AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Human Perception and Performance

ISSN: 0096-1523

2023, Vol. 49, No. 6, 835–851
https://doi.org/10.1037/xhp0001113

Redefining the Decisional Components of Motor Responses: Evidence From Lexical and Object Decision Tasks

Michele Scaltritti[1], Francesca Giacomoni[1], Remo Job[1], and Simone Sulpizio[2, 3]
[1] Dipartimento di Psicologia e Scienze Cognitive, Università degli Studi di Trento
[2] Dipartimento di Psicologia, Università degli Studi di Milano-Bicocca
[3] Milan Center for Neuroscience (NeuroMI), Università degli Studi di Milano-Bicocca

Models of decision making focusing on two-alternative choices have classically described motor-response execution as a nondecisional stage that serially follows the termination of decision processes. Recent evidence, however, points toward a more continuous transition between decision and motor processes. We investigated this transition in two lexical decisions and one object decision task. By recording the electromyographic (EMG) signal associated with the muscle responsible for the manual responses (i.e., button press), we partitioned single-trial reaction times into premotor (the time elapsing from stimulus onset until the onset of the EMG burst) and motor times (the time elapsing from the onset of the EMG burst and the button press), with the latter measuring response execution. Responses were slower for pseudowords and pseudo-objects compared to words and real objects. Importantly, these effects were reliable even at the level of motor time measures. Differently, despite the reliable effect at the level of reaction times and premotor times, there was no difference in motor times between high- and low-frequency words. Although these results, in line with recent evidence, challenge a purely noncognitive characterization of motor-response execution, they further suggest that motor times may selectively capture specific decisional components, which we identify with late-occurring verification and/or control mechanisms.

Public Significance Statement
This study highlights specific decisional components that are still active during action. Even in the context of simple, fast, and discrete manual responses, it appears that part of the decision is still ongoing when we begin to move. Importantly, these motor-decisional components seem to reflect specific cognitive processes, possibly related to response monitoring and/or late verification processes that perform an additional check on difficult items for which we have no preexisting representations stored in long-term memory.

Keywords: decision making, lexical decision, motor-response execution, electromyography, lexicality

Supplemental materials: https://doi.org/10.1037/xhp0001113.supp

Most of our everyday activities stem from cognitive evaluations of the environment yielding decisions about how to act. The constant interaction between cognition and action is possibly what makes human behavior so flexible and adaptable. Yet, in characterizing the relationship between cognition and overt behavior, psychological models often maintain a temporal and functional priority of cognition over action. This *cognition-action thresholding* (Calderon et al., 2018) seems particularly clear within models of decision making focusing on two-alternative choice tasks, in which decisions are based on the sampling of evidence from the stimulus toward a specific response alternative, until reaching an action-triggering boundary.

Michele Scaltritti https://orcid.org/0000-0002-3328-1748
The authors have no conflict of interest to disclose.
This work was supported by a grant of the BIAL Foundation [79/20]. Part of the data were presented at the AMLaP conference, Paris, 2021, at the 62nd Annual meeting of the Psychonomic Society, New Orleans, 2021, and at the 22nd Conference of the European Society for Cognitive Psychology, Lille, 2022. We are grateful to Edoardo Sebastiano De Duro and Isabella Zeni for their help during data collection. Data and materials are publicly available at https://osf.io/6hqk5/.
Michele Scaltritti contributed toward conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, software, supervision, visualization, writing–original draft, writing–review and editing. Francesca Giacomoni contributed toward data curation, formal analysis, investigation, writing–review and editing. Remo Job contributed toward conceptualization, supervision, writing–review and editing. Simone Sulpizio contributed toward conceptualization, formal analysis, funding acquisition, methodology, project administration, supervision, visualization, writing–original draft, writing–review and editing.
Correspondence concerning this article should be addressed to Michele Scaltritti, Dipartimento di Psicologia e Scienze Cognitive, Università degli Studi di Trento, Corso Bettini 31, 38068 Rovereto (TN), Italy. Email: michele.scaltritti@unitn.it, or Simone Sulpizio, Dipartimento di Psicologia, Università degli Studi di Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italy. Email: simone.sulpizio@unimib.it

Although several instances of evidence-accumulation models (e.g., Brown & Heathcote, 2005, 2008; Donkin, Brown et al., 2009, 2011; Ratcliff et al., 2004, 2016; Smith & Vickers, 1988; Usher & McClelland, 2001; Van Zandt et al., 2000) differ significantly in terms of their structure, parameters, and accumulation functions, they share the assumption that motor-response execution is not part of the decisional process, but a separate, discrete stage that serially follows the termination of upstream decisional computations.

A number of empirical findings, however, have questioned this perspective. The analysis of continuous hand movements within choice reaching tasks (e.g., Song & Nakayama, 2009; Spivey & Dale, 2006) shows that the direction of reaching trajectories reflects the dynamic evolution of perceptual (e.g., Resulaj et al., 2009), attentional (e.g., Welsh & Elliott, 2004, 2005), linguistic (e.g., Farmer et al., 2007; Spivey et al., 2005), and decision processes (e.g., McKinstry et al., 2008; see also, e.g., Calderon et al., 2015; Chapman et al., 2010), suggesting that motor responses may be modulated in real time by the progressive unfolding of cognitive states.

Simpler and discrete responses such as the typical button presses may instead hide the cascaded flow of information from cognitive onto motor stages (e.g., Calderon et al., 2018; see also Weindel et al., 2021). Previous studies, in fact, offered rather inconsistent results (for a review, see e.g., Servant et al., 2021; see also Dutilh et al., 2019; Smith & Lilburn, 2020). However, more recent works point to a continuous stream of information that progressively maps stimulus evaluation onto the response channels even in the context of two-alternative choice tasks featuring discrete responses (button presses). In this context, researchers have exploited the electromyographic (EMG) signal to partition the reaction time (RT) into a premotor time (PMT), capturing the time from stimulus onset until the onset of the EMG activity, and a motor time (MT), reflecting the time from the onset of the EMG burst until the button press (Botwinick & Thompson, 1966). The latter provides a measure of response execution, thus enabling the assessment of cognitive/decisional variables at the motor stage. Importantly, as noted by other authors (e.g., Servant et al., 2015, 2021; Weindel et al., 2021), the excellent signal-to-noise ratio of the EMG signal allows to extract these measures at the level of single trials, thus providing precise chronometric indexes that are not blurred by the averaging procedures required by other physiological signals.

This recent evidence, gathered in the field of perceptual decision making, suggests that factors related to the rate of sensory evidence accumulation (e.g., the levels of contrast within Gabor patches or the levels of motion coherence in random dot motion tasks) consistently influence not just PMT, but also MT (Servant et al., 2021; Weindel et al., 2021), thus contradicting the functional characterization of motor-response execution in terms of a nondecision stage. On the contrary, decisions may actually still be unfolding during response execution, with the latter being affected by the same (Servant et al., 2021) or by some (e.g., Weindel et al., 2021) of the variables that shape PMTs.

The Present Study

Moving from the recent advancements reviewed above, we further investigated the assumption of the functional segregation between decisional and motor processes within tasks requiring a sampling of evidence from long-term memory. Specifically, we relied on lexical and object decision paradigms, two cases of two-alternative choice tasks in which participants have to classify each letter string/line configuration as a function of their lexical/object status (word vs. nonword/object vs. nonobject). While during perceptual decision making sensory-perceptual information may be directly mapped onto motor actions through dedicated sensorimotor pathways (e.g., Gordon et al., 2021; Cisek, 2007; Pezzulo & Cisek, 2016; Siegel et al., 2011), during conceptual decision making the link between perception and action is mediated by the activation of a representation stored in memory and by high-level processes underpinning the recognition and identification of complex and—at least for lexical decisions—symbolic stimuli. In this context, we assessed EMG traces associated with button presses, to ascertain whether differences in response latencies are captured solely by the premotor part of the RTs, or, differently, whether the difference is present at the level of MTs as well.

Abrams and Balota (1991) already showed lexicality and word-frequency effects on kinematic parameters of responses provided by moving a handle leftward or rightward (see also Bangert et al., 2012; Barca & Pezzulo, 2012; Moreno et al., 2011). As noted above, this sort of responses, by allowing longer and more continuous movements, also offers additional degrees of freedom for cross-talks and strategic adjustments between decisional and action processes, as advocated within proposals featuring an adaptive flow of information between cognition and action as a function of task- and response-related constraints (Calderon et al., 2018). We thus deem important not to assume an equivalence between these more complex, continuous responses and the more traditional experimental configuration featuring discrete button presses, particularly when considering that the latter is still the response modality used in most of the cognitive and neuroscientific work on decision making.

In Experiment 1, we used the lexical decision task to investigate the cognition-action thresholding, an assumption shared by both models of evidence accumulation and visual word recognition. In the evidence-accumulation framework, both word and nonword decisions are based on the sequential sampling of the same sources of evidence, and variations in the rate of evidence accumulation can account for differences among different types of stimuli (e.g., Ratcliff et al., 2004; for further discussion, Yap et al., 2015). Instead, in models of word recognition word versus nonword decisions are linked to the amount of global and local activation within the orthographic lexicon (e.g., Coltheart et al., 2001; Grainger & Jacobs, 1996), while nonword responses are delivered when a temporal deadline has elapsed and the threshold of lexical activation signaling a word response has not been reached. Importantly, by assuming thresholded decisional process, both frameworks predict that differences in RTs between words and nonwords should only affect the premotor component of response latencies. In the first experiment, we assessed this assumption, which, to anticipate the results, was clearly falsified by lexicality effects on MTs.

In the following experiments, we sought to provide a first general functional characterization of the information processed during motor-response execution. We thus experimentally investigated the generalizability of decisional modulations of motor responses in a different task, as well as across different effects. Specifically, Experiment 2 investigated an object decision task, to assess the generalizability of the effect beyond lexical decisions and linguistic stimuli, and Experiment 3 exploited once more the lexical decision paradigm and investigated the effect of lexical frequency. Other than being one of the major determinants of lexical decision performance across languages (e.g., Yap & Balota, 2009; see also Brysbaert et al., 2016; Ferrand et al., 2010; Keuleers et al., 2012; Yap et al., 2010), lexical frequency represents a particularly interesting test. In the context of the sequential sampling models, this variable is typically mapped onto the rate of evidence accumulation (e.g., Donkin et al., 2009; Gomez & Perea, 2014; Ratcliff et al., 2004; Yap et al., 2012; see also Heathcote & Love, 2012; Rae et al., 2014), with residual nondecisional components of effects solely attributed to perceptual encoding stages (Donkin et al., 2009; Gomez & Perea, 2014). Differently, if evidence accumulates even after response initiation and shapes the unfolding of the motor response (Servant et al., 2021), any robust effect stemming (at least in part) from the rate of evidence accumulation and traditionally detected at the level of RTs should also be sizeable on MTs. The possibility, however, is not trivial. Words and nonwords are considerably different, as only the formers have an existing representation stored within long-term memory systems. This decisional configuration is obviously different from perceptual decision-making tasks. Indeed, researchers have hypothesized the presence of additional processes in the case of nonword decisions (at least in the case of word-like nonwords, i.e., pseudowords) such as late-occurring verification stages (Paap et al., 1982; see Perea et al., 2005; Yap et al., 2015; Ziegler et al., 2001). One possibility is that, in the case of a decision based on information sampled from memory, these specific late-occurring stages take place, at least in part, during actual response execution. If this is the case, one would predict lexicality, but not frequency to affect measures of MT.

Other than calling for specific processing stages, pseudowords and pseudo-objects may require enhanced monitoring and control resources. For example, pseudowords are usually more prone to errors and, in particular, fast, impulsive ones possibly related to lexical capture phenomena (Scaltritti et al., 2021; see also Fernández-López et al., 2022) that might call for additional monitoring and/or control processes (e.g., Ridderinkhof, 2002; van den Wildenberg et al., 2010). Also, pseudowords are more prone to partial errors (Scaltritti et al., 2021)—which consist in a covert activation of the muscle associated with the incorrect choice, before the correct response is delivered, a phenomenon clearly pointing to online monitoring and correction mechanisms operating at the level of motor responses (e.g., Burle et al., 2002). Importantly, chronometric measures of MTs have been consistently associated with an online executive process related to error detection and correction (Allain et al., 2004; see also, Rochet et al., 2014; Smigasiewicz et al., 2020; Weindel et al., 2021). Also, recent proposals suggest that the propagation of evidence accumulation beyond the decisional threshold actually represents a second-order decision variable that is specifically tied to performance monitoring (Desender, Ridderinkhof, et al., 2021). We thus explored whether those manipulations highlighting effects on the MT components also triggered parallel modulations of the indexes of response accuracy associated with monitoring processes.

To summarize, the following experiments had two main aims. The first was to empirically assess a core assumption of decision making and visual word recognition models, according to which motor responses serially follow the termination of decisional processes. Both types of models predict that decisional effects should not percolate onto measures of motor-response duration. To anticipate, this prediction was falsified by our data. As a second aim, we thus sought to provide a functional characterization of the decisional components affecting the motor stage, focusing in particular on late-occurring verification and/or control mechanisms. This was achieved by comparing different experimental manipulations (lexicality and object status vs. words lexical frequency) and converging insights provided by measures of response accuracy (partial as well as fast-impulsive errors).

Experiment 1

Method

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. Data and materials for all the experiments are available at https://osf.io/6hqk5/ (Scaltritti, 2022). Scripts and codes for the analyses are available from the first author. For all the experiments, their design and analyses were not preregistered. All the software programs used to administer the experimental procedures, and for data collection, processing and analyses are reported in the corresponding sections within the reminder of the Method section. Data were collected in 2020–2021.

Participants

Sample size for all the experiments was decided on the basis of recent recommendations in the field (Brysbaert, 2019). We used previous data featuring similar tasks and measures (Scaltritti et al., 2020) and the R package *simR* (version 1.0.6; Green & MacLeod, 2016; Green et al., 2016) to run 200 simulations based on random samples (observed power). The results showed that, with an experiment featuring 28 participants and 100 items per experimental cell, we had a 70% chance (95% CI [63.14%–76.26%]) to detect a significant lexicality effect (alpha = .05) on chronometric measures of motor-response execution. As the experimental plan included new experimental paradigms (Experiment 2), as well as potentially null effects (Experiment 3), we decided to increase sample size (48), as well as the number of observations (128; 64 items repeated twice).

Forty-eight Italian native speakers took part in the experiment (33 females; $M_{\text{age}} = 22.98$; $SD_{\text{age}} = 3.58$). Data from four participants were replaced because of issues during the acquisition of the EMG signal (e.g., faulty electrodes, incorrect placement of the electrodes, detachment of the electrodes during the experiment). Data from three participants were discarded during the analysis due to an excessive number of EMG epochs rejected (see section EMG Recording and Processing).

All participants had normal or corrected-to-normal vision and reported no history of neurological problems or learning disabilities (these criteria were true across all the reported experiments). Using the Edinburgh Handedness Inventory (Oldfield, 1971) 43 participants could be classified as right handed ($M = 80.60$, $SD = 15.43$), whereas five were mixed right handed ($M = 39.02$, $SD = 10.27$).

Participation was compensated with €15. All the procedures received approval from the ethical committee of the University of Trento (protocol number 2020-028), and participants signed an informed consent document before the experiment (these conditions apply also to Experiments 2 and 3).

Stimuli

Sixty-four words were selected from the PhonItalia database version 1.10 (Goslin et al., 2014), and 64 pseudowords were created with the help of the Wuggy software (Keuleers & Brysbaert, 2010). Words and pseudowords were comparable across a series of psycholinguistic variables (see Table 1). Both words and

Table 1
Psycholinguistic Variables of the Stimuli Used in Experiment 1

Variables	Words	Pseudowords	*t*-value
Frequency (log)	3.27	—	—
N. of Letters	7.00	7.00	0
N. of Syllables	2.88	2.88	0
Orthographic *N*	4.16	4.41	0.27
OLD20	2.08	2.22	0.94

Note. N. of Letters = number of letters; N. of Syllables = number of syllables; Orthographic *N* = number of orthographic neighbors; OLD = orthographic Levenshtein distance (Yarkoni et al., 2008). For words, all variables were extracted from the PhonItalia database (Goslin et al., 2014). For pseudowords, the number of orthographic neighbors and OLD were computed on the same database using the *vwr* package (Keuleers, 2013) in R. *t*-values result from independent sample two-tailed *t*-tests conducted to compare words and pseudowords (all $ps > .34$).

pseudowords were partitioned into two subsets for counterbalancing purposes (illustrated below), and items were comparable both within and across subsets in terms of the same variables reported in Table 1.

Apparatus and Procedures

Participants first completed a questionnaire collecting demographic and health-related information. Then, after the installation of the electrodes for the recording of EMG, the experiment began. The experimental procedure and the acquisition of behavioral data were controlled via the E-Prime 2 software (Version 2.0.10.356, Psychology Software Tools) running on a laptop. Participants sat in front of the computer screen at a distance of about 60 cm, holding a joypad in their hands with their thumbs resting on the upper triggers. The joypad could be held either on the table, or resting on the participant's legs, as a function of individual preferences and signal quality (the configuration selected by the participant was typically associated with an increased comfort and a reduction in tonic EMG noise). They were instructed to classify letter strings as words or pseudowords using their thumbs to perform button presses. Speed and accuracy were equally emphasized.

The experiment was divided into four blocks, and the stimulus (word vs. pseudoword)-response (right vs. left hand) mapping was reversed in each following block, to ensure within each participant an equal number of left- and right-hand responses for each category of stimuli. The order of administration of the two stimulus-response hand mappings across blocks was counterbalanced across participants. Each subset of the items was assigned either to the first or the second block (the assignment was counterbalanced across participants). Within participants, the last two blocks were exact repetitions of the first two. Repetition was introduced to increase the number of observations, while keeping a constant number of trials across experiments and given the limitation in the number of items available for Experiment 2.[1] Before each block, participants performed eight practice trials to familiarize with the response mapping. Self-terminated breaks were prompted halfway within each block. The whole experimental session (including the installation of the electrodes and final debriefing) lasted about 90 min.

Stimuli were presented in 25-point Courier New font, in white against a black background. Trials started with a fixation cross (+) and its duration was chosen randomly among five alternatives (400, 450, 500, 550, and 600 ms). Then the stimulus appeared and remained on the screen until participant response or for a maximum of 1500 ms. A blank screen lasting 800 ms was finally presented and served as an inter-trial interval.

EMG Recording and Processing

EMG activity was acquired though an eego sports system (ANT Neuro), with a sampling rate of 1,000 Hz and using two pairs of bipolar electrodes placed about 1.5 cm apart on the thenar eminences of both hands. An additional ground electrode was placed on the pisiform bone of the right hand. The skin was prepared in advance using first isopropyl alcohol and then a mildly abrasive skin preparation gel (Nuprep, Weaver and Company). EMG signal acquisition was monitored online, and participants were asked to relax when tonic noise was detected.

Offline signal processing was performed using EEGLAB (version 14_1_2b; Delorme & Makeig, 2004) functions, as well as custom-made routines. A 5-Hz high-pass filter (order 2 Butterworth) and a 50-Hz notch filter were applied offline to the EMG traces. The signal was then segmented into epochs beginning 500 ms before stimulus onset and lasting until 2,100 ms afterward. Within each epoch, the onset of the EMG activity was detected using an algorithm devised following Liu and Liu (2016; see also Weindel et al., 2021). Specifically, the cumulative sum of the absolute values of the EMG trace is first computed and then subtracted from the straight line that joins the first and the last data points (which would correspond to the cumulative sum of a uniform distribution; see also Liu & Liu, 2016; Weindel et al., 2021). The EMG onset was marked in correspondence to the sample in which the difference reached its minimum value. Notably, the original algorithm by Liu and Liu (2016) was explicitly devised to overcome issues related to spontaneous spike activity within clinical populations, and thus provides a robust solution for EMG onset detection despite potential background noise.

To support artifact rejection, we applied a second algorithm, inspired by Servant et al. (2021). For each epoch, we computed windows of EMG activity by identifying samples in which activity exceeded the threshold of 3.5 *SD*s from the average value in the prestimulus baseline period (from −500 until0 ms). Consecutive windows separated by intervals shorter than 25 ms were merged. From the resulting windows of activity, we discarded those with a duration below 50 ms (arguably reflecting noise or random fluctuations, rather than purposeful EMG activity) as well as windows beginning after the epoch's RT. Epochs displaying more than one window of activity were marked. We then visually inspected all the epochs and retained only those in which the EMG onset was marked in correspondence to the last window of activity before response onset. This was done to discard onsets detected in

[1] All the chronometric analyses for the three experiments were conducted including also the fixed effect of repetition (repeated vs novel item). Importantly, Experiment 2 was the only one to reveal a significant interaction between repetition and the critical experimental manipulation (pseudo-object vs object) at the level of MTs, χ^2 (1) = 7.44, $p = .006$. The effect of object type was however significant for both novel items (*Estimate* = 16.93, *SE* = 2.67, $z = 6.33$), and—albeit reduced—for repeated ones (*Estimate* = 8.92, *SE* = 3.10, $z = 2.87$).

correspondence to noise bursts, drifts or a separate subthreshold EMG burst occurring before response onset (possibly related to hesitations). We also excluded all the epochs in which the onset detection algorithm failed, due to excessive noise or drift in the signal. On average, 8.76% of the epochs were rejected. Datasets ($N = 3$) in which more than 25% of the epochs were rejected were excluded from the analyses.

The two algorithms were finally applied also to the signal corresponding to the hand not involved in the final button press (e.g., the right-hand channel when a left-hand button press was delivered), to detect partial errors and partial correct responses (i.e., trials with an incorrect response, but with a prior subthreshold activation of the correct response hand). Epochs featuring one or more windows of activity were marked. Using visual inspection, epochs were then classified as containing true partial errors or partial correct responses when a visually detectable subthreshold EMG activation was present, and the timing of its onset was accurately detected. On average, partial errors occurred in 6.54% of the trials. Partial correct responses were very few (0.5%) and thus not investigated. Epochs containing partial errors or partial correct responses were dropped from chronometric and accuracy analyses.

All the processing steps reported in this section were consistently applied to the other reported experiments as well.

Measures

Chronometric Measures. Using the EMG traces, we partitioned each single RT into PMT—reflecting the time elapsing from stimulus onset until the onset of the EMG burst—and MT—capturing the time between the onset of the EMG burst and the actual button press. The analysis of these chronometric indexes focused on pure-correct responses (i.e., correct responses with no covert activation of the incorrect response hand).

Accuracy. These analyses focused on pure-correct and pure-error responses (i.e., correct responses and errors with no covert activation of the incorrect or correct response hand, respectively). We also considered conditional accuracy functions (CAFs), reflecting variations in accuracy as a function of response speed. Trials without responses (i.e., time-outs) were excluded from this analysis. Within each participant and within each stimulus category (e.g., words vs. pseudowords), trials were sorted into five quantiles as a function of their RT, with the first quantile capturing the fastest 20% of the responses, the second quantile the next 20%, and so on until the fifth quantile, reflecting the slowest 20% of the responses. The variable quantile was then treated as a fixed effect in the analyses.

Partial Errors. These analyses focused on correct responses and assessed potential variations in the likelihood of partial errors across conditions.

Statistical Analyses

Chronometric measures were analyzed using linear mixed-effects models. Response accuracy and partial errors were analyzed via generalized mixed-effects models due to the binomial nature of the dependent variables. All analyses were conducted using the *lme4* (version 1.1.27.1; Bates et al., 2015) and the *afex* packages (version 28.1; Singman et al., 2021) in R (version 4.2.1; R Core Team, 2015). Figures were made using the ggplot2 package (version 3.3.6; Wickham, 2016).

Fixed effects were assessed by comparing alternative models in which the effect under examination was either present or absent. Fixed terms were retained when likelihood ratio tests revealed that their exclusion would have determined a significant decrease in goodness of fit. In case interactions resulted significant, all the lower-order terms were retained. In this first stage, the random-effect structure was limited to by-participants and by-items random intercepts. Once we identified the significant fixed effects, we then tried to fit the structure of maximal complexity (Barr et al., 2013), including random slopes for all the fixed terms (as well as their correlations with the intercepts). When models failed to converge (due to overparameterization; e.g., Bates et al., 2018; Matuschek et al., 2017), we progressively simplified the random-effect structure by first removing correlations among random terms (i.e., fitting zero-correlation models), then by removing random slopes (or intercepts) associated with the smallest amount of variance (often corresponding to 0).

For the CAFs analyses, we focused on assessing the interaction between the variable quantile and the factor distinguishing between stimulus types (lexicality in Experiments 1 and 3, object type in Experiment 2, and word frequency in Experiment 3). If the interaction resulted significant, we further considered nonlinear relationships using second-order orthogonal polynomials to fit the quantile variables. These nonlinear terms were retained only when they increased goodness of fit. To obtain model convergence, for these analyses the random-effect structure was limited to by-participants and by-items random intercepts.

For all models, information regarding the fixed effects is reported in-text. Details about random effects for all the final models are listed in the online supplemental materials 1 (Tables S1 through S4). Information for all the parameters of CAF models is listed in the online supplemental materials 2 (Table S5 and Table S6). All the procedures outlined in this section were consistently applied to all the experiments.

Results

Chronometric Measures

Trials with errors (4.26% of the total), partial errors (6.54%), or an inaccurate detection of the EMG onset (6.68%) were excluded from the analyses. Results are summarized in Figure 1A.

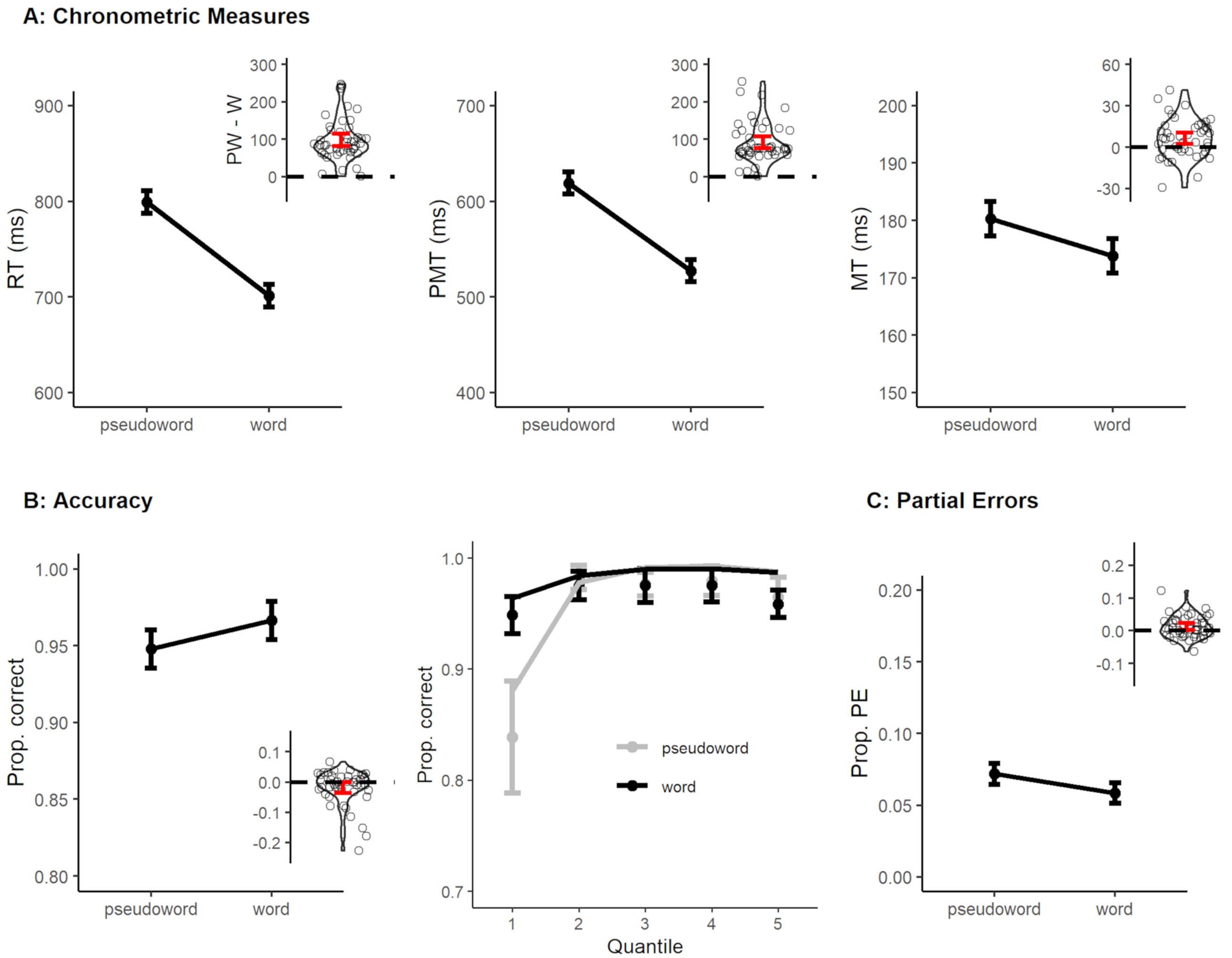
The lexicality effect was significant for measures of RTs, $\chi^2(1) = 71.83$, $p < .001$, PMTs, $\chi^2(1) = 66.03$, $p < .001$, and, crucially, MTs, $\chi^2(1) = 26.40$, $p < .001$. In the final models, we were able to retain the random-effect structure of maximal complexity for all the three measures (Table S1 in the online supplemental materials). Parameters of the fixed effects are listed in Table 3. Words were faster than pseudowords across all the three measures.

Accuracy

There was a significant effect of lexicality, $\chi^2(1) = 7.89$, $p < .001$. The final model (in which we had to drop the correlation between by-participants random intercepts and random slope for the lexicality effect; Table S1 in the online supplemental materials) revealed that responses were more accurate for words compared to pseudowords, $b = 0.55$, $SE = 0.24$, $z = 2.26$ (Figure 1B).

Analyses of CAF revealed a significant lexicality by quantile interaction, $\chi^2(1)$ 34.26, $p < .001$. Fitting the quantile variable with a

Figure 1
Results From Experiment 1



Note. Section A (first row): findings on measures of reaction time (RT), premotor time (PMT), and motor time (MT). Section B (second row, first two columns): findings on accuracy (first panel) and conditional accuracy functions (second panel). For the latter, points represent empirical means, lines represent means predicted by the statistical model. Section C reports findings on partial errors. Error bars reflect 95% confidence intervals. Inset plots provide information about the consistency of the lexicality effect across participants. Points represent individual difference scores between pseudowords and words in the corresponding measure, with the violin plot providing information about the distribution. Red error bars highlight 95% confidence interval of the mean effect for the whole sample. All confidence intervals were adjusted for within-participants variables following Morey (2008). See the online article for the color version of this figure.

quadratic orthogonal polynomial increased goodness of fit, $\chi^2(2) = 93.44$, $p < .001$. As visible in Figure 1B (see also Table S5 in the online supplemental materials), pseudowords were specifically more prone to fast errors (i.e., errors within the first quantile of the RTs distribution), compared to words.

Partial Errors

There was a significant effect of lexicality, $\chi^2(2) = 4.17$, $p = .04$. The final model, retaining the random-effect structure of maximal complexity (Table S1 in the online supplemental materials), highlighted that partial errors were less likely to occur for words compared to pseudowords, $b = -0.29$, $SE = 0.14$, $z = -2.09$ (Figure 1C).

Discussion

Experiment 1 revealed that the lexicality effect (slower responses for pseudowords compared to words) reliably affects both the premotor and, crucially, the motor component of RTs. This result is at odds with the notion that decision is terminated upon motor-response initiation. Had this been the case, there would be no reason to expect a lexicality effect on MTs.

The lexicality effect on MTs might reflect a continuation of the evidence-accumulation processes during response execution (e.g., Servant et al., 2021). An alternative hypothesis is that it would stem from different processes and sources of information. For example, late-occurring verification stages, which have been hypothesized for pseudoword stimuli (e.g., Perea et al., 2005), may still be ongoing after response initiation, thus prolonging MTs duration. Indeed, as for pseudowords there are no representations in long-term memory, decisions may take some extra time to verify that the string really fails to match any lexical entry. Part of this additional search may occur during response execution.

Measures of response accuracy revealed other insights. Pseudoword responses were more likely to yield partial errors as well as fast errors (i.e., errors occurring in the first quantile of the CAF). Both indexes can be linked with a tendency to misidentify pseudowords as words. Potentially, the system may react to this issue by increasing monitoring processes over these more uncertain responses, thus yielding longer MTs (e.g., Allain et al., 2004; Burle et al., 2002).

In the following experiment, we assessed whether these phenomena are exclusively related to lexical decision or can be reproduced with nonlinguistic stimuli.

Experiment 2

Method

Participants

Forty-eight Italian native speakers participated to the experiment (41 females; $M_{age} = 21.79$; $SD_{age} = 2.91$). Data from one participant were replaced as only a few epochs of the practice session were recorded. Data from three participants were discarded during the EMG processing procedure, due to the high number of EMG epochs rejected (>25%). Using the Edinburgh Handedness Questionnaire, 42 participants were classified as right handed ($M = 84.54$, $SD = 14.45$), five as mixed right handers ($M = 38.74$, $SD = 1.36$), and one as left handed (handedness score = −60).

Stimuli

Sixty-four images of manmade objects were selected from the Bank of Standardized Stimuli (BOSS; Brodeur et al., 2010, 2014). The selected pictures had moderately high values of name agreement (.71, $SD = .20$; mean H-value = 1.37, $SD = 0.96$), and depicted highly familiar objects (mean familiarity: 4.13, $SD = .30$; scale 1–5). The 64 images of pseudo-objects consisted in the items of the Novel Object and Unusual Name (NOUN) database (Horst & Hout, 2016). All images were converted to black-and-white images[2] and scaled to a 400 × 400 pixels size. The size of the files in kB, taken as a rough proxy for visual complexity (Székely & Bates, 2000) was comparable across objects and pseudo-objects ($M_{obj} = 69.13$, $SD_{obj} = 16.93$; $M_{pseudo} = 70.56$; $SD_{pseudo} = 12.97$; t [63] = −0.54, $p = .59$). Images for objects and pseudo-objects were partitioned into two subsets, for counterbalancing purposes. The two subsets were comparable for all the variables mentioned above.

Apparatus and Procedures

The same as in Experiment 1. The only differences were that (a) images were presented instead of letter strings, and (b) all the stimuli appeared on a white background.

Results

Chronometric Measures

Errors (8.47% of the total number of trials), partial errors (12.53%), or trials with an inaccurate detection of the EMG onset (4.79%) were excluded from the analyses. Partial correct responses were few (1.49%) and not further analyzed. Results are summarized in Figure 2A.

There were significant effects of object type (pseudo-objects vs. objects) on RTs, $\chi^2(1) = 31.27$, $p < .001$, PMTs, $\chi^2(1) = 22.51$, $p < .001$, and also on MTs, $\chi^2(1) = 48.83$, $p < .001$. The final models for all the three measures retained the random-effect structure of maximal complexity (Table S2 in the online supplemental materials). Parameters of the fixed effects are listed in Table 3.

Accuracy

The effect of object type was not significant, $\chi^2(1) = 3.47$, $p = .06$. CAF analyses revealed no object type by quantile interaction, $\chi^2(1) = 1.24$, $p = .26$ (Figure 2B).

Partial Errors

There was no significant effect of object type, $\chi^2(1) = 0.18$, $p = .67$ (Figure 2C).

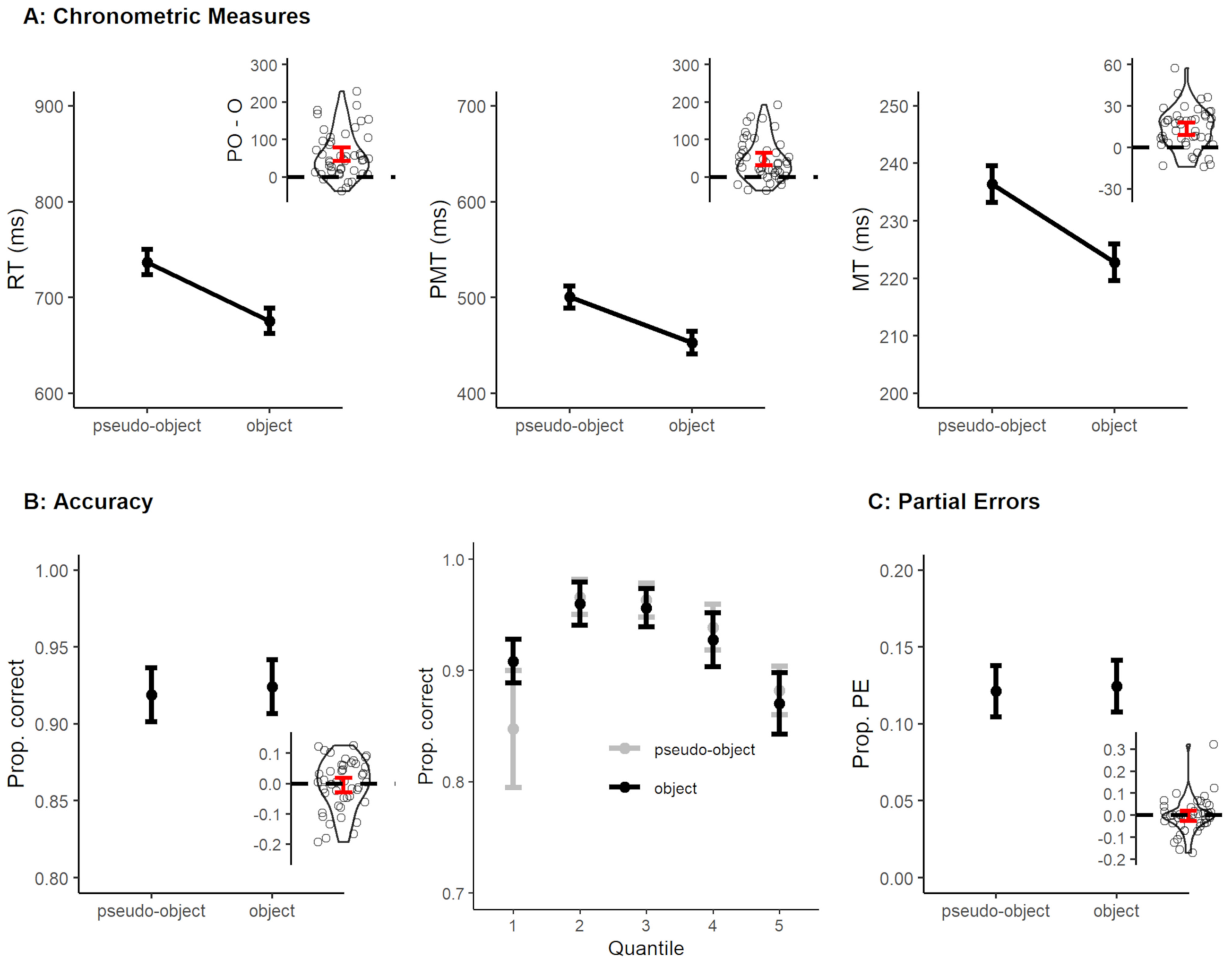
Discussion

Experiment 2 revealed a reliable effect of object type on MTs, which were longer for pseudo-objects compared to real ones. This finding testifies to the generalizability of decisional effects on MTs, beyond the context of lexical decision. However, compared to Experiment 1, measures of response accuracy–which were exploited to functionally characterize the processes occurring during motor-response execution–showed some intriguing differences. In particular, we found no evidence that partial errors were more likely to occur for pseudo-objects compared to objects, in contrast with what we found for pseudowords compared to words. Additionally, the analysis of CAFs failed to reveal a clear object type by quantile interaction. Although fast errors seem qualitatively more likely to occur for pseudo-objects, the lack of a significant interaction warrants against strong conclusions in this sense. Taken together, these results suggest that prolonged MTs may reflect a continuation of evidence-accumulation and/or late-occurring verification processes triggered by stimuli with no preexisting representation in long-term memory (i.e., pseudo-objects).

In the third experiment, we attempted to further specify the functional characterization of the decisional components of motor-response execution. Specifically, we manipulated word frequency within a lexical decision experiment. If the effects on MTs stem from a continuation of evidence accumulation during response execution, we would expect lexical frequency to affect MTs. Differently, the lack of a word-frequency effect paired with a

[2] We first ran a pilot with nine participants using colored stimuli. During the debriefing, participants reported that their decisions relied mainly on differences in color between objects and nonobjects. The pseudo-objects, in fact, had a rather distinctive color-palette compared to real-object pictures. The actual experiment was thus conducted on black-and-white versions of all the images.

Figure 2
Results From Experiment 2



Note. Section A (first row): findings on measures of reaction time (RT), premotor time (PMT), and motor time (MT). Section B (second row, first two columns): findings on accuracy (first panel) and conditional accuracy functions (second panel). Section C reports findings on partial errors. Error bars reflect 95% confidence intervals. Lines were not plotted when the effect under examination was not significant. Inset plots provide information about the consistency of the object-type effect across participants. Points represent individual difference scores between pseudo-objects and real objects in the corresponding measure, with the violin plot providing information about the distribution. Red error bars highlight 95% confidence interval of the mean effect for the whole sample. All confidence intervals were adjusted for within-participants variables following Morey (2008). See the online article for the color version of this figure.

replication of the lexicality effect would be more in line with a verification account, related to the specific features of pseudowords stimuli.

Experiment 3

Method

Participants

Forty-eight participants took part in the experiment (38 females; $M_{age} = 21.02$; $SD_{age} = 2.20$). Data from three participants were replaced because of problems during the acquisition of the signal. Data from the other three participants were excluded during the stage of EMG signal processing due to the high number of EMG epochs rejected. According to the Edinburgh Handedness Questionnaire, 43 participants could be classified as right handed ($M = 83.25$, $SD = 14.35$), three as mixed right handers ($M = 50$, $SD = 0$), one as a mixed left hander (handedness score $= -30$), and one as left handed (score: -88.9).

Stimuli

Sixty-four high-frequency and 64 low-frequency words were selected from the PhonItalia database version 1.10 (Goslin et al., 2014). One-hundred and twenty-eight pseudowords were created with the help of the Wuggy software (Keuleers & Brysbaert, 2010). High- and low-frequency words were comparable for several

Table 2
Psycholinguistic Variables of the Stimuli Used in Experiment 3

Variables	HF	LF	*t*-value	Words	Pseudo.	*t*-value
Frequency (log)	5.87	1.93	19.91	3.90	—	—
Freq. Subtlex (log)	9.11	4.99	19.04	7.05	—	—
Familiarity	7.26	6.31	5.37	6.78	—	—
Imageability	7.24	7.24	0.02	7.24	—	—
Concreteness	6.57	6.87	−1.07	6.72	—	—
Valence	5.57	5.22	1.11	5.40	—	—
Arousal	5.46	5.27	1.17	5.37	—	—
N. of Letters	6.89	6.95	−0.21	6.92	6.92	0.00
N. of Syllables	2.89	2.97	−0.62	2.93	2.95	0.18
Orthographic N	3.00	3.19	−0.29	3.09	3.09	−0.02
OLD20	2.02	2.14	−1.15	2.08	2.14	0.78
Bigr. Freq. Sum	674,018	705,947	−0.58	689,983	665,406	−0.64
Bigr. Freq. Mean	111,813	115,835	−0.67	113,824	108,797	−1.19

Note. HF, high frequency; LF, low frequency; N. of Letters = number of letters; N. of Syllables = number of syllables; Orthographic *N* = number of orthographic neighbors; OLD = orthographic Levenshtein distance (Yarkoni et al., 2008); Bigr. Freq. Sum = summed bigram frequency; Bigr. Freq. Mean = mean bigram frequency. For words, all the surface variables were extracted from the PhonItalia database (Goslin et al., 2014), except for Frequency Subtlex (log), which was extracted from the SUBTLEX-IT database (Crepaldi et al., 2013). Semantic scores (familiarity, concreteness, imageability, valence, arousal) scores were taken from the Italian adaptation (Montefinese et al., 2014) of the Affective Norms for English Words database (Bradley & Lang, 1999). For pseudowords, the number of orthographic neighbors, and OLD were computed on the PhonItalia database using the *vwr* package (Keuleers, 2013) in R. Bigram frequency variables were computed on the same database with a custom-made script. *t*-values result from independent sample two-tailed *t*-tests.

psycholinguistic variables (Table 2). The same was true when comparing words (high- and low-frequency taken together) and pseudowords. High- and low-frequency words were partitioned into two subsets for counterbalancing purposes. The subsets were comparable in terms of the variables listed in Table 2. Pseudowords were similarly partitioned into two subsets, which were comparable with those created for words across the variables reported in Table 2.

Apparatus and Procedures

Apparatus and procedures were the same as in Experiment 1.

Results

Chronometric Measures

Errors (5.29% of the total number of trials), partial errors (13.45%), or trials with an inaccurate detection of the EMG onset (4.16%) were excluded from the analyses. Partial correct responses were very few (0.67%) and thus not further considered. The results are summarized in Figure 3A and Figure 3D. Parameters of the final models are listed in Table 3.

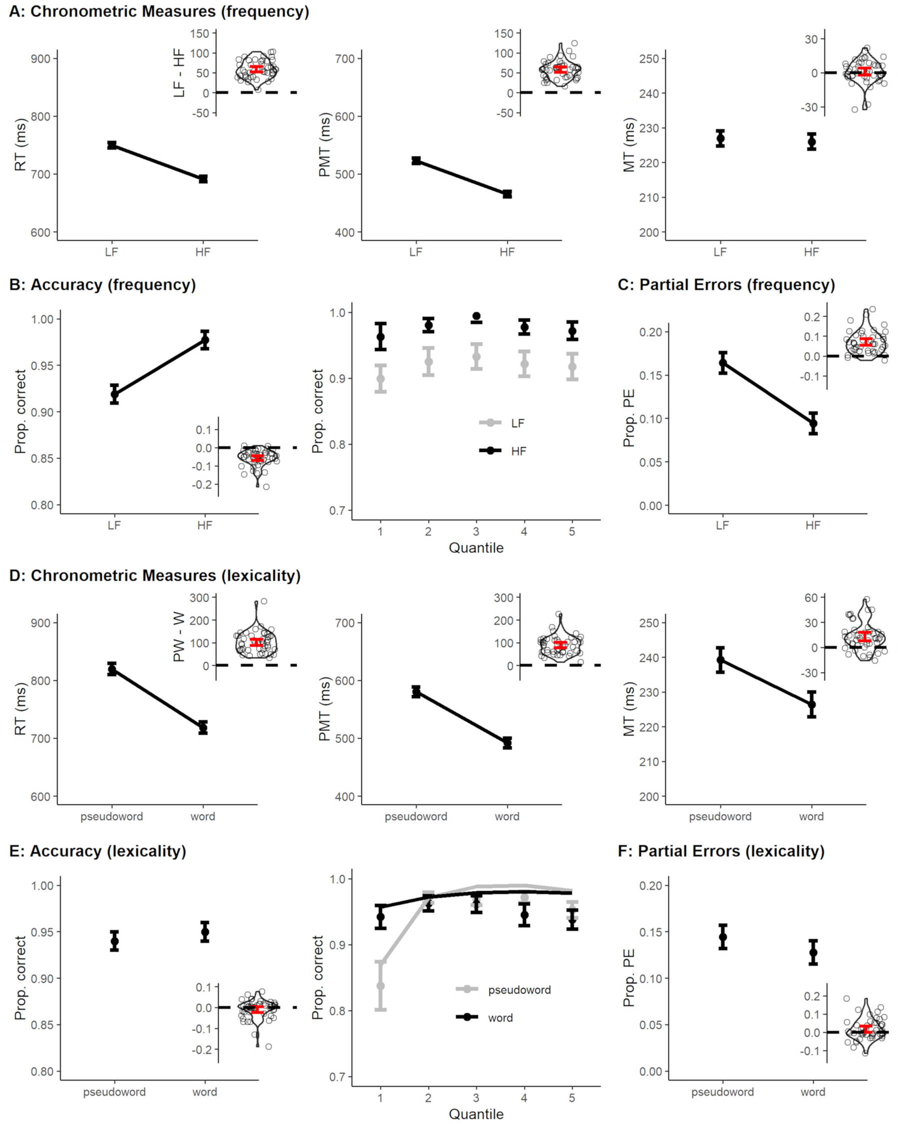
Word Frequency. There were significant frequency effects on RTs, $\chi^2(1) = 47.65$, $p < .001$, and on PMTs, $\chi^2(1) = 48.14$, $p < .001$. The final models, retaining the random effects structure of maximal complexity (Table S3 in the online supplemental materials), revealed that both measures were significantly longer for low- compared to high-frequency words. Differently, there was no frequency effect on MTs, $\chi^2(1) = 0.27$, $p = .60$ ($b = 0.71$, $SE = 1.45$, $t = 0.49$). Given the theoretical relevance of this null effect, we estimated the corresponding Bayes Factor (BF). Specifically, we subtracted the Bayesian Information Criterion (BIC) of the model featuring the fixed effect of word frequency from the one taken from the null model (only random intercepts), thus obtaining the *delta BIC*. The BF was then computed following the formula exp(deltaBIC/2) (Raftery, 1995;

Table 3
Parameters of the Fixed Effects for Models of Chronometric Measures

	RT			PMT			MT		
Fixed effects	Est.	*SE*	*t*	Est.	*SE*	*t*	Est.	*SE*	*t*
Exp 1									
Intercept	802.77	18.97	42.32	622.34	19.52	31.89	180.32	6.18	29.19
Lexicality (word)	−99.67	13.12	−7.59	−92.96	12.83	−7.24	−6.49	2.11	−3.08
Exp 2									
Intercept	742.34	14.23	52.17	506.11	13.59	37.25	236.21	236.21	28.36
Obj. Type (real object)	−59.03	13.67	−4.32	−45.74	12.59	−3.63	−13.34	2.37	−5.63
Exp 3									
Intercept	691.10	14.99	46.10	465.17	12.89	36.09	226.37	6.53	34.66
Frequency (low frequency)	64.10	8.67	7.39	63.09	8.56	7.37	—	—	—
Intercept	822.77	17.62	46.69	583.51	15.63	37.34	239.22	7.17	33.36
Lexicality (word)	−99.89	9.84	−10.15	−86.98	9.13	−9.53	−12.88	2.47	−5.22

Note. RT = reaction time; PMT = premotor time; MT = motor time; *SE* = standard error. When the fixed term failed to increase the goodness of fit, it was excluded from the model and thus parameters are not reported (—).

Figure 3
Results From Experiment 3



Note. Sections A and D (first and third rows): findings on measures of reaction time (RT), premotor time (PMT), and motor time (MT). Sections B and E (second and fourth rows, first two columns): findings on accuracy (first panel) and conditional accuracy functions (second panel). For the latter, points represent empirical means, lines represent means predicted by the statistical model. Sections C and F (last panels in the second and fourth rows) report findings on partial errors. Error bars reflect 95% confidence intervals. Lines were not plotted when the effect under examination was not significant. Inset plots provide information about the consistency of the effects across participants. Points represent individual difference scores between low- and high-frequency words (LF and HF, respectively) or between pseudowords and words in the corresponding measure, with the violin plot providing information about the distribution. Red error bars highlight 95% confidence interval of the mean effect for the whole sample. All confidence intervals were adjusted for within-participants variables following Morey (2008). See the online article for the color version of this figure.

Wagenmakers, 2007). We obtained a BF = 0.042, suggesting that the data provide strong evidence for the null hypothesis.

Lexicality. There were significant lexicality effects on RTs, $\chi^2(1) = 141.4$, $p < .001$, PMTs, $\chi^2(1) = 120.87$, $p < .001$, and MTs, $\chi^2(1) = 111.47$, $p < .001$. All the chronometric measures were significantly slower for pseudowords compared to words (Table 3; for the random-effect structure, see Table S4 in the online supplemental materials).

Accuracy

Word Frequency. The frequency effect was significant, $\chi^2(1) = 39.49$, $p < .001$. The final model, featuring no correlations between random slopes and intercepts (Table S3 in the online supplemental materials), showed that response accuracy was lower for low-compared to high-frequency words, $b = -1.29$, $SE = 0.2$, $z = -6.45$. Analyses of CAFs revealed no significant interaction between Word frequency and quantiles, $\chi^2(1) = 0.07$, $p = .78$ (Figure 3B).

Lexicality. There was a significant effect of lexicality, $\chi^2(1) = 4.51$, $p = .03$. However, once the random slopes were included in the final model (Table S4 in the online supplemental materials), the effect was no longer significant, $b = 0.19$, $SE = 0.19$, $z = 0.98$ (Figure 3E). Analyses of CAFs revealed a significant interaction between lexicality and quantiles, $\chi^2(1) = 81.11$, $p < .001$, driven mostly by fast and impulsive errors for pseudowords (Figure 3E; see also Table S6 in the online supplemental materials).

Partial Errors

Word Frequency. The effect of word frequency was significant, $\chi^2(1) = 53.76$, $p < .001$, with the final model indicating a higher likelihood of partial errors for low- compared to high-frequency words, $b = 0.73$, $SE = 0.1$, $z = 7.59$ (Figure 3C).

Lexicality. There was a significant lexicality effect, $\chi^2(1) = 4.42$, $p = .03$. The reduction in the likelihood of partial errors for words however appeared rather weak when including the random slopes in the final model (Table S4 in the online supplemental materials), $b = -0.16$, $SE = 0.09$, $z = -1.67$ (Figure 3F).

Discussion

The third experiment replicated the lexicality effect on MTs found in Experiment 1. Differently, the effect of lexical frequency remained exclusively bounded to the premotor component of RTs. One trivial possibility is that the lack of a frequency effect on MTs simply reflects a power issue or a scaling effect. However, BF approximation (Wagenmakers, 2007; see also Raftery, 1995) suggests that Experiment 3 provides strong evidence favoring the null hypothesis, that is, that there is no frequency effect on MTs. Moreover, albeit smaller than the lexicality effect, the size of the word-frequency effect on RTs and PMTs is fully comparable to the object type effect reported in Experiment 2 (see Table 3, as well as Figure 2A and Figure 3A). It thus seems that our experiment should have been able to detect a frequency effect on MTs, had there been one.

As we will discuss in more detail below, this dissociation across different experimental manipulations in their ability to reach the motor stage may provide an important constraint with respect to the functional characterization of the decisional components that are active during motor-response execution. In fact, at least in the context of the decision paradigms we have implemented, it seems that the critical factors in determining the effects on the motor component of RTs are either related to the lack of a stimulus representation in long-term memory (pseudowords and pseudo-objects vs. words and objects) and/or to the request of additional control processes for these kinds of stimuli.

General Discussion

We experimentally investigated the boundaries between decision and action within conceptual two-alternative choice tasks featuring discrete button-press responses. Using the EMG signals, RTs were fractioned into a premotor and a motor component (MT, Botwinick & Thompson, 1966) to assess whether decision processes terminate before response initiation—as assumed by prominent models of binary decision making and lexical decision—or, instead, whether they are still at play during motor-response execution. Our results support the latter perspective and reveal important constraints that may further clarify the transition from decision onto action-related processes, at least when evidence is sampled from memory, rather than from sensory input.

Experimental Factors Affecting Versus Not-Affecting MTs Constrain the Functional Interpretation of Motor Responses' Decisional Components

Experiments 1 and 3 revealed that, in lexical decision tasks, the classic lexicality effect can be tracked also during motor-response execution. Experiment 2 additionally revealed a similar effect in an object decision task, suggesting that the phenomenon may not be due to unidentified task-specificities of lexical decision. At first sight, these results seem to fit nicely with the recent ones in the field of perceptual decision making, where the modulation of MTs as a function of the available sensory evidence has been interpreted as a signature of evidence accumulation continuing during motor-response execution (Servant et al., 2021). However, the results on word frequency (Experiment 3) challenge this interpretation, at least in the context of the experimental paradigms we exploited. In fact, albeit RTs were reliably slower for low- compared to high-frequency words, this effect remained bounded within the premotor component of RTs (see also online supplemental materials 3). This does not appear to be compatible with the notion that the unfolding of motor response is shaped by an ongoing evidence-accumulation process relying on the same sources of information that are used during purely cognitive decisional stages (Servant et al., 2021; see also Servant et al., 2015, 2016). In other words, as lexical frequency consistently modulates the rate of evidence accumulation across different models (e.g., Dufau et al., 2012; Heathcote & Love, 2012; Rae et al., 2014; Ratcliff et al., 2004), if we assume that evidence accumulation continues after response onset (Servant et al., 2021), why no frequency effect is detected on MTs?

In the context of perceptual decision making and, in particular, when considering the effects of stimulus-response compatibility, different previous experiments have shown manipulations that selectively affect PMT, while leaving MTs unaffected (e.g., Burle et al., 2002; Hasbroucq et al., 1999; Spieser et al., 2017; but see Servant et al., 2021 for a different perspective). Weindel et al. (2021) have recently reported a number of findings pointing toward the independence of PMTs and MTs. For example, whereas the manipulations of stimulus contrast and speed-accuracy tradeoff modulated the two measures in the same direction, response accuracy revealed opposite-going influences, with errors displaying longer PMTs and shorter MTs compared to correct responses. Further, response force and response side (at least in one experiment) selectively affected MTs. According to the authors, these dissociations support the notion that PMTs and MTs reflect different latent (cognitive) processes.

Similarly, in our experiments, the difference in propagation between frequency and lexicality effect across the motor stages might support the notion that specific processes modulate motor-response execution in lexical decision. In other words, these empirical observations offer some important constraints with respect to the functional characterization of the decisional components observed at the level of motor-response execution. Specifically, the mismatch between the word-frequency effect on the one hand, and the lexicality and the object type effect on the other hand, suggests that decision processes may unfold in different ways as a function of the nature of the stimuli. Slower MTs were selectively found for items with no preexisting representation in long-term memory stores (i.e., pseudowords and pseudo-objects). Differently, when a stored representation was available, albeit less accessible as in the case of low-frequency words, the slowdown of response latencies remained confined to the PMT. This observation paves the way for different functional interpretations of the observed phenomena, as detailed below.

Responses to Nonwords

The differentiation between responses for items that are present versus absent in long-term memory resonates with a critical under-investigated question for any account of lexical decision, concerning what may constitute evidence for a nonword response. The issue has been directly tackled by Dufau et al. (2012), who proposed a leaky competing accumulator (LCA; see Usher & McClelland, 2001) model featuring separate and mutually inhibiting word and nonword nodes. Whereas word responses are driven by lexical evidence, the input of the nonword node is given by a constant value minus the lexical activity, thus envisaging evidence for nonword responses as a function of the time elapsing after stimulus onset (as for deadline models) modulated by the accumulation of lexical evidence through the competitive dynamics of leaky accumulators. Critically, however, this model features a thresholded transition from decision to responses, implemented as the typical action-triggering decisional boundary, which prevents any differentiation between words and nonwords at the level of MT.

Davis (2010), in the context of a more general model of orthographic processing and visual word recognition, implemented lexical decision as a process involving a competition between two different channels, one accumulating evidence for word, and one for nonword responses. A parameter controls lateral inhibition between channels, and the sources of input for the word-response channel are global and local levels of activity in the lexicon. Again, both word and nonword responses are delivered once a decision threshold has been reached. The model thus implements the assumption that responses are made once decisions have terminated, which seems to be questioned by the lexicality effect detected on MTs within our Experiments 1 and 3.

Compared to these notable models, it is worth noticing that in processing word-like nonwords, for which no memory trace is available, additional stages might be uniquely recruited to reach a decision. A potential candidate may be in a late verification stage (e.g., Perea et al., 2005), during which the stimulus is further evaluated in comparison with (a few) lexical units (relatively) close to it. Consistent with this proposal, for example, pseudowords derived from high-frequency words, despite triggering higher levels of lexical activation, yield *faster* response latencies (e.g., Yap et al., 2015; see also Perea et al., 2005; Ziegler et al., 2001), as they are compared against their lexical counterparts to check for deviations from the base stimuli. Assuming that these late verification processes may still be active during response execution, they would particularly increase MTs for pseudowords (for additional exploratory analyses, see online supplemental materials 4). It remains to be investigated whether the increased MTs merely reflect task-specific verification processes or are more broadly connected with mechanisms driving memory search termination processes (e.g., Dougherty et al., 2014).

Response Control and Monitoring

An alternative hypothesis we explored links the MTs to monitoring processes. Indeed, the duration of MTs has been associated—among other factors—to an online mechanism of executive control directed toward error detection and correction via the inhibition of the (erroneous) motor response (Allain et al., 2004). Additionally, recent proposals suggest that evidence-accumulation processes may proceed after a first decisional threshold is met. These would represent a second-order, metacognitive decision variable bounded to performance monitoring (Desender, Ridderinkhof, et al., 2021). Speculatively, MTs effects may be also linked with this continuing process of evidence accumulation beyond a first EMG-triggering boundary, to support an evolving monitoring process on the outcomes of first-order decisional stages (Desender et al., 2021; Pleskac & Busemeyer, 2010; Resulaj et al., 2009; for review and perspectives, see Desender, Ridderinkhof, et al., 2021). Empirically, we focused on how response accuracy changes as a function of response speed (CAFs)—which highlight conditions prone to fast and impulsive errors, thus calling for an allocation of additional control processes (e.g., Ridderinkhof, 2002; van den Wildenberg et al., 2010)—as well as on partial errors—which reflect real-time corrections of motor responses (e.g., Burle et al., 2002). Globally, our data offer mixed evidence of a relationship between MT effects and monitoring-related phenomena. Concerning partial errors, a higher likelihood of these phenomena for a specific class of stimuli does not seem to be necessarily associated with a slowdown in MTs. In Experiment 2, pseudo-objects revealed significantly longer MTs compared to real objects, despite the two types of items yielded statically comparable rates of partial errors. Further, Experiment 3 revealed an enhanced likelihood of partial error phenomena on low-frequency words, despite the lack of any frequency effect on MTs.

The presence of fast-impulsive errors seems more promisingly associated with a slowdown of MTs, at least in the context of the lexical decision task, where pseudowords were consistently more prone to impulsive errors compared to real words, whereas low- and high-frequency words were undistinguishable with respect to this index. However, results from Experiment 2 blur the overall pattern: Albeit pseudo-objects are qualitatively more prone to fast errors (errors in the first quantile of the conditional accuracy function), the lack of a significant object type by quantile interaction hinders any strong conclusion. Note that part of the inconclusiveness of our data on the relation between MT and monitoring processes might be due to the two indexes we adopted, which are either rather indirect (rates of fast errors) or focused on late monitoring components related to the correction of an ongoing behavior (partial errors). Other indexes, such as graded confidence ratings (Desender et al., 2018), or EEG components such as the error-positivity (e.g., Desender, Ridderinkhof, et al., 2021) may offer more direct measures of monitoring.

No Effect of Lexical Frequency on Motor Times

The lack of a word-frequency effect on MTs in lexical decision is apparently at odds with some data available in the literature. As noted in the introduction, Abrams and Balota (1991) reported clear effects of word frequency on responses delivered through left versus rightward movement of a handle. In line with proposals of the adaptive flow of information between cognitive and motor stages (Calderon et al., 2018), we consider that these more complex and time-consuming responses may provide more room for cross-talks between decision and action. Differently, the use of discrete responses such as the button presses used here, other than capitalizing on the traditional experimental setting used in most of the cognitive and neuroscientific experiments, may reveal different insights and dissociations across experimental factors in their ability to modulate response execution.

Further, in a previous study, we also reported that word frequency modulates EEG indexes of effector-selective motor activity (Scaltritti et al., 2020). With respect to this issue, we would like to notice that motor responses are a product of a complex and possibly hierarchical series of processes, involving response selection, planning/programming, and execution (e.g., Rosenbaum et al., 2007; Summers & Anson, 2009). Our previous work highlighted the word-frequency effect at the level of the lateralization of EEG beta activity occurring immediately before response onset and related to the settling of abstract and high-level motor goals (e.g., de Jong et al., 2006; Wheaton et al., 2005). The current experiments, instead, focused on pure measures of motor execution. Different variables might thus propagate their effects at different levels of the motor hierarchy. Importantly, a re-analysis of the previous dataset (Scaltritti et al., 2020) revealed the same pattern highlighted in the current experiments, with fully reliable lexicality effects on both PMT and MT measures, and a selective influence of word frequency on PMTs (supplemental materials 3). Other than corroborating the present findings, these results indeed point toward potential differences in the "cognitive" involvement of the motor hierarchy as a function of specific experimental manipulations and related latent decisional components. Clearly, this line of reasoning requires additional research.

Models of Decision Making

Although the current findings may inspire or even constrain formal models of decision making, we are agnostic with respect to the specific instantiation within the family of evidence-accumulation models that would be better suited to capture our results. Model fits and interpretations, however, may critically depend on the specific model and on the specific setting of its parameters (e.g., Donkin et al., 2011). A systematic comparison across models and parameter settings (e.g., Heathcote & Love, 2012; Rae et al., 2014), however, is beyond the scope of the present research.

Instead, by empirically testing the shared assumption that motor responses serially follow the termination of decisional stages, our investigation questions this core and general construct on which different models rely. Although similar findings have been reported in the field of perceptual decision making (Servant et al., 2021; Weindel et al., 2021), the assessment within different decisional paradigms based on the processing of semantic and lexical evidence sampled from memory revealed novel insights. Specifically, we began to assess different hypotheses concerning the functional characterization of the decisional components that are still active during motor-response execution. The results seem to favor the notion that these motor-decisional components may be related with verification (e.g., Paap et al., 1982; Perea et al., 2005) and/or control and monitoring dynamics (e.g., Allain et al., 2004; Burle et al., 2002; see also Weindel et al., 2021). However, any commitment on our part to one of the many and diverse extant modeling approaches (e.g., Calderon et al., 2018; Desender et al., 2021; Servant et al., 2021) seems premature, as—we believe—the informational content of the motor component still needs to be functionally elucidated, to better identify the linking function mapping the specified psychological processes onto a formal/computational implementation.

For example, the overly focus on the classic drift-diffusion model (Ratcliff, 1978, Ratcliff et al., 2016) is complicated by the fact that a single parameter (*Ter*) jointly captures early encoding stages and motor-response execution, under the assumption that both represent nondecisional processes. The selective contribution of the two stages to the decision process and/or the selective influence of different experimental manipulations on stimulus encoding versus motor-response execution is thus difficult to disentangle (e.g., Vergara-Martínez et al., 2020). Actually, the assessment of the correspondence between the models' parameters and the (presumed) specific cognitive process (for example, via test of selective influence) remains a different, albeit related, research question (as tackled, for example, in Weindel et al., 2021; see also Dutilh et al., 2019; Gomez & Perea, 2014; Heathcote & Love, 2012; Rae et al., 2014).

Differently, when considering the possibility of the post-decisional process of evidence accumulation, different frameworks have been proposed. Some authors (e.g., Servant et al., 2021) suggest that motor responses are informed by a continuation of the same evolving decision variable that shapes premotor stages. However, the differentiation between lexicality and word-frequency effects in their ability to affect MTs does not seem to fit with this perspective. Instead, even when considering models in which post-decisional evidence is explicitly linked to monitoring processes, it remains debated whether these rely on the same sources of information (i.e., evidence) as the ones used during first-order decisional processes (Desender et al., 2021). We thus believe that experimental data such as those highlighting the differences across experimental factors in their ability to affect premotor versus motor components of decision may provide a fertile and complementary ground to inform theories of decision making.

Conclusions

In conclusion, the present experiments show that motor-response execution is not segregated from ongoing decisional dynamics. However, for conceptual decision-making tasks relying on the evidence sampled from memory, the propagation of cognitive/decision processes onto motor responses does not seem to reflect (only) a continuous evolution of the same decision variable informing prior nonmotor stages. In fact, not all the manipulations traditionally ascribed to the rate of evidence accumulation reveal sizeable effects at the level of MTs. It would thus seem that the decision processes unfolding during motor responses are, at least in part, different compared to those driving purely nonmotor ones (Weindel et al., 2021). With respect to the specific informational content of these later processes, we can presently suggest some working hypotheses.

One possibility is that MT effects reflect, at least in part, processes related to performance monitoring for more demanding and confusable stimuli. Although, as mentioned above, our data fail to fully support this interpretation, a dismissal of the monitoring account seems premature at this stage. In our current reading, however, MTs effects yielded by the comparison between words/objects and pseudowords/pseudo-objects may reflect a byproduct of late-occurring verification processes selectively engaged for items featuring no previous representation in long-term memory stores. More broadly, we believe our data may foster a reconsideration on the MT measures. Clearly, MTs cannot be ascribed to purely nondecision components, suggesting that motor-response execution itself reflects also the unfolding of evolving cognitive/decision processes. However, our data also suggest that MTs do not simply mirror PMT/RT measures as not all the effects detected at the level of PMTs and RTs are reflected at the level of MTs (see also Weindel et al., 2021). More specifically, this implies that despite MTs are permeable to cognitive and decisional dynamics, (a) MTs might not be sensitive to all the same factors influencing RTs, and (b) not all PMT-related effects propagate onto MTs. In turn, dissociations among (cognitive and decision related) experimental manipulations in their ability to influence measures of motor-response execution may provide a finer-grained description of the crucial transition from decision onto action, which may instead remain blurred when considering overall RTs measures.

References

Abrams, R. A., & Balota, D. A. (1991). Mental chronometry: Beyond reaction time. *Psychological Science*, *2*(3), 153–157. https://doi.org/10.1111/j.1467-9280.1991.tb00123.x

Allain, S., Carbonnell, L., Burle, B., Hasbroucq, T., & Vidal, F. (2004). On-line executive control: An electromyographic study. *Psychophysiology*, *41*(1), 113–116. https://doi.org/10.1111/j.1469-8986.2003.00136.x

Bangert, A. S., Abrams, R. A., & Balota, D. A. (2012). Reaching for words and nonwords: Interactive effects of word frequency and stimulus quality on the characteristics of reaching movements. *Psychonomic Bulletin & Review*, *19*(3), 513–520. https://doi.org/10.3758/s13423-012-0234-x

Barca, L., & Pezzulo, G. (2012). Unfolding visual lexical decision in time. *PloS One*, *7*(4), Article e35932. https://doi.org/10.1371/journal.pone.0035932

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). *Parsimonious mixed models*. arXiv. https://doi.org/10.48550/arXiv.1506.04967.

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

Botwinick, J., & Thompson, L. W. (1966). Premotor and motor components of reaction time. *Journal of Experimental Psychology*, *71*(1), 9–15. https://doi.org/10.1037/h0022634

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings (technical report C-1)*. Center for Research in Psychophysiology, University of Florida.

Brodeur, M.B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The bank of standardized stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PloS One*, *5*(5), Article e10773. https://doi.org/10.1371/journal.pone.0010773

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of standardized stimuli (BOSS) phase II: 930 new normative photos. *PloS one*, *9*(9), Article e106953. https://doi.org/10.1371/journal.pone.0106953

Brown, S., & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review*, *112*(1), 117–128. https://doi.org/10.1037/0033-295X.112.1.117

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*(3), 153–178. https://doi.org/10.1016/j.cogpsych.2007.12.002

Brysbaert, M.. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*(1). Article 16. https://doi.org/10.5334/joc.72

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). The impact of word prevalence on lexical decision times: Evidence from the Dutch Lexicon Project 2. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(3), 441–458. https://doi.org/10.1037/xhp0000159

Burle, B., Possamai, C. A., Vidal, F., Bonnet, M., & Hasbroucq, T. (2002). Executive control in the Simon effect: An electromyographic and distributional analysis. *Psychological Research*, *66*(4), 324–336. https://doi.org/10.1007/s00426-002-0105-6

Calderon, C. B., Gevers, W., & Verguts, T. (2018). The unfolding action model of initiation times, movement times, and movement paths. *Psychological Review*, *125*(5), 785–805. https://doi.org/10.1037/rev0000110

Calderon, C. B., Verguts, T., & Gevers, W. (2015). Losing the boundary: Cognition biases action well after action selection. *Journal of Experimental Psychology: General*, *144*(4), 737–743. https://doi.org/10.1037/xge0000087

Chapman, C. S., Gallivan, J. P., Wood, D. K., Milne, J. L., Culham, J. C., & Goodale, M. A. (2010). Reaching for the unknown: Multiple target encoding and real-time decision-making in a rapid reach task. *Cognition*, *116*(2), 168–176. https://doi.org/10.1016/j.cognition.2010.04.008

Cisek, P. (2007). Cortical mechanisms of action selection: The affordance competition hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1485), 1585–1599. https://doi.org/10.1098/rstb.2007.2054

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*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Crepaldi, D., Keuleers, E., Mandera, P., & Brysbaert, M. (2013). *SUBTLEX-IT*. Retrieved from http://crr.ugent.be/subtlex-it/

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*(3), 713–758. https://doi.org/10.1037/a0019738

de Jong, R., Gladwin, T. E., & M't Hart, B. (2006). Movement-related EEG indices of preparation in task switching and motor control. *Brain Research*, *1105*(1), 73–82. https://doi.org/10.1016/j.brainres.2006.03.030

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Desender, K., Boldt, A., & Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. *Psychological Science*, *29*(5), 761–778. https://doi.org/10.1177/0956797617744771

Desender, K., Donner, T. H., & Verguts, T. (2021). Dynamic expressions of confidence within an evidence accumulation framework. *Cognition*, *207*, Article 104522. https://doi.org/10.1016/j.cognition.2020.104522

Desender, K., Ridderinkhof, K. R., & Murphy, P. R. (2021). Understanding neural signals of post-decisional performance monitoring: An integrative review. *Elife*, *10*, Article e67556. https://doi.org/10.7554/eLife.67556

Donkin, C., Brown, S., & Heathcote, A. (2011). Drawing conclusions from choice response time models: A tutorial using the linear ballistic accumulator. *Journal of Mathematical Psychology*, *55*(2), 140–151. https://doi.org/10.1016/j.jmp.2010.10.001

Donkin, C., Brown, S. D., & Heathcote, A. (2009). The overconstraint of response time models: Rethinking the scaling problem. *Psychonomic Bulletin & Review*, *16*(6), 1129–1135. https://doi.org/10.3758/PBR.16.6.1129

Donkin, C., Heathcote, A., Brown, S., & Andrews, S. (2009). Non-decision time effects in the lexical decision task. In *Proceedings of the 31st annual conference of the cognitive science society*. Austin, TX: Cognitive Science Society.

Dougherty, M. R., Harbison, J. I., & Davelaar, E. J. (2014). Optional stopping and the termination of memory retrieval. *Current Directions in Psychological Science*, *23*(5), 332–337. https://doi.org/10.1177/0963721414540170

Dufau, S., Grainger, J., & Ziegler, J. C. (2012). How to say "no" to a nonword: A leaky competing accumulator model of lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4):1117–28. https://doi.org/10.1037/a0026948

Dutilh, G., Annis, J., Brown, S. D., Cassey, P., Evans, N. J., Grasman, R. P. P. P., Hawkins, G. E., Heathcote, A., Holmes, W. R., Krypotos, A.- M., Kupitz, C. N., Leite, F. P., Lerche, V., Lin, Y.-S., Logan, G. D., Palmeri, T. J., Starns, J. J., Trueblood, J. S., van Maanen, L. … Donkin, C. (2019). The quality of response time data inference: A blinded, collaborative assessment of the validity of cognitive models. *Psychonomic Bulletin & Review*, *26*, 1051–1069. https://doi.org/10.3758/s13423-017-1417-2

E-Prime (Version 2.0.10.356) [Computer software]. Psychology Software Tools.

Farmer, T. A., Cargill, S. A., Hindy, N. C., Dale, R., & Spivey, M. J. (2007). Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing. *Cognitive Science*, *31*(5), 889–909. https://doi.org/10.1080/03640210701530797

Fernández-López, M., Davis, C. J., Perea, M., Marcet, A., & Gómez, P. (2022). Unveiling the boost in the sandwich priming technique. *Quarterly Journal of Experimental Psychology*, *75*(7), 1382–1393. https://doi.org/10.1177/17470218211055097

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., Augustinova, M., & Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, *42*(2), 488–496. https://doi.org/10.3758/BRM.42.2.488

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*(12), 2455–2466. https://doi.org/10.1080/17470218.2014.937447

Gordon, J., Maselli, A., Lancia, G. L., Thiery, T., Cisek, P., & Pezzulo, G. (2021). The road towards understanding embodied decisions. *Neuroscience and Biobehavioral Reviews*, *131*, 722–736.

Goslin, J., Galluzzi, C., & Romani, C. (2014). Phonitalia: A phonological lexicon for Italian. *Behavior Research Methods*, *46*(3), 872–886. https://doi.org/10.3758/s13428-013-0400-8

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *103*(3), 518–565. https://doi.org/10.1037/0033-295X.103.3.518

Green, P. and MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution* *7*(4): 493–498. https://doi.org/10.1111/2041-210X.12504

Green, P., MacLeod, C. J. and Alday, P. (2016). *Package 'simr'*. Available at https://cran.r-project.org/web/packages/simr/simr.pdf

Hasbroucq, T., Possamaï, C.-A., Bonnet, M., & Vidal, F. (1999). Effect of the irrelevant location of the response signal on choice reaction time: An electromyographic study in humans. *Psychophysiology*, *36*(*4*), 522–526. https://doi.org/10.1017/S0048577299001602

Heathcote, A., & Love, J. (2012). Linear deterministic accumulator models of simple choice. *Frontiers in Psychology*, *3*, Article 292. https://doi.org/10.3389/fpsyg.2012.00292

Horst, J. S., & Hout, M. C. (2016). The novel object and unusual name (NOUN) database: A collection of novel images for use in experimental research. *Behavior Research Methods*, *48*(4), 1393–1409. https://doi.org/10.3758/s13428-015-0647-3

Keuleers, E. (2013). *vwr: Useful functions for visual word recognition research* (R Package Version 0.3.0) [Computer software]. https://CRAN.R-project.org/package=vwr

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627–633. https://doi.org/10.3758/BRM.42.3.627

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods*, *44*(1), 287–304. https://doi.org/10.3758/s13428-011-0118-4

Liu, J., & Liu, Q. (2016). Use of the integrated profile for voluntary muscle activity detection using EMG signals with spurious background spikes: A study with incomplete spinal cord injury. *Biomedical Signal Processing and Control*, *24*, 19–24. https://doi.org/10.1016/j.bspc.2015.09.004

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

McKinstry, C., Dale, R., & Spivey, M. J. (2008). Action dynamics reveal parallel competition in decision-making. *Psychological Science*, *19*(1), 22–24. https://doi.org/10.1111/j.1467-9280.2008.02041.x

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the affective norms for English words (ANEW) for Italian. *Behavior Research Methods*, *46*(3), 887–903. https://doi.org/10.3758/s13428-013-0405-3

Moreno, M. A., Stepp, N., & Turvey, M. T. (2011). Whole body lexical decision. *Neuroscience Letters*, *490*(2), 126–129. https://doi.org/10.1016/j.neulet.2010.12.041

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau. *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64. 10.20982/tqmp.04.2.p061

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation–verification model for letter and word recognition: The word-superiority effect. *Psychological Review*, *89*(5), 573–594. https://doi.org/10.1037/0033-295X.89.5.573

Perea, M., Rosa, E., & Gómez, C. (2005). The frequency effect for pseudowords in the lexical decision task. *Perception & Psychophysics*, *67*(2), 301–314. https://doi.org/10.3758/bf03206493

Pezzulo, G., & Cisek, P. (2016). Navigating the affordance landscape: Feedback control as a process model of behavior and cognition. *Trends in Cognitive Sciences*, *20*(6), 414–424. https://doi.org/10.1016/j.tics.2016.03.013

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review*, *117*(3), 864–901. https://doi.org/10.1037/a0019737

Rae, B., Heathcote, A., Donkin, C., Averell, L., & Brown, S. (2014). The hare and the tortoise: Emphasizing speed can change the evidence used to make decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1226–1243. https://doi.org/10.1037/a0036801

Raftery, A. E. (1995). Bayesian model selection in social research. In P. V. Marsden (Ed.), *Sociological methodology 1995* (pp. 111–196). Blackwell.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/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*(1), 159–182. https://doi.org/10.1037/0033-295X.111.1.159

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

R Core Team. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature*, *461*(7261), 263–266. https://doi.org/10.1038/nature08275

Ridderinkhof, K. R. (2002). Micro- and macro-adjustments of task set: Activation and suppression in conflict tasks. *Psychological Research*, *66*(4), 312–323. https://doi.org/10.1007/s00426-002-0104-7

Rochet, N., Spieser, L., Casini, L., Hasbroucq, T., & Burle, B. (2014). Detecting and correcting partial errors: Evidence for efficient control without conscious access. *Cognitive, Affective, and Behavioral Neuroscience*, *14*, 970–982. https://doi.org/10.3758/s13415-013-0232-0

Rosenbaum, D. A., Cohen, R. G., Jax, S. A., Weiss, D. J., & Van Der Wel, R. (2007). The problem of serial order in behavior: Lashley's legacy. *Human Movement Science*, *26*(4), 525–554. https://doi.org/10.1016/j.humov.2007.04.001

Scaltritti, M. (2022, May 6). *Motor times in lexical and object decision*. https://osf.io/6hqk5/

Scaltritti, M., Job, R., Alario, F. X., & Sulpizio, S. (2020). On the boundaries between decision and action: Effector-selective lateralization of beta-frequency power is modulated by the lexical frequency of printed words. *Journal of Cognitive Neuroscience*, *32*(11), 2131–2144. https://doi.org/10.1162/jocn_a_01606

Scaltritti, M., Job, R., & Sulpizio, S. (2021). Selective suppression of taboo information in visual word recognition: Evidence for cognitive control on semantics. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(7), 934–945. https://doi.org/10.1037/xhp0000917

Servant, M., Logan, G. D., Gajdos, T., & Evans, N. J. (2021). An integrated theory of deciding and acting. *Journal of Experimental Psychology: General*, *150*(12), 2435–2454. https://doi.org/10.1037/xge0001063

Servant, M., White, C., Montagnini, A., & Burle, B. (2015). Using covert response activation to test latent assumptions of formal decision-making models in humans. *Journal of Neuroscience*, *35*(28), 10371–10385. https://doi.org/10.1523/JNEUROSCI.0078-15.2015

Servant, M., White, C., Montagnini, A., & Burle, B. (2016). Linking theoretical decision-making mechanisms in the simon task with electrophysiological data: A model-based neuroscience study in humans. *Journal of Cognitive Neuroscience*, *28*(10), 1501–1521. https://doi.org/10.1162/jocn_a_00989

Siegel, M., Engel, A. K., & Donner, T. H. (2011). Cortical network dynamics of perceptual decision-making in the human brain. *Frontiers in Human Neuroscience*, *5*, Article 21. https://doi.org/10.3389/fnhum.2011.00021

Singman, H., Bolker, B., Westfal, J., Aust, F., & Ben-Shachar, M. S. (2021). *afex: Analysis of Factorial Experiments*. R package version 0.28-1. https://CRAN.R-project.org/package=afex

Smigasiewicz, K., Ambrosi, S., Blaye, A., & Burle, B. (2020). Inhibiting errors while they are produced: Direct evidence for error monitoring and inhibitory control in children. *Developmental Cognitive Neuroscience*, *41*. Article 100742. https://doi.org/10.1016/j.dcn.2019.100742

Smith, P. L., & Lilburn, S. D. (2020). Vision for the blind: Visual psychophysics and blinded inference for decision models. *Psychonomic Bulletin & Review*, *27*(5), 882–910. https://doi.org/10.3758/s13423-020-01742-7

Smith, P. L., & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology*, *32*(2), 135–168. https://doi.org/10.1016/0022-2496(88)90043-0

Song, J. H., & Nakayama, K. (2009). Hidden cognitive states revealed in choice reaching tasks. *Trends in Cognitive Sciences*, *13*(8), 360–366. https://doi.org/10.1016/j.tics.2009.04.009

Spieser, L., Servant, M., Hasbroucq, T., & Burle, B. (2017). Beyond decision! Motor contribtion to speed-accuracy trade-off in decision making. *Psychonomic Bulletin & Review*, *24*(*3*), 950–956. https://doi.org/10.3758/s13423-016-1172-9

Spivey, M. J., & Dale, R. (2006). Continuous dynamics in real-time cognition. *Current Directions in Psychological Science*, *15*(5), 207–211. https://doi.org/10.1111/j.1467-8721.2006.00437.x

Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences*, *102*(29), 10393–10398. https://doi.org/10.1073/pnas.0503903102

Summers, J. J., & Anson, J. G. (2009). Current status of the motor program: Revisited. *Human Movement Science*, *28*(5), 566–577. https://doi.org/10.1016/j.humov.2009.01.002

Székely, A., & Bates, E. (2000). Objective visual complexity as a variable in studies of picture naming. *Center for Research in Language Newsletter*, *12*(2), 3–33. http://crl.ucsd.edu/newsletter

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550–592. https://doi.org/10.1037/0033-295X.108.3.550

Van Den Wildenberg, W. P., Wylie, S. A., Forstmann, B. U., Burle, B., Hasbroucq, T., & Ridderinkhof, K. R. (2010). To head or to heed? Beyond the surface of selective action inhibition: A review. *Frontiers in Human Neuroscience*, *4*, Article 222. https://doi.org/10.3389/fnhum.2010.00222

Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review*, *7*(2), 208–256. https://doi.org/10.3758/bf03212980

Vergara-Martínez, M., Gomez, P., & Perea, M. (2020). Should I stay or should I go? An ERP analysis of two-choice versus go/no-go response procedures in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(11), 2034–2048. https://doi.org/10.1037/xlm0000942

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/bf03194105

Weindel, G., Anders, R., Alario, F., & Burle, B. (2021). Assessing model-based inferences in decision-making with single-trial response time decomposition. *Journal of Experimental Psychology: General*, *150*(8), 1528–1555. https://doi.org/10.1037/xge0001010

Welsh, T. N. & Elliott, D. (2004). Movement trajectories in the presence of a distracting stimulus: Evidence for a response activation model of selective reaching. *Quarterly Journal of Experimental Psychology A*, *57*(6), 1031–1057 https://doi.org/10.1080/02724980343000666

Welsh, T. N. & Elliott, D. (2005).The effects of response priming on the planning and execution of goal-directed movements in the presence of a distracting stimulus. *Acta Psychologica*, *119*(2), 123–142. https://doi.org/10.1016/j.actpsy.2005.01.001

Wheaton, L. A., Shibasaki, H., & Hallett, M. (2005). Temporal activation pattern of parietal and premotor areas related to praxis movements. *Clinical Neurophysiology*, *116*(5), 1201–1212. https://doi.org/10.1016/j.clinph.2005.01.001

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, *60*(4), 502–529. https://doi.org/10.1016/j.jml.2009.02.001

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 53–79. https://doi.org/10.1177/1747021820906566

Yap, M. J., Liow, S. J. R., Jalil, S. B., & Faizal, S. S. B. (2010). The Malay Lexicon Project: A database of lexical statistics for 9,592 words. *Behavior Research Methods*, *42*(4), 992–1003. https://doi.org/10.3758/BRM.42.4.992

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*(3), 597–613. https://doi.org/10.1037/xlm0000064

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), 971–979.

Ziegler, J. C., Jacobs, A. M., & Klüppel, D. (2001). Pseudohomophone effects in lexical decision: Still a challenge for current word recognition models. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(3), 547–559. https://doi.org/10.1037//0096-1523.27.3.547

Received July 25, 2022
Revision received January 2, 2023
Accepted January 6, 2023 ■