Neurocognitive stages of spatial cognitive mapping measured during free exploration of a large-scale virtual environment

Markus Plank,1 Joseph Snider,1 Erik Kaestner,2 Eric Halgren,2,3* and Howard Poizner1,2*
1Institute for Neural Computation, University of California, San Diego, La Jolla, California; 2Interdepartmental Neuroscience Program, University of California, San Diego, La Jolla, California; and 3Departments of Radiology, Neurosciences, and Psychiatry, University of California, San Diego, La Jolla, California

Submitted 10 February 2014; accepted in final form 2 November 2014

Plank M, Snider J, Kaestner E, Halgren E, Poizner H. Neurocognitive stages of spatial cognitive mapping measured during free exploration of a large-scale virtual environment. J Neurophysiol 113: 740–753, 2015. First published November 5, 2014; doi:10.1152/jn.00114.2014.—Using a novel, fully mobile virtual reality paradigm, we investigated the EEG correlates of spatial representations formed during unsupervised exploration. On day 1, subjects implicitly learned the location of 39 objects by exploring a room and popping bubbles that hid the objects. On day 2, they again popped bubbles in the same environment. In most cases, the objects hidden underneath the bubbles were in the same place as on day 1. However, a varying third of them were misplaced in each block. Subjects indicated their certainty that the object was in the same location as the day before. Compared with bubble pops revealing correctly placed objects, bubble pops revealing misplaced objects evoked a decreased negativity starting at 145 ms, with scalp topography consistent with generation in medial parietal cortex. There was also an increased negativity starting at 515 ms to misplaced objects, with scalp topography consistent with generation in inferior temporal cortex. Additionally, misplaced objects elicited an increase in frontal midline theta power. These findings suggest that the successive neurocognitive stages of processing allocentric space may include an initial template matching, integration of the object within its spatial cognitive map, and memory recall, analogous to the processing negativity N400 and theta that support verbal cognitive maps in humans.

as humans and other animals move through their surroundings, they integrate multisensory information from visual, vestibular, and proprioceptive cues to form inner representations of spatial environments (Chance et al. 1998; Klatsky 1998). These representations organize our behavior and form the core context of our subjective experience (O’Keefe and Nadel 1978). Neural firing correlates of spatial representations have been extensively studied in the rodent hippocampal formation as the animals move through environments. These neural correlates take the form of place (O’Keefe et al. 1998), grid (Hafting et al. 2005), boundary (Solstad et al. 2008), and head direction cells (Cho and Sharp 2001)—all of which stabilize rapidly in novel environments (Hafting et al. 2005) and are sensitive to alterations of the environment (Muller et al. 1987; Smith and Mizumori 2006). Theta rhythms organize this spatially related cell firing, gating memory encoding and retrieval and synchronizing activity across multiple cortical areas, including parietal and frontal cortices (Hasselmo and Stern 2014; Siapas et al. 2005).

Consistent with the central role of space in mental experience, O’Keefe and Nadel (1978) interpreted hippocampal place cells, and the entirety of hippocampal recording and lesion studies, as implementing a “cognitive map.” To accommodate the already large literature demonstrating nonspatial memory deficits after medial temporal lesions in humans, they posited that the hippocampus in the dominant hemisphere supported a verbal cognitive map to complement the spatial cognitive map supported by the nondominant hemisphere in humans and both hemispheres in rodents.

In subsequent years, there have been many studies demonstrating the neurocognitive stages that integrate words into the dynamic cognitive context. The most prominent stage is the N400, commonly recorded in the scalp EEG but also apparent in MEG (Halgren et al. 2002) and intracranial local field potentials (Nobre et al. 1994; Smith et al. 1986) and correlated with fMRI (Dale et al. 2000) and unit activity (Heit et al. 1988). The N400 is evoked by words and other semantic stimuli and is smaller when the cognitive context allows the word to be integrated more easily (Brown and Hagoort 1993; Halgren 1990; Holcomb et al. 2002; Kutas and Federmeier 2011). The responses to words can also be modulated at earlier latencies, ~200 ms, in a less contextual fashion, by their lexical frequency (Sahin et al. 2009) or broad semantic category (Chan et al. 2011; Pulvermuller et al. 1999). The N400 is often followed by the P600, especially when initial associations yield erroneous conclusions (Kuperberg 2007) such as puns (Marinkovic et al. 2011), and/or by theta activity. Unlike the N400, theta is larger to actual words and has been interpreted as reflecting sustained memory retrieval processes (Marinkovic et al. 2012). The N200/N400 to words are generated mainly in the left antero-ventral temporal, temporo-parietal, and postero-ventral prefrontal cortices (Halgren et al. 1994a, 1994b, 2002; Nobre et al. 1994; Smith et al. 1986). The P600 and theta appear to be generated in classical language areas, as well as more dorsolateral prefrontal, parietal, and especially cingulate cortices (Wang et al. 2005). In particular, the medial temporal structures in and around the hippocampal formation show strong activation during the N400 (Nobre et al. 1994; Smith et al. 1986), supporting the prediction of O’Keefe and Nadel (1978) that these areas implement a verbal cognitive map, or semantic hub for the cortex (Patterson et al. 2007).
Recording studies of spatial cognitive maps in humans under realistic conditions have been rare, largely because of the head and body restraints usually required in fMRI, MEG, or even EEG recordings. Because spatial awareness requires integration of visual, vestibular, and proprioceptive cues during locomotion (Klatzky 1998), investigating the brain processes underlying spatial learning is best done during fully mobile exploration in naturalistic settings. We have resolved the recording problem and have examined brain activity in fully mobile human participants interacting in an immersive large-scale virtual environment (VE). On day 1, EEG was recorded continuously as participants actively explored a novel environment containing 39 distinctive objects at fixed locations. On day 2, subjects were confronted with objects that were in the expected location or were misplaced. We hypothesized that comparing the response to expected versus misplaced objects would reveal neurocognitive processing stages for spatial cognitive mapping, namely, an early first-pass processing of allocentric space followed by an integration of the object within its spatial cognitive map that would be associated with an N400-like processing negativity.

MATERIALS AND METHODS

Subjects

Twelve young, healthy adults without history of psychiatric or neurological disorder participated in the experiment [3 women, 9 men; age = 24.8 ± 4 yr (mean ± SD)]. All participants were right-handed, as measured with the Edinburgh Handedness Inventory (Oldfield 1971). Participation was reimbursed at a rate of $20/h. The study was approved by the UCSD Human Research Committee and was conducted according to the principles expressed in the Declaration of Helsinki. Written informed consent was obtained from all participants.

Experimental Paradigm

Subjects actively explored a novel VE rendered on a wide-field-of-view head-mounted display (HMD). Subjects physically moved at will in a large-scale, richly textured virtual room (~4.0 × 5.0 m) containing numerous objects resting on shelves, tables, and the floor (see Fig. 1). This virtual room, located on an “aircraft carrier,” was the same size as the real-world space in which subjects were walking.

The experiment comprised two sessions lasting 2 h each, which were conducted on two subsequent days. On day 1, participants first explored the room for 10 min with no instruction except “explore the environment.” Then, during five subsequent blocks, opaque virtual bubbles were placed around 39 different objects in the room. The subjects walked up to one bubble at a time (indicated by turning green) and popped it by touching it with their hand, thereby uncovering the object underneath (the objects had been present during the initial exploration but not obscured by the bubbles). The popping of the virtual bubble was used as a time-locking event for the event-related potential (ERP) analysis and was logged into an event file together with a time stamp that allowed for off-line synchronization and alignment with the EEG system. As a cover task, subjects indicated their interest in the object with a virtual slider. A block ended when all the objects were uncovered and visible. For each of these five “interest” blocks, the order of bubble popping was varied pseudorandomly. However, the objects remained the same, thus ensuring both multiple paths through the VE and opportunity for unsupervised learning of the object locations.

While the first day was dedicated to unsupervised spatial exploration, the second day involved testing subjects’ memory of the environment, in particular whether objects were still placed in their expected spatial locations. Since subjects were completely naive about the memory aspect of the experiment when they performed the interest rating task on day 1, the knowledge of the environment on day 2 must have resulted from unsupervised learning. After an overnight’s sleep to allow for memory consolidation, on day 2 the subjects entered the same virtual room. This time, the locations of a subset of 13 of the 39 objects were randomly interchanged from their positions on day 1.

Fig. 1. A: experimental environment. Participants freely explored a virtual storage room (A1; bird’s eye view, never shown to participants), matched in size (4 × 5 m) to the physical room in which the participant was physically present (A2). The virtual environment was presented via a head-mounted display (A3; ego view). During the session, the continuous EEG of 72 active electrodes was collected (A4; 2-dimensional voltage map). B: trial sequence: the experimental task was to walk toward a green bubble (B1), which disappeared upon being touched with the right hand (B2), also accompanied by a popping sound, thereby revealing the hidden object and providing the time-locking event for neurophysiological analyses. After a randomized time interval of 1,000–1,250 ms, a virtual slider bar appeared in front of the subject (B3). On day 1, subjects used the slider bar to rate “how interesting” the object was. On day 2, subjects again approached and popped green bubbles but used the slider to indicate whether the revealed object was correctly placed or misplaced. After adjusting the slider (no time limit), participants lifted their hand above the slider (B4) to confirm their adjustment, which activated the next green bubble. The time interval between 225 and 25 ms prior to bubble pop (starting at B1) was used as baseline period for the event-related potential (ERP) analysis.
As on day 1, subjects walked to a bubble covering an object when it turned green and touched the bubble, exposing the object. Subjects then indicated how certain they were that the object was the same one that was in that location on day 1, again by adjusting the virtual slider. This process was repeated over five blocks of trials, with each block lasting 5–8 min. A different set of 13 objects were shuffled in each block.

**Details on the VE**

The hardware generating the VE comprised a 24-camera PhaseSpace Impulse active infrared-emitting diode (IRED) motion capture system (http://www.phasespace.com). With this system, the three-dimensional (3D) positions of the right hand, head, and torso were recorded. In addition, an Intersense (www.intersense.com) InertiaCube orientation sensor was mounted on the HMD in order to collect more accurate orientation data and smooth out shifts of the display due to head rotations. Auditory stimulation was provided by means of a 20-speaker Ambisonic Auralizer Sound System (www.worldviz.com). The VE was rendered on a Sensics (www.sensics.com) xSight 6123 HMD, which provided a panoramic, stereo view of the VE (120° horizontal and 45° vertical field of view). The VE was programmed in Vizard 3.0 (www.worldviz.com) with a total end-to-end latency of <40 ms (time from real space movement to update of the visual scene in the HMD) and spatial precision on the order of millimeters (see Snider et al. 2013a for a detailed description of the system). There was an initial constant latency of 15 ms due to the motion capture system, which is well under latencies that would produce a noticeable lag between actual and rendered scenes in the HMD. Additionally, there was a constant 25-ms latency for rendering the visual scene (Snider et al. 2013a), which was subtracted from the latencies of all events.

The dimensions of the virtual room matched those of the real room in which the experiment took place, so that the gain between movements in the real and virtual rooms was 1:1. The virtual room resembled a storage room with a large empty space in the center surrounded by shelves and tables on which a variety of objects rested. The objects were richly detailed and textured 3D renderings, some of which were created from photographs taken on the U.S.S. Midway naval museum, San Diego, CA (http://www.midway.org). All objects were appropriate to the context of an aircraft carrier setting. Exits were obstructed with virtual obstacles except for an open virtual door at which all blocks began. Subjects could look through the door to see a virtual flight deck but were instructed not to pass beyond the limits of the room. The position of the subject’s head, as tracked by the IRED system, and its orientation from the inertial sensor were directly fed into the VE and refreshed the rendering of the main view in real time and with a 1:1 mapping of real onto virtual space. The IRED and inertial sensors were combined to maximize the responsiveness and stability of the system for improved immersion. The subject’s right hand was tracked and rendered as a visible, 5-cm-diameter orange ball that the subjects used to indicate interest (day 1) or certainty (day 2) on a continuously variable slider that appeared in the VE after bubble popping. All of the interactions took place such that touching any object in the real environment was avoided, therefore minimizing any potential haptic-vision mismatch. Further details on the environment can be found in Snider et al. (2013a), which analyzed the periods when the subjects were walking between objects. Here we analyze events that are time-locked to the bubble pops.

**EEG Acquisition and Analysis**

EEG data were recorded from 64 Biosemi Active electrodes placed in a flexible cap according to the extended International 10–20 System (American Electroencephalographic Society 1991) with a sampling rate of 1,024 Hz. Eight external electromyographic (EMG) electrodes were added to help identify muscular artifacts and were placed as follows. Four electrooculogram (EOG) electrodes were mounted on the supra- and infraorbital ridges of the right eye as well as lateral to the outer canthi of right and left eyes. Two electrodes were further attached to the right and left neck at the height of the 7th cervical vertebra, monitoring activity of neck muscles, particularly the trapezius. A final two electrodes were mounted on the left and right mastoids. For all participants, electrode locations were digitized with a FASTRAK system (Polhemus, Colchester, VT) in combination with the Locater Suite (Source Signal Imaging, San Diego, CA).

The participants wore a lightweight backpack that held the battery-driven amplifiers. Active EEG electrodes were connected to the amplifiers via short ribbon cables, comparable to stationary EEG setups. Amplifiers were connected to the recording computer via a long fiber-optic cable, but the already-amplified EEG signals were not contaminated by electrical or mechanical artifacts on this cable.

Data processing was accomplished with custom MATLAB scripts (The MathWorks, Natick, MA) for EEGLAB (Delorme and Makeig 2004) and the Mass Univariate Statistics Toolbox (Groppe et al. 2011). For visualization purposes, BrainVision Analyzer 2 (Brain Products, Gilching, Germany) and EEGLAB were employed. In EEGLAB, data were referenced off-line to average mastoids. Data were high-pass filtered at 1 Hz to remove offset and trend, low-pass filtered at 55 Hz, and epoch time-locked to bubble pop (epochs windows covered the time interval between −2.025 and +1.975 ms). Epoched EEG data were visually inspected for artifactual activity arising from motor processes, punctual bursts of EMG activity, and other nonstereotypical processes, and epochs containing artifacts were excluded from further analyses. On average, 82% of the misplaced and 78% of the correctly placed trials survived the artifact rejection procedure (difference not significant, t-test$_{df} = 1.1 = 0.8$; $P = 0.25$). After removal of data sections containing artifacts identified via visual inspection, EEG data were further inspected for artifacts with independent component analysis (ICA) and dipole analysis (Delorme and Makeig 2004; Hammon et al. 2008). Extended Infomax ICA allowed for identification of temporally maximally independent component (IC) time courses that could be categorized into muscular and oculomotor activity as well as cortically based processes (e.g., see Gwin et al. 2010). ICs were analyzed with respect to scalp topography and frequency characteristics during the epoched time intervals, and those that displayed features indicative of artifacts were excluded from the back-projection to sensor space. Ocular ICs were identified according to the criteria described in the literature (Delorme and Makeig 2004; Jung et al. 1998) as having a spatially focal scalp projection and high power at high frequencies (20–50 Hz and above). Dipole models that best explained the scalp topography were fit to each of the remaining components with the DIPFIT plug-in for EEGLAB and localized within a four-shell spherical boundary element head model (BEM) of the Montreal Neurological Institute standard brain (an average of 152 MRI brain scans of healthy individuals). We morphed the digitized 3D electrode array to the head model by scaling and rotating the head coordinate system so that the digitized anatomical reference points were aligned with the head model anatomical reference points as provided by the DIPFIT 2 plug-in for EEGLAB (see http://sccn.ucsd.edu/wiki/A08: DIPFIT) and as used elsewhere (Gwin et al. 2010, 2011; Gwin and Ferris 2012). All ICs but one were fitted with a single dipole, resulting in minimal residual variances between IC scalp map and dipole scalp projection, following the procedures previously (Delorme et al. 2012; Gwin et al. 2010, 2011; Makeig et al. 2002; Sipp et al. 2013). Across all subjects, the residual variance of one of the occipital ICs could be further reduced by a dual dipole fit (with a symmetrical constraint right/left), since its component topography indicated a bilateral distribution. However, for the clustering (see below), EEGLAB automatically selected only the right or left dipole aspect, preventing the multiple use of a dual dipole for clustering. Only ICs whose dipoles resided within the vicinity of the
eyeballs (ocular ICs) or within the brain volume of the head model and with no more than 15% residual variance between IC scalp map and dipole scalp projection were retained. On average, 9.7 ± 2.8 (mean ± SD) ocular and brain-based ICs per participant were identified (range 4 – 14). Cleaned EEG data were generated by back-projecting the time course of activity within the remaining ICs to the surface electrodes. The ocular ICs were excluded from the backprojection in order to remove artifacts due to eye movements. However, the ocular ICs were retained for the analysis in source space; since the clustering procedure can isolate ICs related to eye movements consistently across subjects into specific clusters, this permitted us to determine whether differences between experimental conditions are related to eye movements (see below). This procedure allows attenuation of artifacts from the EEG without having to reject the entire trial during which an artifact occurred (Jung et al. 2000).

**ERP difference waves.** Artifact-free component activations were back-projected to the electrodes and baseline corrected (−225 to −25 ms). Average ERP waveforms were computed for correctly placed versus misplaced objects, and differences waves were generated. The amplitude difference waves of misplaced minus correctly placed objects across the 64 cap electrodes were submitted to a randomization-based masking procedure referred to as the cluster mass permutation test (Bulmore et al. 1999; Groppe et al. 2011; Maris and Oostenveld 2007; Worsley et al. 1996). This nonparametric testing procedure was selected because of its robustness against outliers and different trial numbers for correctly placed and misplaced objects. In this procedure, the t-score of the actual grand average difference waveform as generated by averaging the individual difference waves between misplaced and correctly placed objects is compared with a distribution of t-scores of difference waves of surrogate grand averages without knowing their assignment to misplaced and correctly placed conditions. This procedure was accomplished as described in Groppe et al. (2011). Differences in the real grand average difference amplitudes were considered significant for t-scores corresponding to a global α-level of 5%.

The large number of comparisons in the point-by-point bootstrapping procedure introduces false positives across both electrodes (see below) and time. We assume that temporal false positives occur as a Poisson-like process, without any pattern, but the already complicated signal is temporally continuous (from earlier filtering) and we cannot expect parametric statistics. Continuity means that “neighboring” significant points are more likely to be significant just because the data have been filtered to remove large differences, even if the significance is spurious. The goal is to estimate the minimum duration that separates “neighbors” (neighbor time). Thus we bootstrap the duration of spurious intervals (10,000 resamples from the difference amplitudes above) by swapping the “correctly placed” and “misplaced” labels within each electrode and finding the longest contiguous time interval with apparent significance. The upper 95% confidence bound of these durations estimates the neighbor time (α = 0.05). The average neighbor time was 26 ± 4 ms (mean ± SD), which corresponds to frequency of 39(6) Hz, or somewhat slower than the high-pass filter at 55 Hz, as we would expect. Any significant intervals in the original masked data with duration shorter than the neighbor time were removed. The approach for temporal bootstrapping described above only applies to multiple comparisons within an electrode, but false positives may still exist across electrodes. To simplify the problem across electrodes we clustered them into 10 regions (see ERP Results in Sensor Space below) and used parametric tests on the peak values at latencies of interest via a linear mixed model in R (version 3.0.1) and the lme4 package version 1.0-6 (Bates et al. 2014). The clustering was necessary for the linear mixed models to converge.

To test for significant differences, models with and without fixed effects were compared in marginal order with a χ²-test on the difference in profiled deviance (log likelihood penalized for the number of terms in the model). Post hoc tests were accomplished by using the multcomp package to control for multiple comparisons across condition, cluster, and latency (Hothorn et al. 2008). Significance was defined as P < 0.05 for all parametric tests.

**Source estimation.** Standardized low-resolution brain electromagnetic tomography (sLORETA, http://www.uzh.ch/keyins/loreta.htm) was used to estimate the cortical 3D distribution of current density based on the scalp topography (Fuchs et al. 2002; Pascual-Marqui et al. 1994). For ERPs, sLORETA was computed using the mean amplitude during a 40-ms interval centered on the peak of the ERP difference wave between misplaced and correctly placed objects. Additionally, the following procedure complemented the surface-based reconstruction of distributed sources: After IC decomposition, all single-trial IC time series (~2,025 to 1,975 ms time-locked to bubble pop) of the identified brain components as well as components capturing oculomotor processes were transformed into a spectrographic image using Morlet wavelets in the frequency range between 2 and 55 Hz (EEGLAB functions std_precomp and newtimef) using the following parameters and values: “ffreq” = [2, 55], “time-limits” = [-2,025 1,975], “baseline” = 0, “cyles” = [3 0.5], “padratio” = 1, and “nmitesout” = 400. Each single-trial component activity time series was then transformed to a spectrographic image, yielding for each trial the total power changes (see David et al. 2006 for details). The event-related spectral perturbations (ERSPs; Kliem et al. 2007) of each trial were baseline corrected and averaged.

To identify sets of similar components across all subjects, IC processes were clustered with a k-means cluster algorithm (Hartigan and Wong 1979) implemented in the EEGLAB STUDY toolbox. For this purpose, an N-dimensional cluster position vector for each IC was created to measure “distances” between all ICs in the defined cluster space based on particular IC characteristics such as IC log spectra, single-trial ERPs, IC inverse weights across the electrodes, equivalent dipole locations, single-trial ERSPs, and single-trial intertrial coherence (ITC), or phase locking factor (Onorton et al. 2006). Parameters were chosen to emphasize 3D location of sources and spectral perturbation/coherence that we expected to be most relevant given the present task: A weighting factor of 1 was assigned to the IC log spectrum (2–55 Hz) and to the inverse weight topography, a weighting factor of 3 was assigned to the single-trial ERP waveforms of the IC, a weighting factor of 4 was assigned to the single-trial ERSP and ITC data, and a weighting factor of 15 was assigned to the dipole location. This set of weighting factors has been shown in the literature to balance the relative influence of the various IC measures, so that no one measure has disproportionate influence on cluster assignment (see Gramann et al. 2010; Plank et al. 2010). The resulting joint vector was reduced to 10 principal dimensions with principal component analysis (PCA) (Jung et al. 2001) and was set to 7 final clusters based on existing studies on spatial cognition using the same ICA and clustering approach (Gramann et al. 2010; Plank et al. 2010). Of the seven defined clusters, two clusters emerged that reflected ocular processes that were common to all subjects (horizontal eye movements and blinks); the remaining five clusters could be localized in the cortical domain. The two ocular clusters were retained and fed into the same processing pipeline as the cortical clusters in order to determine whether differences between experimental conditions would be specifically related to cortical ICs but not eye movements. Components that were located further than 3 standard deviations from any of the cluster centroids were classified as outliers and removed from further analysis. For sLORETA reconstructions of the resulting IC cluster centroids, the mean topographical weights of all ICs contributing to a cluster-based findings are presented in ERSP Results in Source Space.

In contrast to standard EEGLAB functionality, where the computation of ERSP differences simply takes the actual difference between conditions and shuffles the pixels (the number of shuffles is determined by the significance threshold), we masked the actual differences between misplaced and correctly placed objects with the randomization-based mask according to Maris and Oostenveld (2007) described above, revealing differences statistically different from a...
distribution of mean surrogate differences (for each time-frequency bin, or “pixel”) derived from randomized experimental conditions. Whenever the actual ERSP difference exceeded 1.96 standard deviations of the randomized ERSP distribution, we classified the ERSP difference between misplaced and correctly placed objects as statistically significant ($P < 0.05$).

RESULTS

Behavioral Results
Given that on day 2 subjects had to select between one of two alternative choices (correctly placed vs. misplaced), chance performance was 50%. On day 2 subjects performed well above chance, correctly identifying 86.86% ± 5.94% (mean ± SD) of all objects as misplaced or correctly placed compared with day 1 (minimum = 76.60%, maximum = 95.90%). No relationship between interest ratings on day 1 and performance on day 2 was found ($r = -0.0035$, $P = 0.99$; Spearman test). Additionally, an ANOVA was performed to measure dependence of the rated interest of the objects across subjects and showed that no object was consistently rated as either interesting or not ($P = 0.2$, bootstrap Kolmogorov-Smirnov (bKS) test detailed in Snider et al. 2013b). Similarly, no location was more memorable than any other ($P = 0.85$, bKS test). In sum, subjects performed well above chance, despite a delay of 24 h between exposure and testing and the lack of explicit learning instructions on day 1. Furthermore, a subject’s rating of how interesting the individual objects were did not predict subsequent memory of the object/location relations, and there were no marked differences in interest across objects.

ERP Results in Sensor Space

Figure 2 presents the grand average ERP waveforms evoked by correctly placed versus misplaced objects in specified electrode clusters: 1: left occipito-parietal (PO3, PO7, O1), 2: right occipito-parietal (PO4, PO8, O2), 3: left centro-parietal (CP3, CP5, TP7, P3, P5, P7), 4: medial centro-parietal (CPz, Pz), 5: right centro-parietal (CP4, CP6, TP8, P4, P6, P8), 6: left fronto-central (FC1, FC3, FC5, C1, C3, C5), 7: medial fronto-central (FCz, Cz), 8: right fronto-central (FC2, FC4, FC6, C2, C4, C6), 9: left frontal (F1, F3, F5), and 10: right frontal (F2, F4, F6).

To initially identify differences in the ERP waveform to correctly placed or misplaced objects, we identified two peaks of activity in the grand average difference wave centered at around 160 and 535 ms after the object was revealed. We ran a linear mixed model with random effects allowing for subject-level variability in response to condition (correctly placed or misplaced), electrode cluster, and peak latency (N160 or N535). Consistent with the waveforms shown in Fig. 2, the

---

Fig. 2. Grand average ERP waveforms evoked by correctly placed vs. misplaced objects in specified electrode clusters in the time interval −225 to +975 ms time-locked to bubble pop: 1 and 2, left and right occipito-parietal; 3, 4, and 5, left, medial, and right centro-parietal; 6, 7, and 8, left, medial, and right fronto-central; 9 and 10, left and right frontal. While the N160 effect (light gray area) is significant in posterior clusters 1, 2, and 5, the N535 effect (dark gray area) is present in all electrode clusters 1–6 and 8 (all except anterior regions). Significant differences: *$P < 0.05$; **$P < 0.01$. 

J Neurophysiol • doi:10.1152/jn.00114.2014 • www.jn.org
Table 1. Mean amplitudes for N160 and N535 ERP components at each of the 10 electrode clusters

<table>
<thead>
<tr>
<th>No.</th>
<th>Electrode Cluster</th>
<th>N160 Misplaced, µV</th>
<th>N160 Correctly Placed, µV</th>
<th>N535 Misplaced, µV</th>
<th>N535 Correctly Placed, µV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L occipito-parietal</td>
<td>-0.66 ± 0.59**</td>
<td>-1.30 ± 0.49**</td>
<td>-0.41 ± 0.40**</td>
<td>0.23 ± 0.28**</td>
</tr>
<tr>
<td>2</td>
<td>R occipito-parietal</td>
<td>-0.92 ± 0.61**</td>
<td>-1.68 ± 0.51**</td>
<td>-0.15 ± 0.43**</td>
<td>0.45 ± 0.31**</td>
</tr>
<tr>
<td>3</td>
<td>L centro-parietal</td>
<td>-0.22 ± 0.34</td>
<td>-0.51 ± 0.27</td>
<td>-0.90 ± 0.25**</td>
<td>-0.27 ± 0.17**</td>
</tr>
<tr>
<td>4</td>
<td>M centro-parietal</td>
<td>-0.10 ± 0.70*</td>
<td>-0.56 ± 0.55*</td>
<td>-0.69 ± 0.54**</td>
<td>0.06 ± 0.38**</td>
</tr>
<tr>
<td>5</td>
<td>R centro-parietal</td>
<td>-0.29 ± 0.33</td>
<td>-0.79 ± 0.27</td>
<td>-0.45 ± 0.25**</td>
<td>0.07 ± 0.18**</td>
</tr>
<tr>
<td>6</td>
<td>L fronto-central</td>
<td>0.49 ± 0.38</td>
<td>0.40 ± 0.34</td>
<td>-1.17 ± 0.28**</td>
<td>-0.55 ± 0.19**</td>
</tr>
<tr>
<td>7</td>
<td>M fronto-central</td>
<td>0.92 ± 0.76</td>
<td>0.73 ± 0.68</td>
<td>-1.00 ± 0.56</td>
<td>-0.38 ± 0.38</td>
</tr>
<tr>
<td>8</td>
<td>R fronto-central</td>
<td>0.63 ± 0.36</td>
<td>0.37 ± 0.32</td>
<td>-0.86 ± 0.27*</td>
<td>-0.33 ± 0.18*</td>
</tr>
<tr>
<td>9</td>
<td>L frontal</td>
<td>0.56 ± 0.59</td>
<td>0.59 ± 0.57</td>
<td>-1.36 ± 0.41</td>
<td>-0.80 ± 0.28</td>
</tr>
<tr>
<td>10</td>
<td>R frontal</td>
<td>0.74 ± 0.57</td>
<td>0.62 ± 0.54</td>
<td>-1.07 ± 0.42</td>
<td>-0.58 ± 0.28</td>
</tr>
</tbody>
</table>

Values are mean ± SD amplitudes for the N160 and N535 event-related potential (ERP) components at each of the 10 electrode clusters. Significant differences between correctly placed and misplaced objects within each electrode cluster: *P < 0.05, **P < 0.01.

Results of the linear mixed model indicated that the triple interaction of condition, cluster, and peak latency was highly significant [χ²(9) = 28.36, P < 0.001]. Given the structure of the experiment, we interpret this interaction as a difference in activity between conditions (correctly placed or misplaced object) with variable strength depending on electrode cluster and peak latency. Post hoc tests (as indicated in Fig. 2) bore this out, with the earlier N160 component showing an ~0.5-µV increased negativity over lateral parieto-occipital regions for correctly placed objects. The N535 showed widespread differences extending from occipital to fronto-central regions with ~0.5-µV increased negativity for misplaced objects (opposite in polarity compared with the N160 effect). Table 1 provides the mean amplitudes for the N160 and N535 components at each of the 10 electrode clusters.

Decrease in centro-parietal N160 amplitude for misplaced objects. The mean amplitude differences between conditions were examined in more detail with nonparametric tests over the entire space of electrodes and times. Permutation statistics (Fig. 3) revealed a significant difference in ERP amplitudes between correctly placed and misplaced conditions in the time interval 145–175 ms after bubble pop. This effect of a decreased negativity for misplaced objects was strongest in electrodes over central and bilateral occipital/parietal regions. Source estimates of the difference wave as returned by sLORETA were consistent with generation in medial parietal areas, particularly precuneus (BA 7/31; Talairach coordinates X, Y, Z = −9, −52, 29; Fig. 4A).

Increased amplitude of N535 over left hemisphere and central parietal regions for misplaced objects. In addition, there were significant ERP amplitude differences in the time range 515–555 ms apparent in a large number of electrodes (49 of 64 electrodes), peaking in medial parietal scalp leads and expanding toward bilateral central and left frontal electrodes. This difference was due to an increased negativity for misplaced objects in the given time interval. Estimated sources of the difference wave by sLORETA were most prominent in bilateral fusiform gyrus (BA 19, 37) and/or lingual gyrus (BA 18; Talairach coordinates X, Y, Z = −19, −60, −8; Fig. 4B).

ERSP Results in Source Space

In addition to changes in ERPs, we hypothesized that theta would differ between conditions. To test this hypothesis, we examined time-frequency dynamics by means of ERSSPs at the level of clustered ICs. As mentioned in MATERIALS AND METHODS, on average, 9.7 ± 2.8 (mean ± SD) ocular and brain-based ICs per participant were identified (range: 4–14) and used for clustering, resulting in the specifications of the cluster centroids shown in Table 2.

Cluster 1 comprised ICs whose inverse weights projected most strongly toward frontal midline electrodes and exhibited power in the (3–6 Hz) low theta frequency range, displaying well-established characteristics of “frontal midline theta” (Hsieh and Ranganath 2014) (Fig. 5, A and C). By sLORETA, the scalp distributions of these components were consistent with generation in medial frontal cortex (BA 6) and/or the anterior cingulate gyrus (BA 23, 24; Talairach coordinates X, Y, Z = −19, −60, −8; Fig. 4B).
Based on the permutation-based masking procedure described in MATERIALS AND METHODS, group-level ERSPs, that is, event-locked changes in spectral power for this source, could be computed in component space for correctly placed and misplaced objects (Fig. 5D). Figure 5D shows that there is increased delta and theta power shortly after the bubble pop for both expected and misplaced objects. This increase in total power was significantly greater for misplaced objects within the low theta frequency range (3–6 Hz) between 145 and 475 ms after bubble pop. No other cluster exhibited differences between conditions in the theta range, including cluster 2, which localized to BA 18/31 (very close to the previously observed N160). Cluster 2 also showed an ERP difference with increased negativity to correctly placed objects peaking at 170 ms, similar to the N160 component.

DISCUSSION

Basic Findings

The present study investigated neural processes underlying learning of the spatial location of objects when subjects freely walked about a complex, naturalistic, large-scale virtual environment. Because of technical advances, we were able to record event-related EEG in conjunction with behavioral measures during naturalistic, full-body exploration in 3D space. We experimentally probed acquired spatial representations by altering the location of a subset of objects from those previously learned in an unsupervised fashion. Our primary hypothesis was that amplitude modulations during the N400 time interval would differentiate brain responses to objects placed in the expected location versus in an unexpected location. Al-

Table 2. Specifications of IC clusters identified by the k-means clustering procedure

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Talairach Coordinates</th>
<th>Brodmann Area</th>
<th>Ss</th>
<th>ICs</th>
</tr>
</thead>
<tbody>
<tr>
<td>VEOG</td>
<td>X, Y, Z = 12, 14, 31</td>
<td>Vertical eye movement cluster</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>HEOG</td>
<td>X = 0, Y = −22, 22</td>
<td>Horizontal eye cluster</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>1</td>
<td>X = −4, Y = 11, 37</td>
<td>Left anterior cingulate gyrus</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>X = 0, Y = −72, 22</td>
<td>BA 18/31</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>X = 20, Y = −22, 43</td>
<td>Bilateral (pre)cuneus</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>X = −30, Y = 20, 13</td>
<td>Right cingulate gyrus</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>X = −10, Y = −12, 56</td>
<td>Left inferior frontal gyrus</td>
<td>9</td>
<td>17</td>
</tr>
</tbody>
</table>

Values are specifications of the independent component (IC) clusters identified by the k-means clustering procedure, including X, Y, Z locations of the cluster centroid specified within the stereotaxic coordinate system of Talairach and Tournoux (1988) and their anatomical region defined in the Brodmann area system (Brodmann 1925). Nos. of subjects (Ss) and ICs within each cluster are also provided.
though this ERP component has mainly been studied with words (Kutas and Federmeier 2011), it has also been observed to other meaningful complex stimuli such as faces (Eimer 2000), mathematical computations (Galfano et al. 2004), sounds (Aramaki et al. 2010; Painter and Koelsch 2011), and visual scenes (Amoruso et al. 2013; Ganis and Kutas 2003). As N400 amplitude has been conceptualized to reflect cognitive processes coordinating the “feed-forward flow of stimulus-driven activity with a state of the distributed, dynamically active neural landscape that is (semantic) memory” (Kutas and Federmeier 2011), it is seemingly modulated by the contextual congruity of a stimulus with its surroundings. Expanding these conceptualizations to the spatial domain, we predicted an N400-like component to be elicited when objects have to be integrated with their spatial context and its amplitude to be increased as participants’ spatial expectancies of object-location associations were violated and objects were encountered whose locations were interchanged from those learned on the previous day. This hypothesis was supported: We found an N400-like component that resembled lexico-semantic N400s in morphology and scalp topography, with some differences in peak latency. Source estimates were consistent with an increased negativity being generated in ventral posterior temporal areas.

Previous studies in rodents and humans have shown the central role of theta oscillation for active maintenance and retrieval of mental representations (Buzsaki 2005; Cornwell et al. 2008; Klimesch 1999; O’Keefe and Burgess 1999; Sauseng et al. 2002), particularly in tasks involving spatial attention and memory (Bischof and Boulanger 2003; Caplan et al. 2003; de Araujo et al. 2002; Jones and Wilson 2005; Kahana et al. 1999). In humans, frontal-midline theta power is modulated by task demands (Jensen and Tesche 2002), and we expected an increase in theta power when encountering alterations to a previously learned environment. This hypothesis was supported: misplaced objects elicited an increase in frontal-midline theta power, particularly in the 3–5 Hz range. Source estimates of temporally maximally independent components based on single dipole fitting in a BEM as well as distributed source estimation in sLORETA were consistent with this increase originating in anterior cingulate cortex.

Finally, we found a negative-going deflection occurring ~160 ms after presentation of an object that differentiated between correctly placed and misplaced objects, resembling an N200, which has been linked to cognitive processes monitoring stimulus frequency (Sahin et al. 2009; Squires et al. 1975) and broad semantic category (Chan et al. 2011; Pulvermüller et al. 1999). In contrast to semantic studies in which the N200 was found to originate from inferior frontal or temporal regions, source estimates in our paradigm suggested generators in medial parietal cortices.

By analogy with possibly homologous components observed during verbal processing, the early N200-like component may reflect first-pass processing of object-location information.
along the dorsal stream. The later N400-like component would reflect the cognitive comparison of the actual object location with a memory-based representation of its expected location in the environment. Finally, the late theta activity may organize retrieval of spatial information from the mental representation.

An N400-Like ERP Component Reflects Contextual Integration in the Spatial Domain

Our results indicate the presence of an N400-like ERP component peaking in midline parietal contacts and extending toward temporal regions that is elicited by violation of object-location expectations that had been learned in an unsupervised fashion on the previous day. Its morphology, temporal characteristics, and spatial distribution suggest similarities with the N400 for semantic stimuli (Kutas and Federmeier 2011). Neural sources of the N400, as identified by intracranial recordings and MEG, have been localized in temporal areas (e.g., anterior medial, middle superior, and ventral temporal) and dorsolateral prefrontal cortex, bilaterally, but primarily in the left hemisphere (Dale et al. 2000; Halgren et al. 1994a, 1994b, 2002; Nobre et al. 1994; Smith et al. 1986; Tse et al. 2007; Van Petten and Luka 2006). These findings have been complemented by hemodynamic studies suggesting involvement of left superior temporal and left inferior frontal gyri (e.g., Van Petten and Luka 2006).

The present finding that the reconstructed source of our N400-like component was ventral temporal regions may derive from the cognitive demands for learning and recalling spatial locations and associated objects from an internal representation. As noted by O’Keefe and Nadel (1978), the spatial cognitive map is highly abstract and flexible, since the world around us is constantly changing, objects in the environment might vanish, or viewing conditions might change from one encounter to the next (Christou and Bulthoff 2000). We note that subjects in the present task learned novel objects and their locations not simply from sequences of stimuli presented on a two-dimensional screen but rather from active, unsupervised spatial exploration in an immersive virtual reality environment. Subjects approached the objects from different angles and were required to view them from different perspectives in order to accomplish the task. The identification of object location correspondences approached from different viewpoints and cognitive demands for recalling spatial locations required associative processes in order to bind object features such as shape, color, contrast, etc. and extract multidimensional object features such as identity and location. The present source estimates of the N400-like ERP component suggest the involvement of extrastriate areas that play a major role for object identity (Kohler et al. 1998; Moscovitch et al. 1995) and object recognition (Price et al. 1996; Rosier et al. 1997; Stewart et al. 2001; Zelkowicz et al. 1998), in concert with the hippocampal formation further down the ventral stream. Interestingly, several studies have pointed out the critical role of Brodmann area 37 (occipital-temporal cortex) for the detection of violations of spatial expectancies (e.g., Faillenot et al. 1997; Goel et al. 2004; Smith et al. 1995).

Another aspect to be discussed is the difference in latency between the classic N400 component and the N400-like component identified in the present study at 515–555 ms. Kutas and Federmeier (2011) characterize the N400 as a monophasic negativity between 200 and 600 ms, peaking at 400 ms after stimulus onset (relative to a 100 ms prestimulus baseline). In semantic paradigms, the N400 amplitude was shown to begin at around 250 ms, peak at around 400 ms, and return to baseline at around 600 ms after stimulus (e.g., Kutas and Hillyard 1980). In the study of Eimer (2000), negativities were apparent between 300 and 500 ms after face onset. In studies using mathematical computations (Galfano et al. 2004), the core N400 time window was defined between 300 and 400 ms, whereas in paradigms utilizing complex visual scenes (Amoruso et al. 2013; Ganis and Kutas 2003; Sitnikova et al. 2003, 2008), N400-like waves were found to be present in the time range between 325 and 600 ms. One of the major reasons for latency shifts and modulations in the duration of the effect may be stimulus complexity. In studies that utilize rapid stimulus sequences with brief presentations, the N400 seems to occur earlier than in studies using complex stimuli. In the present paradigm, continuous visual input, self-determined choice of perspective on the task-related objects, and subject-driven duration of exploration might have been accompanied by a comparable shift in component latency due to the prolonged presentation of the objects, allowing for extended interaction that unfolded over several hundred milliseconds (see Sitnikova et al. 2008).

A second issue relates to the duration of the classic N400 versus that found in the present study. As mentioned above, depending upon the type of stimuli presented, the duration of the N400 varied from about 200 ms to 400 ms. We found significant differences between conditions of correctly placed and misplaced objects in the time interval 515–555 ms; however, the difference was nonzero for ~200 ms (Fig. 2). This duration is in the same range as that found for other complex stimuli.

Taken together, the functional sensitivity, latency range, topographic distribution, as well as the estimated source locations suggest a classification of the present findings in the functional family of the N400. We agree with the notion put forward by Kutas and Federmeier (2011, p. 623) that the label “N400-like” is solely used as a “heuristic label for stimulus-related brain activity in the 200–600 ms poststimulus-onset window with a characteristic morphology and, critically, a pattern of sensitivity to experimental variables—and hence a common functionality,” which also applies in the context of our paradigm. Since comparable N400 effects have been identified for nonlinguistic stimuli such as faces (Halgren et al. 1994a, 1994b), mathematical computations (Galfano et al. 2004), static visual scenes (Amoruso et al. 2013; Ganis and Kutas 2003; McPherson and Holcomb 1999), and complex video stimuli (Sitnikova et al. 2003, 2008), the present results support an interpretation of the N400 as reflecting a general cognitive process across different semantic domains.

Differences Between Conditions Are Observable as Early as 145 ms After Stimulus

N400 amplitude is influenced by a wide range of distant associations, across sensory modalities, semantic domains, and memory systems (Brown and Hagoort 1993; Halgren 1990; Holcomb et al. 2002; Kutas and Federmeier 2011). Consistent with this broad access to widely distributed information, laminar recordings have found that the N400 is generated by...
activity in upper cortical layers reflecting second-pass associative processing (Halgren et al. 2006). The same sites that generate the N400 often generate a negativity at \( \sim 200 \) ms in layer 4 that reflects first-pass feedforward processing (Halgren et al. 2006). This component can be influenced by word characteristics such as lexical frequency (Hauk et al. 2006; Sahin et al. 2009) and even general semantic category (Chan et al. 2011). In the present study, misplaced objects evoked ERP differences in a component peaking at \( \sim 160 \) ms, suggesting that the expected spatial location of an object may be a fundamental characteristic processed in the first sweep of activity through the cortex. This rapid processing of object location would have clear adaptive advantage in responding rapidly within a changing spatial environment. Medial parietal and retrosplenial areas were estimated as the possible source for this effect. This region, strongly related to both medial temporal and parietal cortices, may play a key role in relating allocentric and egocentric spaces, to guide adaptive interaction with the outside world (Cho and Sharp 2001; Maguire 2001; Vann and Aggleton 2004).

**Frontal Midline Theta Is Sensitive to Object-Location Mismatches**

Our paradigm might be considered as a form of a spatial paired-associate learning task in which participants learned associations between objects and locations. Paired-associate learning has been demonstrated not only when subjects are presented with a stimulus item and must reproduce the associated pair item (Yim et al. 2013) but also when only recognition of the pair item is required (Bunsey and Eichenbaum 1996; Ragland et al. 1995). EEG studies of paired-associate learning have generally identified modulations of the contingent negative variation (CNV) and P300 ERP components over central-parietal areas (Peters et al. 1977). In the present study, theta power was modulated by the experimental manipulation. Although there are clear similarities of our paradigm to paired-associate learning, there are also differences. In our paradigm, learning was accomplished in an unsupervised fashion, i.e., spatial information was acquired incidentally, without explicit allocation of attentional resources toward learning object-location associations. During the learning session in *day 1* of the experiment, subjects did not know that they would later be asked to recall object-location associations. In contrast, in paired-associate learning, learning is accomplished explicitly with foreseeable consequences for successful or unsuccessful retrieval (Schacter et al. 1998).

We observed total spectral power between 3 and 6 Hz in electrodes over midline frontal regions to be modulated by misplacing objects. Specifically, misplaced objects elicited increased synchronizations of high delta/low theta during 145–475 ms after stimulus onset. This finding was based on an across-subject statistical clustering of individual ICs based on specific component properties such as ERPs, ERSPs, TICs, and estimated dipole location in a BEM. Unlike other IC clusters that localized in different brain regions and showed effects in other frequency ranges, this cluster alone exhibited significant power in the theta range and dissociated correctly placed and misplaced objects. Equivalent dipole models of the ICs as well as distributed source estimation in sLORETA were compatible with areas in the anterior cingulate sulcus and/or the overlying prefrontal cortex, consistent with other studies utilizing comparable component-based clustering procedures (Gramann et al. 2010; Michels et al. 2008; Onton et al. 2005; Plank et al. 2010). This localization is distinct from the evoked response that occurred at a similar time (N535) but was located anteriorly in the bilateral fusiform gyrus and/or lingual gyrus. With other experimental paradigms, theta has been found to emanate not only from medial and lateral prefrontal but also from central, parietal, and medial temporal cortices (Cornwell et al. 2008; Hsieh and Ranganath 2014; Kahana et al. 1999; Raghavachari et al. 2001; Wang et al. 2005). EEG studies in humans have estimated (6–8 Hz) theta during memory-related tasks to originate in or near medial prefrontal cortex, particularly in anterior cingulate cortex (Asada et al. 1999; Gavins and Smith 2000; Meltzer et al. 2008), but it has been shown that during spatial processing lower theta and delta effects (3–6 Hz) are also observable (Watrous et al. 2011). This was apparent in Snider et al. (2013b), where spatial autocorrelation during walking was found to be strongest in the 2–8 Hz (delta/theta) range. Moreover, these findings are consistent with human intracranial studies of the hippocampus, where low-frequency (2–6 Hz) modulations also have been observed (Arnolds et al. 1979; Brazier 1968; Jacobs et al. 2007; Mormann et al. 2008; Rutishauser et al. 2010; Watrous et al. 2011).

Theta in general has been associated with a wide range of cognitive processes and has been found to correlate with memory retrieval, locomotion, and goal-directed behavior (Caplan et al. 2003; Cornwell et al. 2008; Ekstrom et al. 2003; Kahana et al. 2001; Klimesch 1999). Particularly in situations with high cognitive demands and working memory load increased levels of theta power have been identified (Bastiaansen et al. 2002; Klimesch 1999; Meltzer et al. 2008; Onton et al. 2005; Raghavachari et al. 2001; Sauseng et al. 2002; Schack et al. 2005; Tesche and Karhu 2000), which can be interpreted in concordance with findings on event-related synchronization (ERS) effects in the theta range (Michels et al. 2008; Pfurtscheller and Aranibar 1977, 1979). In the present study we found increased frontal midline theta power for incorrectly placed objects. This would be consistent with misplaced objects eliciting increased cognitive demands for memory recall and detecting deviations between currently encountered and internally represented environmental configurations in frontal regions. This finding extends the existing literature on increased frontal midline theta power associated with increased cognitive demands.

Furthermore, the crucial role of theta for spatial learning and memory in humans is well documented (Bischof and Boulander 2003; Caplan et al. 2003; de Araujo et al. 2002; Jones and Wilson 2005; Kahana et al. 1999). In an intracranial study of Kahana and colleagues (1999), increased theta was detected during recall compared with learning, as well as when participants were confronted with more complex environments. Bischof and Boulander (2003) also found increased theta in the time range when subjects encountered central navigation points along a previously learned route that were crucial for successful task accomplishment. Theta oscillations are seemingly modulated by the amount of self-motion cues available to the subject. While attenuated in stationary setups or under bodily restraints, theta power increases as subjects are being passively moved along trajectories or experience visual motion cues in the absence of proprioceptive or vestibular signals, with the
strongest theta response being generated as subjects are allowed to move freely, generating consistent visual, vestibular, and proprioceptive input (Russell et al. 2006; Terrazas et al. 2005; Vanderwolf 1975).

We believe that our experimental paradigm allowed subjects to fully utilize multisensory information, allowing for naturalistic integration and consolidation of the acquired knowledge of the environment into a robust internal representation based on unsupervised spatial learning. Since subjects had to recall spatial information on object location traversing the same environment again on day 2, theta activity might establish a dynamic framework for distributed cognitive processing between spatially segregated cortical areas (Mizuhara et al. 2004), allowing for action-oriented retrieval of spatial information from the internal representation. In nonspatial experiments, theta localized to the anterior cingulate has been associated with cognitive conflict (Kovacevic et al. 2012). Our observation of an increase in theta to misplaced objects may thus reflect sustained retrieval and reprocessing of spatial information resulting from conflicts between incoming sensory input and representational systems. The modulations observable in frontal midline theta are certainly distinct from the N400 effects, since they differ in latency and localization of the source estimates. However, future research will have to address how representational recall as represented by frontal midline theta relates to hippocampal-parietal processing of object-location information as reflected by the N400 effects. Given the large literature on sex differences in spatial orientation (Coluccia and Louse 2004; Lawton and Morrin 1999), future studies should also examine potential sex differences using naturalistic data acquisition environments such as that in the present experiment.

Limitations

The analysis pipeline used here was a major component enabling the collection of usable EEG during ambulatory movement, but it does have some limitations. One potential limitation is the assumption that an IC can be localized to a dipolar source within the brain. Indeed, a split occipital IC was observed in the data presented here that could only be fit with two dipoles (and it was). We then removed any other ICs whose best-fit dipole was outside the brain. This processing pipeline followed procedures that have been promulgated by other authors (Delorme et al. 2012; Gwin et al. 2010, 2011; Makeig et al. 2002; Sipp et al. 2013), but this could lead to loss of brain signal (Castellanos and Makarov 2006). While ICA decomposition has been shown to be a highly useful tool for EEG data analysis and noise attenuation, assumptions of ICA and dipole source modeling might not map 1:1 into physiological processes, as discussed by Delorme and colleagues (2012). For example, the assumption that scalp potentials are generated by synchronization across localized cortical patches might represent a simplification of spatiotemporal nonstationary processes, phase-coherent activations across cortical regions or hemispheres, which cannot be handled well with dipole-based reconstruction methods. More accurate head models, for example, based on segmentation of actual structural MRIs of the participants, or frequency-domain complex ICA decomposition techniques might allow for even more individualized IC selection procedures with stronger physiological validity.

We cannot rule out an evoked component to the spectral differences we observed for IC cluster 1 (theta cluster). However, we note three findings in favor of it being induced. First, the theta cluster did not match any evoked component observed in the electrode clusters. Second, frontal midline theta has a long-standing role in the literature (as described above) that is consistent with the role in memory proposed here. Finally, the theta cluster localization is far removed from the localization of the evoked activity at the same time.

Conclusions

Taken together, the present results suggest that the increased negativity for object-location mismatches in the time range between 400 and 700 ms peaking over parietal regions that localizes in temporal regions resembles an N400 in semantic memory paradigms, here reflecting the mismatch between sensory input and representational expectation. Second, the increase in frontal midline theta power in case of violations of object-location expectancies might be related to recalling the actual object as learned on the previous day. Both effects (theta power and N400 amplitude) are most likely generated by an extensive network comprising posterior parietal and anterior cingulate cortex, potentially mediated by hippocampus.

ACKNOWLEDGMENTS

We thank Jamie Lukos and Jason Trees for help with the data collection. Present address of Markus Plank: Brain Products GmbH, Zeppelininstr. 7, 82205 Gilching, Germany.

GRANTS

Supported in part by Office of Naval Research MURI Award No. N00014-10-1-0072 and National Science Foundation Grants SMA-1041755 and ENG-1137279 (EFRI M3C).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: M.P., J.S., E.K., E.H., and H.P. conception and design of research; M.P. and J.S. performed experiments; M.P., J.S., E.K., E.H., and H.P. interpreted results of experiments; M.P. and J.S. prepared figures; M.P., J.S., E.H., and H.P. drafted manuscript; M.P., J.S., E.K., E.H., and H.P. edited and revised manuscript; M.P., J.S., E.K., E.H., and H.P. approved final version of manuscript.

REFERENCES


J Neurophysiol • doi:10.1152/jn.00114.2014 • www.jn.org