Fluid intelligence relates to neural measures of cognitive map formation

Tenderra, R.M.[1]* & Theves, S.[1,2]*

[1] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[2] Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany

*Corresponding Authors: tenderra@cbs.mpg.de and stephanie.theves@ae.mpg.de

Abstract

Psychometric research on intelligence consistently identifies latent factors underlying performance correlations across cognitive tasks, with one general factor (g) explaining most variance and predicting general life outcomes. Their biological basis is yet unresolved, in particular with regard to the neural information processing mechanisms that may underlie intelligence. Here we test the hypothesis that interindividual differences in relational processing, supported by cognitive maps in the hippocampus, are related to fluid intelligence (g_f), a statistical approximation of the general factor. Using standardized cognitive tests and fMRI of mnemonic tasks, we demonstrate a positive correlation specifically between g_f and map-like relational encoding of task information. This relationship was neither present for non-relational mnemonic responses in the hippocampus nor for less g-related factors. These findings offer first empirical support for a link between neural mechanisms related to relational reasoning and general cognitive performance and probe the presumed relevance of hippocampal coding properties for cognition more broadly.

Introduction

The study and definition of human intelligence is a longstanding interdisciplinary endeavor. Psychologists assess general cognitive ability by evaluating performance across diverse cognitive tasks[1]. A major discovery is that positive correlations between task performances can be explained by a small number of latent factors, which load onto one general factor g[2]. G is also the strongest single predictor of general life outcomes[3,4]. Among the second-order factors, the general ability for reasoning or fluid intelligence (g_f) shows a near-perfect correlation to g[5]. While this factor structure is well-replicated, its biological basis remains unclear. Does a confined set of cognitive and neural functions underlie g? Previous neuroimaging research on intelligence has focused on locating interindividual differences in anatomical properties[6–8], structural and functional connectivity[9,10], or general activation levels, with an emphasis on the parieto-frontal cortices[11–13]. However, the mechanisms of neural information processing that may underlie general cognitive abilities are yet unknown. Here we take an information processing-based perspective, asking how individuals differ in the way they encode task information with respect to measures of general cognitive performance. Given the significant contribution of general reasoning ability to g[5,14] and the central role of relational reasoning in human cognition[15–17], it can be hypothesized that the general factor is related to the degree to which individuals encode the relational and structural information of experiences critical to ad-hoc inferences and transfer to novel situations[18]. Empirical and computational work implicates a critical role for representational properties of the hippocampal-entorhinal system in reasoning: The system is thought to integrate spatial and non-spatial relations between experiences[19–25] into domain-general cognitive maps of physical and abstract task spaces[26–28], especially when these are task-relevant[20]. In particular, the hippocampal system may support decision making by compressing multivariate inputs according to task-relevant dimensions[29–31]. This may help capture relational structure from experiences in support of reasoning, i.e., to infer unobserved links[32–34] and states[35] in the same context, or to eventually generalize it to novel contexts. For instance, the hippocampus has been shown to support the inference of unobserved relations in paired associative inference[34,36] and transitive inference tasks[32] which rely on integrating individual associations into a representation of the underlying structure. Structural representations also allow alignment across problem instances, which can afford fast inferences via vector-based processing[36]. In this study, we set out to test whether interindividual differences in relational coding in the hippocampal system, associated with its function in forming and maintaining cognitive maps, predict fluid intelligence. To this end, we conducted comprehensive standardized cognitive tests as well as functional magnetic resonance imaging (fMRI) of relational and non-relational mnemonic tasks in a large sample of participants and evaluated hippocampal processing with respect to intelligence factors. Our findings reveal a significant relation between map-like relational encoding of task information in the hippocampus and general reasoning ability g_f.

Results

Performance in an object location memory paradigm

As a neural measure of map-like relational encoding, we assessed the hippocampal representation of two-dimensional (2D) distances between objects previously encountered in an object location memory

task (Figure 1A-B). We opted for this measure as it is well established in spatial and non-spatial tasks[22,29,37–39], well-controlled, and the task difficulty and duration of learning in the spatial paradigm are sufficiently moderate to ensure the same experimental procedure and low dropout rates for testing a large sample of participants. In the object location memory task, participants learned to replace six cued objects in a circular arena. Each placement was followed by feedback and participants navigated to the true object location (Figure 1 A-B). Each trial began at a random starting position, and participants encoded object positions independently, without navigating between objects. By the end of the learning phase, participants knew the positions of the objects well (performance in the last 10% of trials: mean(SD)=9.36(5.26) % of max. distance; Figure 1D-E).

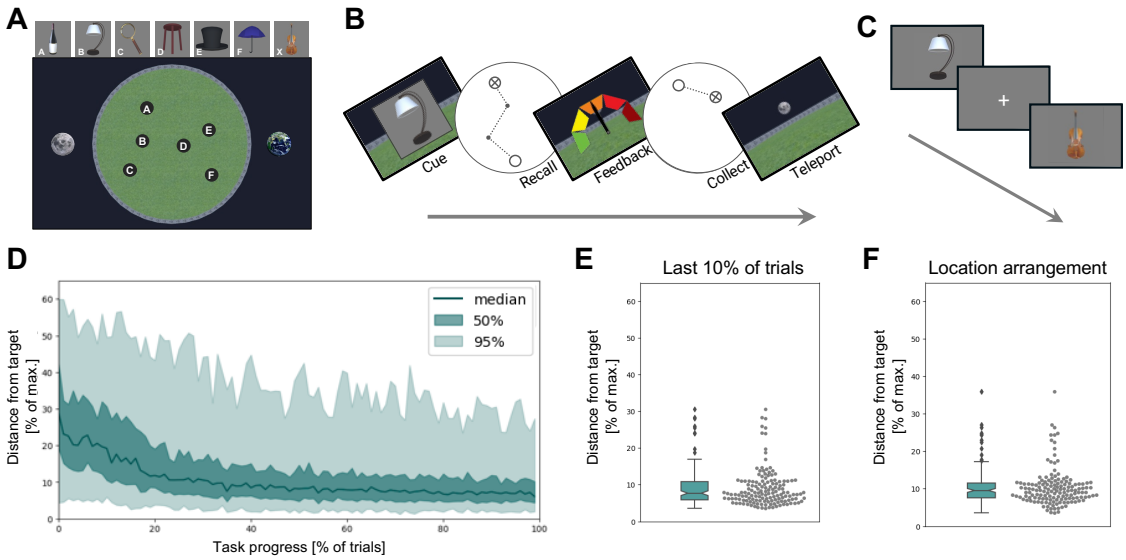

Figure 1. Performance in the object location memory paradigm. (A) Top-down view onto the virtual arena with diametrically opposed landmarks outside the arena. Each of the 6 locations A to F was associated with one object. Participants navigated the arena from a first-person perspective. (B) In each trial of the object location memory task one of the objects had to be placed in the empty arena. Subsequently, feedback was displayed and the object appeared at its true location to be collected by the participant. (C) To measure object specific neural response patterns all objects and an additional catch object were presented individually in the object picture viewing task. (D) Replacement errors during the object location memory task decreased and plateaued over the course of the task. Green line indicates the median error, shaded areas indicate the range covering 50% (dark green) and 95% (light green) of all data points. (E) Replacement error in the last 10% of trials was on average (SD) 9.1(5.04)% of the maximum possible distance (diameter). (F) Replacement errors in a post-learning object arrangement task, in which object locations had to be arranged from a top-down perspective, show a similar distribution.

In object viewing blocks that occurred before and after the location learning task to measure hippocampal representations (Figure 1C), participants detected the presence of target objects with high accuracy, ensuring attention to the stimuli (percentage correct: mean(SD)=98.6(2.7)%). In a final

assessment at the end of the experiment, participants recreated the object location arrangement from a top-down view of the arena. The arranged locations were close to the true locations (mean(SD)=9.13(2.80)% of max. distance; Figure 1F), and the inter-object distances between the arranged locations were highly correlated with the true inter-object distances (Pearson correlation: mean(SD)=.88(.14)).

Map-like relational encoding in the hippocampus scales with g_f

We hypothesized that the factor g_f, the general ability for reasoning or fluid intelligence, is related to the extent to which individuals encode the objects in the task in a map-like, relational manner. The latter would be reflected in a representation of the 2D distances between objects, with higher multivoxel pattern similarity for closer objects (Figure 2A). To test this, we measured hippocampal representations of the objects in object viewing blocks before (baseline) and after location-learning to account for pre-existing object similarities unrelated to the information in the task. Corresponding to group-level analyses in previous reports, we found a significant representation of the 2D inter-object relations in the hippocampus. Hippocampal pattern similarity between objects correlated with the 2D-distance model matrix from the object location arrangement task (right hippocampus: $t(54)=2.47$, $p=.008$; Figure 2B; as the main hypothesis posited a correlation with g_f, and to match the sample size of previous studies, the initial group-level test was performed on the upper third of the g_f-ranked participants). Next, we assessed interindividual differences in these hippocampal mnemonic representations. Corresponding to our main hypothesis, we observed a significant positive correlation between the strength of the hippocampal 2D distance representation and g_f across participants. Higher g_f scores were associated with a stronger correspondence between hippocampal pattern similarity and the 2D distance model (right hippocampus: $r(147)=.17$, $p=.021$; Figure 2C). This correlation also persisted after accounting for behavioral performance in the object location memory, the location arrangement, and the picture viewing tasks ($r_{partial}(147)=.20$, $p=.008$), indicating that variance in the neural representational format is not simply due to variance in the availability of task information. Figure 2D visualizes the neural representational space for the upper and lower thirds of the g_f distribution using Multidimensional Scaling (MDS), depicting lower congruency with the true location arrangement in the lower g_f group. It could be considered whether representations in the lower g_f spectrum reflect noisier, yet map-like relational representations or whether a different way of encoding leads to systematic errors. Here, we found no evidence in support of a map-like representation in the lower g_f group ($t(57)=-.88$, $p=.808$).

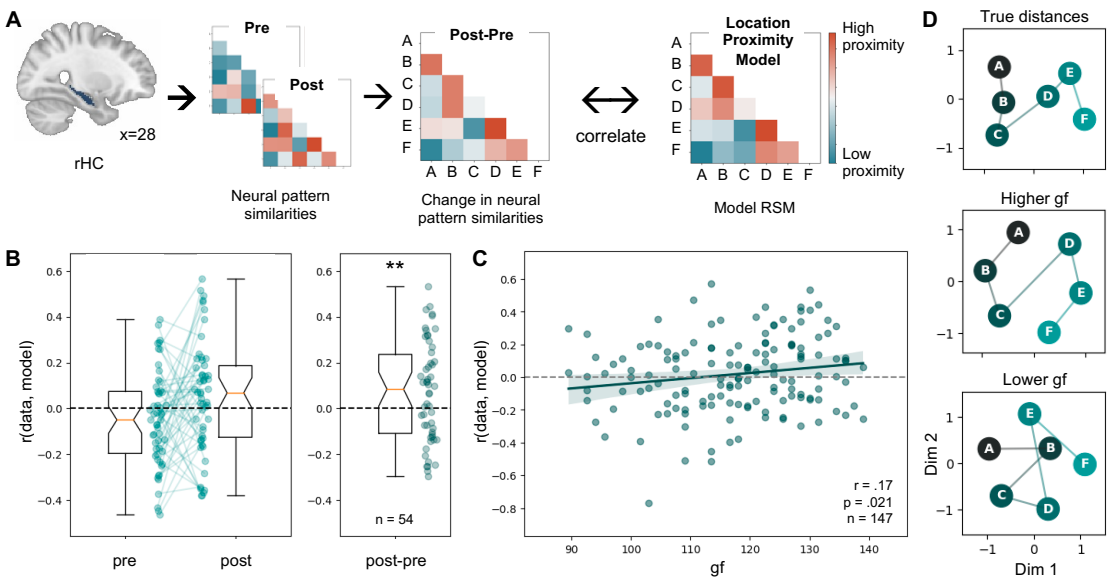

Figure 2. Map-like relational representation in the hippocampus scales with fluid intelligence (g_f). (A) Measuring the representation of 2D distances between objects in the object location memory task. The baseline representational similarity matrix (RSM) acquired before the location learning phase was subtracted from the post-learning RSM to remove similarities unrelated to the information in the task. The resulting neural RSM was correlated with a location proximity model RSM. High correlations indicate that hippocampal response patterns to the objects became more similar for close locations and more different for far locations. (B) Significant 2D distance representation in the right hippocampus. (C) The hippocampal representation of the 2D arrangement of objects correlates positively with g_f. The green line represents a linear regression fit; the green shadow represents the 95% confidence interval. (D) MDS visualizations of the distances between true locations (top) and the average neural pattern similarities for participants in the upper (middle) and lower (bottom) third of the sample's g_f distribution.

Geometrically inconsistent responses may indicate a lack of relational integration with lower gf

The reduced overlap of hippocampal representations with 2D object-relations in the lower g_f group (Figure 2D) could reflect a lack of relational integration at encoding, which can result in geometrially inconsistent beliefs about single object pairs that violate the constraints of a 2D space. To explore this hypothesis, we first compared hippocampal representational dissimilarity matrices (RDMs) between the lower and higher g_f group with respect to how well the neural similarity relationships across object pairs could be expressed as a 2D arrangement. Specifically, we assessed how much variance was explained by the first two Principal Components of a Principal Component Analysis (PCA). Indeed, the first two PCs explained significantly more variance in the higher g_f compared to the lower g_f group ($t(54,55)=1.74$, $p=.042$, Figure 3A). This effect was specific to the combined explained variance of the first two PCs, but did not occur for each PC individually (PC1: $t(54,55)=1.37$, $p=.087$; PC2: $t(54,55)=0.80$, $p=.212$). Next, we asked whether behavior was congruent with this explanation. In addition to the location arrangement task, participants had also performed a pairwise-distance

estimation task (Figure 3B-C). This task is particularly interesting in this context, as it queries object-to-object relations individually and, unlike the location arrangement task, thus does not require to adhere to geometric constraints of 2D space and permits estimates that are overall geometrically inconsistent. Congruent with the neural analysis, also in the behavioral RDMs more variance was explained by the first two PCs in the higher g_f group compared to the lower g_f group ($t(54,55)=3.58$, $p<.001$; cf. Figure 3D). Importantly, while performance differed between both groups ($t(54,55)=3.62$, $p<.001$), participants overall performed well in this task (mean(SD)=.70(.20)). We also replicated the group difference in variance explained by the first two PCs after accounting for the coherence of responses per object pair ($t(54,55)=2.52$, $p=.007$; cf. Figure 3E), assuring it is not based on random responses. Taken together, participants in the lower g_f group did not respond randomly but tended to show responses that are more inconsistent with a 2D geometry of the object relations than participants in the higher g_f group. This may suggest lapses in relational integration at encoding of individual relations, and possibly alternative memorization strategies, in the lower range of the g_f spectrum.

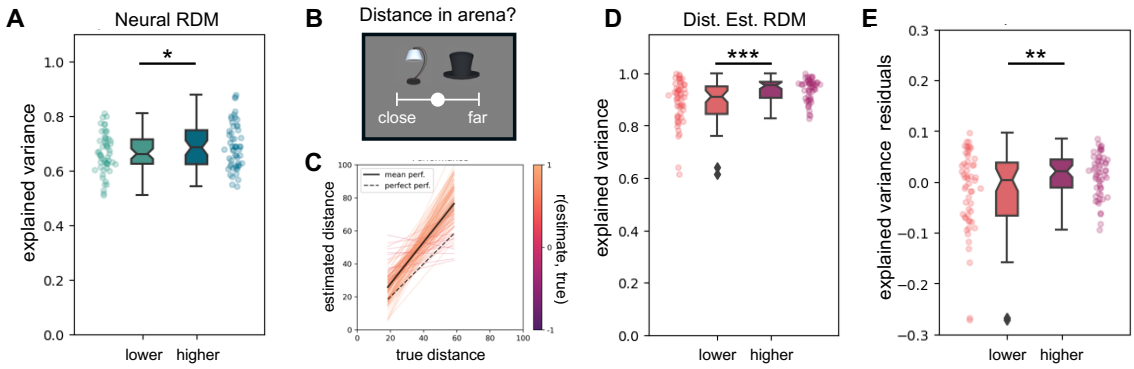

Figure 3. Neural and behavioral RDMs are less consistent with a 2D geometry in the lower g_f group. We assessed how much variance of the neural and behavioral representations of object-to-object relations is explained by the first two principal components (PCs) as an indicator of potential inconsistencies with a 2D arrangement of all relations. The more variance of the RDMs is explained by additional components, the less do the represented pairwise distances reflect a 2D arrangement. (A) In the neural RDMs (Figure 2A), the first two PCs explain more variance in the higher g_f group than in the lower g_f group. (B) In a pairwise distance estimation task participants rated the distance between two presented objects using a slider from 0 (same location) to 100 (opposite ends of arena). (C) The correlations between estimated distance and corresponding true distance are overall high (mean(SD)=.70(.20)). (D) Congruent with (A), also in the behavioral RDMs of the pairwise distance estimation task, the first two PCs explain more variance in the higher g_f compared to the lower g_f group. (E) This is also the case after accounting for response coherence by comparing the groups on the residuals from regressing response coherence on explained variance by the first two PCs.

The relation between intelligence and hippocampal representations is specific to g_f and relational processing

To further evaluate whether the relationship between intelligence and hippocampal processing is specific to the relational component of the representation and particularly related to g_f, we ran two analyses in which we replaced either the neural score by a non-relational mnemonic response or the intelligence score by a less g-loaded second-order factor (c.f. Figure 4A). First, if the relational component of processing indicated by our neural measure underlies the relation to g_f, we would not expect a significant relation to a non-relational mnemonic response in the hippocampus such as an item recognition signal. To test this, participants performed an item recognition memory task during fMRI in which they indicated whether objects had been presented in a preceding encoding task (Figure 4C). As established in previous studies[40–42], we observed a significantly higher BOLD response in the right posterior hippocampus for correctly recognized compared to correctly rejected novel items ($t(173)=4.43$, $p<.001$; Figure 4E). This non-relational memory signal did not significantly correlate with g_f ($r(173)=.04$, $p=.320$). Second, if representational formats that integrate relations support reasoning and human cognition more broadly, they should be particularly related to g_f, which approximates g, and less related to second-order factors with lower g-loading such as processing speed (g_s). We found no significant correlation between relational map-like encoding and g_s ($r(147)=.06$, $p=.228$; Figure 4D).

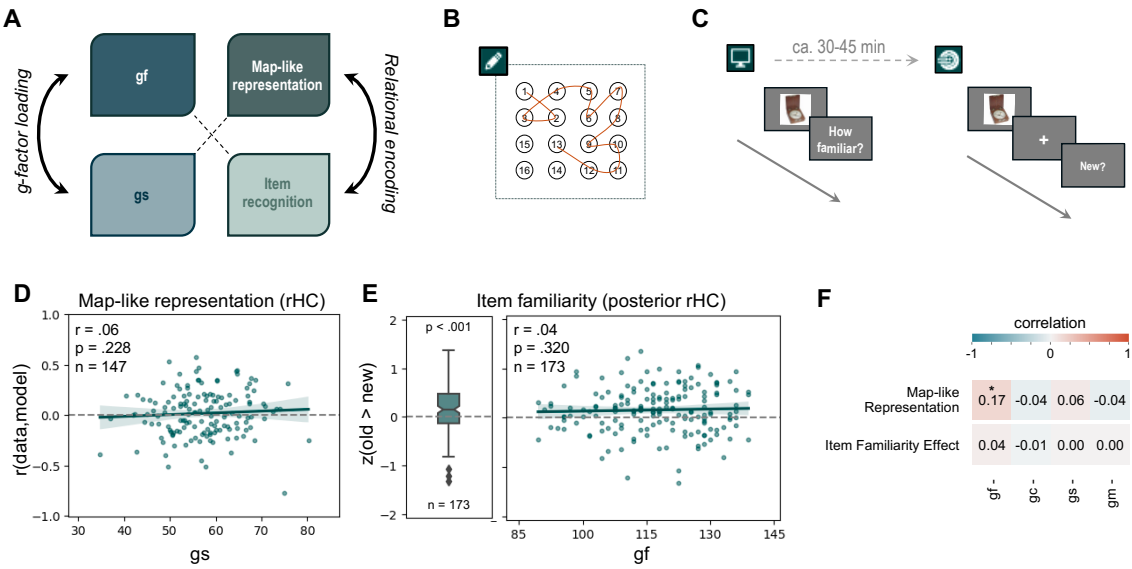

Figure 4. The relation between g_f and relational representations is specific. (A) To test whether the relationship between intelligence and hippocampal processing is specific to relational processing and g_f, we either replaced g_f by a less g-related factor (processing speed g_s) or the neural score by a non-relational memory signal in the hippocampus (item recognition response) (B) g_s was derived from a trail-making test and (C) the hippocampal item recognition signal was measured in an item familiarization paradigm, in which participants indicated whether an item was old or new with respect to a preceding implicit encoding phase (D) The hippocampal distance representation (cf. Figure 2) did not correlate with g_s. (E) g_f did not correlate with non-relational item recognition memory signals

in the hippocampus. (F) Correlations of intelligence test scores (g_f, g_s, further crystallized intelligence (g_c) and memory (g_m) to neural processing measures.

Discussion

How do individuals differ in the way they process information with respect to their general cognitive ability? We set out to examine general factors underlying multi-task performance, in particular general reasoning ability (g_f), for the first time, on the level of neural representational properties. We find that interindividual differences in relational encoding in the hippocampal system, related to its function in forming and maintaining a cognitive map, correlate with g_f. We further show that the relation between intelligence and hippocampal processing is specific to the relational component and to g_f as a highly g-loaded factor.

We use the representation of 2D distances between locations in physical or abstract spaces as an established measure of map-like relational representations (for space[37,38] and concepts[22,29,39]). We found that individual model fits positively correlated with the factor scores for general reasoning ability (g_f; measured via a standardized test-battery), suggesting that relational encoding was more pronounced with higher g_f scores. In particular, this correlation persisted when accounting for interindividual differences in task performance (i.e., in the object location memory and object viewing task), indicating that variance in the neural representational format is not just due to variance in the availability of task information (e.g. if locations were not known, or the object cues not attended). If relational coding in the hippocampus can be considered a g_f/g-related mechanism, there might in principle also be a relation to task performance. However, the effect of interindividual differences in a processing mechanism on task performance can depend on the task difficulty. Here, we designed the task to be sufficiently easy for most participants to finish learning of the object locations within a given time and thus decouple potential interindividual differences in neural coding strategies from the mere availability of information. Apart from integrating object relations in a unified map-like representation, the object positions might also be learned via alternate navigation and memorization strategies (e.g. in reference to an external landmark or remembering a specific path from the center to each object), particularly as objects were encoded and retrieved individually without navigating between them or seeing them all at once. Participants with less 2D-like neural RDMs in the lower g_f group might have used such alternate strategies. Relatedly, interindividual differences in the use of navigation strategies have been shown to relate to spatial and verbal working memory[43].

To better characterize the interindividual processing differences along the g_f scale, we compared the representational geometry of two extreme groups corresponding to the upper and lower third of the g_f distribution. If the relation between map-like representations in the hippocampus and g_f reflects specifically an effect of relational processing, then lower g_f values should be related to lapses in relational integration of separately encoded object locations. Latter can result in representations of single relations that violate the constraints of a 2D geometry. Consistent with this, both the neural and behavioral representations of object-similarities were less well explained by the first two principal components in a PCA in the lower compared to the higher g_f group. Together, these results provide evidence for a link between neural coding of relational structure and g_f.

A further analysis confirmed that the correlation between intelligence and hippocampal representations is specific to relational integration, as it was absent for non-relational item recognition signals in the hippocampus. Likewise, the relation between map-like relational representations in the hippocampus and intelligence factors was most pronounced for g_f, the factor with the highest g-loading among all broad abilities[5,14] while they did not significantly correlate with a less g-loaded factor (g_s). In sum, our results are congruent with a role of neural mechanisms related to relational reasoning in human intelligence.

We focused on the hippocampal-entorhinal system due to its central role in relational processing, organizing spatial and non-spatial relations into domain-general cognitive maps[27,44] affording inference and generalization. In particular, modelling suggests that firing properties of entorhinal cells such as grid cells reflect abstract (specifics-independent) coding of relational structure, which allows its flexible conjunction with the specifics of new situations by hippocampal place cells[28]. It was further observed that the hippocampus can represent task variables in a disentangled geometry, which is optimal for cross-context generalization[45] and can be formed immediately after instruction[46] and swiftly adapt to new goals or contexts[47]. The present findings may be consolidated by further investigations of (the hippocampal role in) structural generalization in humans with respect to intelligence.

Further, the hippocampal-entorhinal system serves as a multimodal convergence hub[48–50], well positioned to interact with domain-specific regions to provide a general processing factor across diverse tasks. Prior neuroimaging research on intelligence has largely highlighted relations of intelligence to the anatomical properties and overall activation levels in parietal and frontal cortices[11–13]. Interestingly, parieto-frontal regions are functionally coupled with the hippocampus, and have shown similar[51,52] as well as complementary[53] representations of relational structure in spatial and nonspatial tasks[54,55].

Our results suggest a specific link between map-like relational representations of task information in the hippocampus and interindividual differences in general a reasoning ability, an approximation of the general factor of intelligence. This may indicate that the general factor is partly related to the degree to which individuals tend to extract relational structure from experiences, which is thought to provide the basis for flexible reasoning[18]. More generally the approach showcases how psychometric insights about cognitive factors might help to probe the general relevance of neural processing mechanism - here hippocampal representations subsumed under the framework of cognitive maps - for cognition.

Acknowledgements

S.T. is supported by a Minerva Fast Track Fellowship of the Max Planck Society. R.T. is enrolled in the Max Planck School of Cognition. We thank all research assistants and technical support staff who supported the data acquisition. We thank Alexander Eperon and Max Hinrichs for sharing code, and Volker Reisner, Misun Kim, and Kim-Laura Speck for technical help on the VR task.

Author Contributions

S.T. conceived the study. R.T. and S.T. designed and implemented the experimental tasks. R.T. lead the data acquisition. R.T. and S.T. analyzed the data. R.T. and S.T. jointly wrote the manuscript.

References

1. VandenBos, G.R. ed. (2015). APA dictionary of psychology (2nd ed.) (American Psychological Association) https://doi.org/10.1037/14646-000.
2. Carroll, J.B. (1993). Human cognitive abilities: A survey of factor-analytic studies (Cambridge University Press).
3. Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. Intelligence *35*, 401–426. https://doi.org/10.1016/j.intell.2006.09.004.
4. Calvin, C.M., Batty, G.D., Der, G., Brett, C.E., Taylor, A., Pattie, A., Cukic, I., and Deary, I.J. (2017). Childhood intelligence in relation to major causes of death in 68 year follow-up: Prospective population study. BMJ *357*, 1–14. https://doi.org/10.1136/bmj.j2708.
5. Caemmerer, J.M., Keith, T.Z., and Reynolds, M.R. (2020). Beyond individual intelligence tests: Application of Cattell-Horn-Carroll Theory. Intelligence *79*, 101433. https://doi.org/10.1016/j.intell.2020.101433.
6. Pietschnig, J., Penke, L., Wicherts, J.M., Zeiler, M., and Voracek, M. (2015). Meta-analysis of associations between human brain volume and intelligence differences: How strong are they and what do they mean? Neurosci. Biobehav. Rev. *57*, 411–432. https://doi.org/10.1016/j.neubiorev.2015.09.017.
7. Penke, L., Maniega, S.M., Bastin, M.E., Valdés Hernández, M.C., Murray, C., Royle, N.A., Starr, J.M., Wardlaw, J.M., and Deary, I.J. (2012). Brain white matter tract integrity as a neural foundation for general intelligence. Mol. Psychiatry *17*, 1026–1030. https://doi.org/10.1038/mp.2012.66.
8. Genç, E., Fraenz, C., Schlüter, C., Friedrich, P., Hossiep, R., Voelkle, M.C., Ling, J.M., Güntürkün, O., and Jung, R.E. (2018). Diffusion markers of dendritic density and arborization in gray matter predict differences in intelligence. Nat. Commun. *9*. https://doi.org/10.1038/s41467-018-04268-8.
9. Van Den Heuvel, M.P., Stam, C.J., Kahn, R.S., and Hulshoff Pol, H.E. (2009). Efficiency of functional brain networks and intellectual performance. J. Neurosci. *29*, 7619–7624. https://doi.org/10.1523/JNEUROSCI.1443-09.2009.
10. Girn, M., Mills, C., and Christoff, K. (2019). Linking brain network reconfiguration and intelligence: Are we there yet? Trends Neurosci. Educ. *15*, 62–70. https://doi.org/10.1016/j.tine.2019.04.001.
11. Jung, R.E., and Haier, R.J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: Converging neuroimaging evidence. Behav. Brain Sci. *30*, 135–154. https://doi.org/10.1017/S0140525X07001185.
12. Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends Cogn. Sci. *14*, 172–179. https://doi.org/10.1016/j.tics.2010.01.004.
13. Duncan, J. (2013). The Structure of Cognition: Attentional Episodes in Mind and Brain. Neuron *80*, 35–50. https://doi.org/10.1016/j.neuron.2013.09.015.
14. Arendasy, M.E., Hergovich, A., and Sommer, M. (2008). Investigating the 'g'-saturation of various stratum-two factors using automatic item generation. Intelligence *36*, 574–583.

https://doi.org/10.1016/j.intell.2007.11.005.

15. Gentner, D., and Smith, L. (2012). Analogical Reasoning 2nd ed. (Elsevier Inc.) https://doi.org/10.1016/B978-0-12-375000-6.00022-7.
16. Gick, M.L., and Holyoak, K.J. (1983). Schema induction and analogical transfer. Cogn. Psychol. *15*, 1–38. https://doi.org/10.1016/0010-0285(83)90002-6.
17. Richland, L.E., Zur, O., and Holyoak, K.J. (2007). Cognitive supports for analogies in the mathematics classroom. Science (80-.). *316*, 1128–1129. https://doi.org/10.1126/science.1142103.
18. Theves, S. (2024). Thinking as analogy-making: towards a neural process account of general intelligence. Manuscript under review.
19. Tavares, R.M., Mendelsohn, A., Grossman, Y., Williams, C.H., Shapiro, M., Trope, Y., and Schiller, D. (2015). A Map for Social Navigation in the Human Brain. Neuron *87*, 231–243. https://doi.org/10.1016/j.neuron.2015.06.011.
20. Aronov, D., Nevers, R., and Tank, D.W. (2017). Mapping of a non-spatial dimension by the hippocampal-entorhinal circuit. Nature *543*, 719–722. https://doi.org/10.1038/nature21692.
21. Bao, X., Gjorgieva, E., Shanahan, L.K., Howard, J.D., Kahnt, T., and Gottfried, J.A. (2019). Grid-like Neural Representations Support Olfactory Navigation of a Two-Dimensional Odor Space. Neuron *102*, 1066-1075.e5. https://doi.org/10.1016/j.neuron.2019.03.034.
22. Theves, S., Fernández, G., and Doeller, C.F. (2019). The Hippocampus Encodes Distances in Multidimensional Feature Space. Curr. Biol. *29*, 1226-1231.e3. https://doi.org/10.1016/j.cub.2019.02.035.
23. Park, S.A., Miller, D.S., and Boorman, E.D. (2021). Inferences on a multidimensional social hierarchy use a grid-like code. Nat. Neurosci. *24*, 1292–1301. https://doi.org/10.1038/s41593-021-00916-3.
24. Samborska, V., Butler, J.L., Walton, M.E., Behrens, T.E.J., and Akam, T. (2022). Complementary task representations in hippocampus and prefrontal cortex for generalizing the structure of problems. Nat. Neurosci. *25*, 1314–1326. https://doi.org/10.1038/s41593-022-01149-8.
25. Gauthier, J.L., and Tank, D.W. (2018). A Dedicated Population for Reward Coding in the Hippocampus. Neuron *99*, 179-193.e7. https://doi.org/10.1016/j.neuron.2018.06.008.
26. Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., and Tanila, H. (1999). The Hippocampus, Memory, and Place Cells: Is It Spatial Memory or a Memory Space? Neuron *23*, 209–226.
27. Behrens, T.E.J., Muller, T.H., Whittington, J.C.R., Mark, S., Baram, A.B., Stachenfeld, K.L., and Kurth-Nelson, Z. (2018). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. Neuron *100*, 490–509. https://doi.org/10.1016/j.neuron.2018.10.002.
28. Whittington, J.C.R., Muller, T.H., Mark, S., Chen, G., Barry, C., Burgess, N., and Behrens, T.E.J. (2020). The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the Hippocampal Formation. Cell *183*, 1249-1263.e23. https://doi.org/10.1016/j.cell.2020.10.024.

29. Theves, S., Fernández, G., and Doeller, C.F. (2020). The Hippocampus Maps Concept Space, Not Feature Space. J. Neurosci. *40*, 7318–7325. https://doi.org/10.1523/JNEUROSCI.0494-20.2020.
30. Park, S.A., Zolfaghar, M., Russin, J., Miller, D.S., O'Reilly, R., and Boorman, E.D. (2023). The geometry of cognitive maps under dynamic cognitive control. bioRxiv. https://doi.org/10.32470/ccn.2022.1023-0.
31. Muhle-Karbe, P.S., Sheahan, H., Pezzulo, G., Spiers, H.J., Chien, S., Schuck, N.W., and Summerfield, C. (2023). Goal-seeking compresses neural codes for space in the human hippocampus and orbitofrontal cortex. Neuron *111*, 3885-3899.e6. https://doi.org/10.1016/j.neuron.2023.08.021.
32. Bunsey, M., and Eichenbaum, H. (1996). Conservation of hippocampal memory function in rats and humans. Nature *379*, 255–257. https://doi.org/10.1038/379255a0.
33. Shohamy, D., and Wagner, A.D. (2008). Integrating Memories in the Human Brain: Hippocampal-Midbrain Encoding of Overlapping Events. Neuron *60*, 378–389. https://doi.org/10.1016/j.neuron.2008.09.023.
34. Zeithamova, D., and Preston, A.R. (2010). Flexible memories: Differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. J. Neurosci. *30*, 14676–14684. https://doi.org/10.1523/JNEUROSCI.3250-10.2010.
35. Schäfer, T.A.J., Thalmann, M., Schulz, E., Doeller, C.F., and Theves, S. (2024). The hippocampus supports interpolation into new states during category abstraction. bioRxiv Neurosci., 1–32.
36. Morton, N.W., Schlichting, M.L., and Preston, A.R. (2020). Representations of common event structure in medial temporal lobe and frontoparietal cortex support efficient inference. Proc. Natl. Acad. Sci. U. S. A. *117*, 29338–29345. https://doi.org/10.1073/pnas.1912338117.
37. Morgan, L.K., MacEvoy, S.P., Aguirre, G.K., and Epstein, R.A. (2011). Distances between real-world locations are represented in the human hippocampus. J. Neurosci. *31*, 1238–1245. https://doi.org/10.1523/JNEUROSCI.4667-10.2011.
38. Deuker, L., Bellmund, J.L., Navarro Schröder, T., and Doeller, C.F. (2016). An event map of memory space in the hippocampus. Elife *5*, 1–26. https://doi.org/10.7554/eLife.16534.
39. Park, S.A., Miller, D.S., Nili, H., Ranganath, C., and Boorman, E.D. (2020). Map Making: Constructing, Combining, and Inferring on Abstract Cognitive Maps. Neuron *107*, 1226-1238.e8. https://doi.org/10.1016/j.neuron.2020.06.030.
40. Strange, B.A., Fletcher, P.C., Henson, R.N., Friston, K.J., and Dolan, R.J. (1999). Segregating the functions of human hippocampus. Proc. Natl. Acad. Sci. U. S. A. *96*, 4034–4039. https://doi.org/10.1073/pnas.96.7.4034.
41. Strange, B.A., and Dolan, R.J. (2001). Adaptive anterior hippocampal responses to oddball stimuli. Hippocampus *11*, 690–698. https://doi.org/10.1002/hipo.1084.
42. Daselaar, S.M., Fleck, M.S., and Cabeza, R. (2006). Triple dissociation in the medial temporal lobes: Recollection, familiarity, and novelty. J. Neurophysiol. *96*, 1902–1911. https://doi.org/10.1152/jn.01029.2005.

43. Weisberg, S.M., and Newcombe, N.S. (2015). How Do (Some) People Make a Cognitive Map? Routes, Places, and Working Memory. J. Exp. Psychol. Learn. Mem. Cogn. *42*, 768–785.
44. Eichenbaum, H. (1999). The hippocampus and mechanisms of declarative memory. Behav. Brain Res. *103*, 123–133. https://doi.org/10.1016/S0166-4328(99)00044-3.
45. Bernardi, S., Benna, M.K., Rigotti, M., Munuera, J., Fusi, S., and Salzman, C.D. (2020). The Geometry of Abstraction in the Hippocampus and Prefrontal Cortex. Cell *183*, 954-967.e21. https://doi.org/10.1016/j.cell.2020.09.031.
46. Courellis, H.S., Mixha, J., Cardenas, A.R., Kimmel, D., Reed, C.M., Valiante, T.A., Salzman, C.D., Mamelak, A.N., Adolphs, R., Fusi, S., et al. (2024). Abstract representations emerge in human hippocampal neurons during inference behavior. Nature.
47. Sarel, A., Palgi, S., Blum, D., Aljadeff, J., Las, L., and Ulanovsky, N. (2022). Natural switches in behaviour rapidly modulate hippocampal coding. Nature *609*, 119–127. https://doi.org/10.1038/s41586-022-05112-2.
48. Manns, J.R., and Eichenbaum, H. (2006). Evolution of Declarative Memory. Hippocampus *16*, 795–808. https://doi.org/10.1002/hipo.
49. Lavenex, P., and Amaral, D.G. (2000). Hippocampal-neocortical interaction: A hierarchy of associativity. Hippocampus *10*, 420–430. https://doi.org/10.1002/1098-1063(2000)10:4<420::AID-HIPO8>3.0.CO;2-5.
50. Reznik, D., Trampel, R., Weiskopf, N., Witter, M.P., and Doeller, C.F. (2023). Dissociating distinct cortical networks associated with subregions of the human medial temporal lobe using precision neuroimaging. Neuron *111*, 2756-2772.e7. https://doi.org/10.1016/j.neuron.2023.05.029.
51. Doeller, C.F., Barry, C., and Burgess, N. (2010). Evidence for grid cells in a human memory network. Nature *463*, 657–661. https://doi.org/10.1038/nature08704.
52. Constantinescu, A.O., O'Reilly, J.X., and Behrens, T.E.J. (2016). Organizing conceptual knowledge in humans with a gridlike code. Science (80-.). *352*, 1464–1468. https://doi.org/10.1126/science.aaf0941.
53. Summerfield, C., Luyckx, F., and Sheahan, H. (2020). Structure learning and the posterior parietal cortex. Prog. Neurobiol. *184*. https://doi.org/10.1016/j.pneurobio.2019.101717.
54. Farzanfar, D., Spiers, H.J., Moscovitch, M., and Rosenbaum, R.S. (2022). From cognitive maps to spatial schemas. Nat. Rev. Neurosci. https://doi.org/10.1038/s41583-022-00655-9.
55. Kaplan, R., Schuck, N.W., and Doeller, C.F. (2017). The Role of Mental Maps in Decision-Making. Trends Neurosci. *40*, 256–259. https://doi.org/10.1016/j.tins.2017.03.002.
56. Amthauer, R., Brocke, B., Liepmann, D., and Beauducel, A. (2007). I-S-T 2000 R: Intelligenz-Struktur-Test 2000 R. Hogrefe.
57. Oswald, W.D. (2016). Der Zahlen-Verbindungs-Test (ZVT). Hogrefe.
58. Brady, T.F., Konkle, T., Alvarez, G.A., and Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. Proc. Natl. Acad. Sci. U. S. A. *105*, 14325–14329. https://doi.org/10.1073/pnas.0803390105.

Methods

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

We recruited 179 healthy adults (90 female, 86 male, 3 diverse; mean age(SD)=26.49(4.85)) who were right-handed, had normal or corrected-to-normal sight, spoke German natively or learned it before age 7, and held a high school diploma or equivalent. Participants provided written informed consent before participation. This study was approved by the ethics commission of the Medical Faculty of the University of Leipzig and adhered to the Declaration of Helsinki.

METHOD DETAILS

Experimental procedure

Participants completed standardized cognitive test batteries to assess intelligence factor scores on a first day, and participated in computerized tasks and MRI recordings on two other days (see below for a detailed description of the tasks).

Cognitive Tests

Participants completed all modules of the Intelligence-Structure-Test (I-S-T 2000 R), a German, standardized intelligence test battery [56], measuring fluid intelligence (g_f) and crystallized intelligence (g_c) across verbal, numerical, and figural domains, as well as memory. Processing speed was assessed via the Zahlen-Verbindungs-Test (ZVT), a standardized trail making test [57].

Object Location Memory Task

The navigation tasks using the circular arena were built using Unity Game Engine (version 2018.1.9f2). All other tasks were implemented using PsychoPy (version 2021.1.3).

For all spatial navigation tasks, participants controlled a virtual character in a circular arena from a first-person perspective on a screen. Two distal landmarks were projected at infinite distance, and a boundary along the arena perimeter served as orientation cues. During familiarization, participants collected traffic cones until they felt comfortable handling the buttons. Subsequently, they learned 6 object locations. The object-location associations were randomized across participants. Participants collected each of the objects once at its true location before the object location memory task began. In each trial participants were presented a target object, which they replaced as close as possible to its true location in the empty arena. Participants received feedback about their replacement error (distance to the true target location). Then the object appeared at its true location, where participants collected it to start the next trial. After collection, participants' characters were teleported to a new start location. Participants completed 3 learning task runs (up to 72 trials/run; 20.5 min max./run).To illustrate the median learning curve, we applied linear interpolation to resample the varying number of drop errors for each participant (n_{mean}(SD)=177.0(28.5)) into a common range of 50 samples each representing a 2% increment in task progress. For each step, we computed the median value as well as the percentiles including 50% and 95% of all datapoints to capture central and broader distribution trends. To assess

learning outcomes, we calculated the average drop error of all trials across the final 10% of trials for each participant (n_{mean}(SD)=18.0(2.8)).

Object Picture Viewing Task

We recorded fMRI data during a pre- and post-learning object picture viewing task from 179 participants. The pre- and post-learning runs used identical stimulus sequences. Participants viewed 6 objects previously encountered in the object location memory task along with one additional catch object. To probe attention, participants were instructed to press a button whenever a catch object appeared. Each object image was presented for 1.5 sec, followed by a fixation cross for an average duration of 4 sec (onset jittered using truncated exponential function). All 7x7 possible image-to-image transitions occurred once before repeating, resulting in 196 trials including 14.14% catch trials. Each run lasted approximately 18 min. Participants were excluded from fMRI data analysis if they exhibited excessive motion (mean framewise displacement > .4), responded incorrectly in more than 75% of catch trials in any run, or had an average response time slower 1.5 s in any run, as these criteria indicated insufficient attention to the stimuli.

Pairwise Object Distance Estimation Task

Following the Object Picture Viewing Task, participants completed behavioral post-learning tasks to assess their knowledge of the object locations in greater detail. In the distance estimation task, each possible combination of object pairs was presented twice, with object positions (left/right) counterbalanced across trials, in a pseudorandom sequence. Participants estimated the distance between objects on a scale ranging from 0 (indicating the objects share the same location in the arena) to 100 (indicating the objects were located at opposite ends of the arena's diagonal). To identify participants who may have responded randomly, we simulated 100,000 datasets with object distance estimates drawn from a uniform random distribution. 2 participants were excluded because their response coherence (mean absolute difference between same-pair distance estimates) did not significantly differ from the distribution of random values, i.e. their coherence was not lower than the 99th percentile of the random distribution. Performance in the distance estimation task is measured as the Pearson correlation between participants' estimated distances and corresponding true distance.

Object Location Arrangement Task

Finally, participants completed an object location arrangement task which provided a top-down view onto the empty arena and landmarks. In Phase 1, participants arranged 6 grey circles to reflect the object positions. The circles initially appeared at the arena's center and could be moved via drag and drop. In Phase 2, the locations were highlighted one by one and participants assigned an object to each location from a dropdown menu listing all 6 objects. To ensure meaningful object location arrangements, we excluded participants if their arrangement did not reflect 6 unique locations. Specifically, 4 participants did not complete the task as instructed failing to remove 2 or more circles from the arena center, and 9 participants assigned the same object to multiple locations. Performance in the object

location arrangement task was determined by calculating for each participant the mean of the absolute distance of between each participant's arranged locations to the corresponding true locations.

Item Recognition Task

We collected fMRI data from 173 participants during an item recognition memory task. Note that due to technical issues, procedural errors, or participant dropouts, not all participants contributed usable data for both fMRI tasks. Participants were initially familiarized with a set of 180 unique everyday items [58] under the guise of a cover task. They were instructed to rate how often they had encountered each item or similar images in everyday life as quickly and intuitively as possible. Each item was displayed for 1 sec, followed by the rating question. Approximately 30–45 minutes later, participants completed the item recognition memory task while undergoing fMRI scanning. In this task, 60 previously seen ("old") and 60 novel ("new") items were presented in a randomized sequence lasting approximately 13 min. Each item appeared for 1 sec, followed by a fixation cross for an average of 0.5 sec, and then a question mark prompting a response within 2 seconds. Participants indicated their recognition response ("yes" or "no") and their confidence ("certain" or "uncertain") via button press. An inter-trial interval followed, during which a fixation cross was displayed for an average of 3 seconds (jittered using a truncated exponential function). The same fMRI exclusion criteria applied as in the object picture viewing task.

MRI

In the MRI scanner tasks were presented using a projector on a screen outside the scanner, which participants could view through a mirror attached to the head coil. Participants responded using MRI compatible button boxes. MRI data was recorded using a 3T Siemens MAGNETOM Skyra Connectom A (Siemens, Erlangen, Germany) using a 32-channel head coil.

Anatomical Imaging Sequence

High-resolution T1-weighted anatomical images were acquired using a Magnetization Prepared 2 Rapid Acquisition Gradient Echo (MP2RAGE) sequence. The following acquisition parameters were used: repetition time (TR) = 5 s, echo time (TE) = 2.84 ms, inversion times (TI1, TI2) = 700 ms and 2500 ms, and flip angles = 4° (first readout) and 5° (second readout). The field of view (FOV) was 256 × 240 mm^2 with 176 sagittal slices, providing an isotropic voxel resolution of 1 mm^3. Full k-space sampling was performed with a partial Fourier factor of 1. To produce the final image the INV1 and INV2 images were combined, then the final T1-weighted image was generated by multiplying the intensity-normalized INV2 image with the MP2RAGE UNI image using an AFNI based script (3dMPRAGEise, https://github.com/srikash/3dMPRAGEise). This approach enhances tissue contrast while suppressing background noise and residual artifacts.

Functional Imaging Sequence

All fMRI data was recorded using a CMRR multiband echo-planar-imaging (EPI) pulse sequence (TR=1.5s, TE=24ms, flip angle=70°, partial Fourier=.75, multiband-factor=4, phase encoding direction=AP, voxel size=2 mm^3 isotropic, FOV=204x204 mm^2, 80 slices, interleaved slice acquisition

order). The measurement volumes were angled parallel to the axis crossing the anterior and posterior commissure (ACPC). For each fMRI recording of a task run a corresponding EPI field map for estimation of the B0 field inhomogeneities was constructed based on 2 EPI reference images with opposite phase encoding direction.

Preprocessing

Raw DICOM files were converted into BIDS format using the dcm2bids package (version 2.1.7, https://github.com/UNFmontreal/Dcm2Bids).Results included in this manuscript come from preprocessing performed using *fMRIPrep* 22.0.1 (Esteban, Markiewicz, et al. (2018); Esteban, Blair, et al. (2018))

Preprocessing of B_0 inhomogeneity mappings

A B_0-nonuniformity map (or *fieldmap*) was estimated based on two (or more) echo-planar imaging (EPI) references with topup (Andersson, Skare, and Ashburner (2003); FSL 6.0.5.1:57b01774).

Anatomical data preprocessing

The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al. 2010), distributed with ANTs 2.3.3 (Avants et al. 2008, RRID:SCR_004757), and used as T1w-reference throughout the workflow. 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 (FSL 6.0.5.1:57b01774, RRID:SCR_002823, Zhang, Brady, and Smith 2001). Volume-based spatial normalization to a standard space (MNI152NLin6Asym) 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 template was selected for spatial normalization: *FSL's MNI ICBM 152 non-linear 6th 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].

Functional data preprocessing

For each of the 3 BOLD runs found per subject (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*. 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 6.0.5.1:57b01774, Jenkinson et al. 2002). The estimated *fieldmap* was then aligned with rigid-registration to the target EPI (echo-planar imaging) reference run. The field coefficients were mapped on to the reference EPI using the transform. BOLD runs were slice-time corrected to 0.702s (0.5 of slice acquisition range 0s-1.41s) using 3dTshift from

AFNI (Cox and Hyde 1997, RRID:SCR_005927). The BOLD reference was then co-registered to the T1w reference using mri_coreg (FreeSurfer) followed by flirt (FSL 6.0.5.1:57b01774, Jenkinson and Smith 2001) with the boundary-based registration (Greve and Fischl 2009) cost-function. Co-registration was configured with six degrees of freedom. 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 (absolute sum of relative motions, Power et al. (2014)) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al. (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. 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. Additional nuisance timeseries are calculated by means of principal components analysis of the signal found within a thin band (*crown*) of voxels around the edge of the brain, as proposed by (Patriat, Reynolds, and Birn 2017). 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*. 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).
Many internal operations of *fMRIPrep* use *Nilearn* 0.9.1 (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.

Regions of interest

Regions of interest (ROI) were created based on a probabilistic map of HC from the Havard-Oxford Atlas (Makris et al. 2006, Frazier et al. 2005, Desian et al. 2006, Goldstein et al. 2007). Probability maps were thresholded at 80%. Anterior and posterior ROIs were created by dividing each HC ROI along the anterior-to-posterior axis of MNI space such that the resulting ROIs contained approximately equal numbers of voxels.

First-Level Modelling of fMRI data using General Linear Models (GLMs)

fMRI data analysis was performed in MNI space. A separate GLM was constructed for each task. For each object picture viewing run the model specified one regressor per object, whereas for the item recognition memory run four regressors were included to capture the conditions of correctly and incorrectly classified *old* and *new* items. Both GLMs incorporated the same set of nuisance regressors

to account for non-task-related variability. These included regressors for button presses, six motion parameters derived from realignment, and cosine drift regressors to implement a high-pass filter with a cutoff frequency of 0.01 Hz.

QUANTIFICATION AND STATISTICAL ANALYSIS

All hypotheses were directional and corresponding test statistics were based on one-sided tests. When analyses were performed on both hemispheres, we corrected the alpha-level accordingly.

Representational similarity analysis (RSA)

Representational Similarity Analysis (RSA) was used to investigate the neural representation of 2D spatial relations between the objects in the task. Object-specific response patterns were extracted from the first-level GLM for both the pre- and post-learning runs. These patterns were used to construct Representational Similarity Matrices (RSMs) for each run. RSMs were computed by calculating the Pearson correlation of neural response patterns for each object pair. The pre-learning RSM was subtracted from the post-learning RSM for each participant to remove pre-existing similarities and capture the learning-related representational changes. Each participant's RSM was then correlated with a model RSM using Spearman's rho. The model RSM entailed the estimated distances from the location arrangement task, accounting for minor variance in the perception of the arena (e.g. some participants reported not perceiving it as perfectly circular).

Multi-Dimensional Scaling (MDS)

In order to visualize the representational spaces in individuals with low and high g_f scores, we split the data into thirds (corresponding to a g_f equal or greater than the 67.7th percentile in the higher g_f group, and a g_f equal or lower than the 33.3rd percentile in the lower g_f group). We used metric MDS to illustrate a 2D embedding that corresponds well to the multidimensional neural distance representations of the higher and lower g_f group.

Dimensionality analysis

To investigate whether neural RDMs in the low relative to the high g_f group reflect inconsistencies of relations with the constraints of a 2D arrangement, we assessed the variance explained by the first two principal components of a Principal Component Analysis (PCA) of each subject's neural RDM. Any 2D arrangement of the object set should result in an RDM in which 100% of variance is explained by the first two PCs. A larger proportion of variance explained by additional components indicates that the inter-object distances are less consistent with a 2D structure. Further we conducted this analysis on the behavioral RDMs of participants' pairwise object distance estimations. Repeated estimates per pair were averaged to construct the RDM. The effect of response coherence on the explained variance by the first two PCs was partialled out using a linear model predicting explained variance with response coherence and repeating the analysis on the residuals.

Item recognition memory

To quantify item familiarity responses in the hippocampus, we contrasted correctly classified old items and correctly classified new items ($old_{correct} > new_{correct}$) and averaged the corresponding z-scores of the ROI for each participant.

Correlation of neural response measures with g factors

For correlation analyses between neural response measures and g factors we computed Pearson correlations.