CrossMark

BRIEF REPORT

The Generality of Dynamic Adjustments in Decision Processes across Trials and Tasks

Andrew J. Aschenbrenner[1] · Melvin J. Yap[2] · David A. Balota[1]

Abstract There is a growing interest in understanding how aspects of binary decision-making change dynamically at the trial level. For example, in lexical decision, there is a well-established interaction between current and previous trial characteristics (e.g., lexicality and stimulus degradation) that suggests that participants adjust their decision processes based on the relative match in signal strength between the current and previous trial. The present study assessed the generality of this finding by examining the presence of such cross-trial adjustments in two new tasks, syntactic classification, and memory scanning. Stimulus degradation is manipulated in both tasks. Results indicate that response latencies are facilitated when salient aspects of the stimulus repeat across trials. These findings are interpreted within the context of a flexible processor that utilizes information from the prior target to prime the relevant processing pathway on the current trial. Candidate models that potentially can accommodate the pattern are briefly discussed.

Keywords Decision making · Reaction time analysis

For nearly 150 years, response time (RT) from binary decision tasks has been used to provide insights into the speed and duration of underlying cognitive operations (Donders, 1868/1969; Sternberg, 1969). By measuring RT differences across levels of an independent variable, one can make inferences about the processes that contribute to performance in a given task. This approach implicitly assumes each trial is a relatively isolated and independent event and that a variable influences RT in the same way regardless of prior events. However, there is a wealth of evidence indicating that responses are influenced by other stimuli or trials in the task. For example, the magnitude of interference in the Stroop task depends on the relative proportion of congruent to incongruent trials in the list (Logan & Zbrodoff, 1979) and the proportion of related primes can change the magnitude of semantic priming in the lexical decision task (Neely, Keefe, & Ross, 1989). Such list-wide proportion effects have been typically interpreted as a global orientation of attention towards aspects of the stimuli that are the most contextually relevant, as determined by prior exposure to the task stimuli.

Additionally, there has been growing interest in how aspects of processing might adjust dynamically across adjacent trials. Returning to the Stroop task, it has been shown repeatedly that the magnitude of interference exhibited on Trial N changes depending on the congruency of the stimulus that was presented on Trial N-1 (Gratton, Coles, & Donchin, 1992). This phenomenon has been interpreted within a number of distinct frameworks. For example, it has been suggested that cognitive control can be regulated based on the demands of the preceding trial (Botvinick, Braver, Barch, Carter, & Cohen, 2001) or that priming of particular stimulus dimensions remains active in the system influencing the subsequent trial (Aschenbrenner & Balota, 2015). Of course, interference tasks are designed to place strong demands on attentional selection and it is unclear whether cross-trial mechanisms will operate in a similar fashion across tasks that predominantly tap other cognitive domains.

Consider for example the lexical decision task (LDT), one of the premier tasks used to study visual word recognition. The influence of multiple variables on the speed of classifying stimuli as a word or nonword has provided powerful

✉ Andrew J. Aschenbrenner
  a.aschenbrenner@wustl.edu

[1] Washington University in St. Louis, St. Louis, MO, USA

[2] National University of Singapore, Singapore, Singapore

🐾 Springer

constraints on extant models (Yap & Balota, 2015). A major focus of the present report is on stimulus quality (SQ). Perceptually degrading a stimulus markedly slows early, feature-level processing and can therefore be used as a diagnostic marker for where in the cognitive processing stream other variables operate, based on patterns of additivity or interactivity. For example, in his classic study, Sternberg (1969) demonstrated that SQ is additive with set size in short-term memory scanning. He interpreted this as evidence that SQ influences an early encoding stage and that set-size influences a later scanning process.

Returning to lexical decision, the effect of specific variables on RT has been shown to be systematically influenced by the parameters of the preceding trial. Specifically, there is now considerable evidence for a highly reliable interaction among the following three variables: SQ of the current target, SQ of the prior target and lexicality of the prior target (Balota, Aschenbrenner, & Yap, 2013, 2016; Masson & Kliegl, 2013; Masson, Rabe, & Kliegl, 2017). This interaction indicates that responses are faster when aspects of the current trial (in this case lexicality and SQ) match those of the previous trial. For example, RTs to clear words are faster when the preceding trial was also a clear word relative to when the previous trial was a degraded word. In contrast, when lexicality changed across trials, changes in SQ have little to no influence on RTs. Balota et al. (2016) extended this work by examining the influence on nonword targets and found a very similar pattern.

These recent studies add to the already established literature on cross-trial lexicality effects. Specifically, in both lexical decision and speeded word naming, responses have been shown to be faster following word targets compared to following nonword targets (Lima & Huntsman, 1997; Perea & Carreiras, 2003; Taylor & Lupker, 2001). This pattern can be easily accommodated by a dynamically adjusted response threshold such that participants selectively slow down following relatively difficult targets (i.e., nonwords). However, this account cannot accommodate the recent findings with SQ, because such a model predicts that all responses after the most difficult items (e.g., degraded items) should be slower than their clearly presented counterparts, which is clearly not the case. Rather, it is the *match* in both lexicality and SQ across trials that predicts modulation of RTs.

Before assessing whether extant cognitive models can accommodate the robust four-way interaction amongst SQ and lexicality of the previous trial with SQ and lexicality of the current trial, it is important to determine whether the results are idiosyncratic to lexical decision. This idiosyncrasy might be expected, because LDT places a premium on familiarity and has other task-specific constraints (Balota & Chumbley, 1984; Besner, 1983; Gomez, 2012). Therefore, these findings may not extend to other tasks that do not have the same constraints. Thus, the primary question addressed in this report is whether the critical interaction among current and previous

SQ and current and previous target reflects a domain-general decision process or is produced by the constraints imposed by LDT. To accomplish this goal, we first examine cross-trial effects from a previously published syntactic classification task (noun vs. verb, Yap & Pexman, 2016). This study is described only briefly, because full details can be found in the original paper. We then extend the findings beyond the realm of visual word recognition using data from a newly collected short term memory scanning experiment. Importantly, both studies involve stimulus degradation as a variable to parallel the analyses from the lexical decision studies discussed above.

Analysis 1: Yap and Pexman (2016)

Method

Participants and Procedure

Thirty-two participants were recruited from the University of Calgary to participate in two separate syntactic classification experiments. The two experiments differed only on the semantic dimensions that were manipulated (concreteness and number of features vs. semantic neighborhood density and ambiguity) and thus were combined for the purposes of the present analysis. Participants were shown a target item in the center of the screen and were asked to indicate if the item was a noun or a verb. A total of 240 nouns and 240 verbs were presented. The nouns varied on several semantic dimensions (concreteness and number of features), which were not considered here. Stimulus quality was manipulated within subjects and degraded items were produced by rapidly alternating a pattern mask consisting of symbols (@#$%&?) with the target item. The pattern mask was randomly generated anew for each item and was constrained to be the same length as the target item. Degradation was counterbalanced across items.

Results

For the analysis of RTs, we first discarded error trials and any trial with latency shorter than 200 ms or longer than 3,000 ms. Following this, we eliminated any RT that was more than 2.5 standard deviations from the mean of the SQ condition for a given subject. We did this to avoid eliminating a disproportionate number of degraded stimuli which are slower on average. This procedure eliminated 4% of the total correct responses. To facilitate comparison across experiments and to control for individual differences in baseline processing speed, all RTs were transformed into z-scores based on each individual's mean and standard deviation across all conditions (Faust,

Balota, Spieler, & Ferraro, 1999). Thus, effects are expressed in standard deviation units.

The z-scores were analyzed with linear mixed effects models using the lme4 package in R (Bates, Maechler, Bolker, & Walker, 2015). The four factors (current and previous SQ and current and previous syntactic category) were coded using −0.5/+0.5 contrasts. We included as many random effects as were supported by the data determined by sequentially adding random intercepts and slopes until a likelihood ratio test no longer indicated improvement in model fit. Significance of fixed effects was evaluated with t-tests of regression coefficients with degrees of freedom estimated using the Satterthwaite approximation in the lmerTest package (Kuznetsova, Brockhott, & Christensen, 2015). We first analyze the full four-way interaction and then perform follow-up comparison within levels of each target status (e.g., nouns vs. verbs). For simplicity, we report only the significant main effects of the four variables and the critical, four-way interaction.[1]

Random intercepts of subjects and items as well as random slopes of current and previous target and current SQ across subjects were included as random effects. Analyses revealed a significant main effect of SQ ($\beta = 0.166$, $t = 4.85$, $p < 0.0001$), indicating slower responses to degraded stimuli, and of previous target ($\beta = 0.138$, $t = 4.35$, $p < 0.0001$), indicating slower responses when the prior trial was a verb. More importantly, the four-way interaction among current and previous trial syntax and current and previous trial SQ was significant ($\beta = -0.479$, $t = -5.14$, $p < 0.0001$). Follow-up analyses indicated that the three-way interaction among previous syntax and current and previous degradation was reliable for both nouns ($\beta = 0.148$, $t = 2.27$, $p = 0.023$) and for verbs ($\beta = -0.323$, $t = -4.84$, $p < 0.0001$).

To visualize this four-way interaction, it is helpful to consider the impact of changing SQ across trials holding the other variables constant (current and previous target). The interaction conceptualized in this fashion is illustrated in Fig. 1. The top panel contains data from noun targets on the current trial, and the bottom panel contains data from the verb targets. The left side of the graph illustrates noun targets on the previous trial, and the right side illustrates verb targets on the previous trial. Thus, each quadrant of the 2 by 2 graph plots alternation versus repetition of SQ within a level of current and previous trial syntax. The slope of the line in each graph is the degradation effect which varies as a function of whether the prior target was clear (solid lines) or degraded (dashed lines). In the top-left panel, the two-way interaction between current and previous trial SQ was reliable ($\beta = -0.131$, $t = -2.86$, $p = 0.004$), which indicates that RTs are faster when SQ repeats across two trials relative to when SQ alternates. Similarly, in

the bottom right panel (verbs on both the current and previous target), the two-way interaction also was reliable ($\beta = -0.288$, $t = -6.10$, $p < 0.001$), again indicating RTs are shorter when SQ repeats relative to when SQ alternates. However, when syntax also alternated across trials (as in the top right and bottom left panels), there was no modulation of RTs for either current noun targets ($\beta = 0.02$, $t = 0.36$, $p = 0.72$), nor for current verb targets ($\beta = 0.04$, $t = 0.76$, $p = 0.45$). The absence of an interaction indicates that RTs on trials that alternated SQ were no different from when SQ repeated, at least when the syntactic category also changed from the prior trial. Thus, when only one dimension changes across trials (top left and bottom right panels) RTs are modulated, but when more than one dimension changes (top right and bottom left) RTs are not influenced, a pattern which produces the four-way interaction.

Interim Discussion

These analyses make it clear that the interaction among current and previous trial SQ with current and previous trial "status" (in this case, syntactic category) is not a phenomenon that is idiosyncratic to the task demands of LDT. Indeed, these patterns may reflect a more general decision making mechanism. However, before speculating on this, we investigate the four-way interaction again in a short-term memory task to examine the domain generality of this effect. We do so using a variant of the classic Sternberg short-term memory scanning paradigm.

Current Experiment

Method

Participants

A total of 160 undergraduates were recruited from the National University of Singapore for this study. The memory scanning task was embedded within a battery of other tasks that included lexical decision, vocabulary, and working memory assessments. Participants were required to have normal or corrected-to-normal vision and be a native speaker of English.

Stimuli and Procedure

The stimuli used for this task were the following consonants: B, D, F, G, H, M, N, Q, R, T. Participants were presented with letters one at a time in the center of the screen. These letters formed the "memory set." Each letter in the set was presented for 500 ms and separated by a 100-ms blank screen ISI. The letters that comprised the memory set were selected at random, without replacement, on each trial. After the entire memory set was displayed, three asterisks (***) were displayed for

---

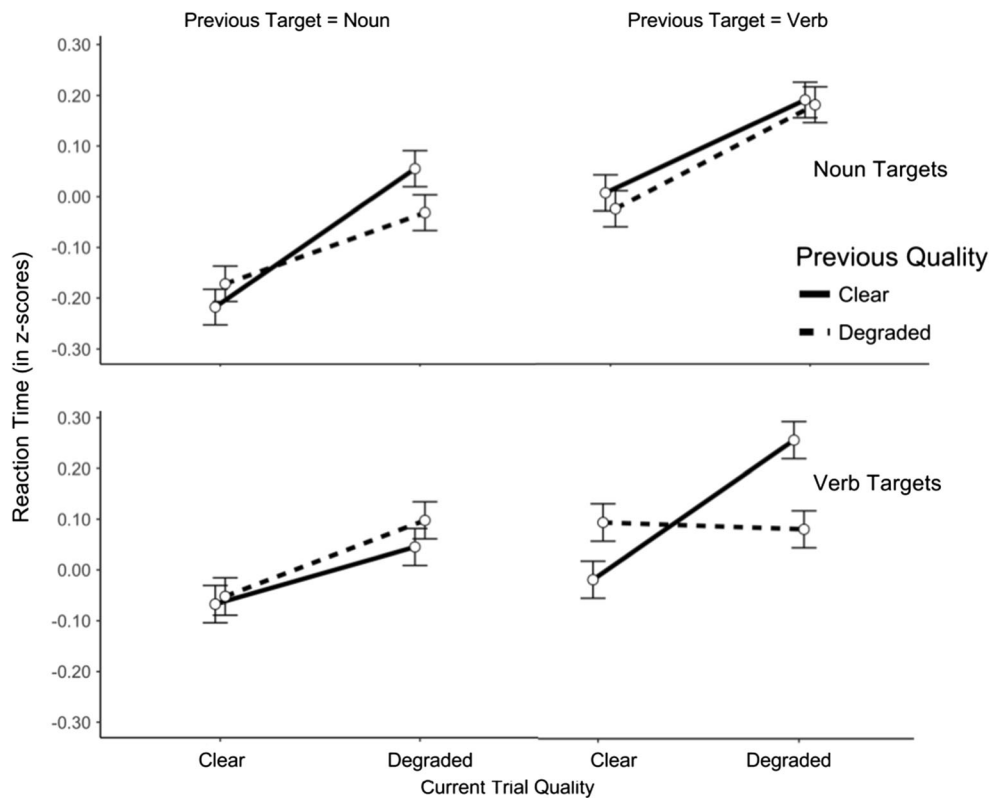[1] Analysis of accuracy rates is provided in the Appendix for the interested reader.

Fig. 1 Interaction among current and previous SQ with current and previous syntax in Experiment 1. *Error bars* represent the standard error of the mean

400 ms. This indicated to the participants that the next item was the memory probe. They were instructed to respond with the "/" key if the probe letter was contained in the memory set and the "z" key if it was not.

Three variables were manipulated in this task. First, we varied the size of the memory set to contain one, two, or four letters. Second, we manipulated whether the memory probe was present in the set or absent. Finally, we varied whether the probe was degraded or clearly presented. Only the probe letter was perceptually degraded; the items in the memory set were always presented in a clear fashion. The method of degradation was the same as described in Experiment 1. Participants completed 24 trials at each level of set size, SQ, and target presence for a total of 288 trials. For the purposes of the present analysis, we collapsed across set size and analyze only current and previous trial SQ and previous and current trial target presence to maintain consistency with prior studies.

Results

The same trimming procedure was used as in Experiment 1 and removed 4% of the correct responses. The random effects structure included random intercepts for subjects and items as well as uncorrelated random slopes of target presence, previous target, and current trial SQ across subjects. The main effects of current target ($\beta = 0.197$, $t = 10.33$, $p < 0.001$), previous target ($\beta = 0.097$, $t = 9.16$, $p < 0.001$), SQ ($\beta = 0.239$, $t = 9.50$, $p < 0.001$), and previous SQ ($\beta = 0.019$, $t = 2.20$, $p = 0.028$) were all reliable. These main effects indicate that RTs on the current trial were faster for present targets compared to absent, were faster when the previous target was present compared to when it was absent, were faster when the current trial was clear rather than degraded and were faster when the previous target was clear rather than degraded, respectively. Most importantly, the four-way interaction was again significant ($\beta = -.241$, $t = -3.52$, $p < .001$). Follow-up comparisons again indicated the three-way interaction was separately reliable both for present targets ($\beta = 0.10$, $t = 2.07$, $p = 0.039$) and for absent targets ($\beta = -0.143$, $t = -2.87$, $p = 0.004$).

This four-way interaction, conceptualized in the same manner as in Experiment 1, is displayed in Fig. 2. Responses to present targets are in the top panels and absent targets are in the bottom. Previous present targets are on the left and previous absent targets on the right. Critically, the overall pattern replicates the prior results. Specifically, the interaction between current previous trial SQ was significant when two sequential targets were present ($\beta = -0.07$, $t = -2.190$, $p = 0.028$) or absent ($\beta = -0.213$, $t = -5.89$, $p < 0.001$) but was not reliable when target presence alternated ($\beta = 0.02$, $t = 0.70$, $p = 0.48$) for present targets (Fig. 2, top right panel) and marginally for absent targets ($\beta = -0.07$, $t = -2.03$, $p = 0.04$).
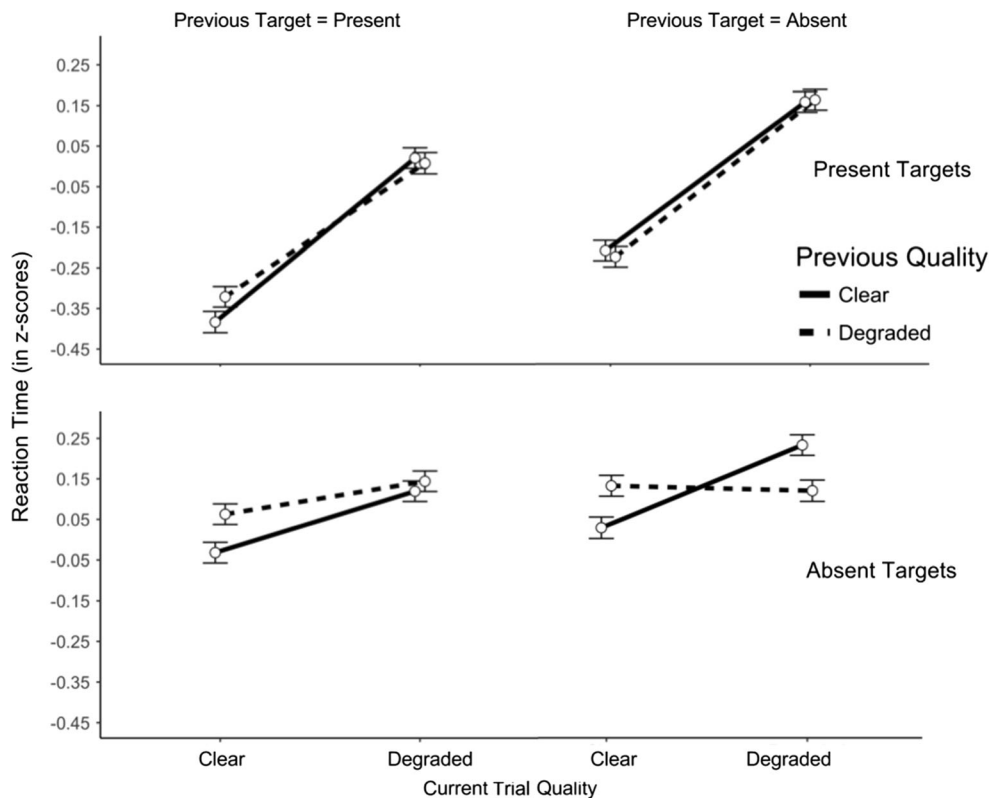
Fig. 2 Interaction among current and previous SQ with current and previous target presence in Experiment 2. *Error bars* represent the standard error of the mean

General Discussion

Across two experiments in different cognitive domains (visual word recognition and short-term memory retrieval) and different university populations, we showed evidence for a reliable and consistent four-way interaction among the SQ and target status of the current trial and SQ and target status of the previous trial. At a general level, this interaction indicates that when target status remains constant across trials, changes in SQ greatly influence RT, but when target status changes, differences in SQ have little or no influence on RT. These modulations in the SQ effect can be quite large. For example, across both experiments, the magnitude of the SQ effect varied by 54% (Experiment 1) and 36% (Experiment 2) of the average SQ difference and, as shown in bottom right panel of Figs. 1 and 2, can be totally eliminated depending upon the cross-trial contingencies. This pattern replicates and extends recent studies investigating SQ and lexicality in LDT. Given the stability of these findings, we now turn to a brief discussion of candidate models that have been proposed, which might be able to accommodate these patterns.

Based on detailed computational modeling of a color categorization task, Little, Wang, and Nosofsky (2016) argued that two separate mechanisms were required to account for their sequential effects. Specifically, there is a bias to switch category responses when stimuli change and a persistence of

activation of recently presented items. It is important to note, however, that their results were dominated by repetition effects (faster RTs to repeated stimuli) and such repetitions were not as prevalent in our experiments (items never repeated in Experiment 1 and were temporally separated in Experiment 2). Similarly, models of sequential effects from simple RT paradigms (Jones, Curran, Mozer, & Wilder, 2013) rely on learning of category base rates and frequency of alternations versus repetitions of stimuli. Thus, these models may not extend to the present tasks in a straightforward manner.

A dynamic signal detection model proposed by Turner, Van Zandt, and Brown (2011) suggests that participants continually update the representations used to produce a binary decision as a function of prior exemplars of the task stimuli. In lexical decision, words and nonwords can be thought of as distributed along a familiarity axis (Balota & Chumbley, 1984). The degree of familiarity then drives the word/ nonword decision. Assuming that perceptual degradation disrupts familiarity, when SQ is included in the task design, the familiarity axis can then be divided into four distinct regions ranging from clear words (highest familiarity) to degraded nonwords (lowest familiarity). Therefore, if a stimulus on the current trial has a similar familiarity value as the previous stimulus, the same response can be quickly (and accurately) executed. Similarly, if a stimulus has a very distal familiarity value from the prior stimulus, the opposite response can

be quickly endorsed. This provides the basis for the RT facilitation across trials that match (or mismatch) in both lexicality and SQ. Alternatively, Masson et al. (2017) suggested that a correct response to a particular stimulus type (e.g., clear words) lowers the response threshold for that type of item. Thus, if another clear word follows, it will have a lower threshold to exceed, thereby leading to faster responses (Dufau, Grainger, & Ziegler, 2012).

Although the Turner et al. and Masson et al. models are useful in accounting for aspects of the current data, these models would appear to have some difficulty accommodating the fact that previous trial variables have not shown consistent interactions with other variables that influence RT on the current trial. For example, the robust word frequency effect is not consistently modulated by previous trial lexicality, frequency, or SQ (Balota et al., 2016; Masson & Kliegl, 2013). Because word frequency influences the rate of processing, lowering a response boundary should affect slower items (e.g., low frequency words) to a greater degree than faster items (high frequency words), leading to an interaction. Similarly, word frequency effects in LDT are largely redundant with familiarity and therefore one would once again expect an interaction within the signal detection framework.

Finally, a model proposed by Annis and Malmberg (2013) suggests that lapses in attentional control cause features or information from the prior stimulus to be combined with the current stimulus to inform decision making. Although their modeling work suggests that such carryover occurs for a relatively small subset of the trials (20-30%), the details of the model are similar to a descriptive account that we have offered of the consistent cross-trial pattern within the context of flex-ible lexical processor (Balota et al., 2016). Specifically, participants become tuned to the most relevant aspects of a given stimulus (e.g., the abstract dimensions of nounness and verbness) and prepare to process the same attributes on the next trial. If other salient attributes (e.g., SQ) change across the trials this priming of attributes is eliminated.

Of course, this mechanism is *post-hoc*, descriptive only, and embodies a host of assumptions which will need to be explicitly tested in future experimentation. One approach would be to use a model of binary choice to better understand these processing constraints. For example, application of the diffusion model (Ratcliff & McKoon, 2008) might be able to determine whether the cross-trial effect is in drift rate, response threshold, or non-decision time. Of course, changes in specific parameters could help adjudicate amongst the models described above. However, such a detailed computational modeling approach is clearly beyond the scope of this brief report. The clear strength of the current series of studies is in extending the four-way interaction to a new lexical task (syntactic decision) where familiarity is not so strongly emphasized and also to a more general short-term memory task.

In conclusion, the present results are consistent with accumulating evidence indicating that the processing system is constantly adjusting on a given trial, based on the attributes of the prior trial. Hence, a snapshot of performance at the trial level does not simply reflect stable architectural constraints of the processing system but a dynamically changing processing system that is being influenced by recent history. How (and how much) the system is changing across time is a critical issue for our understanding of task performance.

Appendix A
Accuracy from Experiment 1

Both the main effects of current target ($\beta = -0.585$, $z = -4.96$, $p < 0.0001$) and SQ ($\beta = 0.153$, $z = 2.97$, $p = 0.003$) were reliable. In addition, the four-way interaction was also highly reliable ($\beta = 2.13$, $z = 6.27$, $p < 0.0001$). Individual cell means are displayed in Appendix Table 1. Most importantly, it is clear that overall accuracy is higher when the attributes completely repeat across trials or completely alternate relative to only partial repetition.

Table 1 Estimated accuracy means (95% confidence intervals) from Experiment 1

Current trial	Previous target = noun		Previous target = verb	
	Previous target clear	Previous target degraded	Previous target clear	Previous target degraded
Noun targets				
Clear	0.961 (0.950-0.970)	0.956 (0.944-0.966)	0.954 (0.941-0.964)	0.96 (0.948-0.969)
Degraded	0.958 (0.946-0.968)	0.963 (0.953-0.972)	0.966 (0.956-0.974)	0.948 (0.934-0.959)
Verb targets				
Clear	0.904 (0.881-0.923)	0.918 (0.897-0.934)	0.937 (0.920-0.951)	0.913 (0.891-0.931)
Degraded	0.943 (0.927-0.955)	0.918 (0.896-0.934)	0.931 (0.913-0.946)	0.951 (0.97-0.962)

Accuracy from Experiment 2

The main effects of current target (β = 0.618, z = 7.37, *p* < 0.0001) and previous target (β = −0.187, z = −4.50, *p* <

0.0001) were reliable. The four-way interaction was again reliable (β = 1.364, z = 4.56, *p* < 0.0001). The means are presented in Appendix Table 2.

Table 2 Estimated accuracy means (95% confidence intervals) from Experiment 2

Current trial	Previous target = present		Previous target = absent	
	Previous target clear	Previous target degraded	Previous target clear	Previous target degraded
Present targets				
Clear	0.958 (0.947-0.967)	0.956 (0.945-0.965)	0.939 (0.925-0.950)	0.953 (0.941-0.962)
Degraded	0.952 (0.941-0.962)	0.955 (0.943-0.964)	0.950 (0.937-0.959)	0.938 (0.924-0.950)
Absent targets				
Clear	0.970 (0.962-0.976)	0.974 (0.967-0.980)	0.969 (0.961-0.976)	0.965 (0.956-0.972)
Degraded	0.980 (0.974-0.985)	0.974 (0.967-0.980)	0.970 (0.962-0.977)	0.976 (0.969-0.982)

References

Annis, J., & Malmberg, K. J. (2013). A model of positive sequential dependencies in judgments of frequency. *Journal of Mathematical Psychology, 57*(5), 225–236. doi:10.1016/j.jmp.2013.06.006

Aschenbrenner, A. J., & Balota, D. A. (2015). Interactive effects of working memory and trial history on Stroop interference in cognitively healthy aging. *Psychology and Aging, 30,* 1–8. doi:10.1037/pag0000012

Balota, D. A., Aschenbrenner, A. J., & Yap, M. J. (2013). Additive effects of word frequency and stimulus quality: The influence of trial history and data transformations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(5), 1563–1571. doi:10.1037/a0032186

Balota, D. A., Aschenbrenner, A. J., & Yap, M. J. (2016). Dynamic adjustment of lexical processing in the lexical decision task: Cross-trial sequence effects. *Quarterly Journal of Experimental Psychology.* doi:10.1080/17470218.2016.1240814

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10*(3), 340–357. doi:10.1037/0096-1523.10.3.340

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

Besner, D. (1983). Basic decoding components in reading: Two dissociable feature extraction processes. *Canadian Journal of Psychology, 37*(3), 429–438. doi:10.1037/h0080739

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624–652. doi:10.1037//0033-295X.I08.3.624

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30,* 412–431. doi:10.1016/0001-6918(69)90065-1 (original work published in 1868)

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), 1117–1128. doi:10.1037/a0026948

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin, 125,* 777–799. doi:10.1037/0033-2909.125.6.777

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

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General, 121,* 480–506. doi:10.1037/0096-3445.121.4.480

Jones, M., Curran, T., Mozer, M. C., & Wilder, M. H. (2013). Sequential effects in response time reveal learning mechanisms and event representations. *Psychological Review, 120*(3), 628–666. doi:10.1037/a0033180

Kuznetsova, A., Brockhott, P. B., & Christensen, R. H. B. (2015). lmerTest: Tests in Linear Mixed Effects Models. *R Package Version.* https://doi.org/http://CRAN.R-project.org/package=lmerTest

Lima, S. D., & Huntsman, L. A. (1997). Sequential dependencies in the lexical decision task. *Psychological Research, 60*(4), 264–269. doi:10.1007/BF00419412

Little, D. R., Wang, T., & Nosofsky, R. M. (2016). Sequence-sensitive exemplar and decision-bound accounts of speeded-classification performance in a modified Garner-tasks paradigm. *Cognitive Psychology, 89,* 1–38. doi:10.1016/j.cogpsych.2016.07.001

Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition, 7,* 166–174. doi:10.3758/BF03197535

Masson, M. E. J., & Kliegl, R. (2013). Modulation of additive and interactive effects in lexical decision by trial history. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 898–914. doi:10.1037/a0029180

Masson, M. E. J., Rabe, M. M., & Kliegl, R. (2017). Modulation of additive and interactive effects by trial history revisited. *Memory & Cognition, 45,* 480–492. doi:10.3758/s13421-016-0666-z

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(6), 1003–1019. doi:10.1037/0278-7393.15.6.1003

Perea, M., & Carreiras, M. (2003). Sequential effects in the lexical decision task: The role of the item frequency of the previous trial. *The Quarterly Journal of Experimental Psychology, 56A,* 385–401.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873–922. doi:10.1162/neco.2008.12-06-420

Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction time experiments. *American Scientist, 57,* 421–457.

Taylor, T. E., & Lupker, S. J. (2001). Sequential effects in naming: A time-criterion account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 117–38. doi:10.1037//0278-7393.27.1.117

Turner, B. M., Van Zandt, T., & Brown, S. D. (2011). A dynamic stimulus-driven model of signal detection. *Psychological Review, 118*(4), 583–613. doi:10.1037/a0025191

Yap, M. J., & Balota, D. A. (2015). Visual word recognition. In A. Pollastek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 26–43). New York: Oxford University Press.

Yap, M. J., & Pexman, P. M. (2016). Semantic richness effects in syntactic classification: The role of feedback. *Frontiers in Psychology, 7.* doi:10.3389/fpsyg.2016.01394