Original Article

QJEP

Quarterly Journal of Experimental Psychology
2020, Vol. 73(3) 375–383

DOI: 10.1177/1747021819878385
qjep.sagepub.com
SAGE

On the limits of familiarity accounts in lexical decision: The case of repetition effects

Manuel Perea[1,2], Ana Marcet[1], Marta Vergara-Martínez[1] and Pablo Gomez[3]

Abstract

Recent modelling accounts of the lexical decision task have suggested that the reading system performs evidence accumulation to carry out some functions. Evidence accumulation models have been very successful in accounting for effects in the lexical decision task, including the dissociation of repetition effects for words and nonwords (facilitative for words but inhibitory for nonwords). The familiarity of a repeated item triggers its recognition, which facilitates 'word' responses but hampers nonword rejection. However, reports of facilitative repetition effects for nonwords with several repetitions in short blocks challenge this hypothesis and favour models based on episodic retrieval. To shed light on the nature of the repetition effects for nonwords in lexical decision, we conducted four experiments to examine the impact of extra-lexical source of information—we induced the use of episodic retrieval traces via instructions and list composition. When the initial block was long, the repetition effect for nonwords was inhibitory, regardless of the instructions and list composition. However, the inhibitory effect was dramatically reduced when the initial block included two presentations of the stimuli and it was even facilitatory when the initial block was short. This composite pattern suggests that evidence accumulation models of lexical decision should take into account all sources of evidence—including episodic retrieval—during the process of lexical decision.

Keywords
Lexical decision; evidence accumulation models; word recognition; repetition

Received: 13 August 2018; revised: 7 August 2019; accepted: 25 August 2019

A key issue in reading research is how repetition affects performance. To tackle this question, we focused on the most representative task in word recognition research (i.e., lexical decision) for which there is a robust tradition of implementing possible explanations of phenomena as process models. Many of these process models assume that the response made in a trial (i.e., "word" vs. "nonword") is a function of the degree of wordness of the letter string (e.g., two-stage model: Balota & Chumbley, 1984; dual-route cascaded model: Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; multiple read-out model: Grainger & Jacobs, 1996; lexical decision account of the diffusion model: Ratcliff, Gomez, & McKoon, 2004; REM-LD model: Wagenmakers, Steyvers, et al., 2004; Bayesian reader model: Norris, 2006).

Previous research suggests that the lexical decision task can be described as an evidence accumulation mechanism, in particular, as a drift-diffusion process (Ratcliff et al., 2004; see also Dutilh, Vandekerckhove, Tuerlinckx, & Wagenmakers, 2009; Gomez, 2012; Gomez & Perea, 2014; Gomez, Perea, & Ratcliff, 2013; Oganian et al., 2016; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008; Yap, Sibley, Balota, Ratcliff, & Rueckl, 2015, among others). The diffusion model has been very successful in accounting for performance in an extensive list of lexical decision manipulations (e.g., perceptual vs. lexical effects, masked vs. unmasked priming, etc.). When accounting for lexical decision data, the assumption implemented in the diffusion model has been that the degree of wordness of the letter string drives the decision process towards a positive

[1]Departamento de Metodología, Universitat de València, Valencia, Spain
[2]Universidad Nebrija, Madrid, Spain
[3]DePaul University, Chicago, IL, USA

Corresponding author:
Manuel Perea, Departamento de Metodología, Universitat de València, Avenue Blasco Ibáñez, 21, 46010 Valencia, Spain.
Email: mperea@uv.es

("word") or negative ("nonword") boundary: this can be captured as the drift-rate parameter in the model.

In the diffusion model,[1] the degree of wordness in the decision process is assumed to integrate "semantic information, phonological information, orthographic information, and other kinds of lexical information" (Ratcliff et al., 2004, p. 170). The higher the wordness a letter string has, the higher its drift rate. The emblematic illustration is the word-frequency effect: words that are encountered frequently (e.g., high-frequency words like MUSIC) generate more positive drift rates than words that are encountered rarely (e.g., DECOY). As a result, for "word" responses, the model predicts shorter response times (RTs), higher accuracy, and less asymmetrical RT distributions for high- than for low-frequency words. Likewise, wordlike nonwords (e.g., COUSE) generate less negative drift rates than non-wordlike nonwords (e.g., UDIAM or FGTER). Therefore, for "nonword" responses, the model predicts longer latencies, lower accuracy, and more asymmetrical RT distributions for wordlike nonwords than for non-wordlike nonwords (see Ratcliff et al., 2004, for fits).

Importantly, the diffusion model account of lexical decision can be used to make predictions not only on word-frequency effects but also about repetition effects. Ratcliff et al. (2004) showed that drift rates increase "with word frequency *and* with repetition" (p. 175, emphasis added). Obviously, the word-frequency effect can be thought of as a special case of the repetition effect: a high-frequency word is an item that has been presented on multiple occasions (i.e., it has a high degree of wordness), thus making its drift rate more positive in a word versus nonword discrimination task. We refer to this as the familiarity hypothesis. An important corollary of this hypothesis is that if repeating an item generates a larger degree of wordness than a nonrepeated item, drift rates for repeated items should be more positive regardless of their lexical status. Therefore, the familiarity hypothesis predicts a dissociation of repetition effects for words and nonwords in lexical decision: repeated words would produce faster responses, higher accuracy, and less skewed RT distributions than nonrepeated words; in contrast, repeated nonwords would produce slower responses, lower accuracy, and more skewed RT distributions than nonrepeated nonwords.

A number of experiments have reported this dissociation (i.e., facilitative repetition effects for words; inhibitory repetition effects for nonwords; see Balota & Spieler, 1999; Bowers, 1994; McKoon & Ratcliff, 1979; Perea, Marcet, Vergara-Martínez, & Gomez, 2016). (For consistency with prior experiments, we will use the term "nonword" even though the nonwords used in the experiments are orthographically legal pseudowords.) As an illustration, Perea et al. (2016) fit the diffusion model to data from a two-block lexical decision experiment. In Block 1, participants were presented with 100 words and 100 nonwords. In Block 2, half of the items were repeated from Block 1 (50 words and 50 nonwords), and the other half was new. The word/nonword dissociation of repetition effects was readily captured by changes in the drift-rate parameter: drift rates were more positive for repeated than for nonrepeated items (words: 0.36 vs. 0.32; nonwords: –0.31 vs. –0.37, respectively).

However, a number of published reports have found a facilitative repetition effect for both words and nonwords in lexical decision when there are several repetitions at short blocks (e.g., Logan, 1990; see Zeelenberg, Wagenmakers, & Shiffrin, 2004, for review). Logan (1990) claimed that this outcome ruled out familiarity accounts (e.g., Balota & Chumbley, 1984, two-stage model). Instead, Logan interpreted his findings as supporting episodic retrieval for word and nonword responses in the framework of the instance theory. A basic tenet of Logan's (1990) instance theory—which "is intended to apply to all initial tasks" (p. 5)—is that, in the initial presentation, each word and nonword creates a new episodic association (e.g., "CABLE is a word"; "LIUSE is a nonword"). In a subsequent presentation, participants could retrieve these associations to speed up both word and nonword responses. This extra-lexical mechanism would contribute not only to word responses but also to nonword responses to both word and nonword responses (i.e., "the instance theory predicts benefit for nonwords"; Logan, 1990, p. 26).

In an attempt to reconcile these contradictory findings, Zeelenberg et al. (2004) suggested that repetition effects for nonwords in lexical decision could reflect a mixture of two opposing processes: (1) an inhibitory process due to an increase in wordness—as posited by familiarity accounts and (2) a facilitative process due to episodic retrieval—as assumed by the instance theory. Zeelenberg et al. suggested that fast lexical decision responses would be (more likely) driven by the wordness of the letter string, whereas slow responses would be (more likely) driven by episodic retrieval (e.g., "LIUSE is a nonword"). To test their proposal, Zeelenberg et al. (2004; Experiment 2) ran a two-block experiment with few stimuli (16 words and 16 nonwords) that were repeated two times in the initial block and manipulated the instructions (i.e., standard vs. speeded instructions). When using standard instructions, Zeelenberg et al. (2004) found a facilitative repetition effect for both words and nonwords—as predicted episodic accounts. However, when participants were asked to prioritise response speed, the repetition effect was facilitative for words but inhibitory for nonwords (45.8 vs. 39.9% of errors for repeated and nonrepeated nonwords, respectively)—as predicted by familiarity accounts (see also Wagenmakers, Zeelenberg, Steyvers, Shiffrin, & Raaijmakers, 2004, for an inhibitory effect of nonword repetition using a signal-to-respond paradigm in lexical decision).

Clearly, the findings from Zeelenberg et al. (2004) with two repetitions and short blocks demonstrate that the interplay between familiarity and episodic retrieval mechanisms with repeated items is dynamic and calls for a more complex balance of these forces when modelling repetition effects in lexical decision. To integrate these changes into a more comprehensive model of word recognition in lexical decision, it is necessary to address how these two components (i.e., familiarity and episodic retrieval) interplay in a broader scenario.[2] To achieve this goal, we conducted a series of lexical decision experiments intended to induce episodic retrieval with long blocks—for comparison purposes, we used a design parallel to that used by Perea et al. (2016). In Experiment 1, participants were told in advance about the presence of repeated stimuli and that this information could help speed up their responses upon presentation of the repeated stimuli. In Experiment 2, we used relatively infrequent words with the same or different background colour as in the initial presentation. In Experiment 3, we induced episodic retrieval by presenting each item (word/nonword) twice in the initial block. The familiarity account would predict a facilitative repetition effect for words and inhibitory for nonwords across experiments. Alternatively, if participants retrieve the specific instances of repeated items during the second presentation of the items (e.g., "LIUSE is a nonword"), the repetition effects would be facilitative not only for words but also for nonwords (see Logan, 1990). In this case, the familiarity accounts would need to be modified to integrate extra-lexical sources of information such as episodic retrieval in the decision processes. Finally, in an effort to obtain a comprehensive view on the word/nonword dissociation of repetition effects in lexical decision, we conducted Experiment 4. It was parallel to Experiment 3 except that the two blocks were much shorter (Block 1 was composed of 20 words and 20 nonwords that were presented twice)—note that this produced a scenario similar to Zeelenberg et al.'s (2004) Experiment 2 with standard lexical decision instructions.

Experiment 1 (long blocks: informed participants)

The materials and design were the same as in the Perea et al. (2016) experiment, except for a key difference: we induced participants to rely on episodic retrieval to make lexical decisions via instructions. Previous research has shown that subtle changes in the information provided in the instructions such as mentioning or not that prime-target pairs are related in a priming lexical decision task have an important impact in the obtained effects (e.g., see Zeelenberg, Pecher, De Kok, & Raaijmakers, 1998).

Method

Participants. Twenty-four students from the University of Valencia volunteered to participate in the experiment. In this and subsequent experiments, all participants were native speakers of Spanish with normal/corrected-to-normal vision and were provided written informed consent before starting the experimental session.

Materials. The materials were the same as in the Perea et al. (2016) experiment. The experimental stimuli were 100 Spanish words of high frequency ($M=195$ per million words in the Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013, subtitle corpus) and 100 orthographically legal nonwords. The average length was 6.1 letters (range 5–8) for both word and nonword stimuli. Block 2 was composed of the 100 experimental words and the 100 experiments nonwords, half of which (50 words and 50 nonwords) were also presented in Block 1 and the other half was new. To create Block 1, which was also composed of 200 trials, we randomly paired half of the stimuli in Block 2 (50 words and 50 nonwords) with a set of filler items in List 1 (50 high-frequency words and 50 nonwords), whereas the other half of experimental stimuli were paired with the filler items in List 2. Half of the participants were assigned to List 1 and the other half to List 2.

Procedure. The procedure was the same as in the Perea et al. (2016) experiment (i.e., a two-block lexical decision experiment using DmDX [Forster & Forster, 2003] software with standard lexical decision instructions), with an important addition: that participants were explicitly informed about the presence of repeated items (words and nonwords) before Block 1 with the following instruction "IMPORTANT: half of the items in this block will be repeated in the second block of the experiment. This information will help you make the decisions in the second block." In addition, they were presented with the following information before the beginning of Block 2: "IMPORTANT: Please remember that half of the items in this block were presented in the first block. This information may help you make the decisions." Participants did not know in advance which specific items would be repeated.

Results and discussion

As in the Perea et al. (2016) experiment, error responses and lexical decision times faster than 250 ms or slower than 1,500 ms (0.47% of the data) were removed from the latency analyses. The mean RTs and error rates per condition in Block 2 are displayed in Table 1. To analyse the data, we employed linear mixed effects models in R (lmer for the latency data, glmer for the accuracy data; Bates, Maechler, Bolker, & Walker, 2015). The fixed factors in the statistical analyses were Repetition (repeated vs. non-repeated) and Lexicality (word vs. nonword), each of which was coded as –.5 and .5 to be zero-centred. RTs were transformed (–1,000/RT) to comply with the Gaussian requirement of linear mixed effects models. We

Table 1. Mean lexical decision times (in ms) and error rates on words and nonwords in Block 2 of Experiment 1.

	Repeated		Nonrepeated		Nonrepeated−repeated	
	RT	ER	RT	ER	RT	ER
Words	533	2.6	549	4.2	16	1.6
Nonwords	614	3.0	597	4.8	–17	–1.8

RT: response time; ER: error rate.

employed the maximal random structure model that converged. In the latency data, this model was −1,000/RT = repetition × lexicality + (repetition × lexicality + 1|subject) + (repetition + 1|item). The p values in the RT analyses were obtained via the Satterthwaite approximation to the degrees of freedom (Kuznetsova, Brockhoff, & Christensen, 2017). In the analyses of the accuracy data, we employed the parallel strategy with generalised linear models (glmer): accuracy data were coded as 1 = correct and 0 = incorrect.

Unsurprisingly, the statistical analyses showed faster lexical decision times for words than for nonwords (541 vs. 606 ms, respectively), $t = 7.60$, $p < .001$. The overall difference between the repeated and nonrepeated items was minimal (575 vs. 574 ms, respectively, $t < 1$). In addition, there was a significant interaction between Repetition and Lexicality, $t = –3.61$, $p < .001$. This reflected a facilitative 16 ms repetition effect for words, $t = −2.59$, $p = .010$, and an inhibitory 17 ms repetition effect for nonwords, $t = 2.51$, $p = .014$.

We also examined whether the magnitude of repetition effects increased across quantiles in the RT distributions. For words, the magnitude of the repetition effect for the .1, .3, .5, .7, and .9 quantiles was 4, 5, 9, 16, and 48 ms, respectively. For nonwords, the magnitude of the repetition effect for the .1, .3, .5, .7, and .9 quantiles was −2, −3, −10, −24, and −49 ms, respectively.

The accuracy data did not show any main effects of repetition or lexicality (both zs < 0.1), but it showed a significant interaction between the two factors, $z = 3.08$, $p = .002$. This interaction emerged due to an inhibitory repetition effect for nonwords, $z = 2.07$, $p = .038$, and a small nonsignificant facilitative repetition effect for words, $z = –1.22$, $p = .224$.

In sum, when participants knew in advance that there would be repeated items in the experiment, the pattern of repetition effects (i.e., facilitation for words and inhibition for nonwords) mimicked that obtained with uninformed participants (e.g., see Balota & Spieler, 1999; Perea et al., 2016). Furthermore, similar to the Balota and Spieler (1999) and the Perea et al. (2016) experiments, the magnitude of the repetition effect was larger in the higher quantiles than in lower quantiles. In other words, explicit instructions about the nature of the experiment (i.e., the presence of repeated items) do not seem to influence the core processes underlying lexical decisions.

Experiment 2 (long blocks: colour information and low-frequency words)

In this experiment, we examined whether the dissociation of word/nonword repetition effects in lexical decision also occurs with low-frequency words. Thus, the research question was whether the word/nonword dissociation in repetition effects changed when responses were slowed down (i.e., when the whole set of word stimuli was relatively infrequent). We used the same design as in the Perea et al. (2016) experiment except that the word stimuli were of low frequency ($M = 7$ occurrences per million)—Perea et al. (2016) used very high–frequency words ($M = 195$ occurrences per million).

The rationale for this manipulation is twofold. First, when only low-frequency words are included in the word set, the overall pace of the experiment is reduced because word/nonword lexical decision responses would be slower. The idea is that slow responses would allow more room for episodic retrieval (i.e., the alleged facilitative component for repeated nonwords; see Zeelenberg et al., 2004). Second, the effect of word-frequency on recognition memory is well known: lower frequency words are easier to recognise than higher frequency words (e.g., Glanzer & Bowles, 1976). While an expected outcome is a larger (facilitative) repetition effect for low-frequency words, efficient use of episodic retrieval cues in lexical decision may also help the responses to nonwords, thus reducing the inhibitory repetition effect for nonwords or even turn it to facilitative.

Besides using low-frequency words, the current experiment has a subtle difference with respect to the Perea et al. (2016) experiment: words/nonwords across Blocks 1 and 2 were presented with either the same background colour (e.g., TUTOR and TUTOR; respectively) or a different one (e.g., and TUTOR; respectively). The rationale is that the use of episodic retrieval cues could be more likely if repeated items share a salient feature from the initial block (yellow background; for example, see Dulsky, 1935, for early evidence of a facilitative role of shared background colour when learning nonsense syllables).

Method

Participants. The sample was composed of 28 students from the same population as in Experiment 1. None of them had taken part in the previous experiment.

Table 2. Mean RTs (in ms) and error rates on words and nonwords in Block 2 of Experiment 2.

	Repeated		Nonrepeated		Nonrepeated – repeated	
	RT	ER	RT	ER	RT	ER
Words						
Yellow background	608	2.1	629	3.3	21	1.2
White background	610	3.6	635	3.7	15	0.1
Nonwords						
Yellow background	688	5.0	672	5.3	–15	0.3
White background	696	5.7	684	5.1	–12	–0.6

RT: response time; ER: error rate. In Block 1, all the to-be-repeated items were presented in yellow background (i.e., "Yellow background" in Block 2 would mean that these items were presented with the same background in the two blocks).

Materials. The set of experimental stimuli was composed of 100 Spanish words of low frequency, all of the nouns ($M=7.1$ per million words in the Duchon et al., 2013, subtitle corpus; range 5.0–9.9) and 100 orthographically legal nonwords created with Wuggy (Keuleers & Brysbaert, 2010). The average length was 6.5 letters (range 5–8) for both word and nonword stimuli. In addition, we employed 50 filler low-frequency words ($M=7.1$ per million words) and 50 filler orthographically legal nonwords for Block 1—the average length was the same as the experimental trials (6.5; range 5–8 letters). The experimental design was the same as in Experiment 1 and in the Perea et al. (2016) experiment, except that in Block 1, the experimental stimuli were presented in yellow background (e.g., TUTOR) and the filler stimuli were presented in (standard) white background (e.g., RIVAL). In Block 2, half of the repeated and nonrepeated words/nonwords were presented in yellow background (i.e., background match) and the other half in white background (i.e., background mismatch). That is, half of the repeated items were presented in the same background colour as the initial presentation, and the other half were presented in a different background colour. We created four experimental lists to rotate the words/nonwords across Repetition and Background colour. The complete set of words and nonwords in this experiment is provided in the Online Supplementary Material.

Procedure. It was exactly the same as in the Perea et al. (2016) experiment (i.e., participants were not informed of the presence of repeated items in the experiment).

Results and discussion

Incorrect responses and RTs shorter than 250 ms or longer than 1,500 ms (0.83% of the data) were removed from the latency analyses. The mean RTs and error rates per condition in Block 2 are presented in Table 2. The statistical analyses paralleled those in Experiment 1, except that we also included background colour (yellow, white) as a fixed factor.

The statistical analyses showed shorter lexical decision times for words than for nonwords, $t=6.00$, $p<.001$. The main effect of repetition did not reach statistical significance, $t=1.93$, $p=.065$. As in Experiment 1, we found a significant interaction between Repetition and Lexicality, $t=-5.40$, $p<.001$: words showed a facilitative 23 ms repetition effect, $t=5.13$, $p<.001$, whereas nonwords showed an inhibitory −14 ms repetition effect, $t=-2.52$, $p=.018$. The main effect of background colour was not significant, $t=1.67$, $p=.093$, and none of the interactions of this factor with repetition or lexicality approached significance (all ts $<.81$, all ps $>.41$).

We also assessed whether the size of repetition effects augmented across quantiles in the RT data. For words, the size of the repetition effect for the .1, .3, .5, .7, and .9 quantiles was 10, 5, 15, 20, and 47 ms, respectively. For nonwords, the size of the repetition effect for the .1, .3, .5, .7, and .9 quantiles was 1, −7, −9, −15, and −47 ms, respectively.

The accuracy data only showed more errors to nonwords than to words, $z=-3.13$, $p=.002$—the other effects had zs <0.9.

In sum, the present lexical decision experiment showed that including only low-frequency words in the set of words does not alter the dissociation of repetition effects for words and nonwords. As occurred in Experiment 1 and in the Balota and Spieler (1999) and the Perea et al. (2016) experiments, the repetition effect was magnified in the tail of the RT distributions. Furthermore, the use of background colour as an episodic cue did not alter the pattern of repetition effects: repetition effects for words and nonwords were similar in size regardless of whether the initial and second presentations shared the background colour.

Experiment 3 (long blocks: two repetitions)

This experiment induced participants to retrieve episodic cues by repeating each item twice in the initial block. We employed the same stimuli as in Experiment 1 except that Block 1 was repeated twice. In this scenario, participants

Table 3. Mean lexical decision times (in ms) and error rates on words and nonwords in Block 2 of Experiment 3.

	Repeated		Nonrepeated		Nonrepeated–repeated	
	RT	ER	RT	ER	RT	ER
Words	541	3.5	565	4.8	24	1.3
Nonwords	606	5.0	606	3.5	0	−1.5

RT: response time; ER: error rate.

would notice the repetition of the items in Block 1, and they could use these episodic cues to speed up the responses to both words and nonwords in Block 2 (see Logan, 1990). If so, one would expect a reduced inhibitory or even a facilitative repetition effect for nonwords. Alternatively, if the repetition of the stimuli in the initial block merely increases their wordness, we would expect a strong inhibitory repetition effects for nonwords—even greater than in Experiment 1. The idea is that nonwords presented twice in Block 1 (e.g., LOUSE; LOUSE) would have a much higher degree of wordness than nonrepeated nonwords, thus hindering "no" lexical decision responses.

Method

Participants. The sample was composed of 24 students from the same population as in Experiments 1 and 2. None of them had taken part in the previous experiments.

Materials. The materials were the same as in Experiment 1. The only difference was that Block 1 was presented twice.

Procedure. The instructions were the same as in Experiment 1, except that participants were not informed of the presence of repeated stimuli.

Results and discussion

As in Experiments 1 and 2, correct RTs beyond the 250 to 1,500 ms cutoffs (0.19% of the data) and error responses were not entered in the latency analyses. The average RTs and error rates per condition in Block 2 are presented in Table 3. The design and inferential analyses were parallel to those of Experiment 1.

The analyses of the lexical decision times showed that responses were, on average, faster for words than for nonwords (553 vs. 606 ms, respectively), $t=6.21$, $p<.001$. In addition, repeated items were responded to faster than unrepeated items, $t=3.13$, $p=.005$ (573.5 vs. 585.5 ms, respectively). This main effect of repetition was modulated by a significant interaction between Repetition and Lexicality ($t=-4.29$, $p<.001$): there was a facilitative 24 ms repetition effect for words, $t=-5.05$, $p<.001$, but there were no signs of a repetition effect for nonwords (less than 1 ms), $t<1$.

Similarly to Experiments 1 and 2, we computed the .1, .3, .5, .7, and .9 quantiles in the RT distributions. Unsurprisingly, the repetition effect for words increased at the higher quantiles: 12, 15, 22, 30, and 45 ms at the .1, .3, .5, .7, and .9 quantiles, respectively. Importantly, unlike previous experiments, there was no well-defined pattern across quantiles for the repetition effect on nonwords: –5, 2, 10, 3, and 1 ms at the .1, .3, .5, .7, and .9 quantiles, respectively.

The analyses of the accuracy data failed to show any main effects of repetition or lexicality (both zs <1), whereas it showed a significant interaction between the two factors, $z=2.54$, $p=.01$: there was a nonsignificant inhibitory repetition effect for nonwords, $z=1.72$, $p=.085$, whereas there was a nonsignificant facilitative repetition effect for words, $z=-1.79$, $p=.07$.

When the items were repeated twice in Block 1, we failed to find any signs of a repetition effect for nonwords in the RT data, whereas there was an inhibitory trend close to significance ($p=.085$) in the error data. This overall null finding poses problems for Logan's (1990) instance theory of repetition effects—it predicted a facilitative effect for nonwords. Importantly, this finding also poses problems for the familiarity account—it predicted stronger (not weaker) inhibitory repetition effects for nonwords.

All in all, Experiments 1 to 3 did not show a facilitative repetition effect for nonwords with long blocks. This raises a question on the divergence between the present data and the Zeelenberg et al. (2004) facilitative repetition effects for nonwords at short blocks with standard lexical decision instructions. To examine this issue, Experiment 4 included much shorter blocks—we selected 20 words and 20 nonwords from the items employed in Experiment 3—in a two-block design, thus mimicking Zeelenberg et al.'s (2004) study. An episodic account would predict a facilitative effect of repetition for both words and nonwords, whereas a familiarity account would predict a facilitative repetition effect for words and an inhibitory repetition effect for nonwords.

Experiment 4 (short blocks; two repetitions)

Method

Participants. The sample was composed of 28 students from the same population as in the previous experiments. None of them had taken part in the previous experiments.

Table 4. Mean lexical decision times (in ms) and error rates on words and nonwords in Block 2 of Experiment 4.

	Repeated		Nonrepeated		Nonrepeated – repeated	
	RT	ER	RT	ER	RT	ER
Words	541	2.0	563	3.4	22	1.4
Nonwords	606	3.8	630	4.3	24	0.5

RT: response time; ER: error rate.

Materials. We randomly selected 40 target words and 40 target nonwords from the materials of Experiment 1. Block 1 was composed of 20 words and 20 nonwords, whereas Block 2 was composed of the same 40 items as in the initial block plus a new set of 40 items (20 words and 20 nonwords). As in previous experiments, we created two sets of lists to counterbalance the materials (i.e., if a word was repeated in List 1, it would be presented only in Block 2 in List 2).

Procedure. The procedure was the same as in Experiment 3 (i.e., Block 1 was presented twice).

Results and discussion

As in Experiments 1 to 3, we removed error responses as well as those correct responses beyond the 250 to 1,500 ms cutoff (0.13% of the data) in the latency analyses. The mean lexical decision times and error rates per condition in Block 2 are presented in Table 4. The statistical analyses were the same as those in Experiments 1 and 3.

The analyses of the latency data showed that responses were, on average, faster for words than for nonwords (552 vs. 618 ms, respectively), $t=7.67$, $p<.001$, and that repeated items were responded to faster than unrepeated items, $t=4.15$, $p<.001$ (573.5 vs. 596.5 ms respectively). The two effects were additive, as deduced by the lack of interaction between Repetition and Lexicality ($t<1$, $p=.37$).

We also computed the .1, .3, .5, .7, and .9 quantiles in the RT distributions. Although the repetition effect for words occurred across the entire RT distribution, it increased slightly at the higher quantiles: 18, 17, 19, 26, and 29 ms at the .1, .3, .5, .7, and .9 quantiles, respectively. In addition, the facilitative repetition effect for nonwords increased at the higher quantiles: 1, 11, 19, 28, and 60 ms at the .1, .3, .5, .7, and .9 quantiles, respectively.

The analyses of the accuracy data failed to show any significant effects (all $ps>.11$).

Thus, when items are repeated twice in a short block, the effects of repetition are facilitative not only for words but also for nonwords (22 vs. 24 ms, respectively), thereby replicating the repetition effect reported by Zeelenberg et al. (2004).

General discussion

The idea of evidence being accumulated to make a lexical decision is appealing and has been implemented successfully in a number of computational process models (e.g., Ratcliff et al., 2004, diffusion model). A critical question in such implementations is the nature and format of the evidence that is being accumulated. The idea that more familiar items have more "wordness" is consistent with how researchers have thought about the rate of evidence accumulation in the lexical decision task. Importantly, when nonwords are repeated, they become more familiar (i.e., more wordlike), but at the same time they become more associated with the nonword response (i.e., less wordlike). The present set of experiments shows that the *+familiarity* yields *+wordness* mechanism is robust and is not easily turned off by strategic (Experiment 1) or contextual cues (Experiment 2). This account can naturally capture the presence of facilitative repetition effects for words and inhibitory effects for nonwords reported in previous studies (Balota & Spieler, 1999; see also Perea et al., 2016, for fits with the diffusion model). However, there are reports of a facilitative effect of repetition effect for nonwords in lexical decision—typically with several repetitions in short blocks, which have been attributed to retrieval of episodic cues (see Logan, 1990, instance theory). The familiarity account cannot account for such facilitative effects for nonwords. Therefore, exploring the conditions that lead to such facilitation is critical to a better understanding of the dynamics of the lexical decision task. In this article, we had a basic question in mind: "what are the minimal strength/types of manipulations that can turn the inhibitory effect for nonwords into facilitation?" With that goal, we designed 4 two-block experiments that examined whether lexical decision responses to repeated nonwords could be modulated by episodic retrieval over and above familiarity. In summary, and regarding the inhibitory repetition effect for nonwords, the episodic cues in Experiments 1 and 2, as induced by the instructions or by items' features (low-frequency words and visually salient features), did not reduce the effect. However, the episodic cues in Experiment 3 (two repetitions of the initial block) produced a net null effect, which turned to facilitative when the blocks were much shorter (Experiment 4).

Thus, the present set of experiments—using standard lexical decision instructions—have revealed a composite pattern of repetition effects for nonwords in lexical decision that go from inhibitory to facilitative. As Zeelenberg et al. (2004) anticipated, there are two opposite factors at play: one is familiarity and the other is episodic retrieval, with the prominence of the former being dependent on the length of the blocks and the number of encountered stimuli (see Balota & Spieler, 1999, for a similar claim). On one hand, the familiarity account can readily capture the pattern of findings when there is only one previous presentation and the blocks are long: we found an inhibitory repetition effect for nonwords for informed/uninformed participants, for items sharing (or not) the background colour, and when the set of words is composed of frequent/infrequent stimuli. However, the familiarity account cannot accommodate the null effect with two presentations and long blocks (Experiment 3) or the facilitative repetition effect for nonwords with two presentations and short blocks (Experiment 4). On the other hand, the episodic account can easily accommodate the findings in Experiment 4. The two short blocks that preceded Block 2 in Experiment 4 could have induced a different response strategy: one in which familiarity information was discounted or one in which an episodic trace of the nonword was retrieved (e.g., this mode of presentation could have made the stimulus–response associations more salient than in longer blocks).[3] Episodic retrieval was also likely at work when nonwords were repeated twice in Block 1 in long blocks (Experiment 3), cancelling out the influence of familiarity and yielding a net null effect. Along this line, an interesting avenue for further research is to examine whether it is possible to obtain facilitative repetition effects for nonwords with an increased number of repetitions of Block 1 (e.g., four repetitions) while varying the length of the blocks. This would offer valuable information on the boundary conditions for the familiarity versus episodic accounts: whether it is the number of items in each block or it is rather the strength of the memory traces relative to the total number of episodic memories—this may be accompanied by a manipulation of the trial distance between repeated items.

To sum up, the present experiments have shown that evidence accumulation models of lexical decision (e.g., Ratcliff et al., 2004, diffusion model) cannot account for the effects of repetition using a unitary mechanism for all manipulations that increase the familiarity of items: the "strong" version of the +*familiarity* yields +*wordness* hypothesis is ruled out because—at least under some circumstances (e.g., several repetitions at short lags)—extralexical episodic information seems to be utilised to make responses. The remaining question is what kind of mechanism could explain the findings of Experiments 3 and 4. Two possibilities come to mind: (1) perhaps the simplest one is that the episodic trace is integrated into the drift rate, yielding more negative (more non-wordlike) evidence for repeated nonwords; the tokens of information being accumulated would hence include the orthographic, familiarity, and the episodic information that are all integrated into a drift rate. Another possibility (2) is that there is a horse race between the response elicited by the episodic information and the familiarity information. The experiments presented here do not allow us to adjudicate between these two mechanisms; however, and quite importantly, both reflect a mixture of episodic and familiarity-based sources of information.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: The research reported in this article was partially supported by Grants PSI2017-86210-P (Manuel Perea) and BES-2015-07414 (Ana Marcet) from the Spanish Ministry of Science, Innovation and Universities.

ORCID iD

Manuel Perea https://orcid.org/0000-0002-3291-1365

Supplementary material

The Supplementary Material is available at: qjep.sagepub.com

Notes

1. The diffusion model, per se, is agnostic about what is the nature of the evidence being accumulated; however, within the literature that has applied the diffusion model to the lexical decision task, the assumption has been that wordness drives the evidence accumulation, hence we refer to this implementation as "the diffusion model" from now on.
2. Keep in mind that repeating twice the items in short blocks makes the repeated items highly salient in memory. Furthermore, the speeded instructions in the Zeelenberg, Wagenmakers, and Shiffrin (2004) experiment led to a much higher error rate (around/above 40%; 50% is chance) than in the typical lexical decision experiment.
3. As a reviewer pointed out, not only the number of trials differed with respect to Experiments 1 to 3, but also the time between the initial presentations and the test presentations (e.g., decay of episodic information cannot be excluded; see Ricker, Vergauwe, & Cowan, 2016, for review).

References

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 340–357. doi:10.1037/0096-1523.10.3.340

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition. *Journal of Experimental Psychology: General*, *128*, 32–55. doi:10.1037/0096-3445.128.1.32

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. doi:10.18637/jss.v067.i01

Bowers, J. S. (1994). Does implicit memory extend to legal and illegal nonwords? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 534–549. doi:10.1037/0278-7393.20.3.534

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256. doi:10.1037/0033-295x.108.1.204

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, *45*, 1246–1258. doi:10.3758/s13428-013-0326-1

Dulsky, S. G. (1935). The effect of a change of background on recall and relearning. *Journal of Experimental Psychology*, *18*, 725–740. doi:10.1037/h0058066

Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E.-J. (2009). A diffusion model decomposition of the practice effect. *Psychonomic Bulletin & Review*, *16*, 1026–1036. doi:10.3758/16.6.1026

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116–124. doi:10.3758/bf03195503

Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 21–31. doi:10.1037/0278-7393.2.1.21

Gomez, P. (2012). Mathematical models of the lexical decision task. In J. Adelman (Ed.), *Visual Word Recognition Volume 1: Models and methods, orthography and phonology* (pp. 70–89). Hove, UK: Psychology Press.

Gomez, P., & Perea, M. (2014). Decomposing encoding and decisional components in visual-word recognition: A diffusion model analysis. *The Quarterly Journal of Experimental Psychology*, *67*, 2455–2466. doi:10.1080/17470218.2014.937447

Gomez, P., Perea, M., & Ratcliff, R. (2013). A diffusion model account of masked vs. unmasked priming: Are they qualitatively different? *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 1731–1740. doi:10.1037/a0032333

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *10*, 3518–3565. doi:10.1037/0033-295x.103.3.518

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*, 627–633. doi:10.3758/brm.42.3.627

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26. doi:10.18637/jss.v082.i13

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, *22*, 1–35. doi:10.1016/0010-0285(90)90002-l

McKoon, G., & Ratcliff, R. (1979). Priming in episodic and semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *18*, 463–480. doi:10.1016/s0022-5371(79)90255-x

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, *113*, 327–357. doi:10.1037/0033-295x.113.2.327

Oganian, Y., Froehlich, E., Schlickeiser, U., Hofmann, M. J., Heekeren, H. R., & Jacobs, A. M. (2016). Slower perception followed by faster lexical decision in longer words: A diffusion model analysis. *Frontiers in Psychology*, *6*, Article 1958. doi:10.3389/fpsyg.2015.01958

Perea, M., Marcet, A., Vergara-Martínez, M., & Gomez, P. (2016). On the dissociation of word/nonword repetition effects in lexical decision: An evidence accumulation account. *Frontiers in Psychology*, *7*, Article 215. doi:10.3389/fpsyg.2016.00215

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, *111*, 159–182. doi:10.1037/0033-295X.111.1.159

Ricker, T. J., Vergauwe, E., & Cowan, N. (2016). Decay theory of immediate memory: From Brown (1958) to today (2014). *The Quarterly Journal of Experimental Psychology*, *69*, 1969–1995. doi:10.1080/17470218.2014.914546

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, *58*, 140–159. doi:10.1016/j.jml.2007.04.006

Wagenmakers, E.-J., Steyvers, M., Raaijmakers, J. G. W., Shiffrin, R. M., van Rijn, H., & Zeelenberg, R. (2004). A model for evidence accumulation in the lexical decision task. *Cognitive Psychology*, *48*, 332–367. doi:10.1016/j.cogpsych.2003.08.001

Wagenmakers, E.-J., Zeelenberg, R., Steyvers, M., Shiffrin, R., & Raaijmakers, J. G. W. (2004). Nonword repetition in lexical decision: Support for two opposing processes. *The Quarterly Journal of Experimental Psychology*, *57*, 1191–1210. doi:10.1037/e537102012-324

Yap, M. J., Sibley, D. E., Balota, D. A., Ratcliff, R., & Rueckl, J. (2015). Responding to nonwords in the lexical decision task: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 597–613.

Zeelenberg, R., Pecher, D., De Kok, D., & Raaijmakers, J. G. W. (1998). Inhibition for nonword primes in lexical decision re-examined: The influence of instructions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1068–1079. doi:10.1037/0278-7393.24.4.1068

Zeelenberg, R., Wagenmakers, E.-J., & Shiffrin, R. M. (2004). Nonword repetition priming in lexical decision reverses as a function of study task and speed stress. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 270–277. doi:10.1037/0278-7393.30.1.270