

Decomposing encoding and decisional components in visual-word recognition: A diffusion model analysis

Pablo Gomez¹ and Manuel Perea²

¹Psychology Department, DePaul University, Chicago, IL, USA

²ERI-Lectura and Department of Methodology, Universitat de València, Valencia, Spain

In a diffusion model, performance as measured by latency and accuracy in two-choice tasks is decomposed into different parameters that can be linked to underlying cognitive processes. Although the diffusion model has been utilized to account for lexical decision data, the effects of stimulus manipulations in previous experiments originated from just one parameter: the quality of the evidence. Here we examined whether the diffusion model can be used to effectively decompose the underlying processes during visual-word recognition. We explore this issue in an experiment that features a lexical manipulation (word frequency) that we expected to affect mostly the quality of the evidence (the *drift rate* parameter), and a perceptual manipulation (stimulus orientation) that presumably affects the nondecisional time (the T_{er} parameter, time of encoding and response) more than it affects the drift rate. Results showed that although the manipulations do not affect only one parameter, word frequency and stimulus orientation had differential effects on the model's parameters. Thus, the diffusion model is a useful tool to decompose the effects of stimulus manipulations in visual-word recognition.

Keywords: Diffusion model; Lexical decision; Encoding; Bayes factor; Reaction time

The processes underlying visual-word identification involve a large set of components that range from the extraction of the relevant sensory information in the visual system to the access of semantic information. In word-identification tasks (e.g., lexical decision), cognitive psychologists collect latency and accuracy data, and through these dependent variables they make inferences about the different components in visual-word recognition. Although achieving the goal of decomposing the contributions of these components is not trivial, a number of models attempt to do so by accounting for accuracy and reaction time (RT) simultaneously (e.g., the diffusion model,

Ratcliff, 1978; the multiple readout model, Grainger & Jacobs, 1996; the Bayesian reader, Norris & Kinoshita, 2008; the leaky competing accumulator, Usher & McClelland, 2001; and the linear ballistic accumulator, Brown & Heathcote, 2008). In the present research, we utilize Ratcliff's (1978) diffusion model; this model has allowed researchers to advance in their understanding of the structure of various two-choice tasks in different domains.

A comprehensive introduction to the diffusion model account of the lexical decision task can be found elsewhere (Gómez, 2012; Ratcliff, Gomez, & McKoon, 2004; Wagenmakers, 2009, for a

Correspondence should be addressed to Pablo Gomez, Department of Psychology, 2219 North Kenmore Avenue, Chicago, IL 60614-3504, USA. E-mail: pgomez1@depaul.edu

We would like to thank Laura González and Ana García-Blanco for help in running the participants.

The research reported in this article has been partially supported by the Spanish Ministry of Economy and Competitiveness [grant number PSI2011-26924].

general review of the model); here, we only discuss the conceptual issues around modelling the lexical decision task with the diffusion model. An important feature of the diffusion model is that it is a process-account (as opposed to a data-account, like structural equation models). When applied to the lexical decision task, the main assumption is that there is noisy accumulation of evidence triggered by the target stimulus. This evidence is accumulated until either a “word” or a “nonword” response is made. Given that the diffusion model is a process-account, its parameters can be linked to underlying psychological processes: The *drift rate* parameter relates to the rate of information extraction, akin to the lexical activation of an item; the *a* parameter (i.e., boundary separation) relates to response caution, or how much evidence in favour of the word or nonword alternative is needed to make a response; the *z* parameter (i.e., starting point of the diffusion process) relates to the a priori response biases (i.e., the expectation by the participant of the likelihood of a word vs. a nonword stimulus being presented); and the T_{er} parameter (i.e., time of encoding and response) relates to the time taken by the encoding of the sensory information and the motor execution of the response.

Within the domain of visual-word recognition, the diffusion model has proved to be a highly useful tool at dealing simultaneously with latency and accuracy data from lexical decision experiments. Ratcliff, Gomez, and McKoon (2004) indicated that, in lexical decision, “the sources of information are combined to provide a single quantity, the degree of *wordness*, which maps into drift rate in the diffusion model” (p. 175). The effects of word frequency, type of nonword, and repetition in the Ratcliff et al. experiments were successfully accommodated by changes in drift rates. However, if all manipulations in the lexical decision task produce changes only in the drift rate, one could think that the whole exercise of diffusion model fitting does not provide researchers with much more information than the traditional analyses on the mean latencies. Nonetheless, further analyses by Donkin, Heathcote, Brown, and Andrews (2009) on the Ratcliff, Gomez, and

McKoon (2004) data revealed a small but systematic underestimation of the word-frequency effect in the .1 quantiles (i.e., the leading edge of the RT distributions) in the experiments with pseudowords as foils. Donkin et al. argued that a model in which T_{er} was allowed to vary across word-frequency conditions produced a more appropriate account of the Ratcliff, Gomez, and McKoon (2004) data. Indeed, if we consider the “encoding” part of T_{er} as “the time taken for the initial perceptual, encoding and lexical access processes” (Donkin et al., 2009, 2902–2903) then it is not surprising that “lexical access” is modulated by word frequency.

In this article we examine whether the diffusion model can be used to effectively decompose the different underlying processes during visual-word recognition. To do so, we manipulated two factors that have (presumably) a different locus—one lexical and the other perceptual. The first factor is word frequency (high vs. low). As indicated above, Ratcliff, Gomez, and McKoon (2004) found that drift rates alone accounted for this lexical effect, although Donkin et al. (2009) claimed that encoding time (T_{er}) can also be affected (although to a lesser degree) by word frequency. The second factor is stimulus orientation (0°, 45°, 90°). Previous research has demonstrated that there is a large reading cost in rotated words (see Koriat & Norman, 1984, 1985; see also Yu, Park, Gerold, & Legge, 2010). It has been suggested that stimulus orientation affects the perceptual encoding of the words so that participants mentally rotate the string to the canonical horizontal orientation, and then they would process the letter string as usual (see Whitney, 2002)—in particular, when the rotation does not fundamentally alter the process of visual-word recognition (e.g., a 45° rotation as opposed to a more extreme 90° rotation). In a diffusion model, this would mean that stimulus orientation would affect the encoding process (T_{er}) to a larger extent than the accumulation of evidence (drift rate). We also explored the behaviour of the *a* parameter as a function of the rotation manipulation. The reason is that prior research in psychophysics has revealed that, in the early stage of visual processing, orientation detection of stimuli occurs simultaneously to the

detection of the presence of the stimulus (e.g., see Lee, Koch, & Braun, 1997; Smith & Ratcliff, 2009; Thomas & Gille, 1979). Hence, participants could adjust the decision criteria based on the string rotation as soon as they detect that the letter string is presented and before the encoding of graphemic information begins. Note, however, that we have not allowed the a parameter to vary as a function of word frequency because a core assumption in the diffusion model is that participants cannot adjust the decision boundaries once the encoding of the relevant information has started.

In sum, we manipulated two factors in a lexical decision experiment: (a) word frequency, which is a factor that in the diffusion model should affect the drift rate parameter rather than T_{er} (Ratcliff, Gomez, & McKoon, 2004); and (b) stimulus orientation, which is a factor that presumably affects the *encoding process*—and possibly the *decision boundaries*—to a greater degree than the drift rates. If word frequency and stimulus orientation differentially affect the parameters of the diffusion model, this would indicate that the model is a powerful tool to test the loci of stimulus manipulations in visual-word recognition—and presumably in other cognitive domains.

Our analyses of behaviour of the diffusion model parameters and of the empirical data utilize the unified framework of Bayes factors instead of the traditional null hypothesis statistical testing (NHST). The shortcomings of NHST have been extensively discussed over the years and have been forcefully pointed out in recent years (e.g., Cohen, 1994; McCloskey & Ziliak, 2012; Wagenmakers, 2007). It is not the objective of the present article to proselytize in favour of a specific method of statistical inference; however, recent developments of toolbox in Bayesian inference are accessible to researchers, and hence we are taking advantage of such tools in this article. Specifically, we present the Bayes factors for three reasons: (a) Bayesian inference overcomes some basic shortcomings of NHST; (b) Bayes factors allow us to measure the support in favour of the null hypothesis, which is particularly relevant in our experiment as the lack of interactions between word frequency and orientation

is an important issue for our modelling efforts; and (c) Bayes factors allow us to have a unified framework to analyse both the empirical data (see [Appendix A](#)) and the diffusion model parameters.

EXPERIMENTAL STUDY

Method

Participants

Twenty-seven undergraduate/graduate students from the University of Valencia participated voluntarily in the experiment. They had normal or corrected-to-normal vision and were native speakers of Castilian Spanish.

Materials

We selected a set of 258 five-letter words in Spanish from the B-Pal database (Davis & Perea, 2005). Half of these words were of higher frequency [average: 77 occurrences per million; range 16–493; mean number of orthographic neighbours (Coltheart's N): 1.0, range: 0–2; mean token frequency of the initial syllable: 405], and the other half were of lower frequency (average: 4.9 occurrences per million; range 1–11; mean Coltheart's N : 0.8, range: 0–2; mean token frequency of the initial syllable: 396). A set of 258 orthographically legal pseudowords (mean Coltheart's N : 1.3, range: 0–2) was created for the purposes of the lexical decision task. The nonwords were created by replacing one or two random letters from a word not in the word set and checking that all the bigrams/trigrams were orthographically legal (see [Appendix B](#) for the materials). An analysis of our items based on the LD1NN algorithm (Keuleers & Brysbaert, 2011) indicates little-to-no-bias in our materials (odds for word = 1.179, $p = .22$). We created three counterbalancing lists so that each word/nonword was presented to all participants, but in one list it would be horizontal (0°), in the other list it would be diagonal (45°), and in the third list it would be vertical (90°).

Procedure

Subjects were tested individually in a silent room. Presentation of the stimuli and recording of data were controlled by DMDX software (Forster & Forster, 2003). On each trial, a fixation point (+) was presented for 500 ms in the centre of the computer screen. Then, the target stimulus was presented until the participant's response or until 2500 ms had elapsed. Target stimuli were presented in upper case, 24-pt Arial—note that both the fixation point and the target stimuli were presented on the centre of the screen so that the fixation point always corresponded to the midpoint of the stimulus regardless of its orientation. Participants were told that words and nonwords would be displayed on the computer monitor in front of them, and that they should press the “yes” (“sí”) button if the letter string was an existing Spanish word and press the “no” button if the letter string was not a word. Participants were instructed to respond as rapidly as possible while keeping a reasonable level of precision. The order of stimuli was randomized for each subject. Each subject received a total of 16 practice trials prior to the experimental phase. The session lasted approximately 25 min.

Results

Although the main focus of the experiment was to examine the fits of the diffusion model, we also conducted the statistical analyses on the mean RTs and error data—the null hypothesis significance testing and a Bayes factor analysis (Morey & Rouder, 2013; Rouder, Morey, Speckman, & Province, 2012) is presented in the supplementary materials—these analyses revealed additive effects of word frequency and stimulus orientation. The mean latencies for correct responses and accuracy are presented in Table 1.

To obtain the fits to the diffusion model, we employed the routines described by Ratcliff and Tuerlinckx (2002) to minimize the chi-square [$\sum (O - E)^2/E$] for the difference between the empirical data and the model's predictions (see Supplemental material for an example in R code). The *observed values* (O) are the empirical proportions

Table 1. Mean lexical decision times and accuracy for words and nonwords in the experiment

Words	Orientation		
	0°	45°	90°
Higher frequency	602 (0.99)	622 (0.99)	728 (0.97)
Lower frequency	664 (0.94)	710 (0.93)	808 (0.91)
Nonwords	751 (0.96)	779 (0.96)	938 (0.96)

Note: Accuracies in parentheses.

of correct responses that fall within a bin bounded by the 0, .1, .3, .5, .7, .9, and 1.0 quantiles, multiplied by the proportion of responses for that choice (e.g., if there is a .95 response proportion for the “word” alternative among say, higher frequency words presented in the nonrotated condition (HF0R), the proportions would be $.95 \times [.1, .2, .2, .2, .2, .1]$. The *expected values* (E) are the model-generated proportions of responses bounded by the empirical RTs at the 0, .1, .3, .5, .7, .9, and 1.0 quantiles. For example, the model could predict a .93 response proportion for the HF0R condition, then the proportion of responses that the model predicts would fall within the empirical RTs at the relevant quantiles would be, for example, $.93 \times [.084, .095, .187, .339, .224, .07]$. Note that given the high accuracy in the experiment, we fitted the error response proportion (and not the error RTs) in those conditions in which the number of error responses was fewer than five (we had 27 subjects and 9 types of items per subject, and from these 243 combinations, only in 22 cases did we have 5 or more error responses).

We conducted the fits on a subject-by-subject basis (e.g., see Perea & Gomez, 2012). Figure 1 displays the model fits across all conditions and all subjects. The latency fits are quite good; however, the model misses some of the accuracy data; specifically, the model tends to underestimate the accuracy, and these misses can be large (the model underestimates the accuracy by more than .3 for about 3% of the data points, by more than .2 for about 8% of the data points, and by more than .1 for 35% of the data points). A visual examination of the top left panel of Figure 1 shows that

DECOMPOSING COMPONENTS IN THE DM

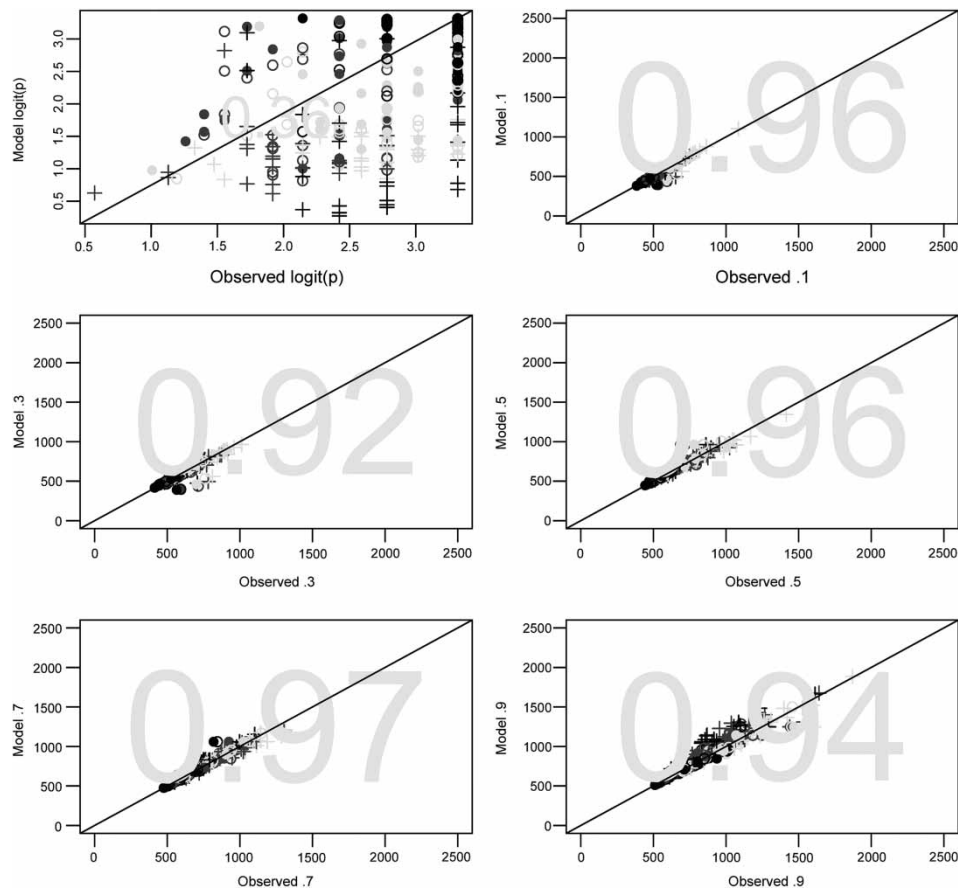


Figure 1. The figure shows scatterplots for the observed versus model-predicted values for logit of the proportion of correct responses and five quantiles in the reaction time (RT) distributions. The different symbols represent the different stimulus categories: The black points represent the higher frequency words, the dark grey points represent the lower frequency words, and the light grey points represent the nonwords. The solid circles represent the 0° rotation, the open circles represent the 45° rotation, and the plus signs the 90° rotation. The numbers printed as watermarks show the correlation between the model and the data.

the misses occur mostly in the higher frequency/90° condition (represented by plus signs in the figure).

We fitted the diffusion model to each subject's data; the values of T_{er} and drift rate were allowed to vary as a function of both orientation and word frequency/lexicality (i.e., there are nine possible values per each participant), and the values of the a parameter were allowed to vary only as a function of orientation. Note that the a parameter is not usually allowed to vary in diffusion model implementations because the model assumes that participants cannot adjust the decisional boundaries while processing the stimulus; however, because

the discrimination of orientation can occur simultaneously to detection of the presentation of the stimulus (Lee et al., 1997) we decided to allow this parameter to vary. All other parameters were allowed to vary across subjects but not across conditions within one participant. Table 2 presents the average parameter values, and Figure 1 displays the quality of the fits of the diffusion model.

After we obtained the point estimates for each parameter for each subject, we used Bayes factors to assess which of the hypotheses (models) that ascribe an effect or no-effect of conditions on diffusion model parameters should be preferred. The

Table 2. Average parameters of the diffusion model across subjects

Condition	a	T_{er} HF word	T_{er} LF word	T_{er} nonword	Drift HF word	Drift LF word	Drift nonword
0°	0.110	0.464	0.473	0.515	0.378	0.190	−0.238
45°	0.108	0.477	0.500	0.535	0.327	0.136	−0.216
90°	0.132	0.470	0.505	0.610	0.220	0.171	−0.168

Note: Other average parameter values: $z/a = 0.583$; $s_z = 0.001$; $\eta = 0.024$; $s_\eta = 0.142$. $\chi^2 = 277.4$. T_{er} = time of encoding and response; HF = higher frequency; LF = lower frequency.

logic of testing by Bayes factors is provided in Rouder, Speckman, Sun, Morey, and Iverson (2009; see also Rouder et al., 2012) and Morey and Rouder (2011, 2013). Bayes factors describe the probability of the data given a position (e.g., H_0) relative to the probability of the data given a different position (e.g., H_1 ; Rouder et al., 2012).

Following Rouder et al.'s (2012) notation, we refer to the Bayes factor as BF_{10} , where the subscripts refer to the models being compared, with the first and second subscripts referring to the model in the numerator (e.g., 1 = alternative) and denominator (0 = null), respectively; for example, a $BF_{10} = 100$ means that there is a 100:1 evidence in favour of the alternative hypothesis, and a $BF_{01} = 7.8$ means that there is a 7.8:1 evidence in favour of the null hypothesis. In this context, *evidence* means that the data are more probable given one model relative to the data given another model¹. In the following analyses we refer to the different models (hypotheses) as:

- *Model 0* is the null hypothesis model (i.e., there are no differences across conditions).
- *Model 1* is the full alternative model (i.e., there are main effects and interactions).
- *Model m* is the main effects only model (i.e., there are main effects but no interactions). When we are testing one-factor manipulations, and hence no interaction is possible, Model $m =$ Model 1.

For all the computations of Bayes factors below, we used Morey and Rouder's (2013) R package, in

which we employed the default settings for the specification of the alternative hypothesis.

Decision boundary (a)

We allowed the a parameter to vary as a function of one factor: orientation. According to the Bayes factor analysis, there is evidence for the alternative hypothesis ($BF_{10} = 3.8e + 06$); this means that the data given the alternative model (i.e., there is an effect of orientation on a) are almost 4 million times more probable than the data given the null model (i.e., there is no effect of orientation on a). In order to assess the effect size, Morey and Rouder's (2013) BF package provides the option of sampling the posterior distribution for each level of each factor. Panel A of Figure 2 shows the posterior distribution of values for the a parameter for the three orientations. In all the panels of Figure 2 the solid line represents the 0° rotation, the dotted line represents the 45° rotation, and the dashed line represents the 90° rotation. The orientation manipulation does not seem to produce a difference in the a parameter between the 0 and 45° conditions (means = 0.110 and 0.108), whereas the effect is quite robust for the 90° condition (mean = 0.132).

Time of encoding (T_{er})

We performed two separate Bayes factor analyses, one for responses to word stimuli, in which the T_{er} parameter was allowed to vary as a function of two factors (word frequency and orientation), and the other for responses to nonword stimuli, in which

¹Jeffreys (1961) provided a scale for the interpretation of Bayes factors; for example, a BF of 1:1 to 3:1, is "barely worth mentioning" while a BF of 100:1 is "decisive". We prefer to not use arbitrary cut-offs that could be construed as a critical value; however, we do use Jeffreys' wording as we present the outcomes of our analyses.

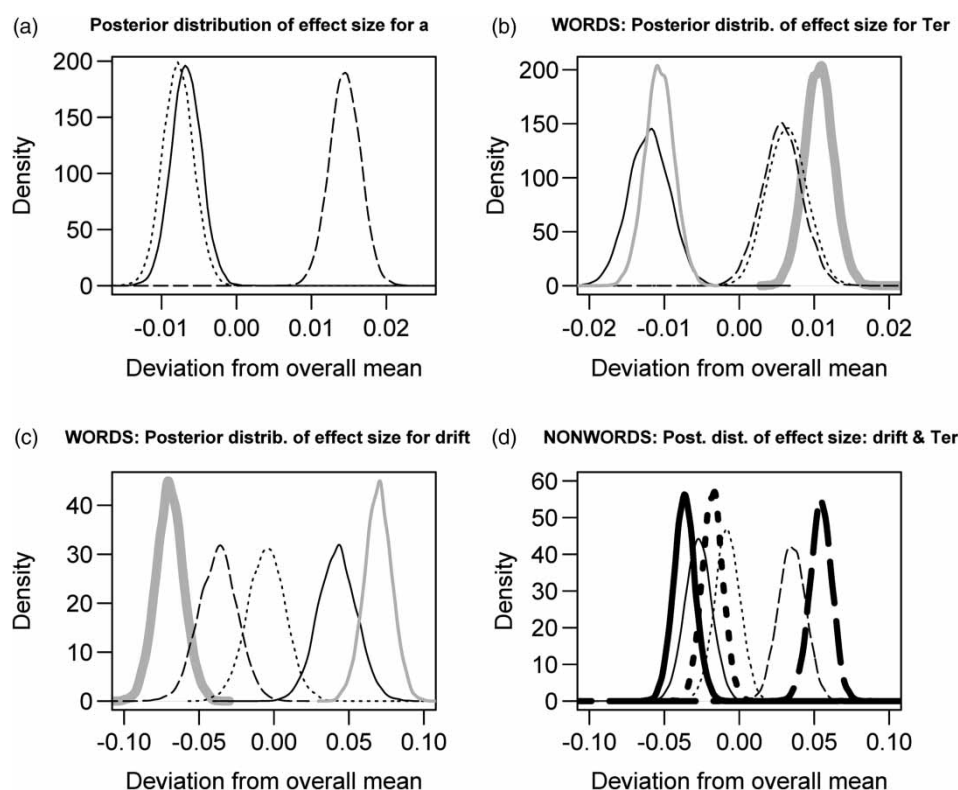


Figure 2. Panels A to C show the posterior distributions for the a parameter, the T_{er} parameter (time of encoding and response) for words, and the drift rates for words, respectively. The grey lines represent the effect of frequency (with the thin lines representing higher frequency words, and the thick lines representing lower frequency words). The type of lines represents the effect of orientation (solid lines for 0° , dotted lines for 45° , and dashed lines for 90°). Panel D shows the posterior distributions for both the drift rate (thin lines) and the T_{er} (thick lines) parameters for nonwords as a function of orientation: 0° solid lines, 45° dotted lines, and 90° dashed lines.

the T_{er} parameter was allowed to vary as a function of orientation. The Bayes factor analyses for word stimuli yielded a preference for the complete model (i.e., there were main effects of word frequency and orientation and an interaction on T_{er}): $BF_{10} = 2.3e + 07$. For the main effect of word frequency, the evidence is $2.2e + 05:1$; averaged across orientation, the T_{er} for higher frequency words is 0.468, and for lower frequency words it is 0.488. For the main effect of orientation, the evidence is 114:1; averaged across frequency, the T_{er} s for 0° , 45° , and 90° are 468, 488, and 488, respectively.

We also compared the full alternative Model 1 to the main effects only Model m; in this case, there is some ambiguity regarding the interaction as the data are only slightly more likely given the

complete model than given the main effects model ($BF_{1m} = 1.9$)—note that this does not imply that there is evidence in favour of Model m either. For the nonword stimuli, $BF_{10} = 1.9e + 08:1$, which indicates that orientation affects the T_{er} parameter for nonwords (0.515, 0.5135, and 0.610 for 0° , 45° , and 90° , respectively). We discuss the samples from the posterior distributions together with the drift rate parameters in a later section.

Drift rates

The analyses were parallel to those on the T_{er} parameter. For the drift rate parameter for word stimuli, there is decisive evidence for the full alternative model ($BF_{10} = 4.8e + 11$). The

evidence in favour of a main effect of orientation is not as decisive but still substantial (4.7:1). Averaged across word frequency, the drift rates for 0°, 45°, and 90° are 0.284, 0.231, and 0.196, respectively. The evidence in favour of a main effect of word frequency is quite large (about 1.3e + 9:1). Averaged across orientation, the drift rates for high- and lower frequency words are 0.308 and 0.166, respectively. Furthermore, there is strong evidence for the full model (i.e., with an interaction) over the main effects only model ($BF_{1m} = 15:1$).

The BF for nonword stimuli examined the effect of orientation on the drift rate parameter for nonwords. The odds in favour of the alternative model are 219:1. It is worth noting that there is an unusual, and most likely spurious, pattern in the drift rate parameters for lower frequency words in the 90° orientation, which is higher than the drift rate for the lower frequency words for the 45° orientation. This might be a consequence of the fact that the a parameter is free to vary, and it trades off with the drift rate.

Effect sizes on the parameter values

The goal of the current research was to decompose the effect of a perceptual manipulation (orientation) and a lexical manipulation (word frequency) using the diffusion model. A superficial assessment might lead to the conclusion that both manipulations affect the same parameters; however, the two manipulations have quite different effect sizes on the parameters. The orientation manipulation yields two distinct a values: a higher one for 90° rotations, and a lower one for 0° and 45° orientations; for the T_{er} parameter there are also two distinct values, but in this case both the 90° and the 45° orientation produce the higher value, while the 0° orientation yields the lower value; for the drift rates, on the other hand, each orientation produces its own value.

A comparison of the effects of word frequency versus orientation reveals that the effects on T_{er} are large for both variables (see Panel B of Figure 2), with word frequency and orientation yielding similar effect sizes. However, the effects on drift rate are qualitatively different from those for T_{er} : The effects of word frequency are larger

than the effects of orientation (see Panel C of Figure 2).

Summary

The effects of our experimental manipulations spread across all parameters of interest in the diffusion model: (a) Subjects adjust their *decision boundaries* for the 90° orientation relative to the items presented with the 0° and 45° orientations; (b) the *encoding time* is affected by both orientation and word frequency to a similar degree, although it is ambiguous as to whether these two factors interact or not; and (c) drift rates are affected by the two factors, and there is strong evidence in favour of an interaction; importantly, the effects on T_{er} are about the same size for orientation and word frequency; on the other hand, the effects on drift rate are much larger for word frequency than for orientation.

DISCUSSION

In previous studies, the diffusion model has been used to study processing differences (including decisional vs. nondecisional components), using the same set of stimuli, for different groups of participants (e.g., *ageing*: Ratcliff, Thapar, Gomez, & McKoon, 2004; *anxiety*: White, Ratcliff, Vasey, & McKoon, 2010), across different tasks (e.g., yes/no vs. go/no-go tasks; Gomez, Perea, & Ratcliff, 2013) and across different experimental settings (e.g., masked vs. unmasked priming; Gomez et al., 2013). The present experiment was designed to examine whether lexical decision data from two distinct manipulations, one perceptual (stimulus orientation) and the other lexical (word frequency), can be decomposed in different parameters (i.e., underlying processes) of the diffusion model. The classical inferential statistics analyses on the empirical data replicated earlier research: a reading cost due to rotation angle and a word-frequency effect (e.g., see Koriat & Norman, 1984, for a similar pattern of additivity). More important, analysing the diffusion model fits using Bayes factors provided us with the possibility of obtaining

samples from the posterior distributions over the effect sizes. These analyses confirmed that word frequency affected drift rate to a higher degree than orientation did, while stimulus orientation and word frequency affected the T_{er} parameter to a similar degree—that is, the present data support Donkin et al.'s (2009) suggestion that encoding time is modulated by word frequency. Notably, the change in T_{er} would shift every quantile, and indeed Donkin et al. also observed that Ratcliff, Gomez, and McKoon (2004) overestimations were not constrained to the .1 quantile; however, given the smaller variability in the .1 quantile they are easier to observe in that quantile than in the higher quantiles. Finally, it is important to note that although the inclusion of the rotated conditions might influence the process of the nonrotated condition, the qualitative pattern of effects in the RT distributions for the nonrotated condition is remarkably similar to the Ratcliff, Gomez, and McKoon (2004) data set.

Thus, the fits provide us with a coherent description of the dissociation between an encoding (nondecisional) component and an evidence accumulation (decisional) component in the lexical decision task. As we varied the orientation of the string from 0° to 45° the encoding process takes a longer time (about 30 ms longer), but the decision criteria do not seem to be affected. The 90° rotation, compared to the 45° rotation, however, barely affects the encoding time, but it makes participants take a more cautious approach. Rotations gradually affect the quality of the lexical/perceptual information, although this effect is not as large as the effect of word frequency.

Ideally, it would be best to have a coherent framework for the whole process, from the parameter estimation to the statistical inference to the model selection; however, this is not practical at this time. In this paper, we obtained the diffusion model's parameters for each participant in a frequentist way—this is the case of all current implementations of the diffusion model. Those parameters become the data analyzed with the Bayesian methods. The Bayesian aspect of our data analysis consists of placing beliefs (see Jackman, 2009, for an introduction) on the plausibility of models (e.g., H_0 , H_1) on

data (the model parameters) and updating our beliefs in light of the data.

This study has important implications for lexical processing research. We have presented an experimental situation in which the manipulations (namely, orientation and word frequency) simultaneously affect, and to different degrees, encoding (T_{er} parameter), lexical (drift rate), and response strategy (a parameter) components. Importantly, this dissociation is a further proof that the diffusion model can be used to determine the loci of stimulus manipulations in lexical decision—and presumably other two-choice tasks. Without a model, it would be impossible to determine the differential contribution (or lack thereof) of these factors on the empirical data. We hope that the present work will have practical implications for research. Data from a laboratory task in cognitive psychology is a function of a combination of factors. Researchers should try to understand what is the contribution of those factors, and one way to do so is by leveraging the capabilities of process models such as the diffusion model. In addition, although diffusion model accounts of lexical decision are not new, the subject-by-subject fits allow for group and individual differences (e.g., see Andrews & Lo, 2012) to be explored within the diffusion model framework.

Supplemental material

Supplemental material is available via the "Supplemental" tab on the article's online page (<http://dx.doi.org/10.1080/17470218.2014.937447>).

REFERENCES

- Andrews, S., & Lo, S. (2012). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 152–163. doi:10.1037/a0024953
- Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear

- ballistic accumulation. *Cognitive Psychology*, 57, 153–178. doi:10.1016/j.cogpsych.2007.12.002
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997–1003. doi:10.1037/0003-066X.49.12.997
- Davis, C. J., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods*, 37, 665–671. doi:10.3758/BF03192738
- Donkin, C., Heathcote, A., Brown, S., & Andrews, S. (2009). Non-decision time effects in the lexical decision task. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st annual conference of the cognitive science society*. Austin, TX: Cognitive Science Society. ISBN 978-0-9768318-5-3
- 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
- Gómez, 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, England: Psychology Press.
- Gomez, P., Perea, M., & Ratcliff, R. (2013). A diffusion model account of masked versus 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*, 103, 518–565. doi:10.1037/0033-295X.103.3.518
- Jackman, S. (2009). *Bayesian analysis for the social sciences* (Vol. 846). Chichester: Wiley.
- Jeffreys, H. (1961). *The theory of probability* (3rd ed.). Oxford, England: Oxford University Press.
- Keuleers, E., & Brysbaert, M. (2011). Detecting inherent bias in lexical decision experiments with the LD1NN algorithm. *The Mental Lexicon*, 6, 34–52. doi:10.1075/ml.6.1.02keu
- Koriat, A., & Norman, J. (1984). What is rotated in mental rotation? *Journal of Experimental Psychology: Learning Memory, and Cognition*, 10, 421–434. doi:10.1037/0278-7393.10.3.421
- Koriat, A., & Norman, J. (1985). Reading rotated words. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 490–508. doi:10.1037/0096-1523.11.4.490
- Lee, D. K., Koch, C., & Braun, J. (1997). Spatial vision thresholds in the near absence of attention. *Vision Research*, 37, 2409–2418. doi:10.1016/S0042-6989(97)00055-2
- McCloskey, D. N., & Ziliak, S. T. (2012). Statistical significance in the new Tom and the old Tom: A reply to Thomas Mayer. *Economics Journal Watch*, 9, 298–308.
- Morey, R. D., & Rouder, J. N. (2011). Bayes Factor approaches for testing interval null hypotheses. *Psychological Methods*, 16, 406–419. doi:10.1037/a0024377
- Morey, R. D., & Rouder, J. N. (2013). Bayes Factor: Computation of Bayes factors for simple designs (Version 0.9.4) [Computer program]. Retrieved from <http://cran.r-project.org/web/packages/BayesFactor/>
- Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology: General*, 137, 434–455. doi:10.1037/a0012799
- Perea, M., & Gomez, P. (2012). Increasing interletter spacing facilitates encoding of words. *Psychonomic Bulletin and Review*, 19, 332–338. doi:10.3758/s13423-011-0214-6
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108. doi:10.1037/0033-295X.85.2.59
- 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
- Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging*, 19, 278–289. doi:10.1037/0882-7974.19.2.278
- Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, 9, 438–481. doi:10.3758/BF03196302
- Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56, 356–374. doi:10.1016/j.jmp.2012.08.001
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t*-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16, 225–237. doi:10.3758/PBR.16.2.225
- Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal

- detection. *Psychological Review*, 116, 283–317. doi:10.1037/a0015156
- Thomas, J. P., & Gille, J. (1979). Bandwidths of orientation channels in human vision. *Journal of the Optical Society of America*, 69, 652–660. doi:10.1364/JOSA.69.000652
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108, 550–592. doi:10.1037/0033-295X.108.3.550
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, 14, 779–804.
- Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, 21, 641–671. doi:10.1080/09541440802205067
- White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Anxiety enhances threat processing without competition among multiple inputs: A diffusion model analysis. *Emotion*, 10, 662–677. doi:10.1037/a0019474
- Whitney, C. (2002). An explanation of the length effect for rotated words. *Cognitive Systems Research*, 3, 113–119. doi:10.1016/S1389-0417(01)00050-X
- Yu, D., Park, H., Gerold, D., & Legge, G. E. (2010). Comparing reading speed for horizontal and vertical English text. *Journal of Vision*, 10, 1–17. doi:10.1167/10.2.21

APPENDIX A

Empirical results

Error responses (4.4% and 4.0% for words and nonwords, respectively) and lexical decision times less than 250 or greater than 2000 ms (less than 1% and 2% for words and nonwords, respectively) were excluded from the latency analyses. The mean latencies for correct responses and accuracy are presented in Table 1.

Although the main focus of the experiment was to examine the fits of the diffusion model, we also conducted the standard statistical analyses for the mean RTs and error data. For word trials, analyses of variance (ANOVAs) based on the subjects' response latencies were conducted on the basis of a 2 (word frequency: high, low) \times 3 (rotation: 0°, 45°, 90°) design. For nonword trials, the ANOVAs were parallel except that only rotation was included as a factor in the analyses. For the word data, we present the traditional null hypothesis significance testing (NHST) analyses and the Bayes factors (see main text for an explanation of Bayes factors). A note about the Bayes factors: We present the Bayes factors for the effects of frequency and rotation for completeness, although for the word-frequency and rotation effects, the null hypotheses are implausible (see Cohen, 1994). For an explanation of the notation of Bayes factors, see the main text.

Word data

The ANOVA on the response times showed that responses to higher frequency words were, on average, 76 ms faster than the responses to lower frequency words, $F(1, 26) = 130.19$, $p < .001$, $BF_{10} = 5e + 32$. There was also a main effect of rotation, $F(2, 52) = 107.61$, $p < .001$, $BF_{10} = 6.25e + 31$.

Bonferroni pairwise comparisons showed that the latencies were 33 ms faster in the 0° condition than in the 45° condition, and 102 ms faster in the 45° condition than in the 90° condition (all $ps < .001$). The interaction between the two factors was not significant, $F(2, 52) = 2.33$, $p > .10$, $BF_{10} = 3.4$. The Bayes factor indicates that the null hypothesis (i.e., there is no interaction) is 3.4 times more likely than the alternative hypothesis (i.e., there is substantial, although not decisive, evidence against the interaction).

The ANOVA on the error rate data showed that participants made fewer errors to higher frequency words than to lower frequency words, $F(1, 26) = 48.79$, $p < .001$, $BF_{10} = 4.5e + 14$. The effect of rotation was also significant, $F(2, 52) = 10.52$, $p < .001$, $BF_{10} = 1.2e + 14$ —there were more errors in the 90° condition than in the 45° condition ($p = .007$) and in the 0° condition ($p = .004$), while there were no trends of a difference between the 0° condition and the 45° condition ($p > .34$). There were no signs of an interaction between the two factors ($F < 1$), and the Bayes factor ($BF_{10} = 8.1$) indicates that the null hypothesis is 8 times more likely than the alternative hypothesis.

Nonword data

The ANOVA on the latency data revealed an effect of orientation, $F(2, 52) = 64.57$, $p < .001$, $BF_{10} = 1.9e + 12$. Bonferroni pairwise comparisons showed that RTs were 28 ms faster in the 0° condition than in the 45° condition, and 159 ms faster in the 45° condition than in the 90° condition (all $ps < .001$).

The ANOVA on the error rate data did not reveal any signs of an effect of orientation: $F < 1$, and the Bayes factor supports the null hypothesis (i.e., there is no effect of orientation on error rates for nonwords), $BF_{01} = 5.8$.

APPENDIX B

Materials

Lower frequency words

CISNE INDIO IDOLO CUPON BOXEO FAENA GRIPE
 AYUNO FLUOR PULPO FIDEO LEGUA ANDEN
 MOLDE LIMON MULTA ELITE VERBO PUÑAL
 LONJA BOINA MOTEL OLIVO VIUDO RURAL
 PROSA TORAX HERIR AGUJA CROMO ROBLE
 MUSGO LINCE BELGA CALIZ JEQUE SESGO
 DANES CUTIS TIFUS FOSIL OVEJA ROSAL HABIL
 FOBIA TIGRE MIXTA VENUS DATIL DIVAN
 HIMNO BUZON ESQUI ABONO LEJIA ARCEN
 ASILO RIÑON ORUGA GRIFO PICOR LEPRAL TILDE
 SECTA PUDOR ROBOT CUTRE FREIR LAPIZ
 TITAN ABETO TIBIO MONJE OBESO REVES
 SOPLO VELOZ CRUDO LICOR ANCLA DOGMA
 TENAZ LUCIR OZONO FIBRA CACAO MIOPE
 FEMUR COFRE BUJIA REUMA ALBUM FUROR
 JERGA MUSLO TRIGO PEAJE LIMBO TECLA
 MATIZ SIDRA ALFIL NAIPE TUTOR ICONO
 ATAUD ERIZO PANEL CRUCE DOCIL OASIS
 TEBEO MOMIA FRESA TAROT SABLE REHEN
 VAGON TRUFA LOGRO FUGAZ OPERA SUIZO
 BLUSA MOVIL JAULA ARROZ CREMA GLOBO

Higher frequency words

NORTE NIVEL VIAJE ERROR AVION VIEJO CIFRA
 COMUN ETAPA ALDEA ORDEN NOCHE HOGAR
 MITAD ANUAL OTOÑO HIELO SIGNO RELOJ
 SUAVE SIGLO LETRA NOVIO LUNES CULPA
 FRASE PLUMA IGUAL DESEO LISTO CIVIL FINAL
 PRIMO HUMOR SEÑAL IDEAL JOVEN ABRIR
 PIEZA DEUDA BOLSA PARED VIRUS TORRE SERIE
 HABLA LIBRE GRAVE TESIS SUCIO NEGRO
 VERDE SABIO MAYOR AYUDA DIETA RIGOR
 FIRMA GOLPE SEÑOR ARABE HOTEL FIRME
 DULCE LEGAL BUENO AUTOR ACTOR CICLO
 CONDE POEMA VIEJA TARDE NARIZ DROGA
 RITMO HONOR CLIMA UNICA DOSIS UNION
 SALUD EXITO CAMPO BARBA PAUSA TRAJE
 ENERO EPOCA FELIZ APOYO MUJER LEJOS
 VUELO HEROE ABRIL FAVOR TEXTO SELVA
 BUQUE MIEDO MOTOR DEBIL PESCA DIGNO
 REINO CLASE REGLA AROMA PAPEL RIVAL
 PADRE TOTAL OREJA VAPOR GENIO SUDOR

CARNE VIDEO PIANO LIDER FACIL METAL
 NIEVE LLAVE BAILE GRUPO ARBOL LENTO

Nonwords

TUBER MEDAO CURNE RUGIL SUADE LILME
 CINSA DOTUS GOCLO VELVO TELIX PIROL
 LUASA SEIZO AUNAZ PILMO OERRA CLANE
 PISNO PIRED SULIT JETEL NOEVE GERSA RIEMO
 PAPLO SOINA BROJO MOGNO SERUS HUMES
 TIESE GELGA LISAC PUARA CURUD BREMO
 DUGIO ARUAL PURVO POMED BORIX VANJA
 HISEN GOITA MILVO TIBLA RUJEZ SEPAR TUNIL
 NASGA ALDER CINJE CLUYA BAVES NORCE
 DEPIR GELPE MARUT LUAGO FURJO LEMBO
 RAZCO OTRIO HOTIN SAICE TECHI CALUR
 MUALA CUMPO JARPE LENZA TAZIZ BOMPO
 CEGLA LUCCLA VITUS TULDO DOVAL BUNTU
 LOCRA GICEL RIMEN FAROZ PURJO CABAC
 ODIR VOSOR SOSIL RUEPO FEBLO TULVA ALGIS
 VENTU DUNSO GUSIA DELCE JARNE TESTU
 SILDO OEDRA DACTA MAORA TINDU NAUDA
 CAZIR ECTRA GAISO ROPIA CEUSA JONOA
 VOPAL JONSO VAIPA TULIN ESCAR CLESE
 GUNSO NORUA VIDO LUSLO ANGRO PRIFA
 DOELA GASAY FIDIS LINZO CLUPO FUSNA
 MURGE REMBO ALZON PUAPO BLEJO RESUA
 LASIN ERIEL NAUCE LAENA VOVIR LAUPA
 YERNE ROEGO PRUZA RUDIL SAGLO BUSCE
 AUGER TALID FLELE FOZON YURMA JUBLO
 MIERY PECHI CASNE BERNE VARBI BURSE
 VAVAZ CELVO LEUDE VINUS FLUVA HAULE
 MORVE UNGAR FURJA BAPTA GEION TAGLA
 TALZA NUORA MAZCA PUCRA ZOTAL CIURA
 FIEPO TOJER CAGLE FEGIZ PRILA CISON TAMB
 FISNO TURAL COMOL MEGIN RUSOR SETIN
 MIGIS CUNSA MUIRA MIOSA BIRDA SULVO FLITO
 TALGE KURDE DUGNA VUANA GEANO SOYON
 FIAGA TRILO DAVED LAERO BATUN MUZAR
 RASGE CRACE BREGE SIUTE MISLA FURSA
 DEPAR FUSTU ESNON BISIR NOCHI JISMO SEBLE
 MATIS CREJE HONJO MIDEZ GULIS HERJA SIRJE
 ZARLO VUBEZ NEMIO VUTIO VULON VORON
 DOSUS BUIBA GULAN FLICO ZATRE ECTAR
 TIRVA LUPOS DRAPA JALVE NAGMA SOGER
 SAMIR BOSEL RUNJA JAROL GAVEA SIMEL
 SOGNO HOSCE NEVIA JAYAS FLOMO TEMOT
 ENDAS