Environmental deformations dynamically shift the spatial metric of the brain

Alex T. Keinath[1], Russell A. Epstein[1], and Vijay Balasubramanian[2]

[1]*Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104*
[2]*David Rittenhouse Laboratories, University of Pennsylvania, Philadelphia, PA 19104*

Abstract

Environmental deformations induce distortions in the time-averaged activity of grid and place cells, which are thought to reflect a rescaling of the spatial metric of the rodent cognitive map. We propose a mechanism for this phenomenon, where input from border cells resets the spatial phase of grid cells, maintaining a learned relationship between grid phase and boundaries. A computational model demonstrated that this mechanism would yield scale-dependent distortions in time-averaged grid fields, as well as stretched, duplicated, and fractured place fields, as observed experimentally. Furthermore, the model yielded a striking prediction: apparent distortions in time-averaged activity actually arise from dynamical, history-dependent "shifts" in grid field locations. We reanalyzed the two classic datasets on grid distortions and found clear evidence of our alternative prediction. Thus, the effects of environmental deformations on spatial representations must be reconceptualized – altering spatial geometry does not distort, but rather dynamically shifts, the hippocampal map.

Introduction

The cognitive map is thought to be a metric representation of space that preserves distances between represented locations [1, 2]. Entorhinal grid cells are hypothesized to generate this metric by maintaining an internally-generated, path-integrated representation of space [3–7]. Results of environmental deformation experiments have led to the widespread belief that this metric is fundamentally malleable. In these experiments, neural activity is recorded as a rat explores deformed versions of a familiar environment where chamber walls have been stretched, compressed, removed, or inserted. Such deformations induce a number of distortions in the time-averaged activity of both grid [8, 9] and hippocampal place cells [10–14]. Often described as 'rescaling', these distortions have been taken to suggest that the metric embedding of the cognitive map can be reshaped by altering environmental geometry [8]. Crucially, however, this interpretation assumes that the distortions observed in the time-averaged rate maps of these cells directly reflect corresponding changes to the underlying spatial code. Here, we propose and provide evidence for an alternative mechanism which challenges this interpretation and instead indicates that the metric embedding of the cognitive map perseveres unchanged during environmental deformations while undergoing dynamical, history-dependent shifts.

We propose that border cells interact with grid cells in familiar environments to maintain learned relationships between grid phase and boundaries [15, 16]. Under our hypothesis, when an enviroment is deformed, border cell input continues to maintain these boundary-phase relationships, inducing shifts in the grid pattern that depend on the most recent boundary encounter. Averaged

over time, these *boundary-tethered grid shifts* lead to apparent distortions of grid patterns. The shifts are inherited by downstream place cells, producing similar changes in place field location. In this view, distortions observed in the time-averaged activity of place and grid cells are an artifact of averaging over paths originating from different boundaries; aligned by the most recently contacted boundary, the grid and place codes should appear shifted, not distorted.

To test this mechanism, we constructed a model where the activity of a grid cell attractor network [17] is shaped by Hebbian-modified input from border cells [16, 18]. Recent work has shown that such border cell input to grid cells can stabilize drift in the grid pattern [15, 16], providing evidence for such interactions in familiar environments. Simultaneously, a population of hippocampal place cells is learned from grid cell output [19, 20]. Our simulations show that during environmental deformations, modeled grid and place units reproduce a number of striking experimentally-observed behaviors, some of which have had no previous explanation: 1) when a familiar environment is rescaled, the firing patterns of large-scale grid cells rescale to match the deformation, while the firing patterns of small-scale grid cells do not [8, 9]; 2) when a familiar environment is stretched, the fields of place units exhibit a mix of stretching, bifurcation, modulation by movement direction, and inhibition [10]; 3) when a familiar linear track is compressed, the place code is updated when a track end is encountered [11]; 4) when a new boundary is inserted in an open environment, place fields exhibit a mix of duplication, inhibition, and perseverance [12–14]. Furthermore, our model generated an additional unique prediction: a signature shift in grid and place field location, dependent on the most recently contacted boundary. To test this prediction, we reanalyzed the datasets from two key environmental deformation experiments [8, 9], and found compelling new evidence for boundary-tethered grid shifts. In sum, our results indicate that deformation of a familiar environment results in dynamic history-dependent shifts in grid and place field locations, rather than distortions of field shape. These results invite a reconceptualization of how environmental deformations affect spatial representation and navigation more broadly.

Results

A model of border, grid, and place cell interactions

Previous recording studies in rodents have shown a number of striking effects of environmental deformations on the time-averaged firing maps of grid cells (e.g., compression and stretching [8, 9]) and place cells (e.g., rescaling, bifurcation, modulation by movement direction, and suppression [10–14]). We sought a single circuit mechanism that could give a unified explanation for these phenomena. To this end, we developed a spiking feedforward model of border, grid, and place cell interactions (Fig. S1).

The border population consisted of 36 units whose activity was designed to mimic the behavior of border cells [18]. (Throughout this paper, we use unit to refer to modeled data, and cell to refer to *in vivo* recorded data.) Each border unit was active only when a boundary was nearby, within 12 cm in a particular allocentric direction [15]. The preferred firing field of each border unit covered $\sim 10\%$ of a boundary length, and maintained proportional coverage if that boundary was deformed [18]. If a new boundary was inserted, the border unit was active at an allocentrically analogous location adjacent to the new boundary.

The grid population was subdivided into 5 modules, each consisting of a neural sheet of size 90×90 units. The internal connectivity and dynamics of each module was based on the attractor network model described in [17], and identical across modules except for a single movement velocity gain parameter controlling the grid scale of each module. This parameter was adjusted to yield a geometric series of scales across modules (scale factor of 1.42), as observed experimentally [9]

and explained theoretically [21–23]. In addition to these connections, each grid unit also received initially random excitatory input from all border units. These connections developed through experience via a Hebbian learning rule in which connections between coactive grid and border units were strengthened at the expense of connections from inactive border units [24]. The place cell population consisted of 64 units receiving initially random excitatory input from 500 random grid units. These connections also developed with experience via Hebbian learning [20, 24]. In combination with uniform recurrent inhibition, these dynamics yield place-cell-like activity at the single unit level.

Model grid units deform with the environment in a scale-dependent fashion.

To explore the effects of environmental rescaling on grid units, we familiarized a naive virtual rat with a 150 cm x 150 cm square environment, during which period the border-grid connectivity self-organized via Hebbian learning (see Methods). Next this rat explored the familiar environment and deformed versions of this environment (chamber lengths between 75 cm to 225 cm in increments of 25 cm; chamber sizes chosen to match experiment [9]). We observed that these deformations induced rescaling of time-averaged rate maps in some grid modules (Fig. 1a). To quantify this module-dependent rescaling, we computed the grid rescaling factor required to stretch or compress the time-averaged rate maps in the deformed environment to best match the rate maps in the familiar environment, separately for each module. We found that the grid patterns of units in large-scale modules morphed with the environment, but pattern of units in small-scale modules tended not to (Fig. 1b). Precisely this behavior is observed experimentally [9], but a mechanistic explanation has remained elusive. Our model is thus the first demonstration of a mechanism capable of reproducing scale dependent distortions of grid activity patterns during environmental deformation.

Our model predicts that whether the rate map of a grid cell appears to rescale is not just an inherent property of its grid scale, but also of the size and type (stretch vs. compression) of the deformation. Preliminary evidence for this prediction comes from the observation of small-scale grid rescaling during less extreme [8] but not during more extreme compression deformations [9]. A detailed exploration of the dependence of grid rescaling on deformation size and type would provide a strong test of our theory.

Model place units distort heterogenously during environmental deformations.

To explore the effects of stretching deformations on model place units, we began by familiarizing the naive virtual rat with a 61 cm x 61 cm square open environment, during which period the border-grid connectivity and grid-place connectivity self-organized via Hebbian learning. Following this familiarization, the virtual rat then again explored the familiar environment, as well as a number of deformed environments (various chamber lengths between 61 cm and 122 cm, chamber widths 61 cm or 122 cm; chamber sizes chosen to match experiment [10]). During these deformations, we observed heterogeneous changes to the time-averaged rate maps of place units. A number of place units exhibited place field stretching in proportion to the rescaling deformation (Fig. 2ai). Other units exhibited place field bifurcations accompanied by progressively lower peak firing rates during more extreme deformations (Fig. 2aii). Finally, many units exhibited emergent modulation by movement direction (Fig. 2aiii), with place fields shifting 'upstream' of the movement direction (Fig. 2b). Precisely this mix of place field distortions is observed experimentally [10].

Next, we examined the effects of compressing a familiar linear track. We first familiarized the naive virtual rat with running laps on a 161 cm long linear track, during which period the border-

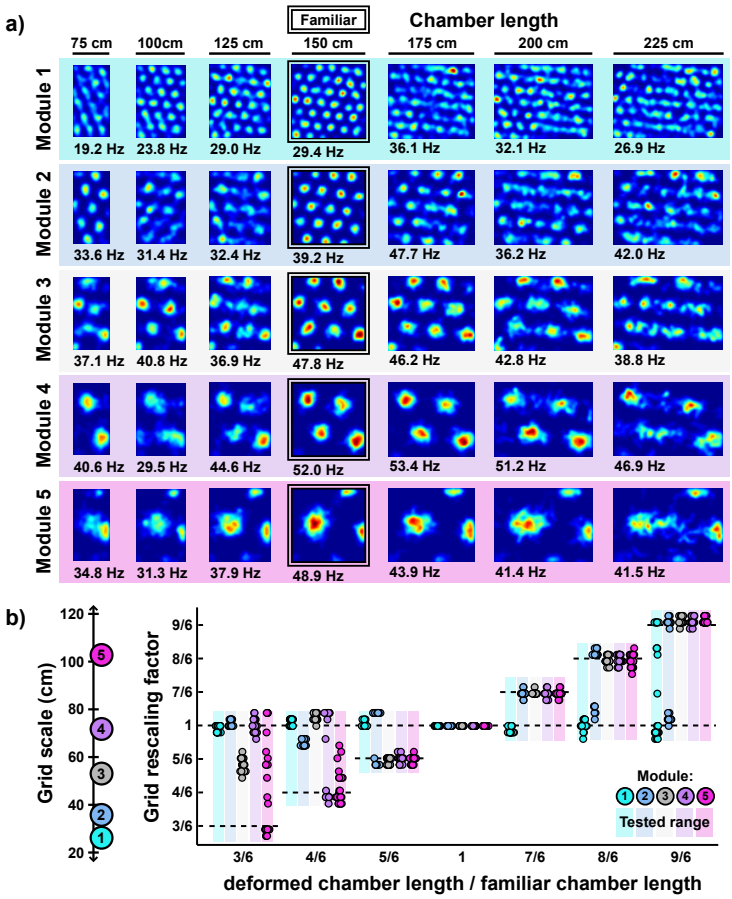

Figure 1: Grid unit response during deformation of a familiar open field environment. a) Rate maps from five example grid units, one from each module, across all tested deformations. Colors normalized to the maximum peak for each set of rate maps. Peak firing rate noted below the lower left corner of each map. b) Grid rescaling factors that best match the familiar enclosure rate maps for 30 randomly chosen grid units from each Module when the familiar open (150 cm x 150 cm) environment is rescaled to various chamber lengths (see Methods). Each dot denotes the grid rescaling factor of a single grid unit, and the color indexes the module (grid scale of each module is indicated at left).

grid connectivity and grid-place connectivity self-organized via Hebbian learning. Following this familiarization, the virtual rat ran laps along both the familiar track and a number of compressed tracks (track lengths between 53 cm to 161 cm; lengths chosen to match experiment [11]). During laps on compressed tracks, place unit activity unfolded as if unaffected by the compression no matter how extreme, until the opposing track end was reached. Once encountered, the place code previously active at this track end during familiarization reemerged (Fig. 2b). Precisely this behavior is observed experimentally [11].

Finally, we explored the effects on model place units of inserting a new boundary, We first familiarized the naive virtual rat with a 65 cm x 65 cm square open environment, during which period the border-grid connectivity and grid-place connectivity self-organized via Hebbian learning. Following this familiarization, the rat explored the familiar environment and a deformed version of this environment containing an additional 40 cm long boundary adjacent to one wall and evenly dividing the space (chosen to match experiment [14]). Again, we observed heterogeneous changes

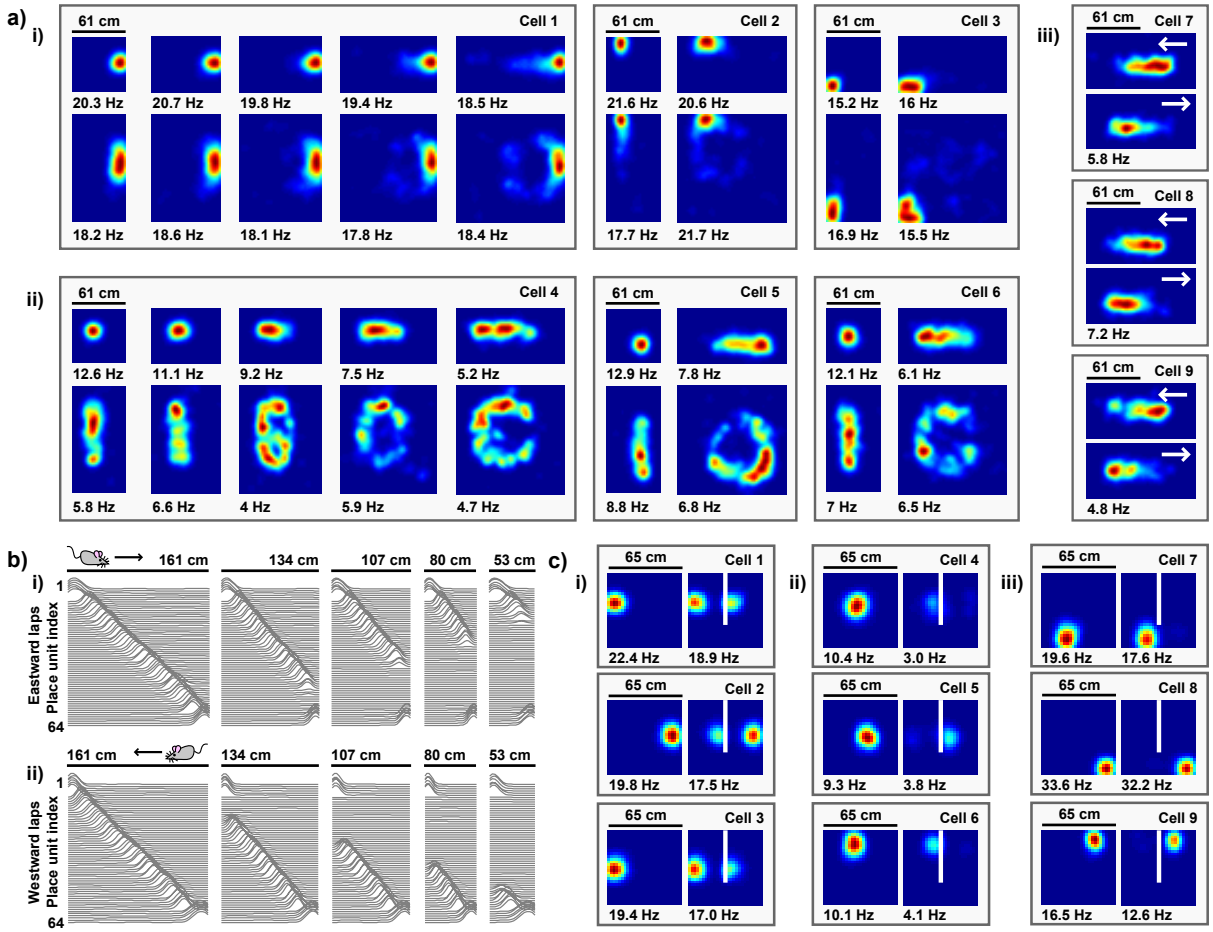

Figure 2: Place unit response during environmental deformations. a) Place unit rate maps when a familiar open (61 cm x 61 cm) environment is stretched. Place fields exhibit (i) stretching, (ii) bifurcation, and (iii) emergent modulation by movement direction (indicated by white arrow). Colors normalized to the peak for each rate map. Peak firing rate noted below the lower left corner of each map. Note that peak firing rate tends to decrease with more extreme deformations for cells with place fields near the center of the environment. b) Place unit activity for all 64 place units during compressions of a familiar (161 cm) linear track, separated by (i) eastward and (ii) westward laps. Each line indicates the firing rate of a single place unit at each location across the entire track during movement in the specified direction, normalized to the familiar track peak rate. Units sorted by place field location on the familiar track. Note that, during compressions, the place code unfolds as if anchored to the beginning of the track until the end of the track is encountered, at which point the familiar end-of-track place units are reactivated. c) Place unit rate maps demonstrating a mix of place field (i) duplication, (ii) inhibition, and (iii) perseverance when a new boundary (white line) is inserted in a familiar open (65 cm x 65 cm) environment. Peak firing rate noted below the lower left corner of each map.

in the time-averaged rate maps of place units. Some units exhibited place field duplication during boundary insertion (Fig. 2ci) while other units exhibited place field inhibition (Fig. 2cii). Still other persevered largely unaffected (Fig. 2ciii). Precisely this mix of responses is observed experimentally during boundary insertions [12–14].

Together, these results demonstrate that our model gives rise to the heterogeneous behaviors of place cells observed across three unique deformation paradigms, including features of these data

that have no other current mechanistic explanation.

Boundary-tethered grid shift underlies model grid and place unit distortions.

How do model interactions give rise to these grid and place unit distortions? During familiarization, Hebbian learning strengthens the connections from active border units to active grid units at the expense of connections from inactive border units (Fig. 3a). Once familiarized, active border units reinstate the grid network state associated with the same pattern of border unit responses during familiarization. This grid reinstatement occurs even when border inputs are activated at a new location, such as when a new or displaced boundary is encountered. In a rescaled open environment, grid reinstatement leads to 'shifts' in the spatial phase of the grid pattern, such that the phase relative to the most recent border input matches the phase entrained during familiarization in the undeformed environment (Fig. 3b,c). Averaged over time, these shifts resemble a rescaling of the grid pattern. Because of the periodic nature of the grid representation, border input can only reset the grid network state to within one grid period. Thus if the grid scale of a unit is small and the extent of the environmental deformation is relatively large, the number of grid fields within the deformed environment will be substantially different, making rescaling a poor description.

As for place cells, in our model place unit activity can be thought of as a thresholded sum of grid unit input [19]. Because of the boundary-tethered grid shifts induced during rescaling deformations, the location of each place field will also shift, maintaining its spatial relationship to the most recently contacted boundary (Fig. 3d). Critically, the likelihood of having most recently encountered a given boundary differs throughout the open environment: locations near the center of the chamber are more likely to have had an even distribution of previously encountered boundaries, while locations near a given boundary are more likely to be visited following an encounter with that boundary (Fig. 3e). When averaged across time, these most recent boundary biases result in a mix of place field stretching (closer to boundaries) and bifurcation distortions (further from boundaries).

The most recently encountered boundary is correlated with the direction of movement: the rat is more likely to have most recently encountered a given boundary when moving away from that boundary (Fig. 3f). Because of this, boundary-tethered place field shift causes place fields to be displaced 'upstream' along the direction of movement. Finally, more extreme rescaling distortions lead to more extreme boundary-tethered shifts and less frequent convergence of grid inputs at the same location, and thus overall systematic decreases in the peak firing rate of grid and place units. Because movement and thus the most recently contacted boundary is constrained when running laps on a linear track, linear track compressions provide an especially clear view of boundary-tethered updating. Until a track end is encountered, grid and place unit activity unfolds according to path integration alone. When a track end is encountered, border input reinstates the grid network state and in turn the place network state that coincided with that track end on the familiar track. This produces the upstream place field shifts observed in Fig. 2c and [11].

Inserting a boundary in an open environment elicits identical border unit activity when either the old boundary or new boundary is nearby in the preferred allocentric direction, inducing boundary-tethered reinstatement of the grid network state at both locations (Fig. S2). This grid shift translates to a duplication of the place unit representation adjacent to the old and inserted boundaries. Because a new representation is now active around the inserted boundaries, the old representation previously active at this location in the familiar environment is no longer activated, leading to an inhibition of place units participating in this old representation. The representations of locations away from the duplicated boundaries often persevere unaffected. Thus, in our model, boundary-tethered grid shift drives the diverse grid and place field distortions observed during

geometric deformations.

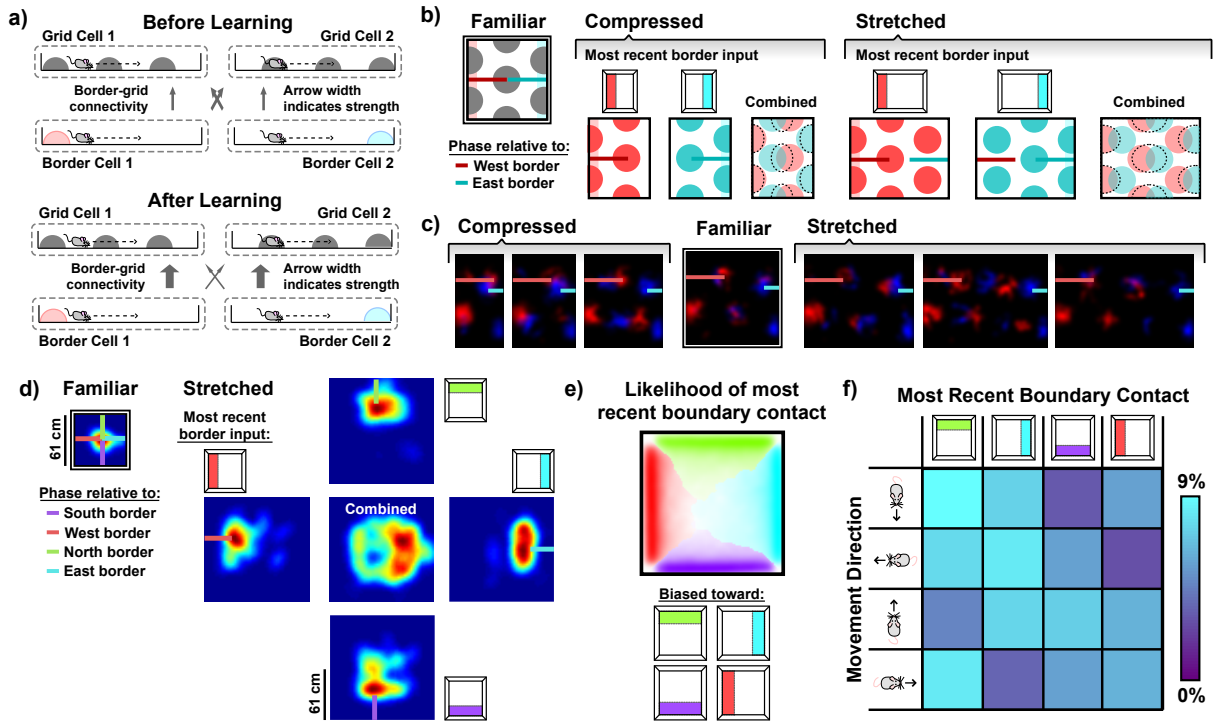

Figure 3: Boundary-tethered grid shift underlies model grid and place unit distortions. a) During familiarization, Hebbian learning strengthens the connections between coactive border and grid cells, at the expense of non-coactive connections. b) During deformations, border input acts to maintain the previously learned relationship between grid phase and the most recent border input. c) Rate map of a Module 4 grid unit when the west border was the most recently contacted boundary (red), overlayed with the rate map of the same unit when the east border was the most recently contacted boundary (blue). The spatial phase relative to the most recent border input (indicated by red/blue bars) is preserved during all deformations. Thus the grid pattern is undistorted when separated by the most recent border input. d) Place fields shift to maintain their previously learned relationships relative to the most recent border input. e) Likelihood of having most recently contacted each border as a function of location in the square environment. Color saturation denotes the strength of the bias. Data from [25]. f) Each square in the grid depicts the joint probability distribution in a square environment of the indicated most recently contacted boundary and movement direction. Data from [25].

Experimental evidence of the predicted boundary-tethered grid shifts.

Our model thus generates a clear and unique prediction: when a familiar open environment is rescaled, the locations of grid and place fields should both shift to preserve their spatial relationship relative to the most recently contacted boundary. To test this prediction, we reanalyzed data from [8] and [9]. In [8], rats were familiarized with either a 100 cm x 100 cm square or a 100 cm x 70 cm rectangular open environment. Following this familiarization, rats were reintroduced to familiar and deformed environments (all combinations of chamber lengths and widths of 70 cm or 100 cm), while the activity of grid cells was recorded (familiar square: 42 grid cells; familiar rectangle: 23 grid cells). In [9], rats were only familiarized with a 150 cm x 150 cm square open environment.

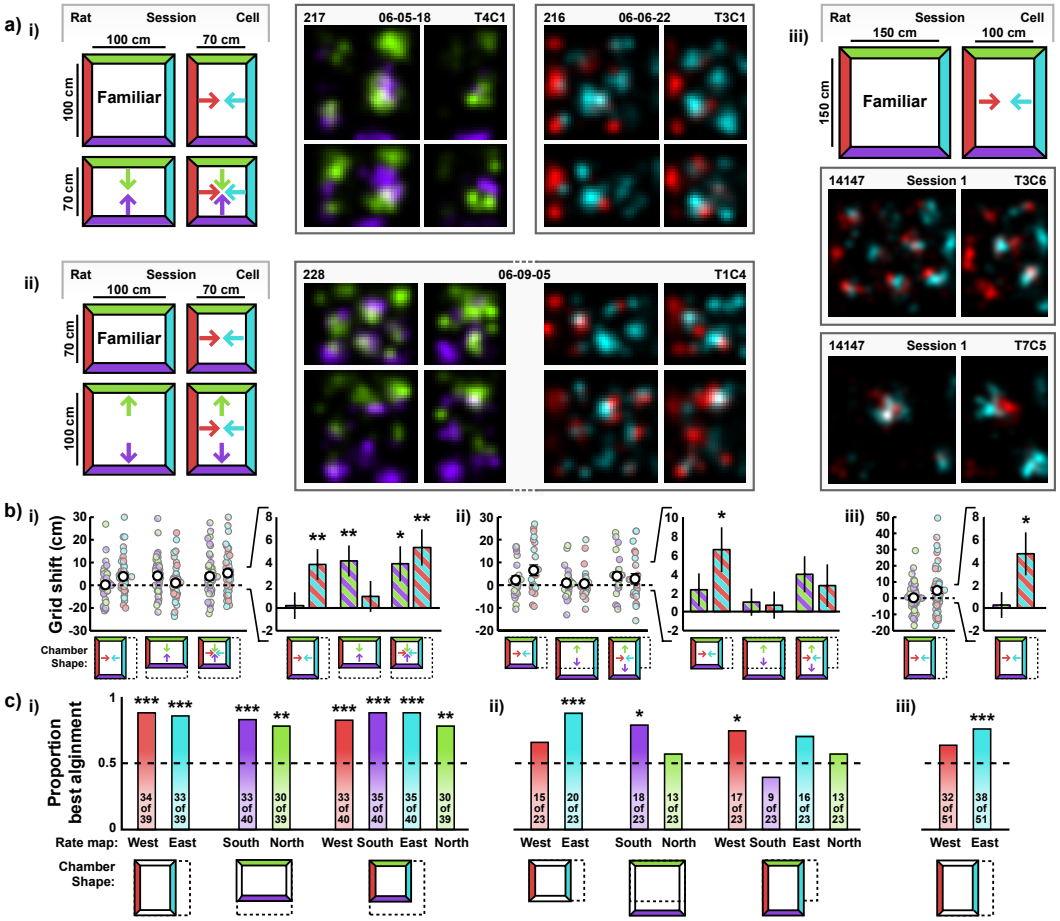

Figure 4: Evidence of predicted grid shifts during environmental deformations from rats trained in (i) a familiar square (data from [8]), (ii) a familiar rectangle (data from [8]), and (iii) a familiar square (data from [9]). a) Examples of grid shifts visible in 'boundary rate maps' (see text) created by dividing spiking data according to the most recently contacted boundary (indicated by rate map color). Maps organized by opposing north-south (green—purple) and east-west (blue—red) boundary pairs. Colored arrows indicate the directions of shifts predicted by our model during each deformation. b) Grid shift was measured as the relative phase between opposing boundary rate maps, minus the average shift during familiar trials (see text). Grid shift is shown for each cell (dots), separated by north-south (green—purple) and east-west (blue—red) boundary rate map pairs. There is an increase in grid shift along deformed, but not undeformed, dimensions, here quantified by the mean of the distribution of shifts (insets, errors bars are ± 1 SEM). Significance markers (asterisks: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$) denote the outcome of an uncorrected 2-tailed 1-sample t-test versus 0 (full statistics in Table S1). Colored arrows indicate the dimensions along which our model predicts positive grid shift during the indicated deformation of the familiar environment (dotted line). c) Proportion of trials for which each boundary rate map was best matched with its familiar environment rate map when aligned by the most recently contacted boundary (as predicted by our model) vs. the opposing boundary (counts shown within the bars). Familiar environment (dashed line), deformed environment (solid walls), and boundary rate map alignment (colored walls) shown in lower insets. Significance markers (asterisks: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$; uncorrected 2-tailed sign test) denote the probability of the measured proportion occurring by chance (i.e., 50% probability of best alignment by either boundary; full statistics inTable S2).

Following this familiarization, the rats were reintroduced to this environment as well as to a 100 cm x 150 cm rectangular enclosure (51 grid cells).

We began by separating the spiking data of each cell by the most recently contacted boundary, either the north, south, east or west, with contact defined as coming within 12 cm of the boundary [15]. From these data, we created four 'boundary rate maps' which summarized the spatial firing pattern of the grid cell after contacting the north, south, east and west boundaries, respectively. Comparison of such rate maps conditioned on contact with opposing boundaries (north-south vs. east-west) revealed clear examples of grid shift along deformed dimensions (Fig. 4a). To quantify these shifts, we first cross-correlated opposing boundary rate maps (i.e., north-south and east-west). Next, we computed the distance from center of the cross-correlation to the nearest peak along the dimension between to the boundary pair. This distance measured the relative shift between the corresponding rate maps. Finally, we subtracted the average shift observed during familiar trials from the shift observed during each deformation trial. This measure revealed an increase in grid shift along deformed, but not undeformed, dimensions across deformation trials (Fig. 4b).

Next we asked whether the grid pattern in each boundary rate map maintained its spatial phase with the corresponding boundary, as our model would predict. To address this question, we compared the full familiar environment rate map to each of the boundary rate maps, while varying their alignment along deformed dimensions. If the spatial relationship relative to the most recently contacted boundary is preserved, then each boundary rate map should be most similar to the familiar environment rate map when the two maps are aligned by the corresponding boundary. If, on the other hand, reshaping a familiar environment rescales the grid pattern uniformly, then the familiar and boundary rate maps should be equally well aligned by any boundary. Consistent with our prediction, we found that a boundary rate map and the familiar environment map were usually best aligned by the corresponding boundary, rather than the opposite boundary (Fig. 4c).

Lastly, we examined firing rate predictions of our model. If, during deformations, grid vertices are shifted to different locations when different boundaries are encountered, then averaging across trajectories originating from different boundaries will necessarily reduce the peak values of the whole-trial rate map. Thus our model predicts a reduction in the *peak firing rate* during environmental deformations (Fig. S3a), as measured by the peak value of the whole-trial rate map. On the other hand, because the density of grid fields within the environment remains unchanged on average, grid shift does not predict a change in *mean firing rate* (Fig. S3b), as measured by the total number of spikes across the entire trial divided by the trial duration. To address this prediction, we computed for each grid cell the ratio of its average peak rate across all deformed environment trials to the average peak rate across familiar environment trials. This ratio was significantly below 1 (0.873 ± 0.028, mean $\pm$ SEM; 2-tailed 1-sample t-test versus 1: $t(114) = 4.6$, $p < 0.001$; Fig. S3a), indicating that peak firing rates were lower during environmental deformations, as predicted. Conversely, the ratio of the mean firing rates during deformed and familiar trials did not significantly differ from 1, suggesting that the mean firing rate did not change during environmental deformations (1.037 ± 0.032, mean $\pm$ SEM; 2-tailed 1-sample t-test versus 1: $t(114) = 1.16$, $p = 0.25$; Fig. S3b). In sum, these results provide convergent evidence of boundary-tethered grid shifts during environmental deformations.

Discussion

Here we proposed that a novel mechanism, boundary-tethered grid shift, underlies the effects of environmental deformations on grid and place cells. We first demonstrated that a model incorporating this mechanism can account for key results from a number of environmental deformation

studies, including features of these data that have remained otherwise unexplained (e.g. scale-dependent rescaling of grid patterns and modulation of grids by movement direction). Next, we reanalyzed two classic environmental deformation datasets, and observed striking evidence of the predicted grid shifts, including the precise alignment of these shifts to the most recently contacted boundary. Together, these results provide new and compelling evidence that a novel mechanism, boundary-tethered grid shift, underlies the distortions observed in grid and place cell activity during environmental deformations. We also showed that at a circuit level the proposed mechanism could be naturally implemented by interactions between border cells and grid cells in the entorhinal cortex.

The apparent rescaling of grid patterns during environmental deformations has been taken as evidence that the mental metric for space maintained by grid cells can be reshaped by altering environmental geometry [8, 26]. Our results indicate instead that the grid pattern continues to reflect unbiased metric path integration during environmental deformations, at least relative to the most recently contacted boundary. However, our results further imply that, during deformations, a particular grid network state no longer coincides with a unique spatial location within the environment. This dissociation between grid network state and spatial location may make environmental deformation paradigms particularly well-suited for dissociating the contributions of grid cells to navigation from other factors [27, 28].

Previous work has revealed striking parallels between deformation-induced distortions of spatial representations in the rat brain and the spatial memory of humans in deformed environments [10, 27, 29, 30], leading to the suggestion that a common mechanism might underlie these effects. Consistent with this, recent evidence suggests that rescaling can be observed in the time-averaged activity of human grid cells [31]. In light of our results, we suggest that boundary-tethered grid shift might be the common mechanism underlying these cross-species similarities, and predict that boundary-anchored shifts in human spatial memory can be observed during environmental deformations.

The phenomenon of boundary-tethered grid shift, predicted by our model and observed in a reanalysis of experimental data, could have various neural circuit realizations. Here we implemented a particular model of interactions between border, grid and place cells that gave rise to these shifts. Our model was feedforward [20], included an attractor network of grid cells [17], and generated place cells from grid cell output alone [19]. Although each of these components was motivated by prior work, our model is not intended as a complete recreation of entorhinal-hippocampal connectivity, but rather demonstrates how border cell input can give rise to the complex dynamics we describe, even in a relatively simple network. As such, our model excludes known connections that are not necessary for these dynamics. For example, our model lacks input to place cells from sources other than grid cells [32], as well as reciprocal connections from place to grid cells [33], both of which play important roles in developing and maintaining a functional spatial code. Nevertheless, our analysis of the experimental data shows that any alternative circuit model of the hippocampal formation should also incorporate a mechanism to produce boundary-tethered grid shifts, as the key phenomenon underlying deformation-induced grid and place field distortions.

The prevailing interpretation of these distortions has been that a rescaled environment induces a rescaled version of the cognitive map [10, 12, 14, 26, 30, 34–39]. Our results suggest an alternative interpretation. Grid and place representations are not rescaled or otherwise distorted during environmental deformations. Rather, these representations are dynamically updated when boundaries are contacted, as if one were feeling one's way through the dark. Our results add to a growing body of literature highlighting the unique ways in which environmental boundaries shape spatial representations in the brain [8–12, 14, 15, 25, 38, 40, 41], and suggest a reconceptualization of how these representations dynamically adapt to a changing world.

Acknowledgments

We are grateful to the laboratories of Edvard Moser, Kate Jeffery and Dori Derdikman for making the data from [9] and [8] available for our re-analysis. We also thank Eli Pollock, Niral Desai and Xuexin Wei for advice on implementing spiking border-grid connections.

References

[1] O'Keefe, J. & Dostrovsky, J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain research* **34**, 171–5 (1971).

[2] O'Keefe, J. & Nadel, L. *The Hippocampus as a Cognitive Map* (Oxford University Press, 1978).

[3] 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).

[4] McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I. & Moser, M.-B. Path integration and the neural basis of the 'cognitive map'. *Nature Reviews Neuroscience* **7**, 663–678 (2006).

[5] Moser, E. I. & Moser, M.-B. A metric for space. *Hippocampus* **18**, 1142–56 (2008).

[6] Buzsáki, G. & Moser, E. I. Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nature neuroscience* **16**, 130–8 (2013).

[7] Moser, E. I. *et al.* Grid cells and cortical representation. *Nature reviews. Neuroscience* **15**, 466–81 (2014).

[8] Barry, C., Hayman, R., Burgess, N. & Jeffery, K. J. Experience-dependent rescaling of entorhinal grids. *Nature neuroscience* **10**, 682–4 (2007).

[9] Stensola, H. *et al.* The entorhinal grid map is discretized. *Nature* **492**, 72–8 (2012).

[10] O'Keefe, J. & Burgess, N. Geometric determinants of the place fields of hippocampal neurons. *Nature* **381**, 425–8 (1996).

[11] Gothard, K. M., Skaggs, W. E. & McNaughton, B. L. Dynamics of mismatch correction in the hippocampal ensemble code for space: interaction between path integration and environmental cues. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **16**, 8027–40 (1996).

[12] Barry, C. *et al.* The boundary vector cell model of place cell firing and spatial memory. *Reviews in the neurosciences* **17**, 71–97 (2006).

[13] Muller, R. U. & Kubie, J. L. The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **7**, 1951–68 (1987).

[14] Lever, C., Burgess, N., Cacucci, F., Hartley, T. & O'Keefe, J. What can the hippocampal representation of environmental geometry tell us about Hebbian learning? *Biological cybernetics* **87**, 356–72 (2002).

[15] Hardcastle, K., Ganguli, S. & Giocomo, L. M. Environmental boundaries as an error correction mechanism for grid cells. *Neuron* **86**, 827–39 (2015).

[16] Pollock, E., Desai, N., Wei, X. & Balasubamanian, V. A mechanism for self-organized error-correction of grid cells by border cells. *COSYNE Abstracts 2017 (Salt Lake City)* (2017).

[17] Burak, Y. & Fiete, I. R. Accurate path integration in continuous attractor network models of grid cells. *PLoS Computational Biology* **5** (2009).

[18] Solstad, T., Boccara, C. N., Kropff, E., Moser, M.-B. & Moser, E. I. Representation of geometric borders in the entorhinal cortex. *Science (New York, N.Y.)* **322**, 1865–8 (2008).

[19] Solstad, T., Moser, E. I. & Einevoll, G. T. From grid cells to place cells: a mathematical model. *Hippocampus* **16**, 1026–31 (2006).

[20] Pilly, P. K. & Grossberg, S. Spiking neurons in a hierarchical self-organizing map model can learn to develop spatial and temporal properties of entorhinal grid cells and hippocampal place cells. *PloS one* **8**, e60599 (2013).

[21] Wei, X.-X., Prentice, J. & Balasubramanian, V. A principle of economy predicts the functional architecture of grid cells. *eLife* **4**, e08362 (2015).

[22] Mathis, A., Herz, A. V. M. & Stemmler, M. B. Multiscale codes in the nervous system: the problem of noise correlations and the ambiguity of periodic scales. *Physical review. E, Statistical, nonlinear, and soft matter physics* **88**, 022713 (2013).

[23] Sanzeni, A., Balasubramanian, V., Tiana, G. & Vergassola, M. Complete coverage of space favors modularity of the grid system in the brain. *Physical Review E* **94**, 062409 (2016).

[24] Grossberg, S. How does a brain build a cognitive code? *Psychological review* **87**, 1–51 (1980).

[25] Stensola, T., Stensola, H., Moser, M.-B. & Moser, E. I. Shearing-induced asymmetry in entorhinal grid cells. *Nature* **518**, 207–12 (2015).

[26] Evans, T., Bicanski, A., Bush, D. & Burgess, N. How environment and self-motion combine in neural representations of space. *The Journal of physiology* **594**, 6535–6546 (2016).

[27] Chen, X., He, Q., Kelly, J. W., Fiete, I. R. & McNamara, T. P. Bias in Human Path Integration Is Predicted by Properties of Grid Cells. *Current biology : CB* **25**, 1771–6 (2015).

[28] Bush, D., Barry, C., Manson, D. & Burgess, N. Using Grid Cells for Navigation. *Neuron* **87**, 507–20 (2015).

[29] Hartley, T., Trinkler, I. & Burgess, N. Geometric determinants of human spatial memory. *Cognition* **94**, 39–75 (2004).

[30] Hartley, T., Burgess, N., Lever, C., Cacucci, F. & O'Keefe, J. Modeling place fields in terms of the cortical inputs to the hippocampus. *Hippocampus* **10**, 369–79 (2000).

[31] Nadasdy, Z. *et al.* Context-dependent spatially periodic activity in the human entorhinal cortex. *Proceedings of the National Academy of Sciences* **114**, E3516–E3525 (2017).

[32] Wills, T. J., Cacucci, F., Burgess, N. & O'Keefe, J. Development of the hippocampal cognitive map in preweanling rats. *Science (New York, N.Y.)* **328**, 1573–6 (2010).

[33] Bonnevie, T. *et al.* Grid cells require excitatory drive from the hippocampus. *Nature neuroscience* **16**, 309–17 (2013).

[34] Raudies, F., Hinman, J. R. & Hasselmo, M. E. Modelling effects on grid cells of sensory input during self-motion. *The Journal of physiology* **594**, 6513–6526 (2016).

[35] Cheung, A. Probabilistic Learning by Rodent Grid Cells. *PLoS computational biology* **12**, e1005165 (2016).

[36] Bush, D. & Burgess, N. A hybrid oscillatory interference/continuous attractor network model of grid cell firing. *J Neurosci* **34**, 5065–5079 (2014).

[37] Bush, D., Barry, C. & Burgess, N. What do grid cells contribute to place cell firing? *Trends in neurosciences* **37**, 136–45 (2014).

[38] Krupic, J., Bauza, M., Burton, S. & O'Keefe, J. Framing the grid: effect of boundaries on grid cells and navigation. *The Journal of physiology* **594**, 6489–6499 (2016).

[39] Hasselmo, M. E., Hinman, J. R., Dannenberg, H. & Stern, C. E. Models of spatial and temporal dimensions of memory. *Current Opinion in Behavioral Sciences* **17**, 27–33 (2017).

[40] Krupic, J., Bauza, M., Burton, S., Barry, C. & O'Keefe, J. Grid cell symmetry is shaped by environmental geometry. *Nature* **518**, 232–5 (2015).

[41] Giocomo, L. M. Environmental boundaries as a mechanism for correcting and anchoring spatial maps. *The Journal of physiology* **594**, 6501–6511 (2016).

[42] Sargolini, F. *et al.* Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science (New York, N.Y.)* **312**, 758–62 (2006).

Methods

Model

Border layer: The border layer consisted of 36 units. First, the area near each wall in 4 allocentric directions (North, South, East, West) was divided into 'bricks' (see [16] for a similar treatment). Each brick extended 12 cm from the wall and covered 11.1% of the total environment length along that dimension. The jth unit received a uniform input $b_j = 0.1$ whenever the simulated rat was within one of three adjacent bricks. This input was was converted to stochastic spiking activity (see below). Each brick was assigned to a single border unit.

Grid layer: The grid layer, derived from the model of [17], consisted of 5 grid 'modules'. Each module consisted of a neural sheet with periodic boundary conditions, visualized as a torus. This neural sheet was composed of 45^2 identical 2 unit x 2 unit tiles (90^2 units per module). Each unit in a tile was associated with a particular direction (North, South, East, West), which determined both the movement-direction-specific excitatory input received, as well as its local connectivity. Movement-direction-specific excitatory input v_j to grid unit j was determined by

$$v_j = \gamma + g_m \left(d \cos\left(\theta - \phi_j\right)\right) \tag{1}$$

where d is the distance moved since the previous timestep, θ is the direction of movement, ϕ_j is the preferred direction of unit j, g_m is a gain factor specific to the module m to which to unit j belongs, and $\gamma = 0.6$ is a constant. Local connections within each module consisted of shifted radial inhibition, in which each unit inhibited all units within a 12 unit radius by a uniform weight of -0.02. The center of this radial inhibition output for each unit was shifted by 2 units away from that unit in a direction consistent with each units preferred direction. In the absence of other inputs, each grid module yields a hexagonal grid-like pattern of activation on the neural sheet, which is translated during movement at a rate proportional to the gain factor. Thus, to model modules with varying grid scales, the gain factor g_m of the mth module was set by

$$g_m = \frac{g_1}{2^{\frac{m-1}{2}}} \tag{2}$$

where $g_1 = 0.92$ is the gain of the smallest-scale module, module 1. This results in a geometric series of biologically-plausible [9] grid scales for each module.

Place layer: The place layer consisted of 64 units, subject to uniform recurrent inhibition from all place layer units with a weight of -0.15.

Border-to-grid connectivity: All grid units received additional excitatory feed-forward projections from all border units. These connections were initialized with random weights uniformly sampled from the range 0 to 0.00222, and developed through experience via Hebbian learning (see below and [16]).

Grid-to-place connectivity: Each place unit received additional excitatory feed-forward projections from 500 random grid units. These connections were initialized with random weights uniformly sampled from the range 0 to 0.022, and developed through experience via Hebbian learning (see below).

Model dynamics

Activation: The dynamics of the network was developed following the methods in [17]. The activation a_j of unit j was determined by first computing the total input b_j to unit j according to

$$b_j = \begin{cases} v_j + \sum_i^I a_i\, w_{ij} & \text{grid} \\ \sum_i^I a_i\, w_{ij} & \text{place} \end{cases} \tag{3}$$

where a_i is a variable quantifying activation of unit i, w_{ij} is the weight from unit i to unit j, and I enumerates all the units. (Note that some weights w_{ij} can be zero.) Also recall from above that a border unit receives a constant input when the rat is in a boundary region associated to the unit. The total input b_i was used to stochastically determine the spiking s_j of each unit j during the current timestep, according to

$$s_j = \begin{cases} 1 & \kappa\,(b_j - \beta_j)\,\mathrm{d}t > \mathrm{unif}(0,1) \\ 0 & \kappa\,(b_j - \beta_j)\,\mathrm{d}t \leq \mathrm{unif}(0,1) \end{cases} \tag{4}$$

where $\kappa = 500$ is a scale factor, β_j (border: $\beta = 0$; grid: $\beta = 0.1$; place: $\beta = 0.05$) is the spike threshold for unit j, $\mathrm{unif}(0,1)$ is a single draw from a random uniform distribution ranging from 0 to 1, and $\mathrm{d}t = 0.001$ sec is the length of each timestep. Finally, this spiking activity was integrated to update the activation variable a_j of unit j after each timestep according to

$$a_j = a_j - a_j \frac{\mathrm{d}t}{c} + \alpha s_j \tag{5}$$

where $\alpha = 0.5$ is a scale factor and $c = 0.02$ sec is the time constant of integration.

Hebbian learning: All Hebbian weights were updated by the competitive learning rule

$$w_{ij} = w_{ij} + \lambda a_j \left(((\xi_j - w_{ij})\, a_i) - (w_{ij} \sum_{n \neq i} a_n) \right) \tag{6}$$

where the sum is only over the set of units with nonzero Hebbian weights to unit j, $\lambda = 0.00001$ is the learning rate, ξ_j is a constant specifc to the connection type (border-to-grid: $\xi = 0.4$; grid-to-place: $\xi = 0.5$) [20, 24]. This rule results in competitive activity-dependent weight changes among incoming Hebbian connections, and leads over time to a total weight of ξ_j across incoming synapses.

Simulation details

Generating simulated rat paths: Because some of the deformed environments that we tested have not been experimentally studied, it was necessary to generate simulated rat paths, rather than using experimentally recorded paths. Open field paths were generated via a bounded random walk model, parameterized by speed and movement direction. At each timestep, unbiased normally-distributed random noise was added to both speed ($\sigma = 0.001$ cm/msec) and movement direction ($\sigma = 1°$/msec). To approximate actual rat exploration, speed was bounded to the range $[0, 40]$ cm/sec. If a step would result in the rat path crossing a boundary, random noise was again added repeatedly to the movement direction until the next step would no longer cross the boundary. Open field paths always began in the center of the environment, with the simulated rat stationary and facing a random direction. Linear track paths were generated as straight end to end laps at a constant speed of 20 cm/sec.

Familiarization: In all simulations, familiarization with the environment was mimicked by allowing the naive simulated rat to explore the environment for 60 min. Prior to familiarization, grid layer activity was allowed to settle into its grid-like attractor state for 2 sec without learning. Initialization of the grid layer was biased so that an axis of the settled grid network state would lie at an angle of -7.5° relative to east, consistent with experiments [25, 40]. Following familiarization, the model weights were saved so that all post-familiarization simulations could begin with the familiarized model.

Post-familiarization testing simulations: The model weights were reset to the state saved after familiarization, and the experienced virtual rat was allowed to explore each tested environment for 30 min. Grid layer activity was also initially reset to the familiar environment state corresponding to the rat's start location. Learning was turned off during the testing phase.

Analysis

Unit sampling: Due to computational constraints and the redundant nature of grid unit activity, only the spikes from 30 randomly chosen grid units in each module were recorded and analyzed during all simulations. All place units were recorded and analyzed.

Rate maps: Rate maps were created by first dividing the environment into 2.5 cm x 2.5 cm pixels. Then the mean firing rate within each pixel was calculated. Finally, this map was smoothed with an isotropic Gaussian kernel with a standard deviation of 1.5 pixels (3.75 cm) and square extent of 9 pixels $\times$ 9 pixels (22.5 cm $\times$ 22.5 cm). Pixels which were never visited were ignored during further analyses.

Autocorrelations and cross-correlations: *Autocorrelations* of rate maps were computed as described in previous reports [42]. Briefly, overlapping bins of the original rate map and a shifted version of itself were correlated at a series of single pixel (2.5 cm) step lags. *Cross-correlations* were computed similarly, except that two different rate maps, rather than two copies of the same rate map, were correlated.

Grid scale: To compute grid scale we first averaged the autocorrelations of all grid units within a module. Next, we computed the mean distance from the center of the autocorrelation to the center of mass of the six closest surrounding peaks. In cases where the grid period was larger than the size of the environment thus obscuring the periodicity, grid scale was instead estimated by multiplying the scale of the next smaller module by $\sqrt{2}$, reflecting the parameters set in the attractor model creating the grid.

Grid rescaling factor: The grid rescaling factor during each deformation trial was computed separately for each unit by comparing rescaled versions of the deformed environment rate map to the familiar environment rate map. Following [9], the deformed rate map was uniformly rescaled to a series of chamber lengths, ranging from 10 cm below the smaller of the deformed and familiar chamber lengths, through 10 cm above the larger of these chamber lengths in 5 cm (2 pixel) increments. This yielded a set of rescaled rate maps for each unit, which were aligned with the familiar environment ate map by each of the two displaced boundaries. For each rescaled map and

alignment, we computed the correlation r between the rescaled and familiar rate maps:

$$r = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(v_{i,j} - \bar{v})(v'_{i,j} - \bar{v'})}{\sqrt{\sum_{i=1}^{I}\sum_{j=1}^{J}(v_{i,j} - \bar{v})^2}\sqrt{\sum_{i=1}^{I}\sum_{j=1}^{J}(v'_{i,j,} - \bar{v'})^2}} \tag{7}$$

where v is the rescaled rate map, v' is the familiar rate map, i, j run over pixels in the overlapping regions of these maps, and $\bar{v}$ and $\bar{v'}$ indicate the mean firing rates across the overlapping pixels. The grid rescaling factor was defined as the ratio between the rescaled chamber length that gave the highest correlation and the the familiar chamber length, across either alignment.

Place field statistics: First, place fields were identified as regions of contiguous pixels with firing rates exceeding 3 Hz. Field size was computed as the total area covered (number of pixels $\times$ the area of a single pixel) by place fields, summing over any disjoint fields. Directional modulation was measured as the difference in peak firing pixel location when comparing rate maps of activity in one movement direction along the chamber length versus the opposite direction.

Reanalysis of experimental data

A complete description of the experiments was provided in [8, 9]. To test these data for the presence of grid shifts during environmental deformations, we first divided the spiking activity of each cell according to the most recent boundary contact (North, South, East, West). Boundary contact was defined as the rat being within 12 cm of a boundary. Spiking activity prior to boundary contact at the beginning of the trial was ignored. Next, four separate rate maps were created, one for each most recently contacted boundary. To quantify grid shift along a particular dimension for each cell, the rate maps of opposing boundaries perpendicular to the chosen dimension were cross-correlated at a series of lags in single pixel steps. Lag range was set by the extent of the largest deformation: ± 12 pixels (± 30 cm) for data from [8], ± 20 pixels (± 50 cm) for data from [9], along both dimensions. The absolute distance from the center to the nearest peak of this cross-correlation along the chosen dimension was computed as the measure of grid shift. The nearest peak was defined by first partitioning the cross-correlation into 'blobs' of contiguous pixels which had correlations of at least 20% of the maximum value. Then, the location with the maximum correlation value within the blob nearest to the center was taken as the nearest peak. Finally, the average shift measured across both dimensions during familiar trials was subtracted from the shifts measured during each deformation trial.

Data and code availability

All simulations were conducted with custom-written Matlab scripts. These scripts, the simulation results presented here, and similar results from an additional simulation of each experiment (to confirm reliability), are available from the authors upon reasonable request. All reanalyzed data are available upon request from the corresponding authors of the relevant papers.

Supplemental Information for:

Environmental deformations dynamically shift the spatial metric of the brain

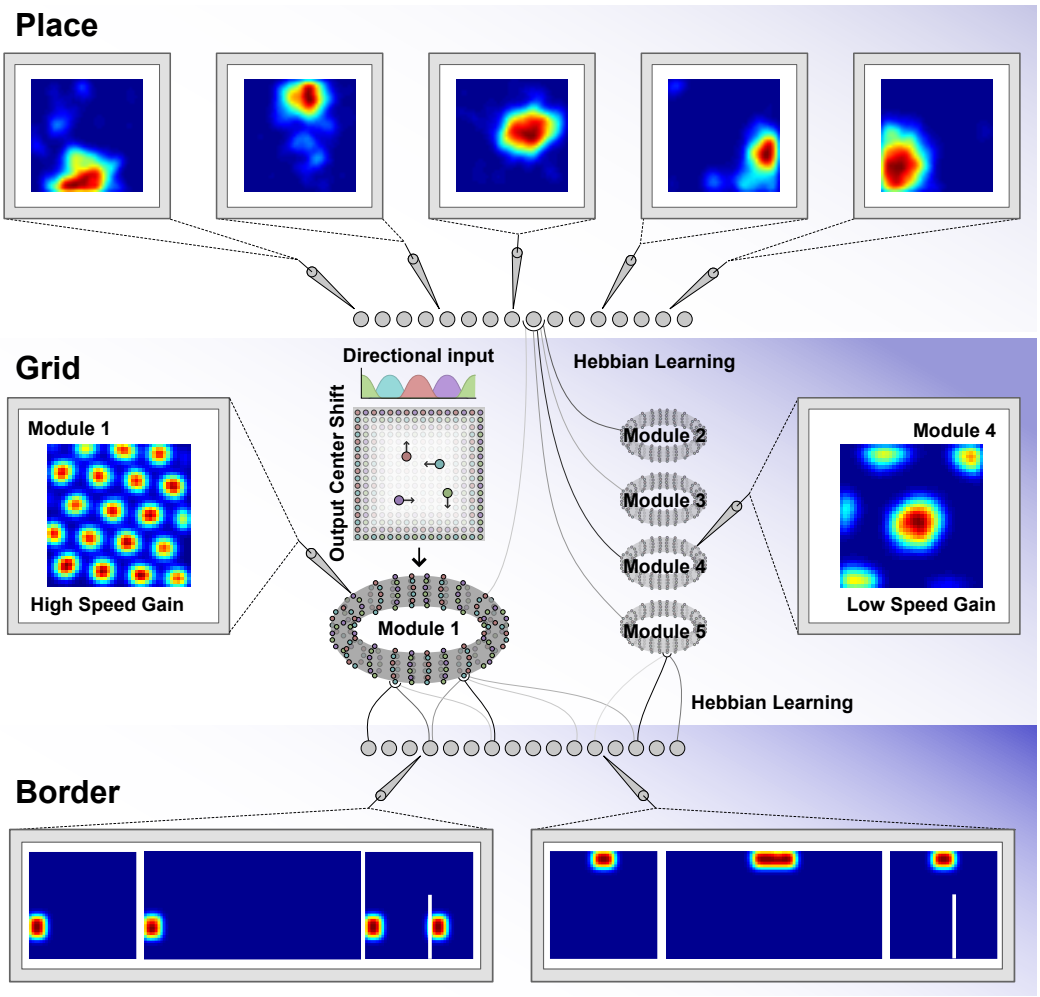

Figure S1: Schematic of the border, grid, and place cell network model. See Methods for a complete description of model interactions. Note the behavior of border units during environmental deformations – stretching the environment stretches the border field, and inserting a barrier duplicates border fields with the same allocentric relationship to the original and inserted boundaries.

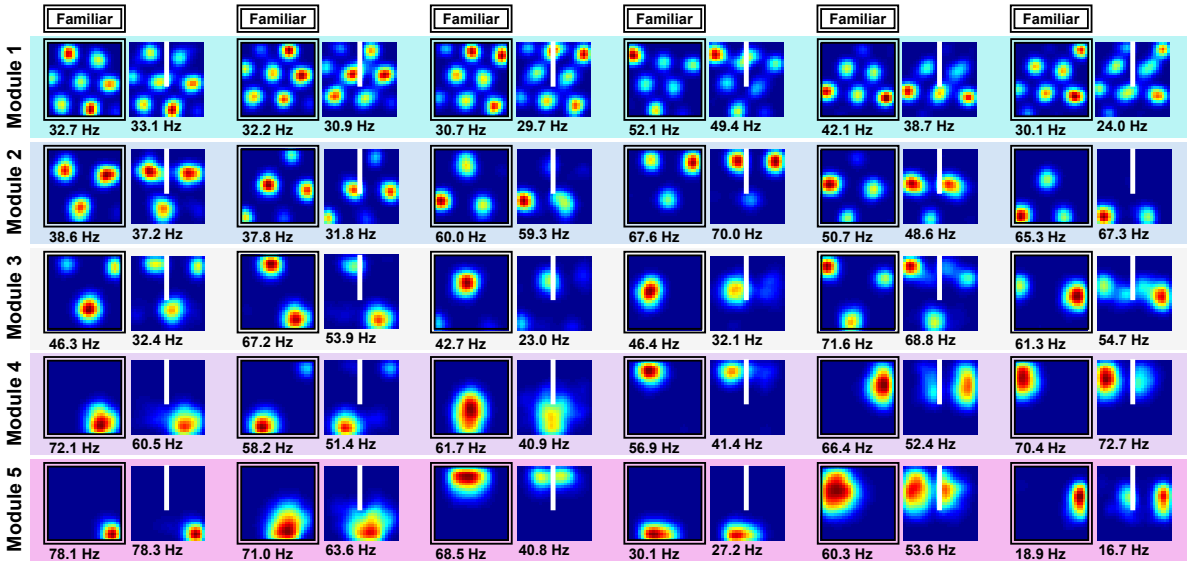

Figure S2: Grid unit activity during insertion of a new boundary in an open environment. Examples of grid unit activity during exploration of a familiar chamber and boundary insertion (white line) – 6 units shown from each module. Distortions are minimal in the time-averaged rate maps of small-scale grid units (matching experiment [18]), but become apparent in the activity of large-scale grid units. Peak firing rate noted below the lower left corner of each map.

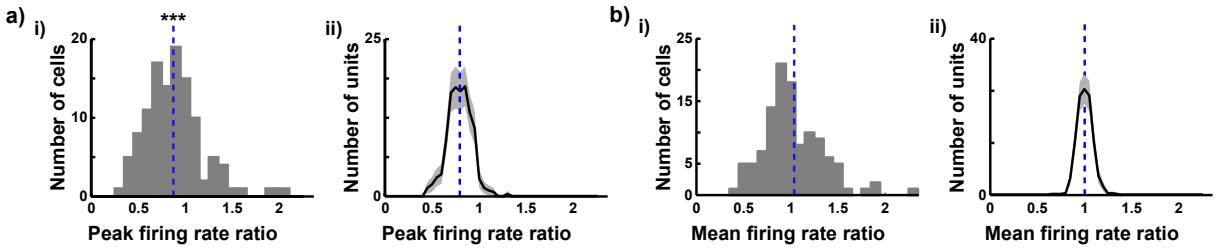

Figure S3: Grid cell peak firing rate is reduced during environmental deformations. Boundary-tethered grid shift predicts a reduction in peak firing rate during chamber deformations (since shifting grid vertices will often misalign), but no change in mean firing rate (as the density of fields remains unchanged). **a**) (i) Histogram of the peak firing rate ratios for each grid cell averaged across all deformations (all data from [8] and [9] combined). Peak firing rate ratio computed for each cell as the peak firing rate (see text) averaged across all deformations, divided by the peak firing rate averaged across familiar environment trials. The distribution of peak firing rate ratios is significantly shifted below 1 (mean = dashed blue line; mean ± SEM = 0.873 ± 0.028), indicating that grid cell peak firing rates are reduced during environmental deformations (one-sample t-test comparing the peak firing rate ratio vs. 1, $t(114) = 4.6$, $p < 0.001$). (ii) Histogram of model grid unit peak firing rate ratios measured under the same deformation conditions as [8] and [9]. Mean (solid black line) ±1 standard deviation (grey shading) across 1000 random draws of model grid units from each experimental condition, with the number of units from each condition matched to the experimental data. The model predicts a decrease in peak firing rate during deformations (mean = dashed blue line) as seen in panel (i). **b**) (i) Histogram of the mean firing rate ratio (see text) for all grid cells and deformations (data from [8] and [9] combined). The distribution of mean firing rate ratios was not significantly shifted away from 1 (mean = dashed blue line; mean ± SEM = 1.037 ± 0.032), indicating that grid cell mean firing rates did not significantly differ during environmental deformations (one-sample t-test comparing the mean firing rate ratio vs. 1, $t(114) = 1.16$, $p = 0.25$). (ii) Histogram of model grid unit mean firing rate ratios measured under the same deformation conditions as [8] and [9]. The model predicts no change in mean firing rates during deformations (mean = dashed blue line) as seen in (i).

Table S1: Statistical information for the grid shift tests depicted in Fig. 4b.

Dataset	Trained in	Deformed dimension(s)	Measured dimension	Test	df	t	p
Barry et al., 2007	Square	east-west	north-south	2-tailed t-test vs. 0	39	0.17	0.864
Barry et al., 2007	Square	east-west	east-west	2-tailed t-test vs. 0	39	2.83	0.007
Barry et al., 2007	Square	north-south	north-south	2-tailed t-test vs. 0	39	3.03	0.004
Barry et al., 2007	Square	north-south	east-west	2-tailed t-test vs. 0	39	0.75	0.459
Barry et al., 2007	Square	both	north-south	2-tailed t-test vs. 0	39	2.53	0.015
Barry et al., 2007	Square	both	east-west	2-tailed t-test vs. 0	39	3.36	0.002
Barry et al., 2007	Rectangle	east-west	north-south	2-tailed t-test vs. 0	22	1.34	0.194
Barry et al., 2007	Rectangle	east-west	east-west	2-tailed t-test vs. 0	22	2.77	0.011
Barry et al., 2007	Rectangle	north-south	north-south	2-tailed t-test vs. 0	22	0.70	0.492
Barry et al., 2007	Rectangle	north-south	east-west	2-tailed t-test vs. 0	22	0.47	0.641
Barry et al., 2007	Rectangle	both	north-south	2-tailed t-test vs. 0	22	2.07	0.050
Barry et al., 2007	Rectangle	both	east-west	2-tailed t-test vs. 0	22	1.22	0.239
Stensola et al., 2012	Square	east-west	north-south	2-tailed t-test vs. 0	50	0.22	0.823
Stensola et al., 2012	Square	east-west	east-west	2-tailed t-test vs. 0	50	2.51	0.015

Table S2: Statistical information for the boundary rate map alignment tests depicted in Fig. 4c.

Dataset	Trained in	Deformed dimension(s)	Boundary rate map	Test	n	sign stat	p
Barry et al., 2007	Square	east-west	West	2-tailed sign test vs. 0.5	39	5	0.000002
Barry et al., 2007	Square	east-west	East	2-tailed sign test vs. 0.5	39	6	0.000010
Barry et al., 2007	Square	north-south	South	2-tailed sign test vs. 0.5	40	7	0.000040
Barry et al., 2007	Square	north-south	North	2-tailed sign test vs. 0.5	39	9	0.001100
Barry et al., 2007	Square	both	West	2-tailed sign test vs. 0.5	40	7	0.000040
Barry et al., 2007	Square	both	South	2-tailed sign test vs. 0.5	40	5	0.000001
Barry et al., 2007	Square	both	East	2-tailed sign test vs. 0.5	40	5	0.000001
Barry et al., 2007	Square	both	North	2-tailed sign test vs. 0.5	39	9	0.001100
Barry et al., 2007	Rectangle	east-west	West	2-tailed sign test vs. 0.5	23	8	0.210000
Barry et al., 2007	Rectangle	east-west	East	2-tailed sign test vs. 0.5	23	3	0.000500
Barry et al., 2007	Rectangle	north-south	South	2-tailed sign test vs. 0.5	23	5	0.010600
Barry et al., 2007	Rectangle	north-south	North	2-tailed sign test vs. 0.5	23	10	0.677600
Barry et al., 2007	Rectangle	both	West	2-tailed sign test vs. 0.5	23	6	0.034700
Barry et al., 2007	Rectangle	both	South	2-tailed sign test vs. 0.5	23	9	0.404900
Barry et al., 2007	Rectangle	both	East	2-tailed sign test vs. 0.5	23	7	0.093100
Barry et al., 2007	Rectangle	both	North	2-tailed sign test vs. 0.5	23	10	0.677600
Stensola et al., 2012	Square	east-west	West	2-tailed sign test vs. 0.5	51	19	0.091900
Stensola et al., 2012	Square	east-west	East	2-tailed sign test vs. 0.5	51	13	0.000621