One view is that conceptual knowledge is organized using the circuitry in the medial temporal lobe (MTL) that supports spatial processing and navigation. In contrast, we find that a domain-general learning algorithm explains key findings in both spatial and conceptual domains. When the clustering model is applied to spatial navigation tasks, so-called place and grid cell-like representations emerge because of the relatively uniform distribution of possible inputs in these tasks. The same mechanism applied to conceptual tasks, where the overall space can be higher-dimensional and sampling sparser, leading to representations more aligned with human conceptual knowledge. Although the types of memory supported by the MTL are superficially dissimilar, the information processing steps appear shared. Our account suggests that the MTL uses a general-purpose algorithm to learn and organize context-relevant information in a useful format, rather than relying on navigation-specific neural circuitry.
Concepts organize experiences to enable generalization and inference. For example, a traveler encountering an unfamiliar bird species would reasonably infer the bird was born from an egg. One longstanding question is the basis for people’s abstract conceptual knowledge. One intuitive idea is that concepts ground in a more basic and concrete substrate, such as sensory-motor experience. For example, abstract concepts such as time may be represented in terms of experience of space. Relatedly, conceptual knowledge may be organized using circuitry in the medial temporal lobe (MTL) that supports navigation.

This view is supported by recent studies that find the brain’s responses to conceptual tasks parallel those previously found in spatial tasks. Place cells in the hippocampus typically have single firing fields at circumscribed locations in a spatial environment, and grid cells in the medial entorhinal cortex (mEC) display multiple regularly-spaced firing fields arranged in a hexagonal pattern covering the environment. These spatially-tuned cells in the MTL are thought to implement a spatial cognitive map for navigation, and recent work suggests these cells also represent conceptual and task spaces. One key question is whether the same brain systems and computations support concept learning, memory, and spatial navigation.

One neglected possibility is that the relation between spatial and conceptual representations has been framed backwards. Perhaps, rather than concepts grounding in the machinery of navigation, spatial concepts are a limiting case of a single, more general learning system. Such a learning system would be tasked with learning all relevant concepts, including those tied to physical space (also see refs. 14,15, and Discussion). This general learning system would support learning concepts, which are typically clumpy in that they consist of clusters of interrelated features in a high-dimensional space. For example, animals that fly tend to be small and have wings (see Fig. 1a). Not all possible combinations of features are relevant and represented. In contrast, many spatial tasks and their conceptual analogs typically involve a uniform and exhaustive sampling of all possible combinations within a low (two-) dimensional space corresponding to locations in an environment (see Fig. 1c, d).

We evaluate whether a domain general account is plausible by applying successful models of human concept learning to spatial contexts. In concept learning studies, these models learn to represent experience in terms of conceptual clusters, which are not uniformly distributed. For example, in a simple case with two clearly separable and internally coherent sets of objects, a clustering model would use one cluster to represent each concept, each of which would be centered amidst its members in representational space (e.g., Fig. 1a).

When the model is presented with a novel item, the closest cluster in representational space is activated, which signals the category membership of the item. An error-monitoring signal gauges how well an item matches this closest cluster in representational space. In these models, only the closest cluster maintains non-zero activation (winner-takes-all), so an error-monitoring signal (entropy term) monitors activation of all existing clusters which indicates how close or far away the current location was from any cluster, acting as a cluster match (or non-match signal).

These clustering representations successfully capture patterns of activity in the MTL and are in accord with the notion that the human hippocampus contains concept cells in which individual cells respond to a specific concept, much like how a cluster in a possibly high-dimensional space can encode a concept. Analogously, place cells respond to a location in a particular two-dimensional spatial context. It is important to note that clusters are abstract entities in the model, and there need not be a one-to-one mapping to single concept or place cell (e.g. a cluster can be represented by a group of place cells with similar tuning (c.f. refs. 22) — a functional mapping of multiple place cells to one cluster, and the place cell population to the whole cluster representation (Fig. 2c). Furthermore, clustering models explain how individual episodes give rise to conceptual knowledge over the course of learning, consistent with both the hippocampus’s importance in memory.

We evaluate whether the same mechanisms also offer a general understanding of place and grid cells, and their relationship to concepts.

To facilitate this evaluation, we simplified the clustering models to only include aspects necessary for this contribution. Clustering models that capture behavior on a trial-by-trial basis typically recruit a new cluster in response to a surprising error. These models also learn attention weights that accentuate task-relevant stimulus dimensions and associate clusters with behavioral responses (e.g., respond “bird”). Without loss of generality, we...
Abstract cluster representations are similar to its newest member (see Methods for full details), much updated its position in representational space to be closer (more particular, the cluster most similar to the current stimulus clusters and limiting learning to updating cluster positions. In NATURE COMMUNICATIONS, we explore how cluster representations can emerge from unsupervised learning, focusing on the case of hippocampal cells or cluster match in the case of mEC cells.

In contrast, the same model applied to an agent exploring a typical laboratory environment leads to clusters that uniformly cover the entire representational space in a hexagonal pattern (see Fig. 1c, d). In the spatial case, there is no salient structure present in the input to the model, which results in clusters covering the representational space, much like how a bunch of tennis balls dropped into a square box will self-organize into a grid-like lattice according to the mathematics of packing. In the spatial case, the clusters function in a similar way to a population of place cells that code for (i.e., discriminate) locations.

In our account, grid-like responses arise from monitoring the match (inverse error) of the clustering solution (Fig. 2a, b). In unsupervised learning, error or uncertainty is simply the inverse of how similar an item is to the best matching cluster. Notice that matching clusters in the spatial case should display a hexagonal pattern because of the hexagonal clustering pattern in representational space, resulting in canonical grid-like receptive fields (see Fig. 2b). In the conceptual case, we predict that typical grid cell firing patterns should not be observed because the clusters (i.e., place cells) do not form a hexagon pattern (Fig. 2a) in representational space. One might object that our account is inconsistent with conceptual learning brain imaging studies that find grid-like response patterns. However, these studies are consistent with the model because they follow the design principles of typical spatial studies—all feature combinations within a 2-dimensional stimulus space are sampled, which would lead to a hexagonal clustering solution (Fig. 2b).

Fig. 2 Cluster representations after learning for conceptual and spatial tasks. a Clusters clump into two groups. Thus, novel bird and mammal stimuli will strongly activate one or the other grouping, which does not lead to a grid response across possible stimuli. b In contrast, for the spatial case, clusters form a hexagonal grid which leads to a grid-like response across possible stimuli when cluster activity is monitored. c Clusters determine the receptive fields for a population of place or concept cells, and the cluster-monitoring/error-monitoring mechanism (grid or spatial cells) reflect the distribution of the clusters. Abstract cluster representations are instantiated by multiple cells in the hippocampus and medial entorhinal cortex (mEC) with similar firing fields to represent the same location (or concept) in the case of hippocampal cells or cluster match in the case of mEC cells.

Clustering solutions match grid patterns in mEC. To relate our account to typical spatial studies, we simulated an agent moving through its environment as in a free-foraging rodent experiment. As expected, learning led to clusters forming a hexagonal pattern (see examples in Fig. 3a, b, left). To assess this quantitatively, we computed the spatial autocorrelograms of the cluster activation maps (Fig. 3a, b, right) to obtain the grid score, which reflects the degree six-fold hexagonal symmetry in the cluster activation pattern across trials (see Methods). We computed grid scores for each time bin during learning and found that grid scores tended to increase over learning in both the square (see Fig. 3c for examples and Fig. S1 for all conditions; mean slope: 0.0044, bootstrap CIs: [0.0040, 0.0048]) and circular environment (mean slope: 0.0042, bootstrap CIs: [0.0038, 0.0046]; see Tables S1 and S2).

Following learning, we evaluated the gridness of the clustering solution (see examples in Fig. 3a, b and Figs. S2 and S3). A substantial proportion of simulations satisfied the criterion for grid-like organization, with 45.3% in the square and 38.6% in the circular environment, which closely match the proportions in empirical results (45% and 38%, respectively; see Supplementary Note 1). The average grid score in both the square environment (mean: 0.277, bootstrap CIs [0.273, 0.0.280]) and circular environment (mean: 0.313, bootstrap CIs [0.309, 0.318]) were greater than zero; see Tables S3 and S4).

Clustering solutions match grid patterns in mEC. To relate our account to typical spatial studies, we simulated an agent moving through its environment as in a free-foraging rodent experiment. As expected, learning led to clusters forming a hexagonal pattern (see examples in Fig. 3a, b, left). To assess this quantitatively, we computed the spatial autocorrelograms of the cluster activation maps (Fig. 3a, b, right) to obtain the grid score, which reflects the degree six-fold hexagonal symmetry in the cluster activation pattern across trials (see Methods). We computed grid scores for each time bin during learning and found that grid scores tended to increase over learning in both the square (see Fig. 3c for examples and Fig. S1 for all conditions; mean slope: 0.0044, bootstrap CIs: [0.0040, 0.0048]) and circular environment (mean slope: 0.0042, bootstrap CIs: [0.0038, 0.0046]; see Tables S1 and S2).

Following learning, we evaluated the gridness of the clustering solution (see examples in Fig. 3a, b and Figs. S2 and S3). A substantial proportion of simulations satisfied the criterion for grid-like organization, with 45.3% in the square and 38.6% in the circular environment, which closely match the proportions in empirical results (45% and 38%, respectively; see Supplementary Note 1). The average grid score in both the square environment (mean: 0.277, bootstrap CIs [0.273, 0.0.280]) and circular environment (mean: 0.313, bootstrap CIs [0.309, 0.318]) were greater than zero; see Tables S3 and S4).

Clustering solutions match grid patterns in mEC. To relate our account to typical spatial studies, we simulated an agent moving through its environment as in a free-foraging rodent experiment. As expected, learning led to clusters forming a hexagonal pattern (see examples in Fig. 3a, b, left). To assess this quantitatively, we computed the spatial autocorrelograms of the cluster activation maps (Fig. 3a, b, right) to obtain the grid score, which reflects the degree six-fold hexagonal symmetry in the cluster activation pattern across trials (see Methods). We computed grid scores for each time bin during learning and found that grid scores tended to increase over learning in both the square (see Fig. 3c for examples and Fig. S1 for all conditions; mean slope: 0.0044, bootstrap CIs: [0.0040, 0.0048]) and circular environment (mean slope: 0.0042, bootstrap CIs: [0.0038, 0.0046]; see Tables S1 and S2).

Following learning, we evaluated the gridness of the clustering solution (see examples in Fig. 3a, b and Figs. S2 and S3). A substantial proportion of simulations satisfied the criterion for grid-like organization, with 45.3% in the square and 38.6% in the circular environment, which closely match the proportions in empirical results (45% and 38%, respectively; see Supplementary Note 1). The average grid score in both the square environment (mean: 0.277, bootstrap CIs [0.273, 0.0.280]) and circular environment (mean: 0.313, bootstrap CIs [0.309, 0.318]) were greater than zero; see Tables S3 and S4).

Clustering solutions match grid patterns in mEC. To relate our account to typical spatial studies, we simulated an agent moving through its environment as in a free-foraging rodent experiment. As expected, learning led to clusters forming a hexagonal pattern (see examples in Fig. 3a, b, left). To assess this quantitatively, we computed the spatial autocorrelograms of the cluster activation maps (Fig. 3a, b, right) to obtain the grid score, which reflects the degree six-fold hexagonal symmetry in the cluster activation pattern across trials (see Methods). We computed grid scores for each time bin during learning and found that grid scores tended to increase over learning in both the square (see Fig. 3c for examples and Fig. S1 for all conditions; mean slope: 0.0044, bootstrap CIs: [0.0040, 0.0048]) and circular environment (mean slope: 0.0042, bootstrap CIs: [0.0038, 0.0046]; see Tables S1 and S2).

Following learning, we evaluated the gridness of the clustering solution (see examples in Fig. 3a, b and Figs. S2 and S3). A substantial proportion of simulations satisfied the criterion for grid-like organization, with 45.3% in the square and 38.6% in the circular environment, which closely match the proportions in empirical results (45% and 38%, respectively; see Supplementary Note 1). The average grid score in both the square environment (mean: 0.277, bootstrap CIs [0.273, 0.0.280]) and circular environment (mean: 0.313, bootstrap CIs [0.309, 0.318]) were greater than zero; see Tables S3 and S4).

Clustering solutions match grid patterns in mEC. To relate our account to typical spatial studies, we simulated an agent moving through its environment as in a free-foraging rodent experiment. As expected, learning led to clusters forming a hexagonal pattern (see examples in Fig. 3a, b, left). To assess this quantitatively, we computed the spatial autocorrelograms of the cluster activation maps (Fig. 3a, b, right) to obtain the grid score, which reflects the degree six-fold hexagonal symmetry in the cluster activation pattern across trials (see Methods). We computed grid scores for each time bin during learning and found that grid scores tended to increase over learning in both the square (see Fig. 3c for examples and Fig. S1 for all conditions; mean slope: 0.0044, bootstrap CIs: [0.0040, 0.0048]) and circular environment (mean slope: 0.0042, bootstrap CIs: [0.0038, 0.0046]; see Tables S1 and S2).

Following learning, we evaluated the gridness of the clustering solution (see examples in Fig. 3a, b and Figs. S2 and S3). A substantial proportion of simulations satisfied the criterion for grid-like organization, with 45.3% in the square and 38.6% in the circular environment, which closely match the proportions in empirical results (45% and 38%, respectively; see Supplementary Note 1). The average grid score in both the square environment (mean: 0.277, bootstrap CIs [0.273, 0.0.280]) and circular environment (mean: 0.313, bootstrap CIs [0.309, 0.318]) were greater than zero; see Tables S3 and S4).
only arise under very specific conditions in which the environment is fairly uniform. The imposition of any structure, including changes to the overall geometry of the environment, should affect the clustering in a manner that makes it less grid-like.

Related, Krupic et al. identified grid cells in rodent mEC in a square box, then placed the animals in a trapezoid environment. They found that activity maps of grid cells became less grid-like in the trapezoid and that the decline was greatest for responses elicited on the narrow side of the trapezoid. To simulate this experiment, the model was first trained in a square and then transferred to a trapezoid environment (see Fig. 4a–d for an example and Fig. S4 for more examples). As in the empirical studies, the model’s overall grid scores declined in the trapezoid environment (Fig. 4e); trapezoid mean grid score: 0.058, bootstrap CIs [0.054, 0.061]; Fig. 4f; square minus trapezoid mean: 0.219, bootstrap CIs [0.214, 0.224]) and the grid scores were higher on the wide than on the narrow side of the trapezoid (Fig. 4g; wide minus narrow mean: 0.133, bootstrap CIs [0.127, 0.139]; see Tables S5–S7).

Discussion

Previous work has explained a wide array of learning and memory phenomena in terms of clustering computations supported by the MTL. Here, this same basic account was shown to account for basic spatial navigation phenomenon, including place and grid cell-like response patterns. Specifically, we showed that a learning mechanism that seeks to minimize error in the task-relevant feature space captures conceptual structure in concept learning tasks and spatial structure in two-dimensional navigation contexts, which lead to place and grid cell-like representations. Rather than spatial mechanisms providing a scaffolding for more abstract conceptual knowledge, the current results suggest that key findings in the spatial literature naturally arise as limiting cases of a more general concept learning mechanism. Whereas concepts can be clumpy, structured, and high dimensional, typical spatial tasks involve exhaustive and uniform sampling of simple two-dimensional environments, which leads to degenerate clustering solutions that pack clusters into a hexagon lattice, giving rise to so-called grid cells (Fig. 3). The clustering account correctly predicted how deviations from these unstructured learning environments should reduce grid-like cell responses (Fig. 4).

Our proposal stands in contrast to other ideas that a dedicated, phylogenetically older spatial navigation system in the MTL supports the newer, higher-level cognitive functions. In particular, we suggest there are no intrinsic ‘place’ or ‘grid’ cells, but instead a flexible system that will represent the relevant variables at hand, including physical space. Emerging evidence shows that cells in the MTL exhibit mixed-selectivity in that they respond to multiple variables, such as place and grid cells that also code for task-relevant sound frequency, routes, objects and context, and time, suggesting a flexible code. Clustering is a flexible mechanism and can learn representations in multi-dimensional space, and therefore is a strong candidate mechanism for organizing multi-modal, complex information for...
consolidation of knowledge for memories and concepts. A growing body of work supports the idea that the MTL system is a key part of a general learning system for organizing knowledge into a useful representation, which can be used for effective behavior and for memory consolidation.

Clustering models organize information and represent concepts in feature space to enable the identification, classification, and generalization of novel objects. These models can be closely linked to episodic memory and accounts of hippocampal function including relational memory, statistical learning and hippocampal representations in combination with sensory inputs arriving via entorhinal and perirhinal cortex, and from anterior inferior temporal cortex to prefrontal cortex to the hippocampus.

Whereas our account holds that place and grid cells emerge from a general learning system, Bellmund and colleagues suggest that the population code of place and grid cells play a role in mapping the dimensions of cognitive spaces in cognitive tasks, and that spatial navigation could serve as a model system to understanding cognitive spaces. Although there are commonalities, their proposal suggests that place and grid cells provide a ’metric’ or distance code for abstract spaces, and that there is a straightforward mapping from neural representations of physical space to abstract space. In our view, when the context involves a significant degree of selective attention to stimulus features or task variables, the representational space can be warped to a different, more effective representation of the context at hand (e.g. reducing dimensionality by attending to the task-relevant dimensions), which does not simply map onto the two-dimensional spatial case.

Our higher-level account provides a general theoretical framework applicable to a large range of tasks, in contrast to lower-level models of place and grid cells which make specific predictions in spatial contexts but have less explanatory power to generalize across contexts. Our model’s contribution is providing a general mechanism that could be used across domains. Here, we provided an algorithmic-level model that links across two different computational accounts of task descriptions (spatial and concept tasks), and connects learning mechanisms from concept...
learning to spatial representations found at the single-cell level. Specifically, we were able to link the model representations to neural measures reported in the spatial literature, closely matching a number of empirical observations.

Our model showed a similar proportion of grid-like cells in found in mEC. Other models either do not capture the hexagonal code (90 degree grids) or need to build in additional constraints (non-negativity constraint changes the 90 degree grid patterns to 60 degree grid patterns). When they find a large proportion of grid cells, they are too good in that all the simulated cells are grid cells. Other work have modeled or analyzed mathematical properties of the grid code (e.g. refs. 31,33), but also do not account for variability in the grid score in mEC cells. Here, we used a simple model from a high-level perspective based on ideas from concept learning and memory and matched the proportion of grid cells with empirical data, suggesting that the constraints of the clustering model matches the constraints the brain uses to build these representations.

Our model also captured the causal relation between place and grid cells. In our account, grid cells play a cluster-match or error-monitoring function where they monitor (connected to and receive input from) place cells, and self-organize over time to produce a hexagonal firing pattern. This is consistent with developmental work, where place cells appear in baby rats very early in life, and grid cells develop shortly after, as they explore and learn about spatial environments during normal development. Furthermore, inactivation of the hippocampus (with place cells) leads to grid cells in mEC losing the periodicity of their firing fields, whereas inactivation of the mEC (with grid cells) only mildly affect hippocampal place fields. Our account provides a different way of thinking about hippocampal-mEC interactions, which makes predictions that can guide future experiments and analyses.

Our account suggests that grid-like responses from the MTL should be the exception, not the rule, when encoding abstract spaces. Outside the typical laboratory study, representational spaces may be high dimensional and not all dimensions or values along dimensions will be equally relevant, nor will all combination of values across dimensions (see Fig. 1a). In support of this characterization, empirical work has shown that grid cells also lose their grid-like properties in more complex environments such as mazes.

Our account made several predictions that matched empirical data, where changes in environmental geometry lead to specific changes in the cluster representation. The model also provides further predictions. First, the mapping from place to grid cells within a context should be predictable. An mEC grid or spatial cell is assumed to receive input from multiple place cells in the hippocampus, and that mEC cell should have fields in the same location as the place cells it receives input from (Fig. 2a, b). Therefore, if place cells that represent a certain location are inactivated, the corresponding fields of the mEC cells that monitor those place cells should also disappear. Since an mEC cell may receive inputs from multiple place cells, a strict test would require inactivation of all (or at least a large proportion of) place cells that represent one location (a cluster in the model), predicting all mEC cells should also lose those fields. Future work with large-scale concurrent recordings in multiple brain regions with specific (e.g. optogenetic) manipulation may allow these predictions to be tested. One novel prediction of our model is that when error is high early in learning for a particular location, mEC cells should show a low firing rate and that best matching place cells should update their tunings to more strongly respond at that location (i.e., cluster updating). Updating a cluster (or recruiting a new cluster) should result in adjustment to the tuning of neighboring clusters, leading to a cascade of changes across place cells.

When error is low, this signifies a good match between the environment and one’s current knowledge (cluster representation) and experience, and little or no update is necessary. Inactivating the mEC should disrupt the error signal, which should disrupt learning in new environments. Recent evidence suggests place and grid cells both move towards goals or rewards, and there seems to be a greater number of place cells recruited near goal locations consistent with more clusters moving towards the goal or more clusters recruited at locations near the goal. Finally, our model predicts that both grid and non-grid spatial cells should perform the same function, in both concept and spatial tasks. There is some evidence in the spatial domain which showed that non-grid spatial cells in mEC contain as much spatial information as grid cells and could serve similar functions.

The primary strength of our account, namely that it offers an algorithmic account of spatial and concept learning tasks, serves to highlight the need for complementary lower-level accounts. There are various open questions such as how place cell remapping occurs across contexts and partial remapping effects with disruption to mEC. Our hope is that our model can eventually link to lower-level models that incorporate biological details such as spiking neurons and incorporate knowledge from memory research that can explain more empirical findings and provide new insights to these questions. Accounts are needed at multiple levels of analysis. We view our model as intermediary (at the algorithmic level) and aim for it to serve as a bridge between the goal of the computation and its implementation. Our model can serve as a guide for how operations such as cluster updating are physically realized.

Building this integrative bridge between the concept, memory, and spatial literatures allows for findings from one domain to inform the other. For example, task goals and attentional mechanisms in the concept literature have been found to shape hippocampal representations. Analogous tasks can readily be constructed to evaluate whether spatial cells support broader information processing functions (cf. ref. 13) and how general learning algorithms shape their response properties (cf. ref. 14). Likewise, the concept literature emphasizes the hippocampus’s interactions with other brain areas, such as medial prefrontal cortex, to assist in encoding task relevant information. When richer spatial tasks are considered, there is a ready set of candidate mechanisms and neural systems that may offer domain general explanations that link across brain, behavior, and computation.

**Methods**

**Simulations.** A simulation run comprised of a learning period with a million trials (training phase) where clusters updated their positions in relation to the agent’s position as it explored the environment. After learning, we quantitatively assessed the regularity of the cluster position arrangements (test phase). We ran 1000 simulation runs for each condition (number of clusters).

**Simulation procedure and model specifications.** At the beginning of the learning phase of each simulation run, we set the number of clusters, number of learning trials, the environment (square, circle), the learning rate, and the learning update batch size. The number of clusters were set (ranging from 10 to 30) and were initiated at random locations in the environment. The shape of the environment was defined by a set of points that could be visited by the agent. The square environment was 50 by 50, where each point was a location specified by a value on the x- and y-axis. The circular environment was defined by drawing a circle in Matlab with a radius of 50, and selecting the points within the bounds of the circle. The starting position and movement trajectory of the agent was then determined as a random walk over one million trials. The agent started at a random position and steps in the horizontal and vertical axes were computed separately. On each trial, the agent could go up, down, or stay on the vertical axis, and left, right, or stay on the horizontal axis. The step was sampled from \([-4, -2, -1, 1, 0, 1, 2, 4]\), where negative values are steps to the left, positive steps are steps to the right, and zero means stay. Movement on the vertical dimension was determined in the same way, but negative values were upward steps and positive values were downwards.
steps. If the generated step brought the agent out of the environment, the step was cancelled and a new step was generated as above.

We considered a simple winner-take-all network in which only the cluster at position \( \mathbf{pos}_i \) closest to stimulus \( x \) (agent’s location) had a non-zero activation. Bold type is reserved for vectors. The distance between \( \mathbf{pos}_i \) and \( x \) is defined as:

\[
\text{dist}_i = ||\mathbf{pos}_i - x||
\]

(1)

In the Kohonen learning rule, cluster \( i \) updates its position \( \mathbf{pos}_i \) to move toward stimulus \( x \) according to:

\[
\Delta \mathbf{pos}_i = \eta_t (x - \mathbf{pos}_i),
\]

(2)

where \( \eta_t \) is the learning rate at time \( t \). In the present simulations, we used batch updating to increase numerical stability in which 200 updates were performed simultaneously. The learning rate \( \eta_t \) for batch time \( t \) followed an annealing schedule:

\[
\eta_t = \frac{\eta_0}{1 + \rho \cdot t}
\]

(3)

where \( \eta_0 \) is the initial learning rate set to 0.25 and \( \rho \) is the annealing rate set to 0.02.

**Assessing regularity of cluster positions.** To assess whether cluster positions formed a regular hexagonal structure with learning in a comparable manner way to grid cells found in the medial entorhinal cortex (mEC), we followed the method of Hafting et al.\(^5\) and Perez-Escobar et al.\(^35\). In Hafting et al.\(^5\), rodents traversed circular and square environments whilst they recorded electrophysiological signals from mEC neurons. They found cells that displayed multiple firing fields and resembled a grid of regularly tessellating triangles spanning the recorded environment. To assess this regularity quantitatively, they computed the spatial autocorrelation of the firing rate map. If the fields were arranged in a regular grid, the center peak of the autocorrelation should be surrounded by six equidistance peaks, forming a regular hexagon. The spatial autocorrelation was computed as follows. With \( \lambda(x, y) \) denoting the cluster activation at location \( (x, y) \), the autocorrelation with spatial lags of \( \tau_x \) and \( \tau_y \) was estimated as:

\[
r(\tau_x, \tau_y) = \frac{n^2 \sum \lambda(x, y) \lambda(x - \tau_x, y - \tau_y) - \sum \lambda(x, y) \sum \lambda(x - \tau_x, y - \tau_y)}{\sqrt{n^2 \sum \lambda(x, y)^2 - (\sum \lambda(x, y))^2}} \times \sqrt{\sum \lambda(x, y)^2 - (\sum \lambda(x, y))^2},
\]

(4)

where \( n \) is the number of activation maps that exceeded the shuffled grid score threshold values (most conservative) thresholds across all conditions in that particular environment.

For each condition, we computed the percentage of activation maps that exceeded the shuffled grid score threshold. We computed the percentage of ‘grid cells’ for each condition (number of clusters) separately and then computed the mean percentage across conditions.

**Gridness in trapezoid environments.** To simulate the effect of asymmetric boundaries in a trapezoid enclosure on gridness,\(^39\) we took cluster positions from simulations after learning in square environments, and ran an additional learning phase for 250,000 trials. In this new learning phase, the shape of the environment was now a trapezoid (the agent could only move to those locations), and the annealed learning rate schedule continued (starting at 0.0025, reducing to 0.002 at the end). The trapezoid dimensions were 5 x 24 x 50 pixels, closely matching the proportions in\(^36\) (0.2 x 0.9 x 1.9 meters; multiplied by (50/1.9) equals to 5.26, 23.7, and 150). In order to test whether the asymmetric boundaries of the trapezoid affected gridness, the trapezoid was split into two halves and we computed the grid score for the spatial autocorrelation on the left (wide) and right (narrow) side of the shape. To deal with this, we made a slight change to the possible steps after generating a step that brings the agent out of the environment, described below. For each trial, the step was generated as before. If the generated step was out of the environment, the step was cancelled, and the next step was determined as follows. If the step generated would have brought the agent out of the bottom of the trapezoid, the next step was sampled from \([0, 0, 1, 1, 1, 0] \) (stay or up). If the step brings the agent out to the top, the next step was sampled from \([-1, -1, 0, 0] \) (down or stay). When the step takes the agent out of the left of the trapezoid, then the next step to be sampled on the horizontal axis were \([0, 1, 1, 2, 4] \), towards the inner portion of the environment. If the step took the agent out of the right side of the trapezoid, the next step was sampled from \([-1, -1, 0, 0, 0, 0] \) (stay or up). This is because when the agent is out of the trapezoid on the horizontal (left-right) axis, the agent could still be in the middle of the shape on the vertical axis, since the shape becomes more narrow as it reaches the right. Finally, when it lands exactly in the middle of the horizontal axis, but is out of the shape (on the horizontal axis), the next step to be sampled from on the vertical axis is \([-1, 0, 1] \).
Code and data availability
All simulation code and plotting scripts are available at https://github.com/robnoskel/code4Cell. Data generated and used for this manuscript is available at https://osf.io/2d3x3/.

Received: 5 October 2018; Accepted: 24 November 2019; Published online: 12 December 2019

References
1. Barsalou, L. W. Grounded cognition. Annu. Rev. Psychol. 59, 617–645 (2008).
2. Boroditsky, L. Metaphoric structuring: understanding time through spatial metaphors. Cognition 75, 1–28 (2000).
3. Buzsáki, G. & Moser, E. I. Memory, navigation and theta rhythm in the hippocampal-entorhinal system. Nature Neuroscience 16, 130–138 (2013).
4. O’Keefe, J. & Nadel, L. The Hippocampus as a Cognitive Map. (Oxford University Press, Oxford, 1978).
5. Hafting, T., Fyhn, M., Molden, S., Moser, M. B. & Moser, E. I. Microstructure of a spatial map in the entorhinal cortex. Nature 436, 801–806 (2005).
6. Doeller, C. F., Barry, C. & Burgess, N. Evidence for grid cells in a human memory network. Nature 463, 657–661 (2010).
7. Horner, A. J., Busby, J. A., Zotow, E., Bush, D. & Burgess, N. Grid-like processing of imagined navigation. Curr. Biol. 26, 842–847 (2016).
8. Tolman, E. C. Cognitive maps in rats and men. Psychol. Rev. 55, 189–208 (1948).
9. Epstein, R. A., Patat, E. Z., Julian, J. B. & Spiers, H. J. The cognitive map in humans: spatial navigation and beyond. Nat. Neurosci. 20, 1504–1513 (2017).
10. Bellmund, J. L. S., Gärdenfors, P., Moser, E. I. & Doeller, C. F. Navigating with complementary coding schemes. Neuron 36, 660–3675 (2016).
11. Burgess, N., Maguire, E. A. & O’Keefe, J. The human hippocampus and spatial and episodic memory. Neuron 35, 625–641 (2002).
12. Constantinescu, A. O., O’Reilly, J. X. & Behrens, T. E. J. Organizing conceptual knowledge in humans with a gridlike code. Science 352, 1464–1468 (2016).
13. Aronov, D., Nevers, R. & Tank, D. W. Mapping of a non-spatial dimension by hippocampal neurons encode information about different types of memory tasks occurring in the same location. Neuron 72, 623–633 (2020).
14. Grieves, R. M., Wood, E. R. & Dudchenko, P. A. Place cells on a maze encode routes rather than destinations. Elife 5, e15986 (2016).
15. Keene, C. S. et al. Complementary functional organization of neuronal activity patterns in the perirhinal, lateral entorhinal, and medial entorhinal cortices. J. Neurosci. 36, 3660–3675 (2016).
16. Eichenbaum, H. On the integration of space, time, and memory. Neuron 95, 1007–1018 (2017).
17. Eichenbaum, H. & Cohen, N. J. From Conditioning to Conscious Recollection: Memory Systems of the Brain. (Oxford University Press, Oxford, 2001).
18. Schapiro, A. C., Rogers, T. T., Cordova, N. L., Turk-Browne, N. B. & Botvinick, M. M. Neural representations of events arise from temporal community structure. Nat. Neurosci. 16, 486–492 (2013).
19. Dayan, P. Improving generalization for temporal difference learning: the successor representation. Neural Comput. 5, 613–624 (1993).
20. Memmennejad, I. & Howard, M. W. Predicting the future with multi-scale successor representations. Preprint at https://www.biorxiv.org/content/10.1101/449470v1 (2018).
21. deCotiis, W. & Barry, C. Neurobiological successor functions for spatial navigation. Preprint at https://www.biorxiv.org/content/10.1101/799412v1 (2019).
22. Whittington, J. C. et al. The Tolman-Eichenbaum Machine: Unifying space and relational memory through generalisation in the hippocampal formation. Preprint at https://www.biorxiv.org/content/10.1101/770495v2 (2019).
23. Catami, M., Howard, R. J., Pajevic, S. & Jones, D. K. Virtual in vivo interactive dissection of white matter fasciculi in the human brain. Neuroimage 17, 77–94 (2002).
24. Martino, J. et al. Cortex-sparing fiber dissection: an improved method for the study of white matter anatomy in the human brain. J. Anat. 219, 531–541 (2011).
25. Nosofsky, R. M. Attention, similarity, and the identification-categorization relationship. J. Exp. Psychol. Gen. (1986).
26. Love, B. C. The algorithmic level is the bridge between computation and brain. Top. Cogn. Sci. 7, 230–242 (2015).
27. Krupic, J., Bauza, M., Burton, S., Lever, C. & O’Keefe, J. How environment geometry affects grid cell symmetry and what we can learn from it. Philos. Trans. R. Soc. B Biol. Sci. 369, 20130188 (2014).
28. Wills, T. J., Cacucci, F., Burgess, N. G. & O’Keefe, J. Development of the hippocampal cognitive map in preweanling rats. Science 328, 1573–1576 (2010).
29. Wills, T. J., Barry, C. & Cacucci, F. The abrupt development of adult-like grid cell firing in the medial entorhinal cortex. Front. Neural Circuits 6, 21 (2012).
30. Bonnevie, T. et al. Grid cells require excitatory drive from the hippocampus. Nat. Neurosci. 16, 309–317 (2012).
31. Halls, J. B. et al. Medial entorhinal cortex lesions only partially disrupt hippocampal place cells and hippocampus-dependent place memory. Cell Rep. 9, 893–901 (2014).
32. Derdakian, D. et al. Fragmentation of grid cell maps in a multicompartament environment. Nat. Neurosci. 12, 1325–1332 (2009).
33. Dupret, D., O’Neill, J., Pleydell-Bouverie, B. & Cisarvá, J. The reorganization and reactivation of hippocampal maps predict spatial memory performance. Nat. Neurosci. 13, 995–1002 (2010).
34. Boccara, C. N., Nardin, M., Stella, F., O’Neill, J. & Cacucci, F. The entorhinal cognitive map is attracted to goals. Science 363, 1443–1447 (2019).
35. Butler, W. N., Hardcastle, K. & Giocomo, L. M. Remembered reward locations restructure entorhinal spatial maps. Science 363, 1447–1452 (2019).
36. Hollup, S. A., Molden, S., Donnett, J. G., Moser, M. B. & Moser, E. I. Accumulation of hippocampal place fields at the goal location in an annular watermaze task. J. Neurosci. 21, 1635–1644 (2001).
37. Hoh, V. et al. Goal-related activity in hippocampal place cells. J. Neurosci. 27, 472–482 (2007).
38. Diehl, G. W., Hon, O. J., Leutgeb, S. & Leutgeb, J. K. Grid and nongrid cells in medial entorhinal cortex represent spatial location and environmental features with complementary coding schemes. Nature 94, 83–92 (2017).
39. Bendor, M. P., Koenig, J., Leutgeb, J. K. & Leutgeb, S. New and distinct hippocampal place codes are generated in a new environment during septal inactivation. Neuron 82, 789–796 (2014).
Acknowledgements
We thank all members of the LoveLab for valuable input. We thank Roddy Grieves for his advice on analysis of grid measures and members of the Institute of Behavioural Neuroscience (IBN) at UCL for valuable input and discussions. This work was funded by the National Institutes of Health [grant number 1P01HD080679]; a Royal Society Wolfson Fellowship [18302] to Bradley C. Love; and a Wellcome Trust Senior Investigator Award [WT106931MA] to Bradley C. Love.

Author contributions
R.M.M.: conceptualization, formal analysis, methodology, software, visualization, writing—original draft, writing—review & editing. B.C.L.: conceptualization, formal analysis, funding acquisition, methodology, resources, supervision, visualization, writing—original draft, writing—review & editing.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-019-13760-8.