge that is not consciously accessible (Reber, 1989; Seger, 1994; Turk-Browne et al., 2005). This contrasts with research on cognitive maps and planning that often relies on instruction-based task knowledge (e.g., Schuck et al., 2016; Constantinescu et al., 2016; Kurth-Nelson et al., 2016). In a statistical learning setting, relationships between events are typically described by pairwise transition probabilities (i.e., the probability that A is followed by B) to which humans show great sensitivity from an early age on (Saffran et al., 1996). Intriguingly, many experiments have shown that humans extract higher-order relational structures among individual events that go beyond pairwise transition probabilities (for reviews, see e.g., Karuza et al., 2016; Lynn and Bassett, 2020). This includes knowledge about ordinal and hierarchical information that structures individual subsequences (Schuck et al., 2012a,b; Solway et al., 2014; Balaguer et al., 2016), graph topological aspects such as bottlenecks and community structure (Schapiro et al., 2013; Karuza et al., 2017; Kahn et al., 2018), and macro-scale aspects of graph structures (Lynn et al., 2020a,b).

A main benefit of abstracted knowledge in the context of transition structures is that it allows to plan multi-step sequences (Miller and Venditto, 2021; Hunt et al., 2021). Specifically, while experienced transition structure can be used to learn about the probability that a given event will be followed by a specific other event, it can also be used to compute long-term visitation probabilities, i.e., which events can be expected over a given future horizon. This idea is formalized in the successor representation (SR) (Dayan, 1993), a predictive map that reflects the (discounted) expected visitations of future events (Garvert et al., 2017; Bellmund et al., 2020; Brunec and Momennejad, 2021; Russek et al., 2021), and can be learned from the experience of individual transitions. Critically, the predictive horizon of the SR depends on a discount parameter γ which determines how far into the future upcoming states are considered (Momennejad and Howard, 2018; Momennejad, 2020). One goal of our study was therefore to investigate whether statistical learning leads to knowledge of expected future visitations over a predictive horizon, as required for mental planning.

The second main interest of our study was to understand whether abstract knowledge derived from statistical learning would be reflected in on-task replay. Replay is characterized by the fast sequential reactivation of neural representations that reflect previously experienced transition structure (see e.g., Wikenheiser and Redish, 2015a; Schuck and Niv, 2019; Wittkuhn et al., 2021; Yu et al., 2021). Replay occurs in hippocampal but also cortical brain areas (Ji and Wilson, 2006; Wittkuhn and Schuck, 2021)

and has been observed during short pauses from the ongoing task in rodents (Johnson and Redish, 2007; Carr et al., 2011) as well as humans (Kurth-Nelson et al., 2016; Tambini and Davachi, 2019). Sequential reactivation observed during brief pauses is often referred to as *online* or *on-task* replay, and likely reflects planning of upcoming choices (Kurth-Nelson et al., 2016; Eldar et al., 2020).

Previous studies have shown that expectations about upcoming visual stimuli elicit neural signals that are very similar to those during actual perception (Kok et al., 2012, 2014; Hindy et al., 2016; Kok and Turk-Browne, 2018) and anticipatory activation sequences have been found in visual cortex following perceptual sequence learning (Xu et al., 2012; Eagleman and Dragoi, 2012; Gavornik and Bear, 2014; Ekman et al., 2017). It remains unknown, however, whether on-task replay mirrors predictive knowledge that is stored in SR-based cognitive maps. In addition, while most research has focused on hippocampal reactivation, the above evidence suggests that statistical knowledge is also reflected in sensory and motor brain areas.

In the present study, we therefore examined whether on-task neural replay in visual and motor cortex reflects anticipation of sequentially structured stimuli in an automatic and incidental statistical learning context. This may elucidate if (non-hippocampal) neural replay during on-task pauses contributes to learning of probabilistic cognitive maps. To this end, participants performed an incidental statistical learning paradigm (cf. Schapiro et al., 2012; Lynn et al., 2020a) in which visual presentation order and motor responses followed statistical regularities that were determined by a ring-like graph structure. The nature of the graph structure allowed us to dissociate knowledge about individual transition probabilities from an SR-based cognitive map that entails long-term visitation probabilities. Moreover, the transition probabilities among the task stimuli changed halfway through the experiment without prior announcement, which allowed us to understand the dynamical updating of task knowledge and replay within the same participants.

Results

Thirty-nine human participants took part in an fMRI experiment over two sessions. Participants were first informed that the experiment involves six images of animals (cf. Snodgrass and Vanderwart, 1980; Rossion and Pourtois, 2004) and six response buttons mapped onto their index, middle, and ring fingers of both hands. Participants then began the first session of magnetic resonance imaging (MRI), during which they learned the stimulus-response (S-R) mappings between images and response buttons through feedback (*recall trials*, Fig. 1a, 8 runs with 60 trials each, 480 trials in total). In recall trials, animal images were shown without any particular sequential order, i.e., all pairwise sequential orderings of the images were presented equally often per run. Participants had to press the correct button in response to briefly presented images (500 milliseconds (ms)) during a response window (800 ms; jittered stimulus-response interval (SRI) of 2500 ms on average). If the response was incorrect, a feedback about the correct button was provided (500 ms; no feedback on correct trials). The trial ended with a jittered inter-trial interval (ITI) of 2500 ms on average.

The second session started with one additional run of recall trials that was followed by five runs of *graph trials* (Fig. 1b, 240 trials per run, 1200 trials in total). As before, participants had to press the correct button in response to each animal. Images were now presented in a faster pace (800 ms per image and 750 ms between images on average), and only on 10% of trials (120 graph trials in total per participant), ITIs were set to 10 seconds (s). Importantly, the order of the images now followed a probabilistic transition structure (see below), about which participants were not informed,

and no feedback was provided. At the end of the second session, participants completed a post-task questionnaire assessing explicit sequence knowledge.

The sequential ordering of images during graph trials was determined by either a *unidirectional* or *bidirectional* ring-like graph structure with probabilistic transitions (Fig. 2a–b; for details, see Methods). In the unidirectional graph condition (Fig. 2a, middle, henceforth *uni*), each image had one frequent transition to the clockwise neighboring node (probability of $p_{ij} = 0.7$), never transitioned to the counterclockwise neighbor ($p_{ij} = 0.0$), and was followed occasionally by the three other nodes ($p_{ij} = 0.1$ each; Fig. 2b, left). In consequence, stimuli most commonly transitioned in clockwise order along the ring shown in Fig. 2a. In the bidirectional graph condition (Fig. 2a, right, henceforth *bi*), transitions to both neighboring nodes (clockwise and counterclockwise) were equally likely ($p_{ij} = 0.35$), and transitions to all other three nodes occurred with $p_{ij} = 0.1$ (Fig. 2b, right), as in the unidirectional graph. Every participant started the task in one of these conditions (uni or bi). Halfway through the third run, transitions began to be governed by the alternative graph, such that all participants experienced both graphs as well as the change between them (Fig. 2c). 12 participants started in the unidirectional condition and transitioned to the bidirectional graph (uni – bi), while 27 participants experienced the reverse order (bi – uni).

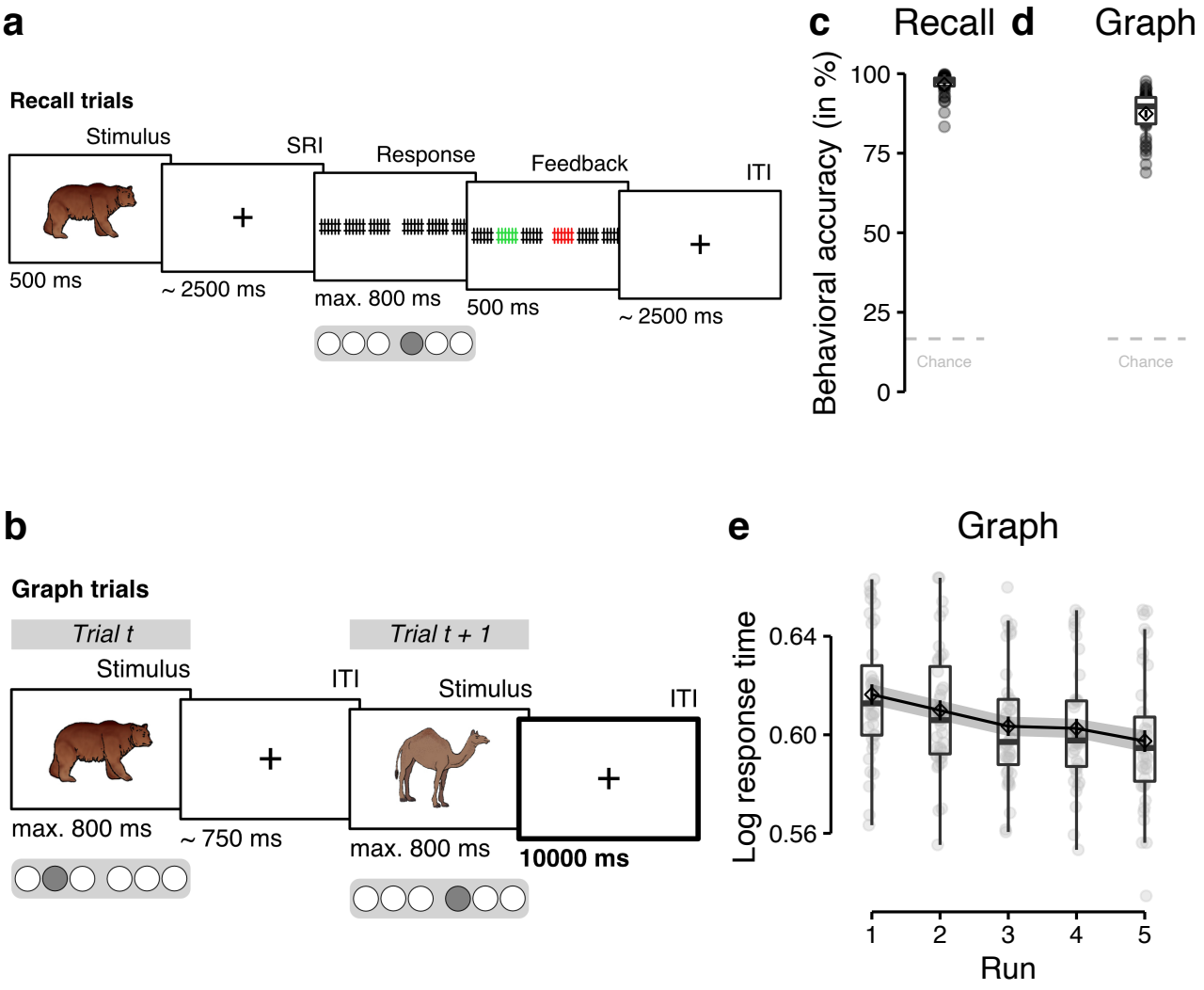


Figure 1: *[see caption on the next page]*

Figure 1: Task design and stimulus-response learning. **(a)** On recall trials, individual images were presented for 500 ms. Participants were instructed to press the correct response button associated with the stimulus during the response interval (time limit of 800 ms). Stimulus presentations and motor responses were separated by SRIs and ITIs which lasted 2.5 s on average (cf. Wittkuhn and Schuck, 2021). Feedback was only presented on incorrect trials. Classifiers were trained on fMRI data from correct recall trials only. **(b)** On graph trials, images were presented for 800 ms, separated by only 750 ms on average. Participants were asked to press the correct response button associated with the presented stimulus as quickly and accurately as possible within 800 ms. On 10% of trials, ITIs lasted 10 s (see ITI in trial $t+1$; highlighted by the thick border, for illustrative purposes only). Classifier trained on fMRI data from correct recall trials were applied to the eight TRs of the 10 s ITIs in graph trials to investigate task-related neural activation patterns during on-task pauses. **(c)** Mean behavioral accuracy (in %; y-axis) across all nine runs of the recall trials. **(d)** Mean behavioral accuracy (in %; y-axis) across all five runs of the graph trials. **(e)** Mean log response time (y-axis) per run (x-axis) in graph trials. Boxplots in (c), (d), and (e) indicate the median and interquartile range (IQR). The lower and upper hinges correspond to the first and third quartiles (the 25^{th} and 75^{th} percentiles). The upper whisker extends from the hinge to the largest value no further than 1.5$*$ IQR from the hinge (where IQR is the interquartile range (IQR), or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5$*$ IQR of the hinge. The diamond shapes show the sample mean. Error bars in (c), (d) and shaded areas in (e) indicate ± 1 standard error of the mean (SEM). Each dot in (c), (d), and (e) corresponds to averaged data from one participant. All statistics have been derived from data of $n = 39$ human participants who participated in one experiment. The stimulus material (individual images of a bear and a dromedary) shown in (a) and (b) were taken from a set of colored and shaded images commissioned by Rossion and Pourtois (2004), which are loosely based on images from the original Snodgrass and Vanderwart set (Snodgrass and Vanderwart, 1980). The images are freely available from the internet at `https://sites.google.com/andrew.cmu.edu/tarrlab/resources/tarrlab-stimuli` under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license (CC BY-NC-SA 3.0; for details, see `https://creativecommons.org/licenses/by-nc-sa/3.0/`). Stimulus images courtesy of Michael J. Tarr, Carnegie Mellon University, (for details, see `http://www.tarrlab.org/`).

Behavioral results

We first asked whether participants learned the stimulus-response (S-R) mapping sufficiently well. Behavioral accuracy on recall trials indeed surpassed chance-level (16.67%) in all runs ($\bar{x} \geq 86.50\%$, CIs $[\geq 80.79, +\infty]$, $t_{38} \geq 20.62$, ps < 0.001 (corrected), ds ≥ 3.30; Figs. 1c, S2b–c). Likewise, during graph trials, participants also performed above chance in all runs ($\bar{x} \geq 85.12$, CIs $[\geq 82.55, +\infty]$, $t_{38} \geq 44.90$, ps < 0.001 (corrected), ds ≥ 7.19; Figs. 1d, S2d), and improved with time (effect of run: $F_{1.00,38.00} = 7.96$, $p = 0.008$, Fig. S2d).

Next, we investigated sequential knowledge. Although participants were not informed that images followed a sequential structure during graph trials, we expected that incidental learning would allow them to anticipate upcoming stimuli during these trials, and thus respond faster with learning. A linear mixed effects (LME) model that tested the effect of task run on response times was broadly in line with this assumption as it showed a significant decrease of response times over the course of learning, $F_{1.00,38.00} = 25.86$, $p < 0.001$ (Figs. 1e, S2e). More directly, we expected that participants would learn the probabilistic transition structure of images and response buttons during graph trials, including the change in transition structure in the middle of the third run. Specifically, we hypothesized that participants would not only learn about one-step transition probabilities, but also form internal maps of the underlying graphs that reflect the higher-order structure of statistical multi-step relationships between stimuli, i.e., how likely a particular stimulus will be experienced in two, three, or more steps from the current time point (cf. Lynn and Bassett, 2020; Lynn et al., 2020a). In our task, this meant that participants might react differently to the three transitions that all have the same one-step transition probability, since they differ in how likely they would occur in multi-step trajectories. For instance, the one-step transition probabilities for A→C, A→D, and A→E were the same in the unidirectional graph, but the two-step probability of A→C was higher than for the other transitions, since the most likely two-step path was A→B→C. This means that participants should react faster

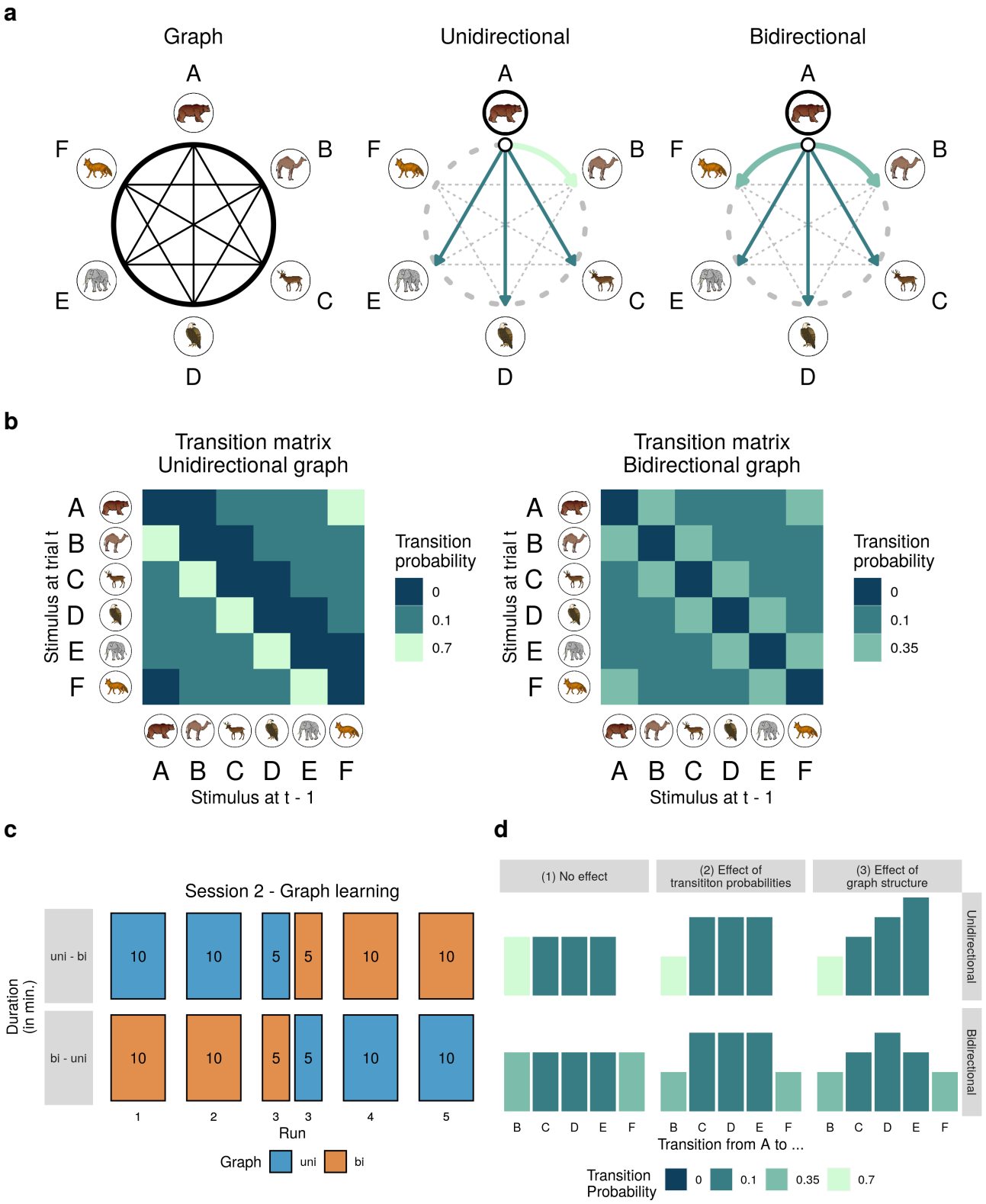


Figure 2: *[see caption on the next page]*

to A→C transitions if they have multi-step knowledge. For simplicity, we will henceforth refer to the A→C transition as having a shorter "node distance", than A→D or A→E (see the rightmost column in Fig. 2d, where colors reflect one-step transition probabilities, and the height of the bars indicate node distance).

Figure 2: Graph learning task. **(a)** The relationships among the six task stimuli depicted as a ring-like graph structure (left). In the unidirectional graph (middle), stimuli frequently transitioned to the clockwise neighboring node ($p_{ij} = p_{AB} = 0.7$), never to the counterclockwise neighboring node ($p_{AF} = 0.0$), and only occasionally to the three other nodes ($p_{AC} = p_{AD} = p_{AE} = 0.1$). In the bidirectional graph (right), stimuli were equally likely to transition to the clockwise or counterclockwise neighboring node ($p_{AB} = p_{AF} = 0.35$) and only occasionally transitioned to the three other nodes ($p_{AC} = p_{AD} = p_{AE} = 0.1$). Transition probabilities are highlighted for node A only, but apply equally to all other nodes. Arrows indicate possible transitions, colors indicate transition probabilities (for a legend, see panel b). **(b)** Transition matrices of the unidirectional (left) and bidirectional (right) graph structures. Each matrix depicts the probability (colors) of transitioning from the stimulus at the previous trial $t - 1$ (x-axis) to the current stimulus at trial t (y-axis). **(c)** Within-participant order of the two graph structures across the five runs of the graph learning task. $n = 12$ participants first experienced the unidirectional, then the bidirectional graph structure (uni – bi; top horizontal panel) while $n = 27$ participants experienced the reverse order (bi – uni; bottom horizontal panel). In both groups of participants, the graph structure was changed without prior announcement halfway through the third task run. Numbers indicate approximate run duration in minutes (min). Colors indicate graph condition (uni vs. bi; see legend). **(d)** Visualization of the relative magnitude of the outcome variable (e.g., behavioral responses or classifier probabilities; y-axis) for specific transitions between the nodes (x-axis) and the two graph structures (uni vs. bi; horizontal panels) under the three assumptions (vertical panels), (1) that there is no difference between transitions (null hypothesis), (2) that response times are only influenced by the one-step transition probabilities between the nodes (colors), or (3) that response times are influenced by the multi-step relationships between nodes in the graph structure (here indicated by node distance). An effect of unidirectional graph structure would be evident in a linear relationship between node distance and the outcome variable, whereas a bidirectional graph structure would be reflected in a U-shaped relationship between node distance and independent measures (possibly inverted, depending on the measure). The stimulus material (individual images of a bear, a dromedary, a dear, an eagle, an elephant and a fox) shown in (a), and (b) were taken from a set of colored and shaded images commissioned by Rossion and Pourtois (2004), which are loosely based on images from the original Snodgrass and Vanderwart set (Snodgrass and Vanderwart, 1980). The images are freely available from the internet at `https://sites.google.com/andrew.cmu.edu/tarrlab/resources/tarrlab-stimuli` under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license (CC BY-NC-SA 3.0; for details, see `https://creativecommons.org/licenses/by-nc-sa/3.0/`). Stimulus images courtesy of Michael J. Tarr, Carnegie Mellon University, (for details, see `http://www.tarrlab.org/`).

A first analysis revealed that participants reacted faster and more accurately to transitions with high compared to low one-step probabilities in the unidirectional graph condition ($p_{ij} = 0.7$ versus $p_{ij} = 0.1$ transition probabilities, ps < 0.001), and in the bidirectional graph condition ($p_{ij} = 0.35$ versus $p_{ij} = 0.1$, ps < 0.001, Fig. 3a–b). In order to investigate whether multi-step transition probabilities also influenced participants' behavior, we then analyzed response times and error rates as a function of the node distance (Fig. 2d; for details, see Methods). Using this analysis approach, we found a significant effect of node distance on response times in both unidirectional, $F_{1.00,115.78} = 44.34$, $p < 0.001$, and bidirectional data, $F_{1.00,38.00} = 57.36$, $p < 0.001$ (Fig. 3c). To further disentangle the effects of one-step and multi-step knowledge, we excluded data of frequent transitions ($p_{ij} = 0.7$ and $p_{ij} = 0.35$ in the uni and bi conditions, respectively). In this case, the effect of node distance on response times in the unidirectional condition disappeared, $F_{1.00,72.32} = 0.43$, $p = 0.51$, but persisted in bidirectional data, $F_{1.00,76.98} = 5.52$, $p = 0.02$ (Fig. 3c). No effects on behavioral accuracy were observed in either of the above analyses (all ps > 0.11).

While these results offer a first indication of incidental learning of multi-step transitions, node distance is only an approximate reflection of the graph structure. A more precise way to express multi-step knowledge is to consider the discounted sum of different n-step probabilities as experienced by participants. This is equivalent to successor representation (SR) models (Dayan, 1993), which assume a representation of each node that reflects the discounted long-term occupation probability of all other nodes starting from the current node. Notably, recent work has shown that SRs can be updated through replay, rather than through online experience alone (Russek et al., 2017). We therefore investigated whether behavior reflected integrated mental SR-based maps of the experienced graph structure.

Specifically, for each node we modeled a vector that reflected the probability that starting from there a participant would experience any of the other nodes over a future-discounted predictive horizon. This vector was dynamically updated following the transitions that participants experienced in the task, using a temporal difference (TD) learning rule as used in SR models (Dayan, 1993; Russek et al., 2017). After experiencing the transition from image s_t to s_{t+1}, the row corresponding to image s_t of the successor matrix $\mathbf{M}$ was updated as

$$\mathbf{M}_{s_t,*} = \mathbf{M}_{s_t,*} + \alpha \left[\mathbf{1}_{s_{t+1}} + \gamma \mathbf{M}_{s_{t+1},*} - \mathbf{M}_{s_t,*} \right] \tag{1}$$

whereby $\mathbf{1}_{s_{t+1}}$ is a zero vector with a 1 in the ${s_{t+1}}^{\text{th}}$ position, and α is a learning rate. Crucially, the discounting parameter γ defined the extent to which multi-step transitions were taken into account, which we will henceforth refer to as the "predictive horizon" (cf. Gershman et al., 2012; Momennejad, 2020). We computed a series of SR models with different predictive horizons between $\gamma = 0$ (no predictive horizon) and $\gamma = 0.95$ (in steps of 0.05), and asked how well response times could be predicted from these individually calculated, time-varying SRs (for details, see Methods). We then compared different LME models of response time, with a Shannon surprise predictor (cf. Shannon, 1948) derived from each participants' SR model, in addition to fixed effects of task run, graph (uni vs. bi) and graph order (uni – bi vs. bi – uni) as well as by-participant random intercepts and slopes. Comparing LME models that contained predictors from SR models with varying predictive horizons (i.e., levels of γ) showed that a discount parameter of $\gamma = 0.3$ resulted in the lowest Akaike information criterion (AIC) score (Fig. 3d), and models with non-zero γ parameters yielded substantially better fits than a model which assumed only knowledge of one-step transitions ($\gamma = 0$, leftmost data point in Fig. 3d). Thus, participants' response times clearly indicated multi-step graph knowledge consistent with SR models.

To investigate if these analyses would differ between the two graph structures (uni vs. bi) and the two graph orders (uni – bi vs. bi – uni), we split the data according to these two factors and repeated a similar analysis of LME models (for details, see Methods). These analyses again showed that models based on a non-zero γ parameter achieved better fits, confirming that participants learned higher-order relationships among the nodes in the graph structure from experiencing sequences of transitions in the task (Fig. 3e). Interestingly, data from the first graph structure were fit best by the same γ parameter ($\gamma = 0.55$), irrespective of graph condition (uni vs. bi; Fig. 3e, left panel column). When considering data from the second graph structure, in contrast, the depth of integration differed markedly depending on whether participants learned the uni- or bidirectional graph structure: participants who transitioned from the uni- to the bidirectional graph condition had a larger predictive horizon ($\gamma = 0.75$; Fig. 3e, top right panel) in the second graph learning phase compared to participants who transitioned from a bi- to a unidirectional graph ($\gamma = 0.3$; Fig. 3e, bottom right panel). These results indicated that the order in which graphs were experienced determined the depth of integration when learning was updated following a change in transition probabilities.

Finally, we assessed whether participants were able to express knowledge of the sequential ordering of stimuli and graph structures explicitly during a post-task questionnaire. Asked whether they had noticed any sequential ordering of the stimuli in the preceding graph task, $n = 19$ participants replied "yes" and $n = 20$ replied "no" (Fig. 3f). Of those participants who noticed sequential ordering ($n = 19$), almost all (18 out of 19) indicated that they had noticed ordering within the first three runs of the task (Fig. 3g), and more than half of those participants (11 out of 19) indicated that they had noticed ordering during the third task run, i.e., the run during which the graph structure was changed.

Thus, sequential ordering of task stimuli remained at least partially implicit in half of the sample, and the change in the sequential order halfway through the third run of graph trials seemed to be one potential cause for the conscious realization of sequential structure. Participants were also asked to rate the transition probabilities of all pairwise sequential combinations of the six task stimuli (30 ratings in total). Interestingly, participants on average reported probability ratings that reflected bidirectional graph structure. Probabilities of transitions to clockwise and counterclockwise neighboring nodes were rated higher than rarer transitions to intermediate nodes, regardless of the order in which participants had experienced the two graph structures immediately before the questionnaire (Fig. 3h).

Figure 3: Behavioral responses are modulated by transition probabilities and graph structure. (a) Behavioral accuracy (y-axis) following transitions with low ($p_{ij} = 0.1$) and high probability (x-axis; $p_{ij} = 0.7$ and $p_{ij} = 0.35$ in the uni and bi conditions, respectively) for both graph structures (panels). Colors as in Fig. 2d. The horizontal dashed lines indicate the chance level (16.67%). **(b)** Log response time (y-axis) following transitions with low ($p_{ij} = 0.1$) and high probability (x-axis; $p_{ij} = 0.7$ and $p_{ij} = 0.35$ in the uni and bi conditions, respectively) for both graph structures (panels). Colors as in panel (a) and Fig. 2d. **(c)** Log response times (y-axis) as a function of uni- or bidirectional (u | b) node distance (x-axis) in data from the two graph structures (colors / panels). **(d)** AIC scores (y-axis) for LME models fit to participants' log response time data using Shannon surprise based on SRs with varying predictive horizons (the discounting parameter γ; x-axis) as the predictor variable. **(e)** AIC scores (y-axis) for LME models fit to participants' log response time data using Shannon information based on SRs with varying predictive horizons (the discounting parameter γ; x-axis) as the predictor variable, separated by graph order (uni – bi vs. bi – uni; horizontal panels) and graph condition (uni vs. bi; panel colors). **(f)** Number of participants (y-axis) indicating whether they had noticed any sequential ordering during the graph task ("yes" or "no", x-axis). **(g)** Number of those participants (y-axis) who had detected sequential ordering indicating in which of the five runs of the graph task (x-axis) they had first noticed sequential ordering. **(h)** Ratings of pairwise transition probabilities (in %; y-axis) as a function of node distance / transition probability, separately for both graph orderings (uni – bi vs. bi – uni; panels). Boxplots in (a), (b), (c), and (h) indicate the median and IQR. The lower and upper hinges correspond to the first and third quartiles (the 25^{th} and 75^{th} percentiles). The upper whisker extends from the hinge to the largest value no further than $1.5*$ IQR from the hinge (where IQR is the interquartile range (IQR), or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most $1.5*$ IQR of the hinge. The diamond shapes in (a), (b), (c), and (h) show the sample mean. Error bars and shaded areas in (a), (b), (c), and (h) indicate ± 1 SEM. Each dot in (a), (b), (c), and (h) corresponds to averaged data from one participant. Vertical lines in (d) and (e) mark the lowest AIC score. All statistics have been derived from data of $n = 39$ human participants who participated in one experiment.

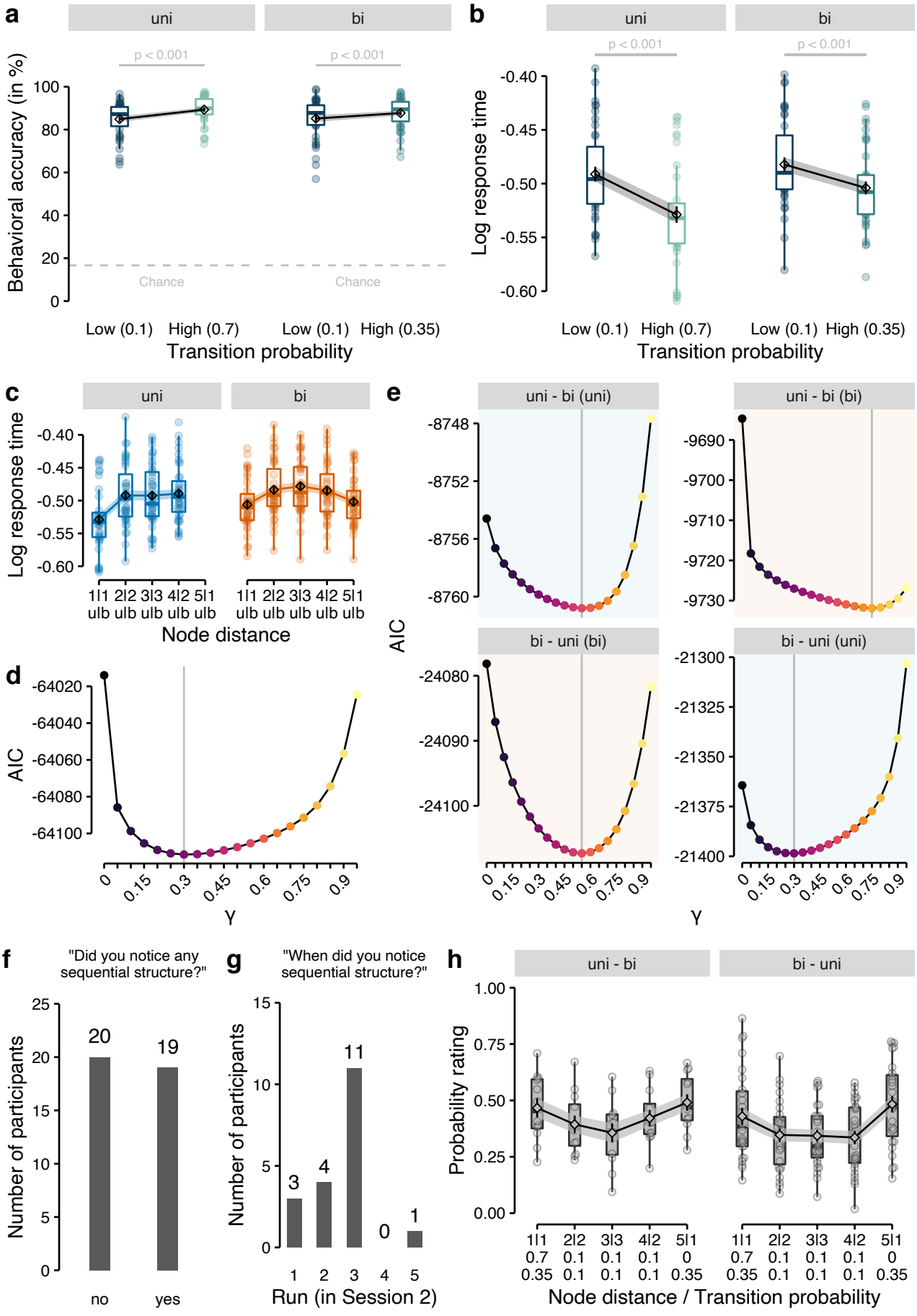


Figure 3: *[see caption on the previous page]*

fMRI results

We next asked whether learning of map-like graph representations was accompanied by on-task replay. First, we trained logistic regression classifiers on fMRI signals related to stimulus and response onsets in correct recall trials (one-versus-rest training; for details, see Methods; cf. Wittkuhn and Schuck, 2021). Separate classifiers were trained on data from gray-matter-restricted anatomical regions of interest (ROIs) of (a) occipito-temporal cortex and (b) pre- and postcentral gyri, which reflect visual object processing (cf. Haxby et al., 2001) and sensorimotor activity (e.g., Kolasinski et al., 2016), respectively. In each case, a single repetition time (TR) per trial corresponding either to the onset of the visual stimulus, or to participants' motor response was chosen (accounting for hemodynamic lag, time points were shifted by roughly 4 s; for details, see Methods). Note, that the order of displayed animals in recall trials was random, and image displays and motor responses were separated by SRIs and ITIs of 2500 ms to reduce temporal autocorrelation (cf. Dale, 1999; Wittkuhn and Schuck, 2021).

The trained classifiers successfully distinguished between the six animals. Leave-one-run-out classification accuracy was $M = 63.08\%$ in occipito-temporal data ($SD = 12.57$, $t_{38} = 23.06$, CI [59.69, $+\infty$], $p < 0.001$, compared to a chance level of 16.67%, $d = 3.69$) and $M = 47.05\%$ in motor cortex data ($SD = 7.79\%$, $t_{38} = 24.36$, CI [44.95, $+\infty$], $p < 0.001$, compared to a chance level of 16.67%, $d = 3.90$, all p-values Bonferroni-corrected, Fig. 4a). We also tested whether the classifiers successfully generalized from session 1 (eight recall runs) to session 2 (one recall run), and found no evidence for diminished cross-session decoding, compared to within-session, $F_{8.00,655.00} = 0.95$, $p = 0.48$ (for details see Methods). Next, we examined the sensitivity of the classifiers to pattern activation time courses by applying them to fifteen TRs following event onsets in recall trials (cf. Wittkuhn and Schuck, 2021). This analysis showed that the estimated normalized classification probability of the true stimulus class given the data peaked at the fourth TR as expected (Fig. 4b), where the probability of the true event was significantly higher than the mean probability of all other events at that time point (difference between current vs. other events; motor: $M = 12.24$, $t_{38} = 32.10$, CI [11.47, 13.01], $p < 0.001$, $d = 5.14$; occipito-temporal: $M = 17.88$, $t_{38} = 21.72$, CI [16.22, 19.55], $p < 0.001$, $d = 3.48$, all p-values Bonferroni-corrected; Fig. 4b).

To address our main questions concerning on-task neural replay, we applied the classifiers to data from the graph trials that included 10 s on-task intervals (ITIs) with only a fixation on screen (120 trials per participant in total; 24 trials per run; 4 trials per stimulus per run; 10 s correspond to 8 TRs). We expected that participants would replay anticipated upcoming events or recently experienced event sequences during these on-task intervals, and that such replay would be evident in the ordering of classification probabilities. Crucially, classifier probabilities should reflect participants' knowledge of one-step transitions, but also their map-like representations that enabled them to form multi-step expectations, as described above. For example, in unidirectional graph trials image A was followed by image B with a higher probability than the other images. Therefore, the probability of decoding image B during an on-task interval following image A should be higher than the classifier probabilities of the other four possible next images (see Fig. 2a). In addition, although images C, D, and E had equal one-step transition probabilities, we expected the corresponding classifier probabilities to be ordered such as to reflect the multi-step SR-model described above. Following our previous work (Wittkuhn and Schuck, 2021), we also assumed that the ordering during the earlier phase of the on-task interval (TRs 1–4) would reflect the true directionality of the replayed sequence and would be reversed in the later phase of the interval (TRs 5–8), reflecting the rising and falling slopes of the underlying hemodynamic response functions (HRFs). As expected, the classifier probability of the

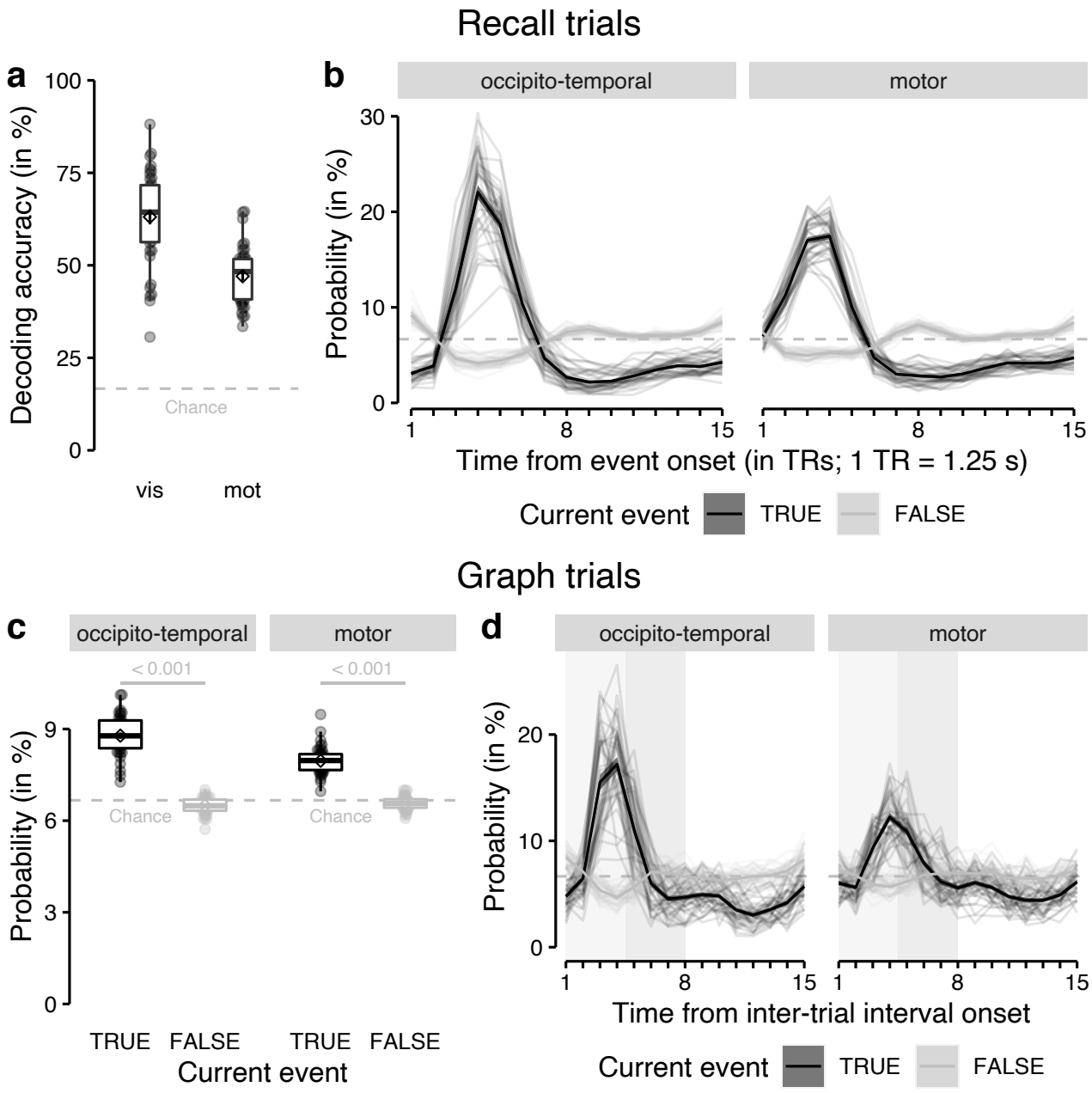


Figure 4: *[see caption on the next page]*

animal displayed in the current trial was higher compared to all other classes (Fig. 4c), and rising and falling slowly as observed in recall trials (Fig. 4d, Fig. 5a; mean probability of current event vs. all others; $ts \geq 17.88$, $ps < .001$, $ds \geq 3.48$, p-values Bonferroni-corrected). Because stimulus-evoked activation was not of interest, we removed probabilities of the current stimulus from all following analyses, considering only (normalized) probabilities from the five classes that did not occur on the current trial.

To investigate replay of experienced or anticipated stimulus sequences, we modeled classifier probabilities of non-displayed stimuli with LME models. LME models contained predictors that reflected node distance, i.e., how likely each stimulus was to appear soon, given either a unidirectional (linear node distance) or bidirectional graph (quadratic node distance, see above). Because linear and quadratic predictors were collinear, corresponding LME models were run separately. Each model included fixed effects of ROIs (occipito-temporal vs. sensorimotor) and ITI phase (early vs. late).

Figure 4: Classification accuracy and probabilistic classifier time courses on recall and graph trials. (a) Cross-validated classification accuracy (in %) in decoding the six unique visual objects in occipito-temporal data ("vis") and six unique motor responses in sensorimotor cortex data ("mot") during task performance. Chance level is at 16.67% (horizontal dashed line). **(b)** Time courses (in TRs from stimulus onset; x-axis) of probabilistic classification evidence (in %; y-axis) for the event on the current recall trial (black) compared to all other events (gray), separately for both ROIs (panels). **(c)** Mean classifier probability (in %; y-axis) for the event that occurred on the current graph trial (black color), shortly before the onset of the on-task interval, compared to all other events (gray color), averaged across all TRs in the on-task interval, separately for each ROI (panels). **(d)** Time courses (in TRs from on-task interval onset; x-axis) of mean probabilistic classification evidence (in %; y-axis) in graph trials for the event that occurred on the current trial (black) and all other events (gray). Each line in (b) and (c) represents one participant. Classifier probabilities in (b), (c), and (d) were normalized across 15 TRs. The chance level therefore is at $100/15 = 6.67\%$ (horizontal dashed line). Gray rectangles in (d) indicate the on-task interval (TRs 1–8). The light and dark gray areas in (d) indicate early (TRs 1–4) and late (TRs 5–8) phases, respectively. Boxplots in (a) and (c) indicate the median and IQR. The lower and upper hinges correspond to the first and third quartiles (the 25^{th} and 75^{th} percentiles). The upper whisker extends from the hinge to the largest value no further than $1.5*$ IQR from the hinge (where IQR is the interquartile range (IQR), or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most $1.5*$ IQR of the hinge. The diamond shapes in (a) and (c) show the sample mean. Error bars and shaded areas indicate ± 1 SEM. Each dot corresponds to averaged data from one participant. All statistics have been derived from data of $n = 39$ human participants who participated in one experiment.

Considering data from runs in which stimulus transitions were governed by the unidirectional graph, an LME model containing the linear node distance predictor indicated a three-way interaction between node distance, ROI and phase $F_{1.00,852.00} = 7.21$, $p = 0.007$. Post-hoc tests revealed an effect of node distance on classifier probabilities in unidirectional data in both ROIs in the early phase (TRs 1–4) of the ITIs, $F_{1.00,810.00} \geq 78.18$, ps < 0.001, akin to backward replay of recently experienced stimuli. Effects in the late phase failed to reach significance (TRs 5–8), ps ≤ 0.11 (Fig. 5c). Considering data from the bidirectional run, we found a corresponding three-way interaction between bidirectional node distance, ROI and phase $F_{1.00,852.00} = 5.59$, $p = 0.02$. Again, post-hoc tests revealed an effect of bidirectional node distance on classifier probabilities in both ROIs, showing a sign reversal when comparing the early to the late phase of the ITIs, $F_{1.00,810.00} \geq 7.09$, ps ≤ 0.008 (Fig. 5c), in line with our expectations about on-task multi-step replay. Although linear and quadratic node distance predictors were collinear and therefore difficult to disentangle, we next tried to assess the specificity of the above effects by testing the linear (unidirectional) node distance on bidirectional data and the quadratic (bidirectional) node distance on unidirectional data. When a linear predictor was used in an LME model of bidirectional data, only a main effect of phase (early vs. late) was observed, $F_{1.00,852.00} = 11.55$, $p < 0.001$, but no main effect of the linear predictor, $F_{1.00,852.00} = 0.27$, $p = 0.60$, or any interactions among the predictor variables, ps ≤ 0.09. Importantly, direct model comparison revealed that the linear model fit better in the unidirectional graph condition and the early phase of the ITI (see Fig. S6a–b). Using the quadratic predictor in the analysis of unidirectional data, we observed a three-way interaction between bidirectional node distance, the ROI, and the phase, $F_{1.00,852.00} = 4.35$, $p = 0.04$. Post-hoc tests revealed an effect of bidirectional node distance on classifier probabilities in unidirectional data only in the occipito-temporal ROI and only in the early phase (TRs 1–4) of the ITIs, $F_{1.00,810.00} \geq 5.56$, ps < 0.02 (Fig. 5c). Yet, model comparison again showed that the the quadratic model fit better in the bidirectional graph condition in both TR phases (differences in AICs were between -31.02 and 162.03, see Fig. S6a–b). Hence, these analyses confirmed that the observed classifier ordering was specific to the currently experienced graph.

The above analysis assumed that replayed sequences would always follow the most likely transitions (assuming a fixed ordering of replay sequences according to the multi-step graph structure). Yet, replay

might correspond more closely to a mental simulation of several possible sequences that are generated from a mental model. Consistent with this idea, the distribution of the observed sequential orders of classifier probabilities indicated a wide variety of replayed sequences (Fig. 5d, distribution over the entire ITI of 8 TRs). We next quantified how likely each possible sequential ordering of 5–item sequences was, based on the transition probabilities estimated by the SR model described above (γ was set to 0.3 in order to approximate to the mean level of planning depth we had estimated based on the behavioral data, see above). To model measurement noise in the observed relative to the predicted sequences, we employed a hidden markov model (HMM) with structured emission probabilities (for details, see Methods). This revealed that during the unidirectional runs, the frequency with which we observed a sequence in brain data during the on-task pauses, strongly related to the probability of that sequence given the unidirectional graph structure (occipito-temporal ROI: $r = .51$, $p < 0.001$; motor ROI: $r = .35$, $p < 0.001$; Fig. 5e). Unexpectedly, this was not the case for the bidirectional runs ($p = 0.21$ and $p = 0.50$, respectively; Fig. 5e).

We then sought to characterize the time courses of evidence for replay of sequences most likely to occur when mentally simulating a given sequence in the two graph structures. To this end, we calculated TR-wise linear regression slopes between the classifier probabilities and the 24 most likely sequences (top 20% of the $5! = 120$ possible permutations), which resulted in an average sequentiality metric for each TR, similar to our previous work (Wittkuhn and Schuck, 2021). This analysis revealed significant backward sequentiality in the earlier phase (TRs 1–4) of the ITIs based on data from the unidirectional graph structure in both ROIs specifically for those sequences that were most likely given the unidirectional graph structure, t_{38}'s ≤ -7.51, ps < 0.001, p-values Bonferroni-corrected (80 – 100%; Fig. 5e). We did not find evidence for sequentiality in the late phase of the interval (TRs 5–8) for either ROI in the unidirectional condition (ps > 0.97). These findings mirror the results from the analysis of classification probabilities (see above) in showing that classifier probabilities in earlier TRs of fMRI data with unidirectional graph structure are ordered backward relative to the sequential ordering implied by the graph structure. In the bidirectional condition, we found forward sequentiality in the earlier phase (TRs 1–4; t_{38}'s ≥ 3.90, ps < 0.02, ds ≥ 0.63) of the ITI and backward sequentiality in the later phase (TRs 5–8; t_{38}'s ≤ -4.31, ps < 0.001, ds ≤ -0.69), in occipito-temporal data for the top 40% most likely sequences (i.e., both 80–100% and 60–80%, p-values Bonferroni-corrected, Fig. 5e). Again, these results were in line with the analyses of classification probabilities, that found an influence on bidirectional graph structure in both early and late TRs.

Together, these results provide evidence that classifier probabilities in ITIs of graph trials are modulated by the multi-step distances between nodes in the graph structure. These effects of multi-step distances are in line with the idea that participants replayed multi-step sequences during brief on-task pauses, which could provide the basis for participants' map-like knowledge of incidentally experienced graph structures. When transition probabilities among stimuli in the task followed a unidirectional graph structure, classifier probabilities are influenced by a linear ordering of nodes that scales with the distance among the nodes in a unidirectional ordering, albeit only in earlier TRs following ITI onset (Fig. 5). When classifier probabilities from trials of the bidirectional graph structure are considered, classifier probabilities are influenced by a quadratic relationship to node distance (modeling a bidirectional ordering of nodes), in both the early (TRs 1–4) and late (TRs 5–8) phases of the ITIs and in both ROIs (Fig. 5). The graph distance effect appeared more pronounced in earlier compared to later TRs, but was present in both occipito-temporal and motor ROIs and followed a similar dynamic with respect to early and late phases of the ITI in both ROIs.

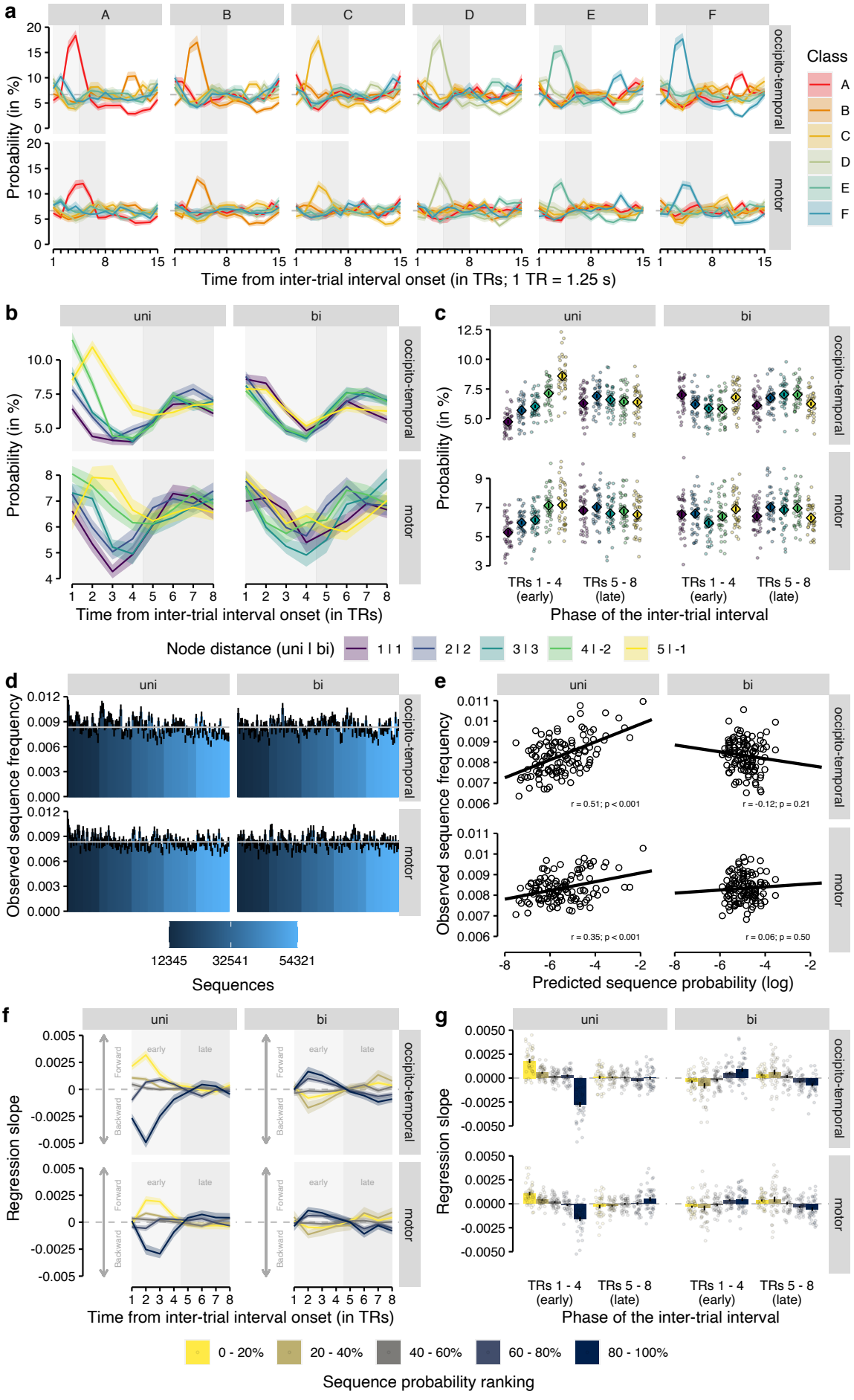


Figure 5: *[see caption on the next page]*

Figure 5: Classifier probabilities during inter-trial intervals (ITIs) of graph trials are modulated by node distances in the graph structure. **(a)** Time courses (in TRs from ITI onset; x-axis) of mean probabilistic classification evidence (in %; y-axis) for each of the six classes (colors) depending on the event of the current trial (vertical panels) and the anatomical ROI (horizontal panels). The event of the current trial (stimulus presentation or motor response) happened a few hundred ms before the onset of the ITI (for the trial procedure of graph trials, see Fig. 1b). **(b)** Time courses (in TRs from ITI onset; x-axis) of mean probabilistic classification evidence (in %; y-axis) for each of the five classes that were not presented on the current trial, colored by node distance in the two graph structures (vertical panels) for both anatomical ROI (horizontal panels). **(c)** Mean probabilistic classification evidence (in %; y-axis) for each node distance (colors) in the unidirectional (left vertical panel) and bidirectional (right vertical panel) graph structures averaged across TRs in the early (TRs 1–4) or late (TRs 5–8) phase (x-axis) for data in the occipito-temporal (top horizontal panels) and motor (bottom horizontal panels) ROIs. **(d)** Relative frequencies (y-axis) of all 120 permutations of probability-ordered 5–item sequences within each TR observed during on-task intervals, separately for both graph structures (vertical panels) and anatomical ROIs (horizontal panels). The horizontal gray line indicates the expected frequency if all sequences would occur equally often ($1/120 = 0.008$). Colors indicate sequence ordering from forward (e.g., `12345`; dark blue) to backward (e.g., `54321`; light blue) sequences. **(e)** Correlations (Pearson's r) between the predicted sequence probability and the observed sequence frequency (120 5–item sequences per correlation), separately for both graph structures (vertical panels) and anatomical ROIs (horizontal panels). Each dot represents one 5–item sequence. **(f)** Regression slopes (y-axis) relating classifier probabilities to sequential positions for both graph structures (vertical panels) and anatomical ROIs (horizontal panels). Sequential orderings were determined based on a hidden markov model (HMM) identifying the most likely sequences based on the two graph structures (colors). Positive and negative slopes indicate forward and backward sequentiality, respectively (cf. Wittkuhn and Schuck, 2021). **(g)** Mean classifier probabilities averaged across all TRs in the early and late phase (x-axis) of the ITIs, separately for both graph structures (vertical panels) and anatomical ROIs (horizontal panels). Each dot in (c) and (g) corresponds to averaged data from one participant. Error bars in (c), (d), and (g) and shaded areas in (a), (b), and (f) represent ± 1 SEM. Gray rectangles in (a), (b), and (d) indicate the on-task interval (TRs 1–8). The light and dark gray areas in (a), (b), and (f) indicate early (TRs 1–4) and late (TRs 5–8) interval phases, respectively. 1 TR in (a), (b), and (f) = 1.25 s. All statistics have been derived from data of $n = 39$ human participants who participated in one experiment.

Discussion

We present results showing on-task cortical replay of future sequences simulated from a mental model of an experienced graph in humans. Replay was detected in visual and sensorimotor cortex while participants briefly paused during an incidental statistical learning task. Statistical regularities in our main task were governed by two graph structures, one of which determined transitions in the first half of the experiment, while the other one determined transitions in the second half. We demonstrate that participants' response times reflect continuous learning of future-discounted predictive expectations that go beyond knowledge of one-step transitions and are captured by temporal difference (TD) learning of a successor representation (SR) model (cf. Dayan, 1993). These behavioral effects are in line with our neural results which indicate on-task replay consistent with sampling from such an SR model. Participants did not receive explicit instructions to learn and about half of participants reported no explicit knowledge of the experienced sequentiality. Learning was therefore automatic and partially implicit.

Our behavioral results are consistent with previous findings showing that humans learn about networks of stimuli beyond one-step transitions (e.g., Schapiro et al., 2013; Karuza et al., 2016, 2017, 2019; Garvert et al., 2017; Kahn et al., 2018; Lynn and Bassett, 2020; Lynn et al., 2020a,b). Our computational modeling establishes a link between these behavioral effects and an online temporal difference (TD) learning mechanism that tracks the long-term visitation probabilities. Our findings add to a growing set of studies that uses models based on SRs (Dayan, 1993) to demonstrate the formation of predictive representations of task structure in human behavioral and neuroimaging data (Garvert et al., 2017; Russek et al., 2017; Momennejad et al., 2017; Momennejad, 2020; Russek et al., 2021). Through model comparisons between SR models that differed in their discounting parameter γ, i.e., their predictive horizon, we found that behavior overall was best explained by a medium deep predictive horizon corresponding to $\gamma = 0.3$ (note, that any model with $\gamma > 0$ suggests that participants formed predictive representations). When we separated the analyses by graph condition and graph order, we found that during learning of the first graph structure, planning depth was deeper, as indicated by a predictive horizon of $\gamma = 0.55$, irrespective of whether transition structure was governed by the uni- or bidirectional graph condition. This finding suggests that, upon entering a novel environment with sequential events, humans might integrate multi-step transition probabilities to a medium depth that is independent from the specific structure of the environment. Interestingly, after the transition structure changed to the second graph structure halfway through the task, this also seemed to influence the predictive horizon in a manner that was dependent on the order in which the two graphs were experienced. In participants who first learned the unidirectional and then the bidirectional graph, the best fitting model was based on an SR with a higher discount parameter of $\gamma = 0.75$. This may indicate a deeper integration of higher-order relationships in the bidirectional graph structure compared to the unidirectional graph structure. In contrast, in participants who experienced the reverse order, the best fitting model during the second half of the experiment was based on an SR with a lower discount parameter of $\gamma = 0.3$. This could indicate a reduced predictive horizon when learning relationships in the unidirectional graph. In sum, these results suggest that participants' predictive horizon interacts with the structure of the task as well as the learning history and indicates that the depth of integration could adapt to changes in the task environment. This idea relates to recent work suggesting that the brain may host SRs at varying predictive horizons in parallel (Momennejad and Howard, 2018; Brunec and Momennejad, 2021).

Analyzing fMRI data recorded during 10 s pauses in-between performing the main task, we found evidence that classification probabilities were modulated by the transition probabilities and multi-step node distances within the two graph structures. Applying our previously developed sequentiality metric (Schuck and Niv, 2019; Wittkuhn and Schuck, 2021), we found evidence for backward sequentiality in unidirectional data and forward sequentiality in bidirectional data in both occipito-temporal and motor ROIs. The sequentiality metric was strongest specifically for those sequential orderings of classification probabilities that were most likely given an SR model of the two graph structures (Fig. 5). Our evidence for on-task replay relates to research in rodents, where time-compressed sequential place cell activations, called theta sequences, occur during active behavior (Foster and Wilson, 2007) and reflect multiple potential future trajectories when the animal pauses at a decision point (Johnson and Redish, 2007), or cycle between future trajectories during movement (Kay et al., 2020) possibly reflecting an online planning process. Similar relationships between hippocampal theta and planning have been observed in human magnetoencephalography (MEG) experiments (Kaplan et al., 2020), which have also yielded evidence for on-task planning in the form of fast sequential neural reactivation (Kurth-Nelson et al., 2016; Eldar et al., 2020). An fMRI study in humans has related on-task prospective neural activation to model-based decision-making (Doll et al., 2015), but the temporal dynamics of the prospective neural representations remained unclear. In contrast to previous studies, participants in our experiment did not engage in any explicit planning process. As mentioned before, participants were not instructed to learn about any sequentiality in the task. Moreover, participants were only told that short pauses may occur during the task, but they were not informed about the purpose of these pauses, and could not predict when the pauses would occur. It therefore seems likely that neural representations during on-task pauses reflect ongoing task representations similar to theta sequences in rodents.

One important aspect of our work is that we focused on cortical replay of predictive representations in visual (occipito-temporal) and sensorimotor (pre- and postcentral gyri) cortex. Previous work has largely focused on the hippocampus as a site of replay and as a potential brain region to host predictive cognitive maps (Garvert et al., 2017; Stachenfeld et al., 2017), while other studies have also emphasized the role of the prefrontal cortex (PFC) (Wilson et al., 2014; Schuck et al., 2016; Badre and Nee, 2018). Several fMRI studies demonstrated that hippocampal activity is modulated by stimulus predictability in sequential learning tasks (Strange et al., 2005; Harrison et al., 2006; Bornstein and Daw, 2012) and is related to the reinstatement of cortical task representations in visual cortex (Bosch et al., 2014; Hindy et al., 2016; Kok and Turk-Browne, 2018). Replay is known to occur throughout the brain (see e.g., Foster, 2017) but the functions of distributed replay events still remain to be further illuminated. Our findings shed light on the distribution of predictive representations and replay in the human brain, and suggest a potential involvement of sensory and motor areas. Yet, which roles the hippocampus and PFC play in this process remains an open question.

Our results suggest that participants formed a predominantly bidirectional representation of the ring-like graph structure, irrespective of the order in which the two graphs were experienced. The influence of node distance on response times was more pronounced and the predictive horizon in SR-based analyses was deeper in bidirectional compared to unidirectional behavioral data. Post-task ratings of transition probabilities were biased by bidirectional node distance, irrespective of graph order. The reversal in the directionality of classifier probabilities from early to late TRs, which is characteristic for sequential neural events in fMRI data (cf. Wittkuhn and Schuck, 2021), was only observed in on-task intervals during bidirectional but not unidirectional graph trials. This dominance

of a bidirectional representation could reflect that transitions in clockwise order in the unidirectional graph (e.g., from A to B; Fig. 2) still allow to infer an associative relationship in the reverse direction (i.e., from B to A), even though this transition actually never occurs during the task.

One remaining challenge for future research is to better understand the sequentiality of replay. We have previously shown that, at the level of classifier probabilities, sequences of neural events first elicit forward followed by backward sequentiality relative to the true sequence of events due to the dynamics of the HRF (Wittkuhn and Schuck, 2021). The fact that we found backward sequentiality in earlier TRs relative to an assumed sequential ordering of classifier probabilities in line with the unidirectional graph structure suggests that the true sequence of neural events at the start of the on-task intervals was indeed backwards. In the bidirectional graph structure, however, sequences can be expected in both directions, i.e., `A-B-C-D-E` and `E-D-C-B-A` sequences are both very likely. It therefore remains unclear whether detecting a replayed sequence of `A-B-C-D-E` reflects forward replay of this sequence or backward replay of its reverse (`E-D-C-B-A`). Previous research has found awake replay in both forward and backward order in rodents (Foster and Wilson, 2006; Diba and Buzsáki, 2007; Gupta et al., 2010) as well as in humans (Liu et al., 2021), and suggested that the directionality of replay may be tied to different functions, such as memory consolidation vs. value learning (e.g., Foster and Wilson, 2006; Ólafsdóttir et al., 2018; Liu et al., 2019; Wittkuhn et al., 2021). Neural sequences that have been associated with a prospective planning function are typically in forward order relative to the experienced sequence (Johnson and Redish, 2007; van der Meer and Redish, 2009; Pfeiffer and Foster, 2013; Wikenheiser and Redish, 2015b). However, as others have pointed out before (Kurth-Nelson et al., 2016), it is plausible to plan backward instead of forward (also see LaValle, 2006), and previous studies also reported backward sequences during theta in rodents (Wang et al., 2020) as well as during value learning in humans (Liu et al., 2021).

Another challenge will be to better understand the relation between changes in neural representations and replay. Repeated exposure to sequences of stimuli has been shown to increase the similarity of neural stimulus representations in the medial temporal lobe (MTL) in both macaques (Miyashita, 1988) and humans (Schapiro et al., 2012). Using fMRI adaptation (cf. Barron et al., 2016), Garvert et al. (2017) showed that the similarity of neural representations of task stimuli decreases with distance between stimuli in a graph structure. This may pose a challenge to classifiers trained on individual stimulus presentations as in the current study, because increases in the similarity of neural representations could increase the confusability of decoded patterns, which in turn may cause biases in the measured sequentiality.

In conclusion, our results provide insights into how the human brain forms predictive representations of the structural relationships in the environment from continuous experience and samples sequences from these internal cognitive maps during on-task replay.

Methods

Participants

44 young and healthy adults were recruited from an internal participant database or through local advertisement and fully completed the experiment. No statistical methods were used to predetermine the sample size but it was chosen to be larger than similar previous neuroimaging studies (e.g., Schuck and Niv, 2019; Momennejad et al., 2018; Tambini and Davachi, 2013). Five participants were excluded from further analysis because they viewed different animals in session 1 and 2 due to a programming error in the behavioral task. Thus, the final sample consisted of 39 participants (mean age = 24.28 years, $SD = 4.24$ years, age range: 18 - 33 years, 23 female, 16 male). All participants were screened for MRI eligibility during a telephone screening prior to participation and again at the beginning of each study session according to standard MRI safety guidelines (e.g., asking for metal implants, claustrophobia, etc.). None of the participants reported to have any major physical or mental health problems. All participants were required to be right-handed, to have corrected-to-normal vision, and to speak German fluently. The ethics commission of the German Psychological Society (DGPs) approved the study protocol (reference number: `SchuckNicolas2020-06-22VA`). All volunteers gave written informed consent prior to the beginning of the experiments. Every participant received 70.00 Euro and a performance-based bonus of up to 5.00 Euro upon completion of the study. None of the participants reported to have any prior experience with the stimuli or the behavioral task.

Task

Stimuli

All visual stimuli were taken from a set of colored and shaded images commissioned by Rossion and Pourtois (2004), which are loosely based on images from the original Snodgrass and Vanderwart set (Snodgrass and Vanderwart, 1980). The images are freely available on the internet at `https://sites.google.com/andrew.cmu.edu/tarrlab/resources/tarrlab-stimuli` under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license (for details, see `https://creativecommons.org/licenses/by-nc-sa/3.0/`) and have been used in similar previous studies (e.g., Garvert et al., 2017). Stimulus images courtesy of Michael J. Tarr at Carnegie Mellon University, (for details, see `http://www.tarrlab.org/`). In total, we selected 24 images which depicted animals that could be expected in a public zoo. Specifically, the images depicted a bear, a dromedary, a deer, an eagle, an elephant, a fox, a giraffe, a goat, a gorilla, a kangaroo, a leopard, a lion, an ostrich, an owl, a peacock, a penguin, a raccoon, a rhinoceros, a seal, a skunk, a swan, a tiger, a turtle, and a zebra (in alphabetical order). For each participant, six task stimuli were randomly selected from the set of 24 the animal images and each image was randomly assigned to one of six response buttons. This randomization ensured that any potential systematic differences between the stimuli (e.g., familiarity, preference, or ability to decode) would not influence the results on a group level (for a similar reasoning, see e.g., Liu et al., 2021). Cages were represented by a clipart illustration of a black fence which is freely available from `https://commons.wikimedia.org/wiki/File:Maki-fence-15.svg`, open-source and licensed under the Creative Commons CC0 1.0 Universal Public Domain Dedication, allowing further modification (for details, see `https://creativecommons.org/publicdomain/zero/1.0/`). When feedback was presented in the training and recall task conditions, correct responses were indicated by a fence colored in green and incorrect responses were signaled by a fence colored in red. The color

of the original image was modified accordingly. All stimuli were presented against a white background.

Hardware and software

Behavioral responses were collected using two 4-button inline fiber optic response pads (Current Designs, Philadelphia, PA, USA), one for each hand, with a linear arrangement of four buttons (buttons were colored in blue, yellow, green, and red, from left to right). The two response pads were attached horizontally to a rectangular cushion that was placed in participants' laps such that they could place their fingers on the response buttons with arms comfortably extended while resting on the scanner bed. Participants were asked to place their index, middle, and ring finger of their left and right hand on the yellow, green, and red buttons of the left and right response pads, respectively. The fourth (blue) button on each response pad was masked with tape and participants were instructed to never use this response button. Behavioral responses on the response pads were transferred to the computer running the experimental task and mapped to the keyboard keys `z`, `g`, `r` and `w`, `n`, `d` for the left and right hand, respectively. The task was programmed in PsychoPy3 (version 3.0.11; Peirce, 2007, 2008; Peirce et al., 2019) and run on a Windows 7 computer with a monitor refresh-rate of 16.7 ms. We recorded the presentation time stamps of all task events (onsets of all presentations of the fixation, stimulus, SRI, response, feedback, and ITI events) and confirmed that all components of the experimental task procedure were presented as expected.

Instructions

After participants entered the MRI scanner during the first study session and completed an anatomical T1-weighted (T1w) scan and a 5 min fMRI resting-state scan, they read the task instructions while lying inside the MRI scanner (for an illustration of the study procedure, see Fig. S1). Participants were asked to read all task instructions carefully (for the verbatim instructions, see Boxes S1 to S15). They were further instructed to clarify any potential questions with the study instructor right away and to lie as still and relaxed as possible for the entire duration of the MRI scanning procedure. As part of the instructions, participants were presented with a cover story in order to increase motivation and engagement (see Box S1). Participants were told to see themselves in the role of a zookeeper in training whose main task is to ensure that all animals are in the correct cages. In all task conditions, participants were asked to always keep their fingers on the response buttons to be able to respond as quickly and as accurately as possible. The full task instructions can be found in the supplementary information (SI), translated to English (see SI, starting on page 7, Boxes S1 to S15) from the original in German (see SI, page 11).

Training trials

After participants read the instructions and clarified all remaining questions with the study instructors via the intercom, they completed the *training* phase of the task. The training condition was designed to explicitly teach participants the assignment of stimuli to response buttons. Each of the six animal stimuli selected per participant was randomly assigned to one of six response buttons. For the training condition, participants were told to see themselves in the role of a zookeeper in training in a public zoo whose task is to learn which animal belongs in which cage (see Box S1). During each trial, participants saw six black cages at the bottom of the screen with each cage belonging to one of the six animals. On each trial, an animal appeared above one of the six cages. Participants were tasked to press the

response button for that cage as fast and accurately as possible and actively remember the cage where the animal belonged (see Box S3 and Box S4). The task instructions emphasized that it would be very important for participants to actively remember which animal belonged in which cage and that they would have the chance to earn a higher bonus if they learned the assignment and responded accurately (see Box S5).

In total, participants completed 30 trials of the training condition. Across all trials, the pairwise ordering of stimuli was set to be balanced, with each pairwise sequential combination of stimuli presented exactly once, i.e., with $n = 6$ stimuli, this resulted in $n * (n - 1) = 6 * (6 - 1) = 30$ trials. In this sense, the stimulus order was drawn from a graph with all nodes connected to each other and an equal probability of $p_{ij} = 0.2$ of transitioning from one node to any other node in the graph. This pairwise balancing of sequential combinations was used to ensure that participants would not learn any particular sequential order among the stimuli. Note, that this procedure only controlled for sequential order between pairs of consecutive stimuli but not higher-order sequential ordering of two steps or more.

On the first trial of the training condition, participants first saw a small black fixation cross that was displayed centrally on the screen for a fixed duration of 300 ms and signaled the onset of the following stimulus. The fixation cross was only shown on the first trial of the training phase, to allow for a short preparation signal before stimulus presentation began. Following the fixation cross, one of the animals was presented in the upper half of the screen above one of six cages that referred to the six response buttons and were presented in the lower half of the screen. The stimuli were shown for a fixed duration of 800 ms which was also the maximum time allowed for participants to respond. Note, that the instructions told participants that they would have 1 s to respond (see Box S4), an actual difference of 200 ms that was likely hardly noticeable. Following the stimulus, participants always received feedback that was shown for a fixed duration of 500 ms. If participants responded correctly, the cage corresponding to the correctly pressed response button, was shown in green. If participants did not respond correctly, the cage referring to the correct response button was shown in green and the cage referring to the incorrectly pressed response button was shown in red. If participants responded too late, the cage referring to the correct response button was shown in green and the German words "Zu langsam" (in English: "Too slow") appeared in large red letters in the upper half of the screen. Finally, a small black fixation cross was shown during an ITI with a variable duration of $M = 1500$ ms. The ITIs were drawn from a truncated exponential distribution with a mean of $M = 1.5$ s, a lower bound of $x_1 = 1.0$ s and an upper bound of $x_2 = 10.0$ s. To this end, we used the `truncexpon` distribution from the `SciPy` package (Virtanen et al., 2020) implemented in Python 3 (Van Rossum and Drake, 2009). The `truncexpon` distribution is described by three parameters, the shape b, the location μ and the scale β. The support of the distribution is defined by the lower and upper bounds, $[x_1, x_2]$, where $x_1 = \mu$ and $x_2 = b * \beta + \mu$. We solved the latter equation for the shape b to get $b = (x_2 - x_1)/\beta$. We chose the scale parameter β such that the mean of the distribution would be $M = 2.5$. To this end, we applied `scipy.optimize.fsolve` (Virtanen et al., 2020) to a function of the scale β that becomes zero when $truncexpon.mean((x_2 - x_1)/\beta, \mu, \beta) - M) = 2.5$. In total, the training phase took approximately 2 min to complete.

Recall trials

After participants finished the training phase of the task in the first experimental session, they completed eight runs of the *recall* condition and another ninth run at the beginning of the second session

(for an illustration of the study procedure, see Fig. S1). The recall condition of the task mainly served two purposes: First, the recall condition was used to further train participants on the associations between animal stimuli and response keys. Second, the recall condition was designed to elicit object-specific neural activation patterns of the presented visual animal stimuli and the following motor response. The resulting neural activation patterns were later used to train the probabilistic classifiers. The cover story of the instructions told participants that they would be tested on how well they have learned the association between animals and response keys during the training phase (see Box S6).

In total, participants completed nine runs of the recall condition. Eight runs were completed during session 1 and an additional ninth run was completed at the beginning of session 2 in order to remind participants about the S-R mappings (for an illustration of the study procedure, see Fig. S1). Each run consisted of 60 trials. As in the training phase, the proportion of pairwise sequential combinations of stimuli was balanced within a run. Across all trials, each pairwise sequential combination of stimuli was presented twice, i.e., with $n = 6$ stimuli, this results in $n * (n - 1) * 2 = 6 * (6 - 1) * 2 = 60$ trials. As for the training trials, the sequential ordering of stimuli was drawn from a graph with all nodes connected to each other and an equal probability of $p_{ij} = 0.2$ of transitioning from one node to any other node in the graph. With 60 trials per run, each of the six animal stimuli was shown 10 times per run. Given nine runs of the recall condition in total, this amounted to a maximum of 90 trials per stimulus per participant of training examples for the classifiers. Including a ninth run at the beginning of session 2 offered two advantages. First, participants were reminded about the associations between the stimuli and response keys that they had learned extensively during session 1. Second, the ninth run allowed to investigate decoding performance across session boundaries. Note, that the two experimental sessions were separated by about one week. Although the pre-processing of fMRI data (for details, see section on fMRI pre-processing below) should align the data of the two sessions, remaining differences between the two sessions (e.g., positioning of the participant in the MRI scanner) could lead to a decrement in decoding accuracy when testing classifiers that were trained on session 1 data to data from session 2. Our decoding approach was designed such that pattern classifiers would be mainly trained on neural data from recall trials in session 1 but then applied to data from session 2.

As in training trials, the first trial of each run in the recall phase started with a black fixation cross on a white background that was presented for a fixed duration of 300 ms. Only the first trial of a run contained a fixation cross, to provide a preparatory signal for participants which would later be substituted for by the ITI. Participants were then presented with one of the six animal stimuli that was presented centrally on the screen for a fixed duration of 500 ms. Participants were instructed to not respond to the stimulus (see instructions in Box S7). To check if participants indeed did not respond during the stimulus or the following SRI, we also recorded responses during these trial events. During the breaks between task runs, participants received feedback about the proportion of trials on which they responded too early. If participants responded too early, they were reminded by the study instructors to not respond before the response screen. A variable SRI followed the stimulus presentation during which a fixation cross was presented again. Including a jittered SRI ensured that the neural responses to the visual stimulus and the motor response could be separated in time and reduce temporal autocorrelation. Following the SRI, the cages indicating the response buttons were displayed centrally on the screen for a fixed duration of 800 ms, which was also the response time limit for participants. If participants responded incorrectly, the cage referring to the correct response button was shown in green and the cage referring to the incorrectly pressed response key was shown

in red. If participants responded too late, the cage referring to the correct response button was shown in green and the German words "Zu langsam" (in English: "Too slow") appeared in large red letters in the upper half of the screen. If participants responded correctly, the feedback screen was skipped. Each trial ended with an ITI with a variable duration of $M = 2.5$ s. Both SRIs and ITIs were drawn from a truncated exponential distribution as on training trials (for details, see description of training trials above).

Graph trials

Following the ninth run of the recall condition in session 2, participants completed five runs of the *graph* condition (for an illustration of the study procedure, see Fig. S1). During graph trials, participants were exposed to a fast-paced stream of the same six animal stimuli as in the training and recall phase. Unbeknownst to participants, the sequential ordering of animal stimuli followed particular transition probabilities.

During the graph task, the sequential order of stimuli across trials was determined by two graph structures with distinct transition probabilities. In the first graph structure, each node had a high probability ($p_{ij} = 0.7$) of transitioning to the next neighboring (i.e., transitioning from A to B, B to C, C to D, D to E, E to F, and F to A). Transitions to all other nodes (except the previous node) happened with equal probability of 0.1. Transitions to the previous node never occurred (transition probability of $p_{ij} = 0.0$). These transition probabilities resulted in a sequential ordering of stimuli that can be characterized by a continuous progression in a unidirectional (i.e., clockwise) order around the ring-like graph structure. We therefore termed this graph structure the *unidirectional graph* (or *uni* in short). The second graph structure allowed sequential ordering that could also progress in counterclockwise order. To this end, stimuli were now equally likely to transition to the next neighboring but also the previous node (probability of $p_{ij} = 0.35$, i.e., splitting up the probability of $p_{ij} = 0.7$ of transitioning to the next neighboring node only in the unidirectional graph structure). As in the unidirectional graph, transitions to all other nodes happened with equal probability of $p_{ij} = 0.1$. Given that stimuli could follow a sequential ordering in both directions of the ring, we refer to this graph structure as the *bidirectional graph* (or *bi* in short).

Participants completed five runs of the graph task condition. Each run consisted of 240 trials. Each stimulus was shown 40 times per run. In the unidirectional graph, for each stimulus the most likely transitions (probability of $p_{ij} = 0.7$) to the next neighboring node occurred 28 times per participant. Per stimulus and participant, 4 transitions to the other three possible nodes (low probability of $p_{ij} = 0.1$) happened. No transitions to the previous node happened when stimulus transitions were drawn from a unidirectional graph structure. Together, this resulted in $28 + 4 * 3 = 40$ presentations per stimulus, run and participant. For the bidirectional graph structure, transitions to the next neighboring and the previous node occurred 14 times per stimulus and to all other nodes 4 times as for the unidirectional graph structure. Together, this resulted in $14 + 14 + 4 * 3 = 40$ presentations per stimulus, run and participant.

As for the other task conditions, only the first trial of the graph phase started with the presentation of a small black fixation cross that was presented centrally on the screen for a fixed duration of 300 ms. Then, an animal stimulus was presented centrally on the screen for a fixed duration of 800 ms, which also constituted the time limit in which participants could respond with the correct response button. Participants did not receive feedback during the graph phase of the task in order to avoid any influence of feedback on graph learning. The stimulus was followed by an ITI with a mean duration

of 750 ms. The ITI in the graph trial phase was also drawn from a truncated exponential distribution with a mean of $M = 750$ ms, a lower bound of $x_1 = 500$ ms and an upper bound of $x_2 = 5000$ ms.

Importantly, during the graph task, we also included long ITIs of 10 s in order to investigate on-task replay. As stated above, participants completed 240 trials of the main task per run. In each run, each stimulus was shown on a total of 40 trials. For each stimulus, every 10^{th} trial on average was selected to be followed by a long ITI of 10 s. This meant that in each of the five main task runs, 4 trials per stimulus were followed by a long ITI. In total, each participant experienced 24 long ITI trials per run and 120 long ITI trials across the entire experiment. The duration of 10 s (roughly corresponding to eight TRs at a repetition time (TR) of 1.25 s) was chosen based on our previous results showing that the large majority of sequential fMRI signals can be captured within this time period (cf. Wittkuhn and Schuck, 2021, their Fig. 3).

Post-task questionnaire

After participants left the scanner in session 2, they were asked to complete a computerized post-task questionnaire consisting of four parts. First, participants were asked to report their handedness by selecting from three alternative options, "left", "right" or "both", in a forced-choice format. Note, that participants were required to be right-handed to participate in the study, hence this question merely served to record the self-reported handedness in addition to the participant details acquired as part of the recruitment procedure and demographic questionnaire assessment. Second, participants were asked whether they noticed any sequential order among the animal stimuli in the main task and could respond either "yes" or "no" in a forced-choice format. Third, if participants indicated that they noticed a sequential order of the stimuli (selecting "yes" on the previous question), they were asked to indicate during which run of the main task they had started to notice the ordering (selecting from run "1" to "5"). In case participants indicated that they did not notice a sequential ordering, they were asked to select "None" when asked about the run. Fourth, participants were presented with all sequential combinations of pairs of the animal stimuli and asked to indicate how likely animal A (on the left) was followed by animal B (on the right) during the Main condition of the task. Participants were instructed to follow their gut feeling in case they were uncertain about the probability ratings. With $n = 6$ stimuli, this resulted in $n * (n - 1) = 6 * (6 - 1) = 30$ trials. Participants indicated their response using a slider on a continuous scale from 0% to 100%. We recorded participants probability rating and response time on each trial. There was no time limit for any of the assessments in the questionnaire. Participants took $M = 5.49$ min ($SD = 2.38$ min; range: 2.23 to 12.63 min) to complete the questionnaire. The computerized questionnaire was programmed in PsychoPy3 (version 3.0.11; Peirce, 2007, 2008; Peirce et al., 2019) and run on the same Windows 7 computer that was used for the main experimental task.

Study procedure

All participants were screened for study and MRI eligibility during a telephone screening prior to participation. The study consisted of two experimental sessions. Upon arrival at the study center in both sessions, participants were first asked about any symptoms that could indicate an infection with the SARS-CoV-2 virus. The study instructors then measured participants' body temperature which was required to not be higher than $37.5°C$. Participants were asked to read and sign all the relevant study documents at home prior to their arrival at the study center.

Session 1 The first MRI session started with a short localizer sequence of ca. 1 min during which participants were asked to rest calmly, close their eyes and move as little as possible. Once the localizer data was acquired, the study personnel aligned the field of view (FOV) for the acquisition of the T1w sequence. The acquisition of the T1w sequence took about 4 min to complete. Using the anatomical precision of the T1w images, the study personnel then aligned the FOV of the functional MRI sequences. Here, the lower edge of the FOV was first aligned to the visually identified anterior commissure - posterior commissure (AC-PC) line of the participant's brain. The FOV was then manually titled by 20 degrees forwards relative to the rostro-caudal axis (positive tilt; for details see the section on "MRI data acquisition" on page 26). Shortly before the functional MRI sequences were acquired, we performed Advanced Shimming. During the shimming period, which took ca. 2 min, participants were again instructed to move as little as possible and additionally asked to avoid swallowing to further reduce any potential movements. Next, we acquired functional MRI data during a resting-state period of 5 min. For this phase, participants were instructed to keep their eyes open and fixate a white fixation cross that was presented on a black background. Acquiring fMRI resting-state data before participants had any exposure to the task allowed us to record a resting-state period that was guaranteed to be free of any task-related neural activation or reactivation. Following this pre-task resting-state scan, participants read the task instructions inside the MRI scanner and were able to clarify any questions with the study instructions via the intercom system. Participants then performed the training phase of the task (for details, see the section "Training trials" on page 21) while undergoing acquisition of functional MRI data. The training phase took circa 2 min to complete. Following the training phase, participants performed eight runs of the recall phase of the task of circa 6 min each while fMRI data was recorded. Before participants left the scanner, field maps were acquired.

Session 2 At the beginning of the second session, participants first completed the questionnaire for MRI eligibility and the questionnaire on COVID-19 symptoms before entering the MRI scanner again. As in the first session, the second MRI session started with the acquisition of a short localizer sequence and a T1w sequence followed by the orientation of the FOV for the functional acquisitions and the Advanced Shimming. Participants were asked to rest calmly and keep their eyes closed during this period. Next, during the first functional sequence of the second study session, participants performed a ninth run of the recall phase of the task in order to remind them about the correct response buttons associated with each of the six stimuli. We then acquired functional resting-state scans of 3 min each and functional task scans of 10 min each in an interleaved fashion, starting with a resting-state scan. During the acquisition of functional resting-state data, participants were asked to rest calmly and fixate a small white cross on a black background that was presented on the screen. During each of the functional task scans, participants performed the graph learning phase of the task (for details, see section "Graph trials" on page 24). Importantly, half-way through the third block of the main task, the graph structure was changed without prior announcement towards the second graph structure. After the sixth resting-state acquisition, field maps were acquired and participants left the MRI scanner.

MRI data acquisition

All MRI data were acquired using a 32-channel head coil on a research-dedicated 3-Tesla Siemens Magnetom TrioTim MRI scanner (Siemens, Erlangen, Germany) located at the Max Planck Institute for Human Development in Berlin, Germany.

At the beginning of each of the two MRI recording sessions, high-resolution T1w anatomical Mag-

netization Prepared Rapid Gradient Echo (MPRAGE) sequences were obtained from each participant to allow co-registration and brain surface reconstruction (sequence specification: 256 slices; TR = 1900 ms; echo time (TE) = 2.52 ms; flip angle (FA) = 9 degrees; inversion time (TI) = 900 ms; matrix size = 192 x 256; FOV = 192 x 256 mm; voxel size = 1 x 1 x 1 mm).

For the functional scans, whole-brain images were acquired using a segmented k-space and steady state T2*-weighted multi-band (MB) echo-planar imaging (EPI) single-echo gradient sequence that is sensitive to the blood-oxygen-level dependent (BOLD) contrast. This measures local magnetic changes caused by changes in blood oxygenation that accompany neural activity (sequence specification: 64 slices in interleaved ascending order; anterior-to-posterior (A-P) phase encoding direction; TR = 1250 ms; TE = 26 ms; voxel size = 2 x 2 x 2 mm; matrix = 96 x 96; FOV = 192 x 192 mm; FA = 71 degrees; distance factor = 0%; MB acceleration factor 4). Slices were tilted for each participant by 20 degrees forwards relative to the rostro-caudal axis (positive tilt) to improve the quality of fMRI signal from the hippocampus (cf. Weiskopf et al., 2006) while preserving good coverage of occipito-temporal and motor brain regions. The same sequence parameters were used for all acquisitions of fMRI data. For each functional task run, the task began after the acquisition of the first four volumes (i.e., after 5.00 s) to avoid partial saturation effects and allow for scanner equilibrium.

The first MRI session included nine functional task runs in total (for the study procedure, see Fig. S1). After participants read the task instructions inside the MRI scanner, they completed the training trials of the task which explicitly taught participants the correct mapping between stimuli and response keys. During this task phase, 80 volumes of fMRI were collected, which were not used in any further analysis. The other eight functional task runs during session 1 consisted of eight runs of the recall condition. Each run of the recall task was about 6 min in length, during which 320 functional volumes were acquired. We also recorded two functional runs of resting-state fMRI data, one before and one after the task runs. Each resting-state run was about 5 min in length, during which 233 functional volumes were acquired.

The second MRI session included six functional task runs in total (for the study procedure, see Fig. S1). After participants entered the MRI scanner, they completed a ninth run of the recall task. As before, this run of the recall task was also about 6 min in length, during which 320 functional volumes were acquired. Participants then completed five runs of the graph learning task. Each run of the five graph learning runs was about 10 min in length, during which 640 functional volumes were acquired. The five runs of the graph learning task were interleaved with six recordings of resting-state fMRI data, each about 3 min in length, during which 137 functional volumes were acquired.

At the end of each scanning session, two short acquisitions with six volumes each were collected using the same sequence parameters as for the functional scans but with varying phase encoding polarities, resulting in pairs of images with distortions going in opposite directions between the two acquisitions (also known as the *blip-up / blip-down* technique). From these pairs the displacement maps were estimated and used to correct for geometric distortions due to susceptibility-induced field inhomogeneities as implemented in the `fMRIPrep` preprocessing pipeline (Esteban et al., 2018) (see details below). In addition, a whole-brain spoiled gradient recalled (GR) field map with dual echo-time images (sequence specification: 36 slices; A-P phase encoding direction; TR = 400 ms; TE1 = 4.92 ms; TE2 = 7.38 ms; FA = 60 degrees; matrix size = 64 x 64; FOV = 192 x 192 mm; voxel size = 3 x 3 x 3.75 mm) was obtained as a potential alternative to the blip-up / blip-down method described above.

We also measured respiration during each scanning session using a pneumatic respiration belt as

part of the Siemens Physiological Measurement Unit (PMU). Pulse data could not be recorded as the recording device could not be attached to the participants' index finger as it would have otherwise interfered with the motor responses.

MRI data preparation

Conversion of data to the brain imaging data structure (BIDS) standard The majority of the steps involved in preparing and preprocessing the MRI data employed recently developed tools and workflows aimed at enhancing standardization and reproducibility of task-based fMRI studies (for a similar data processing pipeline, see e.g., Esteban et al., 2019a; Wittkuhn and Schuck, 2021). Version-controlled data and code management was performed using `DataLad` (version 0.13.0; Halchenko et al., 2019, 2021), supported by the `DataLad` handbook (Wagner et al., 2020). Following successful acquisition, all study data were arranged according to the brain imaging data structure (BIDS) specification (Gorgolewski et al., 2016) using the `HeuDiConv` tool (version 0.8.0.2; freely available from `https://github.com/ReproNim/reproin` or `https://hub.docker.com/r/repronim/reproin`) in combination with the `ReproIn` heuristic (Visconti di Oleggio Castello et al., 2020) (version 0.6.0) that allows automated creation of BIDS data sets from the acquired Digital Imaging and Communications in Medicine (DICOM) images. To this end, the sequence protocol of the MRI data acquisition was set up to conform with the specification required by the `ReproIn` heuristic (for details of the heuristic, see `https://github.com/nipy/heudiconv/blob/master/heudiconv/heuristics/reproin.py`). `HeuDiConv` was run inside a `Singularity` container (Kurtzer et al., 2017; Sochat et al., 2017) that was built from the most recent version (at the time of access) of a `Docker` container (tag 0.8.0.2), available from `https://hub.docker.com/r/repronim/reproin/tags`. DICOMs were converted to the NIfTI-1 format using `dcm2niix` (version 1.0.20190410GCC6.3.0; Li et al., 2016). In order to make personal identification of study participants unlikely, we eliminated facial features from all high-resolution structural images using `pydeface` (version 2.0.0; Gulban et al., 2019, available from `https://github.com/poldracklab/pydeface` or `https://hub.docker.com/r/poldracklab/pydeface`). `pydeface` (Gulban et al., 2019) was run inside a `Singularity` container (Kurtzer et al., 2017; Sochat et al., 2017) that was built from the most recent version (at the time of access) of a `Docker` container (tag 37-2e0c2d), available from `https://hub.docker.com/r/poldracklab/pydeface/tags` and used `Nipype`, version `1.3.0-rc1` (Gorgolewski et al., 2011, 2019). During the process of converting the study data to BIDS the data set was queried using `pybids` (version 0.12.1; Yarkoni et al., 2019a,b), and validated using the `bids-validator` (version 1.5.4; Gorgolewski et al., 2020). The `bids-validator` (Gorgolewski et al., 2020) was run inside a `Singularity` container (Kurtzer et al., 2017; Sochat et al., 2017) that was built from the most recent version (at the time of access) of a `Docker` container (tag v1.5.4), available from `https://hub.docker.com/r/bids/validator/tags`.

MRI data quality control The data quality of all functional and structural acquisitions were evaluated using the automated quality assessment tool `MRIQC`, version 0.15.2rc1 (for details, see Esteban et al., 2017, and the `MRIQC` documentation, available at `https://mriqc.readthedocs.io/en/stable/`). The visual group-level reports of the estimated image quality metrics confirmed that the overall MRI signal quality of both anatomical and functional scans was highly consistent across participants and runs within each participant.

MRI data preprocessing

Preprocessing of MRI data was performed using `fMRIPrep` 20.2.0 (long-term support (LTS) release; Esteban et al., 2018, 2019b, RRID:SCR_016216), which is based on `Nipype` 1.5.1 (Gorgolewski et al., 2011, 2019, RRID:SCR_002502). Many internal operations of `fMRIPrep` use `Nilearn` 0.6.2 (Abraham et al., 2014, RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in `fMRIPrep`'s documentation at `https://fmriprep.readthedocs.io/en/latest/workflows.html`. Note, that version 20.2.0 of `fMRIPrep` is a long-term support (LTS) release, offering long-term support and maintenance for four years.

Preprocessing of anatomical MRI data using `fMRIPrep` A total of two T1w images were found within the input BIDS data set, one from each study session. All of them were corrected for intensity non-uniformity (INU) using `N4BiasFieldCorrection` (Tustison et al., 2010), distributed with Advanced Normalization Tools (ANTs) 2.3.3 (Avants et al., 2008, RRID:SCR_004757). The T1w-reference was then skull-stripped with a Nipype implementation of the `antsBrainExtraction.sh` workflow (from ANTs), using `OASIS30ANTs` as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using `fast` (FMRIB Software Library (FSL) 5.0.9, RRID:SCR_002823, Zhang et al., 2001). A T1w-reference map was computed after registration of two T1w images (after INU-correction) using `mri_robust_template` (FreeSurfer 6.0.1, Reuter et al., 2010). Brain surfaces were reconstructed using `recon-all` (FreeSurfer 6.0.1, RRID:SCR_001847, Dale et al., 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical GM of Mindboggle (RRID:SCR_002438, Klein et al., 2017). Volume-based spatial normalization to two standard spaces (MNI152NLin6Asym, MNI152NLin2009cAsym) was performed through nonlinear registration with `antsRegistration` (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following templates were selected for spatial normalization: FSL's MNI ICBM 152 non-linear 6$^{\text{th}}$ Generation Asymmetric Average Brain Stereotaxic Registration Model (Evans et al., 2012, RRID:SCR_002823; TemplateFlow ID: MNI152NLin6Asym), ICBM 152 Nonlinear Asymmetrical template version 2009c (Fonov et al., 2009, RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym).

Preprocessing of functional MRI data using `fMRIPrep` For each of the BOLD runs found per participant (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of `fMRIPrep`. A B0-nonuniformity map (or fieldmap) was estimated based on two (or more) echo-planar imaging (EPI) references with opposing phase-encoding directions, with `3dQwarp` (Cox and Hyde, 1997, AFNI 20160207). Based on the estimated susceptibility distortion, a corrected echo-planar imaging (EPI) reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using `bbregister` (FreeSurfer) which implements boundary-based registration (Greve and Fischl, 2009). Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using `mcflirt` (FSL 5.0.9, Jenkinson et al., 2002). BOLD runs were slice-time corrected using `3dTshift` from AFNI 20160207 (Cox and Hyde, 1997, RRID:SCR_005927). The BOLD time-series were resampled onto the following surfaces (FreeSurfer reconstruction nomenclature): `fsnative`. The

BOLD time-series (including slice-timing correction) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as preprocessed BOLD in original space, or just preprocessed BOLD. The BOLD time-series were resampled into standard space, generating a preprocessed BOLD run in `MNI152NLin6Asym` space. First, a reference volume and its skull-stripped version were generated using a custom methodology of `fMRIPrep`. Several confounding time-series were calculated based on the preprocessed BOLD: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power et al. (absolute sum of relative motions, 2014) and Jenkinson et al. (relative root mean square displacement between affines, 2002). FD and DVARS are calculated for each functional run, both using their implementations in `Nipype` (following the definitions by Power et al., 2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (`CompCor`, Behzadi et al., 2007). Principal components are estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off) for the two `CompCor` variants: temporal (`tCompCor`) and anatomical (`aCompCor`). `tCompCor` components are then calculated from the top 2% variable voxels within the brain mask. For `aCompCor`, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. (2007) in that instead of eroding the masks by 2 pixels on BOLD space, the `aCompCor` masks are subtracted from a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's `aseg` segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, the masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each `CompCor` decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardized DVARS were annotated as motion outliers. All resamplings can be performed with a single interpolation step by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using `mri_vol2surf` (FreeSurfer).

Additional preprocessing of functional MRI data following `fMRIPrep` Following preprocessing using `fMRIPrep`, the fMRI data were spatially smoothed using a Gaussian mask with a standard deviation (Full Width at Half Maximum (FWHM) parameter) set to 4 mm using an example `Nipype` smoothing workflow (see the Nipype documentation for details) based on the Smallest Univalue Segment Assimilating Nucleus (SUSAN) algorithm as implemented in FSL (Smith and Brady, 1997). In this workflow, each run of fMRI data is separately smoothed using FSL's SUSAN algorithm with the

brightness threshold set to 75% of the median value of each run and a mask constituting the mean functional image of each run.

Multi-variate fMRI pattern analysis

All fMRI pattern classification analyses were conducted using the open-source `Python` (Python Software Foundation, Python Language Reference, version 3.8.6) packages `Nilearn` (version 0.7.0; Abraham et al., 2014) and `scikit-learn` (version 0.24.1; Pedregosa et al., 2011). In all classification analyses, we trained an ensemble of six independent classifiers, one for each of the six event classes. Depending on the analysis, these six classes either referred to the identity of the six visual animal stimuli or the identity of the participant's motor response, when training the classifiers with respect to the stimulus or the motor onset, respectively. For each class-specific classifier, labels of all other classes in the data were relabeled to a common "other" category. In order to ensure that the classifier estimates were not biased by relative differences in class frequency in the training set, the weights associated with each class were adjusted inversely proportional to the class frequencies in each training fold. Given that there were six classes to decode, the frequencies used to adjust the classifiers' weights were $\frac{1}{6}$ for the class of interest, and $\frac{5}{6}$ for the "other" class, comprising any other classes. Adjustments to minor imbalances caused by the exclusion of erroneous trials were performed in the same way. We used separate logistic regression classifiers with identical parameter settings. All classifiers were regularized using L2 regularization. The C parameter of the cost function was fixed at the default value of $C = 1.0$ for all participants. The classifiers employed the `lbfgs` algorithm to solve the multi-class optimization problem and were allowed to take a maximum of $4,000$ iterations to converge. Pattern classification was performed within each participant separately, never across participants. For each example in the training set, we added 4 s to the event onset and chose the volume closest to that time point (i.e., rounding to the nearest volume) to center the classifier training on the expected peaks of the BOLD response (for a similar approach, see e.g., Deuker et al., 2013). At a TR of 1.25 s this corresponded roughly to the fourth MRI volume which thus compromised a time window of 3.75 s to 5.0 s after each event onset. We detrended the fMRI data separately for each run across all task conditions to remove low frequency signal intensity drifts in the data due to noise from the MRI scanner. For each classifier and run, the features were standardized (z-scored) by removing the mean and scaling to unit variance separately for each training and test set.

Classification procedures First, in order to assess the ability of the classifiers to decode the correct class from fMRI patterns, we conducted a leave-one-run-out cross-validation procedure for which data from seven task runs of the recall phase in session 1 were used for training and data from the left-out run (i.e., the eighth run) from session 1 was used for testing the classification performance. This procedure was repeated eight times so that each task run served as the testing set once. Classifier training was performed on data from all correct recall trials of the seven runs in the respective cross-validation fold. In each iteration of the leave-one-run-out procedure, the classifiers trained on seven out of eight runs were then applied separately to the data from the left-out run. Specifically, the classifiers were applied to (1) data from the recall trials of the left-out run, selecting volumes capturing the expected activation peaks to determine classification accuracy, and (2) data from the recall trials of the left-out run, selecting all volumes from the volume closest to the stimulus or response onset and the next seven volumes to characterize temporal dynamics of probabilistic classifier predictions on a single trial basis.

Second, we assessed decoding performance on recall trials across the two experimental sessions. The large majority of fMRI data that was used to train the classifiers was collected in session 1 (eight of nine runs of the recall task), but the trained classifiers were mainly applied to fMRI data from session 2 (i.e., on-task intervals during graph trials). At the beginning of the second experimental session, participants completed another run of the recall task (i.e., a ninth run; for the study procedure, see Fig. S1). This additional task run mainly served the two purposes of (1) reminding participants about the correct S-R mapping that they had learned in session 1, and (2) to investigate the ability of the classifiers to correctly decode fMRI patterns in session 2 when they were only trained on session 1 data. This second aspect is crucial, as the main focus of investigation is the potential reactivation of neural task representations in session 2 fMRI data. Thus, it is important to demonstrate that this ability is not influenced by losses in decoding performance due to decoding across session boundaries. In order to test cross-session decoding, we thus trained the classifiers on all eight runs of the recall condition in session 1 and tested their decoding performance on the ninth run of the recall condition in session 2. Classifiers trained on data from all nine runs of the recall task were subsequently applied to data from on-task intervals in graph trials in session 2. For the classification analyses in on-task intervals of the graph task, classifiers were trained on the peak activation patterns from all correct recall trials (including session 1 and session 2 data) and then tested on all TR corresponding to the graph task ITIs.

Feature selection All participant-specific anatomical masks were created based on automated anatomical labeling of brain surface reconstructions from the individual T1w reference image created with Freesurfer's `recon-all` (Dale et al., 1999) as part of the `fMRIPrep` workflow (Esteban et al., 2018), in order to account for individual variability in macroscopic anatomy and to allow reliable labeling (Fischl et al., 2004; Poldrack, 2007). For the anatomical masks of occipito-temporal regions we selected the corresponding labels of the cuneus, lateral occipital sulcus, pericalcarine gyrus, superior parietal lobule, lingual gyrus, inferior parietal lobule, fusiform gyrus, inferior temporal gyrus, parahippocampal gyrus, and the middle temporal gyrus (cf. Haxby et al., 2001; Wittkuhn and Schuck, 2021). For the anatomical ROI of motor cortex, we selected the labels of the left and right gyrus precentralis as well as gyrus postcentralis. The labels of each ROI are listed in Table 1. Only gray-matter voxels were included in the generation of the masks as BOLD signal from non-gray-matter voxels cannot be generally interpreted as neural activity (Kunz et al., 2018). Note, however, that due to the whole-brain smoothing performed during preprocessing, voxel activation from brain regions outside the anatomical mask but within the sphere of the smoothing kernel might have entered the anatomical mask (thus, in principle, also including signal from surrounding non-gray-matter voxels).

ROI	Freesurfer labels (brain region)
Occipito-temporal	1005, 2005 (cuneus); 1011, 2011 (lateral occipital sulcus); 1021, 2021 (pericalcarine gyrus); 1029, 2029 (superio parietal lobule); 1013, 2013 (lingual gyrus); 1008, 2008 (inferior parietal lobule); 1007, 2007 (fusiform gyrus); 1009, 2009 (inferior temporal gyrus); 1016, 2016 (parahippocampal gyrus); 1015, 2015 (middle temporal gyrus)
Motor	1024, 2024 (left and right gyrus precentralis); 1022, 2022 (left and right gyrus postcentralis)

Table 1: Labels used to index brain regions to create participant-specific anatomical masks of selected ROIs based on Freesurfer's `recon-all` labels (Dale et al., 1999)

Statistical analyses

All statistical analyses were run inside a Docker software container or, if analyses were executed on a high performance computing (HPC), a Singularity version of the same container (Kurtzer et al., 2017; Sochat et al., 2017). All main statistical analyses were conducted using LME models employing the `lmer` function of the `lme4` package (version 1.1.27.1, Bates et al., 2015) in `R` (version 4.1.2, R Core Team, 2019). If not stated otherwise, all models were fit with participants considered as a random effect on both the intercept and slopes of the fixed effects, in accordance with results from Barr et al. (2013) who recommend to fit the most complex model consistent with the experimental design. If applicable, explanatory variables were standardized to a mean of zero and a standard deviation of one before they entered the models. If necessary, we removed by-participant slopes from the random effects structure to achieve a non-singular fit of the model (Barr et al., 2013). Models were fitted using the Bound Optimization BY Quadratic Approximation (BOBYQA) optimizer (Powell, 2007, 2009) with a maximum of $500,000$ function evaluations and no calculation of gradient and Hessian of nonlinear optimization solution. The likelihoods of the fitted models were assessed using Type III analysis of variance (ANOVA) with Satterthwaite's method. A single-step multiple comparison procedure between the means of the relevant factor levels was conducted using Tukey's honest significant difference (HSD) test (Tukey, 1949), as implemented in the `emmeans` package in `R` (version 1.7.0, Lenth, 2019; R Core Team, 2019). In all other analyses, we used one-sample t-tests if group data was compared to a baseline or paired t-tests if two samples from the same population were compared. If applicable, correction for multiple hypothesis testing was performed using the false discovery rate (FDR) (Benjamini and Hochberg, 1995) or Bonferroni (Bonferroni, 1936) correction method. If not stated otherwise, the α-level was set to $\alpha = 0.05$, and analyses of response times included data from correct trials only. When effects of stimulus transitions were analyzed, data from the first trial of each run and the first trial after the change in transition structure were removed.

Statistical analyses of behavioral data In order to test the a-priori hypothesis that behavioral accuracy in each of the nine runs of the recall trials and five runs of the graph trials would be higher than the chance-level, we performed a series of one-sided one-sample t-tests that compared participants' mean behavioral accuracy per run against the chance level of $100\%/6 = 16.67\%$. Participants' behavioral accuracy was calculated as the proportion of correct responses per run (in %). The effect sizes (Cohen's d) were calculated as the difference between the mean of behavioral accuracy scores across participants and the chance baseline (16.67%), divided by the standard deviation of the data (Cohen, 1988). The resulting p-values were adjusted for multiple comparisons using the Bonferroni correction (Bonferroni, 1936).

To examine the effect of task run on behavioral accuracy and response times in recall and graph trials, we conducted an LME model that included all nine task runs of the recall trials (or five runs of graph trials) as a numeric predictor variable (runs 1 to 9 and 1 to 5, respectively) as the main fixed effect of interest as well as random intercepts and slopes for each participant. We also conceived separate LME models that did not include data from the first task run of each task condition. These models only included eight task runs of the recall trials (or four runs of the graph trials) as a numeric predictor variable (runs 2 to 9 and 2 to 5, respectively) as the main fixed effect of interest as well as by-participant random intercepts and slopes.

Analyzing the effect of one-step transition probabilities on behavioral accuracy and response times, we conducted two-sided paired t-tests comparing the effect of high vs. low transition probability

separately for both unidirectional ($p_{ij} = 0.7$ vs. $p_{ij} = 0.1$) and bidirectional ($p_{ij} = 0.35$ vs. $p_{ij} = 0.1$) data. Effect sizes (Cohen's d) were calculated by dividing the mean difference of the paired samples by the standard deviation of the difference (Cohen, 1988) and p-values were adjusted for multiple comparisons across both graph conditions and response variables using the Bonferroni correction (Bonferroni, 1936).

In order to examine the effect of node distance on response times in graph trials, we conducted separate LME models for data from the unidirectional and bidirectional graph structures. For LME models of response time in unidirectional data, we included a linear predictor variable of node distance (assuming a linear increase of response time with node distance; see Fig. 2d top right) as well as random intercepts and slopes for each participant. The linear predictor variable was coded such that the node distance linearly increased from -2 to $+2$ in steps of 1, modeling the hypothesized increase of response time with node distance from 1 to 5 (centered on the node distance of 3). For LME models of response time in bidirectional data, we included a quadratic predictor variable of node distance (assuming an inverted U-shaped relationship between node distance and response time; see Fig. 2d bottom right) as well as by-participant random intercepts and slopes. The quadratic predictor variable of node distance was obtained by squaring the linear predictor variable. We also conducted separate LME models, that did not include data of the most frequent transitions in both the uni- and bi-directional data, but were otherwise specified in the same fashion.

Behavioral modeling based on the successor representation We modeled successor representations (SRs) for each participant depending on the transitions they experienced in the task, including training and recall trials. Specifically, each of the six stimuli was associated with a vector that reflected a *running* estimate of the long-term visitation probability of all six stimuli, starting from the present node. The successor matrix $\mathbf{M}^t$ was therefore a 6-by-6 matrix that contained six predictive vectors, one for each stimulus, and changed over time (hence the index t). The SR matrix on the first trial was initialized with a baseline expectation of $\frac{1}{36}$ for each node. After a transition between stimuli s_t and s_{t+1}, the matrix row corresponding to s_t was updated following a temporal difference (TD) learning rule (Dayan, 1993; Russek et al., 2017) as follows:

$$\mathbf{M}^t_{s_t,*} = \mathbf{M}^t_{s_t,*} + \alpha \left[\mathbf{1}_{s_{t+1}} + \gamma \mathbf{M}^t_{s_{t+1},*} - \mathbf{M}^t_{s_t,*} \right] \tag{2}$$

whereby $\mathbf{1}_{s_{t+1}}$ is a zero vector with a 1 in the ${s_{t+1}}^{\text{th}}$ position, $\mathbf{M}^t_{s_t,*}$ is the row corresponding to stimulus s_t of matrix $\mathbf{M}$. The learning rate α was arbitrarily set to a fixed value of 0.1, and the discount parameter γ was varied in increments of 0.05 from 0 to 0.95, as described in the main text. This meant that the SR matrix would change throughout the task to reflect the experienced transitions of each participant, first reflecting the random transitions experienced during the training and recall trials, then adapting to the first experienced graph structure and later to the second graph structure. In order to relate the SR models to participants' response times, we calculated how surprising each transition in the graph learning task was – assuming participants' expectations were based on the current SR on the given trial, $\mathbf{M}^t$. To this end, we normalized $\mathbf{M}^t$ to sum to 1, and then calculated the Shannon information (Shannon, 1948) for each trial, reflecting how surprising the just observed transition from stimulus i to j was given the history of previous transitions up to time point t:

$$I(j) = -\log_2(\tilde{m}^t_{i,j}) \tag{3}$$

where $\tilde{m}^t_{i,j}$ is the normalized $(i,j)^{\text{th}}$ entry of SR matrix $\mathbf{M}^t$. Using the base-2 logarithm allowed to express the units of information in bits (binary digits) and the negative sign ensured that the information measure was always positive or zero.

The final step in our analysis was to estimate LME models that tested how strongly this trial-wise measure of SR-based surprise was related to participants' response times in the graph learning task, for each level of the discount parameter γ. LME models therefore included fixed effects of the SR-based Shannon surprise, in addition to factors of task run, graph order (uni – bi vs. bi – uni) and graph structure (uni vs. bi) of the current run, as well as by-participant random intercepts and slopes. Separate LME models were conducted for each level of γ, and model comparison of the twenty models was performed using AIC, as reported in the main text. To independently investigate the effects of graph condition (uni vs. bi) and graph order (uni – bi vs. bi – uni), we analyzed separate LME models for each combination of the two factors, using only SR-based Shannon surprise as the main fixed effect of interest, and including by-participant random intercepts and slopes.

Statistical analysis of classification accuracy and single-trial decoding time courses In order to assess the classifiers' ability to differentiate between the neural activation patterns of individual visual objects and motor responses, we compared the predicted visual object or motor response of each example in the test set to the visual object or motor response that actually occurred on the corresponding trial. We obtained an average classification accuracy score for each participant by calculating the mean proportion of correct classifier predictions across all correctly answered recall trials in session 1 (Fig. 4a). The mean decoding accuracy scores of all participants were then compared to the chance baseline of $100\%/6 = 16.67\%$ using a one-sided one-sample t-test, testing the a-priori hypothesis that mean classification accuracy would be higher than the chance baseline. The effect size (Cohen's d) was calculated as the difference between the mean of accuracy scores and the chance baseline, divided by the standard deviation of the data (Cohen, 1988). These calculations were performed separately for each ROI and the resulting p-values were adjusted for multiple comparisons using Bonferroni correction (Bonferroni, 1936).

We also examined the effect of task run on classification accuracy in recall trials. To this end, we conducted an LME model including the task run as the main fixed effect of interest as well as by-participant random intercepts and slopes (Fig. 4c). We then assessed whether performance was above the chance level for all nine task runs and conducted nine separate one-sided one-sample t-tests separately per ROIs, testing the a-priori hypothesis that mean decoding accuracy would be higher than the 16.67% chance-level in each task run. All p-values were adjusted for 18 multiple comparisons (across nine runs and two ROIs) using the Bonferroni-correction (Bonferroni, 1936).

Furthermore, we assessed the classifiers' ability to accurately detect the presence of visual objects and motor responses on a single trial basis. For this analysis we applied the trained classifiers to fifteen volumes from the volume closest to the event onset and examined the time courses of the probabilistic classification evidence in response to the event on a single trial basis (Fig. 4b). In order to test if the time series of classifier probabilities reflected the expected increase of classifier probability for the event occurring on a given trial, we compared the time series of classifier probabilities related to the classified class with the mean time courses of all other classes using a two-sided paired t-test at the fourth TR from event onset. Classifier probabilities were normalized by dividing each classifier probability by the sum of the classifier probabilities across all fifteen TRs of a given trial. Here, we used the Bonferroni-correction method (Bonferroni, 1936) to adjust for multiple comparisons of

two observations. In the main text, we report the results for the peak in classification probability of the true class, corresponding to the fourth TR after stimulus onset. The effect size (Cohen's d) was calculated as the difference between the means of the probabilities of the current versus all other stimuli, divided by the standard deviation of the difference (Cohen, 1988).

Statistical analyses of classifier time courses on graph trials Classifier probabilities on graph trials indicated that the fMRI signal was strongly dominated by the activation of the event on the current trial. In order to test this effect, we calculated the mean classifier probabilities for the current and all other five events of the current trial across all eight TRs in the ITIs. The mean classifier probabilities of the current event were then compared to the mean classifier probabilities of all other events using two two-sided paired t-tests, one for each ROI. The Bonferroni-correction method Bonferroni (1936) was used to correct the p-values for two comparisons. The effect size (Cohen's d) was calculated as the difference between the means of the probabilities of the current versus all other events, divided by the standard deviation of the difference Cohen (1988).

After excluding data from the event of the current trial, we analyzed the effect of node distance on classifier probabilities for all non-displayed items using separate LME models for each graph structure, similar to the analysis of response times described above. Based on our previous findings indicating that the ordering of sequential neural events unfolds in the same order in earlier TRs and in reverse order in later TRs (cf. Wittkuhn and Schuck, 2021), we also included a fixed effect of interval phase (early TRs 1–4 vs. late TRs 5–8). In addition, each model included a fixed effect of ROI (occipito-temporal vs. sensorimotor). As for response times (see above), LME models of classifier probabilities in unidirectional or bidirectional data included a linear or quadratic predictor variable of node distance, respectively, as well as random intercepts and slopes for each participant. In order to examine the effect of a linear predictor in bidirectional data and the effect of the quadratic predictor in unidirectional data, predictor variables were switched accordingly, but otherwise the LME were conducted as before. Finally, we also directly compared the fits of a linear and quadratic model for each graph condition, ROI, and interval phase and quantified the model comparison using AIC.

Predicting sequence probability during on-task intervals We computed how likely it was to observe each 5-item sequence of stimuli under the assumption that participants were internally sampling from an SR model of the unidirectional or bidirectional graph structure. This was done in two steps.

First, we computed an ideal SR representation based on the true transition probabilities for each graph structure. Specifically, we defined the true transition function $\mathbf{T}$, as given by a graph, such that each entry t_{ij} reflected the true probability of transitioning from image i to j. Following the main ideas of the SR, we then calculated the long-term visitation probabilities as the time-discounted 5-step probabilities following the Chapman-Kolmogorov Equation:

$$\hat{\mathbf{M}} = \mathbf{T} + \gamma\mathbf{T}^2 + \gamma^2\mathbf{T}^3 + \gamma^3\mathbf{T}^4 + \gamma^4\mathbf{T}^5 \tag{4}$$

The discount rate γ was set to 0.3. We used five steps since more steps make little practical difference given the exponential discounting. The theoretical sequence probabilities for a given sequence $\mathbf{s}$ were then computed as the product of probabilities for all pairwise transitions (i, j) in the sequence, according to the approximated and normalized SR matrix:

$$p(\mathbf{s}) = \prod_{i,j \in \mathbf{s}} \tilde{\hat{m}}_{i,j} \tag{5}$$

Second, we approximated how likely it was to observe a sequence in the fMRI signal, given a particular sequence event in the brain. Our previous work has investigated which sequences are observed in classifier probabilities for a known true sequence (Wittkuhn and Schuck, 2021), and found that random reordering of items (induced by noise) was most prominent for the middle sequence items, and less severe for the start and end items. To model this effect, we set up a hidden markov model (HMM) in which the emission probabilities for the items that came first or last in a sequence were tuned sharply, sampled from a Gaussian distribution with a standard deviation of 0.5. This meant that the probability to observe the true item was 79%, and the probabilities to observe other items decreased sharply with distance from the true sequence position. The intermediate items had emission probabilities sampled from a Gaussian with a larger standard deviation of 2, yielding a much flatter distribution (probability to observe the true item at these positions was merely 19.9%). Using the HMM framework, we then computed the "forward" probabilities to observe a specific sequence given the transitions of a true sequence and the specified emission probabilities.

Finally, we combined the two probabilities that resulted from steps 1 and 2: (1) how likely a given sequence was to have resulted from a sample of an SR-based internal model of a graph structure, and (2) how likely it was to *observe* a sequence in the fMRI signal, given a specific sequence has been reactivated in the brain. To obtain our final estimates, we multiplied these probabilities for each sequence. This yielded the total probability to *observe* each sequence, assuming a true sequence distribution that results from sampling from the SR model, and a noise model that relates true to observed sequences.

To examine the relationship between predicted sequences based on this approach and observed sequences in fMRI during on-task intervals, we ordered the classes by their classifier probabilities within each TR (removing the class of the stimulus shown on the current trial) to obtain the observed frequencies for each of the possible 120 5–item sequences across all TRs of the on-task intervals during the graph learning task, separately for each participant, ROI and graph condition. The resulting distribution indicated how often classifier probabilities within TRs were ordered according to the 120 sequential 5–item combinations. This distribution was then averaged across participants for each of the 120 sequences and correlated with the sequence probability based on the HMM approach described above, separately for each ROI and graph condition (using Pearson's correlation across 120 data points).

Calculating the TR-wise sequentiality metric To analyze evidence for sequential replay during on-task intervals in graph trials, we calculated a sequentiality metric quantified by the slope of a linear regression between the classifier probabilities and each of the $5! = 120$ possible sequential orderings of a 5-item sequence in each TR, similar to our previous work (Wittkuhn and Schuck, 2021). We next separated the regression slope data based on how likely the permuted sequences were given the transition probabilities of the two graph structures in our experiment. To determine the probabilities of each possible sequential ordering of the 5–item sequences, we used the HMM approach described above to obtain the probability of all the $5! = 120$ sequences, assuming a particular starting position (i.e., the event on the current trial). Next, we ranked the permuted sequences according to their probability given the graph structures which allowed us to separately investigate sequentiality for the most and

the least likely sequences based on the graph structure. We then separated the ranked sequences into quintiles, i.e., five groups of ranked sequences from the least likely to the most likely 20%. Finally, we averaged the regression slopes separately for both ROIs, the two graph structures and the early and late TRs and compared the average slope against zero (the assumption of no sequentiality). The mean slope coefficients of all participants were compared to zero using a series of two-sided one-sample t-test, one for each graph condition, ROI, interval phase and sequence ranking bracket. p-values were adjusted for multiple comparisons using Bonferroni correction (Bonferroni, 1936). The effect size (Cohen's d) was calculated as the difference between the mean of slope coefficients and the baseline, divided by the standard deviation of the data (Cohen, 1988).

Data and code availability statement

Behavioral and MRI data as well as custom code used in this study will be made available upon publication in a peer-reviewed journal.

Acknowledgments

This work was supported by an Independent Max Planck Research Group grant awarded to N.W.S by the Max Planck Society (M.TN.A.BILD0004), and a Starting Grant awarded to N.W.S by the European Union (ERC-2019-StG REPLAY-852669). L.W. is a pre-doctoral fellow of the International Max Planck Research School on Computational Methods in Psychiatry and Ageing Research (IMPRS COMP2PSYCH). The participating institutions are the Max Planck Institute for Human Development, Berlin, Germany, and University College London, London, UK. For more information, see https://www.mps-ucl-centre.mpg.de/en/comp2psych. We also thank Leonardo Pettini for help with task development, Gregor Caregnato for help with participant recruitment and study coordination, Sonali Beckmann, Sam Chien (https://orcid.org/0000-0003-4306-1308), Theresa Fox, Sam Hall-McMaster (https://orcid.org/0000-0003-1641-979X), Nir Moneta (https://orcid.org/0000-0001-6125-4117), Liliana Polyanska (https://orcid.org/0000-0002-0842-8787), Nadine Taube, and Kateryna Yasynska – in alphabetical order of last names – for assistance with MRI data acquisition, Anika Löwe (https://orcid.org/0000-0003-3132-5767) for help with MRI data collection and comments on a previous version of this manuscript, Ondřej Zíka (https://orcid.org/0000-0003-0483-4443) for help with MRI data collection and statistical analyses, Michael Krause for help with high performance computing (HPC), all members of the Max Planck Research Group NeuroCode for helpful discussions about the contents of this manuscript, and all participants for their participation.

Author Contributions

The following list of author contributions is based on the CRediT taxonomy (Brand et al., 2015). For details on each type of author contribution, please see Brand et al. (2015).

Conceptualization: L.W., N.W.S.; Methodology: L.W., L.M.K., N.W.S.; Software: L.W., L.M.K., N.W.S.; Validation: L.W.; Formal analysis: L.W., N.W.S.; Investigation: L.W., L.M.K.; Resources: L.W., N.W.S.; Data curation: L.W., L.M.K.; Writing - original draft: L.W.; Writing - review & editing: L.W., L.M.K., N.W.S.; Visualization: L.W.; Supervision: N.W.S.; Project administration: L.W., N.W.S.; Funding acquisition: N.W.S.

Competing Interests

The authors declare no competing interests.

References

Alexandre Abraham, Fabian Pedregosa, Michael Eickenberg, Philippe Gervais, Andreas Mueller, Jean Kossaifi, Alexandre Gramfort, Bertrand Thirion, and Gaël Varoquaux. Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8, Feb 2014. ISSN 1662-5196. doi:10.3389/fninf.2014.00014. URL http://dx.doi.org/10.3389/fninf.2014.00014.

B. Avants, C Epstein, M. Grossman, and J. Gee. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1):26–41, Feb 2008. ISSN 1361-8415. doi:10.1016/j.media.2007.06.004. URL http://dx.doi.org/10.1016/j.media.2007.06.004.

David Badre and Derek Evan Nee. Frontal cortex and the hierarchical control of behavior. *Trends in Cognitive Sciences*, 22(2):170–188, 2018. ISSN 1364-6613. doi:https://doi.org/10.1016/j.tics.2017.11.005. URL https://www.sciencedirect.com/science/article/pii/S1364661317302450.

Jan Balaguer, Hugo Spiers, Demis Hassabis, and Christopher Summerfield. Neural mechanisms of hierarchical planning in a virtual subway network. *Neuron*, 90(4):893 – 903, 2016. ISSN 0896-6273. doi:https://doi.org/10.1016/j.neuron.2016.03.037. URL http://www.sciencedirect.com/science/article/pii/S0896627316300575.

Dale J. Barr, Roger Levy, Christoph Scheepers, and Harry J. Tily. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3):255–278, 2013. ISSN 0749596X. doi:10.1016/j.jml.2012.11.001. URL http://dx.doi.org/10.1016/j.jml.2012.11.001.

Helen C. Barron, Mona M. Garvert, and Timothy E. J. Behrens. Repetition suppression: a means to index neural representations using bold? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1705):20150355, Oct 2016. ISSN 1471-2970. doi:10.1098/rstb.2015.0355. URL http://dx.doi.org/10.1098/rstb.2015.0355.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015. ISSN 1548-7660. doi:10.18637/jss.v067.i01. URL https://www.jstatsoft.org/v067/i01.

Timothy E.J. Behrens, Timothy H. Muller, James C.R. Whittington, Shirley Mark, Alon B. Baram, Kimberly L. Stachenfeld, and Zeb Kurth-Nelson. What is a cognitive map? organizing knowledge for flexible behavior. *Neuron*, 100(2):490–509, Oct 2018. ISSN 0896-6273. doi:10.1016/j.neuron.2018.10.002. URL http://dx.doi.org/10.1016/j.neuron.2018.10.002.

Yashar Behzadi, Khaled Restom, Joy Liau, and Thomas T. Liu. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1):90–101, Aug 2007. ISSN 1053-8119. doi:10.1016/j.neuroimage.2007.04.042. URL http://dx.doi.org/10.1016/j.neuroimage.2007.04.042.

Jacob L. S. Bellmund, William de Cothi, Tom A. Ruiter, Matthias Nau, Caswell Barry, and Christian F. Doeller. Deforming the metric of cognitive maps distorts memory. *Nature Human Behaviour*, 4(2):177–188, 2020. doi:10.1038/s41562-019-0767-3. URL https://doi.org/10.1038/s41562-019-0767-3.

Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, 57(1):289–300, 1995. ISSN 00359246. URL http://www.jstor.org/stable/2346101.

Carlo Emilio Bonferroni. Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, 8:3–62, 1936.

Aaron M. Bornstein and Nathaniel D. Daw. Dissociating hippocampal and striatal contributions to sequential prediction learning. *European Journal of Neuroscience*, 35(7):1011–1023, 2012. doi:https://doi.org/10.1111/j.1460-9568.2011.07920.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1460-9568.2011.07920.x.

Sander E. Bosch, Janneke F. M. Jehee, Guillén Fernández, and Christian F. Doeller. Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *Journal of Neuroscience*, 34(22):7493–7500, 2014. ISSN 0270-6474. doi:10.1523/JNEUROSCI.0805-14.2014. URL https://www.jneurosci.org/content/34/22/7493.

Amy Brand, Liz Allen, Micah Altman, Marjorie Hlava, and Jo Scott. Beyond authorship: attribution, contribution, collaboration, and credit. *Learned Publishing*, 28(2):151–155, apr 2015. doi:10.1087/20150211. URL https://doi.org/10.1087/20150211.

Iva K. Brunec and Ida Momennejad. Predictive representations in hippocampal and prefrontal hierarchies. *Journal of Neuroscience*, 2021. ISSN 0270-6474. doi:10.1523/JNEUROSCI.1327-21.2021. URL https://www.jneurosci.org/content/early/2021/11/17/JNEUROSCI.1327-21.2021.

Margaret F Carr, Shantanu P Jadhav, and Loren M Frank. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nature Neuroscience*, 14(2):147–153, Jan 2011. ISSN 1546-1726. doi:10.1038/nn.2732. URL http://dx.doi.org/10.1038/nn.2732.

Jacob Cohen. Statistical power analysis for the behavioral sciences. *Lawrence Erlbaum Associates*, 1988.

Alexandra O. Constantinescu, Jill X. OReilly, and Timothy E. J. Behrens. Organizing conceptual knowledge in humans with a gridlike code. *Science*, 352(6292):1464–1468, Jun 2016. ISSN 1095-9203. doi:10.1126/science.aaf0941. URL http://dx.doi.org/10.1126/science.aaf0941.

Robert W. Cox and James S. Hyde. Software tools for analysis and visualization of fmri data. *NMR in Biomedicine*, 10(4-5):171–178, Jun 1997. ISSN 1099-1492. doi:10.1002/(sici)1099-1492(199706/08)10:4/5<171::aid-nbm453>3.0.co;2-l. URL http://dx.doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L.

Anders M. Dale. Optimal experimental design for event-related fmri. *Human Brain Mapping*, 8(2-3):109–114, 1999. ISSN 1097-0193. doi:10.1002/(sici)1097-0193(1999)8:2/3<109::aid-hbm7>3.0.co;2-w. URL http://dx.doi.org/10.1002/(SICI)1097-0193(1999)8:2/3<109::AID-HBM7>3.0.CO;2-W.

Anders M. Dale, Bruce Fischl, and Martin I. Sereno. Cortical surface-based analysis. *NeuroImage*, 9(2):179–194, Feb 1999. ISSN 1053-8119. doi:10.1006/nimg.1998.0395. URL http://dx.doi.org/10.1006/nimg.1998.0395.

Peter Dayan. Improving generalization for temporal difference learning: The successor representation. *Neural Computation*, 5(4):613–624, Jul 1993. ISSN 1530-888X. doi:10.1162/neco.1993.5.4.613. URL http://dx.doi.org/10.1162/neco.1993.5.4.613.

Lorena Deuker, J. Olligs, J. Fell, T. A. Kranz, F. Mormann, C. Montag, M. Reuter, C. E. Elger, and Nikolai Axmacher. Memory consolidation by replay of stimulus-specific neural activity. *Journal of Neuroscience*, 33(49):19373–19383, Dec 2013. ISSN 1529-2401. doi:10.1523/jneurosci.0414-13.2013. URL http://dx.doi.org/10.1523/JNEUROSCI.0414-13.2013.

Kamran Diba and György Buzsáki. Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience*, 10(10):1241–1242, Sep 2007. ISSN 1546-1726. doi:10.1038/nn1961. URL http://dx.doi.org/10.1038/nn1961.

Bradley B Doll, Katherine D Duncan, Dylan A Simon, Daphna Shohamy, and Nathaniel D Daw. Model-based choices involve prospective neural activity. *Nature Neuroscience*, 18(5):767–772, Mar 2015. ISSN 1546-1726. doi:10.1038/nn.3981. URL http://dx.doi.org/10.1038/nn.3981.

S. L. Eagleman and V. Dragoi. Image sequence reactivation in awake V4 networks. *Proceedings of the National Academy of Sciences*, 109(47):19450–19455, Nov 2012. ISSN 1091-6490. doi:10.1073/pnas.1212059109. URL http://dx.doi.org/10.1073/pnas.1212059109.

Matthias Ekman, Peter Kok, and Floris P. de Lange. Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications*, 8(15276):1–9, May 2017. ISSN 2041-1723. doi:10.1038/ncomms15276.

Eran Eldar, Gaëlle Lièvre, Peter Dayan, and Raymond J Dolan. The roles of online and offline replay in planning. *eLife*, 9, Jun 2020. ISSN 2050-084X. doi:10.7554/elife.56911. URL http://dx.doi.org/10.7554/eLife.56911.

Oscar Esteban, Daniel Birman, Marie Schaer, Oluwasanmi O. Koyejo, Russell A. Poldrack, and Krzysztof J. Gorgolewski. MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLoS ONE*, 12(9):e0184661, Sep 2017. ISSN 1932-6203. doi:10.1371/journal.pone.0184661. URL http://dx.doi.org/10.1371/journal.pone.0184661.

Oscar Esteban, Christopher J. Markiewicz, Ross W. Blair, Craig A. Moodie, A. Ilkay Isik, Asier Erramuzpe, James D. Kent, Mathias Goncalves, Elizabeth DuPre, Madeleine Snyder, and et al. fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, 16(1):111–116, Dec 2018. ISSN 1548-7105. doi:10.1038/s41592-018-0235-4. URL http://dx.doi.org/10.1038/s41592-018-0235-4.

Oscar Esteban, Rastko Ciric, Karolina Finc, Ross Blair, Christopher J. Markiewicz, Craig A. Moodie, James D. Kent, Mathias Goncalves, Elizabeth DuPre, Daniel E. P. Gomez, Zhifang Ye, Taylor Salo, Romain Valabregue, Inge K. Amlien, Franziskus Liem, Nir Jacoby, Hrvoje Stojić, Matthew Cieslak, Sebastian Urchs, Yaroslav O. Halchenko, Satrajit S. Ghosh, Alejandro De La Vega, Tal Yarkoni, Jessey Wright, William H. Thompson, Russell A. Poldrack, and Krzysztof J. Gorgolewski. Analysis of task-based functional MRI data preprocessed with fMRIPrep. *bioRxiv*, 2019a. doi:10.1101/694364. URL https://www.biorxiv.org/content/early/2019/07/08/694364.

Oscar Esteban, Christopher J. Markiewicz, Ross W. Blair, Craig A. Moodie, A. Ilkay Isik, Asier Erramuzpe, James D. Kent, Mathias Goncalves, Elizabeth DuPre, Madeleine Snyder, and et al. fMRIPrep 1.2.2., 2019b.

Alan C. Evans, Andrew L. Janke, D. Louis Collins, and Sylvain Baillet. Brain templates and atlases. *NeuroImage*, 62(2):911–922, 2012. ISSN 1053-8119. doi:https://doi.org/10.1016/j.neuroimage.2012.01.024. URL https://www.sciencedirect.com/science/article/pii/S1053811912000419.

Bruce Fischl, André van der Kouwe, Christophe Destrieux, Eric Halgren, Florent Ségonne, David H. Salat, Evelina Busa, Larry J. Seidman, Jill Goldstein, David Kennedy, Verne Caviness, Nikos Makris, Bruce Rosen, and Anders M. Dale. Automatically parcellating the human cerebral cortex. *Cerebral Cortex*, 14(1):11–22, Jan 2004. ISSN 1460-2199. doi:10.1093/cercor/bhg087. URL http://dx.doi.org/10.1093/cercor/bhg087.

VS Fonov, AC Evans, RC McKinstry, CR Almli, and DL Collins. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47:S102, Jul 2009. ISSN 1053-8119. doi:10.1016/s1053-8119(09)70884-5. URL http://dx.doi.org/10.1016/S1053-8119(09)70884-5.

David J. Foster. Replay comes of age. *Annual Review of Neuroscience*, 40(1):581–602, 2017. doi:10.1146/annurev-neuro-072116-031538. URL https://doi.org/10.1146/annurev-neuro-072116-031538.

David J. Foster and Matthew A. Wilson. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 440(7084):680–683, Feb 2006. ISSN 1476-4687. doi:10.1038/nature04587. URL http://dx.doi.org/10.1038/nature04587.

David J. Foster and Matthew A. Wilson. Hippocampal theta sequences. *Hippocampus*, 17(11):1093–1099, 2007. doi:10.1002/hipo.20345. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hipo.20345.

Mona M Garvert, Raymond J Dolan, and Timothy EJ Behrens. A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *eLife*, 6, Apr 2017. ISSN 2050-084X. doi:10.7554/elife.17086. URL http://dx.doi.org/10.7554/eLife.17086.

Jeffrey P Gavornik and Mark F Bear. Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nature Neuroscience*, 17(5):732–737, 2014. doi:10.1038/nn.3683. URL https://doi.org/10.1038/nn.3683.

Samuel J. Gershman, Christopher D. Moore, Michael T. Todd, Kenneth A. Norman, and Per B. Sederberg. The successor representation and temporal context. *Neural Computation*, 24(6):1553–1568, Jun 2012. ISSN 1530-888X. doi:10.1162/neco_a_00282. URL http://dx.doi.org/10.1162/NECO_a_00282.

Chris Gorgolewski, Nell Hardcastle, Teal Hobson-Lowther, David Nishikawa, Ross Blair, Stefan Appelhoff, Suyash, Constellates, Mainak Jas, Chris Holdgraf, Alexander Jones, Rohan Goyal, Robert Oostenveld, Chris Markiewicz, Gregory Noack, Matthew Zito, Joke Durnez, Nicolas Traut, Mikael

Naveau, Parul Sethi, Yaroslav Halchenko, Taylor Salo, Michael Hanke, Dimitri Papadopoulos Orfanos, Horea Christian, Franklin Feingold, Duncan Macleod, Dewarrn1, Brian Grass, and Adam Thomas. bids-standard/bids-validator: 1.4.3, 2020. URL https://zenodo.org/record/3688707.

Krzysztof J. Gorgolewski, Christopher D. Burns, Cindee Madison, Dav Clark, Yaroslav O. Halchenko, Michael L. Waskom, and Satrajit S. Ghosh. Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in Python. *Frontiers in Neuroinformatics*, 5, 2011. ISSN 1662-5196. doi:10.3389/fninf.2011.00013. URL http://dx.doi.org/10.3389/fninf.2011.00013.

Krzysztof J. Gorgolewski, Tibor Auer, Vince D. Calhoun, R. Cameron Craddock, Samir Das, Eugene P. Duff, Guillaume Flandin, Satrajit S. Ghosh, Tristan Glatard, Yaroslav O. Halchenko, and et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, 3(160044), Jun 2016. ISSN 2052-4463. doi:10.1038/sdata.2016.44. URL http://dx.doi.org/10.1038/sdata.2016.44.

Krzysztof J. Gorgolewski, Christopher D. Burns, Cindee Madison, Dav Clark, Yaroslav O. Halchenko, Michael L. Waskom, and Satrajit S. Ghosh. Nipype, 2019.

Douglas N. Greve and Bruce Fischl. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1):63–72, Oct 2009. ISSN 1053-8119. doi:10.1016/j.neuroimage.2009.06.060. URL http://dx.doi.org/10.1016/j.neuroimage.2009.06.060.

Omer Faruk Gulban, Dylan Nielson, Russ Poldrack, John Lee, Chris Gorgolewski, Vanessasaurus, and Satrajit Ghosh. poldracklab/pydeface: v2.0.0, 2019. URL https://zenodo.org/record/3524400.

Anoopum S. Gupta, Matthijs A.A. van der Meer, David S. Touretzky, and Aaron David Redish. Hippocampal replay is not a simple function of experience. *Neuron*, 65(5):695 – 705, 2010. ISSN 0896-6273. doi:10.1016/j.neuron.2010.01.034. URL http://www.sciencedirect.com/science/article/pii/S0896627310000607.

Yaroslav O. Halchenko, Michael Hanke, Benjamin Poldrack, Kyle Meyer, Debanjum Singh Solanky, Gergana Alteva, Jason Gors, Dave MacFarlane, Christian Olaf Häusler, Taylor Olson, Alex Waite, Alejandro De La Vega, Vanessa Sochat, Anisha Keshavan, Feilong Ma, Horea Christian, Jorrit Poelen, Kusti Skytén, Matteo Visconti di Oleggio Castello, Nell Hardcastle, Torsten Stoeter, Vicky C Lau, and Christopher J. Markiewicz. datalad/datalad 0.11.5, 2019. URL https://zenodo.org/record/3233911.

Yaroslav O. Halchenko, Kyle Meyer, Benjamin Poldrack, Debanjum Singh Solanky, Adina S. Wagner, Jason Gors, Dave MacFarlane, Dorian Pustina, Vanessa Sochat, Satrajit S. Ghosh, Christian Mönch, Christopher J. Markiewicz, Laura Waite, Ilya Shlyakhter, Alejandro de la Vega, Soichi Hayashi, Christian Olaf Häusler, Jean-Baptiste Poline, Tobias Kadelka, Kusti Skytén, Dorota Jarecka, David Kennedy, Ted Strauss, Matt Cieslak, Peter Vavra, Horea-Ioan Ioanas, Robin Schneider, Mika Pflüger, James V. Haxby, Simon B. Eickhoff, and Michael Hanke. DataLad: distributed system for joint management of code, data, and their relationship. *Journal of Open Source Software*, 6(63):3262, 2021. doi:10.21105/joss.03262. URL https://doi.org/10.21105/joss.03262.

L.M. Harrison, A. Duggins, and K.J. Friston. Encoding uncertainty in the hippocampus. *Neural Networks*, 19(5):535–546, 2006. ISSN 0893-6080. doi:https://doi.org/10.1016/j.neunet.2005.11.002. URL https://www.sciencedirect.com/science/article/pii/S0893608006000025.

James V. Haxby, M. Ida Gobbini, Maura L. Furey, Alumit Ishai, Jennifer L. Schouten, and Pietro Pietrini. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293(5539):2425–2430, Sep 2001. ISSN 1095-9203. doi:10.1126/science.1063736. URL http://dx.doi.org/10.1126/science.1063736.

Nicholas C Hindy, Felicia Y Ng, and Nicholas B Turk-Browne. Linking pattern completion in the hippocampus to predictive coding in visual cortex. *Nature Neuroscience*, 19(5):665–667, Apr 2016. ISSN 1546-1726. doi:10.1038/nn.4284. URL http://dx.doi.org/10.1038/nn.4284.

L. T. Hunt, N. D. Daw, P. Kaanders, M. A. MacIver, U. Mugan, E. Procyk, A. D. Redish, E. Russo, J. Scholl, K. Stachenfeld, C. R. E. Wilson, and N. Kolling. Formalizing planning and information search in naturalistic decision-making. *Nature Neuroscience*, 2021. doi:10.1038/s41593-021-00866-w. URL https://doi.org/10.1038/s41593-021-00866-w.

Mark Jenkinson, Peter Bannister, Michael Brady, and Stephen Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2): 825–841, Oct 2002. ISSN 1053-8119. doi:10.1006/nimg.2002.1132. URL http://dx.doi.org/10.1006/nimg.2002.1132.

Daoyun Ji and Matthew A Wilson. Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience*, 10(1):100–107, Dec 2006. ISSN 1546-1726. doi:10.1038/nn1825. URL http://dx.doi.org/10.1038/nn1825.

Adam Johnson and Aaron David Redish. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, 27(45):12176–12189, Nov 2007. ISSN 1529-2401. doi:10.1523/jneurosci.3761-07.2007. URL http://dx.doi.org/10.1523/JNEUROSCI.3761-07.2007.

Ari E. Kahn, Elisabeth A. Karuza, Jean M. Vettel, and Danielle S. Bassett. Network constraints on learnability of probabilistic motor sequences. *Nature Human Behaviour*, 2(12):936–947, Nov 2018. ISSN 2397-3374. doi:10.1038/s41562-018-0463-8. URL http://dx.doi.org/10.1038/s41562-018-0463-8.

Raphael Kaplan, Adrià Tauste Campo, Daniel Bush, John King, Alessandro Principe, Raphael Koster, Miguel Ley Nacher, Rodrigo Rocamora, and Karl J. Friston. Human hippocampal theta oscillations reflect sequential dependencies during spatial planning. *Cognitive Neuroscience*, 11(3):122–131, 2020. doi:10.1080/17588928.2019.1676711. URL https://doi.org/10.1080/17588928.2019.1676711. PMID: 31617790.

Elisabeth A. Karuza, Sharon L. Thompson-Schill, and Danielle S. Bassett. Local patterns to global architectures: Influences of network topology on human learning. *Trends in Cognitive Sciences*, 20(8):629–640, 2016. ISSN 1364-6613. doi:10.1016/j.tics.2016.06.003. URL https://www.sciencedirect.com/science/article/pii/S1364661316300717.

Elisabeth A. Karuza, Ari E. Kahn, Sharon L. Thompson-Schill, and Danielle S. Bassett. Process reveals structure: How a network is traversed mediates expectations about its architecture. *Scientific Reports*, 7(1):12733, 2017. doi:10.1038/s41598-017-12876-5. URL https://doi.org/10.1038/s41598-017-12876-5.

Elisabeth A. Karuza, Ari E. Kahn, and Danielle S. Bassett. Human sensitivity to community structure is robust to topological variation. *Complexity*, 2019:1–8, Feb 2019. ISSN 1099-0526. doi:10.1155/2019/8379321. URL http://dx.doi.org/10.1155/2019/8379321.

Kenneth Kay, Jason E. Chung, Marielena Sosa, Jonathan S. Schor, Mattias P. Karlsson, Margaret C. Larkin, Daniel F. Liu, and Loren M. Frank. Constant sub-second cycling between representations of possible futures in the hippocampus. *Cell*, 180(3):552–567.e25, Jan 2020. ISSN 0092-8674. doi:10.1016/j.cell.2020.01.014. URL http://dx.doi.org/10.1016/j.cell.2020.01.014.

Arno Klein, Satrajit S. Ghosh, Forrest S. Bao, Joachim Giard, Yrjö Häme, Eliezer Stavsky, Noah Lee, Brian Rossa, Martin Reuter, Elias Chaibub Neto, and et al. Mindboggling morphometry of human brains. *PLOS Computational Biology*, 13(2):e1005350, Feb 2017. ISSN 1553-7358. doi:10.1371/journal.pcbi.1005350. URL http://dx.doi.org/10.1371/journal.pcbi.1005350.

Peter Kok and Nicholas B. Turk-Browne. Associative prediction of visual shape in the hippocampus. *The Journal of Neuroscience*, 38(31):6888–6899, Jul 2018. ISSN 1529-2401. doi:10.1523/jneurosci.0163-18.2018. URL http://dx.doi.org/10.1523/JNEUROSCI.0163-18.2018.

Peter Kok, Janneke F.M. Jehee, and Floris P. de Lange. Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron*, 75(2):265–270, 2012. ISSN 0896-6273. doi:https://doi.org/10.1016/j.neuron.2012.04.034. URL https://www.sciencedirect.com/science/article/pii/S0896627312004382.

Peter Kok, Michel F. Failing, and Floris P. de Lange. Prior expectations evoke stimulus templates in the primary visual cortex. *Journal of Cognitive Neuroscience*, 26(7):1546–1554, 07 2014. ISSN 0898-929X. doi:10.1162/jocn_a_00562. URL https://doi.org/10.1162/jocn_a_00562.

James Kolasinski, Tamar R. Makin, Saad Jbabdi, Stuart Clare, Charlotte J. Stagg, and Heidi Johansen-Berg. Investigating the stability of fine-grain digit somatotopy in individual human participants. *Journal of Neuroscience*, 36(4):1113–1127, 2016. ISSN 0270-6474. doi:10.1523/JNEUROSCI.1742-15.2016. URL https://www.jneurosci.org/content/36/4/1113.

Lukas Kunz, Lorena Deuker, Hui Zhang, and Nikolai Axmacher. Chapter 26 - tracking human engrams using multivariate analysis techniques. In Denise Manahan-Vaughan, editor, *Handbook of in Vivo Neural Plasticity Techniques*, volume 28 of *Handbook of Behavioral Neuroscience*, chapter 26, pages 481–508. Elsevier, 2018. doi:https://doi.org/10.1016/B978-0-12-812028-6.00026-4. URL https://www.sciencedirect.com/science/article/pii/B9780128120286000264.

Zeb Kurth-Nelson, Marcos Economides, Raymond J. Dolan, and Peter Dayan. Fast sequences of non-spatial state representations in humans. *Neuron*, 91(1):194–204, 2016. ISSN 10974199. doi:10.1016/j.neuron.2016.05.028. URL http://dx.doi.org/10.1016/j.neuron.2016.05.028.

Gregory M. Kurtzer, Vanessa Sochat, and Michael W. Bauer. Singularity: Scientific containers for mobility of compute. *PLoS ONE*, 12(5):e0177459, May 2017. doi:10.1371/journal.pone.0177459. URL http://dx.doi.org/10.1371/journal.pone.0177459.

C. Lanczos. Evaluation of noisy data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, 1(1):76–85, Jan 1964. ISSN 0887-459X. doi:10.1137/0701007. URL http://dx.doi.org/10.1137/0701007.

Steven M. LaValle. *Planning Algorithms*. Cambridge University Press, Cambridge, 2006. ISBN 9780521862059. doi:10.1017/CBO9780511546877. URL https://www.cambridge.org/core/books/planning-algorithms/FC9CC7E67E851E40E3E45D6FE328B768.

Russell Lenth. emmeans: Estimated marginal means, aka least-squares means. 2019. URL https://CRAN.R-project.org/package=emmeans. R package version 1.3.4.

Xiangrui Li, Paul S. Morgan, John Ashburner, Jolinda Smith, and Christopher Rorden. The first step for neuroimaging data analysis: Dicom to nifti conversion. *Journal of Neuroscience Methods*, 264: 47–56, May 2016. ISSN 0165-0270. doi:10.1016/j.jneumeth.2016.03.001. URL http://dx.doi.org/10.1016/j.jneumeth.2016.03.001.

Yunzhe Liu, Raymond J. Dolan, Zeb Kurth-Nelson, and Timothy E.J. Behrens. Human replay spontaneously reorganizes experience. *Cell*, 178(3):640–652, Jul 2019. ISSN 0092-8674. doi:10.1016/j.cell.2019.06.012. URL http://dx.doi.org/10.1016/j.cell.2019.06.012.

Yunzhe Liu, Marcelo G. Mattar, Timothy E. J. Behrens, Nathaniel D. Daw, and Raymond J. Dolan. Experience replay is associated with efficient nonlocal learning. *Science*, 372(6544), 2021. ISSN 0036-8075. doi:10.1126/science.abf1357. URL https://science.sciencemag.org/content/372/6544/eabf1357.

Christopher W. Lynn and Danielle S. Bassett. How humans learn and represent networks. *Proceedings of the National Academy of Sciences*, 117(47):29407–29415, 2020. ISSN 0027-8424. doi:10.1073/pnas.1912328117. URL https://www.pnas.org/content/117/47/29407.

Christopher W. Lynn, Ari E. Kahn, Nathaniel Nyema, and Danielle S. Bassett. Abstract representations of events arise from mental errors in learning and memory. *Nature Communications*, 11(1), May 2020a. ISSN 2041-1723. doi:10.1038/s41467-020-15146-7. URL http://dx.doi.org/10.1038/s41467-020-15146-7.

Christopher W. Lynn, Lia Papadopoulos, Ari E. Kahn, and Danielle S. Bassett. Human information processing in complex networks. *Nature Physics*, Jun 2020b. ISSN 1745-2481. doi:10.1038/s41567-020-0924-7. URL http://dx.doi.org/10.1038/s41567-020-0924-7.

Kevin J Miller and Sarah Jo C Venditto. Multi-step planning in the brain. *Current Opinion in Behavioral Sciences*, 38:29–39, 2021. ISSN 2352-1546. doi:10.1016/j.cobeha.2020.07.003. URL http://www.sciencedirect.com/science/article/pii/S2352154620301054.

Yasushi Miyashita. Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature*, 335(6193):817–820, Oct 1988. ISSN 1476-4687. doi:10.1038/335817a0. URL http://dx.doi.org/10.1038/335817a0.

Ida Momennejad. Learning structures: Predictive representations, replay, and generalization. *Current Opinion in Behavioral Sciences*, 32:155–166, Apr 2020. ISSN 2352-1546. doi:10.1016/j.cobeha.2020.02.017. URL http://dx.doi.org/10.1016/j.cobeha.2020.02.017.

Ida Momennejad and Marc W. Howard. Predicting the future with multi-scale successor representations. *bioRxiv*, 2018. doi:10.1101/449470. URL https://www.biorxiv.org/content/early/2018/10/22/449470.

Ida Momennejad, Evan M. Russek, J. H. Cheong, Matthew M. Botvinick, Nathaniel D. Daw, and Samuel J. Gershman. The successor representation in human reinforcement learning. *Nature Human Behaviour*, 1(9):680–692, Aug 2017. ISSN 2397-3374. doi:10.1038/s41562-017-0180-8. URL http://dx.doi.org/10.1038/s41562-017-0180-8.

Ida Momennejad, A Ross Otto, Nathaniel D Daw, and Kenneth A Norman. Offline replay supports planning in human reinforcement learning. *eLife*, 7:e32548, Dec 2018. doi:10.7554/eLife.32548. URL https://doi.org/10.7554/eLife.32548.

H. Freyja Ólafsdóttir, Daniel Bush, and Caswell Barry. The role of hippocampal replay in memory and planning. *Current Biology*, 28(1):R37–R50, Jan 2018. ISSN 0960-9822. doi:10.1016/j.cub.2017.10.073. URL http://dx.doi.org/10.1016/j.cub.2017.10.073.

Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Edouard Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Jonathan Peirce, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. Psychopy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1):195–203, Feb 2019. ISSN 1554-3528. doi:10.3758/s13428-018-01193-y. URL http://dx.doi.org/10.3758/s13428-018-01193-y.

Jonathan W. Peirce. PsychoPy—psychophysics software in python. *Journal of Neuroscience Methods*, 162(1-2):8–13, may 2007. doi:10.1016/j.jneumeth.2006.11.017. URL https://doi.org/10.1016%2Fj.jneumeth.2006.11.017.

Jonathan W Peirce. Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2, 2008. doi:10.3389/neuro.11.010.2008. URL https://doi.org/10.3389%2Fneuro.11.010.2008.

Brad E. Pfeiffer and David J. Foster. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, 497(7447):74–79, Apr 2013. ISSN 1476-4687. doi:10.1038/nature12112. URL http://dx.doi.org/10.1038/nature12112.

Russell A. Poldrack. Region of interest analysis for fMRI. *Social Cognitive and Affective Neuroscience*, 2(1):67–70, Mar 2007. ISSN 1749-5024. doi:10.1093/scan/nsm006. URL http://dx.doi.org/10.1093/scan/nsm006.

Michael J. D. Powell. Developments of newuoa for unconstrained minimization without derivatives. *Department of Applied Mathematics and Theoretical Physics*, 2007.

Michael J. D. Powell. The bobyqa algorithm for bound constrained optimization without derivatives. *Department of Applied Mathematics and Theoretical Physics*, pages 26–46, 2009.

Jonathan D. Power, Anish Mitra, Timothy O. Laumann, Abraham Z. Snyder, Bradley L. Schlaggar, and Steven E. Petersen. Methods to detect, characterize, and remove motion artifact in resting state fmri. *NeuroImage*, 84:320–341, Jan 2014. ISSN 1053-8119. doi:10.1016/j.neuroimage.2013.08.048. URL http://dx.doi.org/10.1016/j.neuroimage.2013.08.048.

R Core Team. R: A language and environment for statistical computing, 2019. URL https://www.R-project.org/.

Arthur S. Reber. Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118(3):219–235, 1989. ISSN 0096-3445. doi:10.1037/0096-3445.118.3.219. URL http://dx.doi.org/10.1037/0096-3445.118.3.219.

Martin Reuter, H. Diana Rosas, and Bruce Fischl. Highly accurate inverse consistent registration: A robust approach. *NeuroImage*, 53(4):1181–1196, Dec 2010. ISSN 1053-8119. doi:10.1016/j.neuroimage.2010.07.020. URL http://dx.doi.org/10.1016/j.neuroimage.2010.07.020.

Bruno Rossion and Gilles Pourtois. Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, 33(2):217–236, Feb 2004. ISSN 1468-4233. doi:10.1068/p5117. URL http://dx.doi.org/10.1068/p5117.

Evan M. Russek, Ida Momennejad, Matthew M. Botvinick, Samuel J. Gershman, and Nathaniel D. Daw. Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS Computational Biology*, 13(9):e1005768, Sep 2017. ISSN 1553-7358. doi:10.1371/journal.pcbi.1005768. URL http://dx.doi.org/10.1371/journal.pcbi.1005768.

Evan M. Russek, Ida Momennejad, Matthew M. Botvinick, Samuel J. Gershman, and Nathaniel D. Daw. Neural evidence for the successor representation in choice evaluation. *bioRxiv*, 2021. doi:10.1101/2021.08.29.458114. URL https://www.biorxiv.org/content/early/2021/08/31/2021.08.29.458114.

J. R. Saffran, R. N. Aslin, and E. L. Newport. Statistical learning by 8-month-old infants. *Science*, 274(5294):1926–1928, Dec 1996. ISSN 1095-9203. doi:10.1126/science.274.5294.1926. URL http://dx.doi.org/10.1126/science.274.5294.1926.

Theodore D. Satterthwaite, Mark A. Elliott, Raphael T. Gerraty, Kosha Ruparel, James Loughead, Monica E. Calkins, Simon B. Eickhoff, Hakon Hakonarson, Ruben C. Gur, Raquel E. Gur, and Daniel H. Wolf. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage*, 64:240–256, 2013. ISSN 1053-8119. doi:https://doi.org/10.1016/j.neuroimage.2012.08.052. URL https://www.sciencedirect.com/science/article/pii/S1053811912008609.

Anna C. Schapiro and N. Turk-Browne. Statistical learning. In Arthur W. Toga, editor, *Brain Mapping*, volume 3, pages 501–506. Elsevier, 2015. ISBN 9780123973160. doi:10.1016/b978-0-12-397025-1.00276-1. URL http://dx.doi.org/10.1016/B978-0-12-397025-1.00276-1.

Anna C. Schapiro, Lauren V. Kustner, and Nicholas B. Turk-Browne. Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current Biology*, 22(17):1622–1627, Sep 2012. ISSN 0960-9822. doi:10.1016/j.cub.2012.06.056. URL http://dx.doi.org/10.1016/j.cub.2012.06.056.

Anna C Schapiro, Timothy T Rogers, Natalia I Cordova, Nicholas B Turk-Browne, and Matthew M Botvinick. Neural representations of events arise from temporal community structure. *Nature Neuroscience*, 16(4):486–492, Feb 2013. ISSN 1546-1726. doi:10.1038/nn.3331. URL http://dx.doi.org/10.1038/nn.3331.

Nicolas W Schuck and Yael Niv. Sequential replay of nonspatial task states in the human hippocampus. *Science*, 364(6447):eaaw5181, 2019. doi:10.1126/science.aaw5181.

Nicolas W Schuck, Robert Gaschler, and Peter A Frensch. Implicit learning of what comes when and where within a sequence: The time-course of acquiring serial position-item and item-item associations to represent serial order. *Advances in cognitive psychology*, 8(2):83–97, 2012a. doi:10.2478/v10053-008-0106-0. URL https://pubmed.ncbi.nlm.nih.gov/22679464.

Nicolas W. Schuck, Robert Gaschler, Aysha Keisler, and Peter A. Frensch. Position–item associations play a role in the acquisition of order knowledge in an implicit serial reaction time task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(2):440–456, 2012b. doi:10.1037/a0025816. URL https://doi.org/10.1037%2Fa0025816.

Nicolas W. Schuck, Ming Bo Cai, Robert C. Wilson, and Yael Niv. Human orbitofrontal cortex represents a cognitive map of state space. *Neuron*, 91(6):1402 – 1412, 2016. ISSN 0896-6273. doi:10.1016/j.neuron.2016.08.019. URL http://www.sciencedirect.com/science/article/pii/S0896627316305116.

Carol Augart Seger. Implicit learning. *Psychological Bulletin*, 115(2):163–196, 1994. ISSN 0033-2909. doi:10.1037/0033-2909.115.2.163. URL http://dx.doi.org/10.1037/0033-2909.115.2.163.

C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948. doi:10.1002/j.1538-7305.1948.tb01338.x.

Brynn E Sherman, Kathryn N Graves, and Nicholas B Turk-Browne. The prevalence and importance of statistical learning in human cognition and behavior. *Current Opinion in Behavioral Sciences*, 32:15–20, Apr 2020. ISSN 2352-1546. doi:10.1016/j.cobeha.2020.01.015. URL http://dx.doi.org/10.1016/j.cobeha.2020.01.015.

Stephen M. Smith and J. Michael Brady. SUSAN - a new approach to low level image processing. *International Journal of Computer Vision*, 23(1):45–78, May 1997. ISSN 0920-5691. doi:10.1023/a:1007963824710. URL http://dx.doi.org/10.1023/A:1007963824710.

Joan G Snodgrass and Mary Vanderwart. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human learning and memory*, 6(2):174—215, 1980. doi:10.1037/0278-7393.6.2.174. URL https://doi.org/10.1037/0278-7393.6.2.174.

Vanessa V. Sochat, Cameron J. Prybol, and Gregory M. Kurtzer. Enhancing reproducibility in scientific computing: Metrics and registry for singularity containers. *PLoS ONE*, 12(11):e0188511,

Nov 2017. doi:10.1371/journal.pone.0188511. URL http://dx.doi.org/10.1371/journal.pone.0188511.

Alec Solway, Carlos Diuk, Natalia Córdova, Debbie Yee, Andrew G. Barto, Yael Niv, and Matthew M. Botvinick. Optimal behavioral hierarchy. *PLOS Computational Biology*, 10(8):1–10, 08 2014. doi:10.1371/journal.pcbi.1003779. URL https://doi.org/10.1371/journal.pcbi.1003779.

Kimberly L Stachenfeld, Matthew M Botvinick, and Samuel J Gershman. The hippocampus as a predictive map. *Nature Neuroscience*, 20(11):1643–1653, Oct 2017. ISSN 1546-1726. doi:10.1038/nn.4650. URL http://dx.doi.org/10.1038/nn.4650.

Bryan A. Strange, Andrew Duggins, William Penny, Raymond J. Dolan, and Karl J. Friston. Information theory, novelty and hippocampal responses: unpredicted or unpredictable? *Neural Networks*, 18(3):225–230, 2005. ISSN 0893-6080. doi:https://doi.org/10.1016/j.neunet.2004.12.004. URL https://www.sciencedirect.com/science/article/pii/S0893608005000067.

Richard S. Sutton. Dyna, an integrated architecture for learning, planning, and reacting. *ACM SIGART Bulletin*, 2(4):160–163, Jul 1991. ISSN 0163-5719. doi:10.1145/122344.122377. URL http://dx.doi.org/10.1145/122344.122377.

Arielle Tambini and Lila Davachi. Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. *Proceedings of the National Academy of Sciences*, 110(48):19591–19596, Nov 2013. ISSN 1091-6490. doi:10.1073/pnas.1308499110. URL http://dx.doi.org/10.1073/pnas.1308499110.

Arielle Tambini and Lila Davachi. Awake reactivation of prior experiences consolidates memories and biases cognition. *Trends in Cognitive Sciences*, 23(10):876–890, Oct 2019. ISSN 1364-6613. doi:10.1016/j.tics.2019.07.008. URL http://dx.doi.org/10.1016/j.tics.2019.07.008.

Edward C. Tolman. Cognitive maps in rats and men. *Psychological Review*, 55(4):189–208, 1948. ISSN 0033-295X. doi:10.1037/h0061626. URL http://dx.doi.org/10.1037/h0061626.

John W. Tukey. Comparing individual means in the analysis of variance. *Biometrics*, 5(2):99–114, Jun 1949. ISSN 0006-341X. doi:10.2307/3001913. URL http://dx.doi.org/10.2307/3001913.

Nicholas B. Turk-Browne, Justin A. Jungé, and Brian J. Scholl. The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, 134(4):552–564, 2005. ISSN 0096-3445. doi:10.1037/0096-3445.134.4.552. URL http://dx.doi.org/10.1037/0096-3445.134.4.552.

Nicholas J Tustison, Brian B Avants, Philip A Cook, Yuanjie Zheng, Alexander Egan, Paul A Yushkevich, and James C Gee. N4itk: Improved n3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6):1310–1320, Jun 2010. ISSN 1558-254X. doi:10.1109/tmi.2010.2046908. URL http://dx.doi.org/10.1109/TMI.2010.2046908.

Matthijs A A. van der Meer and Aaron David Redish. Covert expectation-of-reward in rat ventral striatum at decision points. *Frontiers in Integrative Neuroscience*, 3, 2009. ISSN 1662-5145. doi:10.3389/neuro.07.001.2009. URL http://dx.doi.org/10.3389/neuro.07.001.2009.

Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009. ISBN 1441412697.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, and et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature Methods*, Feb 2020. ISSN 1548-7105. doi:10.1038/s41592-019-0686-2. URL http://dx.doi.org/10.1038/s41592-019-0686-2.

Matteo Visconti di Oleggio Castello, James E. Dobson, Terry Sackett, Chandana Kodiweera, James V. Haxby, Mathias Goncalves, Satrajit Ghosh, and Yaroslav O. Halchenko. Repronim/reproin 0.6.0, 2020. URL https://zenodo.org/record/3625000.

Adina S. Wagner, Laura K. Waite, Kyle Meyer, Marisa K. Heckner, Tobias Kadelka, Niels Reuter, Alexander Q. Waite, Benjamin Poldrack, Christopher J. Markiewicz, Yaroslav O. Halchenko, Peter Vavra, Pattarawat Chormai, Jean-Baptiste Poline, Lya K. Paas, Peer Herholz, Lisa N. Mochalski, Nevena Kraljevic, Lisa Wiersch, Alexandre Hutton, Dorian Pustina, Hamzah Hamid Baagil, Tristan Glatard, Sarah Oliveira, Giulia Ippoliti, Christian Mönch, Dorien Huijser, and Michael Hanke. *The DataLad Handbook*. Zenodo, 2020. doi:10.5281/ZENODO.3905791. URL https://zenodo.org/record/3905791.

Mengni Wang, David J. Foster, and Brad E. Pfeiffer. Alternating sequences of future and past behavior encoded within hippocampal theta oscillations. *Science*, 370(6513):247–250, 2020. doi:10.1126/science.abb4151. URL https://science.sciencemag.org/content/370/6513/247.

Nikolaus Weiskopf, Chloe Hutton, Oliver Josephs, and Ralf Deichmann. Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: A whole-brain analysis at 3 T and 1.5 T. *NeuroImage*, 33(2):493–504, Nov 2006. ISSN 1053-8119. doi:10.1016/j.neuroimage.2006.07.029. URL http://dx.doi.org/10.1016/j.neuroimage.2006.07.029.

Andrew M Wikenheiser and Aaron David Redish. Decoding the cognitive map: ensemble hippocampal sequences and decision making. *Current Opinion in Neurobiology*, 32:8–15, Jun 2015a. ISSN 0959-4388. doi:10.1016/j.conb.2014.10.002. URL http://dx.doi.org/10.1016/j.conb.2014.10.002.

Andrew M. Wikenheiser and Aaron David Redish. Hippocampal theta sequences reflect current goals. *Nature Neuroscience*, 18(2):289–294, 2015b. doi:10.1038/nn.3909. URL https://doi.org/10.1038/nn.3909.

Robert C. Wilson, Yuji K. Takahashi, Geoffrey Schoenbaum, and Yael Niv. Orbitofrontal cortex as a cognitive map of task space. *Neuron*, 81(2):267–279, 2014. ISSN 0896-6273. doi:10.1016/j.neuron.2013.11.005. URL http://www.sciencedirect.com/science/article/pii/S0896627313010398.

Lennart Wittkuhn and Nicolas W. Schuck. Dynamics of fMRI patterns reflect sub-second activation sequences and reveal replay in human visual cortex. *Nature Communications*, 12(1795), 2021. doi:10.1038/s41467-021-21970-2. URL https://doi.org/10.1038/s41467-021-21970-2.

Lennart Wittkuhn, Samson Chien, Sam Hall-McMaster, and Nicolas W. Schuck. Replay in minds and machines. *Neuroscience & Biobehavioral Reviews*, 129:367–388, 2021. ISSN 0149-7634. doi:10.1016/j.neubiorev.2021.08.002. URL https://www.sciencedirect.com/science/article/pii/S0149763421003444.

Shengjin Xu, Wanchen Jiang, Mu-ming Poo, and Yang Dan. Activity recall in a visual cortical ensemble. *Nature Neuroscience*, 15(3):449–455, Jan 2012. ISSN 1546-1726. doi:10.1038/nn.3036. URL http://dx.doi.org/10.1038/nn.3036.

Tal Yarkoni, Christopher Markiewicz, Alejandro de la Vega, Krzysztof Gorgolewski, Taylor Salo, Yaroslav Halchenko, Quinten McNamara, Krista DeStasio, Jean-Baptiste Poline, Dmitry Petrov, Valérie Hayot-Sasson, Dylan Nielson, Johan Carlin, Gregory Kiar, Kirstie Whitaker, Elizabeth DuPre, Adina Wagner, Lee Tirrell, Mainak Jas, Michael Hanke, Russell Poldrack, Oscar Esteban, Stefan Appelhoff, Chris Holdgraf, Isla Staden, Bertrand Thirion, Dave Kleinschmidt, John Lee, Matteo di Castello, Michael Notter, and Ross Blair. PyBIDS: Python tools for BIDS datasets. *Journal of Open Source Software*, 4(40):1294, aug 2019a. doi:10.21105/joss.01294. URL https://doi.org/10.21105%2Fjoss.01294.

Tal Yarkoni, Christopher J. Markiewicz, Alejandro de la Vega, Krzysztof J. Gorgolewski, Yaroslav O. Halchenko, Taylor Salo, Quinten McNamara, Krista DeStasio, Jean-Baptiste Poline, Dmitry Petrov, Valérie Hayot-Sasson, Dylan M. Nielson, Johan Carlin, Gregory Kiar, Kirstie Whitaker, Adina Wagner, Elizabeth DuPre, Stefan Appelhoff, Alexander Ivanov, Johannes Wennberg, Lee S. Tirrell, Oscar Esteban, Mainak Jas, Michael Hanke, Russell Poldrack, Chris Holdgraf, Isla Staden, Ariel Rokem, Bertrand Thirion, Chadwick Boulay, Dave F. Kleinschmidt, Erin W Dickie, John A. Lee, Matteo Visconti di Oleggio Castello, Michael Philipp Notter, Pauline Roca, and Ross Blair. bids-standard/pybids: 0.9.3, 2019b. URL https://zenodo.org/record/3363985.

Linda Q. Yu, Robert C. Wilson, and Matthew R. Nassar. Adaptive learning is structure learning in time. *Neuroscience & Biobehavioral Reviews*, 128:270–281, 2021. ISSN 0149-7634. doi:https://doi.org/10.1016/j.neubiorev.2021.06.024. URL https://www.sciencedirect.com/science/article/pii/S0149763421002657.

Y. Zhang, M. Brady, and S. Smith. Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1):45–57, 2001. ISSN 0278-0062. doi:10.1109/42.906424. URL http://dx.doi.org/10.1109/42.906424.

Glossary

AC-PC anterior commissure - posterior commissure.

AIC Akaike information criterion.

ANOVA analysis of variance.

ANTs Advanced Normalization Tools.

A-P anterior-to-posterior.

BIDS brain imaging data structure.

BOBYQA Bound Optimization BY Quadratic Approximation.

BOLD blood-oxygen-level dependent.

CSF cerebrospinal fluid.

DGPs German Psychological Society.

DICOM Digital Imaging and Communications in Medicine.

EPI echo-planar imaging.

FA flip angle.

FD framewise displacement.

FDR false discovery rate.

fMRI functional magnetic resonance imaging.

FOV field of view.

FSL FMRIB Software Library.

FWHM Full Width at Half Maximum.

GM gray-matter.

GR gradient recalled.

HMM hidden markov model.

HPC high performance computing.

HRF The hemodynamic response function (HRF) characterizes an fMRI response that results from a brief, spatially localized pulse of neuronal activity.

HSD honest significant difference.

INU intensity non-uniformity.

IQR interquartile range.

ITI inter-trial interval.

LME linear mixed effects.

LTS long-term support.

MB multi-band.

MEG magnetoencephalography.

min minute.

MPRAGE Magnetization Prepared Rapid Gradient Echo.

MRI magnetic resonance imaging.

ms millisecond.

MTL medial temporal lobe.

PFC prefrontal cortex.

PMU Physiological Measurement Unit.

ROI region of interest.

s second.

SEM standard error of the mean.

SI supplementary information.

SR successor representation.

S-R stimulus-response.

SRI stimulus-response interval.

SUSAN Smallest Univalue Segment Assimilating Nucleus.

T1w T1-weighted.

TD temporal difference.

TE echo time.

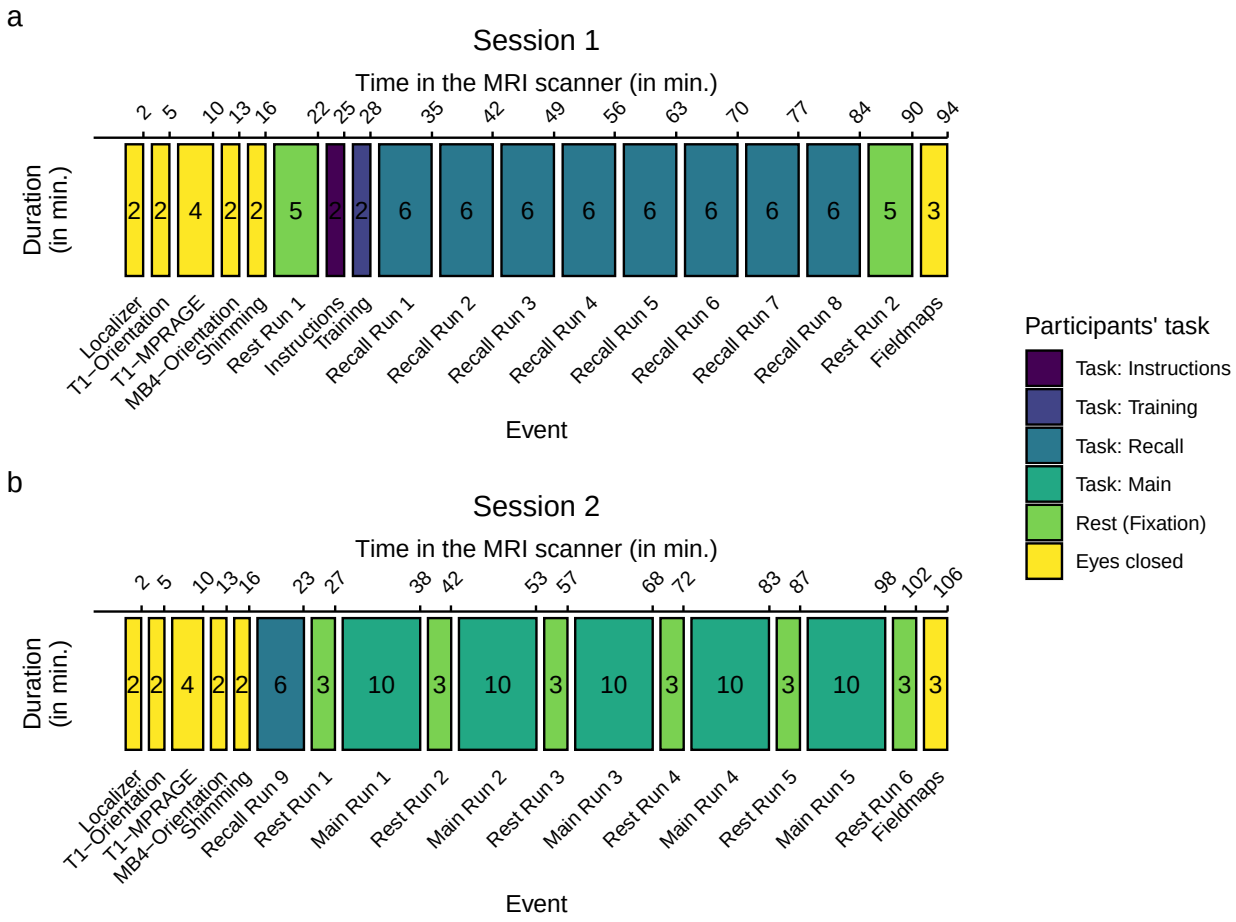
TI inversion time.

TR repetition time.

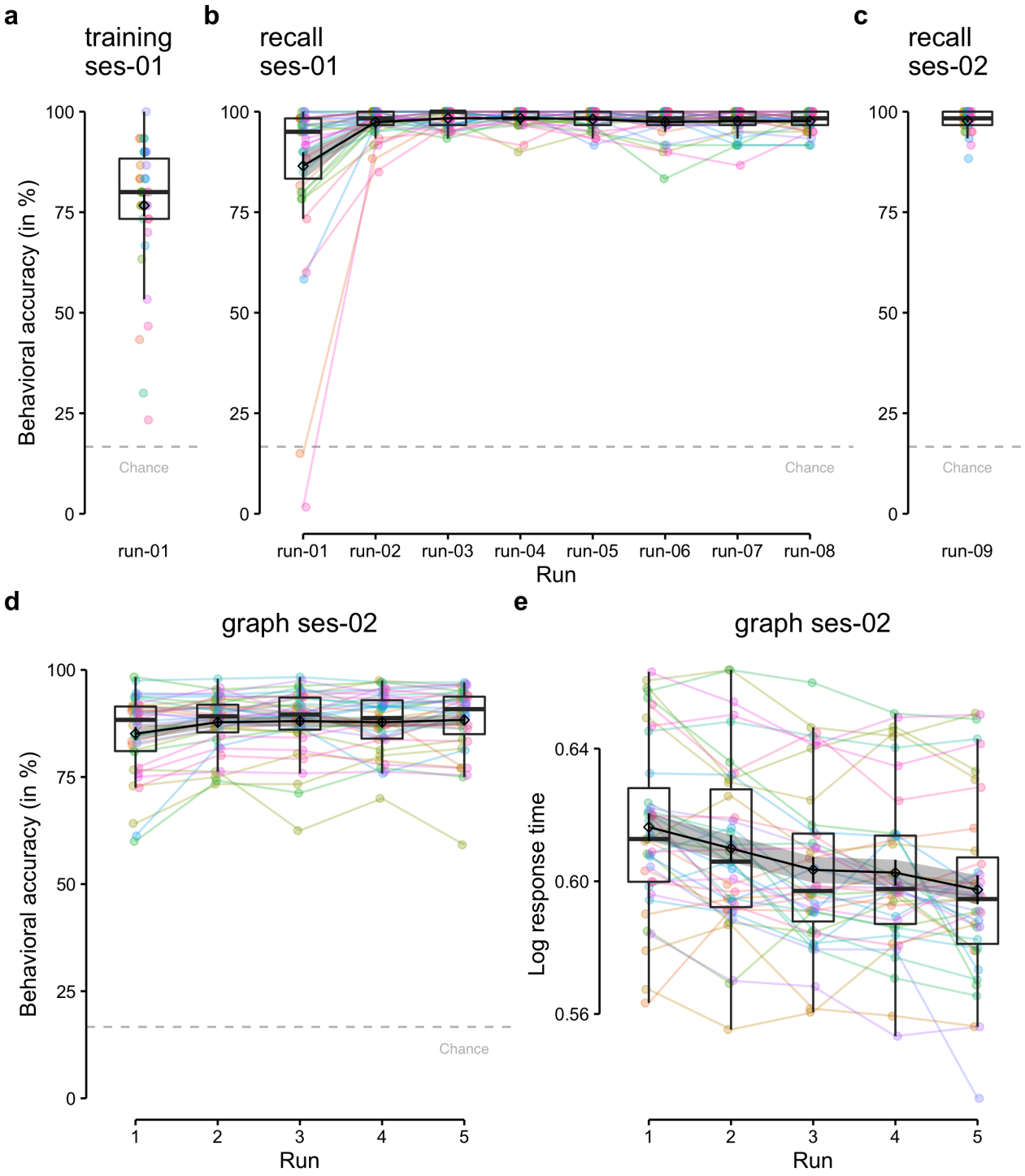
WM white-matter.

Supplementary Information

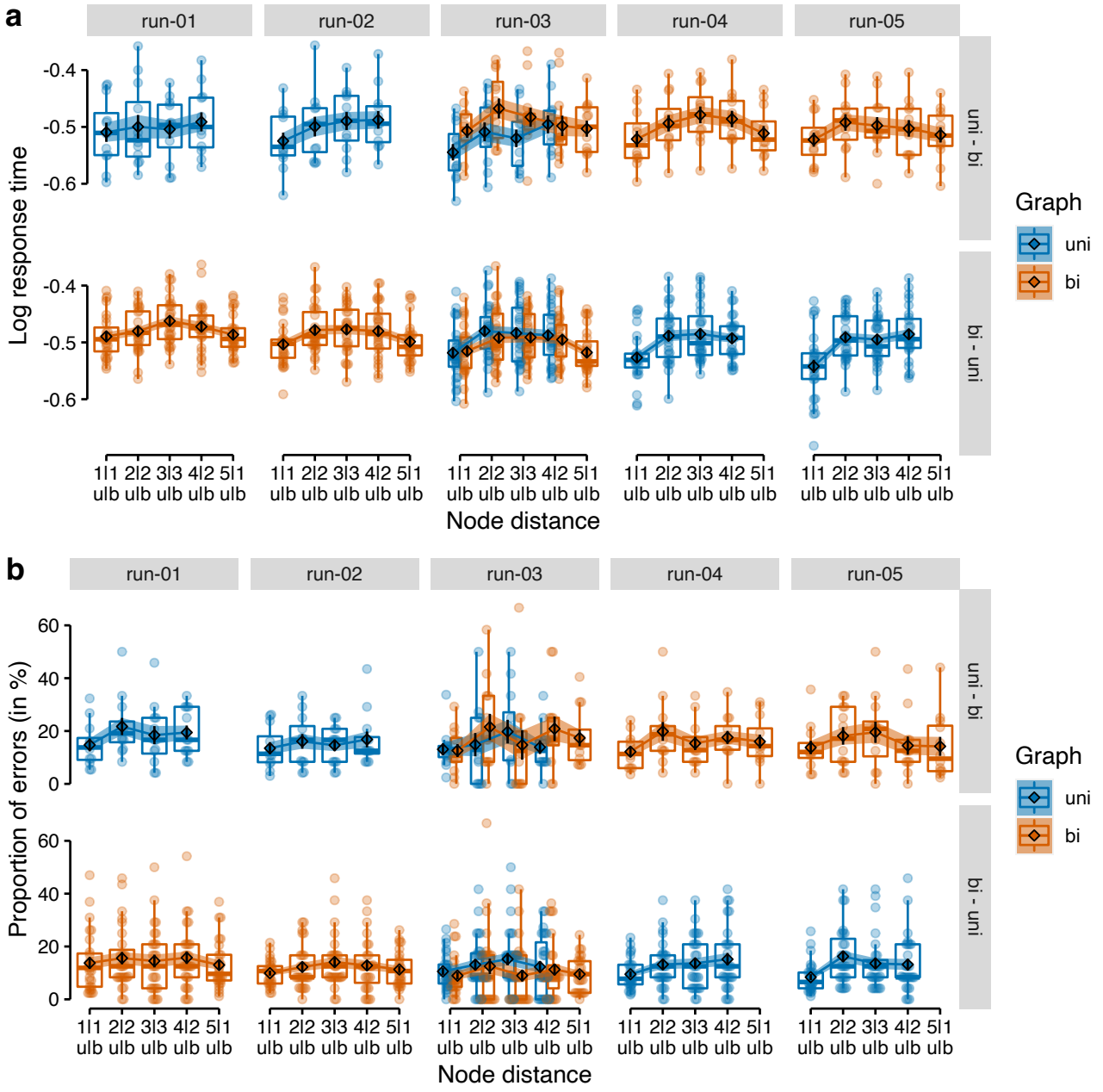
Supplementary Figures



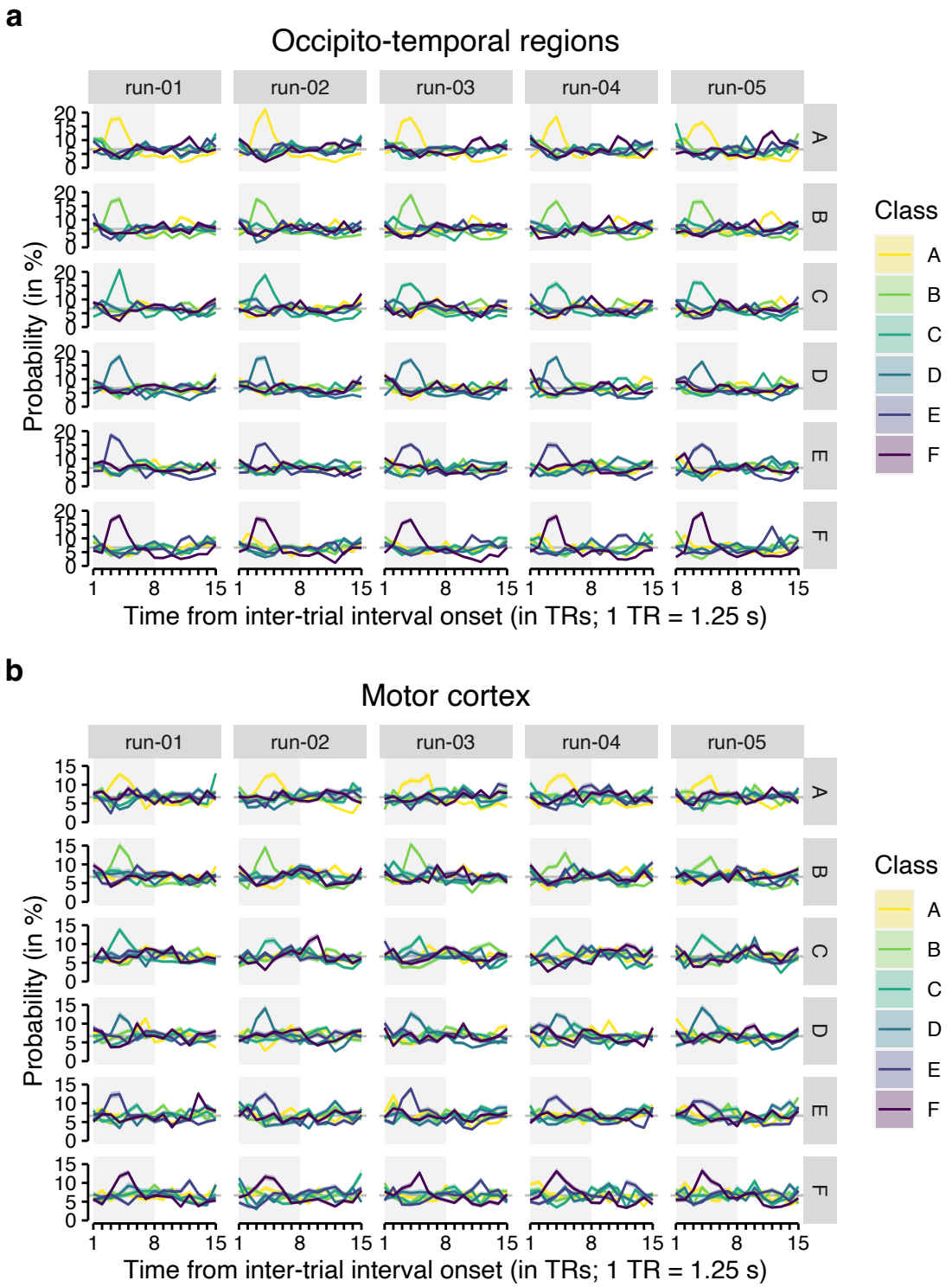
Supplementary Figure S1: Study procedure. (a) Session 1 started with a 5 min resting-state scan before participants read the task instructions and completed the training condition of the task. Participants then completed eight runs of the recall condition of ca. 6 min each before another 5 min resting-state scan was recorded. **(b)** Session 2 started with another run of the recall condition of ca. 6 min. Participants then completed all five runs of the graph learning task of about 10 min each which were interleaved with six resting-state scans of 3 min each. Both experimental sessions started with a short localizer scan and a T1w anatomical scan and ended with the acquisition of fieldmaps. During these scans and additional preparations by the study staff (e.g., orientation of the FOV) participants were asked to keep their eyes closed. Numbers inside the rectangles indicate approximate duration of each step in minutes (mins). Colors indicate participants' task (see legend).



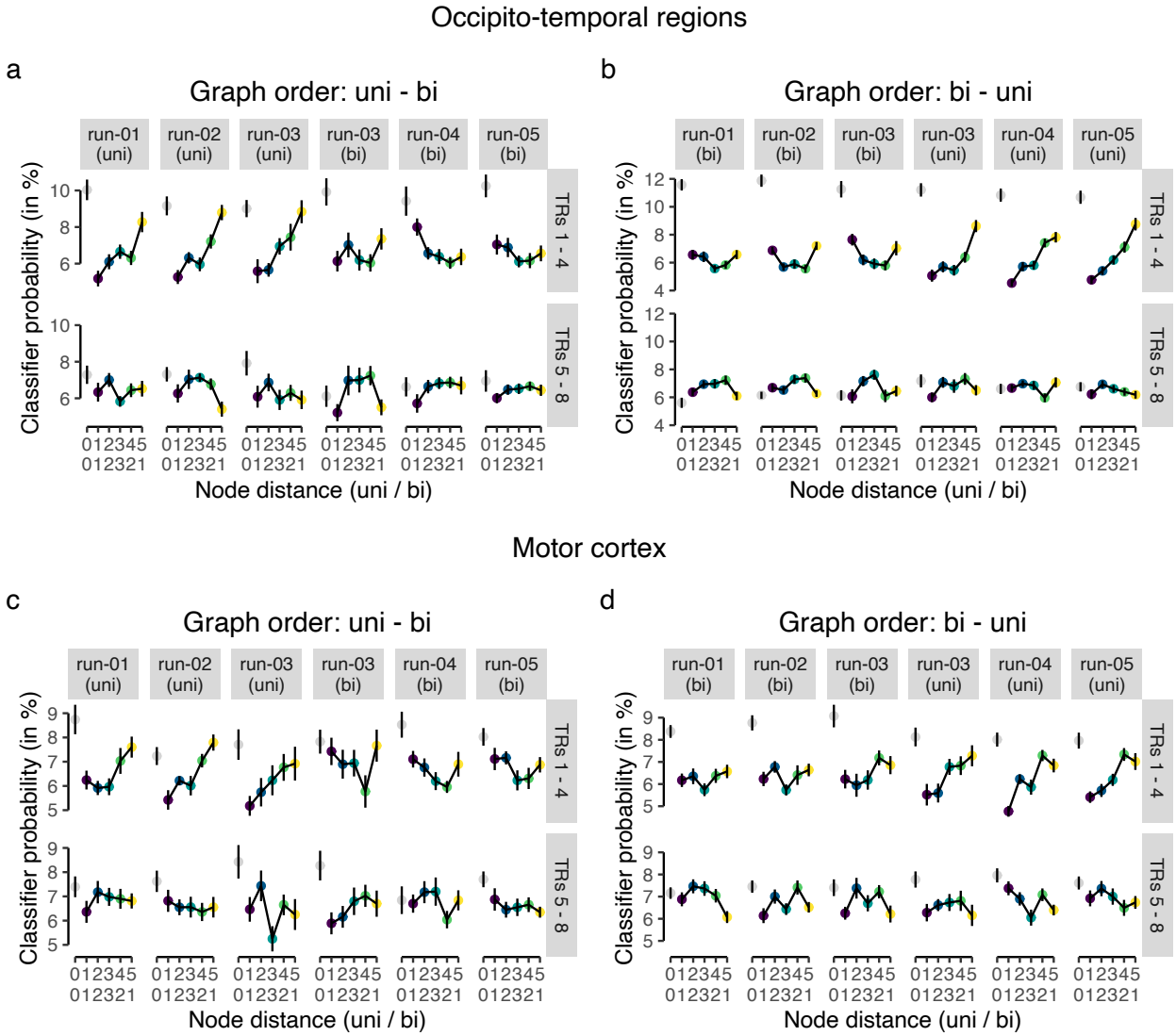
Supplementary Figure S2: Behavioral accuracy and response times per task run in training, recall, and graph trials. Mean behavioral accuracy (in %; y-axis) per task run of the study (x-axis) in **(a)** training trials, **(b)** recall trials in session 1, **(c)** recall trials in session 2, and **(d)** graph trials in session 2. **(e)** Mean log response time (y-axis) per task run of the study (x-axis) in graph trials. The chance-level (gray dashed line) is at 16.67%. Each dot corresponds to averaged data from one participant. Colored lines connect data across runs for each participant. Boxplots indicate the median and IQR. The lower and upper hinges correspond to the first and third quartiles (the 25^{th} and 75^{th} percentiles). The upper whisker extends from the hinge to the largest value no further than $1.5*$ IQR from the hinge (where IQR is the interquartile range (IQR), or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most $1.5*$ IQR of the hinge. The diamond shapes show the sample mean. Error bars and shaded areas indicate ± 1 SEM. All statistics have been derived from data of $n = 39$ human participants who participated in one experiment.



Supplementary Figure S3: Behavioral responses across task runs. (a) Log response times (y-axis) as a function of node distance (x-axis) in the graph structure (colors) for each task run (vertical panels) and graph order (uni – bi vs. bi – uni; horizontal panels). **(b)** Proportion of errors (in %; y-axis; relative to the total number of trials per node distance and run) as a function of node distance (x-axis) in the graph structure (colors) for each task run (vertical panels) and graph order (uni – bi vs. bi – uni; horizontal panels). Boxplots indicate the median and IQR. The lower and upper hinges correspond to the first and third quartiles (the 25^{th} and 75^{th} percentiles). The upper whisker extends from the hinge to the largest value no further than 1.5∗ IQR from the hinge (where IQR is the interquartile range (IQR), or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5∗ IQR of the hinge. The diamond shapes show the sample mean. Each dot corresponds to averaged data from one participant. Error bars and shaded areas represent ±1 SEM. All statistics have been derived from data of $n = 39$ human participants who participated in one experiment.

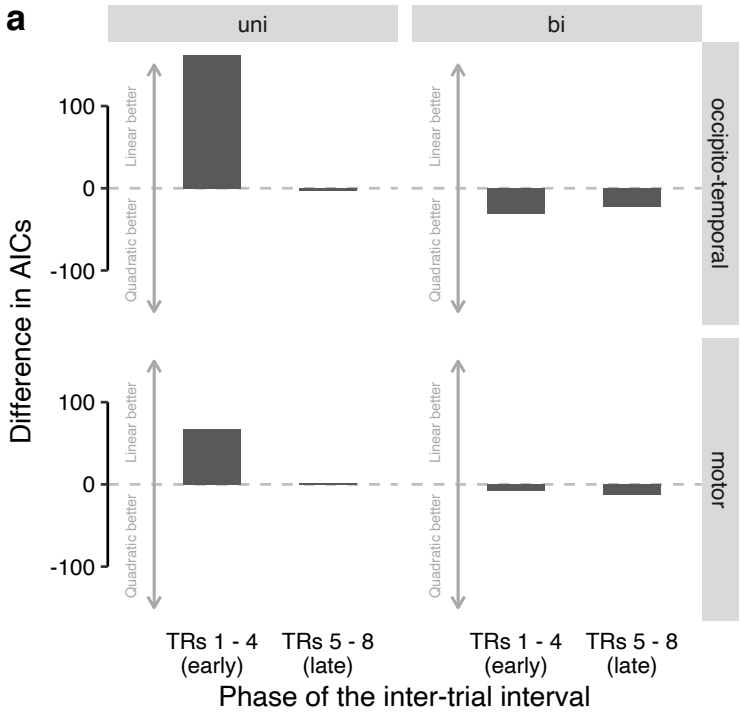


Supplementary Figure S4: Classifier probabilities in long ITIs of graph trials. Time courses (in TRs from the onset of the ITIs; x-axis) of classifier probabilities (in %; y-axis) per class (colors; see legend) and run (vertical panels). Substantial delayed and extended increases in classifier probability were found for the class that occurred on a given trial (horizontal panels) in both occipito-temporal brain regions (a) and motor and somatosensory cortex (b), peaking around the fourth TR following ITI onset, as expected given that classifier were trained on the fourth TR from event onset in fMRI data from recall trials. Each line represented averaged data across all trials of all participants. All shaded areas represent ±1 SEM. Gray rectangles indicate the long ITI (TRs 1–8). All statistics have been derived from data of $n = 39$ human participants who participated in one experiment.



Supplementary Figure S5: Classifier probabilities during graph trials are modulated by node distance in the graph structure. Classifier probabilities (in %; y-axis) as a function of the distance between the nodes in the uni-directional (first line) and bi-directional (second line) graph structure averaged across TRs in the early (TRs 1–4) or late (TRs 5–8) phase (horizontal panels) of the long ITIs of the five runs (vertical panels) in graph trials for data in the occipito-temporal **(a)**, **(b)** and motor cortex **(c)**, **(d)** ROIs. Each dot corresponds to data averaged across participants. Error bars represent ±1 SEM. All statistics have been derived from data of $n = 39$ human participants who participated in one experiment.

a



b

ROI	Graph	Phase	Linear	Quadratic	Diff
occipito-temporal	uni	TRs 1 - 4 (early)	605.78	767.81	162.03
occipito-temporal	uni	TRs 5 - 8 (late)	587.32	583.68	-3.64
occipito-temporal	bi	TRs 1 - 4 (early)	622.04	591.02	-31.02
occipito-temporal	bi	TRs 5 - 8 (late)	598.73	576.13	-22.60
motor	uni	TRs 1 - 4 (early)	593.65	661.04	67.39
motor	uni	TRs 5 - 8 (late)	628.16	629.94	1.78
motor	bi	TRs 1 - 4 (early)	632.90	625.50	-7.40
motor	bi	TRs 5 - 8 (late)	604.99	592.24	-12.75

Supplementary Figure S6: Model comparison of LME models with linear vs. quadratic predictor of classifier probabilities in ITIs of graph trials. (a) Difference in AIC values for LME models including a linear vs. a quadratic predictor for mean classifier probabilities for the two TR phases (early vs. later), the two graph conditions (`uni` vs. `bi`; vertical panels) and the two ROIs (`occipito-temporal` vs. `motor`; horizontal panels). Positive values indicate a better fit of the LME model with the linear predictor and negative values indicate a better fit of the LME model with the quadratic predictor. **(b)** Table of AIC values of LME models with linear and quadratic predictor (and their difference) for all combinations of ROI, graph condition, TR phase. All statistics have been derived from data of $n = 39$ human participants who participated in one experiment with two sessions.

Task instructions in English

Box S1: Screen 1 of instructions for the training condition in session 1

Welcome to the study - Session 1!
Please read the following information carefully. If you have any questions, you can clarify them right away with the study instructor. Please lie as still and relaxed as possible for the entire time.
Press any key to continue.

Box S2: Screen 1 of instructions for the training condition in session 1

Your task:
You are a zookeeper in training and have to make sure that all animals are in the right cages. First you will learn in a training which animal belongs in which cage. We will now explain to you exactly how this task works.
Press any key to continue.

Box S3: Screen 3 of instructions for the training condition in session 1

Training (Part 1)
You want to become a zookeeper and start your training today. First you will learn which animal belongs in which cage. You will see six cages at the bottom of the screen. Each of the six cages belongs to one of six animals. You will select a cage with the appropriate response key. Please keep your ring, middle and index fingers on the response keys the entire time so that you can answer as quickly and accurately as possible.
Press any key to continue.

Box S4: Screen 4 of instructions for the training condition in session 1

During the training, the animals appear above their cages. Press the key for that cage as fast as you can and remember the cage where the animal belongs. Please press the correct button within 1 second. Please answer as quickly and accurately as possible. You will receive feedback if your answer was correct, incorrect or too slow. The correct cage will appear in green and the incorrect cage will appear in red.
Press any key to continue.

Box S5: Screen 5 of instructions for the training condition in session 1

It is very important that you actively remember which animal belongs in which cage. You will get a higher bonus if you remember the correct assignment. The better you remember which animal belongs in which cage, the more money you earn! You will now complete one pass of this task, which will take approximately 2 minutes.

Press any key to continue.

Box S6: Screen 1 of instructions for the recall condition in session 1

Training (part 2)

We will now check how well you have learned the assignment of the animals to their cages. The animals will now appear in the center of the screen. You are asked to remember the correct cage for each animal, and then press the correct key as quickly as possible.

Press any key to continue.

Box S7: Screen 2 of instructions for the recall condition in session 1

This time you respond only after the animal is shown. In each round, the animal will appear first in the center of the screen. Then please try to actively imagine the correct combination of animal, cage and response key. After that, a small cross will appear for a short moment. Then the cages appear and you can respond as quickly and accurately as possible. Please respond as soon as the cages appear, not earlier.

Press any key to continue.

Box S8: Screen 3 of instructions for the recall condition in session 1

You have again 1 second to respond. Please respond again as fast and accurate as possible. You will get feedback again if your response was wrong or too slow. If your response was correct, you will continue directly with the next round without feedback. You will now complete 8 passes of this task, each taking about 6 minutes. In between the rounds you will be given the opportunity to take a break.

Press any key to continue.

Box S9: Screen 1 of instructions for the recall condition in session 2

Welcome to the study - Session 2!

We will check again if you can remember the assignment of the animals to their cages. The animals will appear in the center of the screen again. You are asked to remember again the correct cage for each animal and press the correct key as quickly as possible.

Press any key to continue.

Box S10: Screen 2 of instructions for the recall condition in session 2

You answer again only after the animal has been shown. In each round, the animal appears first in the center of the screen. Then please try to actively imagine the correct combination of animal, cage and answer key. After that, a small cross will first appear for a short moment. Then the cages appear and you can answer as quickly and accurately as possible. Please respond as soon as the cages appear, not earlier.

Press any key to continue.

Box S11: Screen 3 of instructions for the recall condition in session 2

You have again 1 second to respond. Please respond again as fast and accurate as possible. You will get feedback again if your response was wrong or too slow. If your answer was correct, you will proceed directly to the next round without feedback. You will now complete a run-through of this task, which will again take approximately 6 minutes. After the round you will be given the opportunity to take a break. Press any key to continue.

Box S12: Screen 1 of instructions for the graph condition in session 2

You have finished the passage to memory! Well done! You are now welcome to take a short break and also close your eyes. Please continue to lie still and relaxed. When you are ready, you can continue with the instructions for the main task.

Press any key to continue.

Box S13: Screen 2 of instructions for the graph condition in session 2

Main task

Congratulations, you are now a trained zookeeper! Attention: Sometimes the animals break out of their cages! Your task is to bring the animals back to the right cages. When you see an animal on the screen, press the right button as fast as possible to bring the animal back to the right cage. This time you will not get any feedback if your answer was right or wrong. The more animals you put in the correct cages, the more bonus you get at the end of the trial! The main task consists of 5 runs, each taking about 10 minutes to complete.

Press any key to continue.

Box S14: Screen 3 of instructions for the graph condition in session 2

You have again 1 second to respond. In the main task, you again respond immediately when you see an animal on the screen. Again, please respond as quickly and accurately as possible. Between each round you will again see a cross for a moment. Sometimes the cross will be shown a little shorter and sometimes a little longer. It is best to stand by all the time to respond as quickly as possible to the next animal.

Press any key to continue.

Box S15: Screen 4 of instructions for the graph condition in session 2

Resting phases

After all the work as a zookeeper you also need rest. Before, between and after the main task we will take some measurements during which you should just lie still. During these rest periods, please keep your eyes open and look at a cross the entire time. Blinking briefly is perfectly fine. The background of the screen will be dark during the resting phases. Please continue to lie very still and relaxed and continue to try to move as little as possible. Please try to stay awake the entire time.

Please wait for the study instructor.

Task instructions in German

Box S16: Screen 1 of instructions for the training condition in session 1

Willkommen zur Studie - Sitzung 1!
Bitte lesen Sie sich die folgenden Informationen aufmerksam durch. Falls Sie Fragen haben, können Sie diese gleich mit der Versuchsleitung klären. Bitte liegen Sie die gesamte Zeit so ruhig und entspannt wie möglich.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S17: Screen 2 of instructions for the training condition in session 1

Ihre Aufgabe:
Sie sind ein*e Zoowärter*in in Ausbildung und sollen darauf achten, dass alle Tiere in den richtigen Käfigen sind. Zuerst werden Sie in einem Training lernen, welches Tier in welchen Käfig gehört. Wir werden Ihnen jetzt genau erklären, wie diese Aufgabe funktioniert.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S18: Screen 3 of instructions for the training condition in session 1

Training (Teil 1)
Sie wollen Zoowärter*in werden und beginnen heute Ihre Ausbildung. Zuerst lernen Sie, welches Tier in welchen Käfig gehört. Sie werden gleich sechs Käfige im unteren Teil des Bildschirms sehen. Jeder der sechs Käfige gehört zu einem von sechs Tieren. Sie wählen einen Käfig mit der entsprechenden Antworttaste aus. Bitte lassen Sie Ihre Ring-, Mittel- und Zeigefinger die gesamte Zeit auf den Antworttasten, damit Sie so schnell und genau wie möglich antworten können.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S19: Screen 4 of instructions for the training condition in session 1

Während des Trainings erscheinen die Tiere über ihren Käfigen. Drücken Sie die Taste für diesen Käfig so schnell wie möglich und merken Sie sich den Käfig, in den das Tier gehört. Bitte drücken Sie die richtige Taste innerhalb von 1 Sekunde. Bitte antworten Sie so schnell und genau wie möglich. Sie erhalten eine Rückmeldung, wenn Ihre Antwort richtig, falsch oder zu langsam war. Dabei erscheint der richtige Käfig in Grün und der falsche Käfig in Rot.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S20: Screen 5 of instructions for the training condition in session 1

Es ist sehr wichtig, dass Sie sich aktiv merken, welches Tier in welchen Käfig gehört. Sie erhalten einen höheren Bonus, wenn Sie sich an die richtige Zuordnung erinnern. Je besser Sie sich daran erinnern, in welchen Käfig welches Tier gehört, desto mehr Geld verdienen Sie! Sie werden nun einen Durchgang dieser Aufgabe absolvieren, der circa 2 Minuten dauert.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S21: Screen 1 of instructions for the recall condition in session 1

Training (Teil 2)
Wir werden nun überprüfen, wie gut Sie die Zuordnung der Tiere zu ihren Käfigen gelernt haben. Die Tiere werden nun in der Mitte des Bildschirms erscheinen. Sie sollen sich an den richtigen Käfig für jedes Tier erinnern und dann die richtige Taste so schnell wie möglich drücken.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S22: Screen 2 of instructions for the recall condition in session 1

Dieses Mal antworten Sie erst nachdem das Tier gezeigt wurde. In jeder Runde erscheint zuerst das Tier in der Mitte des Bildschirms. Versuchen Sie dann bitte, sich die richtige Kombination von Tier, Käfig und Antworttaste aktiv vorzustellen. Danach erscheint zunächst ein kleines Kreuz für einen kurzen Moment. Dann erscheinen die Käfige und Sie können so schnell und genau wie möglich antworten. Bitte antworten Sie erst sobald die Käfige erscheinen, nicht früher.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S23: Screen 3 of instructions for the recall condition in session 1

Sie haben wieder 1 Sekunde Zeit zu antworten. Bitte antworten Sie wieder so schnell und genau wie möglich. Sie erhalten wieder eine Rückmeldung, wenn Ihre Antwort falsch oder zu langsam war. Wenn Ihre Antwort richtig war, geht es ohne Rückmeldung direkt mit der nächsten Runde weiter. Sie werden nun 8 Durchgänge dieser Aufgabe absolvieren, die jeweils circa 6 Minuten dauern. Zwischen den Durchgängen werden Sie die Möglichkeit bekommen, eine Pause zu machen.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S24: Screen 1 of instructions for the recall condition in session 2

Willkommen zur Studie - Sitzung 2!
Wir werden noch einmal überprüfen, ob Sie sich an die Zuordnung der Tiere zu ihren Käfigen erinnern können. Die Tiere werden wieder in der Mitte des Bildschirms erscheinen. Sie sollen sich wieder an den richtigen Käfig für jedes Tier erinnern und die richtige Taste so schnell wie möglich drücken.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S25: Screen 2 of instructions for the recall condition in session 2

Sie antworten wieder erst nachdem das Tier gezeigt wurde. In jeder Runde erscheint zuerst das Tier in der Mitte des Bildschirms. Versuchen Sie dann bitte, sich die richtige Kombination von Tier, Käfig und Antworttaste aktiv vorzustellen. Danach erscheint zunächst ein kleines Kreuz für einen kurzen Moment. Dann erscheinen die Käfige und Sie können so schnell und genau wie möglich antworten. Bitte antworten Sie erst sobald die Käfige erscheinen, nicht früher.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S26: Screen 3 of instructions for the recall condition in session 2

Sie haben wieder 1 Sekunde Zeit zu antworten. Bitte antworten Sie wieder so schnell und genau wie möglich. Sie erhalten wieder eine Rückmeldung, wenn Ihre Antwort falsch oder zu langsam war. Wenn Ihre Antwort richtig war, geht es ohne Rückmeldung direkt mit der nächsten Runde weiter. Sie werden nun einen Durchgang dieser Aufgabe absolvieren, der wieder circa 6 Minuten dauert. Nach dem Durchgang werden Sie die Möglichkeit bekommen, eine Pause zu machen.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S27: Screen 1 of instructions for the graph condition in session 2

Sie haben den Durchgang zu Erinnerung beendet! Gut gemacht! Sie können jetzt gerne eine kurze Pause machen und dabei auch Ihre Augen schließen. Bitte bleiben Sie weiterhin ruhig und entspannt liegen. Wenn Sie bereit sind, können Sie mit den Instruktionen für die Hauptaufgabe fortfahren.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S28: Screen 2 of instructions for the graph condition in session 2

Hauptaufgabe
Herzlichen Glückwunsch, Sie sind nun ausgebildete*r Zoowärter*in! Achtung: Manchmal brechen die Tiere aus ihren Käfigen aus! Ihre Aufgabe ist es, die Tiere wieder in die richtigen Käfige zu bringen. Wenn Sie ein Tier auf dem Bildschirm sehen, drücken Sie so schnell wie möglich die richtige Taste, um das Tier zurück in den richtigen Käfig zu bringen. Dieses Mal bekommen Sie keine Rückmeldung, ob Ihre Antwort richtig oder falsch war. Je mehr Tiere Sie in die richtigen Käfige bringen, desto mehr Bonus bekommen Sie am Ende der Studie! Die Hauptaufgabe besteht aus 5 Durchgängen, die jeweils circa 10 Minuten dauern.
Drücken Sie eine beliebige Taste, um fortzufahren.

Box S29: Screen 3 of instructions for the graph condition in session 2

Sie haben wieder 1 Sekunde Zeit zu antworten. In der Hauptaufgabe antworten Sie wieder sofort, wenn Sie ein Tier auf dem Bildschirm sehen. Bitte antworten Sie wieder so schnell und genau wie möglich. Zwischen den einzelnen Runden sehen Sie wieder ein Kreuz für einen Moment. Manchmal wird das Kreuz etwas kürzer und manchmal etwas länger gezeigt. Am Besten halten Sie sich die ganze Zeit bereit, um so schnell wie möglich auf das nächste Tier zu reagieren.

Drücken Sie eine beliebige Taste, um fortzufahren.

Box S30: Screen 4 of instructions for the graph condition in session 2

Ruhephasen

Nach der ganzen Arbeit als Zoowärter*in braucht man auch Erholung. Vor, zwischen und nach den Durchgängen der Hauptaufgabe machen wir einige Messungen bei denen Sie einfach nur ruhig liegen sollen. In diesen Ruhephasen sollen Sie bitte Ihre Augen geöffnet halten und die gesamte Zeit auf ein Kreuz schauen. Kurzes Blinzeln ist vollkommen in Ordnung. Der Hintergrund des Bildschirms wird in den Ruhephasen dunkel sein. Bitte liegen Sie weiterhin ganz ruhig und entspannt und versuchen Sie weiterhin sich so wenig wie möglich zu bewegen. Versuchen Sie bitte die gesamte Zeit wach zu bleiben.

Bitte warten Sie auf die Versuchsleitung.