Effects of masked prime duration in an online lexical decision task

Bernhard Angele[1,2], Ana Baciero[2,3], Pablo Gomez[3,4], & Manuel Perea Lara[2,3]

[1] Bournemouth University

[2] Universitat de Valencia, Valencia, Spain

[3] Nebrija University, Madrid, Spain

[4] California State University San Bernardino, Palm Desert Campus

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

Correspondence concerning this article should be addressed to Bernhard Angele, Department of Psychology, Faculty of Science and Technology, Talbot Campus, Fern Barrow, Poole BH12 5BB, UK. E-mail: bangele@bournemouth.ac.uk

Abstract

 Masked priming is one of the most important paradigms in the study of visual word recognition, but it is usually thought to require a laboratory setup with a known monitor and keyboard. To investigate if this technique can be used in an online setting, we conducted two online masked priming lexical decision task experiments using PsychoPy/PsychoJS (Peirce et al., 2019). In particular, we wanted to compare our online results to the data collected by Gomez, Perea, and Ratcliff (2013), who compared masked and unmasked priming. Furthermore, we also tested the role of prime exposure duration effectively in an online experiment (33 vs. 50 ms in Experiment 1 and 16 vs. 33 ms in Experiment 2). We found that our online data are indeed very similar to the masked priming data reported by Gomez, Perea, and Ratcliff (2013). Additionally, we found a clear effect of prime duration, with the priming effect (measured in terms of response time and accuracy) being stronger at 50 ms than 33 ms and no priming effect at 16 ms prime duration. From these results, we can conclude that modern online browser-based experimental psychophysics packages (e.g., Psychopy) can present stimuli and collect responses on standard consumer devices with enough precision. In sum, these findings provide us with confidence that masked priming can be used online, thus allowing us to reach populations that are hard to test in a laboratory.

Keywords: Masked priming, Lexical decision task, Online experiments, PsychoPy, Prime duration

Word count: X

Effects of masked prime duration in an online lexical decision task

Masked priming (Forster & Davis, 1984) is one of the most important paradigms in the study of visual word recognition. Priming refers to the observation that the response to a target stimulus (e.g., responding that "DOCTOR" is a word in a lexical decision task or pronouncing the same stimulus in a naming task) is faster, more accurate, or both when it was preceded by a related stimulus (e.g., "nurse") compared to an unrelated stimulus (e.g., "horse"). In masked priming, the prime stimulus ("nurse" or "horse") is presented very briefly (for less than 60 ms) and is itself preceded by a mask (e.g., "#####") which is presented fora much longer duration (typically 500 ms). This forward mask ensures that participants are usually unaware of the prime (Forster & Davis, 1984). The masked priming paradigm has been used in a large number of studies over the last decades (for a review, see e.g., Kinoshita & Lupker, 2004; Van den Bussche, Van den Noortgate, & Reynvoet, 2009), but so far, virtually all of these studies have been done in a laboratory setting, often using the DMDX software by Forster and Forster (2003).

Given that, as of the time of this writing, researchers in a large part of the world are unable to perform psychology experiments in a laboratory due to the COVID-19 pandemic, with not much clarity about when laboratories will be able to reopen completely, it is an important consideration whether it may be possible to collect masked priming data remotely using online experiments. Even after the current exceptional situation, online data collection has many advantages, such as easy access to a much more diverse population than would be accessible at the typical university research laboratory, independence from laboratory space constraints, and, often, lower costs as participants only need to be compensated for their time on the experiment, not time spent commuting, waiting for the experiment to start, etc. Researchers in decision making and economics have been using online paradigms for several decades now (e.g., Birnbaum & Birnbaum, 2000; Paolacci, Chandler, & Ipeirotis, 2010), but cognitive and experimental psychology have been much

slower in taking up online paradigms, often due to concerns about the validity of the results. Such concerns are not limited to cognitive studies (Aust, Diedenhofen, Ullrich, & Musch, 2013), but they are exacerbated by the reliance on precise presentation times. In masked priming, it is critical for the onset of the mask, the prime, and the target to occur at the intended times. In particular, presenting the prime for too long will counteract the effect of the mask, making the prime consciously visible to the participant.

There have been attempts to address these concerns. For example, a Web version of DMDX (webDMDX) was developed which showed promising results in a trial experiment (Witzel, Cornelius, Witzel, Forster, & Forster, 2013). However, webDMDX was a self-contained Windows executable file that participants had to download and run rather than a "true" online experiment that could be run inside of a browser. A downside of this is that participants often are (and should be) understandably skeptical about downloading and running files from the Internet. Additionally, many participants may not have access to a Windows PC, or may be discouraged from participating by the extra work it takes to deploy the experiment on their computer. As a consequence, the use of webDMDX has been rather limited so far. Fortunately, in recent years, there have been significant improvements in how content can be presented on the World Wide Web. Most notably, the HTML5 standard now makes it possible to use Javascript in order to draw stimuli interactively and monitor participant responses with remarkable flexibility inside the browser. Participants do not have to install any software, and the HTML5 standard is supported by a wide variety of devices, including mobile phones and tablets (Reimers & Stewart, 2015). There is now a variety of software packages taking advantage of the new capabilities to present experimental stimuli and collect data, both commercial, e.g., Gorilla (A. L. Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020) or Testable (Rezlescu, Danaila, Miron, & Amariei, 2020) and open-source, e.g., jsPsych (de Leeuw, 2015) or PsychoJS (the Javascript version of Psychopy 3, Peirce et al., 2019).

Of course, despite the technological advances, there are still concerns about timing

and measurement precision. While the Javascript-based experiments run on the participants' devices and thereby avoid any lag due to connection issues (all stimulus data are usually downloaded before the start of the experiment to avoid delays), experimenters have little control over which devices the experiment is run on beyond the option of explicitly preventing the experiment to run on specific device types such as mobile devices. Moreover, experimenters have no control at all over what other applications are running on the device, screen size and resolution, viewing distance, properties of the keyboard/touchscreen, etc., as all of these are determined by the device or the participants' preferences. As Reimers and Stewart (2015) point out, there are two ways of testing whether timing and response issues are problematic: (1) comparing a Web-based experiment directly with an established lab-based version by measuring presentation timings (using a photodiode) and response timings on various device configurations and (2) attempting to replicate existing lab-based findings using a Web-based paradigm. If the results of the Web-based study are comparable to previous lab-based results, this suggests that whatever the deviations in stimulus and response timing are, they are not severe enough to affect the overall findings in the paradigm in question.

The first approach has the advantage that differences in presentation timings can be objectively recorded and evaluated. A very thorough recent example of this approach is the "timing mega-study" by Bridges, Pitiot, MacAskill, and Peirce (2020), who compared the timing in experiments run in lab-based setups with the timing in online packages run in different browsers. The very similar study by A. Anwyl-Irvine, Dalmaijer, Hodges, and Evershed (2020) compares only online packages and browsers with regard to timing. Overall, Bridges, Pitiot, MacAskill, and Peirce (2020) found that online packages were capable of presenting visual stimuli with "reasonable" precision, although the lab-based packages were slightly better in this regard. This first approach is important in order to establish that a certain level of precision and accuracy can be achieved at all. If this is not possible, there is no point in moving forward to the second approach and replicating

specific paradigms. However, it is of course impossible to test every possible device and configuration that participants might use. On the other hand, some of the differences in precision and accuracy between setups that can be observed using a photodiode may be too small to have an influence on actual participant performance. Therefore, we consider replication of previous lab-based effects a more important test of online paradigms than photodiode measurements. Based on the results by Bridges, Pitiot, MacAskill, and Peirce (2020) and A. Anwyl-Irvine, Dalmaijer, Hodges, and Evershed (2020), we are satisfied that modern Javascript-based stimulus presentation systems are capable of sufficiently fast and precise stimulus presentation. To establish whether masked priming studies can be successfully run online, the next step is now to follow the second approach and implement the masked priming paradigm in one of the online experiment packages tested by Bridges, Pitiot, MacAskill, and Peirce (2020) and A. Anwyl-Irvine, Dalmaijer, Hodges, and Evershed (2020). After evaluating their results, we decided on Psychopy/PsychoJS (Peirce et al., 2019) as it combines relative ease of use with high precision and accuracy across the great majority of platforms.

In the present study, we were particularly interested in manipulating the duration of the prime, as this is what determines whether the stimulus is consciously perceived or not. Indeed, Gomez, Perea, and Ratcliff (2013) performed an experiment in which they presented a prime either for 56 ms preceded by a mask or for 200 ms without a mask and found qualitative differences between masked and unmasked priming, which they explained using the diffusion model (Ratcliff, Gomez, & McKoon, 2004): For masked priming, the priming effects for identical compared to unrelated primes were the same across all quantiles of the response time distribution. In the diffusion model, this corresponds to a shift in a change in the T_{er} parameter, which corresponds to encoding process. However, for unmasked priming, the identity priming effects were stronger in the upper quantiles of the RT distribution than in the lower quantiles. In the diffusion model, this corresponds to a shift in both the T_{er} parameter and the drift rate, the parameter corresponding to the

quality of the information gathered. This qualitative pattern can help us interpret our results as well.

Additionally, if the prime duration can be manipulated very precisely, we may be able to see effects of small differences in prime durations. Ziegler, Ferrand, Jacobs, Rey, and Grainger (2000) found that priming effects were smaller for lower prime durations and interpreted this as evidence for different levels of facilitation (due to phonological overlap) and inhibition (due to visual differences between prime and target). In order to test this, we presented masked primes of different durations in an online experiment and measured response times and accuracy.

If it is possible to present masked primes reliably in an online experiment, our results should be similar to those observed by Gomez, Perea, and Ratcliff (2013) in their masked priming condition for word trials:

1: The priming effects should be approximately the same across quantiles in the reaction time distributions.

2: In terms of the Diffusion model,

a) allowing the drift rate parameter for words to vary as a function of the priming condition should not improve the model (compared to a model that has a single drift rate parameter for all word stimuli regardless of priming condition) , while

b) allowing the non-decision parameter T_{er} to vary as a function of the priming conditions for word stimuli should improve the model (compared to a model that has only one T_{er} parameter for all word stimuli regardless of priming condition).

If it is not possible to achieve consistent presentation times in an online experiment, and the primes are effectively presented for longer durations than intended, we should see a pattern that resembles that observed by Gomez et al. (2013) for unmasked primes:

1': The priming effects would increase in the higher quantiles in the reaction time distributions.

2': In terms of the Diffusion model, both

a) allowing the drift rate parameter for words to vary as a function of the priming condition should improve the model compared to a model that has a single drift rate parameter for all word stimuli regardless of priming condition, and

b) allowing the non-decision parameter T_{er} to vary as a function of the priming conditions for word stimuli should improve the model compared to a model that has only one T_{er} parameter for all word stimuli regardless of priming condition.

Finally, there is the (unlikely) possibility that there are no priming effects at all due to inconsistent stimulus presentation (for example, the frames with the prime are not displayed at all due to timing issues with the participant's device). In this case, we would expect to observe no difference in reaction time between the identical and the unrelated control priming condition for word stimuli.

If the online experiment succeeds in not only presenting the primes fast enough to create masked priming effects, but also in manipulating the prime duration precisely enough that this affects reaction times and accuracy, we expect to find smaller effects for shorter prime presentations, analogous to Ziegler, Ferrand, Jacobs, Rey, and Grainger (2000).

Experiment 1

In the first experiment, we tested whether we could observe reliable effects of masked priming at prime durations of 33 ms and 50 ms (roughly corresponding to two and three frames at a refresh rate of 60 Hz).

We report here how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Method

Participants.   We collected data from 77 participants aged from 18 to 71 (mean age: 31.14) recruited through Prolific (www.prolific.co, 2021). Of the participants, 41 identified as male and 36 identified as female. The experiment was only shown to participants who indicated that English was their first language in the Prolific screening questions. Based on their IP addresses, 47 participants were based in the UK, 23 were based in the US, three participants were based in Canada, and two participants were based in Ireland. Two participants could not be localized in this way. Participants who completed the experiment were paid £1.25 for their participation (corresponding to £5/hour). All participants were naïve to the purpose of the experiment.

Materials.

Procedure.

Data analysis.   We used R [Version 4.0.4; R Core Team (2020)] and the R-packages *dplyr* [Version 1.0.5; Wickham, François, Henry, and Müller (2021)], *forcats* [Version 0.5.1; Wickham (2021a)], *ggplot2* [Version 3.3.3; Wickham (2016)], *papaja* [Version 0.1.0.9942; Aust and Barth (2020)], *purrr* [Version 0.3.4; Henry and Wickham (2020)], *readr* [Version 1.4.0; Wickham and Hester (2020)], *rworldmap* [Version 1.3.6; South (2011)], *sp* [Version 1.4.5; Pebesma and Bivand (2005)], *stringr* [Version 1.4.0; Wickham (2019)], *tibble* [Version 3.1.0; Müller and Wickham (2021)], *tidyr* [Version 1.1.3; Wickham (2021b)], and *tidyverse* [Version 1.3.0; Wickham et al. (2019)] for all our analyses.

Results

Discussion

General Discussion

References

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2020). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods.* https://doi.org/10.3758/s13428-020-01501-5

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown.* Retrieved from https://github.com/crsh/papaja

Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. (2013). Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods*, *45*(2), 527–535. https://doi.org/10.3758/s13428-012-0265-2

Birnbaum, M. H., & Birnbaum, M. O. (2000). *Psychological Experiments on the Internet.* Elsevier. Retrieved from http://books.google.com?id=sw4NHyThFFkC

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, e9414. https://doi.org/10.7717/peerj.9414

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 680–698. https://doi.org/10.1037/0278-7393.10.4.680

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 116–124.

Gomez, P., Perea, M., & Ratcliff, R. (2013). A diffusion model account of masked versus unmasked priming: Are they qualitatively different? *Journal of Experimental Psychology: Human Perception and Performance*, *39*(6), 1731–1740. https://doi.org/10.1037/a0032333

Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools*. Retrieved from https://CRAN.R-project.org/package=purrr

Kinoshita, S., & Lupker, S. J. (2004). *Masked Priming: The State of the Art*. Psychology Press. Retrieved from http://books.google.com?id=VH55AgAAQBAJ

Müller, K., & Wickham, H. (2021). *Tibble: Simple data frames*. Retrieved from https://CRAN.R-project.org/package=tibble

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). *Running Experiments on Amazon Mechanical Turk* (SSRN Scholarly Paper No. ID 1626226). Rochester, NY: Social Science Research Network. Retrieved from https://papers.ssrn.com/abstract=1626226

Pebesma, E. J., & Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News*, *5*(2), 9–13. Retrieved from https://CRAN.R-project.org/doc/Rnews/

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., . . . Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Prolific. (2021). Prolific | online participant recruitment for surveys and market research. Retrieved from https://www.prolific.co/

R Core Team. (2020). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

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

Reimers, S., & Stewart, N. (2015). Presentation and response timing accuracy in Adobe Flash and HTML5/JavaScript Web experiments. *Behavior Research Methods, 47*(2), 309–327. https://doi.org/10.3758/s13428-014-0471-1

Rezlescu, C., Danaila, I., Miron, A., & Amariei, C. (2020). Chapter 13 - More time for science: Using Testable to create and share behavioral experiments faster, recruit better participants, and engage students in hands-on research. In B. L. Parkin (Ed.) (Vol. 253, pp. 243–262). Elsevier. https://doi.org/10.1016/bs.pbr.2020.06.005

South, A. (2011). Rworldmap: A new r package for mapping global data. *The R Journal, 3*(1), 35–43. Retrieved from http://journal.r-project.org/archive/2011-1/RJournal_2011-1_South.pdf

Van den Bussche, E., Van den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: A meta-analysis. *Psychological Bulletin, 135*(3), 452–477. https://doi.org/10.1037/a0015329

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string operations.* Retrieved from https://CRAN.R-project.org/package=stringr

Wickham, H. (2021a). *Forcats: Tools for working with categorical variables (factors).* Retrieved from https://CRAN.R-project.org/package=forcats

Wickham, H. (2021b). *Tidyr: Tidy messy data.* Retrieved from https://CRAN.R-project.org/package=tidyr

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *Dplyr: A grammar of data manipulation.* Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., & Hester, J. (2020). *Readr: Read rectangular text data.* Retrieved from https://CRAN.R-project.org/package=readr

Witzel, J., Cornelius, S., Witzel, N., Forster, K. I., & Forster, J. C. (2013). Testing the viability of webDMDX for masked priming experiments. *The Mental Lexicon, 8*(3), 421–449. https://doi.org/10.1075/ml.8.3.07wit

Ziegler, J. C., Ferrand, L., Jacobs, A. M., Rey, A., & Grainger, J. (2000). Visual and Phonological Codes in Letter and Word Recognition: Evidence from Incremental Priming. *The Quarterly Journal of Experimental Psychology, 53A*(3), 671–692.