NeuroImage 108 (2015) 60–73

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

The discovery of processing stages: Analyzing EEG data with hidden semi-Markov models

Jelmer P. Borst [a,b,*], John R. Anderson [a]

[a] Carnegie Mellon University, Dept. of Psychology, Pittsburgh, USA
[b] University of Groningen, Dept. of Artificial Intelligence, Groningen, The Netherlands

ARTICLE INFO

Article history:
Accepted 10 December 2014
Available online 19 December 2014

Keywords:
Associative recognition
Processing stages
EEG
Hidden semi-Markov models

ABSTRACT

In this paper we propose a new method for identifying processing stages in human information processing. Since the 1860s scientists have used different methods to identify processing stages, usually based on reaction time (RT) differences between conditions. To overcome the limitations of RT-based methods we used hidden semi-Markov models (HSMMs) to analyze EEG data. This HSMM-EEG methodology can identify stages of processing and how they vary with experimental condition. By combining this information with the brain signatures of the identified stages one can infer their function, and deduce underlying cognitive processes. To demonstrate the method we applied it to an associative recognition task. The stage-discovery method indicated that three major processes play a role in associative recognition: a familiarity process, an associative retrieval process, and a decision process. We conclude that the new stage-discovery method can provide valuable insight into human information processing.

Introduction

One of the main goals of cognitive science is to identify processing stages in human information processing. Almost 150 years ago, Donders (1868) already proposed a method to measure the duration of cognitive stages. By subtracting the reaction times (RTs) of two tasks that were hypothesized to share all but one processing stage, the duration of that stage could be calculated. A strong and problematic assumption of Donders' method is the idea that an entire stage can be added without changing the duration of other stages. A century later, Sternberg (1969) proposed the additive-factor method to test whether stages exist in the first place (Roberts and Sternberg, 1993). Although Sternberg avoided an important limitation of Donders' method, RT-based methods are inherently limited. For instance, Sternberg's additive-factor method can only indicate the minimum number of stages in a task, it does not yield the order of discovered stages, and it cannot determine the absolute duration of the stages (Sternberg, 1969, p. 311). To get more insight in stage existence and duration, Sternberg and other have argued for the use of neuroimaging data (e.g., Coltheart, 2011; Henson, 2011; Sternberg, 2011). Following this advice, we propose a new stage-discovery method that uses hidden semi-Markov models (HSMMs; Rabiner, 1989) in combination with EEG data. This method identifies processing stages, and also yields various measures that can be used to infer the cognitive functions of the identified stages, bringing us one step closer to the ultimate goal of understanding how humans process information and perform tasks.

To demonstrate our HSMM-EEG stage-discovery method we will apply it to data of an associative recognition task (Borst et al., 2013). Over the past 50 years, two main classes of theories have been developed to explain associative recognition — our ability to judge whether items were previously experienced together. Whereas single-process theories assume a single memory process between perception and response, dual-process theories assume two qualitatively different memory processes (e.g., Diana et al., 2006; Malmberg, 2008; Wixted, 2007; Yonelinas, 2002). In this paper, we will use the HSMM-EEG method to investigate how many and what kind of processes are involved in associative recognition. In the remainder of this introduction we will first give a high-level overview of the HSMM-EEG method, followed by a description of the associative recognition study to which it was applied.

A new method to identify stages

Rather than inferring the existence of processing stages from RT distributions — as has traditionally been done (e.g., Donders, 1868; Sternberg, 1969) — the goal of the HSMM-EEG method is to find neural signatures that mark the existence of stages and their duration. For this purpose we adapted methods that have been developed to identify stages in fMRI data that last seconds or more (Anderson et al., 2010, 2012a,b; Anderson and Fincham, 2013). For instance, Anderson and Fincham (2013) investigated mathematical problem solving, and discovered four stages: encoding the problems, planning a solution strategy, solving the problems, and entering a response. These methods use

* Corresponding author at: University of Groningen, Dept. of Artificial Intelligence, Nijenborgh 9, 9747 AG Groningen, The Netherlands.
E-mail addresses: j.p.borst@rug.nl (J.P. Borst), ja+@cmu.edu (J.R. Anderson).

http://dx.doi.org/10.1016/j.neuroimage.2014.12.029
1053-8119/

semi-Markov models (Yu, 2010) that model the data as a sequence of discrete states that can have variable duration. Each state is defined by a neural signature, as well as by a gamma distribution describing the state's duration over the trials in the experiment. Although the results were promising, the temporal resolution of fMRI is severely limited by the sluggish nature of the hemodynamic response. EEG, on the other hand, has a millisecond resolution, making the HSMM-EEG analysis suitable for investigating elementary cognitive processes.

Our method consists of four basic steps. First, we fit HSMMs with different numbers of states to EEG data (we equate 'processing stages' with 'HSMM states' for now, see also the discussion). Second, we use leave-one-out-cross-validation (LOOCV) to identify the number of states by comparing the likelihoods of the fitted HSMMs. Third, we inspect the resulting HSMM: the number, order, durations, and neural signatures of the states, as well as how they vary with experimental condition. Fourth, using this information, we deduce the functions of the identified processing stages. Before we turn to a detailed description of these steps in the Methods section, we will now first describe the associative recognition study to which the method was applied.

The associative recognition study

To demonstrate the HSMM-EEG method we used an associative recognition experiment previously reported in Borst et al. (2013). The experiment consisted of two phases: a study phase and a test phase. In the study phase, subjects were asked to memorize 32 word pairs. In the subsequent test phase — during which EEG data were collected — subjects were again presented with word pairs, for which they had to indicate whether they studied them or not. The word pairs could be the same pairs as they learned previously (targets; 'yes' response), rearranged pairs (re-paired foils; 'no' response), or pairs consisting of novel words that were not presented in the study phase (new foils; 'no' response). Although remembering whether the component words were studied (item information) was sufficient to discriminate between targets and new foils, successful discrimination between targets and re-paired foils also required remembering how the words were paired during study (associative information).

Theories of associative recognition are concerned with the cognitive processes in the test phase of this experiment, when subjects make the decision whether they learned the word pairs in the study phase or not. Our goal is therefore to identify the stages — and associated cognitive processes — subjects go through in the test phase. Fig. 1 gives an overview of cognitive processes proposed by three major theories of associative recognition, two single-process theories (ACT-R and global matching) and one dual-process theory, and thus of the processing stages that we might expect to find with the HSMM-EEG method. Although all three theories assume an encoding state and a response stage, the central cognitive stages vary (for recent reviews of associative recognition theories, see Diana et al., 2006; Malmberg, 2008; Wixted, 2007; Wixted and Stretch, 2004; Yonelinas, 2002).

The first model is a computational process model developed in the ACT-R cognitive architecture (Adaptive Control of Thought — Rational; Anderson, 2007). To account for associative recognition it assumes two stages between encoding and response: an associative retrieval stage and a decision stage (e.g., Anderson and Reder, 1999; Danker et al., 2008; Schneider and Anderson, 2012; Sohn et al., 2005; it is a single-process theory in the sense that it assumes only a single memory process). In the associative retrieval stage, the ACT-R model uses the encoded representation to retrieve the best matching word pair from declarative memory. In the following decision stage the retrieved word pair is compared to the encoded word pair. If the pairs match, the model enters the response execution stage and responds "yes", if not it responds "no".

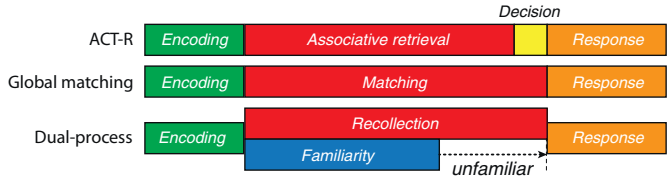

Fig. 1. Cognitive processes stages in three theories of associative recognition.

The second class of models are known as global matching or signal detection models, and assume a single matching process between encoding and response (e.g., Clark and Gronlund, 1996; Gillund and Shiffrin, 1984; Hintzman, 1988; Murdock, 1993; Wixted, 2007; Wixted and Stretch, 2004). After encoding the words on the screen, a compound cue that includes both words is compared to all relevant items in memory. The combined similarity to all items in memory (reflecting familiarity; strength-of-memory in signal detection terms; Wixted and Stretch, 2004) is compared to a certain response criterion, if it exceeds this criterion the model responds "yes", if not it responds "no".

The third class of models assumes two qualitatively distinct memory processes between encoding and response: a familiarity process and a recollection process. These models are therefore known as dual-process theories (e.g., Diana et al., 2006; Malmberg, 2008; Rugg and Curran, 2007; Yonelinas, 2002). The familiarity process is typically thought of as fast and automatic. It gives a continuous index of how familiar an item is, but no information is retrieved from memory. The recollection process is slower and yields qualitative information about a studied item, such as associative information. It is typically assumed that the familiarity process can be used to distinguish between new foils and targets/re-paired foils — directly proceeding to the response stage — while recollection is required to distinguish between targets and re-paired foils (of which the component words are equally familiar).

To facilitate the discovery and interpretation of the stages suggested by these theories, Borst et al. (2013) manipulated three factors in the associative recognition experiment, which were each intended to impact one or more cognitive processes. The first manipulation was intended to tap the encoding process: word length of both words in a pair could either be short (4 or 5 letters) or long (7 or 8 letters; word frequency was matched across all conditions). Word length has a minimal effect on reading speed (Juhasz and Rayner, 2003; Spinelli et al., 2005), but is known to affect the EEG signal first in occipital regions (around 85 ms) and later in prefrontal regions (around 300 ms; Hauk and Pulvermüller, 2004; Hauk et al., 2009; Sudre et al., 2012; Van Petten and Kutas, 1990).

The second manipulation was intended to influence an associative retrieval process. To this end Borst et al. (2013) varied the fan, or associative strength, of the words. Fan refers to the number of pairs a particular word appears in, and thus to the number of other items in memory the word is associated to. Fan is well known to have strong effects on RT and accuracy, with higher fan resulting in longer RTs and lower accuracy (Anderson, 2007; Anderson and Reder, 1999; Schneider and Anderson, 2012). In addition, fan is known to have a broad effect on the EEG signal, ranging from prefrontal to parietal sides (Heil et al., 1997; Khader et al., 2005, 2007; Nyhus and Curran, 2009). Words in the experiment could have a fan of 1 or 2, that is, they could occur in one or two word pairs (both words in a pair had the same fan). New foils had an associative fan of 1, they only appeared in a single word pair.

The third manipulation involved having the three probe types discussed above (targets, re-paired foils, and new foils). This manipulation was expected to impact multiple cognitive processes. First, by contrasting new foils with targets/re-paired foils, it should lead to differences in a familiarity process. Second, by contrasting targets with re-paired foils, it should lead to differences in an associative retrieval process. Third, a decision process would also be affected by the probe manipulation. The probe manipulation was hypothesized to affect RTs (and thus the duration of the stages) and the EEG signal. New foils — which are dissimilar from the other items — are typically responded to faster than either targets or re-paired foils (Gronlund and Ratcliff,

1989; Ratcliff and McKoon, 1989; Rotello and Heit, 2000). On the other hand, re-paired foils are typically responded to slower than targets (e.g., Anderson and Reder, 1999; Gronlund and Ratcliff, 1989; Wixted and Stretch, 2004). In addition, there are well-established EEG correlates of familiarity (new foils vs. targets/re-paired foils) and associative recollection (targets vs. re-paired foils), known as the FN400 and the parietal old/new effect, respectively (for a review, see Rugg and Curran, 2007).

In their study, Borst et al. (2013) used classical EEG analysis methods and a machine learning classifier to identify processing stages. The machine-learning classifier was used to investigate when information regarding the experimental manipulations was present over the course of a trial. Word length information was on average available between 100 and 500 ms, peaking between 200 and 250 ms. Information on the associative fan of the word pairs increased sharply at 350 ms and reached its peak between 400 and 500 ms, but remained present until the response. Finally, information about the target/foil distinction gradually increased from around 600 ms until the response. We will compare these results to the results of the HSMM-EEG approach in the general discussion.

Methods

The experiment

The general design of the experiment was discussed above; detailed methods can be found in Borst et al. (2013). To summarize: 20 subjects were asked to memorize 32 word pairs in the study phase of the experiment. In the subsequent test phase directly after the study phase, during which EEG were measured, subjects had to determine whether word pairs were previously studied pairs (targets), re-arranged pairs (re-paired foils), or new words (new foils). Three factors were manipulated: probe type, associative fan, and word length. This amounted to 10 conditions: 2 (probe: target or re-paired foil) × 2 (fan: 1 or 2) × 2 (word length: short or long) + short and long new foils. Word frequency was matched across all conditions. Subjects completed a total of 13 blocks with 80 trials per block. All 10 conditions occurred equally often in random order in each block, resulting in 104 trials per condition during the test phase. Targets and re-paired foils were each repeated 13 times (one time in each block), while new foils were never repeated. Because of these repetitions, participants might have switched strategies during the experiment: "although participants initially had to perform associative recognition by remembering whether words had been studied together, they might have eventually switched to a strategy in which they determined when during the experiment (training vs. test phase) the words had been presented together" (Borst et al., 2013, p. 2162). However, Borst et al. did not find any evidence for such a strategy shift in the EEG data. In addition, there were still clear differences in RT between fan 1 and fan 2 items at the end of the experiment. This is in line with a long history of research on associative recognition, including a study that showed that even after 25 days of repetition the effects still persist (Pirolli and Anderson, 1985).

EEG recording and preprocessing

EEG were recorded in the test phase of the experiment from 32 Ag–AgCl sintered electrodes (10–20 system) at 250 Hz. In addition, electrodes were placed on the right and left mastoids. The right mastoid served as the reference electrode, and scalp recordings were algebraically re-referenced offline to the average of the right and left mastoids. The vertical EOG was recorded as the potential between electrodes placed above and below the left eye, and the horizontal EOG was recorded as the potential between electrodes placed at the external canthi. The EEG and EOG signals were amplified by a Neuroscan bioamplification system with a bandpass of 0.1–70.0 Hz and were digitized at 250 Hz. Electrode impedances were kept below 5 kΩ.

The EEG recording was first visually inspected for artifacts; samples containing artifacts were removed from the data. The EEG recording was then decomposed into independent components using the EEGLAB infomax algorithm (Delorme and Makeig, 2004). Components associated with eye blinks were visually identified and projected out of the EEG recording. Following artifact and eye-blink removal, a 0.5–30 Hz band-pass filter was applied to attenuate high-frequency noise. Trials were extracted from the continuous recording and baseline-corrected using a linear baseline (cf. Anderson and Fincham, 2013), which was defined as the slope between the average of −200 to 0 ms before stimulus onset and the average of 80–160 ms after the response and subtracted from the data in the trial (visual inspection showed no condition differences during these intervals). We applied such a linear baseline on a trial-by-trial basis to remove random signal drift within trials. In typical ERP analyses such drift would disappear by the averaging procedure, but as the HSMM-EEG analysis uses single-trial data as input it is important to remove random variation between trials.[1] Trials containing voltages above +75 μV or below −75 μV were excluded from further analysis.

General analysis

The HSMM-EEG analysis was applied to correct trials only, with RTs under three standard deviations from the mean per condition per subject and shorter than 3000 ms. In total, 12.2% of the trials was excluded for the current analysis.

HSMM-EEG analysis[2]

An HMM simulates a system that is at any given time in one of a set of distinct states, between which it transitions at certain times (Rabiner, 1989). In our analysis, each state represents a processing stage in the task. A state i is associated with a brain signature M_i that represents the average EEG-activation pattern during this processing stage, and with gamma distributions G_{in} that represent the state's durations (and variability) for the conditions n of the experiment. Thus, while we estimated a single brain signature for all conditions in a state, we estimated separate gamma distributions for each condition, allowing for different duration estimates per condition. In this way we measure how the experimental conditions affect the duration of a stage defined by the constant brain signature. Because the duration in the states is variable, this type of HMM is referred to as a variable-duration HMM (Rabiner, 1989) or a hidden semi-Markov model (HSMM; Yu, 2010), which is the terminology we adopted. For current purposes we only considered HSMMs with linear structures, that is, state 1 always transitions to state 2, state 2 to state 3, etc.

Since the HSMM-EEG method assumes that the brain signatures are the same for each condition (conditions only differ in their duration estimates), conditions should be analyzed jointly only if they are hypothesized to consist of the same sequence of cognitive states. If not, conditions should be committed to separate analyses. For the current dataset this means that targets and re-paired foils should be analyzed together — all three theoretical accounts discussed above assume the same states for these conditions — but new foils should be analyzed separately, as they probably do not involve an associative retrieval state.

An example of a four-state HSMM is shown in Fig. 2 (for display purposes each state shows the average of the gamma distributions for the different conditions). At the top of the figure EEG data is shown for three EEG-channels over three trials of the experiment. The first step in the analysis was to create so-called 'snapshots' that spanned 80 ms. A snapshot was created by restructuring the data. From twenty 4-ms samples in 32 real channels we created one 80-ms snapshot with 640 virtual channels (each real channel appeared 20 times in a snapshot,

[1] Alternatively, as suggested by one of our reviewers, one could use a higher high-pass filter (1 Hz) or remove the DC component from the epochs.

[2] Matlab code for the HSMM-EEG analysis is available at http://www.jelmerborst.nl/models under the title of this paper.

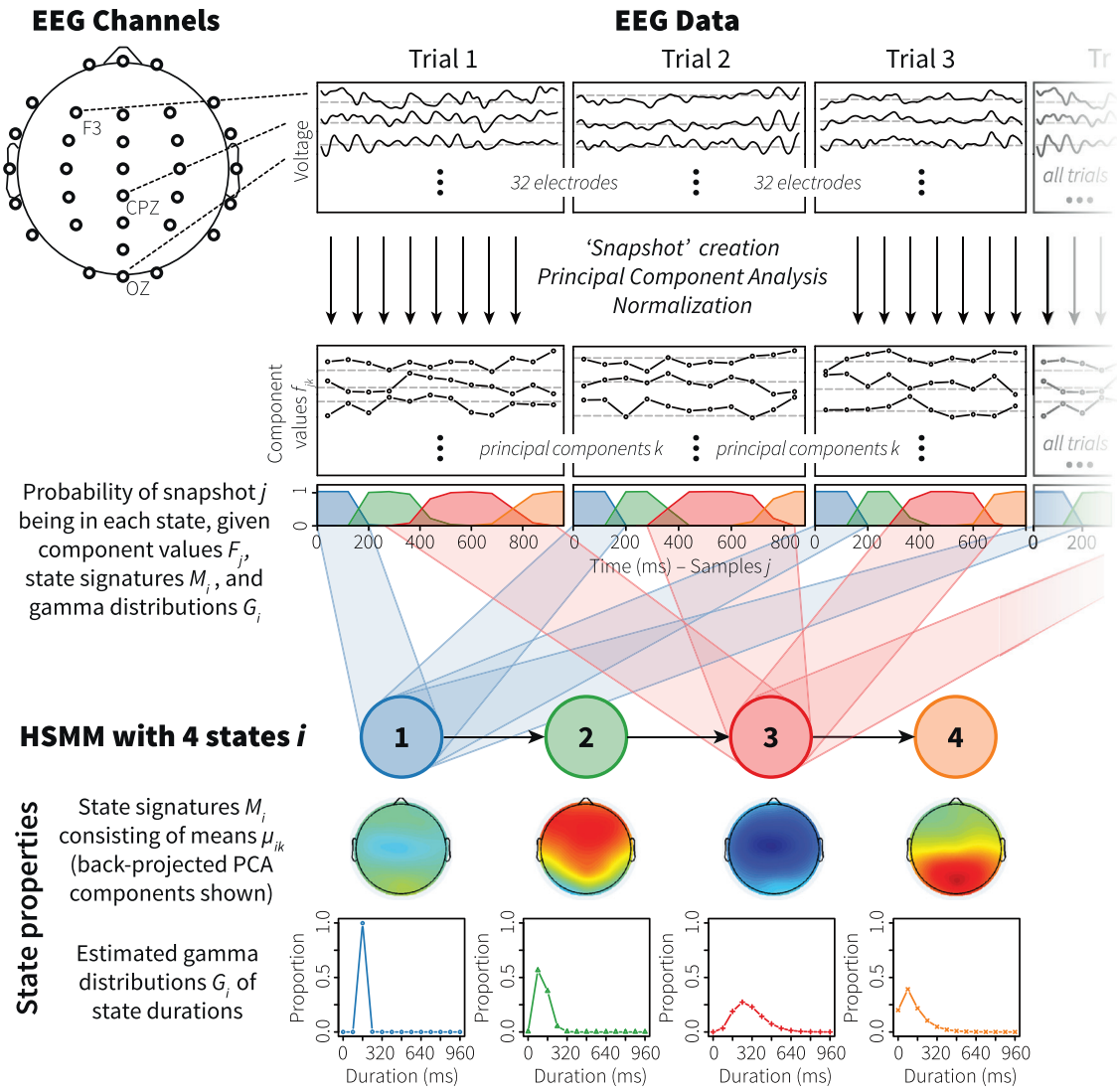

Fig. 2. Overview of the HSMM-EEG analysis. EEG data comes in at the top and is preprocessed into PCA components. At the bottom a fitted four-state HSMM is shown with state signatures and gamma distributions (averaged over the conditions of the experiment). The center graph shows the likelihood of each sample *j* being in each state. The connections between sample likelihoods and states are shown for states 1 and 3 but exist for all states.

representing the different time points in the 80 ms period). Another way of describing this is that we vectorized every 20 consecutive time points in all 32 channels (80 ms), and then concatenated those vectors to create one snapshot. Note that we did not resample the data; in Appendix A we provide a simple example of the procedure. Creating snapshots was done for two reasons: 1) presenting the temporal profile of the data to the HSMM estimation procedure, and 2) reducing the amount of data points in a trial for the estimation process (the duration of the HSMM-estimation is for a large part dependent on the number of data points in a trial, see Eq. (B5)). Our previous work with a machine-learning classifier indicated that crucial information is contained in the temporal profile of the EEG data (Borst et al., 2013). By creating snapshots, each data point not only contained information about the mean voltage in each channel, but also about whether this voltage increased or decreased over the 80 ms interval for each channel. In addition, it reduced the amount of data points for the estimation procedure, which would have taken over a month with a 4-ms time resolution.[3]

Each trial consisted of a certain number of these snapshots (we used all snapshots that fitted entirely between the stimulus and the response on each trial, i.e. snapshots that crossed into the post-response period were excluded). Note the difference with typical ERP analyses, which are either stimulus-locked or response-locked and only look at a fixed period. The next step in the analysis was to normalize each virtual channel to a mean of 0 and a standard deviation of 1, and to apply a spatial principle component analysis (PCA) to the 640 channels. The results of the PCA were again normalized. This procedure was performed across all subjects, leaving us with a single set of PCA coefficients and PCA components that matched across subjects. We used the first 100 PCA components for the analyses in this paper, which accounted for 98.9% of the total variance in the data.[4]

[3] In addition to the snapshot matrix, the algorithm was also provided with four vectors, which indicated the subject, trial, snapshot-in-trial, and experimental condition for each snapshot.

[4] Using 100 components gave representative results. We performed initial analyses with 5, 10, 20, 30, 40, 50, 60, 80, 100, 120, 140, 160, 320, 480, and 640 PCA components. The average of these analyses indicated a 6-state solution, as did the 100-component solution. Furthermore, the LOOCV-pattern of the 100-component solution was very similar to the average of all solutions. Finally, the 6-state solutions of these different analyses were all very similar: the conclusions of the paper would have remained the same independent of the number of PCA components that we used. The LOOCV-results of all these analyses are shown in Inline Supplementary Fig. S1 (cf. Fig. 4).

Inline Supplementary Fig. S1 can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.12.029.

Using HSMM algorithms, one can calculate brain signatures M_i and gamma distributions G_{in} that maximize the likelihood of the data given a particular number of states. As described above, each state duration is modeled as a gamma distribution. Because the 100 PCA factors are basically distributed as independent normals, each brain signature M_i is described by a vector of 100 means μ_{ik} that represent independent normal distributions with unit variance.[5] In Appendix B we describe this parameter estimation process in detail. The HSMM with associated signatures and gamma distributions in Fig. 2 is the result of such an optimization procedure. Given this HSMM, the likelihood that each belongs to a state is depicted in the center of the figure. As expected, the first snapshots in each of the three trials most likely belong to state 1 (blue), the next snapshots to state 2 (green), etc. In addition, in these trials state 1 is always two snapshots long, matching the gamma distribution for this state. State 3 and 4, on the other hand, are much more variable, ranging between 1 and 9 snapshots in length.

The aim of this research was to determine the number of states and their properties that provide the best characterization of the data. Because HSMMs with more states have more parameters to fit the data, they will typically yield a better fit. To test if the extra parameters explain sufficient extra variance to be warranted, we applied leave-one-out cross validation. Our LOOCV-method estimated maximum-likelihood parameters for an r-state HSMM using the data of all but one of the 20 subjects. We then split the data of the remaining 20th subject in half, and estimated maximum-likelihood gamma distributions for the first half of the 20th subject's data given the brain signatures of the other subjects. Using these gamma distributions and the state signatures of the other subjects, the log-likelihood of the second half of the 20th subject's data was calculated. We repeated this for the other half of the 20th subject's data, and averaged the two log-likelihoods (in effect keeping the likelihood estimation independent of fitting the model while accommodating for speed differences between subjects). This process was repeated for all subjects and for HSMMs with different numbers of states.

To select the HSMM with the optimal number of states we used a sign-test: the highest k-state model that fitted the data of a significant number of subjects better than all $(l < k)$-state models was chosen. Even if the true number of states is k, a $(k + 1)$-state model will fit the data of $n - 1$ subjects better in the estimation phase because it has more parameters. However, it is at least as likely to fit the nth subject worse (Anderson and Fincham, 2013). A sign test provides an upper bound on the probability that the $(k + 1)$-state model is better than the k-state model by chance. With 20 subjects, a significant increase is reached when at least 15 subjects improve ($p = .04$). Note that this LOOCV-procedure finds evidence for a certain number of states by showing it performs better than fewer states. However, this does not mean that solutions with fewer states are not to be trusted — they may just be at a different level of aggregation.

After determining the optimal number of states we computed the properties of the identified states: effects of experimental condition on stage duration, brain signatures, and differences in brain activity between conditions. To this end, we first estimated a single HSMM using the data of all subjects. We used the state signatures of this model to estimate gamma distributions for each subject and condition. These gamma distributions were used to calculate the average state duration for each subject and condition, which were used in subsequent ANOVAs to determine which states changed in duration with condition. In addition, the subject-specific models gave us the probability for each snapshot to be in a certain state (center of Fig. 2). This was used to estimate brain activity for each state, condition, and subject; we multiplied the EEG data of each sample with its assigned probability, summed over the resulting values, and divided by the summed probability over all samples. This created an 'EEG pattern' for each subject, condition, and state. Note that these patterns are simply a weighted average of the EEG signal. To locate differences between conditions we then averaged over subjects and subtracted the signatures of the different conditions (even though the HSMM-EEG analysis ensures that the state signatures are similar on a global level compared to the other states, there might be remaining differences between conditions). In addition, we subjected the resulting values to t-tests for each electrode, p-values were corrected for multiple comparisons by applying the False Discovery Rate (FDR; Genovese et al., 2002).

Results

Behavior

Behavioral results of the experiment were reported in its primary publication (Borst et al., 2013). However, because new foils were excluded in that paper and we applied a stricter outlier criterion for the HSMM-EEG analysis, we now report a separate analysis of the RT-data used in this paper. Fig. 3 shows the results. RTs ranged between 414 and 3000 ms after the outlier rejection. A repeated-measures ANOVA on the data of the targets/re-paired foils indicated significant main effects of Probe ($F(1,19) = 69.7$, $p < .001$, $\eta_p^2 = .79$) and Fan ($F(1,19) = 168.3$, $p < .001$, $\eta_p^2 = .90$), as well as an interaction between Probe and Fan ($F(1,19) = 28.8$, $p < .001$, $\eta_p^2 = .60$). Thus, RTs were longer for fan 2 items than for fan 1 items and for re-paired foils than for targets. Additionally, the effect of fan was larger for re-paired foils than for targets. No other effects were significant for targets/re-paired foils. New foils were responded to faster than targets/re-paired foils. In addition, long words resulted in longer RTs than short words ($t(19) = -4.72$, $p < .001$) for the new foils. These results are in agreement with the literature, making it a suitable task for demonstrating the HSMM-EEG analysis.

HSMM-EEG analysis

To get more insight into the cognitive processes involved in this task, we applied the HSMM-EEG analysis. Because new foils differed from the other probe types — no associative information can be retrieved for new

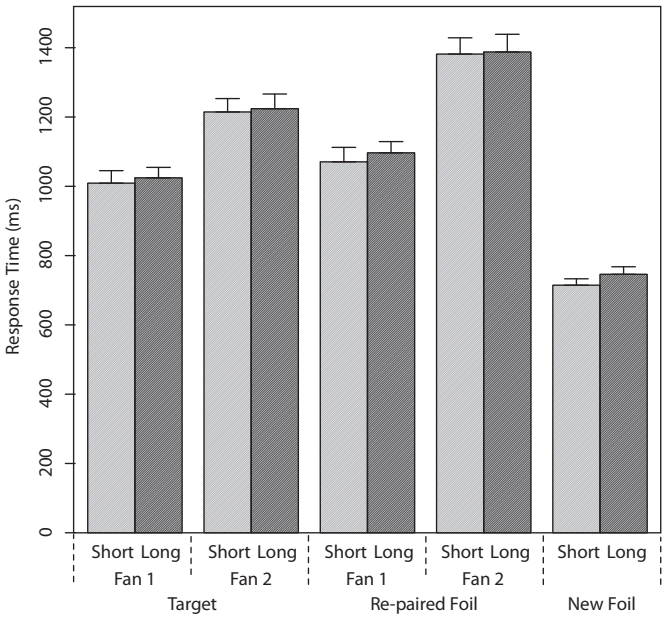

Fig. 3. Response times. Error bars indicate standard errors. Short/long indicate word length.

[5] The PCA factors are orthogonal, but not perfectly independent. However, it seems a close enough approximation.

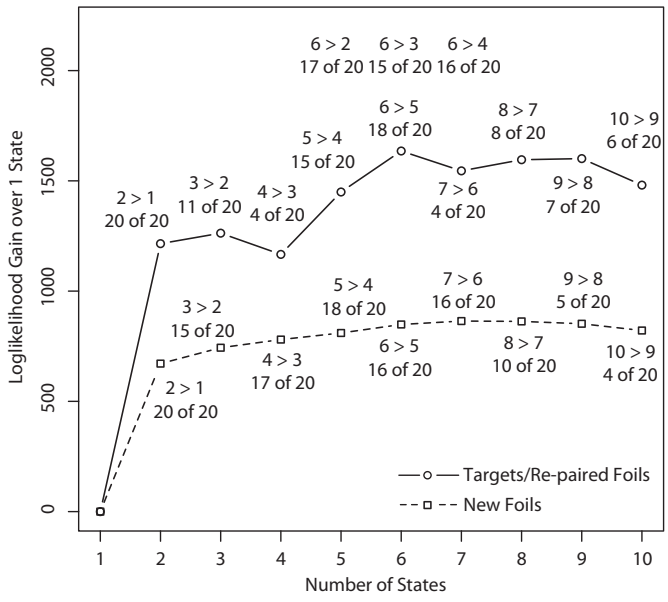

Fig. 4. Results of the LOOCV-procedure: the average gain in log-likelihood of an *r*-state HSMM over a 1-state HSMM. The figure additionally shows for how many subjects the log-likelihood increased from state to state: for instance '3 > 2, 11 of 20' means that for 11 subjects the log-likelihood of a 3-state HSMM was higher than the log-likelihood of a 2-state model. According to a sign-test with 20 subjects, a significant increase is reached when 15 out of 20 subjects improve ($p = .04$), whereas 18 out of 20 corresponds to a *p*-value of .0004.

foils and they were not present during the training phase — we performed independent analyses for targets/re-paired foils and new foils. For targets/re-paired foils a 6-state HSMM was identified as providing the best account of the data. Fig. 4 shows the results of the LOOCV procedure in detail: it depicts the average gain in log-likelihood of an *r*-state HSMM over a 1-state HSMM, as well as the number of subjects that improved from state to state. Note that the average log-likelihood can decrease with more states because the likelihood estimation was kept independent from fitting the models (see the section on HSMM-EEG analysis for details). The 6-state HSMM for targets/re-paired foils was better for at least 15 subjects than HSMMs with fewer states, which corresponds to a significance level of $p = .04$, and HSMMs with more states had a lower log-likelihood.

For the new foils, a 5-state model dominated models with fewer states, but the winning model was a 7-state HSMM (it was better for at least 16 subjects than HSMMs with fewer states; see also Fig. 4). However, the 5-state solution corresponded better to the 6-state solution of targets/re-paired foils. The 6- and 7-state solutions for the new foils seemed to give a more detailed account of the second and fifth state of the HSMM for targets/re-paired foils (see Fig. C1). While it might be possible to sub-divide such states into smaller states, we can be secure in the assumption that there are at least 6 states for targets/re-paired foils and 5 states for new foils, and in whatever conclusions these assumptions lead to. We will therefore focus on the 5-state solution for new foils. The 6- and 7-state solutions are reported for comparison in Appendix C.

The properties of the identified stages are shown in Figs. 5–7. Fig. 5 shows brain signatures,[6] gamma distributions averaged point-wise over conditions and subjects, RT-distributions and convolved gamma distributions. Fig. 6 shows state durations adding up to entire trials, in effect showing overall RTs, and state durations by condition, highlighting the effect of condition on states. Finally, Fig. 7 shows the voltage differences between conditions within states — e.g., we subtracted the 'short word signatures' from the 'long words signatures' for the first row of Fig. 7 (see the last paragraph of the section on HSMM-EEG analysis for details). Table 1 reports effects of Fan, Probe, and Word Length on state durations of the targets/re-paired foils (because there are 6 states we consider results significant when the *p*-values were below a Bonferroni-corrected threshold of .05/6 = .008), Table 2 lists effects of Word Length on the state durations of the new foils (as indicated by *t*-tests). In addition, Table 3 lists the correlations between the state signatures of the HSMMs for targets/re-paired foils and new foils, and Table 4 list the correlations between single-trial RTs and state durations.

Comparing the signatures of the new foils to the signatures of the target/re-paired foils (Fig. 5, Table 3) suggests that the first three states of the targets/re-paired foils correspond to the first three states of the new foils. In addition, the last two states of both solutions are similar. We therefore depicted the solution of the new foils as having states 1–3 and 5–6, while state 4 is missing. In addition, we compared the EEG activity in the last and the prior-to-last state of the new foils to the last and prior-to-last state of the targets/re-paired foils (Fig. 7).

To check whether the targets/re-paired and the new foils indeed go through the same stages, we additionally performed a joint analysis in which we fitted a single HSMM to all three conditions. This joint analysis confirmed the results of the independent analyses: a 6-state HSMM accounted best for the data (Figs. D1–D3), and resulted in the same states as the original 6-state solution of the targets/re-paired foils and the 5-state solution of the new foils (cf. Fig. 5). Although the joint solution indicated that there is some evidence for state 4 in the new foils, this state was very brief (Fig. D3), and was skipped in about 67% of the trials. As shown in the separate analysis of the new foils above, there was not sufficient evidence for this state when only taking the data of the new foils into account. Therefore, our match between the 6-state solution of the targets/re-paired foils and the 5-state solution of the new foils seems reasonable. We report the results of the joint analysis in Appendix D.

Functional interpretation of the processing stages

The reason for wanting to identify processing stages is to gain insight in how tasks are performed, in this case associative recognition. In this section we present our interpretation of the discovered processing stages.

The first and second stage of the HSMM seem to reflect encoding the words on the screen. Not only is that the first thing subjects have to do, but these stages also had very similar brain signatures for targets/re-paired foils and new foils (they both correlated with .98, see Table 3) and did not vary in duration with experimental condition, except for a three-way interaction between Probe type, Fan and Word Length on the duration of the second stage (Fig. 6 and Table 1), which has no relationship to trial length (Table 4). The lack of effects in the first two stages suggests that nothing is processed yet in relation to the associative information, which differed strongly between conditions. The idea that stage 2 is involved in word encoding — in contrast to encoding both words in stage 1 — is supported by two observations: (1) there was no main effect of fan on stage 2, which might be expected as soon as both words are encoded, and (2) this stage was not shorter for the new foils than for the targets/re-paired foils ($t(19) < 1$), which might be expected if stage 2 would be a pure familiarity stage. On the other hand, we do not believe that each stage represents encoding a single word, as they have very different brain signatures, and as single word recognition results in similar ERP curves as we observed in these stages (e.g., Hauk and Pulvermüller, 2004; Hauk et al., 2006). However, there was a substantial difference in EEG signal between the new foils and the targets/re-paired foils in stage 2 (Fig. 7, bottom two rows), and also a smaller difference between new foils and re-paired foils in

[6] These signatures were created by multiplying the unprocessed EEG samples (no PCA, no snapshots) with their assigned probability for a certain state, summing over the resulting values, and dividing by the summed probability over all samples (i.e., we took a weighted average). One could also base signatures directly on the PCA components; the results are virtually indistinguishable.

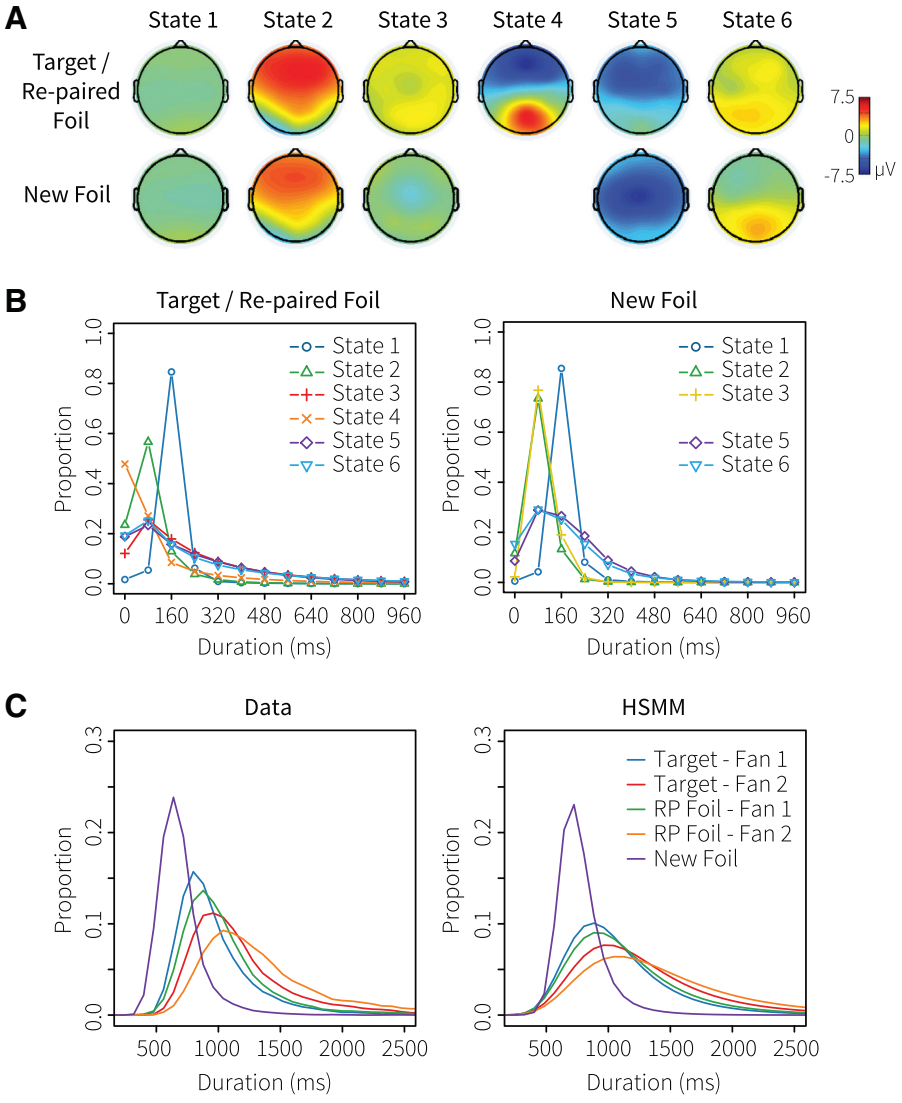

Fig. 5. State signatures (A), gamma distributions (B), distribution of RTs in the data (C—left), and convolved gamma distributions (C—right). For panel C we collapsed over short and long words; the data were smoothed with a 3-point running mean.

stage 1. This suggests that familiarity of the words begins to register in stage 2 and this is what distinguished it from stage 1.

We hypothesize that stage 3 reflects a memory retrieval stage. Stage 3 is by far the shortest for the new foils: as soon as it becomes clear that the encoded words did not match any of the studied words in memory they could be rejected (Gronlund and Ratcliff, 1989; Rotello and Heit, 2000). Second, the duration of stage 3 varied strongly with fan and probe (Table 3), and was a strong predictor of total RT (Table 4), which is known to be sensitive to memory retrieval processes. Stage 3 was longer for fan 2 items than for fan 1 items, and longer for re-paired foils than for targets, which is typical for fan data (Anderson, 2007; Anderson and Reder, 1999; Danker et al., 2008; Schneider and Anderson, 2012; Sohn et al., 2005). Existing models attribute these effects to declarative memory processes, implying that this stage involved retrieval of associative information (see Anderson and Reder, 1999; Diana et al., 2006; Schneider and Anderson, 2012, for detailed accounts of these effects). In addition, fan affected the EEG signal in this stage (Fig. 7), also suggesting that associative information is processed.

Stage 4 does not exist for new foils, and was skipped on average in 48% of the trials for targets/re-paired foils (Fig. 5B, left; there was no difference in the skipping percentage between targets and re-paired foils, $t(19) = 1.09$, $p = .29$). There seemed to be two groups of subjects with respect to skipping this stage: 12 subjects skipped it often with an average of 70%, and 8 subjects skipped it in relatively few trials, with an average of 22%. However, there was no performance difference between these groups, neither with respect to RT ($F < 1$) or accuracy ($F < 1$). Finally, this stage hardly varied with condition or total RT (Fig. 6 and Table 4). We interpret this stage as working memory consolidation of the retrieved associative information from stage 3. Nothing was retrieved for the new foils — so nothing can be consolidated — and working memory consolidation is not required, but possible in this task. This might explain why the stage only occurred for a subset of our participants; it has been argued before that some people automatically consolidate information in working memory, while others apply a more reactive strategy and only consolidate information when required by the task (e.g., Colzato et al., 2008; Taatgen et al., 2009).

Stage 5 is again a shared stage between the new foils and the targets/re-paired foils. For targets/re-paired foils, the duration of this stage increased with fan and probe, with fan 1 targets being as fast as the new foils ($t(19) < 1$). In addition, in this stage old items showed a stronger positive-going wave over parietal sides than new items (Fig. 7, rows 3–4; similar to the parietal old/new effect in conventional ERP analyses, e.g., Rugg and Curran, 2007). Although the conventional parietal old/new effect is typically though to be an indicator of recollection of associative information (for an alternative decision-interpretation, see Finnigan et al., 2002), we do not think that stage 5 reflects recollection

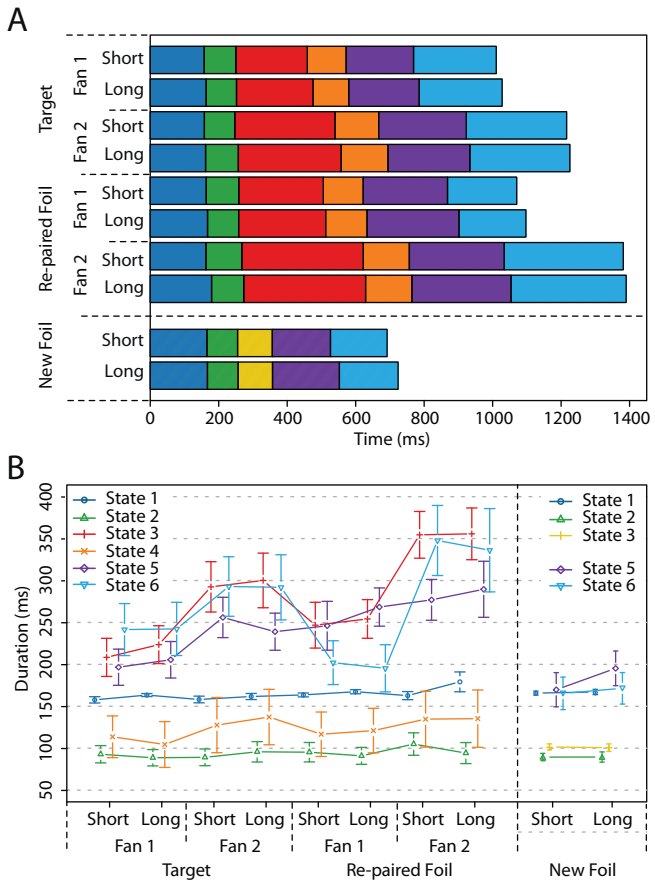

Fig. 6. State durations: A shows how the state durations add up to form complete trials; B shows more clearly how state durations are affected by condition. Error bars indicate standard errors.

because it also occurred for the new foils, for which no associative information exists. Although this stage could reflect a failed recollection process for the new foils — with the same duration as a successful recollection process for fan 1 targets — this seems unlikely in combination with the short stage 3 for new foils, which we hypothesized to reflect the difference in memory retrieval between new foils and targets/re-paired foils (see below for additional discussion of this issue). Instead, we hypothesize that stage 5 is a decision stage, in which the recollected information from stage 3 is used in combination with the encoded words to determine what the subject should respond. If this were conceptualized as an evidence accumulation process, unfamiliar items (new foils) and target-fan-1 retrievals (no alternatives in memory) would yield most evidence, and thus lead to the fastest responses. Fan 2 items would yield less evidence, as the encoded words are also connected to other relevant facts in memory and therefore spread less activation to the retrieved pair, slowing down the decision process. Following the idea of a recall-to-reject process (e.g., Anderson and Reder, 1999; Rotello and Heit, 2000), for re-paired foils a retrieval would be made in stage 3, but the retrieved information would be incongruent with the encoded words, slowing down evidence accumulation and thus RTs.

Stage 6 is the final stage in the task. We therefore assume that it is the response stage. Its duration varied with fan, and showed an interaction

Table 1
Effects on state durations for targets/re-paired foils. We consider results significant if their p-value was below a Bonferroni-corrected threshold of .05/6 = .008.

Source	$F(1,19)$	p	η_p^2	$F(1,19)$	p	η_p^2
	State 1			State 4		
Fan	<1	–	–	7.46	.01	.28
Probe	5.12	.03	.22	1.80	.20	.09
Word length	4.56	.05	.19	<1	–	–
Fan × Probe	1.09	.31	.05	<1	–	–
Fan × Word length	<1	–	–	1.34	.26	.07
Probe × Word length	<1	–	–	<1	–	–
Fan × Probe × Word length	<1	–	–	<1	–	–
	State 2			State 5		
Fan	1.31	.27	.06	**14.83**	**.001**	**.44**
Probe	1.74	.20	.08	**35.97**	**<.001**	**.65**
Word length	1.01	.33	.05	<1	–	–
Fan × Probe	2.08	.17	.10	2.86	.11	.13
Fan × Word length	<1	–	–	1.18	.29	.06
Probe × Word length	2.15	.16	.10	2.85	.11	.13
Fan × Probe × Word length	**9.10**	**.007**	**.32**	<1	–	–
	State 3			State 6		
Fan	**57.18**	**<.001**	**.75**	**33.74**	**<.001**	**.64**
Probe	**22.41**	**<.001**	**.54**	<1	–	–
Word length	<1	–	–	<1	–	–
Fan × Probe	2.00	.17	.10	**25.10**	**<.001**	**.57**
Fan × Word length	<1	–	–	<1	–	–
Probe × Word length	<1	–	–	<1	–	–
Fan × Probe × Word length	<1	–	–	<1	–	–

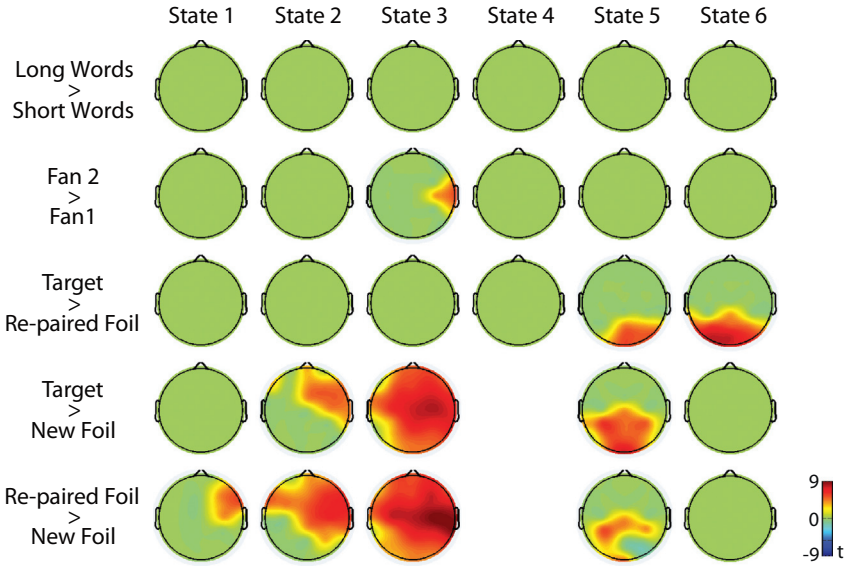

Fig. 7. Differences between conditions in states. The maps show *t*-values for FDR-corrected *p*-values <.01.

Table 2
Word length effects on state durations for new foils.

State	$t(19)$	p
1	<1	–
2	<1	–
3	<1	–
5	−3.58	.002
6	<1	–

Table 4
Correlations between single trial RTs and state durations.

State	Correlation
1	.06
2	.00
3	.58
4	.08
5	.32
6	.53

between fan and probe. For new foils this last stage was shorter than for the other conditions. We interpreted these duration differences as an effect of response confidence. Subjects responded faster and more accurately to new foils than to targets/re-paired foils, and faster and more accurately to fan 1 items than to fan 2 items — indicating they might have been more confident in those responses (e.g., Ratcliff and Murdock, 1976; Wixted and Stretch, 2004). The only effect on the EEG signal — a positive parietal effect when comparing targets to re-paired foils — supports this confidence interpretation. A similar conventional ERP effect, the P300, is known to vary with response confidence, with higher confidence leading to more positive P300 responses (e.g., Nieuwenhuis et al., 2005; Sutton et al., 1982; Wilkinson and Seales, 1978). However, the conventional P300 occurs 300–400 ms post-stimulus. Although we did not observe a classic P300, the signatures in stage 6 might be due to a similar effect, perhaps in response to associative retrieval in stage 3, which comes 300–500 ms before stage 6.

Discussion

In this paper we introduced the HSMM-EEG stage-discovery method, and demonstrated it by applying it to associative recognition. We will now discuss this new method in more detail, followed by potential implications of our current analysis for theories of associative recognition.

HSMM-EEG stage discovery

The main goal of this paper was to demonstrate the HSMM-EEG analysis by applying it to associative recognition. A similar analysis has previously been applied to fMRI data of mathematical problem solving (Anderson and Fincham, 2013). It successfully identified four stages in their task, leading to new insights into how subjects solved the problems. However, the temporal resolution of fMRI is severely limited, both by having scan durations in the order of 2 s, as well as by the sluggishness of the hemodynamic response. To investigate associative recognition — with RTs under 2 s — we turned to EEG.

The analysis resulted in a detailed account of the current experiment. The first thing to note is that the identified gamma distributions for the different conditions in the experiment resulted in the same RTs as we reported for the behavioral data (compare Fig. 6A to Fig. 3). In addition, convolving the gamma distributions within each condition (Fig. 5C, right) showed that the gamma distributions also captured most of the variability in the data (Fig. 5C, left). The reality of the discovered stages is supported by the fact that the stages and the order of duration effects in our dataset seem sensible. For instance, although it might seem obvious that the first stage has a fixed duration and reflects perceptual processing, it could as well have varied in duration with fan or probe type, making the perceptual interpretation problematic. Alternatively, all stages could have varied in duration with overall RT, making the analysis much less informative. In addition, the independent analysis of the new foils resulted in similar perceptual and response stages as the analysis of the targets/re-paired foils, but in different central memory stages. This provides a simple form of cross-validation of the HSMM-EEG method: the perceptual and response stages should be similar across all conditions, and the central stages should be different. Given that the HSMM part of the analysis has been validated before (Anderson and Fincham, 2013) and that the current results are in line with the existing theoretical accounts discussed in the introduction (Fig. 1), we argue that the HSMM-EEG analysis provides an interesting new method to analyze data with a high temporal resolution. However, the current analysis should be seen as a proof-of-concept of the HSMM-EEG method, rather than as a general assessment.

Table 3
Correlations between the state signatures M_i (PCA components) of the 6-state HSMM for targets/re-paired foils (horizontal axis) and the 5-state HSMM for new foils (vertical axis). Bold values indicate the highest correlation in each row.

State	1	2	3	4	5	6
1	**0.98**	−0.31	−0.04	−0.27	−0.08	−0.12
2	−0.28	**0.98**	0.25	−0.17	−0.19	−0.29
3	−0.23	0.06	**0.51**	−0.10	−0.04	−0.29
5	−0.01	−0.28	−0.20	0.02	**0.40**	−0.24
6	−0.22	−0.28	−0.20	0.01	−0.13	**0.64**

The dataset that we used was previously analyzed with standard EEG methods and a machine-learning classifier (Borst et al., 2013). The classifier indicated when certain information was present in the EEG signal. As described in the introduction, word length information was available between 100 and 500 ms (peaking between 200 and 250 ms), associative fan information from 350 ms until the response (peaking between 400 and 500 ms), and information about the target/foil distinction from 600 ms until the response (gradually increasing). Providing another form of cross-validation, these measures are in agreement with the discovered stages and our interpretation of them. Word length should become available during the encoding stages (0–250 ms), which fits with the classifier peak between 200 and 250 ms. If stage 3 (250–600 ms) indeed reflects associative retrieval, one would expect information about associative fan to become gradually available during this stage, and to be very clear when stage 3 finishes for low-fan items while it still continues for high-fan items (different processes should make it very easy for the classifier to distinguish between high and low-fan items). Information on fan became available from 350 ms onwards and reached its peak around 450–500 ms, which coincides with the end of stage 3 for fan 1 items. Finally, the target/foil distinction should match the decision and response stages (from 600 ms onwards). Indeed, the classifier indicated that his information was available from 600 ms and became stronger towards the response.

The basic idea behind the HSMM-EEG analysis is to find similar brain signatures for the different conditions in the experiment. This approach contrasts sharply with most ERP analyses (e.g., Luck, 2005), which aim at finding differences between conditions. The assumption of the HSMM-EEG method is that differences between conditions within stages are relatively small (except in duration) as compared to differences between stages. This is consistent with typical ERP data, in which the main effects on voltage are similar between conditions, but in which there are small differences in amplitude between conditions. That being said, one could imagine estimating different state signatures for each condition and state. While this would vastly increase the parameter space of the HSMMs, it might provide additional insight in how a task is performed.

Because our main interest was associative recognition, and since it matched well to the 6-state solution for targets/re-paired foils, we

reported a 5-state solution for new foils. However, the LOOCV-procedure suggested a 7-state solution for new foils (reported in Appendix C). Although the 7-state solution for new foils might provide a more detailed account of the new foils — it effectively sub-divided two states of the 5-state solution — it matched less well to the solution of the targets/re-paired foils. This highlights a difficulty of the analysis: depending on the noise level in the data one can discover fewer or more states. In addition, we know that EEG is not sensitive to everything that happens in the brain, and we can therefore not discover stages that are not measured by EEG. For instance, while we now identified two perceptual stages, we might as well have discovered a single longer perceptual stage (if there was more noise), and with better neuroimaging techniques it might be possible to discover sub-stages of perceptual processing. All those solutions provide insight into how a task is performed, but differ in their level of detail. The LOOCV-procedure indicates for how many states there is evidence in the data. However, in some cases using fewer states may lead to a better understanding of the data. This seemed to be the case with the 5-state solution for new foils in the current paper, which collapsed pairs of states from the 7-state solution. Although we used states and stages interchangeably throughout this paper, it might therefore be that a single psychological stage consists of multiple HSMM states, or vice versa. For example, in the current analysis HSMM states 1 and 2 both reflect encoding, but in state 2 this encoding is mixed with a familiarity process. The psychological encoding stage is therefore a combination of HSMM states 1 and 2.

A related issue is how to choose the right snapshot length. In the current paper we used 80-ms snapshots, mostly because of the computational advantages compared to shorter snapshots (see Appendix B). However, with the HSMM-EEG analysis it is possible to identify states that are shorter than the snapshot length. For each snapshot, the analysis identifies the probability that it belongs to each state with Eqs. (B1)–(B4) — these probabilities sum to one. Note that it is possible that the summed probability of a state in a trial is less than 1. For instance, if a trial consists of four snapshots, and the probability for state 1 is 0.1 in each snapshot, its summed probability is 0.4 over the trial. This results in a predicted duration of $0.4 * 80\ \text{ms} = 32\ \text{ms}$. Given that the average of this summed probability can be less than 1 over all trials, a state's duration can also be less than the snapshot length on average. For example, in the joined analysis in Appendix D, state 4 of the New Foils is 50 ms long. Although we used 80-ms snapshots in the current analysis, we also explored some solutions with 40-ms snapshots. Interestingly, with 40-ms snapshots the identified states did not split into shorter substates.

Perhaps the analysis method that is most closely related to the HSMM-EEG method is the ERP microstate segmentation method (e.g., Brunet et al., 2011; Pascual-Marqui et al., 1995; see for a related approach Britz and Michel, 2011; Lehmann et al., 1998; Van De Ville et al., 2010). The goal of the microstate analysis is to divide ERPs into quasi-stable microstates that last ~80–120 ms. Each microstate is defined by a fixed EEG scalp topography — what we have termed brain signature — but the amplitude of this topography is allowed to vary. The different topographies, and therefore the optimal number of different microstates, are typically based on group-averaged data, after which each time point in the original data is labeled with a certain microstate. Thus, each trial will contain multiple microstates, and the same state can repeat several times. For instance, Pascual-Marqui et al. (1995) divided the trials in one condition of their experiment into the following microstates: C (86 ms), E (112 ms), B (20 ms), A (116 ms), C (68 ms), and D (146 ms).

Although the goals of the microstate analysis and the HSMM-EEG analysis are similar — dividing each trial into states — there are several clear differences between the methods. First, the microstate analysis allows state signatures to vary in amplitude, while the HSMM-EEG method uses fixed signatures. Second, the microstate analysis focuses on the total number of different states over trials, while the HSMM-EEG analysis tries to determine the number of states within each trial. This is related to the third and perhaps most important difference. The microstate analysis typically determines the number of states and their properties on averaged data (Brunet et al., 2011), and then applies the results to all data. In contrast, the HSMM-EEG takes all data into account during the estimation of the states, and then aggregates the results. This is important because averaging, especially of longer trials, can obscure important properties of the signal and might lead to fewer identified states (particularly further away from fixed time points).

Implications for theories of associative recognition

In this section we will discuss possible implications of the current analysis for theories of associative recognition. Fig. 8 shows the cognitive processes that occur in associative recognition as indicated by the HSMM-EEG analysis. The filled boxes indicate theoretical processes; the dashed lines and the numbers between them indicate HSMM states of the targets/re-paired foils. We will refer to this model as the HSMM-EEG model. It starts with an encoding stage to read the word pair on the screen. As soon as some information is encoded, the model starts a familiarity process for the first word. In case the words on the screen are unfamiliar (new foils), the model directly proceeds to the decision stage (note that encoding a single word is sufficient for this decision). Otherwise, the model starts an associative retrieval process after encoding the words. The duration of this retrieval stage is sensitive to the fan of the word pair, as well as to its probe status, with fan 2 items and re-paired foils leading to longer durations (the model uses a recall-to-reject strategy; Rotello and Heit, 2000). Following the associative retrieval process — or the familiarity process in case of new foils — a decision stage is initiated. This stage is sensitive to the information collected in the previous stages, with shorter durations for unambiguous results (unfamiliar items, fan 1 targets) and longer durations for fan 2 items and re-paired foils, as explained above. After the decision stage the model enters a response stage, which duration depends on response confidence, with more confident responses leading to shorter durations.

None of the models discussed in the introduction match this pattern exactly, although elements of all of them are part of the new model. On the basis of a wide range of behavioral and neuroimaging data, global matching models and dual-process models assumed a familiarity process (e.g., Clark and Gronlund, 1996; Diana et al., 2006; Gronlund and Ratcliff, 1989; Rotello and Heit, 2000; Yonelinas, 2002). The HSMM-EEG analysis confirmed the existence of such a familiarity process, by showing differences in EEG signal between new foils and targets/re-paired foils in the first two stages of the HSMMs. As we do not believe associative information to be available during these stages — there are

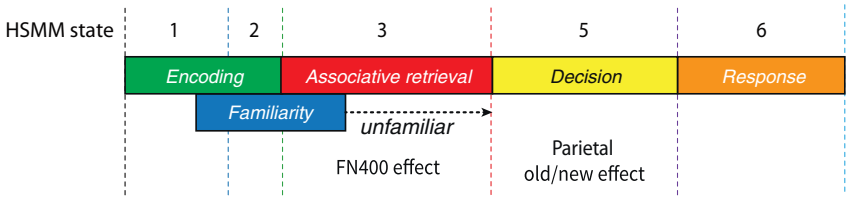

Fig. 8. Cognitive processes in associative recognition based on the HSMM-EEG analysis and associated EEG effects. The numbers indicate corresponding states in the HSMM-EEG analysis of the targets/re-paired foils. Note that state 4 is left out.

no effects of fan or probe type — this implies a separate process. In addition, an item familiarity process elegantly explains the fast rejections of new foils (Clark and Gronlund, 1996; Gronlund and Ratcliff, 1989; Rotello and Heit, 2000). Both the ACT-R model and the dual-process models assume an associative retrieval process, which was confirmed by the analysis. In addition, the ACT-R model gives a detailed explanation of the duration effects we observed in stage 3 (the dual-process SAC model assumes similar mechanisms; Diana et al., 2006; Reder et al., 2000). Finally, although the ACT-R model includes a decision stage, this stage is of a different nature than the decision stage proposed by the HSMM-EEG model. Whereas the decision stage in ACT-R is a fixed 50 ms, the decision stage in the HSMM-EEG model varies in duration based on the available evidence, and was at least 200 ms long in the current task.

There are several elements of this interpretation that warrant discussion. First, one might wonder whether the associative retrieval process starts in parallel with the familiarity process (as soon as the words are encoded), as depicted in Fig. 8. There are at least two alternatives. In the first alternative, the familiarity process finishes at the end of stage 2, and is followed in all conditions by a recollection process, which fails quickly for the new foils. Thus, in that interpretation stage 3 is pure recollection. We think this is unlikely, because there seems to be no reason to start a recollection process for new foils after a familiarity process already indicated that the encoded words are new. In the second alternative, the familiarity process extends to the end of stage 3 of the new foils, and is only followed by a recollection process for targets/re-paired foils. Thus, in that case the first 100 ms of stage 3 reflect pure familiarity. However, in that case we would have expected to find evidence for two different stages. Thus, for now we assume that these processes operate in parallel, and that a recollection process starts as soon as the words are encoded (in all conditions, also for the new foils). In case the familiarity process indicates that the encoded words are new — at the end of stage 3 of the new foils — the model proceeds directly to the decision stage without waiting for the results of the recollection process.

The second point of discussion is the absence of stage 4 of the HSMM in the final model (Fig. 8). We did not include this stage because we are not sure what it entails, and because it only occurs in about half the trials. However, including it in the model would not change much: there would still be three major processes between encoding and response.

Third, one could interpret the results of the HSMM-EEG analysis in line with dual-process theories, by assuming that stage 3 of the HSMM reflects a familiarity process and stage 5 a recollection process. However, we believe this interpretation to be incorrect, because we observed duration effects on stage 3 of fan and of probe type. The difference between targets and re-paired foils in this stage cannot be explained by the familiarity of the individual words in the pairs, as they were equally familiar. However, it could be explained by the target pairs being more familiar than re-paired foil pairs. Thus, in that interpretation more familiar pairs — targets — lead to a shorter familiarity stage than the less familiar re-paired foils (note that this interpretation goes against the general agreement in the literature that familiarity does not contribute to associative recognition, e.g., Diana et al., 2006; Rotello and Heit, 2000; Yonelinas, 2002). However, the effect of fan goes in the opposite direction. Fan 2 pairs and words were presented significantly more often during training than fan 1 pairs and words (Borst et al., 2013), which makes fan 2 items more familiar than fan 1 items. If stage 3 reflects a familiarity process, it should be shorter for fan 2 pairs than for fan 1 pairs, not longer. On the other hand, the duration effects on stage 3 are elegantly explained by both the ACT-R and the SAC processes of recollection (e.g., Anderson and Reder, 1999; Diana et al., 2006; Reder et al., 2000; Schneider and Anderson, 2012). Taken together, this seems to imply that stage 3 involves recollection, not familiarity.

Interestingly, if this idea were to be confirmed by future studies, it would indicate that the ERP signatures related to familiarity and recollection in dual-process theories, the FN400 and the parietal old/new effect, respectively, should be related to recollection and decision processes instead. The FN400 is thought to occur between 300 and 500 ms over mid-frontal electrodes. This time course fits stage 3 of the HSMM. The parietal old/new effect is assumed to occur between 500 and 800 ms, roughly matching the decision stage. Although it goes against the typical interpretation of dual-process theories, the notion that the parietal old/new effect indexes a decision process, instead of recollection, has been suggested before (Finnigan et al., 2002). Furthermore, the idea that the FN400 indexes recollection instead of familiarity would explain a puzzling observation in the literature. Speer and Curran (2007) showed that the FN400 was sensitive to associative information, while dual-process theories typically assume that the familiarity process does not aid associative recognition (e.g., Diana et al., 2006; Yonelinas, 2002). In addition, Nyhus and Curran (2009) have shown that associative fan modulates both the FN400 and the parietal old/new effect — again indicating that associative information is processed between 300 and 500 ms, which is inconsistent with a pure familiarity account. However, these observations are consistent with the HSMM-EEG model in which the FN400 indicates recollection.

Conclusion

Compared to classical stage-discovery methods, such as Sternberg's additive factor method (1969), it seems to us that the addition of neuroimaging data — in this case EEG, but one could alternatively use fMRI, MEG, or eye movements — is very valuable (see also Coltheart, 2011; Henson, 2011; Sternberg, 2011). Researchers have been investigating associative recognition for decades (e.g., Diana et al., 2006; Malmberg, 2008; Wixted and Stretch, 2004; Yonelinas, 2002), and have come up with single- and dual-process models to explain associative recognition. The HSMM-EEG analysis indicated the possible involvement of a decision stage, and although this idea will have to be confirmed in future experiments, we think it is an interesting new direction. Concluding, we believe that the HSMM-EEG analysis has the potential to provide new insights in human information processing.

Acknowledgments

This research was supported by the National Institute of Mental Health grant MH068243 and a James S. McDonnell Foundation (220020162) scholar award. We thank Katja Mehlhorn for comments on the paper.

Appendix A. Snapshot procedure

As explained in the main text, the first step in the HSMM-EEG analysis was to create 'snapshots' that spanned 80 ms. To create those snapshots, we took all 20 original samples in each of the 32 channels in an 80 ms period, and re-labeled those samples to represent one time point in 640 channels in a single snapshot. To give a simple example, if we had restructured 10 4-ms samples (40 ms period) in two channels (Table A1) into 2 20-ms snapshots, this would have resulted in 10 virtual channels (Table A2).

Table A1
Ten normal samples (40 ms) in two channels. 0_1 indicates a sample at time 0 in channel 1.

Time	Channel 1	Channel 2
0	0_1	0_2
4	4_1	4_2
8	8_1	8_2
12	12_1	12_2
16	16_1	16_2
20	20_1	20_2
24	24_1	24_2
28	28_1	28_2
32	32_1	32_2
36	36_1	36_2

Table A2
The samples from Table A1 restructured into two 20-ms snapshots.

Time	VC 1	VC 2	VC 3	VC 4	VC 5	VC 6	VC 7	VC 8	VC 9	VC 10
8	0_1	0_2	4_1	4_2	8_1	8_2	12_1	12_2	16_1	16_2
28	20_1	20_2	24_1	24_2	28_1	28_2	32_1	32_2	36_1	36_2

Appendix B. HSMM estimation

As mentioned in the main text, one can calculate optimal brain signatures M_i and gamma distributions G_{in} to describe the data given an HSMM with a particular number of states. Here we will explain this parameter estimation procedure in detail. For clarity we will describe the procedure assuming a single gamma distribution G_i per state, instead of separate gamma distributions G_{in} for each condition n and each subject. In addition, we will use 'sample' instead of 'snapshot' to describe the parameter estimation process.

Because the PCA factors are basically distributed as independent normals, we can calculate the likelihood for each set F_j of 100 component values f_{jk} at sample j, given state signature M_i (shown below the HSMM in Fig. 2) as follows:

$$P\left(F_j|M_i\right) = \prod_{k=1}^{100} \text{Normal}\left(f_{jk}, \mu_{ik}, 1\right) \tag{B1}$$

which is the product of the likelihoods of the component values f_{jk}, given normal distributions with 100 means μ_{ik} of state signature M_i of state i, where k is the index of the PCA components.

Underneath the state signatures in Fig. 2 the estimated gamma distributions of the state durations are shown. Because the gamma distributions were discretized to the nearest number of samples, and each sample is 80 ms in length, the likelihood of spending m samples in state i can be calculated as:

$$G(m|v_i, a_i) = \int_{80m-40}^{80m+40} \text{Gamma}(t|v_i a_i)dt \tag{B2}$$

where v_i and a_i are the shape and scale parameters of gamma distribution G_i of state i. Following Anderson and Fincham (2013), these integrals were approximated by using their midpoints. Note that we do allow for spending 0 samples in a state, in effect skipping the state.

Based on Eqs. (B1) and (B2) the likelihood of a trial given state signatures M_i and gamma distributions G_i can be calculated. It is the sum of the likelihoods of all ways in which the trial can be interpreted given the number of states and the number of samples. An interpretation of a trial is a way of breaking the m samples in a trial into r states. That is, an interpretation j is defined by $m_1 + m_2 + \ldots + m_r = m$, where m_i is the number of samples in state i. The likelihood of such an interpretation j is given by:

$$P_j(m_1, \ldots, m_r|M, v, a) = \prod_{i=1}^{r} \left(G(m_i|v_i\, a_i) \prod_{k=1}^{m_i} P(F_k|M_i) \right) \tag{B3}$$

which is the product of the likelihoods of the states. The likelihood of a state is the product of the likelihood of the state's duration given its gamma distribution and the likelihoods of the sample values in the state. The likelihood of a trial is the summed likelihood of all possible interpretations J of a trial:

$$P_{trial}(M, v, a) = \sum_{j=1}^{J} P_j(m_1, \ldots, m_r|M, v, a). \tag{B4}$$

HSMM algorithms can efficiently calculate this summed likelihood (Rabiner, 1989; Yu, 2010), even though the number of interpretations J increases rapidly with m and r:

$$J = \frac{(r+m-1)!}{m!(r-1)!}. \tag{B5}$$

Thus far, we have considered the likelihood of an HSMM with certain state signatures M_i and gamma distributions G_i. However, what we are interested in is finding parameters M_i and G_i that yield the optimal interpretation of the data given an HSMM with a certain number of states. In other words, finding the optimal r-state HSMM to model the data. To this end we used an expectation maximization algorithm (Dempster et al., 1977). We start the algorithm with neutral parameters and re-estimate these parameters until the likelihood of the data given the HSMM no longer increases. For the gamma distributions, both the shape and the scale parameters were initialized at $\sqrt{t/r}$, where t is the length of the trial and r the number of states (giving the states mean durations of t/r). The state signatures were initialized at 0, because the PCA components were normalized and thus on average 0. Since these probabilities can become too small for machine precision, we use log-likelihoods in our estimations and these are the measures we report. To calculate the solution we adapted software developed by Yu and Kobayashi (2003, 2006) to minimize the summed log-likelihood of all trials.

To ensure that our starting parameters resulted in the optimal solution, we also performed 1000 iterations with random starting parameters. The shape and scale parameters of the gamma distributions were initialized at 1 and t_{random_trial}/r, respectively, where t_{random_trial} represents the duration of one randomly drawn trial. In addition, instead of initializing the brain signatures at 0, for each of the states and factors we drew a random number from a standard normal distribution, divided by 10 to reduce extreme values. However, as none of these solutions outperformed the original solution for a significant number of subjects, we report the original solution in the paper.

The explanation above concerns an HSMM with a single gamma distribution per state. This model has i states $*$ $(100 + 2)$ parameters (100 means for the signature + 2 parameters for the gamma distribution). However, the full model with a separate gamma distribution for each condition and each subject has i states $*$ $(100 + 2 * n$ conditions $*$ subjects) parameters.

Appendix C. 6- and 7-state HSMMs for new foils

The new foils showed evidence for 7 states (Fig. 4). However, because the 5-state solution corresponded better to the 6-state solution of targets/re-paired foils, we focused on this solution in the main text. Here we report the 6- and 7-state solutions for comparison. Fig. C1 shows the state durations and signatures. The 6-state solution seems to divide the first state of the 5-state HSMM into two

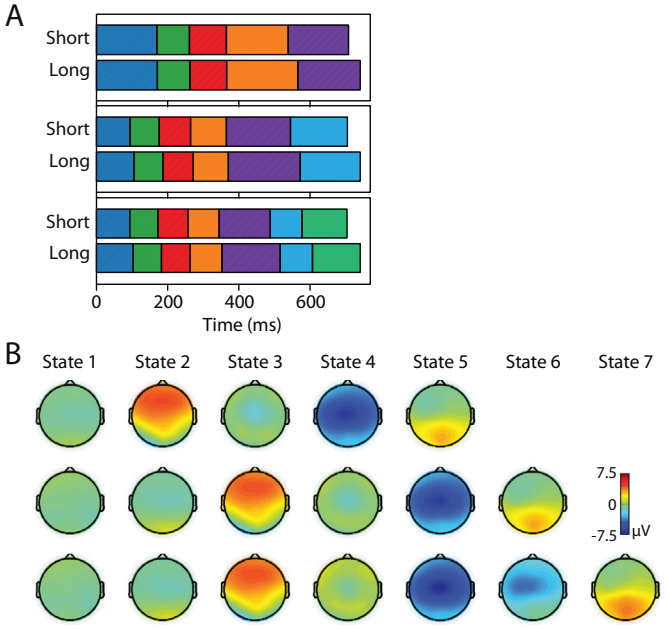

Fig. C1. State durations (A) and EEG signatures (B) for 5-, 6-, and 7-state HMMs for new foils.

separate states. The 7-state solution additionally divided the fourth state of the 5-states solution into two states. Given that the first two states, as well as the fourth state, of the 5-state solution mapped well on the states of the HSMM for targets/re-paired foils (Figs. 5 and 6, Table 3), we decided to concentrate on the 5-state solution.

Appendix D. Joint analysis of targets, re-paired foils, and new foils

In this appendix we report the results of a joint analysis on the targets, re-paired foils, and new foils. As explained in the main text (HSMM-EEG analysis section), conditions should be analyzed jointly only if they are hypothesized to consist of the same sequence of cognitive states. Thus, because the new foils probably go through different states than the targets/re-paired foils we committed them to an independent analysis in the main text. However, it turned out that the targets/re-paired foils and the new foils share most of their stages (Figs. 5 and 6; stage 4 was skipped by the new foils). We therefore performed an additional joint analysis, in which we fitted a single HSMM for the targets, re-paired foils and new foils.

This joint analysis confirmed the results of the independent analyses reported in the main text. Fig. D1 shows that a 6-state HSMM outperformed all HSMMs with fewer states, and no HSMM with more states had a higher log-likelihood for a significant number of subjects. The six states correspond to the states identified by the independent analyses, as is shown in Figs. D2 and D3. Fig. D2 shows the state signatures and associated gamma distributions, while Fig. D3 depicts the duration of the states for the conditions in the experiment.

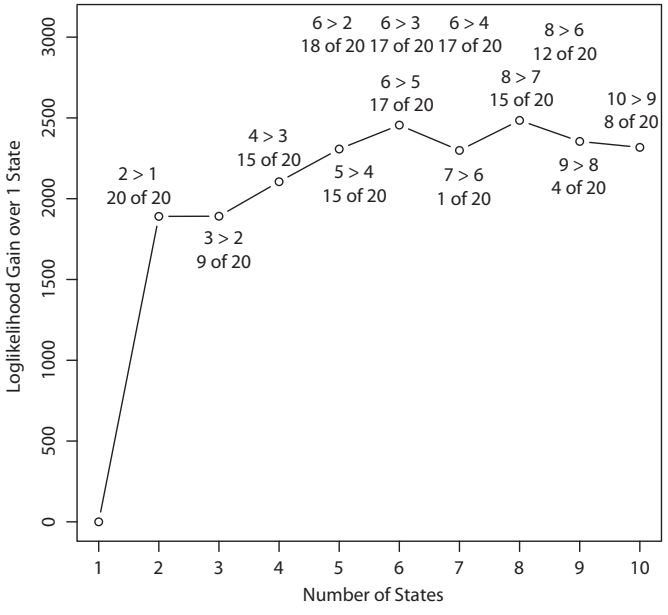

Fig. D1. Results of the LOOCV-procedure for the joint analysis of targets, re-paired foils, and new foils: the average gain in log-likelihood of an *r*-state HSMM over a 1-state HSMM. The figure additionally shows for how many subjects the log-likelihood increased from state to state: for instance '3 > 2, 11 of 20' means that for 11 subjects the log-likelihood of a 3-state HSMM was higher than the log-likelihood of a 2-state model. According to a sign-test with 20 subjects, a significant increase is reached when 15 out of 20 subjects improve ($p = .04$), whereas 18 out of 20 corresponds to a *p*-value of .0004.

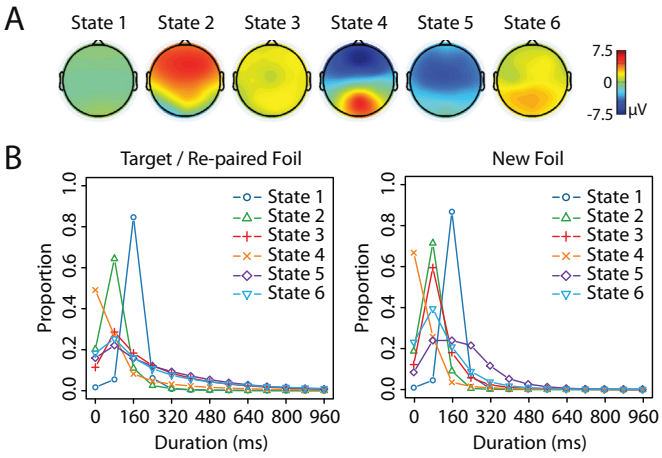

Fig. D2. State signatures (A) and gamma distributions (B) of the joint analysis.

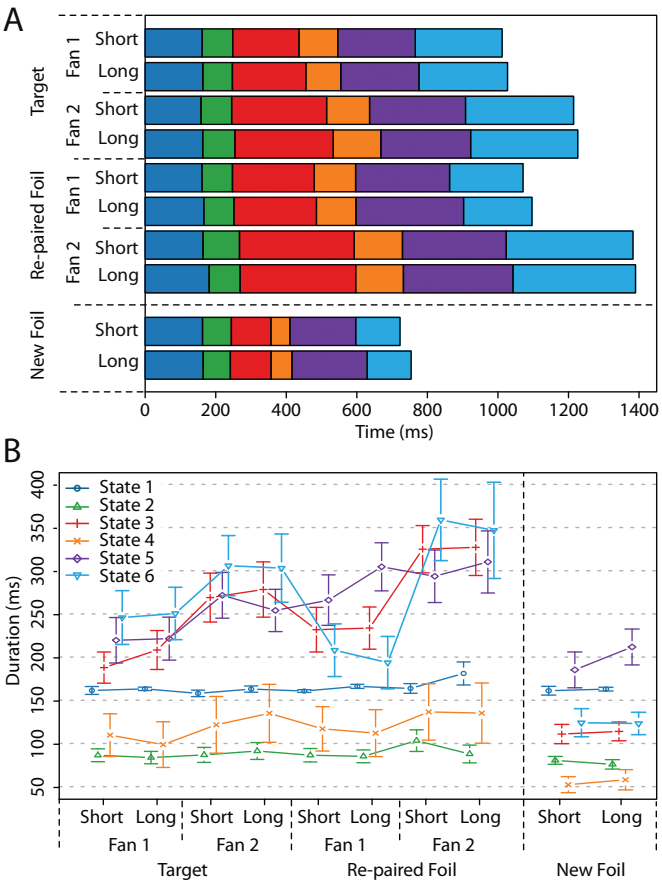

Fig. D3. State durations of the joint analysis: A shows how the state durations add up to form complete trials; B shows more clearly how state durations are affected by condition. Error bars indicate standard errors.

References

Anderson, J.R., 2007. How Can the Human Mind Occur in the Physical Universe? Oxford University Press, New York

Anderson, J.R., Fincham, J.M., 2013. Discovering the sequential structure of thought. Cogn. Sci. 37 (6), 1–31.

Anderson, J.R., Reder, L.M., 1999. The fan effect: new results and new theories. J. Exp. Psychol. Gen. 128 (2), 186–197. http://dx.doi.org/10.1037/0096-3445.128.2.186.

Anderson, J.R., Betts, S., Ferris, J.L., Fincham, J.M., 2010. Neural imaging to track mental states while using an intelligent tutoring system. Proc. Natl. Acad. Sci. U. S. A. 107 (15), 7018–7023. http://dx.doi.org/10.1073/pnas.1000942107.

Anderson, J.R., Betts, S., Ferris, J.L., Fincham, J.M., 2012a. Tracking children's mental states while solving algebra equations. Hum. Brain Mapp. 33 (11), 2650–2665. http://dx.doi.org/10.1002/hbm.21391.

Anderson, J.R., Fincham, J.M., Schneider, D.W., Yang, J., 2012b. Using brain imaging to track problem solving in a complex state space. NeuroImage 60 (1), 633–643. http://dx.doi.org/10.1016/j.neuroimage.2011.12.025.

Borst, J.P., Schneider, D.W., Walsh, M.M., Anderson, J.R., 2013. Stages of processing in associative recognition: evidence from behavior, electroencephalography, and classification. J. Cogn. Neurosci. 25 (12), 2151–2166.

Britz, J., Michel, C.M., 2011. State-dependent visual processing. Front. Psychol. 2. http://dx.doi.org/10.3389/fpsyg.2011.00370.

Brunet, D., Murray, M.M., Michel, C.M., 2011. Spatiotemporal analysis of multichannel EEG: CARTOOL. Comput. Intell. Neurosci. 2011, 813870. http://dx.doi.org/10.1155/2011/813870.

Clark, S.E., Gronlund, S.D., 1996. Global matching models of recognition memory: how the models match the data. Psychon. Bull. Rev. 3 (1), 37–60. http://dx.doi.org/10.3758/BF03210740.

Coltheart, M., 2011. Methods for modular modelling: additive factors and cognitive neuropsychology. Cogn. Neuropsychol. 28 (3–4), 224–240. http://dx.doi.org/10.1080/02643294.2011.587794.

Colzato, L., Bajo, M., Van Den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., Hommel, B., 2008. How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. J. Exp. Psychol. Learn. Mem. Cogn. 34 (2), 302–312. http://dx.doi.org/10.1037/0278-7393.34.2.302.

Danker, J.F., Gunn, P., Anderson, J.R., 2008. A rational account of memory predicts left prefrontal activation during controlled retrieval. Cereb. Cortex 18 (11), 2674–2685. http://dx.doi.org/10.1093/cercor/bhn027.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21. http://dx.doi.org/10.1016/j.jneumeth.2003.10.009.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B Methodol. 1–38.

Diana, R.A., Reder, L.M., Arndt, J., Park, H., 2006. Models of recognition: a review of arguments in favor of a dual-process account. Psychon. Bull. Rev. 13 (1), 1–21. http://dx.doi.org/10.3758/BF03193807.

Donders, F.C., 1868. De snelheid van psychische processen (On the speed of mental processes).

Finnigan, S., Humphreys, M.S., Dennis, S., Geffen, G., 2002. ERP "old/new" effects: memory strength and decisional factor(s). Neuropsychologia 40 (13), 2288–2304. http://dx.doi.org/10.1016/S0028-3932(02)00113-6.

Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage 15 (4), 870–878. http://dx.doi.org/10.1006/nimg.2001.1037.

Gillund, G., Shiffrin, R.M., 1984. A retrieval model for both recognition and recall. Psychol. Rev. 91 (1), 1–67.

Gronlund, S.D., Ratcliff, R., 1989. Time course of item and associative information: implications for global memory models. J. Exp. Psychol. Learn. Mem. Cogn. 15 (5), 846–858.

Hauk, O., Pulvermüller, F., 2004. Effects of word length and frequency on the human event-related potential. Clin. Neurophysiol. 115 (5), 1090–1103.

Hauk, O., Davis, M.H., Ford, M., Pulvermüller, F., Marslen-Wilson, W.D., 2006. The time course of visual word recognition as revealed by linear regression analysis of ERP data. NeuroImage 30 (4), 1383–1400. http://dx.doi.org/10.1016/j.neuroimage.2005.11.048.

Hauk, O., Pulvermüller, F., Ford, M., Marslen-Wilson, W.D., Davis, M.H., 2009. Can I have a quick word? Early electrophysiological manifestations of psycholinguistic processes revealed by event-related regression analysis of the EEG. Biol. Psychol. 80 (1), 64–74. http://dx.doi.org/10.1016/j.biopsycho.2008.04.015.

Heil, M., Rösler, F., Hennighausen, E., 1997. Topography of brain electrical activity dissociates the retrieval of spatial versus verbal information from episodic long-term memory in humans. Neurosci. Lett. 222 (1), 45–48.

Henson, R.N., 2011. How to discover modules in mind and brain: the curse of nonlinearity, and blessing of neuroimaging. A comment on Sternberg (2011). Cogn. Neuropsychol. 28 (3–4), 209–223. http://dx.doi.org/10.1080/02643294.2011.561305.

Hintzman, D.L., 1988. Judgments of frequency and recognition memory in a multiple-trace memory model. Psychol. Rev. 95 (4), 528–551. http://dx.doi.org/10.1037/0033-295X.95.4.528.

Juhasz, B.J., Rayner, K., 2003. Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. J. Exp. Psychol. Learn. Mem. Cogn. 29 (6), 1312–1318. http://dx.doi.org/10.1037/0278-7393.29.6.1312.

Khader, P., Heil, M., Rösler, F., 2005. Material-specific long-term memory representations of faces and spatial positions: evidence from slow event-related brain potentials. Cogn. Psychol. 43 (14), 2109–2124. http://dx.doi.org/10.1016/j.neuropsychologia.2005.03.012.

Khader, P., Knoth, K., Burke, M., Ranganath, C., Bien, S., Rösler, F., 2007. Topography and dynamics of associative long-term memory retrieval in humans. J. Cogn. Neurosci. 19 (3), 493–512. http://dx.doi.org/10.1162/jocn.2007.19.3.493.

Lehmann, D., Strik, W.K., Henggeler, B., Koenig, T., Koukkou, M., 1998. Brain electric microstates and momentary conscious mind states as building blocks of spontaneous thinking: I. Visual imagery and abstract thoughts. Int. J. Psychophysiol. 29 (1), 1–11. http://dx.doi.org/10.1016/S0167-8760(97)00098-6.

Luck, S.J., 2005. An Introduction to the Event-Related Potential Technique. MIT Press.

Malmberg, K.J., 2008. Recognition memory: a review of the critical findings and an integrated theory for relating them. Cogn. Psychol. 57 (4), 335–384. http://dx.doi.org/10.1016/j.cogpsych.2008.02.004.

Murdock, B.B., 1993. TODAM2: a model for the storage and retrieval of item, associative, and serial-order information. Psychol. Rev. 100 (2), 183–203. http://dx.doi.org/10.1037/0033-295X.100.2.183.

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. http://dx.doi.org/10.1037/0033-2909.131.4.510.

Nyhus, E., Curran, T., 2009. Semantic and perceptual effects on recognition memory: evidence from ERP. Brain Res. 1283, 102–114. http://dx.doi.org/10.1016/j.brainres.2009.05.091.

Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1995. Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Trans. Biomed. Eng. 42 (7), 658–665.

Pirolli, P.L., Anderson, J.R., 1985. The role of practice in fact retrieval. J. Exp. Psychol. Learn. Mem. Cogn. 11 (1), 136.

Rabiner, L.R., 1989. A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE 77 (2), 257–286.

Ratcliff, R., McKoon, G., 1989. Similarity information versus relational information: differences in the time course of retrieval. Cogn. Psychol. 21 (2), 139–155. http://dx.doi.org/10.1016/0010-0285(89)90005-4.

Ratcliff, R., Murdock, B.B., 1976. Retrieval processes in recognition memory. Psychol. Rev. 83 (3), 190.

Reder, L.M., Nhouyvanisvong, A., Schunn, C.D., Ayers, M.S., Angstadt, P., Hiraki, K., 2000. A mechanistic account of the mirror effect for word frequency: a computational model of remember–know judgments in a continuous recognition paradigm. J. Exp. Psychol. Learn. Mem. Cogn. 26 (2), 294–320. http://dx.doi.org/10.1037/0278-7393.26.2.294.

Roberts, S., Sternberg, S., 1993. The meaning of additive reaction-time effects: tests of three alternatives. Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience vol. 14, pp. 611–653.

Rotello, C.M., Heit, E., 2000. Associative recognition: a case of recall-to-reject processing. Mem. Cogn. 28 (6), 907–922. http://dx.doi.org/10.3758/BF03209339.

Rugg, M.D., Curran, T., 2007. Event-related potentials and recognition memory. Trends Cogn. Sci. 11 (6), 251–257. http://dx.doi.org/10.1016/j.tics.2007.04.004.

Schneider, D.W., Anderson, J.R., 2012. Modeling fan effects on the time course of associative recognition. Cogn. Psychol. 64 (3), 127–160. http://dx.doi.org/10.1016/j.cogpsych.2011.11.001.

Sohn, M.H., Goode, A., Stenger, V.A., Jung, K.J., Carter, C.S., Anderson, J.R., 2005. An information-processing model of three cortical regions: evidence in episodic memory retrieval. NeuroImage 25 (1), 21–33. http://dx.doi.org/10.1016/j.neuroimage.2004.11.001.

Speer, N.K., Curran, T., 2007. ERP correlates of familiarity and recollection processes in visual associative recognition. Brain Res. 1174, 97–109. http://dx.doi.org/10.1016/j.brainres.2007.08.024.

Spinelli, D., De Luca, M., Di Filippo, G., Mancini, M., Martelli, M., Zoccolotti, P., 2005. Length effect in word naming in reading: role of reading experience and reading deficit in Italian readers. Dev. Neuropsychol. 27 (2), 217–235. http://dx.doi.org/10.1207/s15326942dn2702_2.

Sternberg, S., 1969. The discovery of processing stages: extensions of Donders' method. Acta Psychol. 30, 276–315. http://dx.doi.org/10.1016/0001-6918(69)90055-9.

Sternberg, S., 2011. Modular processes in mind and brain. Cogn. Neuropsychol. 28 (3–4), 156–208. http://dx.doi.org/10.1080/02643294.2011.557231.

Sudre, G., Pomerleau, D., Palatucci, M., Wehbe, L., Fyshe, A., Salmelin, R., Mitchell, T., 2012. Tracking neural coding of perceptual and semantic features of concrete nouns. NeuroImage 62 (1), 451–463. http://dx.doi.org/10.1016/j.neuroimage.2012.04.048.

Sutton, S., Ruchkin, D.S., Munson, R., Kietzman, M.L., Hammer, M., 1982. Event-related potentials in a two-interval forced-choice detection task. Psychon. Bull. Rev. 32 (4), 360–374. http://dx.doi.org/10.3758/BF03206242.

Taatgen, N.A., Juvina, I., Schipper, M., Borst, J.P., Martens, S., 2009. Too much control can hurt: a threaded cognition model of the attentional blink. Cogn. Psychol. http://dx.doi.org/10.1016/j.cogpsych.2008.12.002.

Van De Ville, D., Britz, J., Michel, C.M., 2010. EEG microstate sequences in healthy humans at rest reveal scale-free dynamics. Proc. Natl. Acad. Sci. U. S. A. 107 (42), 18179–18184. http://dx.doi.org/10.1073/pnas.1007841107.

Van Petten, C., Kutas, M., 1990. Interactions between sentence context and word frequency in event-related brain potentials. Mem. Cogn. 18 (4), 380–393.

Wilkinson, R.T., Seales, D.M., 1978. EEG event-related potentials and signal detection. Biol. Psychol. 7 (1), 13–28. http://dx.doi.org/10.1016/0301-0511(78)90039-X.

Wixted, J.T., 2007. Dual-process theory and signal-detection theory of recognition memory. Psychol. Rev. 114 (1), 152.

Wixted, J.T., Stretch, V., 2004. In defense of the signal detection interpretation of remember/know judgments. Psychon. Bull. Rev. 11 (4), 616–641. http://dx.doi.org/10.3758/BF03196616.

Yonelinas, A.P., 2002. The nature of recollection and familiarity: a review of 30 years of research. J. Mem. Lang. 46 (3), 441–517. http://dx.doi.org/10.1006/jmla.2002.2864.

Yu, S.Z., 2010. Hidden semi-Markov models. Artif. Intell. 174 (2), 215–243. http://dx.doi.org/10.1016/j.artint.2009.11.011.

Yu, S.Z., Kobayashi, H., 2003. An efficient forward-backward algorithm for an explicit-duration hidden Markov model. IEEE Signal Process. Lett. 10 (1), 11–14.

Yu, S.Z., Kobayashi, H., 2006. Practical implementation of an efficient forward–backward algorithm for an explicit-duration hidden Markov model. Signal Process. IEEE Trans. 54 (5), 1947–1951.