Title

Individual-specific versus shared cognitive states differently support complex semantic and perceptual judgments

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Vijayarajah, Sagana
Schlichting, Margaret L

Publication Date

2022

Peer reviewed

Individual-specific versus shared cognitive states differently support complex semantic and perceptual judgments

Sagana Vijayarajah (sagana.vijayarajah@mail.utoronto.ca)
Department of Psychology, University of Toronto, Toronto, ON, Canada

Margaret L. Schlichting (meg.schlichting@utoronto.ca)
Department of Psychology, University of Toronto, Toronto, ON, Canada

Abstract

Cognitive processes that underpin performance on a given task may vary both within and across individuals. Yet, it is unclear how individual-specific versus shared cognitive processes each support behaviour. Here, we used a functional magnetic resonance imaging (fMRI) pattern classifier approach to ask how individual-specific and shared neural cognitive states differently relate to an individual's ability to detect consecutive repeats in semantic (story) meaning versus perceptual (artist style) dimensions of illustrations that depicted well-known stories. Both states were related to participants' task performance overall but differently for story versus artist style behaviours: individual-specific states were related to story performance, whereas shared states were related to artist style performance. These findings suggest that behaviours relying upon prior knowledge—likely varying across individuals—may be supported by idiosyncratic versus shared states. In contrast, unfamiliar judgments associated with a smaller number of eligible strategies may be supported by a state shared across individuals.

Keywords: semantic; perceptual; attention; fMRI; MVPA

Introduction

A single perceptual experience can evoke a variety of cognitive processes, with the tendency to adopt one over another varying both across people and within a given person over time. For example, imagine a visitor to a museum stops to view a painting. The viewer may focus on discerning the meaning of the painting by identifying its composite elements or relating it to their own past experiences. In contrast, other viewers might focus on elements of the painter's style such as the vibrant colours and broad brushstrokes. While some of these experiences may elicit a cognitive state that is common across different viewers, others may be idiosyncratic to a particular person—ultimately suggesting a mechanism by which a given event can elicit either shared or uniquely personal experiences. Here, we suggest that cognitive process will be reflected in a person's "brain state," or distributed pattern of neural activation measured using fMRI. We reasoned that directly comparing such neural patterns across people would shed light on how different individuals approach a given task—such that shared and idiosyncratic states would yield high and low across-participant similarity, respectively—as well as how such approaches are related to behavioural differences.

Neural decoding analyses have shown that an individuals' engagement of a cognitive process evokes activation patterns that are stable even across different experiences (i.e., brain states; Hanke et al., 2009; for reviews see Herz et al., 2020; Norman et al., 2006) and relate to subsequent behavior (Carlson et al., 2003; Kuhl et al., 2012; Kuhl & Chun, 2014; Mostert et al., 2015; Sudre et al., 2012). Furthermore, cognitive state as assessed by neural decoding also predicts ongoing behaviour: For example, fluctuating attentional states for face versus scene images relates to sustained attention performance (DeBettencourt et al., 2015). However, decoding states within individuals—as has been the approach in most past work using this technique—does not reveal anything about how an individual's state compares with others', or how any neural divergence from the group might be related to behaviour.

Recent studies have shown that certain cognitive processes evoke brain states that can be decoded across individuals. For example, researchers were able to "read out" from neural patterns the particular memory strategy participants used during an initial experience—and predict later behaviour—using other participants' brain states as a reference (Richter et al., 2016). It has been proposed that such states reflect the common knowledge that people rely upon or extract from a new experience. For instance, viewers tend to evoke similar neural representations when processing semantic themes from movies or narratives (Baldassano et al., 2017; Chen et al., 2017; Hasson et al., 2008; Honey et al., 2012; Meer et al., 2020; Regev et al., 2013; Zadbood et al., 2017), which has been taken as evidence that shared knowledge supports a shared experience engaging with the story. As further support for this idea, other work has shown that disrupting participants' reliance upon common knowledge impacted their ability to extract a shared event representation: event representations were less similar across individuals who had extracted different meanings from the same event (Nguyen et al., 2019) or when the new events could not be linked to prior knowledge (Lerner et al., 2011). These findings suggest that the degree to which an individual's state aligns with others' may reflect the similarity of their interpretations.

Despite this suggested link between a shared interpretation of stimulus meaning and a consistent brain state across people, how such states relate to interpretations evidenced in ongoing behaviours remains unclear. More broadly, few studies have directly compared the behavioural relevance of brain states that are shared versus idiosyncratic across people, making it an open possibility that each state is beneficial in a

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

particular task context. Given the work described above (Baldassano et al., 2017; Chen et al., 2017; Hasson et al., 2008; Honey et al., 2012; Lerner et al., 2011; Meer et al., 2020; Nguyen et al., 2019; Regev et al., 2013; Zadbood et al., 2017), one might expect that an orientation towards the meaning of an experience could elicit similar processes across people as they rely upon their shared semantic knowledge to arrive at an interpretation. Conversely however, past work also suggests an individual's own neural engagement is important for processing meaning-based features (Chadwick et al., 2016; Linde-Domingo et al., 2019; Liuzzi et al., 2020; Sudre et al., 2012)—for instance in showing that individual-specific neural representations were more predictive of meaning-related memory errors than those shared across people (Chadwick et al., 2016). Therefore, it remains unknown whether shared or individual-specific brain states might be more predictive of behaviours reflecting arrival at a common interpretation.

Here we explore how individual-specific versus shared brain states relate to performance in an ongoing task. Motivated by past behavioural paradigms that manipulated participant's cognitive orientation towards semantic (meaning) versus perceptual features to examine processing related to common knowledge (Challis et al., 1996; Craik & Lockhart, 1972; Lockhart, 2002), we examined how brain states for attention to meaning versus perceptual features differently relate to their respective behaviours. In this investigation we contrasted meaning with perceptual style-based orientations because perceptual features are less likely to promote processing in relation to shared knowledge or past experiences (Craik & Lockhart, 1972), but still require complex (style) judgments of the same stimuli. Specifically, we cued participants to make judgments about the story meaning or perceptual artist style of illustrations depicting artists' renditions of well-known storybook stories. We then used a classification analysis to ask if we can reliably decode individual-specific and shared brain states evoked by participants as they made these story meaning and artist style judgments. We then related both orientation states to people's ability to accurately detect consecutive repeats in story versus artist style to ask which brain states were important for behaviour. We predicted that states shared across individuals would support the extraction of stimulus (story) meaning. In contrast, we anticipated that novel perceptual discriminations of artist style that were less related to prior knowledge would be supported by individual-specific states.

Method

Participants

Forty-two right-handed adults participated in this experiment (28 females, 14 males; mean age=19.8 years, SD=2.4 years; 18-30 years old). This sample size was chosen *a priori* to achieve 80% power to detect an effect size of d=0.45 based on previous work (Aly & Turk-Browne, 2015). All participants provided written consent. The experimental protocol was approved by our university's ethics board.

Design

Stimuli Participants viewed 144 storybook-style illustrations that varied in story theme (story depicted) and artist style (artist creator). A subset of these illustrations were repeat pairs in either story or artist—i.e., they depicted the same story but were created by different artists or depicted different stories created by the same artist, respectively. Illustrations were organized into 18 blocks of 8 illustrations each.

Attention cues preceded illustration blocks to indicate the upcoming task. The cues were simple black shapes (a diamond, star, and square) that participants were pre-trained to associate with the different tasks (described in detail below).

Task and behavioural analysis Participants completed two different tasks with the illustration blocks during fMRI scanning: the artist task and the story task. Participants also completed blocks from an unrelated baseline task which did not use the illustration stimuli (not discussed here). The experiment was divided into three runs of equal length, yielding three blocks from each task per run. Fixation was included at the beginning (3s) and end (9s) of each run to allow for stabilization and lag of the MR signal, respectively.

Before each block, an attention cue was presented (2500ms with a 500ms interstimulus interval [ISI]; Figure 1A) to indicate the task for the upcoming block. The assignment of the cues to tasks was counterbalanced across six groups of participants to ensure our neural decoding of shared states could not be attributed to specific cues. Following each cue, illustrations were presented one at a time in blocks for 2500ms with a 500ms ISI.

In the artist and story tasks, participants performed a modified 1-back judgment in which they made a button box response to indicate whether an illustration was or was not a consecutive repeat along the cued dimension (i.e., artist style repeats in artist, and story theme repeats in story). Importantly, the structure of blocks was held constant across tasks while participants' cognitive orientation varied: most blocks (12 blocks) contained one artist style repeat ("artist repeats"), and one story theme repeat ("story repeats"); the remaining 4 illustrations depicted unique stories and artist styles (Figure 1A). Half of these blocks were assigned to each task, with the task assignment counterbalanced across participants to control for stimulus-specific differences between tasks. The remaining six blocks contained an additional repeat at the end of the block purely to reduce task predictability (not considered in subsequent analyses), with two instead of four illustrations depicting unique stories and artist styles so the block still contained 8 illustrations. These blocks were always assigned to the task that aligned with the additional repeat (three blocks per task). Altogether, each participant viewed nine blocks in each task.

Because our block structure was consistent across tasks, we summarized performance as participants' ability to detect repeats specifically along the cued dimension in both tasks (e.g., making a repeat response to artist but not story repeats in the artist task). We then related their behavioural

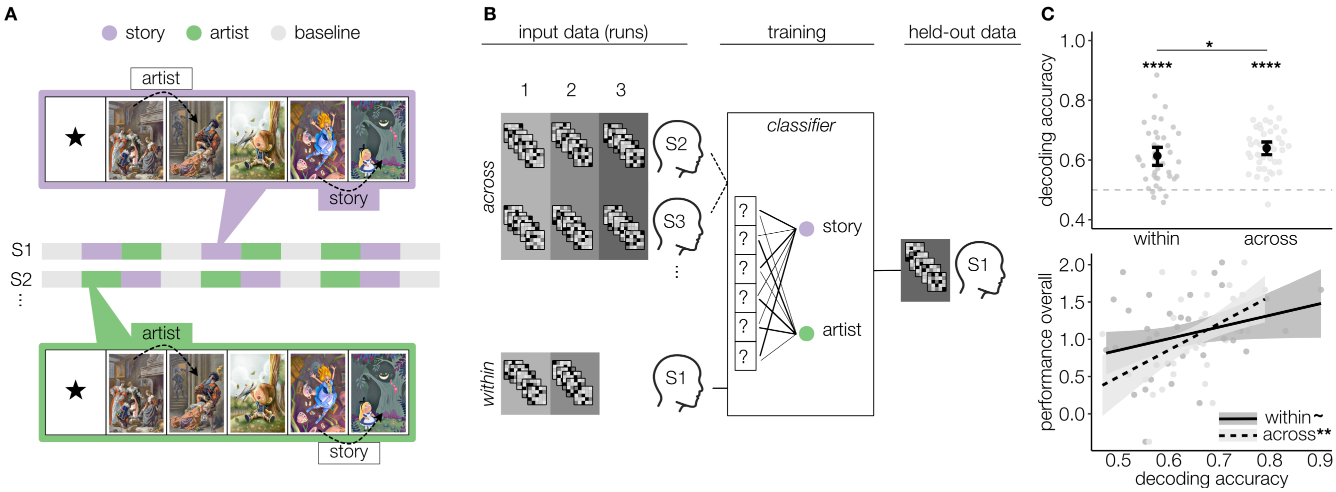

Figure 1: A) Task. Example of partial illustration block in either the story (purple, top, Subject 1 [S1]) or artist (green, bottom, S2) task across participants. Each block contained at least one each artist and story repeats (denoted with "artist" and "story" above the curved arrows, respectively). Each horizontal bar depicts a run that included baseline (light grey) every third block, with one each artist and story blocks (order counterbalanced) in between. Attention cues (depicted; star) preceded each block. Here, we show that the same cue shape indicated story for S1 and artist for S2 because cue shapes were counterbalanced across participants. B) Classification analysis within- (bottom) and across-participants (top). Input data (left; either from multiple or a single participant) was used to train a classifier to discriminate between artist and story orientations to illustration features (middle). The trained classifier was then applied to held-out data (right) to calculate the accuracy of the classifier predictions of state. C) Top, Decoding accuracy within- (left) and across- (right; mean center dots with 95% confidence bars, individual participant data points as smaller dots) participants was well above chance (dashed line), and greater for across than within. Bottom, Decoding accuracy within- (solid; dark grey) and across- (dashed; light grey) participants was related to behavioural performance (ribbons depict 95% CI). ~ $p<0.1$, * $p<0.05$, ** $p<0.01$, **** $p<0.0001$

performance to brain states to ask whether the brain states participants engaged while viewing illustrations related to their ability to detect cued repeats in both tasks.

MRI data collection and preprocessing

Data acquisition Imaging data were collected using a 3.0T Siemens Prisma MRI scanner. Sixty-nine functional slices angled between 15–30° (oblique axial; orientation chosen to maximize participants' brain coverage; repetition time [TR]=1500ms, echo time [TE]=28.0ms, flip angle=71°, 220 x 220 x 138mm matrix, 2mm isotropic voxels, multiband acceleration factor=3, GRAPPA factor=2) were collected using a multi-band echo-planar imaging (EPI) sequence. A structural T1-weighted 3D magnetization-prepared rapid gradient echo (MPRAGE; 256 x 256 x 160mm matrix, 1mm isotropic voxels) volume was collected for co-registration and spatial normalization into standard template space, along with a field map to correct for susceptibility distortion (TR=700ms, TE=4.92/7.38 ms, flip angle=60°, 220 x 220 x 138mm matrix, 2mm isotropic voxels).

Data preprocessing Volumes were preprocessed with the fMRIprep version 1.1.4 automated pipeline (Esteban et al., 2018). Structural images were corrected for intensity non-uniformity (N4BiasFieldCorrection; Tustison et al., 2010), skull-stripped (ANTs brain extraction tool), and normalized into 2mm isotropic MNI152NLin2009cAsym template space (Fonov et al., 2009). Volumes were also segmented into cerebrospinal fluid, white-matter, and gray-matter (GM) (FSL FAST; Zhang et al., 2001) to create T1-weighted GM tissue-probability maps.

Functional data were corrected for motion (FSL MCFLIRT) and susceptibility distortion (FSL FUGUE) before normalization into template space (nonlinear registration; ANTs antsApplyTransforms and Lanczos interpolation; boundary-based registration; 9 degrees of freedom; FreeSurfer BBREGISTER; Greve & Fischl, 2009).

Decoding cognitive orientation towards artist style versus story features

Regions of interest We performed all our neural decoding analyses using a group whole-brain GM mask. This mask was created by first generating participant-specific GM masks using their T1-weighted GM tissue-probability maps (threshold=0.5) transformed into template space. These masks were then merged and thresholded to include only GM voxels in at least 50% of all participants (N=21+) to create the group mask used in our analyses.

Classification approach Before performing the classification analyses, participants' unmodeled neural data was shifted by four timepoints (6s) in each run to account for hemodynamic lag in the fMRI signal. We also removed three timepoints at the onset of each attention cue to exclude cue to task transition periods from our analyses. Non-task fixation time that occurred at the start and end of each run was also excluded from the data. Using the remaining data from the artist and story tasks, we then trained sparse multinomial logistic regression (SMLR) classifiers in PyMVPA (default parameters; no feature selection was performed; Hanke et al., 2009) to decode task-evoked states within each participant, and those that were shared across participants (Figure 1B).

Characterizing participant-specific states We used neural data from the story and artist tasks to train a classifier to decode neural patterns evoked by each participant in these tasks. A leave-one-run-out cross-validation approach was used for classifier training: we trained the classifier on task-labeled timepoints (story/artist task) from two of the three runs and tested the classifier's accuracy on the held-out run (Figure 1B, bottom). This process was repeated in three folds such that each run was held out once.

Characterizing across-participants states We also examined whether artist and story states were shared across participants. We trained another classifier to decode artist and story states consistent across participants by this time using a leave-one-participant-out cross-validation approach. In this approach, we trained the classifier on task-labeled timepoints from all but six participants—one from each of the six counterbalancing groups—and tested the classifier's accuracy on the six held-out participants (Figure 1B, top). We excluded one participant from each counterbalancing group from training to ensure the training set was balanced in terms of the assignment of attention cues to tasks. This process was repeated seven times (42/6) so that all participants were used to test the classifier's accuracy once.

Decoding statistical analyses The classification analyses provided the following metrics derived from both each participant's own neural data, as well as with data across all other participants: 1) predicted, binary (story/artist) task labels for each timepoint; and 2) continuous estimates of the degree to which neural patterns reflect artist and story states, on a timepoint-by-timepoint basis. We first tested whether we could reliably decode artist and story states within and across participants by calculating the accuracy of the classifiers' predicted task labels for timepoints across artist and story tasks, and separately within each task. We then correlated decoding accuracy to behavioural performance to test for evidence of a relationship between these decoded states and all task behaviours in general, and each task specifically.

Once we established that we could decode artist and story states that were also generally related to behaviour, we used a more fine-grained approach to ask how within- versus across-participants states support trial-to-trial variability in task behaviours: we related the continuous estimates of artist and story state evidence from consecutive illustrations to participants' performance in the artist versus story tasks. Specifically, the continuous estimates of story and artist state evidence were log odds transformed to correct for any non-normality in the distribution of classifier estimates (Richter et al., 2016). Then, linear-mixed effects models (R statistical package version 4.0.4; R Core Team, 2021; lme4 package version 1.1-26; Bates et al., 2015) were used to ask whether brain states in the moments leading up to the presentation of a cued repeat were predicted by task accuracy (1=correct, hit; 0=incorrect, miss), on a trial-by-trial basis (while accounting for within-participant variance).

Results

Participants correctly modulated their behaviour in response to the attention cues

We compared the proportion of repeat responses to cued repeats (hits) versus the alternate repeat type (false alarms) to assess participants' behaviour (18 repeats each, per participant; half from each task). Because we were interested in how brain states relate to individual differences in behavioural performance overall and within artist versus story, we examined performance averaged across tasks and then separately in each task. Participants' task performance was well above chance overall (d'; $t(41)=13.5$, $p<0.001$, Cohen's $d=2.08$) and in both tasks separately (d'; t-test versus 0; story: $t(41)=8.24$, $p<0.001$, $d=1.27$; artist: $t(41)=13.7$, $p<0.001$, $d=2.11$). Performance was also significantly better in the artist versus story task ($t(41)=2.57$, $p=0.014$, $d=0.397$) due to elevated false alarms in story over artist ($t(41)=3.14$, $p=0.003$, $d=0.484$); there was no difference in hits between tasks ($p=0.969$). Therefore, participants did correctly modulate their behaviour to the attention cues, with potentially less difficulty in the artist task.

Successful decoding of neural artist versus story states within and across participants

We assessed if we could reliably decode artist and story states within participants, and if these states were shared across participants (Figure 1C, top). Indeed, decoding of artist and story tasks was reliably above chance for both within (mean=0.612; 95% CI [0.594, 0.630]; t-test versus 0.50; $t(41)=7.54$, $p<0.001$, $d=1.16$) and across participants (mean=0.639; 95% CI [0.621, 0.657]; t-test versus 0.50; $t(41)=13.2$, $p<0.001$, $d=2.04$). A direct comparison of within- versus across-participants decoding accuracy showed higher accuracy with across-participants states (paired t-test; $t(41)=2.31$, $p=0.040$, $d=0.328$). We separately considered the decoding of artist versus story blocks to further characterize this accuracy difference and found that the across-participants classifier demonstrated greater accuracy in artist than story ($t(41)=2.12$, $p=0.040$, $d=0.327$). In other words, the classifier was more accurate to identify the cued state when it was an artist over story block. There was no reliable difference in within-participants decoding accuracy for artist versus story ($p=0.198$). Therefore, while we successfully decoded whole-brain artist style and story states both within and across participants, there were differences in the artist and story orientations characterized in these states.

Lastly, we assessed whether decoding of within- and across-participants states were related to individual differences in participants' task behaviour overall (Figure 1C, bottom). Indeed, across decoding was reliably related to task performance overall (d'; r=0.485, t(40)=3.51, p=0.001), with within decoding demonstrating moderate evidence for the same relationship (r=0.300, t(40)=1.99, p=0.053). Thus, in subsequent analyses, we further characterized how within- versus across-participants decoding support behaviour by examining artist and story task judgments separately.

Within-participants decoding was related to story judgments

To interrogate the relationship between state decoding and artist versus story judgments, we assessed how classifier accuracy in decoding artist and story blocks separately was related to performance on their respective tasks (cued repeat detection d'). Specifically, is the degree to which the classifier can accurately decode story blocks related to story but not artist task performance, and vice versa? Importantly, we found no evidence for a correlation between artist and story task performance (p=0.137), suggesting that it was not the case that participants who performed well in one task also performed well in the other task. Thus, we can consider participants' artist and story task performance separately. With respect to story, within- (r=0.335, t(40)=2.25, p=0.030) but not across-participants (p=0.159) story block decoding accuracy was related to story task performance (Figure 2A, left); however, the difference between these relationships was not reliable (p=0.859; Figure 2A, left). The relationship between within-participants story block decoding and story task performance was only present for the respective (story) task: we found no evidence of this relationship with artist task performance (p=0.158, Figure 2A, right), although there was no reliable difference between the relationships to story versus artist task performance (p=0.433; Figure 2A, left versus Figure 2A, right). In contrast to story block decoding, there was no *specific* relationship between artist block decoding and its respective task. Across- (r=0.404, t(40)=2.79, p=0.008) but not within-participants (p=0.432) artist block decoding accuracy was related to artist task performance (Figure 2B, right). The relationship to artist task performance was also moderately greater for across- versus within-participants artist block decoding accuracy (t=1.74, p=0.086; Figure 2B, right). However, across-participants artist block decoding accuracy was also related to story task performance (r=0.507, t(40)=3.71, p=0.001, Figure 2B, left), and there was no reliable difference between the relationships to story versus artist task performance (p=0.331; Figure 2B, right versus Figure 2B, left), suggesting that this decoding relationship was not specific to the respective task. Therefore, while both within- and across-participants states are related to task performance overall, within-participants states demonstrate a relationship between story block decoding and behavioural story judgments.

Differences in across-participants states preceding correct versus incorrect artist judgments

We next examined if variability in the engagement of artist

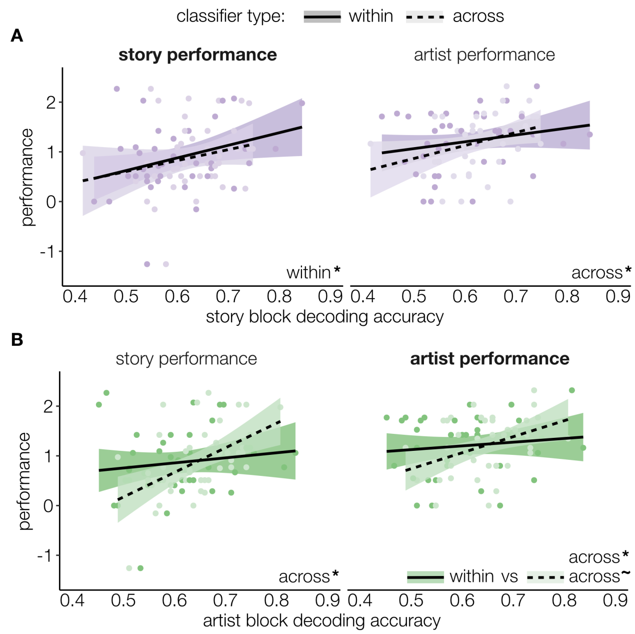

Figure 2: Relating within- and across-participants story versus artist block decoding accuracy to behavioural performance in story (left) versus artist (right) tasks. A) Left, Within- (solid line, dark purple) not across- (dashed line, light purple) participants story block decoding accuracy was related to story task performance. Right, No relationship between within-participants story block decoding and artist performance (solid line, dark purple). Across-participants story block decoding was related to artist performance (dashed, light purple). B) Right, Across- (dashed line, light green) but not within-participants (solid line, dark green) artist block decoding accuracy was related to artist performance. Left, Across-participants artist block decoding was also related to story performance (dashed line, light green). Ribbons represent the 95% CI. Smaller dots represent individual participant data points. ~ p<0.1, * p<0.05

and story states over illustrations can be predicted by trial-wise behavioural accuracy, in the respective tasks. In other words, do participants demonstrate differential evidence for the cued state leading up to correct versus incorrect repeat detection trials (Figure 3A)? For story states, neither within (p=0.128) nor across-participants (p=0.806) states differed between correct and incorrect responses. In contrast, there was reliably more across-participants artist state evidence across illustrations for upcoming correct versus incorrect responses (β=0.837, SE=0.325, t=2.57, p=0.011; Figure 3B). Within-participants artist states did not show the same relationship (p=0.141). Therefore, across-participants states may support participants' accurate artist style judgments.

Discussion

Here, we successfully decoded cognitive orientation to story versus artist both within and across individuals, and found that these states were related to different behaviours: While individual-specific and shared brain states were related to overall task performance, interrogating story versus artist

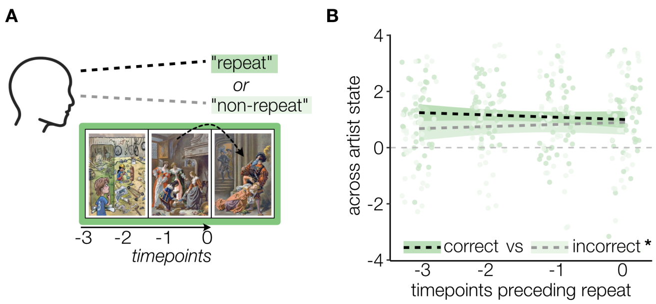

Figure 3: A) Depiction of how we related changes in cued state evidence to upcoming correct versus incorrect responses. We examined whether continuous estimates of state evidence from consecutive illustrations (example illustrations stream depicted; 4 timepoints in total; 1.5s each) that preceded the presentation of cued repeats (illustration following the 0) differed for correct ("repeat") versus incorrect ("non-repeat") responses. B) Estimated marginal means for across artist state evidence as a function of correct versus incorrect responses. Across consecutive illustrations, across-participants artist state evidence was above chance (dashed grey line at y=0) and greater preceding correct (dashed black line; dark green) versus incorrect (dashed grey line; light green) responses. Ribbons represent the 95% CI. Smaller dots represent individual participant data points. * $p<0.05$

style behaviours showed that individual-specific states were related to meaning-based story judgments, while shared states were related to accurate perceptual style judgments.

Participants' tendency to consistently engage the same state across trials was related to their ability to make story judgments. While these results may appear to contradict past work that suggests similar neural representations across individuals underlie the processing of the same semantic themes (Baldassano et al., 2017; Chen et al., 2017; Hasson et al., 2008; Honey et al., 2012; Meer et al., 2020; Regev et al., 2013), one speculative possibility is that our findings reflect that participants may have used a variety of different approaches to accomplish the story task. Consistent with this idea, past work has shown less neural consistency across individuals who ultimately extract different abstract meanings from the same event (Nguyen et al., 2019)—suggesting that in our case, neural divergence across participants may indicate the multitude of different approaches and ultimate interpretations participants might have had in the story task.

Variation in the strategies participants used to make meaning-based judgments may stem from differences in their prior story knowledge, or the qualitative nature of meaning-oriented processing. Firstly, if participants are relying on different types of prior story knowledge when making story judgments, they may have used divergent strategies to perform the task—e.g., while some may have focused on recognizing main characters from their limited knowledge of the story, others may have attempted to identify key story events from their more extensive knowledge. Post-experiment self-report measures suggest there was indeed large variability in participants' prior story knowledge. Such differences in knowledge could give rise to different elaborative processes across people, as cuing an individual's orientation toward semantic over shallow perceptual features may encourage more elaborative processing and connections with semantic (Craik & Tulving, 1975; Craik & Lockhart, 1972; Fisher & Craik, 1980; Moscovitch & Craik, 1976) or autobiographical memories (Pasupathi et al., 2007; Warren et al., 2016). Differences in the elaborative connections participants made to their past story knowledge and experiences may evoke differences in meaning-based states across participants. Future investigations that systematically assess participants' self-reported strategies for the story and artist tasks would be needed to explore this speculation.

In contrast to story judgments, artist style judgments evoked a more consistent state across individuals that benefited artist task performance. A general lack of expertise with the artist style dimension among our participants may be the reason for such a benefit. Although artist style features were intended to be less connected to participants' past knowledge than story themes, less expertise in making artist style discriminations may result in participants having less diverse strategies available when performing the artist task. This is consistent with previous work that has shown greater experience with a particular skill can increase the number of strategies an individual can use to perform the skill (Chase & Simon, 1973; Ericsson & Lehmann, 1996; Gobet & Waters, 2003). Therefore, a lack of familiarity with the artist style dimension across individuals may explain why a shared brain state was related to artist style judgments. It may also be the case that the representation of perceptual information has greater alignment across subjects (Haxby et al., 2020), such that it is better decoded across subjects.

In addition to the unfamiliarity of artist style judgments, lower distinctiveness of artist versus story features may also support neural convergence across participants in the artist task. Past work has suggested that low-level perceptual features are susceptible to memory errors because they are less distinct than semantic features (Elias & Perfetti, 1973; Hunt, 2013; Lockhart, 2002)—e.g., the broad brushstrokes in an artist style may appear in many illustrations versus the story features that may be more unique. Therefore, our finding of shared brain states is also consistent with common processing of these constituent perceptual features.

Together, these findings suggest that the degree to which individual-specific and shared states can benefit behaviour depends on the nature of that behaviour. When faced with a task that affords solutions in myriad potential ways, consistently engaging an individual-specific state may support behaviour. In contrast, we speculate that judgments along an unfamiliar dimension may offer a fixed set of approaches and therefore converge across people.

Acknowledgements

We thank M. Mack, A. Blumenthal, A. Finn, M. Moscovitch, X. Fang, and members of the Budding Minds lab for helpful discussions. This work was supported by a

Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN-04933-2018), Canada Foundation for Innovation JELF, Ontario Research Fund (36876), and University of Toronto startup funds to MLS; and a NSERC Postgraduate Doctoral Scholarship, NSERC Graduate-Master's Program Scholarship, and Ontario Graduate Scholarship to SV.

References

Aly, M., & Turk-Browne, N. B. (2015). Attention Stabilizes Representations in the Human Hippocampus. *Cerebral Cortex*, *26*(2).

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, *95*(3), 709-721

Carlson, T. A., Schrater, P., & He, S. (2003). Patterns of Activity in the Categorical Representations of Objects. *Journal of Cognitive Neuroscience*, *15*(5), 704–717.

Chadwick, M. J., Anjum, R. S., Kumaran, D., Schacter, D. L., Spiers, H. J., Hassabis, D., Designed Research, D. H., & Performed, R. S. A. (2016). *Semantic representations in the temporal pole predict false memories*. *113*(36), 10180–10185.

Challis, B. H., Velichkovsky, B. M., & Craik, F. I. M. (1996). Levels-of-processing effects on a variety of memory tasks: New findings and theoretical implications. *Consciousness and Cognition*, *5*(1–2), 142–164.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55–81.

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience*, *20*(1), 115.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294.

Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684.

DeBettencourt, M. T., Cohen, J. D., Lee, R. F., Norman, K. A., & Turk-Browne, N. B. (2015). Closed-loop training of attention with real-time brain imaging. *Nature Neuroscience 2015 18:3*, *18*(3), 470–475.

Elias, C. S., & Perfetti, C. A. (1973). Encoding task and recognition memory: The importance of semantic encoding. *American Psychological Association*, *99*(2), 151–156.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of Maximal Adaptation to Task Constraints. *Annu. Rev. Psychol*, *47*, 273–305.

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2018). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods 2018 16:1*, *16*(1), 111–116.

Fisher, R. P., & Craik, F. I. M. (1980). The effects of elaboration on recognition memory. In *Memory & Cognition* (Vol. 8, Issue 5).

Fonov, V., Evans, A., McKinstry, R., Almli, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, *47*, S102.

Gobet, F., & Waters, A. J. (2003). The Role of Constraints in Expert Memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *29*(6), 1082–1094.

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, *48*(1), 63–72.

Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, *7*(1), 37–53.

Hasson, U., Furman, O., Clark, D., Dudai, Y., & Davachi, L. (2008). Enhanced Intersubject Correlations during Movie Viewing Correlate with Successful Episodic Encoding. *Neuron*, *57*(3), 452–462.

Haxby, J. V., Guntupalli, J. S., Nastase, S. A., & Feilong, M. (2020). Hyperalignment: Modeling shared information encoded in idiosyncratic cortical topographies. *ELife*, *9*, 1–26.

Herz, N., Baror, S., & Bar, M. (2020). Overarching States of Mind. *Trends in Cognitive Sciences*, *24*(3), 184–199.

Honey, C. J., Thompson, C. R., Lerner, Y., & Hasson, U. (2012). Not lost in translation: Neural responses shared across languages. *Journal of Neuroscience*, *32*(44), 15277–15283.

Hunt, R. R. (2013). Precision in Memory Through Distinctive Processing. *Current Directions in Psychological Science*, *22*(1), 10–15.

Kuhl, B. A., & Chun, M. M. (2014). Successful Remembering Elicits Event-Specific Activity Patterns in Lateral Parietal Cortex. *Journal of Neuroscience*, *34*(23), 8051–8060.

Kuhl, B. A., Rissman, J., & Wagner, A. D. (2012). Multi-voxel patterns of visual category representation during episodic encoding are predictive of subsequent memory. *Neuropsychologia*, *50*(4), 458–469.

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic Mapping of a Hierarchy of Temporal Receptive Windows Using a Narrated Story. *Journal of Neuroscience*, *31*(8), 2906–2915.

Linde-Domingo, J., Treder, M. S., Kerrén, C., & Wimber, M. (2019). Evidence that neural information flow is reversed between object perception and object reconstruction from memory. *Nature Communications*, *10*(1), 179.

Liuzzi, A. G., Aglinskas, A., & Fairhall, S. L. (2020). General and feature-based semantic representations in the semantic network. *Scientific Reports 2020 10:1*, *10*(1), 1–12.

Lockhart, R. S. (2002). Levels of processing, transfer-appropriate processing, and the concept of robust encoding. *Memory*, *10*(5–6), 397–403.

Meer, J. N. van der, Breakspear, M., Chang, L. J., Sonkusare, S., & Cocchi, L. (2020). Movie viewing elicits rich and reliable brain state dynamics. *Nature Communications 2020 11:1*, *11*(1), 1–14.

Moscovitch, M., & Craik, F. I. M. (1976). Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. *Journal of Verbal Learning and Verbal Behavior*, *15*(4), 447–458.

Mostert, P., Kok, P., & de Lange, F. P. (2015). Dissociating sensory from decision processes in human perceptual decision making. *Scientific Reports 2015 5:1*, *5*(1), 1–13.

Nguyen, M., Vanderwal, T., & Hasson, U. (2019). Shared understanding of narratives is correlated with shared neural responses. *NeuroImage*, *184*, 161–170.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. v. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–430.

Pasupathi, M., Mansour, E., & Brubaker, J. R. (2007). Developing a Life Story: Constructing Relations between Self and Experience in Autobiographical Narratives. *Human Development*, *50*, 85–110.

Regev, M., Honey, C. J., Simony, E., & Hasson, U. (2013). Selective and Invariant Neural Responses to Spoken and Written Narratives. *Journal of Neuroscience*, *33*(40), 15978–15988.

Richter, F. R., Chanales, A. J. H., & Kuhl, B. A. (2016). Predicting the integration of overlapping memories by decoding mnemonic processing states during learning. *NeuroImage*, *124*, 323–335.

Sudre, G., Pomerleau, D., Palatucci, M., Wehbe, L., Fyshe, A., Salmelin, R., & Mitchell, T. (2012). Tracking neural coding of perceptual and semantic features of concrete nouns. *NeuroImage*, *62*(1), 451–463.

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320.

Warren, M. W., Hughes, A. T., & Tobias, S. B. (2016). Autobiographical Elaboration and Memory for Adjectives, *60*(1), 55–

Zadbood, A., Chen, J., Leong, Y. C., Norman, K. A., & Hasson, U. (2017). How We Transmit Memories to Other Brains: Constructing Shared Neural Representations Via Communication. *Cerebral Cortex*, *27*, 4988–5000.

Zhang, Y., Brady, M., & Smith, S. (2001). 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.