

# Tracking neural correlates of successful learning over repeated sequence observations



Natalie A. Steinemann<sup>a</sup>, Clara Moisello<sup>b</sup>, M. Felice Ghilardi<sup>b</sup>, Simon P. Kelly<sup>a,c,\*</sup>

<sup>a</sup> Biomedical Engineering Department, The City College of The City University of New York, 160 Convent Ave, New York, NY 10031, USA

<sup>b</sup> Sophie Davis School of Biomedical Education, The City College of The City University of New York, 160 Convent Ave, New York, NY 10031, USA

<sup>c</sup> School of Electrical and Electronic Engineering, University College Dublin, Belfield, Dublin 4, Ireland

## ARTICLE INFO

### Article history:

Received 24 February 2016

Revised 20 April 2016

Accepted 2 May 2016

Available online 5 May 2016

### Keywords:

Event-related potential

Expectancy

Memory

Sequence learning

Surprise

## ABSTRACT

The neural correlates of memory formation in humans have long been investigated by exposing subjects to diverse material and comparing responses to items later remembered to those forgotten. Tasks requiring memorization of sensory *sequences* afford unique possibilities for linking neural memorization processes to behavior, because, rather than comparing across different items of varying content, each individual item can be examined across the successive learning states of being initially unknown, newly learned, and eventually, fully known. Sequence learning paradigms have not yet been exploited in this way, however. Here, we analyze the event-related potentials of subjects attempting to memorize sequences of visual locations over several blocks of repeated observation, with respect to pre- and post-block recall tests. Over centro-parietal regions, we observed a rapid P300 component superimposed on a broader positivity, which exhibited distinct modulations across learning states that were replicated in two separate experiments. Consistent with its well-known encoding of surprise, the P300 deflection monotonically decreased over blocks as locations became better learned and hence more expected. In contrast, the broader positivity was especially elevated at the point when a given item was newly learned, i.e., started being successfully recalled. These results implicate the Broad Positivity in endogenously-driven, intentional memory formation, whereas the P300, in processing the current stimulus to the degree that it was previously uncertain, indexes the cumulative knowledge thereby gained. The decreasing surprise/P300 effect significantly predicted learning success both across blocks and across subjects. This presents a new, neural-based means to evaluate learning capabilities independent of verbal reports, which could have considerable value in distinguishing genuine learning disabilities from difficulties to communicate the outcomes of learning, or perceptual impairments, in a range of clinical brain disorders.

© 2016 Elsevier Inc. All rights reserved.

## Introduction

Despite many significant advances, our understanding of the neural mechanisms underlying the commitment of information to memory remains incomplete. While the finer details of the plastic changes underlying memory trace formation are best examined on synaptic and neuronal levels (Kandel and Schwartz, 1982; Martin et al., 2000), lower-resolution human neuroimaging and neurophysiology can offer valuable, complementary insights owing to the feasibility of linking neural signals to complex behaviors that can readily be manipulated and measured. In human studies of memory formation, subsequent-recall reports have been the major behavioral currency, and many fundamental insights have been gained by comparing neural responses to

items that were later remembered versus forgotten (Fernandez et al., 1999; Gonsalves and Paller, 2000; Karis et al., 1984; Paller et al., 1987b; Wagner et al., 1999).

In the present study we examine from a new perspective the particular case of memorization of sensory sequences over repeated observations, which naturally provides experimental traction in the study of neural correlates of memory formation in several important ways. First, rather than making comparisons among items that may vary widely in memorability, semantic content, and/or sensory characteristics, comparisons within a sequence can be made among highly simplified tokens which minimize these and other factors known to influence learning success, such as item-to-item interactions (Gobet et al., 2001). Second, states of learning can in fact be compared *within-item* because the exact same stimuli are presented repeatedly over the course of learning, allowing endogenous variance in the efficacy of memory formation across and within subjects to be clearly distinguished from inter-stimulus variability. Third, learning state transitions can be traced gradually, from an individual element being completely unknown, while it is being learned, to being fully committed to memory. This

\* Corresponding author at: School of Electrical and Electronic Engineering, University College Dublin, Belfield, Dublin 4, Ireland.

E-mail addresses: [nsteinemann@ccny.cuny.edu](mailto:nsteinemann@ccny.cuny.edu) (N.A. Steinemann), [cmoisello@gmail.com](mailto:cmoisello@gmail.com) (C. Moisello), [lice.mg79@gmail.com](mailto:lice.mg79@gmail.com) (M.F. Ghilardi), [simon.kelly@ucd.ie](mailto:simon.kelly@ucd.ie) (S.P. Kelly).

gradual aspect is particularly important because it provides the ability to examine the potentially incremental nature of memory formation (Schlaghecken et al., 2000), which cannot be readily done using dichotomous subsequent recall outcomes. Fourth, sequence learning over repeated observations systematically influences stimulus expectancy, thus allowing us to relate these gradual changes in sequence knowledge to stimulus-bound surprise and its established neural correlates (Duncan-Johnson and Donchin, 1977; Kolossa et al., 2012; Mars et al., 2008; Sutton et al., 1965). To realize these advantages in the present study, we examined event-related potentials (ERP) of subjects attempting to memorize sequences of spatial locations – possibly the most elementary type of sensory information – over several blocks of repeated observation, with recall tests before and after each block.

Previous research has successfully employed sequences to shed light on issues related to plasticity in primary sensory areas through learning (e.g. Gavornik and Bear, 2014), differences between explicit and implicit learning (e.g. Aizenstein et al., 2004; Baldwin and Kutas, 1997; Rüsseler et al., 2003), and primacy and recency effects (e.g. Rushby et al., 2002; Wiswede et al., 2007). Some studies have examined spectral changes exhibited after versus before learning (Moisello et al., 2013), or more finely over the course of repeated presentations (Madhavan et al., 2015; Zhuang et al., 1997), in the latter cases linking the changes to motor execution (Zhuang et al., 1997), or memory retrieval (Madhavan et al., 2015) rather than memory formation. To our knowledge, however, no study has yet fully exploited the possibilities outlined above, and tracked neural correlates of memory formation on an element-by-element basis as they develop through discrete learning state transitions.

Neural signal changes that could systematically relate to element-wise learning as well as predict overall learning success can be hypothesized on the basis of two different strands of literature. First, seminal ERP studies of memory found broad, late centro-parietal positivities (300 ms after stimulus onset and onwards) to be increased for items subsequently remembered relative to those forgotten (Karis et al., 1984; Neville et al., 1986; Paller et al., 1987b), which has since been widely replicated in tasks ranging from implicit sequence learning (e.g. Eimer et al., 1996; Schlaghecken et al., 2000) to imagining visual scenes (Gonsalves and Paller, 2000), with greater effects being indicative of stronger memories (Müntz et al., 1988; Paller et al., 1988). The same approach has been applied in neuroimaging studies, indicating the involvement of several areas, including left inferior frontal gyrus (De Chastelaine and Rugg, 2014; Kim, 2011; Wong et al., 2013), left fusiform gyrus (De Chastelaine and Rugg, 2014; Kim, 2011), and bilateral hippocampus (Brewer et al., 1998; Kim, 2011; Schendan et al., 2003), in the encoding of memories. There have, however, been indications in functional magnetic resonance imaging (Otten et al., 2001) and electroencephalography (EEG) studies (Paller et al., 1987a; Sanquist et al., 1980; Weyerts et al., 1997) that such neural activation patterns may not be linked primarily to the memorization itself, but more directly to mediating factors such as associative or semantic processing depth, and stimulus complexity (Van Petten and Senkfor, 1996). It is therefore uncertain whether the same subsequent memory effects can be found for simple stimuli inspiring as little associations as the spatial position elements used in the present sequence learning paradigm. Neural correlates of the most basic mechanisms of active memory formation should, in theory, be observed for the learning of information regardless of its semantic content; further, to the extent that their endogenous variation bears on learning efficacy, they should be decreased when learning efforts fail; and finally, assuming they are actively employed specifically for memorization, they should be absent for already well-learned items. A sequence learning paradigm of the kind employed in the present study is uniquely amenable to applying such strong identifying criteria.

A second relevant line of literature concerns the P300 component, a parietal positivity evoked by task relevant stimuli with a peak latency that varies widely (300–1500 ms) as a function of the temporal

requirements of stimulus processing (e.g. McCarthy and Donchin, 1988; Twomey et al., 2015). Most important for the current purposes, its amplitude has been found to decrease as a function of subjective stimulus expectancy, or equivalently, increase with stimulus-bound “surprise,” across a diverse range of paradigms (Donchin, 1981; Duncan-Johnson and Donchin, 1977; Fu et al., 2013; Horst et al., 1980; Kolossa et al., 2012; Mars et al., 2008; Rüsseler et al., 2003; Schlaghecken et al., 2000; Sutton et al., 1965). In the context of sequence learning, sequence knowledge can be equated to correct stimulus expectancy. Hence, we can hypothesize that P300 amplitude will decrease over the course of learning as sequence knowledge strengthens. More specifically, assuming that increments of sequence knowledge are always gained in varying amounts from one observation to the next, and seldom *lost*, we predict that this P300 decrease will be strictly monotonic. This is distinct from the prediction for signals of active memory trace formation which, as we argued above, will be especially elevated at the point where an element begins to be correctly recalled, stronger than both the preceding and the following sequence observations. Both the abovementioned subsequent memory effects and the P300 tend to be focused at similar centro-parietal topographic sites, which in general complicates their separate measurement. However, the use of the simplest sequence tokens in the present study ensured that stimulus processing was both fast and minimally variable. As a result, the evoked P300 signature was highly transient in the average ERP and could be distinguished to a reasonable degree from less time-restricted processes associated with memory-formation based on temporal scale. We additionally explored the potential for using these neurophysiological components to predict learning success across individuals in a group, as well as across sequence observations within a given individual. Such prediction from electroencephalographic measures alone, independent of behavioral reports, would pave the way towards the development of new diagnostic tools for learning disabilities that involve significant communication difficulties. To facilitate these various analyses, as well as to replicate the principal effects to demonstrate reliability, we analyzed the data of two separate experiments, selectively leveraging variance across and within subjects, respectively.

## Materials and methods

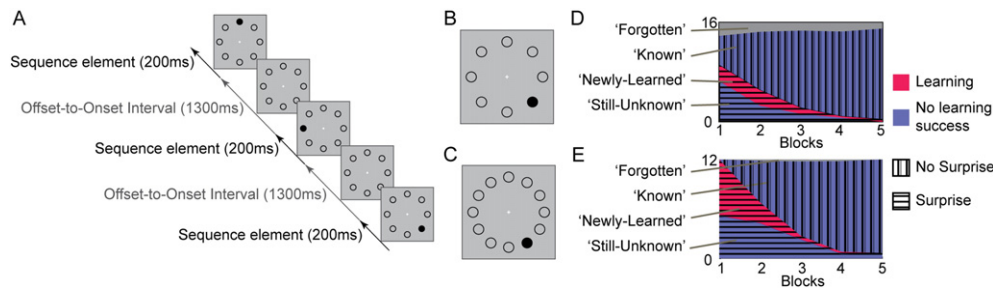
### Participants

Twenty-one right-handed participants volunteered for **Experiment I**. One participant was excluded from the analysis because the sequence was memorized perfectly by the end of the first block, and one further participant was excluded due to an excessive number of artifacts (>95% of trials). This yielded a final sample of 19 participants for **Experiment I** (mean age  $\pm$  standard deviation =  $24.2 \pm 4.8$  years, 13 male). A different set of six subjects volunteered for **Experiment II** (mean age  $\pm$  standard deviation =  $22.5 \pm 3.67$  years, 2 male), with no overlap between the experiments.

All subjects reported normal or corrected-to-normal vision and no current psychiatric diagnosis or history of head injury. All participants gave written informed consent. Procedures were approved by the Institutional Review Boards of the City College of New York and were in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

### Spatial sequence learning task

In all conditions of both experiments, participants were asked to view a series of identical stimuli presented, one at a time, at a sequence of radial locations. Visual stimuli consisted of filled black circles with a diameter of 1 cm. The set of possible spatial locations lay equidistant around a ring of fixed eccentricity (2 cm; Fig. 1A), and were continuously marked by static circular outlines. Stimuli were presented for



**Fig. 1.** Spatial sequence learning task. (A–C) Participants were asked to memorize the order of appearance of spatial sequences of elements presented at eight (B, Experiment I) or twelve (C, Experiment II) equidistant locations. (D–E) Based on verbal reports before and after each sequence block, trials were categorized according to whether a given element was ‘still-unknown’, ‘newly-learned’, or ‘known’. Panels D and E display the development of the average number of sequence elements in each category over sequence repetition number for Experiment I and II, respectively. For Experiment I (D) averages are taken across subjects, while the trial numbers are further averaged across sequences for Experiment II (E). The shading indicates learning-states in which subjects were in the process of successfully committing a new element to memory (red) or not (purple), while the pattern orientation indicates learning states in which subjects remain relatively surprised upon an element’s presentation (horizontal) or have relatively accurate expectancy (vertical grating). Sequence elements reported correctly before but incorrectly after a block (‘forgotten’; shown in gray) were excluded from all analyses.

200 ms with an offset-to-onset interval of 1300 ms. Throughout the task, subjects were asked to hold eye fixation on a central fixation cross.

### Experiment I

In Experiment I, spatial elements were presented at eight equidistant locations (Fig. 1B). Experimental sessions began with three ‘random’ blocks, each containing a pseudo-random sequence of 48 stimuli presented among the eight locations, which subjects were asked to merely observe while avoiding eye movements. Subjects then performed fifteen consecutive ‘sequence learning’ blocks, in each of which a fixed sequence of 16 stimuli, with each of the eight locations coming up twice, was presented three times, resulting in 48 stimuli in total per block. Subjects were asked to memorize this 16-element sequence. As the original purpose of this dataset was partially to examine motor learning, two out of every three of these blocks required subjects to perform temporally unconstrained reaching movements towards the targets during viewing (Ghilardi et al., 2003a, 2003b, 2007, 2008). Since the present study focused on the learning of information during passive observation, and to avoid motor-related brain activity and potential movement artifacts, we only analyzed the third, sixth, ninth, twelfth, and fifteenth blocks, in which subjects only viewed the stimuli and did not make any movements in response to them. Before and after each ‘sequence learning’ block, subjects were asked to recall the 16-element sequence to the best of their memory, by verbal reports to the experimenter with reference to static displays of the 8 locations in different colors on the screen. Two different color arrays were used in alternation to ensure that subjects learned stimulus locations without reference or association to a color. Because the visual-only blocks were interleaved with movement blocks, the results of the pre-block recall tests for any given visual block were not necessarily the same as those of the post-block reports for the previous visual block. No feedback on correctness of recall was given to the subject after verbal recall, so that learning took place only during sequence presentation.

### Experiment II

While the data of Experiment I lend themselves to analysis of across-subject variance in learning, trial counts for each learning condition were relatively low within each subject, and the alternation of visual observation blocks with motor-learning blocks precluded analysis of seamless learning increments. Therefore, we chose to analyze the data of a second experiment (Experiment II), in which there was no motor component and multiple sequences learned by each of a fewer number of individuals, in order to validate the results of Experiment I and to leverage within-subject variability. To that end, six additional participants, who did not participate in Experiment I, were each asked to learn eight different spatial sequences, one sequence at a time. In this experiment, each sequence consisted of twelve elements presented at

twelve equidistant locations, one after another (Fig. 1C). One participant ended the session after six sequences due to fatigue, but the data for those sequences was nevertheless sufficient to be included in the analyses. Within each ‘sequence learning’ block, the same sequence was repeated three times, giving a total of 36 stimuli per block. As in Experiment I, before and after each block, subjects gave verbal reports of the spatial sequence to a static display of the twelve locations in different colors on the screen, with no feedback given. Each sequence was presented for eight consecutive blocks or until a subject recalled it with perfect accuracy on three blocks, whichever occurred first. This resulted in a total of between 36 and 42 sequence learning blocks per subject. In Experiment II, the first ‘learning’ block of each sequence was preceded by one ‘random’ block of 72 spatial elements, which participants were asked to observe without any learning requirement (Fig. 3G + H).

### Behavioral data analysis

Based on verbal reports before and after each sequence learning block, sequence elements were categorized as (1) ‘still-unknown’ elements, whose stimulus location was reported incorrectly both before and after a block, (2) ‘newly-learned’ elements, for which subjects made incorrect reports before a block, but a correct report afterwards, (3) ‘known’ elements, whose location was reported correctly before and after a block, and (4) ‘forgotten’ elements, which were reported correctly before the sequence presentation, but incorrectly afterwards (Figs. 1D–E). This fourth category included an average of 0.7 items per sequence repetition for Experiment I and 0.1 items per sequence repetition for Experiment II. As it is impossible to determine which of these sequence elements were truly forgotten and which ones were merely guessed correctly before the sequence repetition, these trials were left out of all further analyses. In this trial categorization, stimulus transitions were taken into account, so that if two consecutive stimulus locations were reported in the correct sequence, the second element was categorized as reported correctly even if not reported at the exactly correct sequence position. Thus, if the sequence had been 1–2–3–4–5–... and a subject had reported 1–3–4–5–2–..., the first, fourth, and fifth item presented were categorized as reported correctly. This was based on the assumption that subjects learn the sequence by forming associations between consecutive sequence elements, rather than by memorizing the absolute sequence position. Although ongoing computational research on the mechanisms of sequence learning indicates that this strategy is by no means the only one (Sun, 2003), in our two experiments subjects tended to learn the first elements initially and proceed through the sequence, so that alternative schemes for judging correctness of recall would produce very few differences in categorization. Any deviation from this strategy would have the effect of decreasing



the power of our analysis, thus tending to heighten type II, but not type I, error.

For analyses linking neural signals to behavior across blocks (Fig. 5), we derived block-wise behavioral indices that, like the event-related potential measures, take into account only artifact-free trials. Having categorized the sequence elements as ‘still-unknown’, ‘newly-learned’, and ‘known’ as above, we defined the “Knowledge index” as the proportion of trials in a given block for which the subject would have had accurate prior knowledge of the upcoming stimulus. We estimated the number of such trials by combining the number of ‘known’ trials  $n_K(S, b)$  with ‘newly-learned’ trials  $n_{NL}(S, b)$ , with the latter weighted by 0.5 following the assumption that elements that transitioned from unknown to known during a block could on average be classed as ‘known’ 50% of the time:

$$\text{KnowledgeIndex}(S, b) = \frac{n_K(S, b) + 0.5 \cdot n_{NL}(S, b)}{n_T(S, b)} \quad (1)$$

where  $n_T(S, b)$  represents the total number of artifact-free trials in sequence block  $b$  of subject  $S$ . The “Learning index” was quantified as the number of artifact-free trials of newly-learned elements during a sequence block divided by the total number of artifact-free trials during the same block.

$$\text{LearningIndex}(S, b) = \frac{n_{NL}(S, b)}{n_T(S, b)} \quad (2)$$

#### Electrophysiological recordings and data analyses

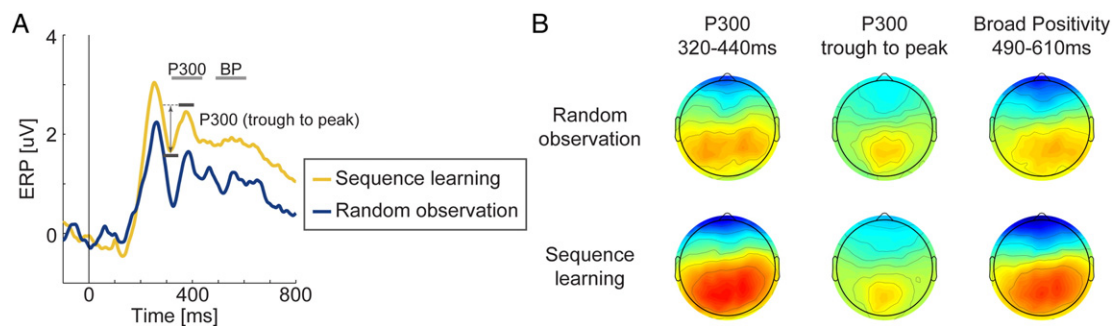
Continuous EEG was acquired using a 256-channel Hydrocel net system (Electrical Geodesics Inc.), digitized at 500 Hz. If any channel's variance in amplitude differed from the mean by more than four standard deviations of the distribution across electrodes, it was considered excessively noisy and was interpolated using spherical splines. Data were re-referenced off-line to the average of all 256 EEG channels. EEG traces of 73 channels located on the cheeks and neck were excluded from topographic plots. We verified that average-referencing the data based on all but the 73 cheek/neck channels yielded the same qualitative results, although it produced a slightly lower signal-to-noise ratio (average centro-parietal P300 amplitude divided by variance in 100-ms prestimulus baseline). Within-block drift was removed by linearly detrending the data before applying a band-pass filter between 0.05 and 45 Hz to all channels. All channels were subject to an artifact rejection criterion of 90  $\mu\text{V}$  to exclude trials with ocular motion, blinks or other noise transients. To minimize trial rejection from frequent blinks, data were segmented into relatively short epochs extending from  $-100$  ms to 800 ms relative to stimulus onset, spanning just half of the inter-stimulus interval. Changes in spectral power during learning

and signatures of prior learning in resting-state data recorded after completion of the task in Experiment I have been reported in a separate article (Moisello et al., 2013). The present study focused on event-related potentials elicited by the sequence elements.

In order to first establish measurement time-windows and sites, we sought to identify the processes of interest in the grand-average ERP waveforms collapsed across learning states in the sequence-learning blocks, compared against random blocks in which the same stimuli were viewed but no learning was required (Fig. 2). Since we used the simplest possible feature in our task, we were able to identify a clear, sharp, centro-parietally focused P300 peak at a relatively early latency of approximately 380 ms (see Fig. 2A) in both the sequence-learning and the random blocks. We therefore took the integrated amplitude in a 120 ms time window centered on this peak (320–440 ms) at a cluster of four centro-parietal electrodes centered on standard site CPz for statistical testing of the P300. An earlier positive deflection at approximately 250 ms is evident preceding the P300 but had a distinct fronto-central topographic focus, and having no hypothesis regarding any specific role in learning, was not tested for effects.

The P300 appeared as a rapid deflection superimposed atop a broader centro-parietal positivity, which distinguishes sequence-learning from random blocks over a much broader interval. Given this superimposition, absolute measurements of P300 amplitude relative to trial baseline would likely reflect variation in the broader process to a considerable degree. We thus additionally tested all P300 effects using a trough-to-peak measure, i.e., the difference in amplitude between the P300 peak (360–400 ms) and the preceding trough (302–342 ms) to minimize the influence of the temporally broader process on P300 measurements. To the extent that the broader component remains relatively constant across these two closely neighboring ‘trough’ and ‘peak’ intervals, its contribution to the trough-to-peak P300 measurement would effectively be cancelled-out. To measure the broader positivity itself, we used a second 120-ms timeframe centered on the next positive peak following the P300 in the grand average waveforms (490–610 ms; Figs. 2A–B). This interval lay sufficiently beyond the rapid P300 component but well within the broad time frame over which centro-parietal activity differed during active sequence learning compared to passively viewed random blocks. Statistical test results were not influenced by the exact choice of measurement time window.

We began our analyses by testing for differences in average ERP waveforms across the three learning states of ‘still-unknown’, ‘newly-learned’, and ‘known’ for each of the two signal timeframes (P300 and Broad Positivity) in both experiments. We predicted that the P300, in line with its suggested role in the encoding of surprise, would show greater amplitudes for ‘unknown’ and ‘newly-learned’ sequence elements compared to ‘known’ elements, which would likely be expected. For the broader positivity, we hypothesized that, if it is involved in active memorization, it would decrease when learning efforts fail (‘still-unknown’), and be absent when memorization efforts cease once a



**Fig. 2.** Spatio-temporal characterization of centro-parietal signal components. (A) Grand average centro-parietal evoked potentials during sequence learning blocks (yellow) and while subjects were merely observing random sequences (blue). Gray bars indicate time windows during which measures for the P300, and Broad Positivity (BP) were taken. (B) Topographies in both the early and late timeframe, as well as the trough-to-peak P300 deflection, exhibited a centro-parietal focus.

stimulus location has already been committed to memory ('known'). One-way ANOVAs were computed for both experiments to test differences between learning states, followed by pairwise comparisons. For [Experiment I](#), four subjects were excluded from this learning state specific analysis, due to low trial counts. The remaining 15 subjects presented with an average of  $132.93 \pm 41.46$  'known',  $16.93 \pm 8.51$  'newly-learned', and  $31.27 \pm 22.75$  'still-unknown' trials per subject. Greater trial numbers in [Experiment II](#) ( $411.67 \pm 252.97$ ,  $182 \pm 70.89$ , and  $160.67 \pm 93.21$  for 'known', 'newly-learned', and 'still-unknown' elements, respectively) enabled validation of the results of [Experiment I](#).

We next verified that changes in both ERP components over blocks of learning occur only while subjects are engaged in memorizing the sequence, but not during passive observation ("random" blocks). This is particularly important in the case of the P300 since it is hypothesized to monotonically decrease over learning, which could in theory occur simply as a function of time on task. We thus carried out a two-way ANOVA for each experiment with factors of repetition number (1 through 3 for [Experiment I](#), and 1 through 4 for [Experiment II](#)) and memorization effort ('sequence learning' vs 'random observation'). Our hypothesized link between P300 amplitude and sequence knowledge would be supported by a significant interaction between sequence repetition and memorization effort in the two-way ANOVA, driven by a significant decrease of P300 amplitude across sequence repetitions for 'sequence' but not 'random' blocks. Since [Experiment II](#) involved learning multiple different sequences, we conducted an additional control analysis to verify that while P300 amplitude decreases across repetitions of a given sequence, it does not change across sequences. We tested the latter using linear contrasts with the single factor of sequence, and carried out these tests for the later time frame (490–610 ms) as well as the P300. In the analyses of effects both across learning-state and over blocks, ANOVAs were corrected for non-sphericity using Greenhouse–Geisser corrections. In follow-up pairwise comparisons carried out where ANOVAs were significant, we did not correct for multiple comparisons, on the basis that within this single study we test for each of these effects in both of two independent datasets ([Experiment I](#) and [II](#)), thus greatly reducing the risk of type I error.

As participants learned multiple sequences in [Experiment II](#), within-subject signal estimates were sufficiently robust to allow examination of the gradual learning process with yet finer temporal resolution than the above comparison of learning categories. We thus conducted a more detailed temporal analysis aligned to the point of first successful recall of each element (newly-learned category), where we more finely categorized 'still-unknown', and 'known' trials according to how many blocks in the past or into the future they occurred relative to this point. We considered only up to two blocks prior ('-2') to the newly-learned point ('0'), and up to 4 blocks following ('+4') that point because there were insufficient numbers of trials outside of this timeframe. For this analysis we predicted that P300 amplitude would decrease monotonically, whereas the Broad Positivity, if a marker of active memorization, should achieve its maximum at the point where a given element is successfully committed to memory ('block 0'). To conduct a well-powered test of this in [Experiment II](#), which had many sequences for relatively few subjects, we fit linear mixed-effects models to the averaged amplitudes within each sequence and subject for a given component, with 'distance from first accurate recall' as a fixed effect and 'subject' as a random effect. Importantly, we defined the 'distance from the first accurate recall' in two different ways — the signed distance from first accurate recall ('signed distance'), and the absolute distance from the first accurate recall ('absolute distance'). A monotonic decrease in P300 amplitude with increasing sequence knowledge should be reflected in a significant negative fixed-effect of the 'signed distance' to the first accurate recall, while a marker of learning that peaks during the trial of first successful recall should show a significant negative fixed-effect of the 'absolute distance' to the first accurate recall. We statistically established which model (signed or absolute distance)

better fit the P300 and Broad Positivity amplitude data, respectively, using a Likelihood ratio test (function 'compare' applied to the output of 'fitlme' in Matlab). Since this analysis directly compares the learning-aligned time course of the P300 and Broad Positivity, we used the trough-to-peak (360–400 ms relative to 302–342 ms) P300 measure to minimize the conflation of one with the other.

We next investigated whether these links between neural signals and behavior could be exploited to predict sequence knowledge within and across participants. These analyses focus on the data of [Experiment I](#) which had greater across-subject variance. We first conducted three across-block correlation analyses in which we tested whether existing knowledge can be predicted by average P300 amplitudes, and whether *change* in knowledge, i.e. learning rate, can be predicted by change in P300 within a block as well as the average Broad Positivity amplitude, within subjects.

We first determined the correlation between the Knowledge Index and the average amplitude in the time range of the P300 (320–440 ms) at the centro-parietal electrode cluster across the five sequence learning blocks of each subject. A negative correlation specific to centro-parietal electrodes would indicate that P300 can be employed as a predictor of learning success across blocks. We also estimated the change of P300 amplitude within blocks, by computing the slope of the decreasing trend of P300 amplitude across trials within every block for each subject. To do this, we averaged the same P300 amplitude measure across the 16 trials in each of the three sequence presentations (omitting the artifact trials) within each block and fit a line to the three values we obtained. We reasoned that, if P300 is a predictor of sequence knowledge, this measure should decrease more steeply during blocks in which more effective learning takes place, i.e. there is a relatively large increase in the number of 'fully' learned, and thus accurately expected, stimuli. Therefore, there should be a negative correlation between the slope of the line fit to centro-parietal electrodes and the number of newly learned sequence elements within a block. To examine the ability to predict learning rate from the Broad Positivity across blocks, we computed the correlation between the Learning Index and the average amplitude in the 490–610 ms timeframe for the centro-parietal electrodes across the five sequence learning blocks for each subject. In this case we expected to observe greater amplitudes at centro-parietal sites during blocks in which more new elements were successfully committed to memory. To verify the topographic specificity of these effects by visual inspection, we simply repeated these measures for all electrodes and plotted their scalp distributions.

For all three correlations, statistical significance was assessed by comparing the group average of the actual correlation factor against 1000 such correlation factors obtained from randomly permuted assignments of behavioral indicators to trial amplitudes. This computation of correlations within each subject was more appropriate than pooling all blocks regardless of subject, because inter-individual differences in ERP amplitude were much greater than within-subject changes.

We next turned to predicting learning success across subjects. To quantify a subject's overall learning speed we fit exponential functions to the individual learning curves by exhaustive search. Sequence knowledge was constrained to be zero before the first sequence presentation (no sequence knowledge), and to asymptotically approach 1 (full sequence knowledge). The time point when full sequence knowledge was to be attained was, however, not constrained to lie within the number of repetitions performed during the experiment:

$$\text{Sequenceknowledge} = 1 - e^{-\frac{\text{sequencerepetition}}{\tau}} \quad (3)$$

In this equation, the time constant  $\tau$  quantifies after how many blocks a subject has learned  $1 - 1/e$  or 63% of all sequence elements. Hence, lower time constants identify faster learners. As most learning took place within the first two sequence learning blocks, we then correlated this time constant with the decrease in P300 amplitude from the first to

the second sequence repetition block. A negative correlation in this measure would be in line with our predicted relationship between P300 amplitude and sequence knowledge.

## Results

### *Main effects of stimulus expectation and learning*

In two separate experiments, participants were asked to learn the order of appearance of spatial sequence elements (Figs. 1 A–C). Verbal reports before and after each sequence presentation block were employed to identify trials associated with the learning states of (1) ‘still-unknown’ (location was reported incorrectly both before and after the block), (2) ‘newly-learned’ (incorrect reports before, but a correct report after a block), and (3) ‘known’, (location reported correctly before and after a block). The logic behind our analyses can be understood by considering the distribution of these learning states across sequence repetitions (Figs. 1D–E). Both still-unknown and newly-learned trials are classed as ‘surprising’ (horizontal grating), since the elements were reported incorrectly before the sequence presentation and therefore could not have been fully expected. Memorization processes, on the other hand, can be assumed to have been most efficiently deployed in the block where the participant ultimately succeeded in committing a given sequence element to memory (‘newly-learned’, red shading), and can be assumed to be weaker both when learning was unsuccessful or had already been completed (purple shading).

As expected, recall performance derived from verbal reports increased significantly as a function of sequence repetition, both in *Experiment I* ( $F(4, 72) = 56.8, p < 1 \times 10^{-10}, \eta^2 = 0.759$ ) and *Experiment II* ( $F(7, 35) = 29.4, p < 5 \times 10^{-9}, \eta^2 = 0.855$ ). Learning rates, measured as the proportion of newly-learned elements per block, decreased over sequence repetition (*Experiment I*:  $F(4, 72) = 11.5, p < 5 \times 10^{-5}, \eta^2 = 0.391$ ; *Experiment II*:  $F(7, 35) = 15.1, p < 5 \times 10^{-4}, \eta^2 = 0.752$ ).

Fig. 3 shows the centro-parietal ERP waveforms, and amplitudes extracted to measure the P300 (320–440 ms) and Broad Positivity (490–610 ms), as a function of learning state (Fig. 3C–D), and as a function of sequence repetition block (Fig. 3E–F). With remarkable qualitative similarity across Experiments I and II, the P300 again manifests as a rapid deflection superimposed over the broader positivity, and each demonstrates systematic changes as a function of both learning state and block. One-way ANOVAs revealed that P300 amplitude differed significantly across learning states for both *Experiment I* (Fig. 2C;  $F(1.991, 27.872) = 14.6, p < 1 \times 10^{-3}, h^2 = 0.511$ ), and *Experiment II* (Fig. 2D;  $F(1.895, 9.474) = 13.5, p < 0.002, h^2 = 0.729$ ). Post-hoc comparisons indicated a significantly larger amplitude for ‘still-unknown’ than ‘known’ (*Experiment I*:  $t(14) = 4.0191, p < 1 \times 10^{-3}$ ; *Experiment II*:  $t(5) = 3.0145, p < 0.05$ ), and also for ‘newly-learned’ relative to ‘known’ (*Experiment I*:  $t(14) = 5.0239, p < 1 \times 10^{-4}$ ; *Experiment II*:  $t(5) = 5.8002, p < 5 \times 10^{-3}$ ) trials, while the difference between ‘still-unknown’ and ‘newly-learned’ did not reach significance in either experiment ( $p > 0.05$ ). When measured as the difference between P300 peak and the preceding trough, differences between learning-states were smaller, but remained significant in *Experiment I* between ‘still-unknown’ and ‘known’ (*Experiment I*:  $t(14) = 2.6104, p < 0.05$ ; *Experiment II*:  $t(5) = 1.6422, p = 0.0807$ ), as well as ‘newly-learned’ and ‘known’ (*Experiment I*:  $t(14) = 1.9831, p < 0.05$ ; *Experiment II*:  $t(5) = 1.4775, p = 0.0998$ ) trials. These differences are in line with our hypothesis that P300 amplitude reflects stimulus-bound surprise, which is decreased for sequence elements that are already ‘known’. The temporally broader centro-parietal positivity also differed significantly across learning states for both *Experiment I* ( $F(1.816, 25.424) = 6.54, p < 0.01, h^2 = 0.318$ ) and *Experiment II* ( $F(1.480, 7.399) = 18.6, p < 5 \times 10^{-3}, h^2 = 0.788$ ). As predicted by our second hypothesis, amplitudes were greatest when the presented sequence

element was being newly learned. This was confirmed by significant post-hoc differences between all three experimental conditions (‘newly-learned’ > ‘unknown’ > ‘known’; *Experiment I*: ‘newly-learned’ vs ‘known’:  $t(14) = 3.5563, p < 0.005$ , ‘newly-learned’ vs ‘unknown’:  $t(14) = 1.4539, p = 0.0840$ , ‘unknown’ vs ‘known’:  $t(14) = 2.3497, p < 0.05$ ; *Experiment II*: ‘newly-learned’ vs ‘known’:  $t(5) = 9.537, p < 5 \times 10^{-4}$ , ‘unknown’ vs ‘known’:  $t(5) = 2.9896, p < 0.05$ , ‘newly-learned’ vs ‘unknown’:  $t(5) = 2.3528, p < 0.05$ ). It is worth pointing out that for both neural correlates, effect sizes indicate that ‘learning state’ accounts, on average, for more than half of the variability across subjects.

Having established these learning state effects in planned statistical analyses, we explored the temporal and topographic specificity of the effects by plotting a series of topographies of the F-value quantifying the main effect of learning state (Supplementary Fig. 1). This confirmed that learning state effects common to both experiments are topographically restricted to midline centro-parietal regions. Importantly, no systematic effects of learning state were visible at electrodes surrounding the eyes, counting against any potential confounding influence of eye movement artifacts caused by saccade behaviors that may change over learning. An earlier positive deflection at approximately 250 ms, evident preceding the P300 in Figs. 2 and 3, had a distinct fronto-central topographic focus, and although we had no a priori hypotheses regarding its role in learning, this exploratory analysis revealed an effect of learning state in experiment 2 only. Whereas learning states in this timeframe were ordered in the same way as the later Broad Positivity timeframe (newly-learned greater than both known and unknown) at centro-parietal sites, the fronto-central effect was driven by lower amplitude in the unknown condition relative to both newly-learned and know, confirming that these effects were distinct in nature, and that the early onset of the Broad Positivity in *Experiment II* is not a result of a fronto-central effect spreading posteriorly.

### *Development across sequence repetition*

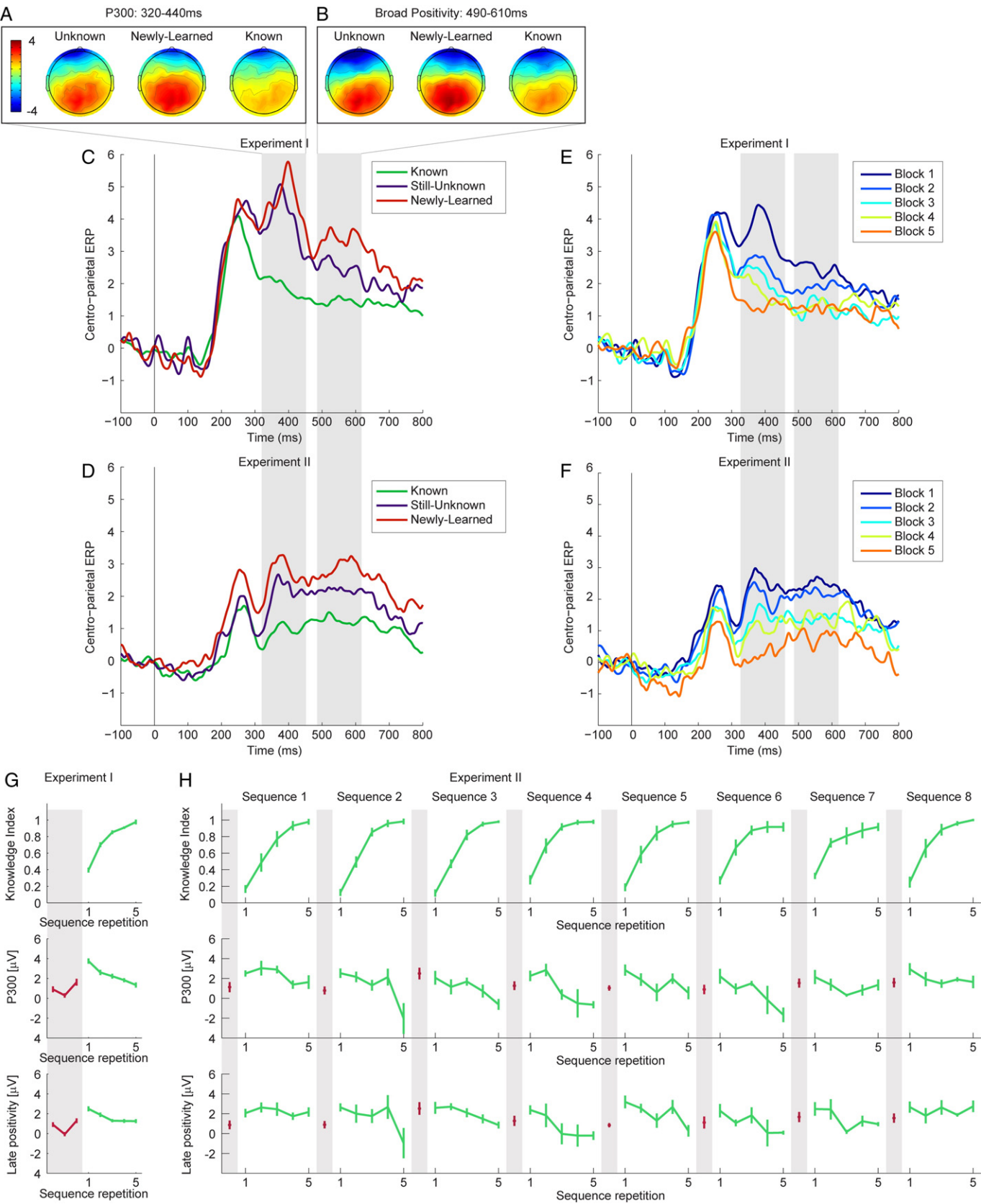
As learning state has an obvious dependence on repetition number, it was important to confirm that differences across learning states are not mere effects of time on task. In *Experiment I*, only sequence repetitions 3, 6, 9, 12, and 15 were available for this analysis, as all other repetitions involved unconstrained movements towards targets (See *Materials and methods*). We thus compared changes in amplitude over visual blocks occurring during sequence learning (collapsed across elements regardless of learning state) to passive ‘random’ blocks (Figs. 3G–H). Two-way repeated-measures ANOVAs with the factors of sequence repetition number and memorization effort (learning versus random blocks) showed significant interactions for both components in both experiments (for P300, *Experiment I*:  $F(2, 108) = 6.9433, p < 0.005$ ; *Experiment II*:  $F(3, 40) = 3.6811, p < 0.05$ ; for Broad Positivity, *Experiment I*:  $F(2, 108) = 6.0288, p < 0.001$ ; *Experiment II*:  $F(3, 40) = 3.0982, p = 0.05686$ ), which in all cases was driven by a significant decrease over sequence learning blocks (for P300, *Experiment I*:  $F(2, 36) = 6.6, p < 0.01, \eta^2 = 0.268$ ; *Experiment II*:  $F(3, 15) = 6.65, p < 0.05, \eta^2 = 0.571$ ; for Broad Positivity, *Experiment I*:  $F(2, 36) = 5.46, p < 0.01, \eta^2 = 0.233$ ; *Experiment II*:  $F(3, 15) = 7.91, p < 0.001, \eta^2 = 0.613$ ) but not random blocks (all  $p > 0.1$ ). There were also significant main effects of memorization effort driven by larger amplitudes during sequence learning than random blocks, which were more reliable for the Broad Positivity (for P300, *Experiment I*:  $F(1, 108) = 48.2792, p < 5 \times 10^{-6}$ ; *Experiment II*:  $F(1, 40) = 2.3572, p = 0.1853$ ; for Broad Positivity, *Experiment I*:  $F(1, 108) = 26.5419, p < 1 \times 10^{-4}$ , *Experiment II*:  $F(1, 40) = 25.5440, p < 0.005$ ). Further discounting the possibility of a time on task confound, neither the changes in P300 (Fig. 3H, middle) nor in the Broad Positivity (bottom) significantly varied as a function of sequence in *Experiment II* ( $p > 0.1$ ).



The gradual process of learning on an element-by-element basis

To provide a stronger test of the distinct roles hypothesized for the P300 and Broad Positivity, we examined the finer temporal

development of the signals with respect to the point of first accurate recall of a given sequence element (newly-learned point). To this end, we sorted trials of Experiment II based on distance (number of blocks) from that point. The ‘newly-learned’ trials themselves thus have a distance of



zero, while ‘unknown’ and ‘known’ trials have negative and positive distances, respectively. This revealed gradual, but markedly distinct, time courses for the P300 deflection and broader positivity over learning-aligned blocks (Fig. 4). The Broad Positivity peaked at the newly-learned point, rose gradually during the preceding sequence presentations, and fell precipitously immediately after. In contrast, P300 amplitude, measured as a trough-to-peak difference to isolate it as much as possible from the broader positivity, shows a progression that is much more consistent with a monotonic decrease with increasing sequence knowledge (Fig. 4B). To statistically test these effects we fit linear mixed-effects models to the average component amplitudes of each sequence learned by each subject, with ‘distance from the block of first accurate recall’ as a fixed effect (*signed distance* or *absolute distance*) and *subject* as a random effect.

The *signed distance* to the first accurate recall had a significant effect on P300 amplitude ( $p < 0.001$ ), and the model approximating P300 as linear functions of the *signed distance* was a significantly better fit to the data than the same model based on the *absolute distance* to the first accurate recall supporting our hypothesis that P300 amplitude decreases monotonically with sequence knowledge ( $p < 1 \times 10^{-5}$ ). The amplitude of the Broad Positivity, on the other hand, is well-predicted by the *absolute distance* to the block of first accurate recall ( $p < 5 \times 10^{-5}$ ), and amplitudes are significantly better explained by the absolute than the signed distance to the first accurate recall ( $p < 1 \times 10^{-5}$ ), indicating that the Broad Positivity is especially elevated while a sequence element is being actively memorized relative to previous blocks in which learning was unsuccessful (‘-1’), or the blocks following consolidation (‘+1’). These distinct time courses are consistent, respectively, with the hypothesized role for the Broad Positivity in active memory formation, and the link between the P300 and the results of that memory formation process (i.e., knowledge).

#### Learning success predictions from P300 and Broad Positivity

Having gained these insights into the involvement of the P300 and broad positive components in sequence learning, we next sought to examine how they could be used to predict learning success within and across subjects. These analyses capitalized on the greater inter-subject variance of Experiment I.

To examine within-subject prediction, we computed correlations between neural signals in each timeframe and behavioral recall indices across blocks within each individual subject. Given the inverse P300-knowledge relationship, we predicted that (1) the mean P300 amplitude of a sequence learning block should be anti-correlated with the amount of sequence knowledge a participant has within that block (‘Knowledge Index’, Eq. 1), and (2) the decrease in P300 amplitude across the three sequence presentations within a block should be correlated with the decrease in surprise and hence the proportion of ‘newly-learned’ sequence elements (‘Learning Index’, Eq. 2). Based on the link between Broad Positivity and active memorization, we further predicted a positive correlation between the Learning Index and the average Broad Positivity. Averaging across the same centro-parietal electrodes as above, we found a significant correlation between P300 amplitude and sequence knowledge ( $p < 0.001$ ; Fig. 5A), a weaker negative correlation between the P300 decrease within sequence blocks

and Learning Index that did not reach significance ( $p > 0.1$ , Fig. 5B), and a strong positive correlation between Broad Positivity amplitude and Learning Index ( $p < 0.001$ ; Fig. 5C). Repeating the tests at all 183 scalp electrodes revealed that the effects were topographically specific (Figs. 5A–C), and that the weak relationship between within-block P300 reduction and Learning Index reached significance at several individual centro-parietal electrodes.

To facilitate the prediction of learning success across subjects in Experiment I, we fit an exponential function to each individual’s learning curve. An example of such a fit is displayed in Fig. 5D, where observed data are indicated by black circles and the exponential fit by a red line. We then correlated the time constant  $\tau$  of this exponential (the time at which the curve reaches  $1 - 1/e \approx 63\%$ ), with two potential neural indicators of learning efficiency: 1) the average Broad Positivity amplitude during the first sequence block, and 2) the decrease in P300 amplitude from the first to the second visual sequence presentation block. The Broad Positivity- $\tau$  correlation showed the expected negative trend, but failed to reach significance ( $p = 0.38$ ). However, there was a strong and significant correlation between the time constant and the initial P300 change ( $r = 0.57$ ;  $p < 0.05$ ), indicating that change in P300 amplitude can be used as a behavioral report-independent measure to assess subjects’ capability to learn sequences of simple stimuli.

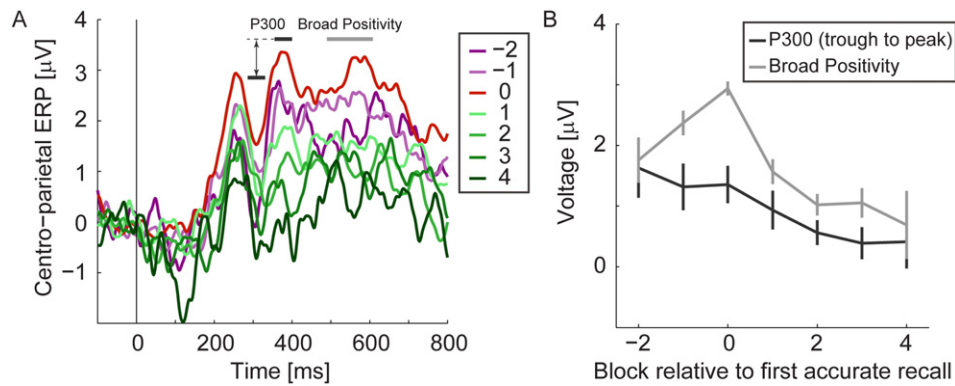
#### Discussion

Sequence learning paradigms, in which a given sequence of elements is memorized over repeated observations, naturally afford the possibility to track the progress of gradual memory formation through regular behavioral recall. In the present study we exploited this ability in order to 1) establish electrophysiological correlates of the gradual transitioning of elements through learning states from unknown to known, and 2) examine the possibility to predict learning success solely from such electrophysiological indices. Using the data of two experiments that respectively emphasized across- and within-subject variance in learning, we characterized two distinct neural signals in terms of their relationship to the point of first successful recall. A rapid P300 process monotonically declined over the course of learning in accordance with the decreasing surprise associated with increasingly confident sequence knowledge. Meanwhile, a temporally broader potential was especially elevated at the point of first successful recall. The learning-related changes in P300 amplitude were so robust as to enable significant prediction of which participants were the fastest learners across the cohort. Such predictive power could have significant implications for the evaluation of true learning capacity in clinical groups with impaired ability to communicate the results of their learning.

The examination of graded transitions through learning states with respect to the point of first accurate recall of each individual element allowed us to gain novel insights into the gradual nature of learning that are precluded in traditional approaches utilizing a remembered-forgotten comparison. Progressive adjustments underlying memory formation have in the past predominantly been studied through human (e.g. Schlaghecken et al., 2000) and animal (e.g. Cooke et al., 2015; Jaffard et al., 1996; Messinger et al., 2001) behavior, and to a lesser extent in animal electrophysiology (e.g. Jaffard et al., 1996; Messinger et al., 2001; Petersen et al., 1998; see Frankland and

**Fig. 3.** Centro-parietal event-related potentials compared across learning state and sequence repetition. (A + B) Scalp topographies of the P300 (320–440 ms) and Broad Positivity (490–610 ms) timeframes for each learning state. For generating topographies, ERPs in each experiment were normalized before averaging across experiments. (C + D) Centro-parietal ERPs for trials of each of the three learning states during sequence learning blocks. P300 amplitudes were significantly greater for unexpected sequence elements (‘newly-learned’ and ‘unknown’) than when a location had already been committed to memory (‘known’), replicated in both Experiment I (C) and Experiment II (D). Broad Positivity amplitude was greater when sequence elements were being ‘newly-learned’ than when learning was unsuccessful (‘unknown’), or they had been committed to memory previously (‘known’). (E + F) P300 deflection decreased significantly over sequence repetitions in both Experiment I (E) and Experiment II (F). The Broad Positivity showed a similar, but less consistent decrease. Block number for Experiment I (E) refers to visual-only blocks. (G + H) Across the repetitions of a single sequence, average recall performance increased monotonically in Experiment I (G: top) as well as Experiment II (H: top). In Experiment I, the first purely visual presentation of a sequence was preceded by a motor learning block of the same sequence, and hence elements can be ‘known’ during the first visual block of Experiment I, but not Experiment II, where there was no motor component. Both P300 amplitude (middle, green traces) and the Broad Positivity (bottom, green traces), decrease across sequence repetitions, but not across sequences (Experiment II: bottom right). There was no significant change in P300 (middle, red traces) or Broad Positivity (bottom, red traces) over random blocks. Error bars quantify the standard error of the mean.





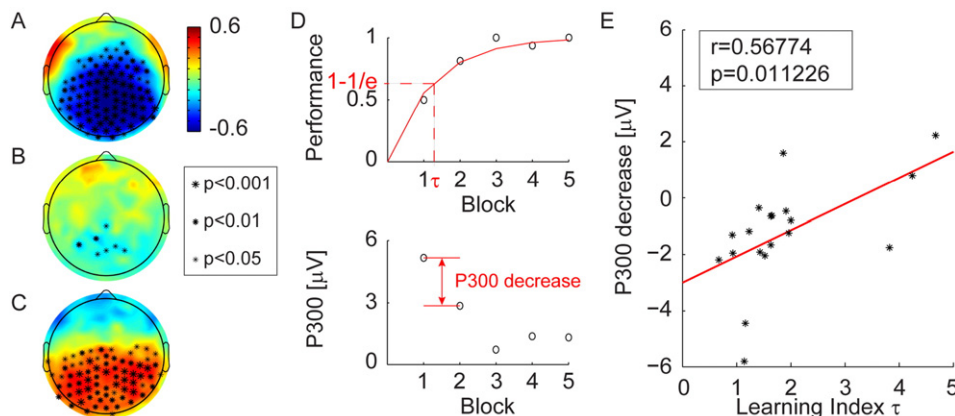
**Fig. 4.** Gradual changes over the course of learning. (A) Centro-parietal ERPs were sub-divided on the basis of temporal distance relative to the block during which a given sequence element was first accurately recalled (block '0', red). Negative distances (purple hues) indicate 'still-unknown' trials, while positive numbers (green hues) are 'known' trials. Over a wide post-stimulus timeframe, the Broad Positivity undergoes clear, gradual changes over the course of learning, reaching its highest amplitude at the point of first accurate recall ('0'). In contrast, the P300, which manifests as a rapid deflection superimposed over the broader positivity peaking at ~380 ms, decreases in size monotonically across the entire learning range '-2' to '+4'. (B) To examine these distinctions in time course more closely, we plot the learning-aligned temporal development of measures of P300 (trough-to-peak) and Broad Positivity extracted from the waveforms in (A). While the P300 (black) shows a monotonic decrease with increasing sequence knowledge, the Broad Positivity (gray) gradually increases towards the point of first accurate recall, and decreases with each additional sequence presentation thereafter.

Bontempi, 2005 for review). Although changes in response to new, unexpected items within increasingly familiar sequences have been reported in the human EEG (Ferdinand et al., 2008), such adjustments could not be related to individual learning curves due to the implicit nature of the task, and whether implicit knowledge can be classed as merely a weaker representation of explicit knowledge is a matter of ongoing debate (Paller et al., 1987b; Reber and Squire, 1998; Schlaghecken et al., 2000). Only animal studies have directly related graded learning curves to neural signals (Jaffard et al., 1996; Messinger et al., 2001), but in very different tasks and during learning processes taking several days rather than the scale of minutes examined in our case.

Although our main hypotheses regarding gradual learning progression were simply that a maximum would be achieved at the point of first accurate recall for the memorization process whereas the surprise-encoding P300 would monotonically decrease, other aspects of the time courses are potentially interesting and may warrant experimental follow-up. For example, the strength of the Broad Positivity one block prior to the point of first accurate recall ('-1') indicates that commission to memory is predicted by not just the current block but efforts

made in the previous block too, and more generally highlights how knowledge is accumulated over several observations of the same information. The sudden drop in the Broad Positivity at '+1' indicates that memorization efforts are immediately reduced upon gaining confident knowledge, and more generally suggests that the process is under strategic, volitional control. At the same time, the continued decrease in P300 amplitude after first successful recall indicates that learning does not altogether cease at this point, and that knowing an element enough to correctly select it among 12 or 16 alternatives is not tantamount to having no doubt about it.

Although never previously characterized in the same manner as in our sequence learning paradigm, both the P300 and a later positive component have been implicated in a long line of EEG studies of learning and memory. A positive relationship between broad, positive ERP waves (sometimes called the "late positive complex") and learning was first described in seminal studies comparing evoked responses to words or images subsequently remembered to those of stimuli later forgotten (Karis et al., 1984; Münte et al., 1988; Neville et al., 1986; Paller et al., 1987b, 1988), and have since been found in a broad variety of



**Fig. 5.** Within and across subject prediction of learning success. (A–C). Within-subject correlations of ERP amplitude and behavior affirm relationships between P300 and stimulus expectancy, and between Broad Positivity and learning. (A) Correlations across sequence repetition block of average amplitude in the time frame of 320–440 ms and Knowledge Index. Significant negative correlations exhibit a topography similar to that of the P300 itself. Small ( $p < 0.05$ ), medium ( $p < 0.01$ ), and large ( $p < 0.001$ ) stars indicate electrodes with a significant correlation between amplitude and sequence knowledge as determined by permutation statistics. (B) An equivalent computation was carried out to correlate the decrease in P300 amplitude (320–440 ms) within a sequence learning block with the number of elements learned during that block. Negative correlations were more focalized, but still significant at centro-parietal electrodes. (C) Across sequence repetition blocks, there is a positive correlation between parietal late positive amplitude (490–610 ms) and Learning Index. The topography of this correlation is consistent with that of the Broad Positivity itself. (D–E) For across-subject predictions of learning success, exponential functions were fit to participants' learning curves. The decrease in average P300 amplitude from sequence block 1 to block 2 was found to predict the time constant  $\tau$  of this fit. The exponential fit (D, top) and P300 decrease (D, bottom) are displayed for one example subject. (E) Across subject correlations showed that we can predict which subjects are fast versus slow learners (small versus big time constants) based on electrophysiological measures alone ( $r = 0.57$ ;  $p < 0.05$ ).

paradigms (e.g. Ferdinand et al., 2008; Gonsalves and Paller, 2000; Schlaghecken et al., 2000). A number of negative results involving tasks of 'shallow' processing (Paller et al., 1987a; Sanquist et al., 1980; Weyerts et al., 1997) or stimuli of low complexity (Van Petten and Senkfor, 1996), however, resulted in remaining doubt as to whether such neural correlates could be observed for simplistic stimuli. To our knowledge, our study is the first to observe correlates of memory formation for such simple visual stimuli with minimal semantic value. Stimuli of low semantic content have previously been used to study recognition rather than encoding (Voss and Paller, 2009), and in this as well as other studies, neural signatures of recognition have tended to be broadly similar in timing and topography to those of encoding (see also Curran, 2004; Curran and Cleary, 2003; Friedman, 1990; Gonsalves and Paller, 2000; Paller and Kutas, 1992). However, since memory extraction processes would occur alongside additional decision processes required for subjects to report on recognition of the image itself as well as indicate their confidence (Curran, 2004; Voss and Paller, 2009), the exact relationship between these correlates of recognition and those of encoding remains unclear. Future work that brings together the complementary advantages of each of these approaches may provide a means to dissociate neural correlates of the constituent processes underlying both memory formation and retrieval. Considering memory formation alone, our findings related to the broad positivity are broadly consistent with the extant literature on the role of late positive ERP activity in encoding memories, and extend our understanding by confirming its role not just during the "study" phase of classical memory paradigms, but also in active memory formation for sequences encoded more gradually over repeated observations.

Our tracking of P300 amplitude as an index of stimulus-bound surprise builds on previous work examining changes in stimulus-expectation as a function of recent trial history as well as overall estimations of stimulus probability, including recent model-based approaches (Duncan-Johnson and Donchin, 1977; Kolossa et al., 2012; Mars et al., 2008; Squires et al., 1976). We are not the first to capitalize on this sensitivity in the study of sequence learning. Several studies, especially those in which subjects learned intentionally and gained explicit knowledge, have shown greatly enhanced P300 amplitudes for *deviant* stimuli – stimuli that do not obey the structure of a learned sequence and are therefore unexpected – compared to expected, standard sequence elements (Fu et al., 2013; Miyawaki et al., 2005; Rüsseler et al., 2003; Schlaghecken et al., 2000). In some cases, this effect was exclusive to later recalled or recognized portions of the sequence (Miyawaki et al., 2005; Schlaghecken et al., 2000), or dependent on whether subjects were learning the sequence intentionally (Rüsseler et al., 2003; Schlaghecken et al., 2000). These effects led the investigators to suggest that the number of learned sequence elements may be predicted from this difference in P300 amplitude (Miyawaki et al., 2005; Rüsseler et al., 2003; Schlaghecken et al., 2000), although none of these studies attempted to test this idea. A critical distinction of the present study is that we have demonstrated the feasibility of such prediction based on the actual content to be learned, without the use of disruptive deviants to probe the state of learning of that content.

Although the learning-related P300 changes mediated by surprise, which we predicted and observed in the present study, would seem logical based on the well-established sensitivity of the P300 to stimulus probability, it was far from a foregone conclusion since discrepant views on the role of the P300 in learning have long permeated the literature (Donchin and Coles, 1988; Polich 2007; Verleger 2008; Verleger, 1988). Based on its sensitivity to subjective stimulus probability, task relevance, and later remembering, and the variation of its peak latency depending on the time required for stimulus evaluation, the P300 was initially assumed to play an active role in memory formation (Donchin and Coles, 1988). Dominant theories have asserted that the P300 is associated with the memory-related concept of "context updating" (Donchin and Coles, 1988; Duncan-Johnson and Donchin, 1982), or the allocation of attentional resources in response to a change in

context, which in turn facilitates memory formation (Polich 2007). More recently it was asserted that the P300's attentional role is such that it should increase for stimuli as they are better stored in memory (Jongsma et al., 2012), the polar opposite prediction to a surprise-based account. Mirroring this lack of theoretical convergence, findings of subsequent memory effects on the P300 have not been consistent (e.g. Chapman et al., 1981; Fabiani et al., 1986, 1990; Rushby et al., 2002), and as was pointed out early in the debate, may express a dependency on 'memorability' rather than an involvement in memorization (Donchin, 1981). Our results stand to shed light on this array of discrepant findings because they suggest that there may be two dissociable processes of similar topography that systematically change with learning, but in distinct ways which are revealed when the progress of learning is traced on an element-by-element basis according to recall performance. Further, recent work has provided strong evidence for an earlier proposal that the P300 plays a role in decision formation (Hillyard et al., 1971; Kutas et al., 1977; Nieuwenhuis et al., 2005; Verleger et al., 2005), by demonstrating dynamics consistent with accumulation of sensory evidence to a decision bound (Kelly and O'Connell 2013; O'Connell et al., 2012; Twomey et al., 2015). This provides an explanation not only for the scaling of the P300 with surprise in terms of modulated distance to bound depending on priors (Ratcliff and McKoon, 2008), but also for the well-known dependence of peak latency on the amount of processing required for full stimulus evaluation (Kutas et al., 1977; Magliero et al., 1984; McCarthy and Donchin 1988; Verleger 1997). For the highly reduced, elementary location elements of our task, the P300 peaked rapidly and sharply in the stimulus-aligned average waveforms. This was critical to our ability to dissociate it from the broader positivity based on their distinct time scales, and despite their apparent overlap. An important general implication of the P300's link with decision formation is that reduced amplitude can result not just from high prior probability but alternatively from a tendency not to make a decision at all on some trials. This could be a significant consideration for some aspects of learning and memory; for example, where strong primacy effects are observed in sequence learning (Azizian and Polich, 2007; Grune et al., 1996; Rushby et al., 2002; Wiswede et al., 2007), plateau elements may not actually be processed to the point of full identification initially, in favor of first committing the beginning elements to memory.

If the P300 represents identification of the current item, a plausible interpretation of the broader positivity would be of a temporally less constrained representation of previous and future sequence elements, possibly alongside a prolonged representation of the current one. Sustained activity has long been regarded as the neural basis of holding items in working memory (e.g. D'Esposito and Postle 2015; Fuster and Jervey, 1982; Miller et al., 1996; Postle, 2011), and influential theories of memorization have posited that the strength and duration of working memory representations strongly determines longer-term memorization (Atkinson and Shiffrin, 1968; Davachi et al., 2001; Jensen and Lisman, 2005). Studies using transcranial magnetic stimulation show that disrupting these prolonged representations may hinder successful memorization (Rossi et al., 2011). In the particular case of serial events, the working memory representations required to form associations would include preceding as well as current stimuli (Beiser and Houk, 1998). In fact, it has been suggested that forming associations between stimuli more generally involves forging a representation of each stimulus upon the presentation of the other (Messinger et al., 2001, 2005). Work mainly done in rodents has shown that neural representations of spatial location sequences are replayed in hippocampal areas during resting periods after finishing the sequence (Huber et al., 2004; Pavlides and Winson, 1989; Peigneux et al., 2004; Wilson and McNaughton, 1994; reviewed in Pezzulo et al., 2014), and intracranial recordings in humans indicate that similar cell populations exist in the human hippocampus (Ekstrom et al., 2003). Although we cannot say with certainty that the same mechanisms are at play here, it is plausible that some types of learning require the re-representation of stimuli to

commit them to memory and that these re-representations are reflected at similar cortical levels as the original stimulus representation formed during identification, thus resulting in similar activation patterns when measured noninvasively (Ester et al., 2013; Ester et al., 2015). In essence, these ideas parallel interpretations of the Broad Positivity that have been made in the past, implicating deep semantic (Friedman et al., 1996; Paller et al., 1987a) or even imaginary representations (Gonsalves and Paller, 2000), the common thread being that prolonged, internally-driven, associative representations connected to a stimulus – in our case of the neighboring elements – create a necessary representational context in which to integrate new information and ultimately aid remembering.

A large body of literature has focused on localizing the neural sources of the P300 in the brain, using a range of techniques including fMRI (e.g. Stevens et al., 2000), EEG source analysis (e.g. Mulert et al., 2004), human lesion studies (e.g. Knight et al., 1989; Verleger et al., 1994), and intracranial recordings in both humans (Halgren et al., 1980) and animals (e.g. Katayama et al., 1985), and across these studies a wide range of areas have been implicated (Nieuwenhuis et al., 2005). Most of this work has considered not just the centro-parietal P300, or ‘P3b’ as measured in the current study, but also the so-called ‘P3a’ which, in contrast to the P3b, activates especially in response to salient, novel stimuli that need not be task-relevant (Polich 2007). Consistent with this functional relationship, the P3a has been linked to parietal regions including the temporoparietal junction (TPJ), which is known to play an important role in orienting to salient events (Corbetta and Shulman 2002), as well as prefrontal regions (Linden, 2005; Mulert et al., 2004). The P3b has the distinction of being selectively evoked by task-relevant stimuli, which would appear to be incompatible with brain areas that respond to purely bottom-up forms of salience; however, lesion studies have shown that the P3b is compromised by some of the same salience-sensitive regions that have been linked to the P3a, including the TPJ (Knight et al., 1989; Verleger et al., 1994). Many have concluded that this is but one node of a widely distributed network of cortical areas contributing to the P300 (Nieuwenhuis et al., 2005), on the basis of fMRI and source analysis (Mulert et al., 2004), EEG-magnetoencephalography comparisons (Ahlfors et al., 2010), and human intracranial recordings from several different regions including medial temporal lobe structures (e.g. Halgren et al., 1980; Halgren et al., 1998), thalamus (Yingling and Hosobuchi, 1984), superior temporal sulcus, ventrolateral prefrontal cortex, and posterior superior parietal cortex (Halgren et al., 1998).

The close spatiotemporal overlap of functionally distinct ‘P3a’ and ‘P3b’ components has motivated the development of spatiotemporal Principal Component Analysis (PCA) approaches to parse the “late positive complex” in order to separate these components (e.g. Dien et al., 2004; Dien, 2012). Such approaches are highly effective in such decomposition problems where, despite significant overlap, there is a systematic difference in both topography (P3a is more frontal) and timing (P3a is 60–80 ms earlier). However, the decomposition of the P300 from the Broad Positivity presents a somewhat different challenge in that their topographies are highly similar, thus precluding spatial PCA or other linear spatial projections (see e.g. Parra et al., 2005), and, to the extent that we have correctly interpreted the Broad Positivity as an endogenously generated representation of information other than the immediate element, it would likely be both broader and much more variable in timing across trials, precluding the use of temporal PCA which relies on relatively invariant post-stimulus timing (Dien, 2012). Our solution was a very elementary one based on the premise that two superimposed components whose trial-average differs markedly in temporal scale can be measured with minimal mutual conflation by using a temporally local baseline for the faster process (P300 peak relative to preceding trough) and simply using a window well beyond the faster process to measure the broader one. As these components become better characterized, no doubt more sophisticated decomposition approaches can be developed to more cleanly isolate them and will be important for clinical applications.

Across participants, individual learning success, measured in the time constant of the exponential fit to participants’ learning curves, was significantly predicted by the decrease in P300 amplitude from the first to the second block. Although a similar analysis based on Broad Positivity amplitude failed to reach significance, a critical difference between the two measures is that the P300 *change*, being a within-subject differential measure, is self-normalized and hence is more resilient to spurious factors of inter-individual variability (e.g. brain geometry dictating electrical field projections on the scalp) in a way that the absolute measure of Broad Positivity would not be. Being able to predict learning success across subjects independent of verbal reports may prove a valuable tool in diagnosing learning abilities in children. If patients are reluctant or unable to report their learning success, it is often challenging to determine whether the underlying cause is the disability to learn, anxiety to communicate learning success, or difficulties in stimulus perception. Further studies in target populations will be needed to illuminate whether the P300 as a neural correlate of stimulus identification and broader positivities indicating learning can indeed distinguish the three conditions more reliably.

If learning success could be predicted from electroencephalographic measures alone and completely independent of behavioral reports made by the participants, this would be a great step towards the development of new diagnostic tools for learning disabilities, as the ability to commit elements to memory could be determined even if communication of progress was inhibited by other factors.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2016.05.001>.

## Acknowledgements

This work was supported by the National Science Foundation (BCS-1358955). Individual support was granted by the McDonnell Foundation; and the National Institute of Health (P01 NS-083514) to MFG, and by the German National Academic Foundation (ERP-Fellowship) to NS.

## References

- Ahlfors, S.P., Han, J., Lin, F.-H., Witzel, T., Belliveau, J.W., Hämäläinen, M.S., Halgren, E., 2010. Cancellation of EEG and MEG signals generated by extended and distributed sources. *Hum. Brain Mapp.* 31 (1), 140–149.
- Aizenstein, H.J., Stenger, V.A., Cochran, J., Clark, K., Johnson, M., Nebes, R.D., Carter, C.S., 2004. Regional brain activation during concurrent implicit and explicit sequence learning. *Cereb. Cortex* 14, 199–208.
- Atkinson, R.C., Shiffrin, R.M., 1968. Human memory: a proposed system and its control processes. In: Spence, K.W., Spence, J.T. (Eds.), *The psychology of learning and motivation: Advances in research and theory* vol. 2. Academic Press, New York (NY).
- Azizian, A., Polich, J., 2007. Evidence for attentional gradient in the serial position memory curve from event-related potentials. *J. Cogn. Neurosci.* 19 (12), 2071–2081.
- Baldwin, K.B., Kutas, M., 1997. An ERP analysis of implicit structured sequence learning. *Psychophysiology* 34 (1), 74–86.
- Beiser, D.G., Houk, J.C., 1998. Model of cortical-basal processing: encoding the serial order of sensory events. *J. Neurophysiol.* 79, 3168–3188.
- Brewer, J.B., Zhao, Z., Desmond, J.E., Glover, G.H., Gabrieli, J.D., 1998. Making memories: brain activity that predicts how well visual experience will be remembered. *Science* 281, 1185–1187.
- Chapman, R.M., McCarthy, J.W., Chapman, J.A., 1981. Memory processes and evoked potentials. *Can. J. Psychol.* 35 (2), 201–212.
- Cooke, S.F., Komorowski, R.W., Kaplan, E.S., Gavornik, J.P., Bear, M.F., 2015. Visual recognition memory, manifested as long-term habituation, requires synaptic plasticity in V1. *Nat. Neurosci.* 18 (2), 262–271.
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nature* 3, 201–215.
- Curran, T., 2004. Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia* 42 (8), 1088–1106.
- Curran, T., Cleary, A.M., 2003. Using ERPs to dissociate recollection from familiarity in picture recognition. *Cogn. Brain Res.* 15 (2), 191–205.
- Davachi, L., Maril, A., Wagner, A.D., 2001. When keeping in mind supports later bringing to mind: neural markers of phonological rehearsal predict subsequent remembering. *J. Cogn. Neurosci.* 13 (8), 1059–1070.
- D’Esposito, M., Postle, B.R., 2015. The cognitive neuroscience of working memory. *Annu. Rev. Psychol.* 66, 115–142.
- De Chastellaine, M., Rugg, M.D., 2014. The relationship between task-related and subsequent memory effects. *Hum. Brain Mapp.* 35 (8), 3687–3700.



- Dien, J., 2012. Applying principal components analysis to event-related potentials: a tutorial. *Dev. Neuropsychol.* 37 (6), 497–515.
- Dien, J., Spencer, K.M., Donchin, E., 2004. Parsing the late positive complex: mental chronometry and the ERP components that inhabit the neighborhood of the P300. *Psychophysiology* 41 (5), 665–678.
- Donchin, E., 1981. Surprise!... Surprise? *Psychophysiology*, 18(5), 493–513.
- Donchin, E., Coles, M.G.H., 1988. Is the P300 component a manifestation of context updating? *Behav. Brain Sci.* 11 (03), 357–374.
- Duncan-Johnson, C.C., Donchin, E., 1977. On quantifying surprise: the variation of event-related potentials with subjective probability. *Psychophysiology* 14 (5), 456–467.
- Duncan-Johnson, C.C., Donchin, E., 1982. The P300 component of the event-related brain potential as an index of information processing. *Biol. Psychol.* 14 (1–2), 1–52.
- Eimer, M., Goschke, T., Schlaghecken, F., Stürmer, B., 1996. Explicit and implicit learning of event sequences: evidence from event-related brain potentials. *J. Exp. Psychol. Learn. Mem. Cogn.* 22 (4), 970–987.
- Ekstrom, A.D., Kahana, M.J., Caplan, J.B., Fields, T.A., Isham, E.A., Newman, E.L., Fried, I., 2003. Cellular networks underlying human spatial navigation. *Letters to Nature* 425, 184–187.
- Ester, E.F., Anderson, D.E., Serences, J.T., Awh, E., 2013. A neural measure of precision in visual working memory. *J. Cogn. Neurosci.* 25 (5), 754–761.
- Ester, E.F., Sprague, T.C., Serences, J.T., 2015. Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron* 87, 1–13.
- Fabiani, M., Karis, D., Donchin, E., 1986. P300 and recall in an incidental memory paradigm. *Psychophysiology* 23 (3), 298–308.
- Fabiani, M., Karis, D., Donchin, E., 1990. Effects of mnemonic strategy manipulation in a Von Restorff paradigm. *Electroencephalogr. Clin. Neurophysiol.* 75 (2), 22–35.
- Ferdinand, N.K., Mecklinger, A., Kray, J., 2008. Error and deviance processing in implicit and explicit sequence learning. *J. Cogn. Neurosci.* 20 (4), 629–642.
- Fernandez, G., Effern, A., Grundwald, T., Pezer, N., Lehnertz, K., Dimpelmann, M., Van Roost, D., Elger, C.E., 1999. Real-time tracking of memory formation in the human rhinal cortex and hippocampus. *Science* 285 (5433), 1582–1585.
- Fu, Q., Bin, G., Dienes, Z., Fu, X., Gao, X., 2013. Learning without consciously knowing: evidence from event-related potentials in sequence learning. *Conscious. Cogn.* 22 (1), 22–34.
- Frankland, P.W., Bontempi, B., 2005. The organization of recent and remote memories. *Nat. Rev. Neurosci.* 6, 119–130.
- Friedman, D., 1990. ERPs during continuous recognition memory for words. *Biol. Psychol.* 30, 61–87.
- Friedman, D., Ritter, W., Snodgrass, J.G., 1996. ERPs during study as a function of subsequent direct and indirect memory testing in young and old adults. *Brain Res. Cogn. Brain Res.* 4 (1), 1–13.
- Fuster, J.M., Jervey, J.P., 1982. Neuronal firing in the inferotemporal cortex of the monkey in a visual memory task. *J. Neurosci.* 2 (3), 361–375.
- Gavornik, J.P., Bear, M.F., 2014. Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nat. Neurosci.* 17 (5), 732–737.
- Ghilardi, M.F., Carbon, M., Silvestri, G., Dhawan, V., Tagliati, M., Bressman, S., Ghez, C., Eidelberg, D., 2003a. Impaired sequence learning in carriers of the DYT1 dystonia mutation. *Ann. Neurol.* 54 (1), 102–109.
- Ghilardi, M.F., Eidelberg, D., Silvestri, G., Ghez, C., 2003b. The differential effect of PD and normal aging on early explicit sequence learning. *Neurology* 60 (8), 1313–1319.
- Ghilardi, M.F., Feigin, A.S., Battaglia, F., Silvestri, G., Mattis, P., Eidelberg, D., Di Rocco, A., 2007. L-Dopa infusion does not improve explicit sequence learning in Parkinson's disease. *Parkinsonism Relat. Disord.* 13(3), 146–151.
- Ghilardi, M.F., Silvestri, G., Feigin, A., Mattis, P., Zgaljardic, D., Moissello, C., Crupi, D., Di Rocco, A., Eidelberg, D., 2008. Implicit and explicit aspects of sequence learning in pre-symptomatic Huntington's disease. *Parkinsonism Relat. Disord.* 14 (6), 457–464.
- Gobet, F., Lane, P.C.R., Croker, S., Cheng, P.C.H., Jones, G., Oliver, I., Pine, J.M., 2001. Chunking mechanisms in human learning. *Trends Cogn. Sci.* 5 (6), 236–243.
- Gonsalves, B., Paller, K.A., 2000. Neural events that underlie remembering something that never happened. *Nat. Neurosci.* 3 (12), 1316–1321.
- Grune, K., Metz, A.M., Hagendorf, H., Fischer, S., 1996. Information processing in working memory and event-related brain potentials. *Int. J. Psychophysiol.* 23 (1–2), 111–120.
- Halgren, E., Marinkovic, K., Chauvel, P., 1998. Generators of the late cognitive potentials in auditory and visual oddball tasks. *Electroencephalogr. Clin. Neurophysiol.* 106, 156–164.
- Halgren, E., Squire, N.K., Wilson, C.L., Rohrbaugh, J.W., Babb, T.L., Crandall, P.H., 1980. Endogenous potentials generated in the human hippocampal formation and amygdala by infrequent events. *Science* 210, 803–805.
- Hillyard, S.A., Squires, K.C., Bauer, J.W., Lindsay, P.H., 1971. Evoked potential correlates of auditory signal detection. *Science* 25 (172), 1357–1360.
- Horst, R.L., Johnson, R., Donchin, E., 1980. Event-related brain potentials and subjective probability in a learning task. *Mem. Cognit.* 8, 476–488.
- Huber, R., Ghilardi, M.F., Massimini, M., Tononi, G., 2004. Local sleep and learning. *Letters to Nature* 430, 78–81.
- Jaffard, R., Vouimba, R.M., Marighetto, A., Garcia, R., 1996. Long-term potentiation and long-term depression in the lateral septum in spatial working and reference memory. *J. Physiol. Paris* 90 (5–6), 339–341.
- Jensen, O., Lisman, J.E., 2005. Hippocampal sequence encoding driven by a cortical multi-item working memory buffer. *Trends Cogn. Sci.* 28 (2), 67–72.
- Jongsma, M.L.A., Gerrits, N.J.H.M., van Rijn, C.M., Quian Quiroga, R., Maes, J.H.R., 2012. Event related potentials to digit learning: tracking neurophysiologic changes accompanying recall performance. *Int. J. Psychophysiol.* 85, 41–48.
- Kandel, E.R., Schwartz, J.H., 1982. Molecular biology of learning: modulation of transmitter release. *Science* 218 (4571), 433–443.
- Karis, D., Fabiani, M., Donchin, E., 1984. “P300” and memory: individual differences in the von Restorff effect. *Cogn. Psychol.* 16 (2), 177–216.
- Katayama, Y., Tsukiyama, T., Tsubokawa, T., 1985. Thalamic negativity associated with the endogenous late positive component of cerebral evoked potentials (P300): recordings using discriminative aversive conditioning in humans and cats. *Brain Res. Bull.* 14, 223–226.
- Kelly, S.P., O'Connell, R.G., 2013. Internal and external influences on the rate of sensory evidence accumulation in the human brain. *J. Neurosci.* 33 (50), 19434–19441.
- Kim, H., 2011. Neural activity that predicts subsequent memory and forgetting: a meta-analysis of 74 fMRI studies. *NeuroImage* 54, 2446–2461.
- Knight, R.T., Scabini, D., Woods, D.L., Clayworth, C.C., 1989. Contributions of temporal-parietal junction to the human auditory P3. *Brain Res.* 502, 109–116.
- Kolossa, A., Fingscheidt, T., Wessel, K., Kopp, B., 2012. A model-based approach to trial-by-trial p300 amplitude fluctuations. *Front. Hum. Neurosci.* 6, 359.
- Kutas, M., McCarthy, G., Donchin, E., 1977. Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science* 197, 792–795.
- Linden, D.J., 2005. The P300: where in the brain is it produced and what does it tell us? *Neuroscientist* 11 (6), 563–576.
- Madhavan, R., Millman, D., Tang, H., Crone, N.E., Lenz, F.A., Tierney, T.S., Madsen, J.R., Kreiman, G., Anderson, W.S., 2015. Decrease in gamma-band activity tracks sequence learning. *Front. Syst. Neurosci.* 8, 1–16.
- Magliero, A., Bashore, T.R., Coles, M.G.H., Donchin, E., 1984. On the dependence of P300 latency on stimulus evaluation processes. *Psychophysiology* 21 (2), 171–186.
- Mars, R.B., Debener, S., Gladwin, T.E., Harrison, L.M., Haggard, P., Rothwell, J.C., Bestmann, S., 2008. Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *J. Cogn. Neurosci.* 28 (47), 12539–12545.
- Martin, S.J., Grimwood, P.D., Morris, R.G.M., 2000. Synaptic plasticity and memory: an evaluation of the hypothesis. *Annu. Rev. Neurosci.* 23(1), 649–711.
- McCarthy, G., Donchin, E., 1988. A metric for thought: a comparison of P300 latency and reaction time. *Science* 211, 77–80.
- Messinger, A., Squire, L.R., Zola, S.M., Albright, T.D., 2001. Neuronal representations of stimulus associations develop in the temporal lobe during learning. *Proc. Natl. Acad. Sci.* 98 (21), 12239–12244.
- Messinger, A., Squire, L.R., Zola, S.M., Albright, T.D., 2005. Neural correlates of knowledge: Stable representation of stimulus associations across variations in behavioral performance. *Neuron* 48 (2), 359–371.
- Miller, E.K., Erickson, C.A., Desimone, R., 1996. Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J. Neurosci.* 16 (16), 5154–5167.
- Miyawaki, K., Sato, A., Yasuda, A., Kumano, H., Kuboki, T., 2005. Explicit knowledge and intention to learn in sequence learning: an event-related potential study. *Neuroreport* 16 (7), 705–708.
- Moissello, C., Mezziane, H., Kelly, S.P., Perfetti, B., Kvint, S., Voutsinas, N., Blanco, D., Quartarone, A., Tononi, G., Ghilardi, M.F., 2013. Neural activations during visual sequence learning leave a trace in post-training spontaneous EEG. *PLoS One* 8 (6), e65882.
- Mulert, C., Pogarell, O., Juckel, G., Rujescu, D., Giegling, I., Rupp, D., Mavroggiorgou, P., Bussfeld, P., Gallinat, J., Möller, Hegerl, U., 2004. The neural basis of the P300 potential. Focus on the time-course of the underlying cortical generators. *Eur. Arch. Psychiatry. Clin. Neurosci.* 254, 190–198.
- Münste, T.F., Heinze, H.J., Scholz, M., Künkel, H., 1988. Effects of a cholinergic nootropic, WEB 1881 FU on event-related potentials recorded in incidental and intentional memory tasks. *Neuropsychobiology* 19, 158–168.
- Neville, H.J., Kutas, M., Chesney, G., Schmidt, A.L., 1986. Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. *J. Mem. Lang.* 25, 75–92.
- Nieuwenhuis, S., Aston-Jones, G., Cohen, J.D., 2005. Decision Making, the P3, and the Locus Coeruleus-Norepinephrine System. *Psychol. Bull.* 131 (4), 510–532.
- O'Connell, R.G., Dockree, P.M., Kelly, S.P., 2012. A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat. Neurosci.* 15 (12), 1729–1735.
- Otten, L.J., Henson, R.N.A., Rugg, M.D., 2001. Depth of processing effects on neural correlates of memory encoding: relationship between findings from across- and within-task comparisons. *Brain* 124, 399–412.
- Paller, K.A., Kutas, M., Mayes, A.R., 1987a. Neural correlates of encoding in an incidental learning paradigm. *Electroencephalogr. Clin. Neurophysiol.* 67 (4), 360–371.
- Paller, K.A., Kutas, M., Shimamura, A.P., Squire, L.R., 1987b. Brain responses to concrete and abstract words reflect processes that correlate with later performance on a test of stem-completion priming. *Electroencephalogr. Clin. Neurophysiol.* 40, 360–365.
- Paller, K.A., McCarthy, G., Wood, C.C., 1988. ERPs predictive of subsequent recall and recognition performance. *Biol. Psychol.* 26 (1–3), 269–276.
- Paller, K.A., Kutas, M., 1992. Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *J. Cogn. Neurosci.* 4 (4), 375–392.
- Pavlices, C., Winson, J., 1989. Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. *J. Neurosci.* 9 (8), 2907–2918.
- Parra, L.C., Spence, C.D., Gerson, A.D., Sajda, P., 2005. Recipes for the linear analysis of EEG. *NeuroImage* 28, 326–341.
- Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., Phillips, C., Degueldre, C., Del Fiore, G., Aerts, J., Luxen, A., Maquet, P., 2004. Are spatial memories strengthened in the human hippocampus during slow wave sleep? *Neuron* 44, 535–545.
- Petersen, C.C.H., Malenka, R.C., Nicoll, R.A., Hopfield, J.J., 1998. All-or-none potentiation at CA3-CA1 synapses. *Proc. Natl. Acad. Sci.* 95, 4732–4737.
- Pezzuolo, G., van der Meer, M.A.A., Lansink, C.S., Pennartz, C.M.A., 2014. Internally generated sequences in learning and executing goal-directed behavior. *Trends Cogn. Sci.* 18 (12), 647–657.
- Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118 (10), 2128–2148.

- Postle, B.R., 2011. What underlies the ability to guide action with spatial information that is no longer present in the environment? In: Vandierendonck, A., Szmalec, A. (Eds.), *Spatial Working Memory*. Psychology Press, Hove (UK).
- Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20 (4), 873–922.
- Reber, P.J., Squire, L.R., 1998. Encapsulation of implicit and explicit memory in sequence learning. *J. Cogn. Neurosci.* 10 (2), 248–263.
- Rossi, S., Innocenti, I., Polizzotto, N.R., Feurra, M., De Capua, A., Olivelli, M., Bartolini, S., Cappa, S.F., 2011. Temporal dynamics of memory trace formation in the human prefrontal cortex. *Cereb. Cortex* 21, 368–373.
- Rushby, J.A., Barry, R.J., Johnstone, S.S., 2002. Event-related potential correlates of serial-position effects during an elaborative memory test. *Int. J. Psychophysiol.* 46 (1), 13–27.
- Rüsseler, J., Kuhlcke, D., Münte, T.F., 2003. Human error monitoring during implicit and explicit learning of a sensorimotor sequence. *Neurosci. Res.* 47 (2), 233–240.
- Sanquist, T.F., Rohrbaugh, J.W., Syndulko, K., Lindsley, D.B., 1980. Electrocorical signs of levels of processing: perceptual analysis and recognition memory. *Psychophysiology* 17 (6), 568–576.
- Schendan, H.E., Searl, M.M., Melrose, R.J., Stern, C.E., 2003. An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron* 37, 1013–1025.
- Schlaghecken, F., Stürmer, B., Eimer, M., 2000. Chunking processes in the learning of event sequences: Electrophysiological indicators. *Mem. Cognit.* 28, 821–831.
- Squires, K.C., Wickens, C., Squires, N.K., Donchin, E., 1976. The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science* 193, 1142–1146.
- Stevens, A.A., Skudlarski, P., Gatenby, C., Gore, J.C., 2000. Event-related fMRI of auditory and visual oddball tasks. *J. Magn. Reson. Imaging* 18, 495–502.
- Sun, R., 2003. Introduction to sequence learning. In: Sun, R., Giles, C.L. (Eds.), *Sequence Learning: Paradigms, Algorithms, and Applications*. Springer, Heidelberg.
- Sutton, S., Braren, M., Zubin, J., John, E.R., 1965. Evoked-potential correlates of stimulus uncertainty. *Science* 150 (3700), 1187–1188.
- Twomey, D.M., Murphy, P.R., Kelly, S.P., O'Connell, R.G., 2015. The classic P300 encodes a build-to-threshold decision variable. *Eur. J. Neurosci.* 42 (1), 1636–1643.
- Van Petten, C., Senkfor, A.J., 1996. Memory for words and novel visual patterns: repetition, recognition, and encoding effects in the event-related brain potential. *Psychophysiology* 33, 491–506.
- Verleger, R., 1988. Event-related potentials and cognition: a critique of the context updating hypothesis and an alternative interpretation of P3. *Behav. Brain Sci.* 11, 343–427.
- Verleger, R., Heide, W., Butt, C., Kömpf, D., 1994. Reduction of P3b in patients with temporo-parietal lesions. *Cogn. Brain Res.* 2, 103–116.
- Verleger, R., 1997. On the utility of P3 latency as an index of mental chronometry. *Psychophysiology* 34, 131–156.
- Verleger, R., Jaskowski, P., Wascher, E., 2005. Evidence for an integrative role of P3b in linking reaction to perception. *J. Psychophysiol.* 19 (3), 165–181.
- Verleger, R., 2008. Letter to the Editor of *Clinical Neurophysiology* 119, 968–990.
- Voss, J.L., Paller, K.A., 2009. An electrophysiological signature of unconscious recognition memory. *Nat. Neurosci.* 12 (2), 349–355.
- Wagner, A.D., Koutstaal, W., Schacter, D.L., 1999. When encoding yields remembering: insights from event-related neuroimaging. *Philos. Trans. R. Soc. B.* 354 (1387), 1307–1324.
- Weyerts, H., Tendolkar, I., Smid, H.G., Heinze, H.J., 1997. ERPs to encoding and recognition in two different inter-item association tasks. *Neuroreport* 8 (7), 1583–1588.
- Wilson, M.A., McNaughton, B.L., 1994. Reactivation of hippocampal ensemble memories during sleep. *Science* 265 (5172), 676–679.
- Wiswede, D., Rüsseler, J., Münte, T., 2007. Serial position effects in free memory recall – An ERP-study. *Biol. Psychol.* 75, 185–193.
- Wong, J.X., de Chastelaine, M., Rugg, M.D., 2013. Comparison of the neural correlates of encoding item-item and item-context associations. *Front. Hum. Neurosci.* 7, 436.
- Yingling, C.D., Hosobuchi, Y., 1984. A subcortical correlate of P300 in man. *Electroencephalogr. Clin. Neurophysiol.* 59, 72–76.
- Zhuang, P., Toro, C., Grafman, J., Manganotti, P., Leocani, L., Hallett, M., 1997. Event-related desynchronization (ERD) in the alpha frequency during development of implicit and explicit learning. *Electroencephalogr. Clin. Neurophysiol.* 102, 374–381.