



The congruent, the incongruent, and the unexpected: Event-related potentials unveil the processes involved in schematic encoding

Gerrit Höljtje*, Bente Lubahn, Axel Mecklinger

Experimental Neuropsychology Unit, Department of Psychology, Saarland University, Saarbrücken, Germany



ARTICLE INFO

Keywords:

Schema
Semantic congruency
Prediction
Language
Recognition memory
Event-related potentials
Subsequent memory effect
Frontal positivity

ABSTRACT

Learning is most effective when new information can be related to a preexisting knowledge structure or schema. In the present study, event-related potentials (ERPs) were used to investigate the temporal dynamics of the processes by which activated schemata support the encoding of schema-congruent information. Participants learned category exemplar words that were either semantically congruent or incongruent with a preceding category cue phrase. Congruent words were composed of expected (high typicality, HT) and unexpected (low typicality, LT) category exemplars. On the next day, recognition memory for the exemplars and the category cues they were presented with was tested. Semantically related lures were used in order to ascertain that memory judgements were based on episodic memory for specific category exemplars. Generally, congruent (HT and LT) exemplars were remembered better than incongruent exemplars. ERPs recorded during the encoding of the exemplar words were compared for subsequently remembered and forgotten items. Subsequent memory effects (SME) emerged in the N400 time window at frontal electrodes and did not differ between congruent and incongruent exemplars. In the same epoch, an SME with a parietal distribution was specific for congruent exemplars, suggesting that activated schemata strengthened memory for congruent exemplars by supporting the encoding of item-specific details. Subsequently remembered LT exemplars were associated with a late frontal positivity that is assumed to reflect expectancy mismatch-related processing such as the contextual integration of an unexpected word by the suppression of strongly expected words. A correlation analysis revealed that the greater the involvement of the processes reflected by the frontal positivity, the lower the level of false positive memory responses in the test phase one day later. These results suggest that the contextual integration of schema-congruent but unexpected events involves a weakening of the representations of semantically related, but unstudied items in memory and by this benefits subsequent memory.

1. Introduction

Schemata are superordinate knowledge structures that are thought to consist of associative connections extracted from commonalities of multiple single episodes (see Gilboa and Marlatte, 2017, for a recent review). Once a schema is activated, it can exert an influence on the processing of newly encountered information. According to a recently proposed framework for schema-based memory, i.e. the SLIMM framework (van Kesteren et al., 2012), schematic encoding mainly relies on the medial prefrontal cortex (mPFC) and is associated with better memory for schema-congruent than for incongruent information (van Kesteren et al., 2013a). Furthermore, memory for schema-congruent information is thought to benefit from a fast consolidation of schema-congruent memory traces (van Kesteren et al., 2012). In support of this view, schema effects on item memory (i.e. better memory for schema-

congruent than incongruent items) have been found to increase after a period of consolidation (Durrant et al., 2015; van der Linden et al., 2017; van Kesteren et al., 2013b). As a consequence of the fast consolidation and the concomitant accelerated neocortical integration, schema-congruent memory traces become independent from the medial temporal lobe (MTL) particularly fast (Tse et al., 2007) and can be characterized as being more semantic and less detailed than memory traces that rely more strongly on the MTL (Spalding et al., 2015; van der Linden et al., 2017; Winocur et al., 2010).

While a considerable number of studies have explored the brain regions involved in schema-based learning (Brod et al., 2015; Spalding et al., 2015; Staresina et al., 2009; van der Linden et al., 2017; van Kesteren et al., 2013a; van Kesteren et al., 2010) and the brain regions mediating schematic encoding have been disclosed (Preston and Eichenbaum, 2013; van Kesteren et al., 2012), relatively little is known

* Corresponding author. Address: Experimental Neuropsychology Unit, Department of Psychology, Campus A2.4, Saarland University, 66123 Saarbrücken, Germany.
E-mail address: gerrit.hoeltje@uni-saarland.de (G. Höljtje).

about the temporal characteristics of the processes involved in schema-based learning. Event-related potentials (ERPs) are well-suited to address this question because of their high temporal resolution. A long-standing line of ERP research has made use of the subsequent memory paradigm (Sanquist et al., 1980) to examine neural activity indicative of successful memory encoding. In this approach, activity recorded during the encoding of items is compared for subsequent hits and misses, i.e. items that are remembered versus forgotten on a subsequent memory test (see Cohen et al., 2015, or Paller and Wagner, 2002, for reviews).

In an illustrative recent study by Packard et al. (2017), ERPs were used to investigate schematic memory encoding in a subsequent memory paradigm. Packard and colleagues used a DRM task (Deese, 1959; Roediger and McDermott, 1995) to compare recognition memory for category exemplar word lists that were preceded by a semantically congruent or incongruent category cue. They found that congruent targets were associated with increased hit rates, but also with increased false alarm rates (false memories) for semantically related lures. As semantic congruency increased true and false memories there was no indication that memory accuracy was higher for congruent than for incongruent exemplars. Therefore, it is possible that the processing of congruent exemplars in the study phase, rather than promoting episodic encoding, was associated mostly with increased activation for all exemplars belonging to the studied categories (see Tibon et al., 2017, for a similar argument). The same objection applies to a second ERP experiment conducted by Packard et al. (2017, Exp. 4), in which no semantically related lures were used in the recognition test. An alternative account for the data, namely that the participants' decisions were guided by the semantic congruency of an item with a studied category cue rather than reflecting episodic memory, cannot be excluded in either experiment. Comparing ERPs for subsequently remembered and forgotten items, Packard et al. (2017) found a congruency by subsequent memory interaction. The difference between subsequently remembered and forgotten congruent targets emerged at around 400 ms and by this 200 ms earlier than the corresponding difference for incongruent targets. Based on these results, Packard et al. (2017) concluded that semantic congruency accelerates the onset of ERP correlates of memory formation. However, because it cannot be ruled out that memory decisions were mostly based on the semantic congruency of an item with a studied category, an alternative interpretation of the ERP data seems more likely, namely that congruent words judged as "old" in the memory test were associated with attenuated N400 responses, known to be sensitive to semantic congruency relations (see Kutas and Federmeier, 2011, for a review), in the study phase. Thus, the early on setting SME for subsequently remembered congruent targets reported by Packard et al. (2017) may not reflect episodic memory formation for a specific event, but rather semantic integration processes as indexed by the N400 attenuation, that were contingent on the semantic congruency relations between category cues and study words (Tibon et al., 2017).

In summary, even though the Packard et al. (2017) study yielded some important insights, the shortcomings described do not allow for a conclusive interpretation of the (ERP) data, and uncertainty remains regarding the temporal characteristics of the processes involved in schema-based learning. Therefore, the aim of the present study was to investigate the ERP correlates of the processes by which activated schemata support the encoding of schema-congruent information. For this purpose, it is important to employ a design that allows to rule out alternative interpretations in terms of increased category activation for congruent items, for example by contrasting memory decisions for studied words and semantically related lures. Under this precondition, showing that a congruent semantic context during encoding boosts memory for studied words, without increasing false alarms to related lures to a similar extent, would allow to identify ERP activity related to episodic memory formation and to compare SME for schema-congruent and -incongruent words.

Another goal of the present study was to explore the role of

expectancy mismatches during the encoding of schema-congruent information and their ensuing mnemonic consequences. Given that schema knowledge is activated by context information, it can be assumed that expectancies (predictions) regarding the upcoming input are built up (Ghosh and Gilboa, 2014). In some situations, expectancies can be disconfirmed by plausible, but unexpected information, and this type of mismatch processing may benefit subsequent memory for these events. In an illustrative study, Federmeier et al. (2007) compared expected and unexpected sentence ending words for sentences inducing a strong or weak expectation about the final word of the sentence. In support of the aforementioned view, they found that plausible, but unexpected sentence ending words after sentence frames that induced a strong expectation (e.g. the word "collection" after a sentence frame like "He bought her a pearl necklace for her ...") were associated with particularly good recognition memory and a late frontal positivity following the N400. Notably, it has been shown that a late frontal positivity is not only elicited by unexpected sentence endings, but also by semantically congruent, untypical category exemplar words following short phrases (e.g. the word "ash" after a phrase like "a kind of tree"; Federmeier et al., 2010). In contrast, incongruent exemplars (e.g. the word "tin" after the same phrase) did not elicit a frontal positivity, presumably because these words were not interpretable in the given context and expectancies were not disconfirmed.

Even though the functional significance of this ERP component has not yet been completely disclosed, it has been proposed that such late frontal positivities are elicited when the processing and contextual integration of an unexpected word demands the suppression of strongly expected and competing words (DeLong et al., 2011; Federmeier et al., 2007; Kutas, 1993; see Van Petten and Luka, 2012, for a review). It is less clear, however, how the processes reflected by the frontal positivity benefit subsequent memory for unexpected events. The suppression hypothesis suggests that, rather than reflecting processes that directly contribute to memory formation for unexpected target representations, the late frontal positivity reflects post-encoding processes that act upon the representations of strongly expected events. This implies that successfully encoded (i.e. subsequently remembered) unexpected events should be associated with strong expectancy mismatch-related processing as reflected in the late frontal positivity.

The present study

In this study ERPs were used to explore schema-based online memory encoding. On every trial of the learning phase, participants were presented a semantic category cue (e.g. "A four-footed animal"), followed by a congruent ("dog") or incongruent ("sapphire") category exemplar. To compare expected and unexpected words, congruent exemplars were either of high or low typicality ("dog" or "wolf", respectively). One day after the encoding session, memory for the exemplar words and the category cues they were presented with was tested in a recognition test. Semantically related lures were used in the recognition memory test in order to ascertain that memory decisions were based on episodic memory for the category exemplars. Whereas semantically unrelated lures could be rejected based on memory for a semantic category presented in the study phase, semantically related lures can only be rejected based on detailed episodic memory for a specific category exemplar.

If a congruent semantic context supports memory encoding, memory accuracy should be higher for congruent than for incongruent exemplars (van Kesteren et al., 2013a). If this congruency effect in memory is related to an accelerated onset of memory formation for congruent exemplars, as proposed by Packard et al. (2017), ERP subsequent memory effects should differ in their onset between congruent and incongruent exemplars.

Congruent, but unexpected (low typicality) exemplars were expected to be associated with particularly high memory accuracy and a late frontal positivity (Federmeier et al., 2010, 2007). Because the late frontal positivity probably does not reflect memory encoding processes per se, we did not predict larger SME for low versus high typicality

exemplars. Rather, as stated above, successfully encoded (i.e. subsequently remembered) unexpected exemplars should be associated with strong expectancy mismatch-related processing as reflected in the late frontal positivity.

2. Material and methods

2.1. Participants

Twenty-eight young adults (21 female) participated in the experiment. Their age ranged between 19 and 29 years ($Mdn = 23$). They were all German native speakers, had normal or corrected-to-normal vision and no neurological or psychiatric conditions. Before the experiment, participants gave their informed consent and completed the Edinburgh Handedness Inventory (Oldfield, 1971) confirming their right-handedness. After testing, a post-enquiry and debriefing followed, and participants received money (8 euros per hour) or course credit as a compensation for their participation. Four participants were excluded from all analyses because of an insufficient number of EEG segments per condition (see 2.5). Thus, all analyses are based on a sample of $N = 24$ subjects that was completely balanced regarding list and key assignments.

2.2. Stimuli

Stimuli consisted of 60 phrasal category cues and 480 category exemplar words, examples of which are given in Table 1. For every category cue, four congruent and four incongruent exemplars were retrieved from rating studies based on and complementing the Battig and Montague (1969) category norms (Hellerstedt et al., 2012; Mannhaupt, 1983; Van Overschelde, Rawson and Dunlosky, 2004). The four congruent exemplars were composed of two high typicality (HT) and two low typicality (LT) exemplars. HT exemplars were the two most frequently generated responses to the category cues ($M = 0.70$, $SD = 0.20$); except for a few cases in which the most frequent response was related to another category. LT exemplars were two items that were generated in response to the cue with substantially lower probability, $M = 0.06$, $SD = 0.03$. Incongruent (IC) exemplars were generated by randomly redistributing four exemplars of medium congruency from every category to others. The exemplar word lists (HT, LT, and IC) were composed of nouns and matched for word length ($M = 6.20$, $SD = 0.93$) across conditions, all p -values $> .055$. Lemma frequencies were taken from the dlexDB database (Heister et al., 2011) and higher for HT ($M = 4410.38$, $SD = 6909.35$) than for LT exemplars ($M = 1869.56$, $SD = 3341.96$), $t(59) = 3.20$, $p < .01$, $d = 0.43$, but did not differ between IC exemplars ($M = 2562.93$, $SD = 8311.24$) and both HT and LT exemplars (IC vs. HT: $t(59) = 1.46$, $p = .15$, $d = 0.24$; IC vs. LT: $t(59) = 0.63$, $p = .53$, $d = 0.11$).

2.3. Procedure

The experiment consisted of a study (ca. 45 min) and a test phase (ca. 55 min) taking place on two consecutive days. In the study phase,

Table 1

Example stimuli. Please note that in this example, the assignment of incongruent exemplars illustrates that exemplars from one category were used as incongruent exemplars for other categories.

Category Cue	High Typicality	Low Typicality	Incongruent
A four-footed animal	Dog	Fox	Pepper
A fruit	Apple	Apricot	Hair
A vegetable	Carrot	Zucchini	Puzzle
A metal	Steel	Zinc	Bear
A part of the human body	Arm	Tongue	Tin
A toy	Doll	Marble	Mango

participants learned associations between the category cues and exemplars and rated the congruency of the presented category-exemplar pairings. 24 h later, recognition memory of the exemplars was assessed. For items judged as “old”, a 3-alternative forced-choice (3AFC) recognition memory test additionally assessed memory for the associated category. All experimental tasks were presented using E-Prime 2 software (Psychology Software Tools, Inc.), and participants used a keyboard for their responses.

2.3.1. Study phase

When participants arrived at the laboratory, approximately 45 min were spent on preparation for the EEG recording. Thereafter, participants were seated in front of a 19” computer screen with a resolution of 1280 x 1024 pixels in an electrically shielded and sound-attenuated booth.

Overall, the study phase consisted of 240 trials, divided into six blocks. Each of the 60 category cues was presented with one HT, one LT and two IC exemplars. The remaining 240 exemplars would later be used as related lures in the test phase. The assignment of exemplars to the study or lure function was balanced across subjects. At the beginning of each block, participants were instructed to remember the pairs for a test 24 h later, but were not told what aspects their memory would be tested on. The main task was to rate the congruency of every exemplar to the category shown before. Participants were given five practice trials.

Every trial started with a central fixation cross (500 ms), followed by the category cue (3000 ms). After a blank screen (1750 ms) and a fixation cross (jittered between 500 and 750 ms) the target exemplar was presented for 2000 ms and followed by a congruency rating asking about the perceived fit between category and exemplar. Subjects indicated their answer by key press on a 4-point scale ranging from “good” via “possible” and “rather not” to “not”. The presentation of the rating scale continued until a reaction was registered. Pairs were presented pseudorandomly, with no more than three adjacent incongruent or congruent exemplar trials.

2.3.2. Test phase

Approximately 24 h later (range: 23–25.5, $M = 24.05$, $SD = 0.53$), recognition memory for the studied exemplars and the associated category cues was assessed. The test phase took place in another laboratory, where the experiment was presented on a 24” screen (1920 x 1080 pixels resolution).

In the recognition memory test, the 240 studied exemplars were presented randomly intermixed with 240 unstudied exemplars in six blocks. Subjects were instructed to discriminate old items from new items (lures) and, for items judged as “old”, to identify the category an exemplar was presented with at study in a 3AFC recognition memory test. Participants were given five practice trials.

Trials started with the presentation of a fixation cross (1000 ms). Thereafter, the exemplar was presented for 1500 ms and the response interval started. The six-point confidence scale (“sure old”, “probably old”, “maybe old”, “maybe new”, “probably new” and “sure new”) appeared after a 1000 ms blank screen together with the question “Old or New?”.

For every “old” judgement, the request “please choose the category this exemplar was studied with” was presented together with three category cues - always one congruent and two incongruent ones (see van Kesteren et al., 2013a, for a similar procedure). For congruent exemplars, the correct congruent category cue and two randomly assigned incongruent category cues were used. For incongruent exemplars the correct (incongruent) category was shown together with a second incongruent and a congruent category (both incorrect).

2.4. Behavioral data

To quantify the subjective congruency ratings at study, mean rating

scores were calculated for every level of the Congruency factor, with a value of four (zero) indicating the highest (lowest) possible score. In order to quantify item recognition performance, probability of true recognition (Pr) scores (Snodgrass and Corwin, 1988) were calculated as the difference between the proportions of correct and incorrect “old”-decisions (hits and false alarms). For this purpose, the corresponding three steps of the confidence scale were collapsed into “old”- and “new”-decisions. High-confidence Pr scores were calculated based on “sure”- and “probably”- decisions only. Hit rates were calculated for each level of the Congruency factor. Because congruency relations could only be established in relation to the category cue phrases and not on the item level, and in the test phase all exemplars were congruent to at least one semantic category, a single false alarm rate was calculated for all lure exemplars.

Associative memory performance was calculated as the proportion of correctly recognized exemplars for which the category was correctly identified, irrespective of associative memory confidence. Because the associative decisions showed a strong bias in favor of choosing the congruent category cue, hit rates were corrected for the probability of choosing the correct category by guessing. For the congruent conditions, this probability was equivalent to the proportion of item recognition false alarms across congruency conditions (i.e., lure exemplars judged as “old”) for which the congruent category cue was chosen. For the incongruent condition, because always two incongruent category cues were given as options, the guessing probability was equivalent to the proportion of item recognition false alarms across congruency conditions for which the incongruent category cue was chosen, divided by two. Corrected hit rates were calculated by subtracting the corresponding guessing probability from the hit rate. Means and standard errors of the mean (SEM) for all behavioral measures are given in Table 2.

2.5. EEG recording and analysis

The EEG was recorded during the learning phase from 28 Ag/AgCl scalp electrodes embedded in an elastic cap with positions according to the 10–20 electrode system (Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC3, FCz, FC4, FC6, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P7, P3, Pz, P4, P8, O1, O2, and A2). Blinks and eye movements were monitored horizontally and vertically from four electrodes placed above and below the right eye and at the canthi of the left and right eyes. The electrodes were online referenced to a left mastoid electrode (A1). AFz was used as a ground electrode. The EEG was amplified with a BrainAmp DC amplifier (Brain Products GmbH) from 0.016 to 250 Hz and digitized at 500 Hz. For off-line processing of the EEG data, Brain Vision Analyzer 2 software (Brain Products GmbH) was used. Electrodes were re-referenced to the average of the left and right mastoid electrodes. The data were lowpass-filtered at 40 Hz using a second order zero phase shift Butterworth filter. Independent component analysis (ICA) was applied to the continuous data to correct for ocular artefacts. Components associated with blinks and eye movements were rejected using a semi-automatic algorithm implemented in BrainVision Analyzer 2 (Ocular Correction ICA). The algorithm identified components that showed high correlations with

Table 2
Means (SEM) for behavioral measures.

	HT	LT	IC
Congruency Rating	3.88 (0.02)	3.41 (0.05)	1.16 (0.04)
Hit rate	0.76 (0.03)	0.79 (0.03)	0.57 (0.03)
False alarm rate	0.24 (0.02)		
Pr score	0.53 (0.03)	0.55 (0.03)	0.33 (0.03)
High-confidence hit rate	0.63 (0.04)	0.67 (0.03)	0.38 (0.03)
High-confidence false alarm rate	0.11 (0.01)		
High confidence Pr score	0.52 (0.04)	0.56 (0.03)	0.28 (0.03)
Corrected associative hit rate	0.21 (0.02)	0.23 (0.03)	0.21 (0.02)

vertical and horizontal eye channel activity. In a second step, the component topographies were checked manually. Components that exhibited a typical blink or eye movement topography were rejected. For ERPs, segments were extracted from 200 ms before the onset of a study phase exemplar to 2000 ms thereafter and baseline-corrected based on the 200 ms prestimulus activity. Segments containing artefacts were rejected using the following criteria: A maximal allowed voltage step of 50 $\mu\text{V}/\text{ms}$, a maximal difference of values of 200 μV during intervals of 200 ms, and minimal and maximal allowed total amplitude of $\pm 100 \mu\text{V}$. On average, 7.1% of segments were rejected. Subsequent hits were defined as targets that were correctly judged as “old” (item hits) and for which the correct associated category cue was chosen (associative hits) in the test phase. Subsequent misses consisted of targets that were either judged as “new” (item misses) or for which the correct category cue was not identified (associative misses) in the test phase. Subject averages for combinations of the factors Congruency (HT, LT, IC) and Memory (hits, misses) were calculated based on a number of at least seven artifact-free trials per condition (for SME studies using a similar criterion for trial selection, see Kamp et al., 2017, 2018; Otten and Donchin, 2000). The mean and range of trial numbers per condition were as following: $M = 37$, range 14–51 (HT hits); $M = 19$, range 9–45 (HT misses); $M = 39$, range 17–53 (LT hits); $M = 17$, range 7–36 (LT misses); $M = 26$, range 11–55 (IC hits); $M = 85$, range 38–105 (IC misses). Grand average waveforms were low-pass filtered at 12 Hz for illustration purposes only.

2.6. Statistical analyses

All statistical analyses were conducted using IBM SPSS Statistics 22.0 software at a 5% significance level. Behavioral measures and ERP mean amplitudes were analyzed using repeated measures ANOVAs and dependent t -tests. Greenhouse-Geisser corrected degrees of freedom and p -values are reported whenever the assumption of sphericity was violated. Significant effects were decomposed using lower level ANOVAs and dependent t -tests. As measures of effect sizes, partial eta squared (η_p^2) are reported for ANOVA results. For independent t -tests, Cohen's d was calculated. For dependent t -tests, d was calculated according to Dunlap et al. (1996), taking into account the correlations between measurements.

3. Results

3.1. Behavioral results

3.1.1. Congruency ratings

Congruency ratings for the three predefined levels of Congruency (HT, LT, IC) differed significantly, $F(2,46) = 1835.79$, $p < .001$, $\eta_p^2 = .99$, whereby dependent t -tests revealed that the perceived congruency was higher for HT than for LT exemplars, $t(23) = 11.08$, $p < .001$, $d = 2.02$, and higher for LT than for IC exemplars, $t(23) = 41.55$, $p < .001$, $d = 10.18$. The subjects' perception of the fit between category cues and exemplars therefore corresponded to the assignment of exemplars to the congruency conditions.

3.1.2. Item recognition

A one-way ANOVA across the three levels of the Congruency factor revealed that Pr scores differed as a function of Congruency, $F(2,46) = 86.17$, $p < .001$, $\eta_p^2 = .79$. Paired t -tests showed that congruent exemplars were remembered better than incongruent ones (HT vs. IC, $t(23) = 11.16$, $p < .001$, $d = 1.33$; LT vs. IC, $t(23) = 10.87$, $p < .001$, $d = 1.71$), whereas HT and LT exemplars did not differ, $t(23) = 1.53$, $p = .14$, $d = 0.18$. High confidence Pr scores also differed as a function of Congruency, $F(2,46) = 115.26$, $p < .001$, $\eta_p^2 = .83$, and subsidiary t -tests yielded a similar pattern as for the Pr scores: HT vs. IC, $t(23) = 12.03$, $p < .001$, $d = 1.41$; LT vs. IC, $t(23) = 12.95$, $p < .001$, $d = 1.96$; HT vs. LT, $t(23) = 2.04$, $p = .05$, $d = 0.22$. Accordingly,

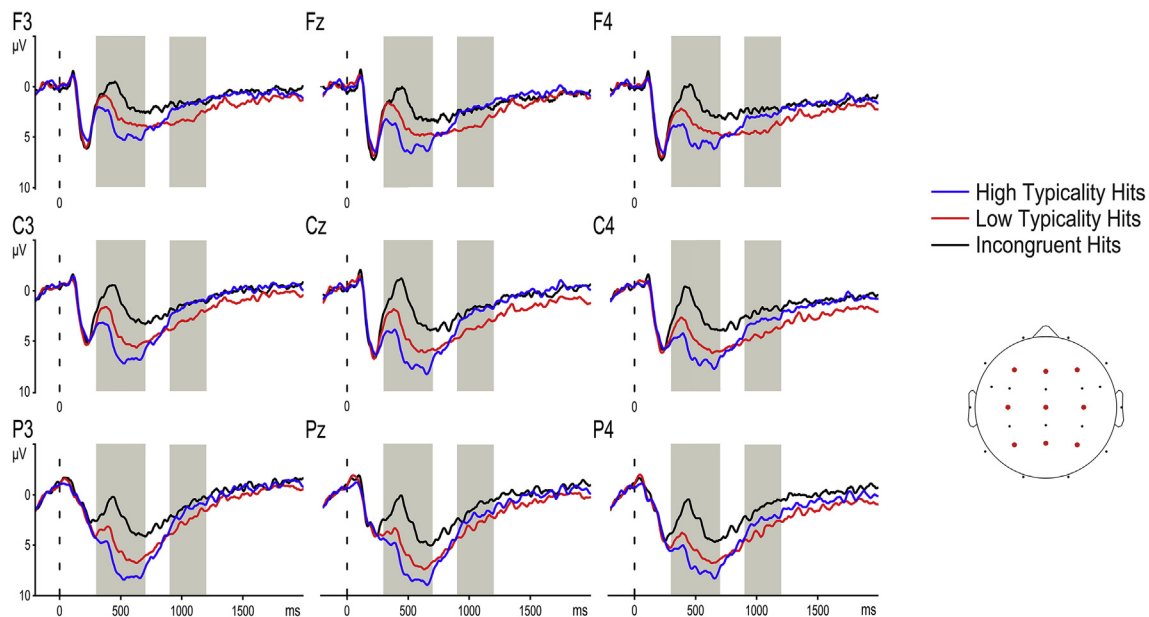


Fig. 1. Waveforms for subsequently remembered HT, LT, and IC exemplars at representative scalp electrodes. The N400 (300–500 ms), post-N400 (500–700 ms), and frontal slow wave (900–1200 ms) time windows are marked in grey.

congruent exemplars were remembered better than incongruent exemplars, whereas typicality did not additionally modulate remembrance.

3.1.3. Associative recognition

Associative decisions showed a strong bias in favor of choosing the congruent category cue. Therefore, we calculated and analyzed hit rates that were corrected for the probability of choosing the correct category cue by guessing (see 2.4 for details). A one-way ANOVA with the factor Congruency (HT, LT, IC) showed that after correcting for guessing there were no significant differences between the congruency conditions, $F(1.351, 31.07) < 1$, $p_{\text{corr}} = .59$, $\eta_p^2 = .02$.

3.2. ERP results

Fig. 1 shows the grand average ERP waveforms elicited by subsequently remembered words in the three congruency conditions (HT, LT, and IC hits). Congruency effects for the N400 emerged around 300 ms and were largest at posterior electrode sites. In this time interval, amplitudes were most negative for IC hits, intermediate for LT and most positive for HT hits. A positive deflection that was most pronounced at parietal electrode sites followed at around 500 ms. In a later time window starting around 900 ms, LT hits exhibited a sustained positivity relative to HT and IC hits that persisted until approximately 1200 ms. This effect was more pronounced at frontal than at parietal electrode sites.

Furthermore, Fig. 2, depicting the subsequent memory effect (SME) for the three conditions, shows that across all three conditions, subsequently remembered exemplars elicited more positive-going waveforms than subsequently forgotten ones. At frontal electrodes, an SME emerged for all three congruency conditions around 300 ms and persisted until the end of the epoch. In contrast, qualitative differences of the SME between the two congruent and the incongruent condition were evident at parietal electrodes, where SMEs were obtained for congruent (HT and LT) exemplars. For incongruent exemplars subsequent hits and misses did not differ.

These observations were confirmed by a series of statistical analyses. Mean amplitudes were analyzed in the 300–500 ms time window, where N400 effects are typically largest (Kutas and Federmeier, 2011), and in the adjacent 500–700 ms time window. These two time windows

were chosen for the analyses of memory and congruency effects because they cover the early time window (370–680 ms) in which Packard et al. (2017) found an SME for semantically congruent, but not incongruent words. As subsequent memory effects are usually largest at anterior recordings sites (Otten and Donchin, 2000), especially under deep (semantic) encoding conditions or when elaborative encoding strategies are used (Kamp et al., 2017; Mecklinger and Müller, 1996; Otten and Rugg, 2001; Van Petten and Senkfor, 1996), whereas effects of semantic congruency for the N400 are most pronounced at posterior electrodes (Kutas and Federmeier, 2011), the electrode montage consisted of 12 electrodes that cover anterior and posterior brain regions, divided into two electrode clusters (anterior: F3, Fz, F4, FC3, FCz, FC4; posterior: CP3, CPz, CP4, P3, Pz, P4). As previous ERP studies exploring the processing of contextual information have consistently found that the late positivity is largest at anterior electrode sites (DeLong et al., 2011; Federmeier et al., 2010, 2007; Thornhill and Van Petten, 2012; see Van Petten and Luka, 2012, for a review), the late frontal positivity was analyzed between 900 and 1200 ms in the anterior electrode cluster.

3.2.1. N400 (300–500 ms) and post-N400 (500–700 ms) time windows

An ANOVA including the factors Time Window (300–500 ms, 500–700 ms), Antpos (anterior, posterior), Congruency (HT, LT, IC), and Memory (hits, misses) yielded significant main effects of Congruency, $F(2,46) = 28.33$, $p < .001$, $\eta_p^2 = .55$, and Memory, $F(1,23) = 13.56$, $p < .01$, $\eta_p^2 = .37$. No significant interaction with the factor Time Window occurred, but there were interactions between Congruency and Antpos, $F(2,46) = 4.02$, $p < .05$, $\eta_p^2 = .15$, and Memory and Antpos, $F(1,23) = 7.78$, $p < .05$, $\eta_p^2 = .25$. The three-way interaction between Antpos, Congruency, and Memory was not significant, $F(2,46) = 1.31$, $p = .28$, $\eta_p^2 = 0.05$. To follow up the significant Congruency by Antpos and Memory by Antpos interactions, mean amplitudes in both time windows were analyzed separately at anterior and posterior electrodes.

3.2.1.1. 300–700 ms, anterior electrodes. At anterior electrodes, mean amplitudes differed as a function of Congruency, $F(2,46) = 17.95$, $p < .001$, $\eta_p^2 = .44$. Mean amplitudes were more negative for incongruent exemplars than for congruent exemplars (HT: $M = 3.86$, $SEM = 0.92$ μV ; LT: $M = 2.64$, $SEM = 0.93$ μV ; IC: $M = 0.95$, $SEM = 0.70$ μV ; IC vs. LT: $t(23) = 3.40$, $p < .01$, $d = 0.38$; IC vs. HT: $t(23) = 6.27$,

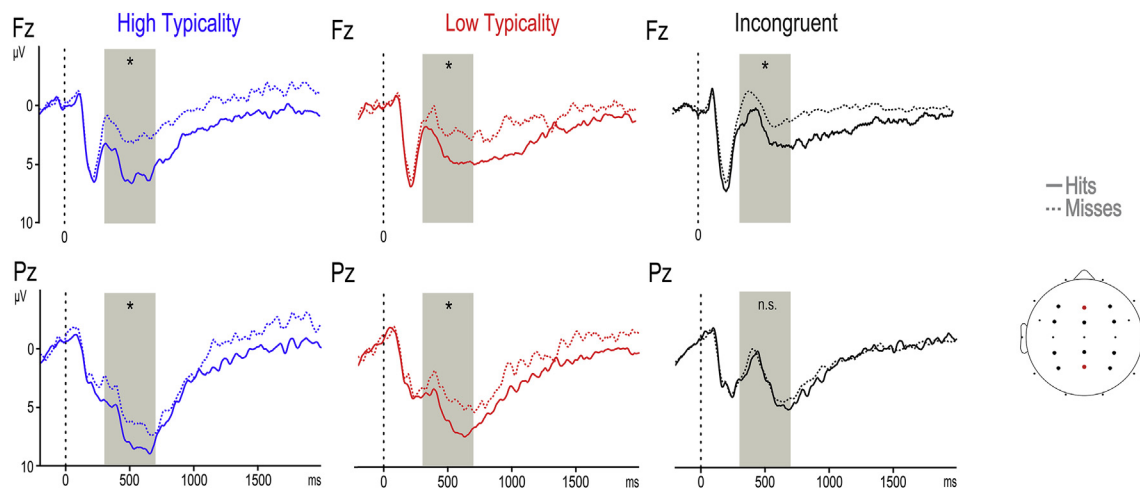


Fig. 2. Subsequent Memory Effects at two electrodes representative for the frontal and parietal electrode clusters (Fz and Pz), separate for congruent (HT, LT) and incongruent (IC) exemplars. Conditions in which significant SME were obtained are marked with an asterisk (n.s. = not significant).

$p < .001$, $d = 0.65$), and more negative for low typicality than for high typicality exemplars, $t(23) = 2.44$, $p < .05$, $d = 0.27$. A significant main effect of Memory, $F(1,23) = 14.64$, $p < .01$, $\eta_p^2 = .39$, indicated that subsequently remembered exemplars ($M = 3.41$, $SEM = 0.88 \mu V$) were associated with more positive waveforms than subsequent misses ($M = 1.56$, $SEM = 0.81 \mu V$). No Congruency by Memory interaction was obtained.

3.2.1.2. 300–700 ms, posterior electrodes. At posterior electrodes, significant main effects of Congruency, $F(2,46) = 37.59$, $p < .001$, $\eta_p^2 = .62$, and Memory, $F(1,23) = 11.25$, $p < .01$, $\eta_p^2 = .33$, were obtained. Notably, the main effect of Memory, was qualified by a significant Congruency by Memory interaction, $F(2,46) = 3.36$, $p < .05$, $\eta_p^2 = .13$. To disentangle this interaction, subsequent memory effects were calculated as the difference between subsequent hits and misses for each level of the Congruency factor. SME differed from zero for congruent exemplars (HT: $M = 1.89$, $SEM = 0.63 \mu V$, $t(23) = 2.98$, $p < .01$, $d = 0.61$; LT: $M = 1.77$, $SEM = 0.54 \mu V$, $t(23) = 3.29$, $p < .01$, $d = 0.67$), but not for incongruent exemplars ($M = 0.28$, $SEM = 0.49 \mu V$, $t(23) = 0.58$, $p = .57$, $d = 0.12$). Paired t -tests revealed that the SME was larger for congruent than for incongruent exemplars (HT vs. IC: $t(23) = 2.87$, $p < .01$, $d = 0.57$; LT vs. IC: $t(23) = 2.30$, $p < .05$, $d = 0.59$), but did not differ between high and low typicality exemplars, $t(23) = 0.14$, $p = .89$, $d = 0.04$.

In sum, while only main effects of Congruency and Memory were obtained at anterior recording sites, in the posterior cluster SME were obtained for congruency exemplars but not for incongruent ones in the 300–700 ms time period, as reflected by a Memory by Congruency interaction.

3.2.2. Late frontal positivity time window (900–1200 ms)

As evident from Fig. 1, congruency effects in this late time interval showed a different pattern than in the earlier 300–700 ms epoch. At frontal electrodes, low typicality hits were associated with more positive-going waveforms than high typicality and incongruent hits. To confirm our assumption that subsequently remembered unexpected exemplars should be associated with strong expectancy mismatch-related processing as reflected in the late frontal positivity, mean amplitudes in the anterior electrode cluster (see 3.2) for subsequent hits were analyzed in a one-way ANOVA that yielded a significant main effect of Congruency, $F(2,46) = 5.77$, $p < .01$, $\eta_p^2 = .20$. Subsidiary t -tests revealed that LT hits exhibited more positive-going waveforms than HT and IC hits (HT hits: $M = 2.28$, $SEM = 0.86 \mu V$; LT hits: $M = 3.73$, $SEM = 0.78 \mu V$; IC hits: $M = 1.88$, $SEM = 0.84 \mu V$; LT vs. IC hits: $t(23) = 3.33$, $p < .01$, $d = 0.46$; LT vs. HT hits: $t(23) = 2.98$,

$p < .01$, $d = 0.36$; HT vs. IC hits: $t(23) = 0.60$, $p = .56$, $d = 0.10$).

If expectancy mismatch-related processing as reflected in the frontal positivity involves the suppression of the representations of strongly expected exemplars, the accessibility of these exemplars in the subsequent test phase may be lowered. As a consequence, the likelihood of false positive memory decisions for these words in the subsequent memory test may be reduced. Thus, after having confirmed that subsequently remembered LT exemplars were on average associated with a larger late frontal positivity than HT and IC hits, a complementary analysis was conducted to test the assumption that the late frontal positivity was predictive of individual differences in recognition memory performance.

Correlations were computed between standardized mean amplitudes at electrode Fz in the 900–1200 ms time interval, and behavioral measures of memory performance (hit rates and false alarm rates). Three participants had to be excluded from this analysis because their mean amplitudes deviated by more than two standard deviations from the sample mean. As evident from Fig. 3, the late frontal positivity elicited by subsequent hit responses to congruent exemplars (HT and LT), but not IC exemplars showed significant negative correlations with the false alarm rate. In other words: the more pronounced the late positive slow wave to successfully encoded congruent (LT and HT) events, the less false positive memory responses occurred in the ensuing test phase. No such correlations were obtained for hit responses. The complete pattern of correlations is listed in Table 3.

4. Discussion

In the present study, ERPs were used to investigate the processes by which activated schemata support the encoding of new events. Based on the previous literature, it was predicted that schema-congruent information should be remembered better than incongruent information (van Kesteren et al., 2013a). A second aim was to explore the role of expectancy during the processing of schema-congruent events. The assumption was that activated schema knowledge leads to the built up of expectancies and that in particular the disconfirmation of these expectancies by plausible, but unexpected events is associated with strong expectancy mismatch-related processing at a post-encoding stage that boosts memory performance for these events (Federmeier et al., 2007). To assess the role of expectancy mismatches, memory for events which were plausible but of low typicality in a given context was explored and contrasted with memory for high typicality events.

Importantly, incongruent exemplars were not expected to elicit strong expectancy mismatches. Previous studies comparing memory for semantically congruent and incongruent word pairs have consistently

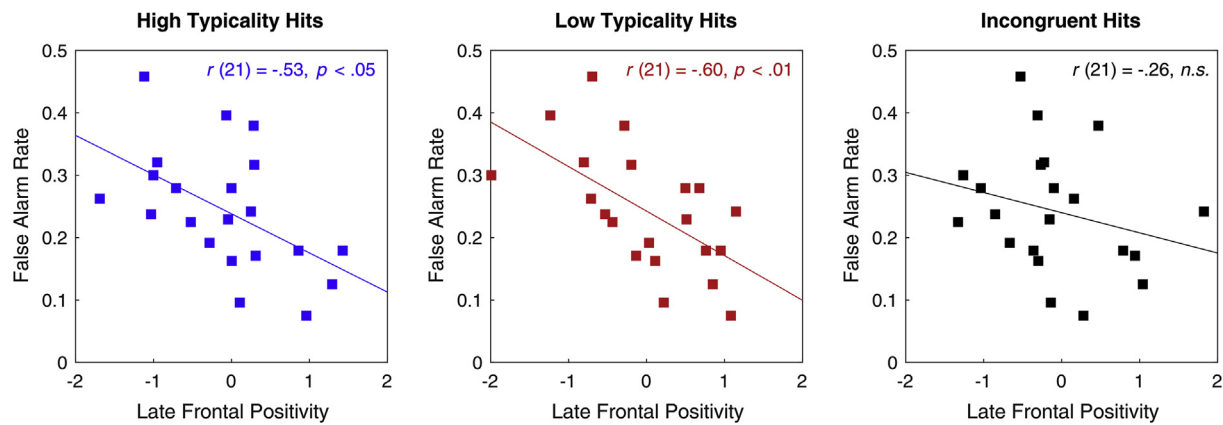


Fig. 3. Scatterplots of the correlations between mean amplitudes in the late frontal slow wave time window (900–1200 ms) at electrode Fz and the false alarm rate.

Table 3

Pearson's correlation coefficients between Hit/ False alarm rates and standardized ERP mean amplitudes between 900 - 1200 ms at electrode Fz. Note: * = $p < .05$, ** = $p < .01$, two-tailed.

	HR HT	HR LT	HR IC	FAR
Fz HT Hits	.16	-.07	-.07	-.53*
Fz LT Hits	.13	-.21	-.21	-.60**
Fz IC Hits	-.06	-.14	-.22	-.26

found better memory for congruent events (e.g., Bein et al., 2015; Schulman, 1974). Even though expectancy mismatches associated with incongruent events can also benefit learning and memory (see Brod et al., 2018 and Greve et al., 2017, for two recent examples), at least in the case of semantic incongruency this effect is strongly contingent on the distinctiveness of these events (Reggev et al., 2017). Because congruent and incongruent items were presented equally often in the present study, memory for incongruent exemplars could not have benefited from a higher distinctiveness of the latter items. Also, our procedure to define incongruent events by rearranging words which originally belonged to different studied categories may have additionally lowered expectancy mismatches for incongruent trial in the present study. Thus, incongruent exemplars should best be regarded as a control condition in which the activated schema did not support memory encoding.

In the study phase, participants rated the congruency relation between the exemplar and the semantic category cue. The rating results confirm that the perceived fit between exemplars and categories corresponded to the experimental congruency conditions. To make sure that memory decision were not based on semantic category information but were reliant on detailed episodic memory for a particular event, studied items had to be discriminated from semantically related lures in a recognition memory test that was conducted one day after the study phase. Consistent with previous studies investigating schema-based encoding (e.g. Bein et al., 2014; van Kesteren et al., 2013a), congruent (high and low typicality) exemplars were associated with superior memory compared to incongruent exemplars. Importantly, and different from the findings of Packard et al. (2017), congruent exemplars were clearly associated with higher memory accuracy than incongruent exemplars when correcting for false alarms. This suggests that schema congruency may not have induced a semantic decision bias (i.e., respond old whenever a word from a study category is presented). Rather, memory for words learned in a semantically congruent context may have benefited from a preferential encoding of these words, leading to an increased accessibility of these traces in the subsequent memory test. Unexpectedly, we did not find that low typicality exemplars were remembered better than high typicality ones, a result which was reported by Federmeier et al. (2007) who used sentence contexts to induce

expectancies for the sentence final word (e.g., “There were brightly colored pictures on every”). It is conceivable that the category exemplars used in the present study did not elicit expectancy violations as strong as those provoked by unexpected words in strongly constraining sentences, presumably because the category cue phrases (e.g., “A four-footed animal”, see Table 1) were shorter and less likely to induce as strong expectations as the sentence frames used in the Federmeier et al. (2007) study.

Previous studies have found that schematic encoding benefits associative memory more than item memory (van Kesteren et al., 2013). In contrast, we found strong effects of semantic congruency on item recognition for exemplar words, but no differences regarding memory for associations between exemplars and category cues.¹ This may be due to the fact that, different from previous studies, items and category cues were not presented simultaneously in the study phase. Instead, the exemplar word was always presented several seconds after the presentation of the category cue, which may have fostered the encoding of item-specific details and dampened the processing of the item-category cue relationship. Consistent with this view, ERPs elicited by congruent words showed a parietal SME that has been associated with the encoding of item-specific details, a process that heightens the distinctiveness of an item and thereby enhances its recall (Fabiani et al., 1986; Kamp et al., 2017; Karis et al., 1984). Thus, the results of the present study indicate that differential effects of schematic encoding on item and associative recognition may at least partly depend on the parameters of the encoding situation.

ERPs elicited by the onset of the exemplar words in the study phase showed a typical N400 semantic congruency effect that emerged around 300 ms. It was broadly distributed but most pronounced over posterior electrode sites. Consistent with previous studies (Federmeier et al., 2010; Heinze et al., 1998), N400 amplitudes decreased linearly with increasing semantic congruency. Incongruent exemplars elicited the largest N400, low typicality exemplars intermediate, and high typicality exemplars smallest N400 amplitudes.

In contrast to Packard et al. (2017), where an SME emerged ca. 200 ms earlier for semantically congruent words, in the present study SME elicited by congruent and incongruent exemplars were present in

¹ Recently, Brod and Shing (2019) proposed a schema bias score that reflects the influence of prior knowledge on associative memory decisions. To examine the extent to which the participants' memory decisions relied on prior knowledge, we calculated the schema bias score similar to Brod and Shing (2019) as the proportion of erroneously chosen congruent category cues in the incongruent condition for correctly recognized words (item hits). We found that the mean schema bias score was 0.70 (SEM = 0.03) and by this significantly larger than one would expect if memory decision were unbiased (0.50), $t(23) = 5.91$, $p < .001$, $d = 1.21$, indicating that as in the Brod and Shing (2019) study, associative memory decisions were strongly biased by prior knowledge.

the 300–700 ms time period at frontal recording sites and showed highly similar temporal and topographical characteristics. This finding is consistent with a number of studies that have found frontally-distributed SME with an onset between 300 and 400 ms (Kamp et al., 2017; Mecklinger and Müller, 1996; Otten and Donchin, 2000; Otten and Rugg, 2001; Van Petten and Senkfor, 1996) and challenges the view that semantic congruency accelerates the onset of memory formation (Packard et al., 2017). Our results rather suggest that the processes reflected by the early frontal SME contributed in a highly similar way to the formation of memory traces for schema-congruent and -incongruent events. Notably, as congruent (HT and LT) exemplars were associated with higher hit rates than incongruent exemplars, the highly similar pattern of SME across congruency conditions cannot account for the superior memory performance for congruent exemplars.

In further contrast to the Packard et al. (2017) study, in which SME for semantically congruent and incongruent words showed a temporal dissociation, the present study revealed a topographical difference of the SME as a function of semantic congruency. While the SME did not differ between congruent and incongruent events at anterior electrodes, there was an SME in the same epoch which was most prominent over parietal electrode sites. As revealed by a Memory by Congruency interaction, this SME was specific for congruent exemplars and virtually absent for incongruent ones. Parietal subsequent memory effects of similar kind are thought to occur when episodic details are successfully encoded and memory for items is probed at test, and less so for the successful encoding of relational information (Fabiani et al., 1986; Kamp et al., 2017; Karis et al., 1984). The pattern of results therefore suggests that activated schemata supported memory formation for schema-congruent exemplars by enhancing the successful encoding of item specific details. This kind of online memory formation may render single items distinctive in later memory test and boost memory performance, as evident from the superior memory accuracy for congruent events.

Of note, in this study, subsequent hits were defined as target words that were correctly judged as “old” and for which the correct category cue was chosen. This procedure has the advantage of reducing the number of correctly guessed “old” decisions included in the averaged ERP waveform. Because participants showed a strong bias in favor of choosing the congruent category cue, congruent hits presumably contained more correct category guesses than incongruent hits. Importantly, correct and incorrect category guesses should not differ at the electrophysiological level. Thus, if a higher proportion of correct category guesses affected the ERP data, then this should have reduced the difference between subsequent hits and misses in the congruent conditions. This is not what we found: The early frontal SME was highly similar in the congruent and incongruent conditions, and the parietal SME was specific for the congruent conditions. Because these findings are hard to reconcile with the view outlined above, we feel safe to conclude that the higher proportion of category guesses in the congruent conditions probably did not bias the pattern of ERP results. If, however, the ERP results are biased by a higher proportion of correct category guesses in the congruent conditions, then this would mean that the SME obtained in the congruent conditions were underestimated in the present study.

Consistent with previous studies, ERPs elicited by subsequently remembered low typicality exemplars were associated with a late frontal positivity that is thought to reflect expectancy mismatch-related processing at a post-encoding stage that is caused by the discrepancy between a constraining context and an unexpected though plausible event (Federmeier et al., 2010, 2007). Interestingly, a complementary correlation analysis revealed that the late frontal positivity elicited by successfully encoded congruent events showed a negative correlation with the false alarm rate in the test phase. This indicates that the higher the involvement of the processes reflected in the late frontal positivity, the better are participants able to avoid false positive memory decisions by detecting semantically related lures. This finding is consistent with

functional interpretations of the late frontal positivity which assume that it reflects the contextual integration of an unexpected event by the suppression of strongly expected ones (DeLong et al., 2011; Kutas, 1993; Van Petten and Luka, 2012).

The notion that the late frontal positivity reflects the suppression of expected events implies that the accessibility of these events should be reduced in a memory test performed after the assumed suppression. Research on retrieval-induced forgetting provides ample evidence for this view (Hanslmayr et al., 2010; Johansson et al., 2007; Staudigl et al., 2010). Interestingly, however, whereas ERPs indicated that only low typicality exemplars elicited a large frontal positivity, the analysis of individual differences showed that the negative correlation pattern between the frontal positivity and the false alarm rate was less specific: The positivity to both low and high typicality events was correlated with false memory responses. As in the present task all lures were semantically related to the studied categories, this finding suggests that the processes reflected by the late frontal positivity do not lead to a selective suppression of memory representations of highly expected events but – more likely – attenuate false memory responses by means of a general suppression of semantically related representations at the post-encoding stage.

Taken together, the ERP results suggest that two processes contributed to the pattern of memory results observed in the present study: Firstly, activated schemata have a strong positive effect on the encoding of schema-congruent item-specific details and make the memory representation of these newly encoded items more accessible in later memory tests. This process was reflected by the increased hit rates for schema congruent exemplars and electrophysiologically supported by the parietal SME, a measure of successful item-specific encoding, which was selectively revealed for congruent events. Secondly, the processing of unexpected exemplars fosters expectancy mismatch-related processing and enhances the contextual integration of unexpected exemplars presumably through the suppression of semantically related exemplars, a processing stage that could be reflected in the late frontal positivity. A correlation analysis revealed that the greater the involvement of the processes reflected by the frontal positivity, the lower the level of false positive memory responses in the ensuing test phase. The present study thereby provides further evidence in support of the view that the contextual integration of unexpected, but plausible words involves a suppression of the representations of semantically related, but unstudied words in memory. This inhibitory process lowers the accessibility of semantically related words in memory and leads to a lower rate of false positive memory decisions for these words. The results of the correlation analysis are preliminary and further studies are needed for a thorough functional characterization of the late frontal positivity and the detailed mechanisms by which this type of expectancy mismatch-related processing affects memory formation for unexpected events.

5. Declaration of interest

None.

CRedit authorship contribution statement

Gerrit Hölftje: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration. **Bente Lubahn:** Methodology, Formal analysis, Investigation, Resources, Writing - original draft, Visualization. **Axel Mecklinger:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Supervision, Funding acquisition.

Acknowledgments

We wish to thank Yoana Vergilova and two anonymous reviewers for helpful comments on earlier versions of the manuscript. This study

was funded by the German Research Foundation as part of the collaborative research center SFB 1102 “Information Density and Linguistic Encoding”, project A6.

References

- Battig, W.F., Montague, W.E., 1969. Category norms of verbal items in 56 categories: A replication and extension of the Connecticut category norms. *J. Exp. Psychol.* 80 (3), 1–46. <https://doi.org/10.1037/h0027577>.
- Bein, O., Livneh, N., Reggev, N., Gilead, M., Goshen-Gottstein, Y., Maril, A., 2015. Delineating the effect of semantic congruency on episodic memory: the role of integration and relatedness. *PLoS One* 10 (2), 1–24. <https://doi.org/10.1371/journal.pone.0115624>.
- Bein, O., Reggev, N., Maril, A., 2014. Prior knowledge influences on hippocampus and medial prefrontal cortex interactions in subsequent memory. *Neuropsychologia* 64, 320–330. <https://doi.org/10.1016/j.neuropsychologia.2014.09.046>.
- Brod, G., Hasselhorn, M., Bunge, S.A., 2018. When generating a prediction boosts learning: the element of surprise. *Learn. Instr.* 55, 22–31. <https://doi.org/10.1016/j.learninstruc.2018.01.013>.
- Brod, G., Lindenberger, U., Werkle-Bergner, M., Shing, Y.L., 2015. Differences in the neural signature of remembering schema-congruent and schema-incongruent events. *Neuroimage* 117, 358–366. <https://doi.org/10.1016/j.neuroimage.2015.05.086>.
- Brod, G., Shing, Y.L., 2019 February 25. A Boon and a Bane: Comparing the Effects of Prior Knowledge on Memory across the Lifespan. *Developmental Psychology*. Advance online publication. <https://doi.org/10.1037/dev0000712>.
- Cohen, N., Pell, L., Edelson, M.G., Ben-Yakov, A., Pine, A., Dudai, Y., 2015. Peri-encoding predictors of memory encoding and consolidation. *Neurosci. Biobehav. Rev.* 50, 128–142. <https://doi.org/10.1016/j.neubiorev.2014.11.002>.
- Deese, J., 1959. On the prediction of occurrence of particular verbal intrusions in immediate recall. *J. Exp. Psychol.* 58 (1), 17–22. <https://doi.org/10.1037/h0046671>.
- DeLong, K.A., Urbach, T.P., Groppe, D.M., Kutas, M., 2011. Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology* 48 (9), 1203–1207. <https://doi.org/10.1111/j.1469-8986.2011.01199.x>.
- Dunlap, W.P., Cortina, J.M., Vaslow, J.B., Burke, M.J., 1996. Meta-analysis of experiments with matched groups or repeated measures designs. *Psychol. Methods* 1 (2), 170–177. <https://doi.org/10.1037/1082-989X.1.2.170>.
- Durrant, S.J., Cairney, S.A., McDermott, C., Lewis, P.A., 2015. Schema-conformant memories are preferentially consolidated during REM sleep. *Neurobiol. Learn. Mem.* 122, 41–50. <https://doi.org/10.1016/j.nlm.2015.02.011>.
- Fabiani, M., Karis, D., Donchin, E., 1986. P300 and recall in an incidental memory paradigm. *Psychophysiology*. <https://doi.org/10.1111/j.1469-8986.1986.tb00636.x>.
- Federmeier, K.D., Kutas, M., Schul, R., 2010. Age-related and individual differences in the use of prediction during language comprehension. *Brain Lang.* 115 (3), 149–161. <https://doi.org/10.1016/j.bandl.2010.07.006>.
- Federmeier, K.D., Wlotko, E.W., De Ochoa-Dewald, E., Kutas, M., 2007. Multiple effects of sentential constraint on word processing. *Brain Res.* 1146 (1), 75–84. <https://doi.org/10.1016/j.brainres.2006.06.101>.
- Ghosh, V.E., Gilboa, A., 2014. What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia* 53 (1), 104–114. <https://doi.org/10.1016/j.neuropsychologia.2013.11.010>.
- Gilboa, A., Marlatte, H., 2017. Neurobiology of schemas and schema-mediated memory. *Trends Cognit. Sci.* 21 (8), 618–631. <https://doi.org/10.1016/j.tics.2017.04.013>.
- Greve, A., Cooper, E., Kaula, A., Anderson, M.C., Henson, R., 2017. Does prediction error drive one-shot declarative learning? *J. Mem. Lang.* 94, 149–165. <https://doi.org/10.1016/j.jml.2016.11.001>.
- Hanslmayr, S., Staudigl, T., Aslan, A., Bäuml, K.H., 2010. Theta oscillations predict the detrimental effects of memory retrieval. *Cognit. Affect. Behav. Neurosci.* 10 (3), 329–338. <https://doi.org/10.3758/CABN.10.3.329>.
- Heinze, H.J., Muentze, T.F., Kutas, M., 1998. Context effects in a category verification task as assessed by event-related brain potential (ERP) measures. *Biol. Psychol.* 47 (2), 121–135. [https://doi.org/10.1016/S0301-0511\(97\)00024-0](https://doi.org/10.1016/S0301-0511(97)00024-0).
- Heister, J., Würzner, K.M., Bubenzner, J., Pohl, E., Hanneforth, T., Geyken, A., Kliegl, R., 2011. dlexDB - Eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychol. Rundsch.* 62 (1), 10–20. <https://doi.org/10.1026/0033-3042/a000029>.
- Hellerstedt, R., Rasmussen, A., Johansson, M., 2012. Swedish category norms, third ed. *Lund Psychological Reports, vol. 12 Department of Psychology, Lund University*.
- Johansson, M., Aslan, A., Bäuml, K.H.T., Gäbel, A., Mecklinger, A., 2007. When remembering causes forgetting: electrophysiological correlates of retrieval-induced forgetting. *Cerebr. Cortex* 17 (6), 1335–1341. <https://doi.org/10.1093/cercor/bhl044>.
- Kamp, S.-M., Bader, R., Mecklinger, A., 2017. ERP subsequent memory effects differ between inter-item and utilization encoding tasks. *Front. Hum. Neurosci.* 11, 30. <https://doi.org/10.3389/fnhum.2017.00030>.
- Kamp, S.-M., Bader, R., Mecklinger, A., 2018. Utilization of word pairs in young and older adults: encoding mechanisms and retrieval outcomes. *Psychol. Aging* 33 (3), 497–511. <https://doi.org/10.1037/pag0000256>.
- Karis, D., Fabiani, M., Donchin, E., 1984. “P300” and memory: individual differences in the von restorff effect. *Cogn. Psychol.* 16 (2), 177–216. [https://doi.org/10.1016/0010-0285\(84\)90007-0](https://doi.org/10.1016/0010-0285(84)90007-0).
- Kutas, M., 1993. In the company of other words: electrophysiological evidence for single-word and sentence context effects. *Lang. Cogn. Process.* 8 (4), 533–572. <https://doi.org/10.1080/01690969308407587>.
- Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. <https://doi.org/10.1146/annurev.psych.093008.131123>.
- Mannhaupt, H.-R., 1983. Produktionsnormen für verbale Reaktionen zu 40 geläufigen Kategorien. *Sprache & Kognition* 2 (4), 264–278.
- Mecklinger, A., Müller, N., 1996. Dissociations in the processing of “what” and “where” information in working memory: an event-related potential analysis. *J. Cogn. Neurosci.* 8 (5), 453–473. <https://doi.org/10.1162/jocn.1996.8.5.453>.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9 (1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4).
- Otten, L.J., Donchin, E., 2000. Relationship between P300 amplitude and subsequent recall for distinctive events: dependence on type of distinctiveness attribute. *Psychophysiology* 37 (5), 644–661. <https://doi.org/10.1017/S004857720098171X>.
- Otten, L.J., Rugg, M.D., 2001. Electrophysiological correlates of memory encoding are task-dependent. *Cogn. Brain Res.* 12 (1), 11–18. [https://doi.org/10.1016/S0926-6410\(01\)00015-5](https://doi.org/10.1016/S0926-6410(01)00015-5).
- Packard, P.A., Rodríguez-Fornells, A., Bunzeck, N., Nicolás, B., de Diego-Balaguer, R., Fuentes-Milla, L., 2017. Semantic congruence accelerates the onset of the neural signals of successful memory encoding. *J. Neurosci.* 37 (2), 291–301. <https://doi.org/10.1523/JNEUROSCI.1622-16.2017>.
- Paller, K.A., Wagner, A.D., 2002. Observing the transformation of experience into memory. *Trends Cognit. Sci.* 6 (2), 93–102. [https://doi.org/10.1016/S1364-6613\(00\)01845-3](https://doi.org/10.1016/S1364-6613(00)01845-3).
- Preston, A.R., Eichenbaum, H., 2013. Interplay of hippocampus and prefrontal cortex in memory. *Curr. Biol.* 23 (17), R764–R773. <https://doi.org/10.1016/j.cub.2013.05.041>.
- Reggev, N., Sharoni, R., Maril, A., 2017. Distinctiveness benefits novelty (and not familiarity), but only up to a limit: the prior knowledge perspective. *Cogn. Sci.* 1–26. <https://doi.org/10.1111/cogs.12498>.
- Roediger, H.L., McDermott, K.B., 1995. Creating false memories: remembering words not presented in lists. *J. Exp. Psychol. Learn. Mem. Cogn.* 21 (4), 803–814. <https://doi.org/10.1037/0278-7393.21.4.803>.
- Sanquist, T.F., Rohrbaugh, J.W., Sydulko, K., Lindsley, D.B., 1980. Electrocorical signals of levels of processing: perceptual analysis and recognition memory. *Psychophysiology* Retrieved from papers3://publication/uuid/81D4F0D4-6FCD-4D6B-83C8-71282783EF24.
- Schulman, A.L., 1974. Memory for words recently classified. *Mem. Cogn.* 2 (1A), 47–52.
- Snodgrass, J.G., Corwin, J., 1988. Pragmatics of measuring recognition memory: applications to dementia and amnesia. *J. Exp. Psychol. Gen.* 117 (1), 34–50. <https://doi.org/10.1037/0096-3445.117.1.34>.
- Spalding, K.N., Jones, S.H., Duff, M.C., Tranel, D., Warren, D.E., 2015. Investigating the neural correlates of schemas: ventromedial prefrontal cortex is necessary for normal schematic influence on memory. *J. Neurosci.* 35 (47), 15746–15751. <https://doi.org/10.1523/JNEUROSCI.2767-15.2015>.
- Staresina, B.P., Gray, J.C., Davachi, L., 2009. Event congruency enhances episodic memory encoding through semantic elaboration and relational binding. *Cerebr. Cortex* 19 (5), 1198–1207. <https://doi.org/10.1093/cercor/bhn165>.
- Staudigl, T., Hanslmayr, S., Bäuml, K.-H.T., 2010. Theta oscillations reflect the dynamics of interference in episodic memory retrieval. *J. Neurosci.* 30 (34), 11356–11362. <https://doi.org/10.1523/JNEUROSCI.0637-10.2010>.
- Thornhill, D.E., Van Petten, C., 2012. Lexical versus conceptual anticipation during sentence processing: frontal positivity and N400 ERP components. *Int. J. Psychophysiol.* 83 (3), 382–392. <https://doi.org/10.1016/j.ijpsycho.2011.12.007>.
- Tibon, R., Cooper, E., Greve, A., 2017. Does semantic congruency accelerate episodic encoding, or increase semantic elaboration? *J. Neurosci.* 37 (19), 4861–4863. <https://doi.org/10.1523/JNEUROSCI.0570-17.2017>.
- Tse, D., Langston, R.F., Kakeyama, M., Bethus, I., Spooner, P.A., Wood, E.R., et al., 2007. Schemas and memory consolidation. *Science* 316 (5821), 76–82. <https://doi.org/10.1126/science.1135935>.
- van der Linden, M., Berkers, R.M.W.J., Morris, R.G.M., Fernández, G., 2017. Angular gyrus involvement at encoding and retrieval is associated with durable but less specific memories. *J. Neurosci.* 37 (39), 3603–3616. <https://doi.org/10.1523/JNEUROSCI.3603-16.2017>.
- van Kesteren, M.T.R., Beul, S.F., Takashima, A., Henson, R.N., Ruiters, D.J., Fernández, G., 2013a. Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: from congruent to incongruent. *Neuropsychologia* 51 (12), 2352–2359. <https://doi.org/10.1016/j.neuropsychologia.2013.05.027>.
- van Kesteren, M.T.R., Rijpkema, M., Ruiters, D.J., Fernández, G., 2010. Retrieval of associative information congruent with prior knowledge is related to increased medial prefrontal activity and connectivity. *J. Neurosci.* 30 (47), 15888–15894. <https://doi.org/10.1523/JNEUROSCI.2674-10.2010>.
- van Kesteren, M.T.R., Rijpkema, M., Ruiters, D.J., Fernández, G., 2013b. Consolidation differentially modulates schema effects on memory for items and associations. *PLoS One* 8 (2). <https://doi.org/10.1371/journal.pone.0056155>.
- van Kesteren, M.T.R., Ruiters, D.J., Fernández, G., Henson, R.N., 2012. How schema and novelty augment memory formation. *Trends Neurosci.* 35 (4), 211–219. <https://doi.org/10.1016/j.tins.2012.02.001>.
- Van Overschelde, J.P., Rawson, K.A., Dunlosky, J., 2004. Category norms: an updated and expanded version of the Battig and Montague (1969) norms. *J. Mem. Lang.* 50 (3), 289–335. <https://doi.org/10.1016/j.jml.2003.10.003>.
- Van Petten, C., Luka, B.J., 2012. Prediction during language comprehension: benefits, costs, and ERP components. *Int. J. Psychophysiol.* 83 (2), 176–190. <https://doi.org/10.1016/j.ijpsycho.2011.09.015>.
- Van Petten, C., Senfor, A.J., 1996. Memory for words and novel visual patterns: repetition, recognition, and encoding effects in the event-related potential. *Psychophysiology* 33, 491–506.
- Winocur, G., Moscovitch, M., Bontempi, B., 2010. Memory formation and long-term retention in humans and animals: convergence towards a transformation account of hippocampal-neocortical interactions. *Neuropsychologia* 48 (8), 2339–2356. <https://doi.org/10.1016/j.neuropsychologia.2010.04.016>.