Received: 10 June 2022 | Revised: 15 November 2022 | Accepted: 13 December 2022

DOI: 10.1111/psyp.14258

ORIGINAL ARTICLE

PSYCHOPHYSIOLOGY SPR WILEY

Task context dissociates the FN400 and the N400

Regine Bader[1] | Luca Tarantini[1,2] | Axel Mecklinger[1]

[1]Experimental Neuropsychology Unit, Department of Psychology, Saarland University, Saarbrücken, Germany

[2]Department of Psychiatry and Psychotherapy, Saarland University Medical Center, Homburg, Germany

Correspondence
Regine Bader, Experimental Neuropsychology Unit, Department of Psychology, Saarland University, Campus A2.4, 66123 Saarbrücken, Germany.
Email: regine.bader@mx.uni-saarland.de

Abstract

In event-related potential studies, familiarity-based recognition has been associated with the FN400, that is, more positive-going waveforms for old items than new items 300–500 ms post-stimulus onset, maximal at frontal electrodes. We tested the proposition that the FN400 reflects the attribution of unexpected processing fluency to familiarity. This implies that the FN400 is greater when fluency is less expected, that is, for less familiar stimuli. Moreover, the FN400 should be modulated by the goal of remembering and only elicited when fluency is correctly attributed to the past, that is, by correct old responses in recognition memory tests. In the absence of a retrieval task, enhanced fluency for repeated items should be associated with an N400 attenuation as no episodic attribution takes place. In an incidental study-test design with words of low and high life-time familiarity, participants made pleasantness judgments for half of the studied words. The other half re-appeared in a recognition test. Only in the latter task, participants had the goal of remembering. As both tasks included also new words, we could compare old/new effects under conditions in which both effects are driven by increased fluency for repeated words. We did not find the expected differences in the FN400 for low vs. high life-time familiarity items. However, as expected, we found a frontally distributed FN400 in the recognition test whereas the old/new effect in the pleasantness task resembled an N400 effect. This supports the view that the FN400 occurs when fluency is attributed to familiarity during a recognition decision.

KEYWORDS

ERPs, familiarity, fluency attribution, FN400, N400, recognition memory, retrieval intention

1 | INTRODUCTION

Judgments of prior occurrence of an event can be accomplished by recollection for contextual information or rely on familiarity for the event. According to dual-process models of recognition memory (see Yonelinas, 2002, for a review), these two types of mnemonic experiences are associated with different neurocognitive processes (e.g., Diana et al., 2007; Montaldi & Mayes, 2010). In ERP studies of recognition memory, correctly remembered old items elicit more positive-going waveforms than correctly rejected new items. These so-called old/new effects can be

Regine Bader and Luca Tarantini contributed equally to this study.

differentiated based on their temporal and topographical characteristics (see Rugg & Curran, 2007, for a review). The left-parietal old/new effect or late positive component (LPC) has functionally been associated with recollection and is most pronounced between 500 and 700 ms after stimulus onset over (left) parietal recording sites. The positivity associated with successful recollection co-varies with the amount of contextual details remembered (e.g., Vilberg et al., 2006) and is greater when the old judgment is accompanied by a subjective judgment of recollection (Woodruff et al., 2006; Yu & Rugg, 2010). The mid-frontal old/new effect or the FN400 is maximal from 300 to 500 ms after stimulus onset and has a frontal scalp distribution. It has been associated with familiarity-based recognition judgments as it is graded according to familiarity strength (Woodruff et al., 2006; Yu & Rugg, 2010) and can be observed for responses under time pressure, for which the LPC does not emerge (Mecklinger et al., 2010). However, its exact functional significance has been a matter of debate (e.g., Mecklinger et al., 2012; Paller et al., 2012).

Several studies concluded that the FN400 reflects differences in conceptual fluency between old and new items rather than an episodic familiarity signal and is therefore functionally indistinguishable from the N400 (Voss & Federmeier, 2011), which, in language studies, is typically attenuated when conceptual processing is facilitated (Kutas & Federmeier, 2000, 2011). However, other studies demonstrated that the FN400 and the N400 can topographically and functionally be dissociated (Bader & Mecklinger, 2017; Bridger et al., 2012; Stróżak et al., 2016; Woollams et al., 2008). While the N400 with its typical posterior distribution was associated with semantic priming in these studies, the more frontally distributed FN400 was linked to episodic recognition decisions (Rugg et al., 1998). Recently, we argued that familiarity has oftentimes been only loosely defined in opposition to recollection as a feeling of knowing without memory for contextual information. We inferred that this might be one reason for the difficulty in linking the FN400 to familiarity-based recognition judgments (Mecklinger & Bader, 2020). Therefore, we proposed a framework which provides functional interpretations for the FN400 and N400 drawing on earlier ideas about different uses of familiarity signals (Mandler, 1980) and about mnemonic attributions (Whittlesea & Williams, 2001a, 2001b). Mandler (1980) distinguished between absolute and relative familiarity with the former being the baseline familiarity strength for an item accumulated over all previous exposures and the latter reflecting the relative increment in familiarity strength induced by a specific episodic encounter. Importantly, since this episodic familiarity signal is computed relative to the baseline familiarity strength associated with an item before the encounter, relative familiarity for stimuli with low baseline familiarity is greater than for stimuli with high baseline familiarity.

To overcome the limitations of dual process models that define familiarity merely as memory in the absence of recollection, Whittlesea and Williams (2001a, 2001b) assume that familiarity results from an attribution process by which processing fluency is ascribed to the past. They proposed that old items in a recognition test are processed more fluently than new items. However, only when this experience of fluent processing is unexpected, fluency is attributed to prior occurrence and a strong feeling of familiarity arises. Combining the ideas of Mandler (1980) and Whittlesea and Williams (2001a, 2001b), we recently proposed an unexpected fluency-attribution account of familiarity (Mecklinger & Bader, 2020). According to this framework, an episodic relative familiarity signal is greater for rare stimuli because fluent processing of these stimuli is less expected than for frequent stimuli and therefore a mnemonic attribution is more likely (Coane et al., 2011; Mecklinger & Bader, 2020). Furthermore, according to our account, N400 variations reflect differences in absolute familiarity or conceptual processing fluency per se while the FN400 is an episodic relative familiarity signal which only occurs when processing fluency is surprisingly high and therefore attributed to prior occurrence (see also Leynes et al., 2017, for a similar idea). In line with this idea, the FN400 old/new effect is greater for rare than frequent stimuli (Bridger et al., 2014; Stenberg et al., 2009). Moreover, Leynes and Mok (2020) found a centro-parietal N400 attenuation for (well-known) name-brand products compared to (unknown) off-brand products in a life-time familiarity test, in line with higher absolute familiarity for the former compared to the latter. A topographically distinct FN400, however, was found when the same name-brand products had to be distinguished from new name-brand products in a subsequent recognition test (see also Bridger et al., 2014, for a similar dissociation within a recognition test). Finally, the old/new effect is only frontally distributed when subjects respond "old" to the old items, that is, when they make a correct attribution to oldness. In contrast, when they incorrectly respond "new" (i.e., in miss trials), the topographical distribution is shifted to parietal sites (Rugg et al., 1998).

The unexpected fluency attribution account allows deriving clear predictions on whether differences between repeated (old) and new items in a recognition test should be accompanied by an FN400 or an N400. Crucially, we proposed that in contrast to fluency signals, which are determined by stimulus characteristics (including previous experience with a stimulus), familiarity attributions presuppose an intention to retrieve episodic information and can be modulated by top-down processes. Therefore, we claimed that the FN400 should only occur when participants are required to make episodic decisions while differences in fluency between old and new items should be evident in an N400 effect in situations in which non-mnemonic judgments are made. In support of this prediction, factors that are closely related to the recognition decision affect the size of the FN400. For example, we recently showed that the FN400 can be influenced by the test format (Bader et al., 2020). When old items have to be discriminated from highly similar foil items, familiarity is not useful to identify old items because the difference between the familiarity distributions for old items and foils is smaller than the overall variance in familiarity across items (Migo et al., 2009; Norman & O'Reilly, 2003). However, when an old target item is presented simultaneously with its corresponding similar foil in a forced-choice format, the two familiarity signals can be directly compared, and the slightly more familiar item can be identified as the old item. In contrast, the forced-choice format does not support familiarity-based judgments when targets are presented next to a foil that corresponds to another studied picture. We found the FN400 only in the former condition, in which familiarity-based judgments were supported, but not in the latter condition, in which familiarity was not useful. Relatedly, Ecker and Zimmer (2009) found a larger FN400 for such similar foils compared to new items in a yes/no-recognition test when participants had the task to judge similar foils as "old." In contrast, waveforms elicited by similar foils were more comparable to those elicited by completely new items when similar foils had to be judged as "new" (for similar results with faces see Guillaume & Tiberghien, 2013). In summary, these studies support the view that the FN400 is modulated by factors pertaining to the recognition decision and co-varies with the familiarity-based judgment of prior occurrence rather than the absolute familiarity strength at a given time.

The studies described so far provide evidence that the FN400 is closely tied to correct "old" judgments in episodic recognition judgments and that the FN400 can be dissociated from the N400 effect when the latter was induced by priming manipulations or differences in pre-experimental familiarity. The focus of prior studies on the FN400 was the investigation of two alternating accounts of the FN400 (conceptual fluency and familiarity). Hence, these studies do not directly assess how the FN400 and the N400 are modulated by the goal of remembering. In contrast, the present study investigates whether the FN400 and the N400 effect can be dissociated by the presence or absence of an episodic retrieval task when both effects are driven by differences in processing fluency between old and new items. One recent study by Yang et al. (2019) investigated recognition memory for words denoting object concepts with low and high life-time familiarity (LTF). They found a centro-parietal N400 effect for LTF, that is, a larger N400 for low than high LTF items. Moreover, both the FN400 and the N400 were sensitive to how often a word had previously occurred during the experiment with more positive-going waveforms for more frequent words. This recency-related N400 effect had a similar centro-parietal distribution as the LTF N400 effect and was assessed in the final part of the study phase, in which some words had appeared more often than others and the number of occurrences was task-irrelevant (participants had to make animacy judgments, i.e., they did not have the goal of remembering). In contrast, the FN400 effect was found in a subsequent test phase, in which frequency judgments for recent laboratory exposures had to be made, that is, participants were engaged in an episodic retrieval task. Although this FN400 effect was more anteriorly distributed than the N400 effect in the study phase, it was not as frontally distributed as usually observed in recognition memory studies. This might be explained by the unusual task of judging the frequency of recent laboratory exposures rather than making standard old/new judgments. Moreover, the N400 and the FN400 in this study were not directly comparable as the study and the test phase did not only differ in task demands but also with respect to how often the words had been presented and how much time participants had already spent in the experiment.

In the current study, we also used low and high LTF words as a manipulation of the pre-experimental baseline familiarity but chose a more powerful design to dissociate FN400 and N400 effects, both elicited by differences in fluency for old compared to new words but emerging in different tasks. We adopted a relatively simple study-test design in which participants incidentally encoded concrete nouns in a naming task. Thereafter, they were tested on half of the words in a recognition memory test, in which they had to discriminate old from new words. For the other half of the words, they performed a pleasantness task, in which they had to judge each word on its subjective pleasantness. Crucially, this task did not only comprise old but also new words and hence, was matched in old/new status with the recognition memory test. Task order was counterbalanced in order to keep study/test distance constant between tasks across participants. Greater fluency for old than new words should be attributed to familiarity only in the recognition test. Conversely, in the pleasantness task, fluency should be attributed to pleasantness and give rise to a mere exposure effect (Jacoby et al., 1989; Whittlesea & Price, 2001; Zajonc, 1968). Behaviorally, we expected higher hit and lower FA rates for low relative to high LTF items in the recognition task, resembling the word-frequency mirror effect (Glanzer & Adams, 1985). In the pleasantness task, we predicted shorter RTs for items with high LTF than low LTF and for old words compared to new words.

Moreover, if processing fluency is attributed to pleasantness in the pleasantness rating, old words should be associated with higher pleasantness ratings than new words. The same was expected when high LTF items are compared to low LTF items. For the ERPs, we assumed that if the N400 reflects conceptual processing fluency irrespective of task context, it should be attenuated for high compared to low LTF items in the study phase and for high compared to low LTF new items in both test tasks. The latter contrast was restricted to new items because unstudied items should not be influenced by episodic familiarity (Bader & Mecklinger, 2017; Bridger et al., 2014). A similar N400 effect was expected for old relative to new words in the pleasantness rating, wherein episodic retrieval was not required. In contrast, old/new differences in the recognition test should be reflected in an FN400 effect. As the increased fluency signal for studied items should be less expected for low than for high LTF items, we also predicted that the FN400 would be larger for low than for high LTF items. Finally, we expected all effects to be present in the same time interval but that the FN400 effect in the recognition test would be more frontally distributed than the N400 contrasts.

2 | METHOD

The study was pre-registered on the OSF platform at https://osf.io/s67rx.

2.1 | Participants and sample size

The final sample comprised twenty-four native German speakers (16 female) from the student population of Saarland University via advertisement posters on campus and in social networks. All participants had normal or adjusted to normal vision, were right-handed as assessed by the Edinburgh Handedness Inventory (laterality quotient ≥50; Oldfield, 1971) and had no known neurological or psychiatric issues. Mean age was 24 years (range = 19–30 years). One additional subject was excluded and replaced as there were not enough artifact-free trials (<13) in the conditions of interest. Informed consent was required before the experiment, and we debriefed participants after the experiment. We compensated participants with either €10/h or course credit. Note that we based sample size on a power analysis for a repeated-measures ANOVA (not for the linear mixed-effects models we used) using G*Power 3.1 software (Faul et al., 2009). This analysis yielded a required sample size of $n = 21$ resulting in a final sample size of $n = 24$ for counterbalancing purposes. The study was approved by the ethics committee of the Faculty of Human and Business Sciences at Saarland University and adhered to the Declaration of Helsinki.

2.2 | Stimuli

Stimuli were 400 German nouns ranging from high to low life-time familiarity, taken from a normative German database (Schröder et al., 2012). The words belonged to eleven different semantic categories (mammals, birds, clothes, fruits, furniture, music instruments, professions, sports, tools, vegetables, vehicles) and varied between 3 and 17 characters in length. No two words shared the same word stem. Items were divided into four sets, each consisting of 50 words with high and 50 words with low life-time familiarity. Semantic categories were similarly distributed across sets. There were no differences between sets for life-time familiarity ($p = .821$), lemma frequency ($p = .951$) according to dlexDB (Heister et al., 2011), or word length ($p = .813$). Across lists, low LTF items had a LTF mean rating of 2.48 (scale 1–5) and high LTF items a mean rating of 3.80. Ratings for low and high LTF items differed significantly, $t(369) = 30.32$, $p < .001$. Correspondingly, low LTF items had a mean normalized lemma frequency of 3.01 and high LTF items a frequency of 11.12, which was also significantly different from each other, $t(246.34) = 5.23$, $p < .001$. Participants studied a list of 200 words comprising two of the word sets. In each of the two subsequent tasks, participants saw 100 old words (one of the studied sets) and 100 new words (one of the unstudied sets), so that every item occurred only once across both tasks. The assignment of the word sets to old/new status was counterbalanced across both tasks, but not across task orders. Female and male participants were distributed equally on the counterbalancing conditions. Study and test lists were individually randomized for each participant. Stimuli were always presented in black on gray background (font: Arial, 28 pt).

2.3 | Procedure

We programmed and conducted the experiment using E-Prime 2.0 (Psychology Software Tools). After assessing demographical variables and inclusion criteria, the EEG cap was fitted. The experimental session lasted approximately 1.5 h and was conducted on a standard PC. Participants were seated in front of a 19-in. monitor with a resolution of 1280 × 1024 px at a distance of approximately 80 cm, inside a sound-attenuated and electrically shielded chamber. Procedure of the study and both test phases is illustrated in Figure 1. Each phase of the experiment began with a short practice phase (10 trials each) to familiarize participants with the task.

PSYCHOPHYSIOLOGY

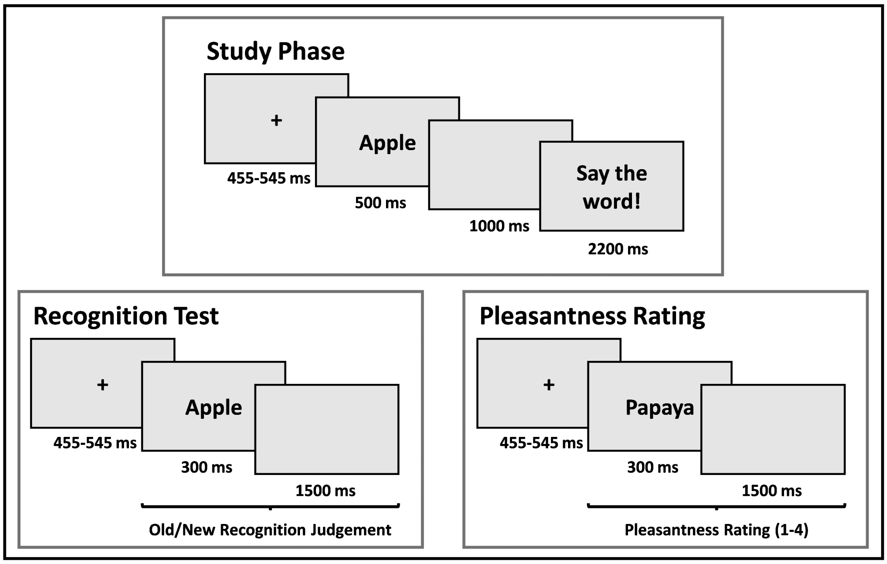


FIGURE 1 Schematic of the experimental procedure in the study phase (top), the recognition test (bottom left), and pleasantness task (bottom right). After the study phase, all participants performed the recognition test and the pleasantness task sequentially in counterbalanced order.

The study phase always occurred first and was followed by a short distractor task (4 min). Temporal order of the recognition test and pleasantness task was counterbalanced between participants. During the incidental study phase, participants were told that the experiment investigated neurocognitive processes underlying speech processing and that they would have to pronounce the presented words. We chose this task following Bridger et al. (2014) and because we wanted to ensure that the study task was equally dissimilar to both ensuing test phase tasks. A microphone was placed in the experimental chamber to increase authenticity. Each study trial began with a fixation cross jittered in steps of 15 ms from 455–545 ms, which was replaced by the study word, presented for 500 ms. The screen was then blanked for 1000 ms, which was followed by the instruction "Wort sagen!" (German for "Say the word!") for 2200 ms. Participants were instructed that the word should not be uttered until they were explicitly required to, in order to measure artifact-free ERPs during the presentation of the word and the blank interval. Participants took a self-paced break after every 50 trials. Following the study phase, participants completed a short distractor task (math equations in a paper-pencil setting and counting backwards) until the temporal limit of 4 min was reached. Each trial of the recognition test (which directly followed the distractor task in half of the cases) started with a fixation cross (jittered around 500 ms as in study phase). Thereafter, the test word was presented for 300 ms followed by a blank screen for 1500 ms. Participants were required to make a binary old/new judgment using two separate keys on a response box as soon as the test word appeared. The mapping of old/new responses onto left/right keys was counterbalanced across participants and temporal test order so that half of the participants from each test order group pressed the left key for "old" responses. Trial timing of the pleasantness task was the same as in the recognition test. Participants were required to press one of four response box keys to make a pleasantness judgment in the following steps: 1 (*very unpleasant*), 2 (*unpleasant*), 3 (*pleasant*), 4 (*very pleasant*). The mapping of pleasantness ratings onto keys was counterbalanced across participants and temporal test order. We informed participants that words from the naming task were intermixed with new words in the pleasantness task but that it was not relevant for the pleasantness rating whether the words were old or new. In both tasks, a cardboard stand-up reminded participants of the correct response-key mapping and we instructed them to respond fast and spontaneously. Participants took a self-paced break after every 50 trials.

2.4 | Electrophysiological recording & preprocessing

Continuous EEG was recorded from 28 AG/AgCl scalp electrodes mounted in an elastic cap (Easycap) and labeled according to the extended 10–20 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). Electrode AFz served as ground electrode and EEG was acquired referenced to the left mastoid electrode (A1), using a 16bit BrainAmp amplifier (Brain Products). Signals were band-pass filtered from 0.016–250 Hz and digitized at a sampling rate of 500 Hz. Impedances were kept below 5 kΩ. Electrooculogram (EOG) was recorded with four additional electrodes on the outer canthi and above and below the right eye. Offline processing was performed with Brain Vision Analyzer 2.1 (Brain Products). First, we visually identified and removed excessive (i.e., muscular) artifacts from the raw data. Afterward, we applied a 0.05–30 Hz butterworth-filter (order: 4) and a notch-filter (50 Hz). For EOG and cardiac artifact correction, we employed independent component analysis, using the classic biased restricted info-max algorithm implemented in Brain Vision Analyzer 2.1 (Brain Products). The signal was re-referenced to the average of both mastoids. Epochs from −200 to 1000 ms around stimulus-onset were constructed and baseline-corrected to the 200 ms before stimulus onset. Segments with artifacts such as voltage steps greater than 30 microvolts per ms, voltage differences greater than 100 microvolts per 200 ms, and absolute amplitudes larger/smaller than +/−70 microvolts were automatically rejected. Finally, we excluded all segments that contained excessive alpha-activity as detected by manual inspection. Finally, we built participant-wise averages for each condition (recognition hits, recognition correct rejections, pleasantness old, pleasantness new) and computed grand-average waveforms. ERP waveforms were plotted with "ggplot2" from the "tidyverse" package in R (see below) after exporting the low-pass filtered (12 Hz) grand averages.

2.5 | Statistical analyses

Overall, analyses were conducted with R version 3.6.3 (R Core Team, 2019) in RStudio (RStudio Team, 2019), especially using the packages "tidyverse" (Wickham et al., 2019), "psych" (Revelle, 2019), "reshape" (Wickham, 2007), "nlme" (Pinheiro et al., 2020) and "ez" (Lawrence, 2016). Significance level was set to $\alpha = .05$. Marginally significant effects are reported in the results section but are not further dissolved. For behavioral analyses, trials that contained reaction time (RT) outliers were identified subject-wise and excluded at the recommendation of Tukey (1977), that is, values below first quartile − (1.5 * interquartile range) and above third quartile + (1.5 * interquartile range). Mean numbers of excluded outlier trials in all conditions and respective combinations (Item Status × LTF × Task) ranged from 0.08 to 1.92 (mean = 0.96). For inferential statistics, we used mixed ANOVAs with the between-subjects factor test order (pleasantness first, recognition first) and the within-subjects factor item status (old, new) and LTF (high, low) for the dependent variables proportions of hits and false alarms. Response times for hits and correct rejections (CRs) in the recognition test were analyzed separately in two-way ANOVAs with the factors test order and LTF. For the pleasantness task, we used a mixed ANOVA to compare rating scores and response times including the between-subjects factor test order and the within-subjects factor item status and LTF. For ERP analyses, we included all items of the study phase, hits and CRs in the recognition test as well as old and new items for the pleasantness task. Mean proportion of artifact-free trials (range) was as follows: study phase: 87.75 (76–97) for low LTF items, 87.08 (76–97) for high LTF items; recognition test: 28.88 (18–42) for low LTF hits, 27.33 (13–42) for high LTF hits, 41.25 (26–49) for low LTF CRs, 39.58 (21–50) for high LTF CRs; pleasantness task: 46.58 (30–50) for low LTF old items, 46.17 (35–50) for high LTF old items, 46.04 (31–50) for low LTF new items, 46.29 (36–50) for high LTF new items. Time windows for inferential statistics on mean amplitudes were defined as follows (Luck, 2014): We computed the grand average (GA) over all phases (study phase and both test tasks) and both levels of item status and LTF. Afterward, we averaged the local peak in the time window from 300 to 500 ms after stimulus onset (as this is the typical FN400 and N400 time window) across all electrodes considered in the analyses. We then defined the empirical time window from −100 ms to +100 ms around the time point of this local peak. This procedure resulted in an empirical time window from 300 to 500 ms. In order to maximize the probability of finding topographical differences, we built five topographic electrode clusters of three electrodes each: a frontal cluster (F3, Fz, F4), a fronto-central cluster (FC3, FCz, FC4), a central cluster (C3, Cz, C4), a centro-parietal cluster (CP3, CPz, CP4), and a parietal cluster (P3, Pz, P4). Moreover, we used linear mixed-effects models for the analyses of the ERP data as this approach does not assume sphericity, which is mostly violated in ERP data. The dependent variable was the mean amplitude difference between high and low LTF items for the study phase and the new items analyses in the test phase, between hits and CRs in the recognition test and between old and new words in the pleasantness task (in the following referred to as old/new difference for both test tasks). In each model, subject was treated as random factor. LTF (low, high), topography (frontal, fronto-central, central, centro-parietal, parietal), task (recognition/pleasantness), test order and possible interactions were treated as fixed factors. As we expected a linear decrease of the old/new difference in the recognition test from frontal to parietal electrodes, we modeled the factor topography with four polynomial contrasts in order to assess the linear effect.[1] LTF was coded as −1

[1]Note that the way we modeled the factor topography deviates from the pre-registration where we intended to model this factor with two levels (anterior, posterior). However, we think that the linear contrast is a more powerful approach to test our hypothesis of a more frontally focused effect in the recognition test because it tests more specifically for the continuous linear decrease of the FN400 effect from frontal to parietal sites across all clusters – even if overall differences between frontal and parietal clusters are small.

for low and 1 for high, task was coded as −1 for pleasantness and 1 for recognition, and test order was coded as −1 for pleasantness task first and 1 for recognition test first. Note that as we used amplitude-difference scores as dependent variables, the test for the intercept of each model indicates whether the high-low difference (study phase and new item comparison) or the old-new difference (test phase) is different from zero. All remaining effects can be interpreted as interactions with this effect. In order to dissociate the ERP effects of interest topographically and functionally, we compared their rescaled difference scores using the vector-scaling method (McCarthy & Wood, 1985) and re-ran the relevant analyses. To examine whether the between-subjects factor test order (pleasantness first vs. recognition first) affects the results, we included it in all analyses. For all (marginally) significant inferential statistics, we report generalized eta-squared (η_G^2) (Bakeman, 2005), Hedges' g_s, and Hedges' g_{av} (Lakens, 2013) as indicators of effect sizes for ANOVAs, two-samples, and paired *t* tests, respectively.

3 | RESULTS

3.1 | Behavioral results

Behavioral results of the recognition test are displayed in Table 1. The three-way ANOVA on hit and false alarm rates revealed only a significant effect of item status $F(1,22) = 261.08$, $p < .001$, and a trend for an item status by LTF interaction, $F(1,22) = 3.18$, $p = .089$. In planned separate analyses, the ANOVA on hit rates revealed a marginally significant effect of task order, $F(22) = 3.64$, $p = .069$, $\eta_G^2 = .12$, with those participants tending to perform better who did the recognition test first. The LTF effect ($p = .214$) and the LTF x Test order ($p = .864$) interaction were not significant. For false alarms, there were no significant effects ($ps > .343$). Thus, we did not find the predicted mirror effect between high and low LTF items.

Using reaction times as dependent variables in an exploratory analysis, there was a significant effect of test order for hits, $F(1,22) = 14.64$, $p < .001$, $\eta_G^2 = .39$, with those participants being faster who did the recognition test first. Moreover, across test order groups, high LTF hits were made significantly faster than low LTF hits, $F(1,22) = 10.83$, $p = .003$, $\eta_G^2 = .02$. The interaction was not significant ($p = .446$). For CRs, there was a marginally significant effect of test order, $F(1,22) = 3.70$, $p = .067$, $\eta_G^2 = .14$, again with those participants being faster who did the recognition test first. The other effects were not significant.

Mean rating scores and response times in the pleasantness task are displayed in Table 2. A mixed ANOVA on the ratings with the between-subjects factor test order (pleasantness first, recognition first) and the within-subjects factors item status (old, new) and LTF (low, high) revealed a significant effect of test order, $F(1,22) = 5.51$, $p = .028$, $\eta_G^2 = .16$, no significant effect for item status ($p = .447$), but a significant effect of LTF, $F(1,22) = 55.48$, $p < .001$, $\eta_G^2 = .20$, a significant interaction of test order by item status, $F(1,22) = 5.04$, $p = .035$, $\eta_G^2 = .024$, and a significant 3-way interaction Test order x LTF x Item status, $F(1,22) = 4.32$, $p = .049$, $\eta_G^2 = .01$. Dissolving the interactions involving the test order factor, we ran separate ANOVAs with item status and LTF as within-subjects factors. These ANOVAs revealed that the item status effect was not significant within both test order groups (pleasantness first: $p = .162$; recognition first: $p = .111$). However, high LTF items were judged as more pleasant than low LTF items in both groups (pleasantness first: $F(1,11) = 39.52$, $p < .001$, $g_{av} = .419$; recognition first $F(1,11) = 17.91$, $p = .001$, $g_{av} = .094$). The interaction was not significant in both groups

TABLE 1 Mean proportions of hits and false alarms and mean response times for correct responses in the recognition test separated according to task order (pleasantness task first vs. recognition test first) and lifetime familiarity (low LTF vs. high LTF).

	Pleasantness first		Recognition first	
	Low LTF	**High LTF**	**Low LTF**	**High LTF**
Memory performance				
Hits	.55 (0.04)	.51 (0.04)	.64 (0.04)	.62 (0.04)
FAs	.10 (0.03)	.12 (0.03)	.14 (0.03)	.15 (0.05)
Response times				
Hits	907 (36)	868 (37)	741 (20)	717 (24)
CRs	820 (42)	835 (46)	733 (20)	738 (18)

Note: Numbers in parentheses represent the standard error of the mean.

TABLE 2 Mean pleasantness rating scores and mean response times for old and new words in the pleasantness task separated according to task order (pleasantness task first vs. recognition test first) and lifetime familiarity (low LTF vs. high LTF).

	Pleasantness first		Recognition first	
	Low LTF	**High LTF**	**Low LTF**	**High LTF**
Pleasantness rating scores				
Old	2.42 (0.05)	2.73 (0.06)	2.79 (0.09)	2.91 (0.09)
New	2.49 (0.05)	2.75 (0.04)	2.63 (0.09)	2.88 (0.09)
Response times				
Old	967 (55)	921 (52)	852 (35)	831 (35)
New	959 (50)	922 (48)	891 (39)	850 (38)

Note: Numbers in parentheses represent the standard error of the mean.

($ps > .101$). Using response times as dependent variable, a mixed ANOVA with the same factors as above yielded no significant effect of test order ($p = .181$), but a significant effect of item status, $F(1,22) = 5.54$, $p = .028$, $\eta_G^2 = .002$, and LTF, $F(1,22) = 83.34$, $p < .001$, $\eta_G^2 = .015$, and a significant interaction of Test order x Items status, $F(1,22) = 8.68$, $p = .007$, $\eta_G^2 = .003$. No other effects were significant ($ps > .129$). Dissolving the interaction, we found that there was no difference in response times to old and new words ($p = .723$), for those participants who did the pleasantness task first. However, when the recognition test came first, participants responded significantly faster to old than new words, $F(1,11) = 21.09$, $p < .001$, $g_{av} = .014$. RTs to high LTF items were significantly faster than to low LTF items in both groups (pleasantness first: $F(1,11) = 66.86$, $p < .001$, $g_{av} = .02$ recognition first: $F(1,11) = 25.84$, $p < .001$, $g_{av} = .02$).

To sum up, participants were faster in the recognition test when they did it before the pleasantness task as compared to the other way round. In addition, they were faster making high than low LTF hits. Moreover, for the whole sample, there was no evidence that exposure in the study phase increased pleasantness as old words were not judged as more pleasant than new words. However, pleasantness ratings to high LTF items were higher and were made faster than to low LTF items. Finally, a speed advantage for old words over new words in the pleasantness task was only found when participants did the pleasantness task after the recognition test.

3.2 | Study phase ERPs

ERP waveforms for the study phase are depicted in Figure 2 (left panel). High and low LTF items start to differ slightly around 300 ms, albeit the differences between conditions are small. A mixed effects model with the mean high/low difference in the 300–500 ms time window as dependent variable, subjects as random effect and topography as fixed effect revealed a trend for the intercept ($p = .077$) and a marginally significant positive linear contrast, $b = .19$, $t(92) = 1.89$, $p = .062$.

3.3 | Test phase ERP old/new effects

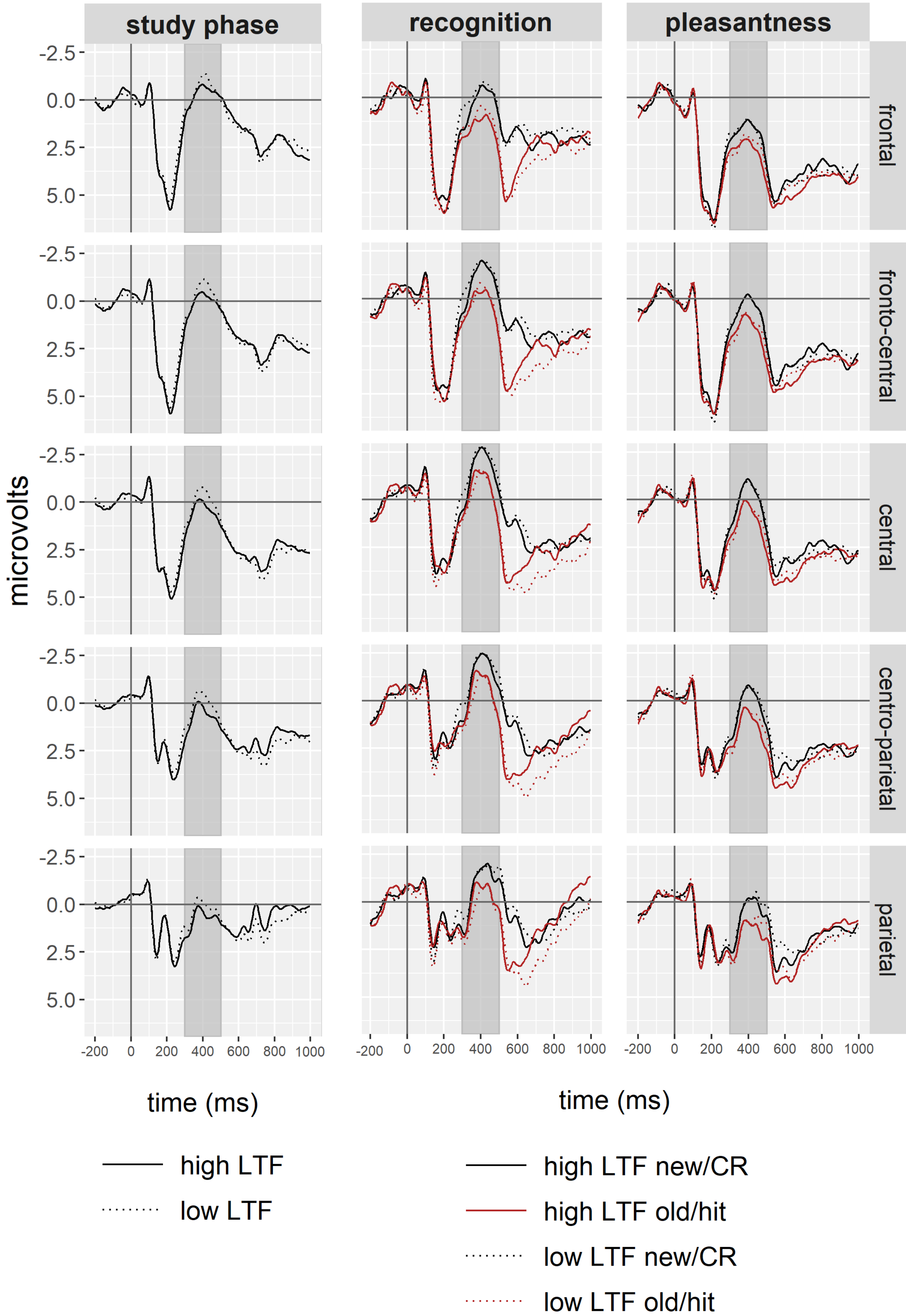
ERP waveforms associated with hits vs. CRs and old vs. new items in the recognition test and the pleasantness task, respectively, for low and high LTF items are depicted in the middle and right panel of Figure 2. In both tasks, the two item types start to differ from about 300 ms onwards across the whole scalp with hits/old items eliciting more positive-going ERPs. The difference between low and high LTF items is negligible. However, the difference in the FN400 time window is more frontally focused in the recognition test than in the pleasantness task, as is most evident in the topographic maps in Figure 3 (see also Figure 4 for mean amplitudes across the scalp). In the subsequent time window, old/new differences seem to be larger in the recognition test than in the pleasantness task.

In the full linear mixed effect model with the mean old/new difference in the 300–500 ms time window as dependent variable, we included subjects as random effect and task, LTF, topography, and test order as fixed effects. This model revealed a significant intercept, $b = 1.11$, $t(352) = 6.37$, $p < .001$, indicating a significant old/new effect across tasks, LTF levels and test orders. Confirming our predictions, there was a significant interaction between the linear topographic contrast and task, $b = -0.19$, $t(352) = -2.80$, $p = .005$. None of the other interactions between the polynomial contrasts and task were significant (range of ps: .567–.925). There was also a significant LTF × linear contrast interaction, $b = -0.16$, $t(352) = -2.40$, $p = .017$, a significant LTF × test order interaction, $b = .31$, $t(44) = 2.32$, $p = .025$, and a significant Task × LTF × Test order interaction, $b = 0.38$, $t(44) = 2.85$, $p = .007$. No other main effects or interactions were significant.

Dissolving the interactions involving the LTF factor, we ran separate analyses for high and low LTF items. We report only those effects and interactions that are subject to interpretation because of the higher order interactions. For high LTF items, there was a significant intercept, $b = 1.03$, $t(176) = 4.56$, $p < .001$, a marginally significant effect of test order, $b = 0.44$, $t(22) = 2.02$, $p = .064$, and a significant negative linear trend, $b = -0.23$, $t(176) = -2.20$, $p = .029$. Moreover, there was a Task × Test order interaction, $b = 0.52$, $t(22) = 2.42$, $p = .024$. Dissolving this interaction, we ran separate analyses on each level of task for the high LTF items. For the recognition task, there was a significant intercept, $b = 1.03$, $t(88) = 2.89$, $p = .005$, and a significant effect of test order, $b = 0.96$, $t(22) = 2.70$, $p = .013$. For the pleasantness task, only the intercept was significant, $b = 1.03$, $t(88) = 3.94$, $p < .001$.

FIGURE 2 ERP waveforms in the study phase (left panel), recognition test (middle panel), and the pleasantness task (right panel). High LTF items are displayed as solid lines and low LTF items as dotted lines. For the middle and right panel hits/old items are displayed in red and correct rejections/new items in black. ERPs are displayed for five electrode clusters (from top to bottom: frontal, fronto-central, central, centro-parietal, and parietal). Waveforms were low-pass-filtered (12 Hz) for illustrative purposes.

PSYCHOPHYSIOLOGY SPR



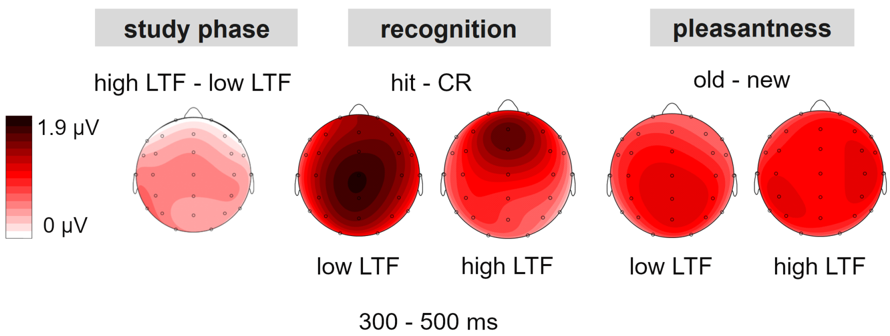


FIGURE 3 Topographic maps for the difference waveforms between high and low LTF items for the study phase, between hits and CRs in the recognition test and between old and new words in the pleasantness task in the 300–500 ms time window.

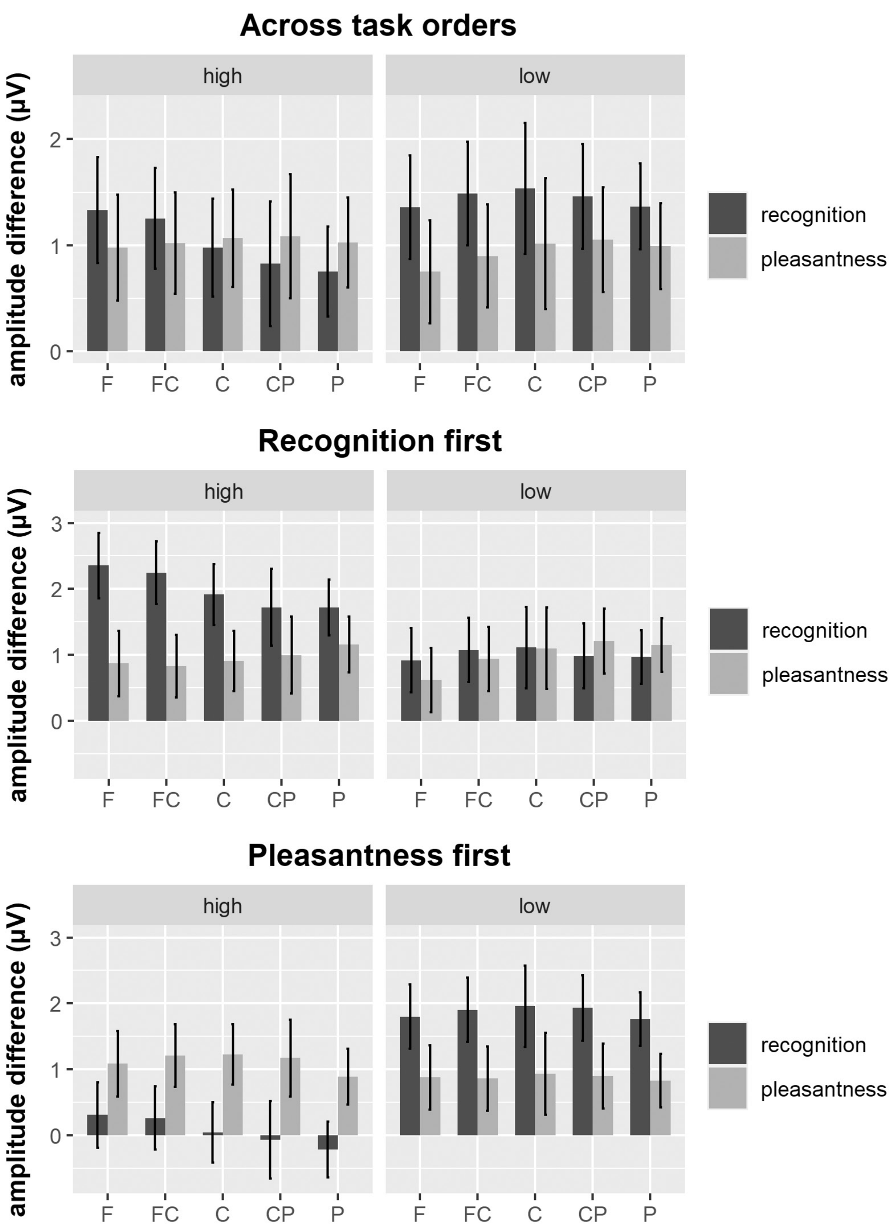


FIGURE 4 The mean amplitude differences from 300–500 ms for hits minus CRs and old minus new items in the recognition test (dark) and pleasantness task (light), respectively, at the five clusters (frontal, fronto-central, central, centro-parietal, and parietal). "High" and "low" refer to high and low LTF items, respectively. Upper panel: Means were calculated across all participants irrespective of task order. Middle panel: Means for the recognition test first group. Lower panel: Means for the pleasantness first group. Error bars represent the standard error of the mean difference between tasks.

For low LTF items, there was only a significant intercept, $b = 1.19$, $t(176) = 5.43$, $p < .001$. Thus, for the high LTF items, the FN400 in the recognition test was frontally distributed and larger when the recognition test came first whereas for the low LTF items there was a more broadly distributed old/new effect in both test tasks and no moderation by test order.

Finally, we dissolved the interaction of task and linear contrast with two separate mixed effects models for each task with LTF, topography and test order as fixed effects.

We only report the intercept and the effect of topography as the two-way interaction was not moderated by any other factors in the full model. For the recognition test, we found a significant intercept, $b = 1.23$, $t(176) = 4.46$, $p < .001$, and a significant negative linear trend, $b = -0.25$, $t(176) = -2.36$, $p = .019$, but no other significant topographic effects, range of *p*s: .589–.905. The negative linear trend indicates that the old/new difference linearly decreases in size from frontal to parietal electrode clusters. For the pleasantness task, there was a significant intercept, $b = 0.99$, $t(176) = 4.67$, $p < .001$, but no significant linear topographic effect, $b = 0.13$, $t(176) = 1.52$, $p = .130$, or any other topographic effect (range of *p*s: .579–.959), in line with the broader distribution of this effect across the scalp (see Figures 3 and 4).

In summary, test order effects were generally not very pronounced. Moreover, contrary to our expectations, there was a frontally focused old/new effect for high LTF items and a broadly distributed old/new effect for low LTF items in the recognition test. In line with better memory performance for those participants who did the recognition test first, the FN400 for high LTF items was larger in this group. The distribution of the old/new difference in the pleasantness task was generally more wide-spread across the scalp. Importantly, as predicted, across all levels of LTF and test order, the old/new effect was more frontally focused in the recognition test than in the pleasantness task.

3.4 | Test phase ERP LTF effects for new items

We started with the full mixed effects model with the mean high-low difference in the 300–500 ms time window as dependent variable, subjects as random effect and task, topography, and test order as fixed effects. In this model, there was only a marginally significant test order effect, $b = -0.49$, $t(22) = -1.89$, $p = .072$, indicating that the LTF difference tends to be larger for the group who started with the pleasantness task. Moreover, there was a marginally significant two-way interaction between task and the linear topographic factor, $b = -0.14$, $t(176) = -1.66$, $p = .098$, suggesting that the LTF difference was more pronounced at frontal electrode sites but only in the recognition task. Overall, evidence for a difference between high and low LTF items that have no learning history, that is, new items, is weak and not significant.

3.5 | ERP topographic dissociation

In order to confirm that the significant two-way interaction between task and linear contrast did not arise because of overall amplitude differences between the two tasks, we re-ran the model using vector-scaled amplitude differences as the dependent variables. We applied a mixed effects model with the old–new difference as dependent variable, subjects as random effect and task, topography, and test order as fixed effects. It revealed a significant intercept, $b = 1.00$, $t(176) = 6.44$, $p < .001$, as well as a significant interaction between task and linear contrast, $b = -0.17$, $t(176) = -2.61$, $p = .010$. The three-way interaction of task, linear contrast, and test order was only marginally significant, $b = -0.11$, $t(176) = -1.78$, $p = .076$.

Moreover, we compared the topographic distribution of the recognition FN400 effect also with the (only marginally significant) N400 effect from the study phase. We applied a mixed-effects model with the difference between high vs. low LTF (study task) or hits vs. CRs (recognition task) as dependent variable, subjects as random effect and task, topography, and test order as fixed effects. The model revealed a significant intercept, $b = 0.98$, $t(176) = 3.16$, $p = .002$, as well as a significant interaction between task and linear contrast, $b = 0.32$, $t(176) = 2.52$, $p = .013$, but no further significant results.

Next, we applied a mixed-effects model with the difference wave of high vs. low LTF (study task) or old vs. new items (pleasantness task) as dependent variable, subjects as random effect and task, topography, and test order as fixed effects. The model revealed a significant intercept, $b = 0.98$, $t(176) = 3.33$, $p = .001$, but no interaction between task and topography (range of *p*s: .214–.985). The three-way interaction of task, linear contrast, and test order was marginally significant, $b = -0.22$, $t(176) = -1.77$, $p = .079$.

Taken together, topographic dissociation analyses yielded no evidence that the (marginally significant) N400-LTF effect in the study phase was topographically distinct from the N400-old/new effect in the pleasantness task as both are pronounced at centro-parietal electrode sites. Crucially, the FN400-old/new effect in the recognition task is more frontally distributed than both N400 effects.

4 | DISCUSSION

Recently, we proposed that the FN400 is linked to the attribution of surprisingly high fluency to familiarity (Mecklinger & Bader, 2020). We further claimed that only as long as a person has adapted an episodic retrieval intention, familiarity attribution takes place and in turn, the FN400 is elicited. In contrast, an N400 effect can be observed between items differing in conceptual fluency, as for example, words with low and high life-time familiarity (LTF) or when old and new items are compared, and no episodic decision has to be made. The present research's aim was to dissociate the FN400 from the N400 by manipulating the presence or absence of an episodic retrieval task under otherwise identical testing conditions. Moreover, we wanted to investigate the influence of expectedness on the size of the FN400 by varying LTF. For this purpose, we implemented a study-test design, in which the test phase comprised a recognition test and a pleasantness task in counterbalanced order.

Behaviorally, although there were numerically more hits to low than high LTF items and more false alarms to high than low LTF items, we found no significant mirror effect in the recognition test. However, in line with faster processing of those words that have been encountered more frequently in life, high LTF hits were made faster. Interestingly, this difference in processing speed for words differing in LTF was not significant for CRs as has been observed in combined priming and recognition studies (Bader & Mecklinger, 2017; Woollams et al., 2008) and was interpreted as conflicting fluency signals from priming and oldness. Moreover, participants were generally faster and had (marginally significantly) more hits, when they did the recognition test before the pleasantness task than when they did it afterward. When the recognition test was administered last, participants were presumably more exhausted, the retention interval was longer and there was interference from the preceding pleasantness task. This might have led to impoverished memory representations and a more difficult classification of the old words (Hockley, 1991; Yonelinas & Levy, 2002).

In the pleasantness task, we found faster responses for high compared to low LTF items consistent with the RT pattern for hits in the recognition test. Moreover, there were faster responses for old words compared to new words, however, only when this task was administered last. It is possible that we found repetition priming for old words only in this case because then the pleasantness task was more challenging due to task-switching requirements and because participants were more exhausted. Consistent with this view, it has been shown that priming effects are larger when target processing is more demanding (Hines et al., 1986; Horner & Henson, 2009). It is conceivable that if the task is difficult, participants need to “rely more heavily on resources made available by other stimulus encounters” (Hughes & Whittlesea, 2003, p. 402).

Regarding the pleasantness rating scores, we observed higher ratings for high compared with low LTF items, but there was no conclusive pattern for the old vs. new comparison as there was an interaction between item status and task order, but within-group comparisons yielded no significant results. Thus, frequent exposures accumulated during the lifetime enhanced the subjective pleasantness of words, but an additional single study phase exposure did not. The reason for the absence of an effect induced by study exposure on the pleasantness ratings in our experiment might lie in the stimulus materials. In contrast to the majority of mere exposure studies, which use previously unknown stimuli, we used words denoting familiar concepts. One more exposure to a word which has been encountered thousands of times before by an individual (which is presumably true also for the low LTF items) is unlikely to affect the individual's preference for this word (see Butler et al., 2004; Zajonc, 1968, for a similar view).

Contrary to our assumptions, analyses of the LTF contrast in the ERPs suggest that differences in conceptual fluency between high and low LTF items were relatively small since the effects were only marginally significant as in the study phase or virtually absent as for the new items comparison in the test phase tasks. In line with that, the FN400 was not larger for low than for high LTF items. The frontal focus was even more pronounced for high LTF items. However, as can be observed in Figure 4, this did not result from a larger old/new difference at frontal sites but from a smaller old/new difference at parietal sites. Hence, we believe that our manipulation of LTF turned out to be not as effective as intended and as manipulations of word frequency (Bridger et al., 2014; Stenberg et al., 2009) have been in the past. Our intention to use words for which familiarity norms are available might have diluted the difference between the two levels of LTF. As an illustrative example, the difference in mean normalized lemma frequency (Heister et al., 2011) between low and high LTF items was 3 vs. 11 in our study and 2 vs. 192 in Bridger et al. (2014). Hence, our high LTF category might have been not familiar enough. This explanation could also account for the absence of the behavioral mirror effect in the recognition test (see above). Differences in reaction times and pleasantness ratings between high and low LTF items—in the absence of appreciable N400 differences—might also origin from priming on other levels than the conceptual level, as for example the lexical level.

In line with our assumption that only in the recognition test, participants needed to make episodic memory judgments, we found a frontally focused difference between old and new items (i.e., hits and CRs) only in the recognition test, consistent with an FN400 effect. In contrast, the old/new difference in the pleasantness task resembled a broadly distributed N400 effect. Moreover, the recognition FN400 effect was not only topographically dissociable from the N400 effect in the pleasantness task but also from the (only marginally significant) N400 effect in the study phase, corroborating the dependency of the FN400 on episodic fluency attributions. Hence, our study adds to the existing evidence that the FN400 and the N400 can be topographically and functionally dissociated as has been shown by studies that contrasted recognition tests with priming manipulations (Bader & Mecklinger, 2017; Bridger et al., 2012;

Stróżak et al., 2016; Woollams et al., 2008) or life-time familiarity judgments (Leynes & Mok, 2020). Importantly, we showed that this dissociation holds also for two situations in which exactly the same stimulus materials with the same old/new status are presented but only the tasks differ. The relevance of the task design can also explain why in other studies that used the two tasks within one trial (e.g., Leynes & Addante, 2016; Voss & Federmeier, 2011), the difference between the two components might have been obscured. Yang et al. (2019) have provided first and preliminary evidence for such a dissociation. However, we used an improved design with standard recognition instructions, in which the distance between study and the two tasks was held constant across subjects. Moreover, in contrast to the Yang et al. study, in our study, all old words were only repeated once in the recognition test and in the pleasantness task. This provides a more controlled framework for old/new decisions and renders differences between old and new items more salient.

One drawback of this within-subjects task manipulation are the differences in task order between subjects. We indeed found an influence of task order on some of our behavioral measures emphasizing the importance of controlling this factor when attempting a task dissociation. For the ERPs, the FN400 effect for high LTF items was the only effect moderated by test order in that it was larger for those participants who started with the recognition test. This can be accounted for by the notion that familiarity-based memories suffer from interference (Sadeh et al., 2016). Importantly, the task by topography interaction was found across task orders.

Although the current results could not speak to the role of expectedness in familiarity-based recognition, they confirmed the hypothesis about the importance of the task context derived from the unexpected fluency attribution account (Mecklinger & Bader, 2020). Moreover, they resonate well with other accounts of familiarity. Leynes et al. (2017) emphasize that whether fluency is perceived as fluency or attributed to another source depends on expectations that are modulated by task and context (albeit they do not explicitly distinguish relative from absolute familiarity). The current results are consistent with Leynes et al.'s notion inasmuch as greater fluency of repeated words elicited an FN400 only when the task entails the goal to remember and requires an attribution to oldness. The importance of an attribution mechanism for familiarity judgments is also acknowledged by Bastin et al. (2019) in their integrative memory model. This framework posits that reactivation of representations in the entity representation core system, which is located in structures along the visual ventral pathway, results in a fluency signal, which needs to be attributed to familiarity (or another source) by a separate attribution system. In Bastin et al.'s model, the attribution system provides the meta-cognitive knowledge that fluent processing can result from prior occurrence and can therefore be used to guide recognition judgments, a task taken over by the medial prefrontal cortex. Finally, the dorsolateral prefrontal cortex is proposed to monitor whether fluency signals are relevant to the task at hand and to transform the fluency signal into a familiarity judgment. Integrating this suggestion with the view that the FN400 is the electrophysiological marker of fluency attribution to familiarity dovetails with the proposition that the dorsolateral prefrontal cortex is one of the neural generators of the FN400 (Hoppstädter et al., 2015). Moreover, the FN400 can be moderated by top-down and attentional processes (Ecker et al., 2007; Rosburg et al., 2013) as would be predicted by the Bastin et al. model. Therefore, the unexpected fluency-attribution account (Mecklinger & Bader, 2020) constitutes a valuable extension to Bastin et al.'s model as it allows making inferences concerning the temporal characteristics of mnemonic attributions. It should be noted that Bastin et al. (2019) assume a more general attribution system which also attributes signals from a hippocampus-centered representational core system to recollection. The temporal and topographical dissociation of the ERP effects related to familiarity and recollection (Rugg & Curran, 2007) suggest at least partly differing attributional systems for the two processes. Therefore, future research needs to find out whether the whole set of ERP old/new effects can be mapped onto the integrative memory model proposed by Bastin et al. (2019).

The unexpected fluency attribution account makes the specific prediction that the presence of a retrieval intention is the prerequisite for an FN400 to be elicited. However, in this study, we compared an episodic recognition memory test with a non-mnemonic pleasantness task, which differ in more aspects than just the presence or absence of a retrieval intention (e.g., response format, decision criteria, or task difficulty). If we take response times as an indicator of task difficulty, our results could be reconciled with the view that the FN400 occurred only in the easier task. However, Ecker and Zimmer (2009) found an FN400 for similar foils only in a condition for which these foils had to be judged as "old." In this condition, response times were longer and not shorter than in the condition where similar foils had to be judged as "new" and where no FN400 for these items was elicited. Rosburg et al. (2013) also report an FN400 only in a retrieval situation in which target information is difficult to retrieve. This argues against the view that the FN400 is mainly elicited in easy task situations. Therefore, future research has to determine which aspects of a recognition decision are most relevant to elicit an FN400.

A final point that needs consideration is that amplitudes, both of the N400 as a reflection of conceptual fluency and of the FN400 as a reflection of an episodic familiarity signal, are not interpretable in an absolute sense. Due to component overlap, the ERPs in a particular time interval are also influenced by other aspects of the task. Therefore, both components need to be examined relative to a baseline measured under highly similar conditions (i.e., within the same task). Consistent with this view, we observed general ERP differences between tasks (generally more positive waveforms in the pleasantness task). This implies that theoretical assumptions about ERP components do not only pertain to the old items. Thus, for the FN400 we not only assume attribution to oldness as the underlying mechanism. Rather, we think deciding that there is no erroneous attribution to oldness for correct rejections also plays a role in generating the FN400. Subsequent studies might disentangle the separate contributions of old and new items to the FN400 and N400 effects.

To conclude, the current study dissociated the FN400 and the N400 effects by comparing an episodic recognition test with a pleasantness task that made no reference to memory under conditions in which both effects were driven by increased fluency signals for repeated old words. The results added to the evidence that the FN400 is dependent on the use of familiarity during a recognition decision and most likely reflects a relative familiarity signal for items with a surprisingly high fluency. Therefore, this study contributes to the understanding of familiarity-based recognition and its electrophysiological correlates.

AUTHOR CONTRIBUTIONS

Regine Bader: Conceptualization; data curation; formal analysis; methodology; supervision; validation; visualization; writing – original draft; writing – review and editing. **Luca Tarantini:** Conceptualization; data curation; formal analysis; methodology; validation; visualization; writing – original draft; writing – review and editing. **Axel Mecklinger:** Conceptualization; methodology; writing – review and editing.

ACKNOWLEDGMENT

Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data and analyses scripts are available on the Open Science Framework (OSF) platform at https://osf.io/f9hsa/?view_only=c1f731b5e12c49e39c0fccfedde599ff.

ORCID

Regine Bader https://orcid.org/0000-0001-7460-1720

REFERENCES

Bader, R., & Mecklinger, A. (2017). Separating event-related potential effects for conceptual fluency and episodic familiarity. *Journal of Cognitive Neuroscience*, *29*(8), 1402–1414. https://doi.org/10.1162/jocn_a_01131

Bader, R., Mecklinger, A., & Meyer, P. (2020). Usefulness of familiarity signals during recognition depends on test format: Neurocognitive evidence for a core assumption of the CLS framework. *Neuropsychologia*, *148*, 107659. https://doi.org/10.1016/j.neuropsychologia.2020.107659

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379–384. https://doi.org/10.3758/BF03192707

Bastin, C., Besson, G., Simon, J., Delhaye, E., Geurten, M., Willems, S., & Salmon, E. (2019). An integrative memory model of recollection and familiarity to understand memory deficits. *Behavioral and Brain Sciences*, *42*, e281. https://doi.org/10.1017/S0140525X19000621

Bridger, E. K., Bader, R., Kriukova, O., Unger, K., & Mecklinger, A. (2012). The FN400 is functionally distinct from the N400. *NeuroImage*, *63*(3), 1334–1342. https://doi.org/10.1016/j.neuroimage.2012.07.047

Bridger, E. K., Bader, R., & Mecklinger, A. (2014). More ways than one: ERPs reveal multiple familiarity signals in the word frequency mirror effect. *Neuropsychologia*, *57*, 179–190. https://doi.org/10.1016/j.neuropsychologia.2014.03.007

Butler, L. T., Berry, D. C., & Helman, S. (2004). Dissociating mere exposure and repetition priming as a function of word type. *Memory & Cognition*, *32*(5), 759–767. https://doi.org/10.3758/BF03195866

Coane, J. H., Balota, D. A., Dolan, P. O., & Jacoby, L. L. (2011). Not all sources of familiarity are created equal: The case of word frequency and repetition in episodic recognition. *Memory & Cognition*, *39*(5), 1–15. https://doi.org/10.3758/s13421-010-0069-5

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends in Cognitive Sciences*, *11*(9), 379–386. https://doi.org/10.1016/j.tics.2007.08.001

Ecker, U. K. H., & Zimmer, H. D. (2009). ERP evidence for flexible adjustment of retrieval orientation and its influence on familiarity. *Journal of Cognitive Neuroscience*, *21*(10), 1907–1919. https://doi.org/10.1162/jocn.2009.21135

Ecker, U. K. H., Zimmer, H. D., Groh-Bordin, C., & Mecklinger, A. (2007). Context effects on familiarity are familiarity effects of context—An electrophysiological study. *International Journal of Psychophysiology*, *64*(2), 146–156. https://doi.org/10.1016/j.ijpsycho.2007.01.005

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition*, *13*(1), 8–20. https://doi.org/10.3758/BF03198438

Guillaume, F., & Tiberghien, G. (2013). Impact of intention on the ERP correlates of face recognition. *Brain and Cognition*, *81*(1), 73–81.

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). DlexDB–eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, *62*, 10–20. http://econtent.hogrefe.com/doi/full/10.1026/0033-3042/a000029

Hines, D., Czerwinski, M., Sawyer, P. K., & Dwyer, M. (1986). Automatic semantic priming: Effect of category exemplar level and word association level. *Journal of Experimental Psychology: Human Perception and Performance*, *12*(3), 370–379. https://doi.org/10.1037/0096-1523.12.3.370

Hockley, W. E. (1991). Recognition memory for item and associative information: A comparison of forgetting rates. In S. Lewandowsky & W. E. Hockley (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 227–248). Psychology Press.

Hoppstädter, M., Baeuchl, C., Diener, C., Flor, H., & Meyer, P. (2015). Simultaneous EEG–fMRI reveals brain networks underlying recognition memory ERP old/new effects. *NeuroImage*, *116*, 112–122. https://doi.org/10.1016/j.neuroimage.2015.05.026

Horner, A. J., & Henson, R. N. (2009). Bindings between stimuli and multiple response codes dominate long-lag repetition priming in speeded classification tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 757–779. https://doi.org/10.1037/a0015262

Hughes, A. D., & Whittlesea, B. W. (2003). Long-term semantic transfer: An overlapping-operations account. *Memory & Cognition*, *31*(3), 401–411. https://doi.org/10.3758/BF03194398

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger, III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Erlbaum.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lawrence, M. A. (2016). *Ez: Easy analysis and visualization of factorial experiments. R package version 4.4-0*. https://CRAN.R-project.org/package=ez

Leynes, P. A., & Addante, R. J. (2016). Neurophysiological evidence that perceptions of fluency produce mere exposure effects. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(4), 754–767. https://doi.org/10.3758/s13415-016-0428-1

Leynes, P. A., Bruett, H., Krizan, J., & Veloso, A. (2017). What psychological process is reflected in the FN400 event-related potential component? *Brain and Cognition*, *113*, 142–154. https://doi.org/10.1016/j.bandc.2019.103579

Leynes, P. A., & Mok, B. A. (2020). Context influences the FN400 recognition event-related potential. *International Journal of Psychophysiology*, *158*, 16–26. https://doi.org/10.1016/j.ijpsycho.2020.09.006

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252–271. https://doi.org/10.1037/0033-295X.87.3.252

McCarthy, G., & Wood, C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, *62*, 203–208. https://doi.org/10.1016/0168-5597(85)90015-2

Mecklinger, A., & Bader, R. (2020). From fluency to recognition decisions: A broader view of familiarity-based remembering. *Neuropsychologia*, *146*, 107527. https://doi.org/10.1016/j.neuropsychologia.2020.107527

Mecklinger, A., Brunnemann, N., & Kipp, K. (2010). Two processes for recognition memory in children of early school age: An event-related potential study. *Journal of Cognitive Neuroscience*, *23*(2), 435–446. https://doi.org/10.1162/jocn.2010.21455

Mecklinger, A., Frings, C., & Rosburg, T. (2012). Response to Paller et al.: The role of familiarity in making inferences about unknown quantities. *Trends in Cognitive Sciences*, *16*(6), 315–316. https://doi.org/10.1016/j.tics.2012.04.009

Migo, E., Montaldi, D., Norman, K. A., Quamme, J., & Mayes, A. (2009). The contribution of familiarity to recognition memory is a function of test format when using similar foils. *The Quarterly Journal of Experimental Psychology*, *62*(6), 1198–1215. https://doi.org/10.1080/17470210802391599

Montaldi, D., & Mayes, A. R. (2010). The role of recollection and familiarity in the functional differentiation of the medial temporal lobes. *Hippocampus*, *20*(11), 1291–1314. https://doi.org/10.1002/hipo.20853

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, *110*(4), 611–646. https://doi.org/10.1037/0033-295X.110.4.611

Paller, K. A., Lucas, H. D., & Voss, J. L. (2012). Assuming too much from 'familiar' brain potentials. *Trends in Cognitive Sciences*, *16*(6), 313–315. https://doi.org/10.1016/j.tics.2012.04.010

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2020). *nlme: Linear and nonlinear mixed effects models. R package version 3.1-149*. https://CRAN.R-project.org/package=nlme

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Revelle, W. (2019). *psych: Procedures for personality and psychological research. R package version 1.9.12*. https://CRAN.R-project.org/package=psych

Rosburg, T., Johansson, M., & Mecklinger, A. (2013). Strategic retrieval and retrieval orientation in reality monitoring studied by event-related potentials (ERPs). *Neuropsychologia*, *51*(3), 557–571. https://doi.org/10.1016/j.neuropsychologia.2012.11.014

RStudio Team. (2019). *RStudio: Integrated development environment for R*. RStudio, Inc. http://www.rstudio.com/

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251–257. https://doi.org/10.1016/j.tics.2007.04.004

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*, 595–598. https://doi.org/10.1038/33396

Sadeh, T., Ozubko, J. D., Winocur, G., & Moscovitch, M. (2016). Forgetting patterns differentiate between two forms of memory representation. *Psychological Science*, *27*(6), 810–820. https://doi.org/10.1177/0956797616638307

Schröder, A., Gemballa, T., Ruppin, S., & Wartenburger, I. (2012). German norms for semantic typicality, age of acquisition, and concept familiarity. *Behavior Research Methods*, *44*, 380–394. https://doi.org/10.3758/s13428-011-0164-y

Stenberg, G., Hellman, J., Johansson, M., & Rosén, I. (2009). Familiarity or conceptual priming: Event-related potentials in name recognition. *Journal of Cognitive Neuroscience*, *21*(3), 447–460. https://doi.org/10.1162/jocn.2009.21045

Stróżak, P., Abedzadeh, D., & Curran, T. (2016). Separating the FN400 and N400 potentials across recognition memory experiments. *Brain Research*, *1635*, 41–60. https://doi.org/10.1016/j.brainres.2016.01.015

Tukey, J. W. (1977). *Exploratory data analysis*. Addison Wesley.

Vilberg, K. L., Moosavi, R. F., & Rugg, M. D. (2006). The relationship between electrophysiological correlates of recollection and amount of information retrieved. *Brain Research*, *1122*(1), 161–170. https://doi.org/10.1016/j.brainres.2006.09.023

Voss, J. L., & Federmeier, K. D. (2011). FN400 potentials are functionally identical to N400 potentials and reflect semantic processing during recognition testing. *Psychophysiology*, *48*(4), 532–546. https://doi.org/10.1111/j.1469-8986.2010.01085.x

Whittlesea, B. W., & Price, J. R. (2001). Implicit /explicit memory versus analytic/nonanalytic processing: Rethinking the mere exposure effect. *Memory & Cognition*, *29*(2), 234–246. https://doi.org/10.3758/BF03194917

Whittlesea, B. W., & Williams, L. D. (2001a). The discrepancy-attribution hypothesis: I. The heuristic basis of feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 3–13. https://doi.org/10.1037/0278-7393.27.1.3

Whittlesea, B. W., & Williams, L. D. (2001b). The discrepancy-attribution hypothesis: II. Expectation, uncertainty, surprise, and feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 14–33. https://doi.org/10.1037/0278-7393.27.1.14

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, *21*, 1–20. https://doi.org/10.18637/jss.v021.i12

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., & Hester, J. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research*, *1100*, 125–135. https://doi.org/10.1016/j.brainres.2006.05.019

Woollams, A. M., Taylor, J. R., Karayanidis, F., & Henson, R. N. (2008). Event-related potentials associated with masked priming of test cues reveal multiple potential contributions to recognition memory. *Journal of Cognitive Neuroscience*, *20*(6), 1114–1129. https://doi.org/10.1162/jocn.2008.20076

Yang, H., Laforge, G., Stojanoski, B., Nichols, E. S., McRae, K., & Köhler, S. (2019). Late positive complex in event-related potentials tracks memory signals when they are decision relevant. *Scientific Reports*, *9*(1), 1–15. https://doi.org/10.1038/s41598-019-45880-y

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., & Levy, B. J. (2002). Dissociating familiarity from recollection in human recognition memory: Different rates of forgetting over short retention intervals. *Psychonomic Bulletin & Review*, *9*(3), 575–582. https://doi.org/10.3758/BF03196315

Yu, S. S., & Rugg, M. D. (2010). Dissociation of the electrophysiological correlates of familiarity strength and item repetition. *Brain Research*, *1320*, 74–84. https://doi.org/10.1016/j.brainres.2009.12.071

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2p2), 1–27. https://doi.org/10.1037/h0025848

How to cite this article: Bader, R., Tarantini, L., & Mecklinger, A. (2023). Task context dissociates the FN400 and the N400. *Psychophysiology*, *00*, e14258. https://doi.org/10.1111/psyp.14258