

A

**Diagram of the two session task structure**

Session 1		Session 2			
<b>Treadmill Training</b> - Walking on treadmill - Walking in virtual environments	<b>Practice City</b> - Navigation (4 rounds) - JRD pointing task (4 rounds) - <b>Criterion:</b> mean JRD error on round 3 + 10° - Questionnaire (cybersickness)	<b>Training Phase</b> - Navigation - JRD pointing task - Until criterion - Repeat for next city	<b>Re-exposure Phase</b> - Navigation (all 3 cities)	<b>Pre-fMRI Test Phase</b> - JRD task (all 3 cities) - Ensure performance is better than chance for all 3 cities - if not: re-expose and re-test - if true: go to fMRI	<b>fMRI Tasks</b> - JRD task - Active baseline task - Resting state

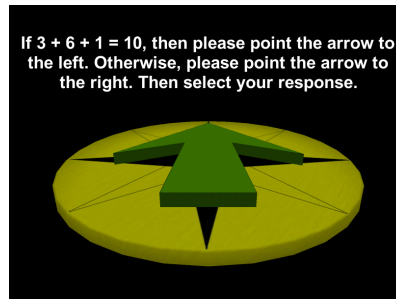
B

**JRD Task (3 cities)**



C

**Active Baseline**



D

**Resting State**

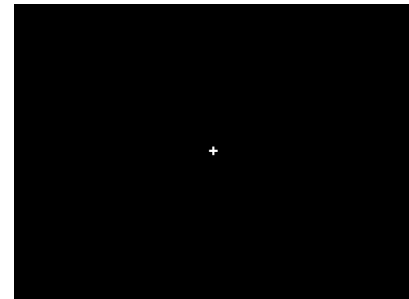


Figure S1 related to Figure 1. (A) A diagram of the two session task structure. During fMRI scanning, participants performed three tasks (shown in panels B-D). (B) The judgments of relative directions (JRD) task in which participants were asked to making pointing responses regarding the locations of landmarks in the environment. (C) Our active baseline task in which the visual features were matched to the JRD task but participants in this task were asked to make math judgments (note, the direction of the correct answer was randomized on each trial). (D) The resting-state task, in which participants were instructed to keep their eyes open and to fixate on the central fixation cross but to feel free to think about whatever came to mind.

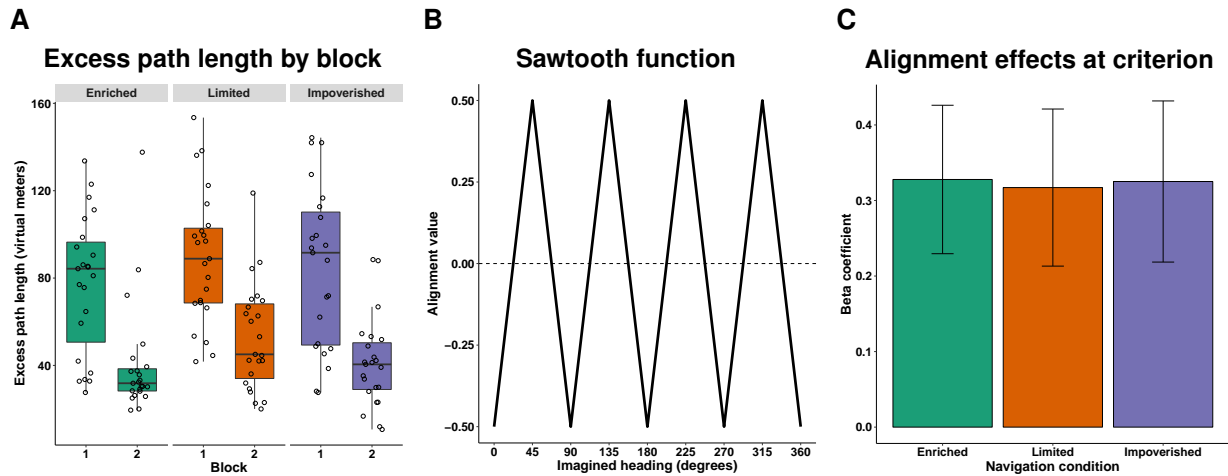


Figure S2 related to Figure 1. Additional behavioral analyses support the notion that performance was similar across body-based cues conditions. (A) Excess path length decreased from the first to the second block of the task ( $\beta = -0.63$ , null model AIC = 7,078.5, block model AIC = 7,031.6,  $X^2(1, N = 23) = 48.9$ ,  $p < 0.0001$ ), suggesting that participants rapidly acquired navigationally relevant spatial representations. However, the interaction between block and body-based cues conditions did not reach significance (block model AIC = 7,031.6, block  $\times$  cues interaction model AIC = 7,034.6,  $X^2(4, N = 23) = 5.0$ ,  $p = 0.29$ ), and the effect of block was stable across all three body-based cues conditions (enriched:  $\beta = -0.65$ , null model AIC = 2,326.9, block model AIC = 2,311.0,  $X^2(1, N = 23) = 17.9$ ,  $p < 0.0001$ ; limited:  $\beta = -0.53$ , null model AIC = 2,412.9, block model AIC = 2,403.4,  $X^2(1, N = 23) = 11.5$ ,  $p = 0.0007$ ; impoverished:  $\beta = -0.73$ , null model AIC = 2,351.5, block model AIC = 2,332.2,  $X^2(1, N = 23) = 21.4$ ,  $p < 0.0001$ ). Each dot represents the mean excess path length for a single participant (note, the data were analyzed using a generalized linear mixed-effects model with a gamma function for our dependent variable and a log link function; see STAR Methods; Lo et al., 2015). (B) We used a sawtooth function for our analysis of alignment effects. (C) We observed a significant main effect of alignment on pointing error ( $p < 0.0001$ ). The alignment  $\times$  body-based cues interaction did not reach significance ( $p = 0.27$ ), and the main effect of alignment was stable across body-based cues conditions (all  $p \leq 0.005$ ; figure shows the beta coefficients with standard errors of the estimates).

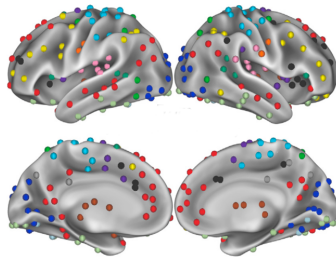
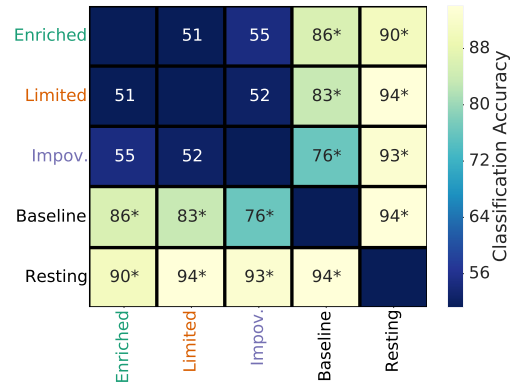
**A****Power nodes****B****Classification results: task in time-series**

Figure S3 related to Figure 2. (A) We used nodes from Power et al. (2011) in our analysis (figure from Cole et al., 2014). (B) We also ran our classification analysis without regressing out task-based activation because such activations undoubtedly contain task-relevant information. Thus, leaving such activations in the time-series could potentially increase the sensitivity of discriminating between task conditions (e.g., between the different JRD task blocks). However, the pattern of classification results was the same as the background correlativity analysis. We observed significant classification accuracy between the resting state and all other task states (active baseline task: 94%,  $p < 0.0001$ ; JRD-enriched: 90%,  $p = 0.0002$ ; JRD-limited: 94%,  $p < 0.0001$ ; JRD-impooverished: 93%,  $p < 0.0001$ ) and between the active baseline task and the JRD retrieval tasks (JRD-enriched: 86%,  $p < 0.0001$ ; JRD-limited: 83%,  $p < 0.0001$ ; JRD-impooverished: 76%,  $p = 0.0008$ ). Note, the classification accuracies observed here were, on average, numerically better than those in the background correlativity analysis, thus suggesting that this approach could potentially be more sensitive to task-related changes in network interactions. The comparison between JRD task blocks as a function of body-based cues, however, did not produce significant classification accuracy (enriched vs. limited: 51%,  $p = 0.89$ ; enriched vs. impooverished: 55%,  $p = 0.31$ ; limited vs. impooverished: 52%,  $p = 0.70$ ; please compare this figure to Figure 2A). \* indicates two-tailed permutation  $p < 0.05$  (note, these also survived FDR correction for multiple comparisons).

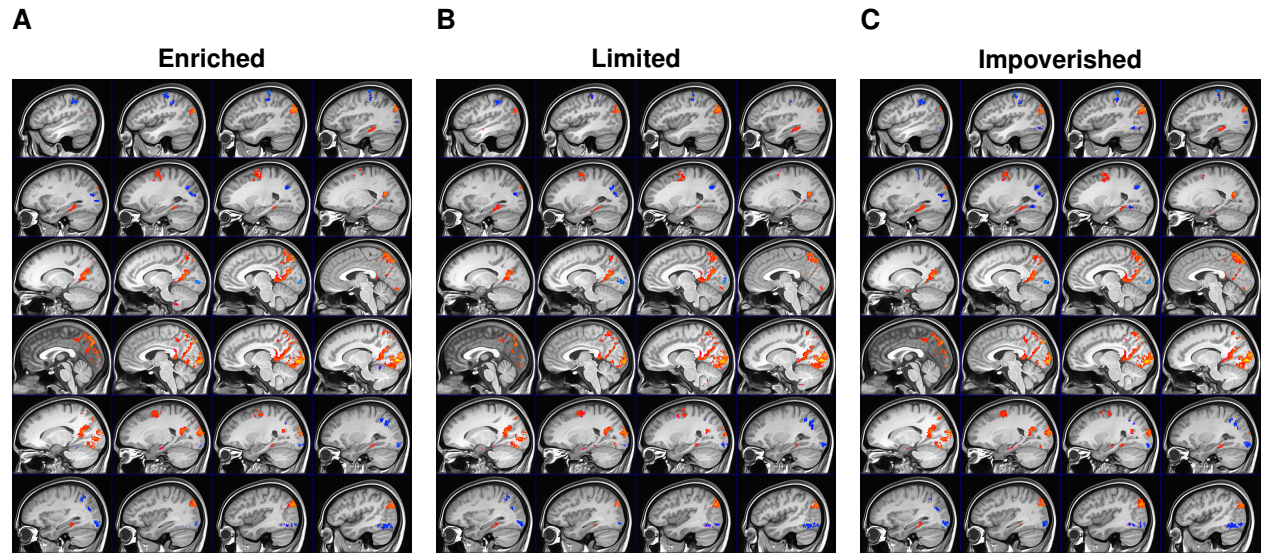


Figure S4 related to Figure 4. The activation maps for the JRD task vs. the active baseline task were highly similar across the three body-based cues conditions (all dice overlaps  $\geq 0.69$ ; voxel-wise  $p < 0.001$  [corresponds to FDR  $q$  of 0.030, 0.032, 0.030 for the Enriched, Limited, and Impoverished conditions, respectively], cluster-corrected threshold  $p < 0.01$  [15 voxel extent]).

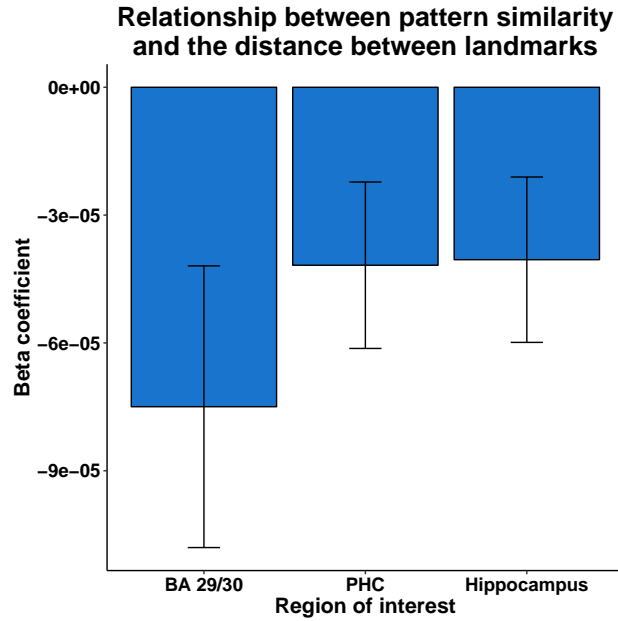


Figure S5 related to Figure 5. Evidence that regions of the spatial network contain information related to distances between landmarks but that this relationship is not significantly different between body-based cues conditions. A mixed-model revealed a significant relationship between pattern similarity and the Euclidean distance between landmarks in the retrosplenial cortex (BA 29/30), parahippocampal cortex (PHC), and a functionally defined hippocampal region of interest (hippocampus; all  $p < 0.05$ ) but the interaction between distance and body-based cues did not reach significance in any of these regions (all  $p \geq 0.19$ ). The figure shows the beta coefficients with standard errors of the estimates.