



**HAL**  
open science

## Effects of non-adjacent letter repetition in the same-different matching task

Jonathan Mirault, Stéphanie Massol, Jonathan Grainger

► **To cite this version:**

Jonathan Mirault, Stéphanie Massol, Jonathan Grainger. Effects of non-adjacent letter repetition in the same-different matching task. *Language, Cognition and Neuroscience*, 2022, 37 (10), pp.1292-1302. 10.1080/23273798.2022.2061022 . hal-04582830

**HAL Id: hal-04582830**

**<https://hal.science/hal-04582830v1>**

Submitted on 22 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Effects of non-adjacent letter repetition  
in the same-different matching task

Jonathan Mirault<sup>1</sup>, Stéphanie Massol<sup>2</sup>, Jonathan Grainger<sup>1</sup>

<sup>1</sup> Aix-Marseille University & Centre National de la Recherche Scientifique,  
Laboratoire de Psychologie Cognitive, Marseille, France

<sup>2</sup> Université de Lyon2,  
Laboratoire d'Études des Mécanismes Cognitifs, Bron, France

Short title: Letter repetition effects

**Corresponding author:**

Jonathan Grainger

Laboratoire de Psychologie Cognitive

Aix-Marseille Université

3 place Victor Hugo

13331 Marseille, France

[i.jonathan.grainger@gmail.com](mailto:i.jonathan.grainger@gmail.com) / [jonathan.grainger@univ-amu.fr](mailto:jonathan.grainger@univ-amu.fr)

**Abstract**

Effects of non-adjacent character repetition were investigated using the same-different matching task. We compared matching of strings of 5 consonants with matching of 5 symbols in order to distinguish letter-specific mechanisms from generic order encoding mechanisms. Character repetition was found to facilitate “same” responses to both types of stimuli, but crucially the interfering effect of repetition on “different” responses was only significant with letter strings. However, the latter effect only emerged in response times (as well as error rates) when the repetition occurred in the target string as opposed to the reference string. We conclude that inhibitory effects of letter repetition reflect the operation of a letter-specific order encoding mechanism, and that letter repetition in the target string has a greater influence than letter repetition in the reference string when making “different” responses in the same-different matching task.

**Keywords:** Reading; Orthographic processing; Letter order encoding; Same-different matching

## Introduction

Kerr et al. (2021) made an interesting informal observation that they thought might shed light on basic mechanisms involved in orthographic processing during reading. That is, that detecting the spelling mistake indicated by a red underline in MS Word is particularly difficult when that mistake is caused by the repetition of a letter that is already in the word (e.g., silencne). This observation attracted the interest of Kerr and colleagues because the open-bigram model of orthographic processing (Grainger & van Heuven, 2004; Whitney, 2001) provides a principled mechanism for accounting for this perceived difficulty. The core mechanism in this model is the mechanism used to encode for location-invariant letter order via an unordered set of ordered contiguous and non-contiguous letter pairs referred to as “open-bigrams” (e.g., word = od, wd, or, wo, wr, rd). According to this coding scheme, there is only one open-bigram in the nonword “silencne” that is incompatible with the incorrectly written word “silence” – that is the bigram “cn”. On the other hand, if the typographical error is caused by the insertion of a letter that is not already present in the targeted word (e.g., silencre), then the number of incompatible open-bigrams is greater (er, nr, cr, re), hence making it easier to detect the error (here we apply the parameters of the Grainger & van Heuven, 2004, model in these calculations).

Although letter repetition effects have been investigated using various paradigms in the past (e.g., Bjork & Murray, 1977; Gomez, Ratcliff, & Perea, 2008; Kanwisher, 1991; Mozer, 1987; Schoonbaert & Grainger, 2004), two studies in particular have provided key evidence in favor of an inhibitory impact of non-adjacent<sup>1</sup> letter repetition on visual word recognition. First, Trifonova and Adelman (2019) performed a regression analysis on several

---

<sup>1</sup> Following Kerr et al. (2021), we draw an important distinction between adjacent and non-adjacent letter repetitions, given the evidence for a special status for adjacent repetitions (Caramazza & Miceli, 1990; Fischer-Baum, 2017).

mega-studies of lexical decision and word naming (Balota et al., 2007; Brysbaert et al., 2016; Ferrand et al., 2010). They found a small but significant inhibitory influence of repeated letters when the repetition did not involve adjacent letters. The effects were modulated by the distance separating the repeated letters, being strongest with 1-3 intervening letters. Prompted by the findings of Trifonova and Adelman (2019) and the informal observation noted above, Kerr et al. (2021) examined the effects of non-adjacent letter repetition on responses to nonword stimuli in a lexical decision task. They tested two types of nonword generated from the same baseword: 1) nonwords created by repeating one of the letters in the baseword (e.g., silence => silencne); and 2) nonwords created by inserting a letter that is not present in the baseword (e.g., silencre). According to open-bigram coding, nonwords created by repeating a letter are more similar to their baseword than nonwords created by inserting a letter, and this should make it harder to reject letter repetition nonwords than letter insertion nonwords. This is exactly the pattern that was found by Kerr et al. (2021).

In the present study we extend the investigation of non-adjacent letter repetition effects to the same-different matching task. In this task, a first stimulus, the reference, is briefly presented, and immediately followed by a second stimulus, the target, and participants have to judge as rapidly and as accurately as possible if the two stimuli are the same or different. Reference and target locations differ slightly in order to avoid purely physical overlap in the “same” response condition. When used with character strings, the characters have to be in the same order to be judged identical. One key finding here is that it is harder to classify two strings of consonants as being different when the difference is caused by transposing two letters compared with substituting two letters with different letters (Duñabeitia et al., 2012; Massol et al., 2013; see Grainger & Hannagan, 2014, for a review). The fact that these transposition effects have found to be greater for letter stimuli compared

with other kinds of character strings<sup>2</sup> has been taken as evidence that same-different judgements on strings of stimuli reveal processing at the level of a location-invariant order encoding mechanism that differs depending on the stimuli being tested.

More directly related to the present work is the study of Gomez et al. (2008) using a variant of the same-different matching task – the match-to-sample task – in which a target stimulus is briefly presented and immediately followed by two stimuli (one to the left the other to the right), one of which is the target and the other a foil. Participants have to select the target by pressing the appropriate response key (left or right). In their Experiment 4, Gomez et al. (2008) found that accuracy in selecting the target dropped when the target contained a letter repetition compared with targets that did not contain a repeated letter. The most relevant comparison (out of a total of 37 conditions - see Table 6 in Gomez et al., 2008) with respect to the present work is the following contrast (foil – target – accuracy): ABCDE – ABCBE (.692) vs. ABCXE – ABCDE (.762).<sup>3</sup> Thus, non-adjacent repetition of internal letters would appear to impair performance in the match-to-sample task.

In light of this prior research, we therefore hypothesized that same-different judgments would be sensitive to letter repetition effects. According to open-bigram theory (Grainger & van Heuven, 2004; Whitney, 2001), letter repetitions should be harder to detect in different responses (e.g., KGLBS – KGLGS) compared with a different letter substitution condition (e.g., KGLBS – KGLTS) because there are more incompatible bigrams in the substitution condition. On the contrary, same responses to letter stimuli with repetitions (e.g.,

---

<sup>2</sup> We note that the most systematic difference in the size of character transposition effects in same-different matching has been seen in the contrast between letter stimuli and keyboard symbols, hence the choice of these stimuli in the present study.

<sup>3</sup> Gomez et al. (2008) tested pronounceable pseudowords. The ABC notation they employed is used to indicate the positions of the letters involved in the difference between foil and target. Accuracy is expressed as probability correct.

KGLGS – KGLGS) should be easier to make than same responses to stimuli without repetitions (e.g., KGLTS – KGLTS), because there are fewer bigrams to match in the former. Experiment 1 was designed to test these predictions. In comparison to the prior work of Gomez et al. (2008), here we were able to test the opposing predictions of open-bigram coding on “same” and “different” responses while measuring RTs as well as accuracy. Furthermore, the use of consonant strings allowed us to rule out any contribution from phonology and facilitates the comparison with the symbol strings tested in the follow-up experiments.

However, as acknowledged by Kerr et al. (2021), noisy slot-coding models, such as the overlap model (Gomez et al., 2008), can also account for letter repetition effects. Nevertheless, one important difference between open-bigram coding and noisy slot-coding is that the proponents of the latter claim that the same mechanism should apply to strings of all kinds of characters. We therefore compared character repetition effects in same-different matching obtained with strings of 5 consonants and strings of 5 symbol stimuli (e.g., % £ ? £ &). This was done using a between-participant design in Experiment 1 (i.e., different participants tested with letters and symbols) and a within-participant design in Experiments 2 and 3.

## **Experiment 1**

### **Methods**

#### *Participants*

We recruited 510 participants (255 per group) via the Prolific platform (Palan & Schitter, 2018). Prior to analysis, we excluded 64 participants whose first language was not written with the Roman alphabet. The remaining sample was composed of 191 participants (108 males) who were more than 18 years old ( $M = 26.80$ ;  $SD = 8.31$ ). Prior the beginning of the experiment, participants were informed that data would be collected anonymously. A power analysis (Green & MacLeod, 2016) indicated we achieved a power of 100% (95% CI = 3.62) with this sample size.<sup>4</sup> Prior to the beginning of the experiment, participants were informed that their data would be collected anonymously.

### *Design & Stimuli*

A same-different matching task was used with the presentation of a reference sequence followed by a target sequence, and participants had to decide if the two were identical or not, that is, formed of the same characters in the same order. The character sequences were either letters or symbols, and this Stimulus Type factor was manipulated between participants. We first selected the letter stimuli which were formed of two sets of 40 reference sequences all formed of 5 consonants drawn from the following set: G, N, D, K, F, T, S, B, and L. The first set of reference sequences contained no repeated letter (e.g., KGLBS). From this set we constructed 2 types of different target sequence by replacing the letter at position 4 of the reference sequence with a different letter that could either be the letter at position 2 of the reference sequence (repeated letter condition – e.g., KGLGS) or a letter not present in the reference sequence (substituted letter condition – e.g., KGLTS). The 40 reference sequences in the “same” response condition were formed of the target sequences of the “different” response condition and therefore either contained a repeated letter at positions 2 and 4 (e.g., KGLGS) or had no repeated letters (e.g., KGLTS). The exact same conditions were generated

---

<sup>4</sup> We ran a power analysis on the LMEs with over 100 simulations over the data for *same* responses and *different* responses, and they both reached 100% of power.



for symbol stimuli by replacing each of the 9 letters (G, N, D, K, F, T, S, B, L) with the following symbols (% , ? , ! , & , + , £ , € , \$). The symbol strings were then formed by simply substituting each letter of the letter strings with the corresponding symbol character. The different experimental conditions are summarized in Table 1. The effects of the Repetition factor (repeated character vs. substituted character) and the Stimulus Type factor (letters vs. symbols) were analyzed separately for “same” and “different” responses.

Table 1 about here

### *Apparatus*

The experiment was created using LabVanced (Finger et al., 2017) and displayed on the personal computer screen of participants.

### *Procedure*

We first informed participants that the data would be collected anonymously, and we asked participants to select their native language, their gender and their age. Then the experimental instructions were presented on the screen. After receiving 8 practice trials, the main experiment began. The first trial was preceded by a countdown of 3 seconds, then on each trial a fixation cross was presented in the middle of the screen for 300 ms, followed by the reference which was displayed 1 line above the center of the screen for 300 ms. After a gap of 100 ms the target was presented 1 line below the center of the screen and remained on the screen until participants' response (i.e., reference and target strings were separated by 1 line). Participants were requested to sit approximately 80 cm from their screen. LabVanced adapts the size and resolution of stimuli to each participant's screen, such that a viewing distance of 80cm systematically gives 3.6 degrees of visual angle for a 5-character string. At this viewing

distance the vertical separation of reference and target stimuli (center-to-center) was approximately 2.5 degrees of visual angle. Participants were instructed to give their responses using the right and left arrows of their computer keyboard, as quickly and as accurately as possible. Following each response, a short feedback dot (500 ms) was presented, in green if the response was correct or in red otherwise. The trial ended with a gap of 200 ms. Figure 1 illustrates the procedure for one trial.

Figure 1 about here

### *Analysis*

We used linear mixed-effects models (LMEs) to analyse response times (RTs) with items and participants as crossed random effects. Generalized (logistic) linear mixed-effects models (GLMEs) were used to analyse the error rates. For each analysis we retained the model with the maximum level of randomness that converged (Baayen, Davidson, & Bates, 2008; Barr, Levy, & Tily, 2013). This included by-item and by-participant random intercepts and with random slopes. The models were fitted with the `lmer` (for LMEs) and `glmer` (for GLMEs) functions from the `lme4` package (Bates et al., 2015) in the R statistical computing environment (R Core Team, 2018). We report regression coefficients ( $b$ ), standard errors (SE), and  $t$ -values (for LMEs) or  $z$ -values (for GLMEs). Fixed effects were deemed reliable if  $|t|$  or  $|z| > 1.96$  (Baayen, 2008). RTs were log-transformed ( $\log_{10}(\text{RT})$ ) prior to analysis for the purpose of normalization. All the models, in this and the following experiments, included random intercepts for the fixed effects and interaction terms, and also random slopes whenever the model converged with this addition.

### **Results**

We recorded response time (RT - the time between onset of the target sequence and participants' response) and response accuracy on each trial. After excluding RTs less than

100 ms and greater than 3000 ms (1.02%) the dataset was composed of 30,247 observations. Data for “same” responses and “different” responses were analyzed separately.

### *Response Time*

Prior to analysis, we excluded trials with incorrect responses (23.54 %). We then excluded values from +/- 2.5 standard deviations from the grand mean (2.72 %). The remaining dataset was composed of 22,493 observations. The condition means are shown in Table 2.

Table 2 about here

There was a significant effect of Repetition on “different” responses ( $b = 0.01$ ;  $SE = 0.002$ ;  $t = 6.10$ ), with longer RTs in the repetition condition compared with the substitution condition, and a significant effect of Stimulus Type ( $b = 0.02$ ;  $SE = 0.002$ ;  $t = 9.87$ ), with longer RTs to symbols strings compared with letter strings. Crucially, we observed a significant interaction between the two factors ( $b = 0.01$ ;  $SE = 0.003$ ;  $t = 3.41$ ), with a greater inhibitory effect of Repetition with letter stimuli ( $b = 0.01$ ;  $SE = 0.002$ ;  $t = 4.85$ ) compared with symbols ( $b = 0.005$ ;  $SE = 0.003$ ;  $t = 1.60$ ). There was a significant facilitatory effect of Repetition on “same” responses ( $b = 0.01$ ;  $SE = 0.002$ ;  $t = 5.03$ ) and a significant effect of Stimulus Type ( $b = 0.02$ ;  $SE = 0.002$ ;  $t = 10.02$ ). The interaction was not significant ( $b = 0.004$ ;  $SE = 0.003$ ;  $t = 1.43$ ).

### *Error rates*

The condition means for error rates are shown the Table 3. There was a significant effect of Repetition on “different” responses ( $b = 0.57$ ;  $SE = 0.06$ ;  $z = 8.43$ ), with more errors in the repetition condition than in the substitution condition, and a significant effect of Stimulus Type ( $b = 0.62$ ;  $SE = 0.03$ ;  $z = 18.47$ ), with more errors to symbol strings compared with letter strings. Crucially, there was also a significant interaction between the two factors ( $b = 0.63$ ;  $SE = 0.04$ ;  $z = 13.39$ ), with a clear inhibitory effect of Repetition on responses to letter stimuli ( $b = 0.61$ ;  $SE = 0.09$ ;  $z = 6.76$ ) contrasting with the non-significant facilitation seen with symbol stimuli ( $b = 0.07$ ;  $SE = 0.09$ ;  $z = 0.78$ ). There was a significant effect of Repetition on “same” responses ( $b = 0.24$ ;  $SE = 0.05$ ;  $z = 4.69$ ) and a significant effect of Stimulus Type ( $b = 0.09$ ;  $SE = 0.04$ ;  $z = 2.34$ ). The interaction was not significant ( $b = 0.10$ ;  $SE = 0.05$ ;  $z = 1.77$ ).

Table 3 about here

### **Discussion**

The results of Experiment 1 are in line with the predictions of the open-bigram model of letter order encoding. We predicted that the difference in two letter strings presented in a same-different matching task would be harder to detect when the difference is generated by a letter repetition than a difference generated by substituting a letter for a different letter. We also predicted that detecting that two strings are the same would be easier when they contain a repeated letter. The results revealed exactly this pattern in both RTs and error rates. Crucially, in a different group of participants tested with strings of symbol characters, there

was no significant effect of character repetition on “different” responses, although we did observe a facilitatory effect on “same” responses.

## **Experiment 2**

Experiment 2 was designed to correct for two potential problems in Experiment 1: that the diverging pattern of effects seen in “different” responses to letters and symbols was obtained with different groups of participants; and due to the particular design used in these experiments (see Table 1) a repetition in the reference was always associated with a “same” response. Experiment 2 therefore applied a within-participant design and switched the reference and target strings (i.e., the reference became the target and vice versa – see Table 4 for a description of the new conditions).

### **Methods**

#### *Participants*

Two hundred and sixty participants were recruited online. Seventy-four participants were removed prior to analysis because their reported native language did not use the roman alphabet. The remaining group was composed of 186 participants (84 female) ranging in age from 18 to 65 years ( $M = 27.49$  years;  $SD = 9.46$ ).

#### *Design & Stimuli*

These were the same as in Experiment 1 except that the reference stimuli now became targets and the targets became the references (see Table 4), and that Experiment 2 applied a within-participant manipulation of Stimulus Type (letters vs. symbols). Letters and symbols were presented in different blocks, and block order was counterbalanced across participants.

Table 4 about here

### *Procedure and Analyses*

Same as in Experiment 1.

### **Results**

From the raw data of the 186 participants, we first excluded 0.50 % of trials with RTs less than 100 ms or greater than 3,000 ms, leaving a dataset composed of 59,222 observations.

### *Response times*

Prior to analyzing the RTs, we removed trials with incorrect responses (21.50 %) and trials with values lying beyond  $\pm 2.5$  SD from the grand mean (2.62 %). Condition means are shown in Table 5.

Table 5 about here

For “different” responses, the main effect of Repetition was not significant ( $b = 0.001$ ;  $SE = 0.003$ ;  $t = 0.33$ ), and neither was the effect of Stimulus Type ( $b = 0.001$ ;  $SE = 0.003$ ;  $t = 0.54$ ). The interaction was not significant ( $b = 0.003$ ;  $SE = 0.004$ ;  $t = 0.79$ ). For “same” responses, we observed a significant effect of Repetition ( $b = 0.01$ ;  $SE = 0.002$ ;  $t = \mathbf{6.12}$ ), with faster RTs in the repetition condition ( $M = 651$ ;  $SD = 26.50$ ) compared to the substitution condition ( $M = 667$ ;  $SD = 26.72$ ). The main effect of Stimulus Type was not significant ( $b = 0.004$ ;  $SE = 0.002$ ;  $t = 1.54$ ), and neither was the interaction ( $b = 0.0003$ ;  $SE = 0.003$ ;  $t = 0.09$ ).

### *Error rates*

Condition means are shown in Table 6. For “different” responses, there was a main effect of Stimulus Type ( $b = 0.55$ ;  $SE = 0.10$ ;  $z = 5.26$ ) with more errors made to symbols ( $M = 32.02\%$ ;  $SD = 6.70$ ) compared to letters ( $M = 27.21\%$ ;  $SD = 6.39$ ). Crucially, we observed a significant interaction between Stimulus Type and Repetition ( $b = 0.63$ ;  $SE = 0.15$ ;  $z = 4.12$ ) with a significant inhibitory effect of Repetition with letter stimuli ( $b = 0.30$ ;  $SE = 0.09$ ;  $z = 3.36$ ) contrasting with the significant facilitatory effect obtained with symbols ( $b = 0.34$ ;  $SE = 0.12$ ;  $z = 2.68$ ). For “same” responses, there was a significant effect of Repetition ( $b = 0.33$ ;  $SE = 0.05$ ;  $z = 5.87$ ) with fewer errors in the repetition condition ( $M = 12.01\%$ ;  $SD = 4.67$ ) compared to the substitution condition ( $M = 14.73\%$ ;  $SD = 5.09$ ). There was also a significant effect of Stimulus Type ( $b = 0.21$ ;  $SE = 0.04$ ;  $z = 4.31$ ), with more errors being made to symbol stimuli ( $M = 14.91\%$ ;  $SD = 5.12$ ) compared to letters ( $M = 11.83\%$ ;  $SD = 4.64$ ). The interaction failed to reach significance ( $b = 0.14$ ;  $SE = 0.07$ ;  $z = 1.94$ ).

Table 6 about here

### **Discussion**

Experiment 2 tested for repetition effects in letters and symbols using a within-participant design, and without the potential response bias that was present in Experiment 1 (whenever the reference sequence had a repeated letter / symbol the response was always “same”). We found basically the same pattern as in Experiment 1, but this time the critical Stimulus Type X Repetition interaction in “different” responses was only found in error rates, whereas it was found in RTs and error rates in Experiment 1. It is important to note that the RTs showed no

sign of a speed-accuracy trade-off in Experiment 2, and the interaction in error rates followed the same pattern as Experiment 1, with an inhibitory repetition effect of 5.5% obtained with letter stimuli and a facilitatory effect of 6.6% with symbols.

In an attempt to understand why we failed to find the critical Stimulus Type X Repetition interaction in “different” response RTs in Experiment 2, we performed a number of post-hoc analyses. We first examined if this might depend on speed of responding, with faster RTs on a given trial and/or overall faster responding per participant causing the repetition effect to emerge mainly in error rates. We therefore analyzed the distribution of the RT effects (delta plots for the difference between the repetition and substitution conditions for “different” responses) as well as median-split analyses based on all trials and also based on mean RT by participant. None of these analyses revealed an impact of speed of responding on repetition effects. Finally, we examined whether block order might have impacted on the results of Experiment 2, with repetition effects emerging in RTs when letters were seen first. This was not the case. We tentatively conclude that it was the shift from repetition in the target string (Experiment 1) to repetition in the reference string (Experiment 2) that was the most likely cause of the change in results. Experiment 3 therefore manipulated character repetition in the “different” response condition in the target string while using a within-participant manipulation of Stimulus Type and inserting filler trials in order to counteract the potential response bias that might have been present in Experiment 1.

### **Experiment 3**

Experiment 3 implemented a different means for countering the potential response bias in Experiment 1 (i.e., that a repetition in the reference string was always associated with a



“same” response). To do so we added filler trials (using different combinations of the same set of letters and symbols as used in the critical trials) where there was a repetition in the reference string associated with a “different” response.

## **Methods**

### *Participants*

Two hundred participants were recruited online. Forty-nine participants were removed prior to analysis because their reported native language did not use the roman alphabet. The remaining group was composed of 151 participants (70 female) ranging in age from 19 to 60 years ( $M = 30.89$  years;  $SD = 9.12$ ).

### *Design & Stimuli*

Stimuli were the same as in Experiment 1 except for the addition of the following filler stimuli.

For each block of trials (letters / symbols) we added 40 filler trials where a repetition in the reference string was associated with a “different” response (e.g., NGBGF – NGBTF). Furthermore, in order to have the same number of “different” and “same” response trials we added 40 filler trials associated with a “same” response and with no character repetition (e.g., NBKLF – NBKLF). The design was the same as Experiment 2.

### *Procedure and Analyses*

Same as the previous experiments.

## **Results**

From the raw data of the 151 participants, we first excluded 0.76 % of trials with RTs less than 100 ms or greater than 3,000 ms, leaving a dataset composed of 59,938 observations.

### *Response times*

Prior to analyzing the RTs, we removed trials with incorrect responses (32.36 %) and trials with values lying beyond  $\pm 2.5$  SD from the grand mean (2.96 %). Condition means are shown in Figure 2.

Figure 2 about here

For “different” responses, there was a main effect of Repetition ( $b = 0.01$ ;  $SE = 0.002$ ;  $t = 5.49$ ), and no effect of Stimulus Type ( $b = 0.0004$ ;  $SE = 0.002$ ;  $t = 0.13$ ). Crucially, the Stimulus Type X Repetition interaction was significant ( $b = 0.004$ ;  $SE = 0.004$ ;  $t = 2.68$ ). Follow-up analyses revealed a significant inhibitory effect of Repetition for letter stimuli ( $b = 0.01$ ;  $SE = 0.003$ ;  $t = 5.17$ ) and a non-significant effect for symbol stimuli ( $b = 0.004$ ;  $SE = 0.004$ ;  $t = 1.28$ ). For “same” responses, we observed a significant effect of Repetition ( $b = 0.01$ ;  $SE = 0.002$ ;  $t = 5.69$ ), with faster RTs in the repetition condition compared to the substitution condition. There was also a main effect of Stimulus Type ( $b = 0.007$ ;  $SE = 0.002$ ;  $t = 3.05$ ) with faster RTs to letter strings compared to the symbol strings. The interaction was not significant ( $b = 0.0003$ ;  $SE = 0.003$ ;  $t = 0.09$ ).

### *Error rates*

Condition means are shown in Figure 3. For “different” responses, there was a main effect of Repetition ( $b = 0.58$ ;  $SE = 0.09$ ;  $z = 6.20$ ), with more errors in the repetition than the substitution condition, and a main effect of Stimulus Type ( $b = 0.54$ ;  $SE = 0.09$ ;  $z = 5.79$ )

with more errors made to symbols compared to letters. Crucially, we observed a significant interaction between Stimulus Type and Repetition ( $b = 0.58$ ;  $SE = 0.13$ ;  $z = 4.39$ ) with a significant inhibitory effect of Repetition with letter stimuli ( $b = 0.59$ ;  $SE = 0.07$ ;  $z = 7.58$ ) contrasting with a non-significant effect obtained with symbols ( $b = 0.001$ ;  $SE = 0.11$ ;  $z = 0.98$ ). For “same” responses, there was a significant effect of Repetition ( $b = 0.12$ ;  $SE = 0.06$ ;  $z = 2.07$ ) with fewer errors in the repetition condition compared to the substitution condition. There was also a significant effect of Stimulus Type ( $b = 0.22$ ;  $SE = 0.05$ ;  $z = 3.89$ ), with more errors being made to symbol stimuli compared to letters. The interaction was not significant ( $b = 0.06$ ;  $SE = 0.08$ ;  $z = 0.73$ ).

Figure 3 about here

## Discussion

Experiment 3 resorted to a manipulation of character repetition in the target string, as in Experiment 1, rather than in the reference string (Experiment 2) while maintaining the within-participant manipulation of Stimulus Type as in Experiment 2 and controlling for potential response biases by adding filler stimuli. The results of Experiment 3 perfectly replicate those of Experiment 1, with significant interactions between Stimulus Type and Repetition in both RTs and error rates to “different” responses (see Figures 2 & 3), and no significant interactions in “same” responses. We can therefore conclude that in the same-different matching task, the effects of letter repetition on “different” responses are stronger (observed in both RTs and error rates) when the repetition occurs in the target string rather than the reference string.

## General Discussion

In three same-different matching experiments we examined the reading-specific nature of non-adjacent letter repetition effects observed in prior experiments (e.g., Kerr et al., 2021; Gomez et al., 2008; Trifonova & Adelman, 2019). All three experiments examined character repetition effects in same-different matching in consonant strings and symbols strings. Based on the interpretation of the letter repetition effects reported in Kerr et al. (2021), we expected to observe inhibitory effects of repetition in “different” responses, and, on the contrary, a facilitatory impact on “same” responses. That is, we predicted that it would be harder (more errors and longer RTs) to judge that the reference-target pair KGLBS-KGLGS are different letter sequences compared with KGLBS-KGLTS, where the difference is created by inserting a different letter rather than repeating a letter. We also predicted that it would be easier to determine that target and reference are the same when they contain a repeated letter (e.g., KGLGS-KGLGS) compared with letter sequences without a repetition (e.g., KGLTS-KGLTS). Crucially, according to our preferred interpretation of letter repetition effects (i.e., open-bigram coding), these effects are thought to be primarily driven by reading-specific mechanisms. We therefore predicted that the inhibitory effect of character repetition would only be found for letter stimuli, and not for symbol stimuli. On the other hand, noisy generic order-encoding models, such as the overlap model (Gomez et al., 2008), predicted that similar patterns of character repetition effects should be seen with letters and symbols, since these effects are thought to be driven by a generic order-encoding mechanism that applies noise to positional information for any kind of object sequence. The present study tested these predictions with strings of 5 consonants and strings of 5 keyboard symbols.

The results of all three experiments provide support for our account of inhibitory effects of non-adjacent letter repetition as revealing a letter-specific order encoding mechanism (Kerr et al., 2021). These inhibitory effects only emerged with letter strings and were not present in symbol strings, and the crucial interaction between Stimulus Type and Repetition was significant in both RTs and error rates in Experiments 1 and 3, and in the error rates in Experiment 2. On the other hand, the predicted facilitatory effect of repetition on “same” responses with letter stimuli was also found with symbol stimuli in all Experiments. What might be driving the facilitatory repetition effect seen in “same” responses to symbol sequences? We suspect that it might be due to the fact that there is less information to compare with repeated character sequences. More precisely, when the reference and the target sequence contain a repetition (e.g., % £ ? £ &), then the matching procedure is simplified by the fact that fewer distinct symbols are involved. Of course, this might also be the mechanism driving repetition effects on “same” responses to letter strings. Therefore, the key result of the present study is the distinct pattern of repetition effects seen for letters and symbols when making “different” decisions in the same-different matching task.

Thus, the main conclusion to be drawn from the present results is that the inhibitory effects of non-adjacent letter repetition reported by Trifovona and Adelman (2019) and Kerr et al. (2021) are driven by a reading-specific order-encoding mechanism, and not a generic order-encoding mechanism such as implemented in the overlap model (Gomez et al., 2008). Our results also allow us to rule-out an interpretation of prior observations of letter-repetition effects as some form of generic spatial repetition blindness (e.g., Kanwisher, 1991; Luo & Caramazza, 1996), as discussed in Kerr et al. (2021). The results obtained with letter stimuli

in Experiment 1 were predicted by the open-bigram coding account of non-adjacent letter repetition effects described in Kerr et al. This account accurately predicted an inhibitory effect of letter repetitions on “different” judgments, and a facilitatory effect on “same” judgments. Crucially, the present results show that these were not trivial predictions that would apply to the process of making same-different judgments to any kind of character sequence.

Our account of the letter-specific inhibitory effect of character repetition in same-different matching appeals to mechanisms used to encode for identity and order information in strings of characters that are hypothesized to be processed in parallel and to involve a location-invariant encoding of the ordering of character identities (e.g., Ratcliff, 1981; Gomez et al., 2008). In prior work we have hypothesized that the processing involved in same-different matching is quite different from the processing involved in tasks where participants have to search for, or identify, a single character in a string of characters. However, the results of studies using such character-in-string search tasks have shown that the serial position functions (i.e., differences in performance as a function of the target character’s position in the string) differ for letter stimuli compared with non-letter stimuli (e.g., Mason, 1982; Schmitt & Lachmann, 2020). This has also been shown for post-cued character-in-string identification when contrasting letter stimuli with symbol stimuli as in the present study (e.g., Tydgate & Grainger, 2009). Crucially, however, in post-cued target-in-string identification tasks, although the pattern of results seen with letter stimuli is distinct from that seen with symbol stimuli, it aligns with that of digit stimuli (e.g., Tydgate & Grainger, 2009), which is not the case in same-different matching (e.g., Duñabeitia et al., 2012; Massol et al., 2013). This has been taken as key evidence in favor of distinct levels of processing involved in single target search/identification in strings of characters compared

with the matching of whole strings of characters (Grainger, 2018; Grainger & Hannagan, 2014). In order to account for this diverging pattern of findings across tasks, we have hypothesized that tasks that involve detecting a single target character in a string of characters are based on activity in gaze-centered representations, where the location of a given character identity is specified relative to eye fixation location. On the other hand, the same-different judgment task would involve processing at the level of location-invariant order information, where the order of character identities is processed independently of eye-fixation location (see Massol & Grainger, 2022, for a recent description of this theoretical framework).

Finally, our results point to the importance of distinguishing between manipulations made in the reference string as opposed to the target string in the “different” response condition of same-different matching. Our non-adjacent letter repetition manipulation clearly had a larger impact on same-different judgments when the manipulation occurred in the target string (Experiments 1 & 3) compared with the reference string (Experiment 2). Although letter repetition effects were significant in error rates in Experiment 2, contrary to Experiments 1 and 3 they were not significant in RTs. This might well be due to the fact that in the procedure used in the present study target strings remained on screen until participants’ response, whereas reference strings were presented for 300 ms. Only future research in which the durations of the reference and target strings are equated (or manipulated) will enable a test of this tentative explanation for the different pattern seen in RTs across experiments in the present study.

## **Conclusions**

We set-out to test the predictions derived from the open-bigram coding account of non-adjacent letter repetition effects reported in Kerr et al. (2021). We predicted opposite effects on “same” vs. “different” decisions in the same-different matching task, with facilitatory effects of repetition on “same” responses and inhibitory effects on “different” responses, and we indeed found that pattern. Crucially, by using the same-different matching task in the present study we were able to examine the letter-specific nature of these effects, as predicted by open-bigram coding, by testing sequences of symbol stimuli under the exact same conditions as used with letter stimuli. We found clear evidence for the letter-specific nature of the inhibitory effects of repetition on “different” responses, while symbol stimuli showed the same facilitatory effect on “same” responses as seen with letter stimuli. Non-adjacent letter repetition effects are therefore at least partly reading-specific in nature.



**Acknowledgments**

This research was supported by ERC grant 742141

**Open Practices**

All scripts for data analysis and the raw data are available at: <https://osf.io/h29jv/>

**Pre-registration**

The letter block of Experiment 1 was pre-registered on OSF the 27th of February, 2021. See: <https://osf.io/dv8hk>

The symbol block of Experiment 1 was subsequently added to complete this experiment. Experiments 2 & 3 were designed on the basis of the results obtained in Experiment 1.

## References

- Baayen, R. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.
- Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., ... & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*, 445-459.
- 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*,
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1-48.
- Bjork, E. L. & Murray, J. T. (1977). On the nature of input channels in visual processing. *Psychological Review*, *84*, 472-484.
- 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*, 441.
- Caramazza, A., & Miceli, G. (1990). The structure of graphemic representations. *Cognition*, *27*, 243-297.
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*, 42-45.
- Duñabeitia, J.A., Dimitropoulou, M., Grainger, J., Hernández, J.A., & Carreiras, M. (2012). Differential sensitivity of letters, numbers and symbols to character transpositions. *Journal of Cognitive Neuroscience*, *24*, 1610-1624.
- Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., ... & Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, *42*, 488-496.
- Finger, H., Goeke, C., Diekamp, D., Standvoß, K., and König, P. (2017). "LabVanced: a unified JavaScript framework for online studies," in *International Conference on Computational Social Science IC2S2* (Cologne).
- Fischer-Baum, S. (2017). The independence of letter identity and letter doubling in reading. *Psychonic Bulletin & Review*, *24*, 873-878.
- Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*, 577-600.
- Grainger, J. (2018). Orthographic processing: A "mid-level" vision of reading. *Quarterly Journal of Experimental Psychology*, *71*, 335-359.

- Grainger, J. & Hannagan, T. (2014). What is special about orthographic processing? *Written Language & Literacy*, 17, 225-252.
- Grainger, J. & van Heuven, W. J. B. (2004). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *Mental Lexicon*. New York: Nova Science Publishers.
- Green, P. & MacLeod, C. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7, 493-498.
- Kanwisher, N. (1991). Repetition blindness and illusory conjunctions: Errors in binding visual types with visual tokens. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 404-421.
- Kerr, E., Mirault, J., & Grainger, J. (2021). On non-adjacent letter repetition and orthographic processing: Lexical decisions to nonwords created by repeating or inserting letters in words. *Psychonomic Bulletin & Review*, 28, 596-609.
- Luo, C. R. & Caramazza, A. (1996). Temporal and spatial repetition blindness: Effects of presentation mode and repetition lag on the perception of repeated items. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 95-113.
- Mason, M. (1982). Recognition time for letters and nonletters: Effects of serial position, array size, and processing order. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 724-738.
- Massol, S., Duñabeitia, J.A., Carreiras, M., & Grainger, J. (2013). Evidence for letter-specific position coding mechanisms. *PLoS ONE*, 8(7).
- Massol, S. & Grainger, J. (2022). *Effects of horizontal displacement and inter-character spacing on transposed-character effects in same-different matching*. Manuscript submitted for publication.
- Mozer, M. C. (1987). *Early parallel processing in reading: A connectionist approach*. Lawrence Erlbaum Associates, Inc.
- Palan, S., & Schitter, C. (2018). Prolific. A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 22-27.
- Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, 88, 552.
- R Core Team. (2018). *R: A Language and Environment for Statistical Computing*.
- Schmitt, A., & Lachmann, T. (2020). Skilled readers show different serial-position effects for letter versus non-letter target detection in mixed-material strings. *Acta Psychologica*, 204, 103025.

Schoonbaert, S. & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes, 19*, 333–367.

Tydgat, I. & Grainger, J. (2009). Serial position effects in the identification of letters, digits, and symbols. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 480-498.

Trifonova, I.V. & Adelman, J.S. (2019). A delay in processing for repeated letters: Evidence from megastudies. *Cognition, 189*, 227-241.

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review 8*, 221–243.