American Psychological Association

Journal of Experimental Psychology: Learning, Memory, and Cognition

© 2018 American Psychological Association
0278-7393/19/$12.00

2019, Vol. 45, No. 4, 677–688
http://dx.doi.org/10.1037/xlm0000617

Forced Fixations, Trans-Saccadic Integration, and Word Recognition: Evidence for a Hybrid Mechanism of Saccade Triggering in Reading

Elizabeth R. Schotter
University of South Florida

Titus von der Malsburg
University of Potsdam

Mallorie Leinenger
Denison University

Recent studies using the gaze-contingent boundary paradigm reported a *reversed preview benefit*—shorter fixations on a target word when an unrelated preview was easier to process than the fixated target (Schotter & Leinenger, 2016). This is explained via *forced fixations*—short fixations on words that would ideally be skipped (because lexical processing has progressed enough) but could not be because saccade planning reached a point of no return. This contrasts with accounts of preview effects via *trans-saccadic integration*—shorter fixations on a target word when the preview is more similar to it (see Cutter, Drieghe, & Liversedge, 2015). In addition, if the previewed word—not the fixated target—determines subsequent eye movements, is it also this word that enters the linguistic processing stream? We tested these accounts by having 24 subjects read 150 sentences in the boundary paradigm in which both the preview and target were initially plausible but later one, both, or neither became implausible, providing an opportunity to probe which one was linguistically encoded. In an intervening buffer region, both words were plausible, providing an opportunity to investigate trans-saccadic integration. The frequency of the previewed word affected progressive saccades (i.e., forced fixations) as well as when trans-saccadic integration failure increased regressions, but, only the implausibility of the target word affected semantic encoding. These data support a hybrid account of saccadic control (Reingold, Reichle, Glaholt, & Sheridan, 2012) driven by incomplete (often parafoveal) word recognition, which occurs prior to complete (often foveal) word recognition.

Keywords: parafoveal processing, word recognition, regressive saccades, eye movements, reading

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

It is quite remarkable how quick the reading process is (i.e., skilled reading progresses at about 200–400 words per minute) given how many complex cognitive processes are involved (e.g., word recognition, sentence parsing, eye movement control, etc.; Rayner, Schotter, Masson, Potter, & Treiman, 2016). Part of this remarkable speed comes from the opportunity for readers to obtain information from a word before directly viewing it, when the word is in parafoveal vision (i.e., during parafoveal preview). Although it is clear that parafoveal preview allows the reading system to gain a head-start on processing a word, it is not entirely clear how this happens. For example, is information gained during parafoveal preview used to facilitate reading on its own, or by way of integration with subsequently obtained foveal information, or both? In this article, we will consider three accounts of the relationship between linguistic processing and eye movement control in reading. We review the existing evidence for and against these accounts and then describe a novel experiment that investigates them, particularly with respect to parafoveal preprocessing of linguistic information and what happens when it differs from foveal information on the subsequent fixation.

Traditional Accounts of Parafoveal Preview Effects: Trans-Saccadic Integration

The phenomenon of parafoveal preview is most clearly demonstrated experimentally with a gaze-contingent boundary paradigm (Rayner, 1975) in which a parafoveal preview stimulus changes to a different foveal target word during the saccade to the target (readers are blind to this change if it occurs during the saccade due to saccadic suppression; Matin, 1974). Reading time (e.g., fixation duration on the target) is compared between conditions in which the preview is available (i.e., valid) compared to when it is masked with a different stimulus (i.e., invalid) and, on average, fixation durations are shorter by about 20–50 ms for valid compared to invalid preview conditions (see Rayner, 2009; Reingold et al., 2012; Schotter, Angele, & Rayner, 2012; Vasilev & Angele, 2017). For decades, effects of parafoveal preview have been explained by trans-saccadic linguistic integration accounts, which posit that parafoveal preview information is merged or compared

This article was published Online First July 12, 2018.

Elizabeth R. Schotter, Department of Psychology, University of South Florida; Titus von der Malsburg, Department of Linguistics, University of Potsdam; Mallorie Leinenger, Department of Psychology, Denison University.

Correspondence concerning this article should be addressed to Elizabeth R. Schotter, Department of Psychology, University of South Florida, 4202 East Fowler Avenue PCD 4118G, Tampa, FL 33620. E-mail: eschotter@usf.edu

with foveal target information once directly fixated. On these accounts, processing is easier when preview and target information are more similar and therefore more easily integrated; consequently, fixation durations on the target word are shorter, leading to a preview benefit (e.g., Pollatsek, Lesch, Morris, & Rayner, 1992; Rayner, 1975; see Cutter et al., 2015). Alternatively, processing are more difficult when preview and target information are more dissimilar and less easily integrated; consequently, fixation durations are longer on the target word, leading to a preview cost (e.g., Kliegl, Hohenstein, Yan, & McDonald, 2013; Marx, Hawelka, Schuster, & Hutzler, 2015).

Another Account of Parafoveal Preview Effects: Forced Fixations

Trans-saccadic integration accounts have been successful at accounting for the majority of findings obtained using the gaze-contingent boundary paradigm (see Cutter et al., 2015; Schotter et al., 2012), but not all findings are compatible with the idea of trans-saccadic integration. For example, recent work by Schotter and Leinenger (2016) used the boundary paradigm with unrelated high- and low-frequency target words (e.g., "The boy found a red *phone*/*scarf* on his way to school.") and previews that were either identical to the target (e.g., *phone* was a preview for *phone* or *scarf* was a preview for *scarf*) or the other member of the pair (i.e., the unrelated higher- or lower-frequency word, *phone* was a preview for *scarf* or *scarf* was a preview for *phone*, respectively). In the high-frequency target condition, the unrelated preview led to longer fixation durations on the target than the identical preview—a *standard preview benefit*. However, in the low-frequency target condition, the unrelated preview led to shorter fixation durations on the target than the identical preview—a *reversed preview benefit*. Because trans-saccadic integration accounts are predicated on the idea of better processing with similar compared to dissimilar previews, they cannot explain this reversed preview benefit. Trans-saccadic integration accounts likewise cannot explain recently reported preview plausibility benefits, where readers spend less time on a target when a completely unrelated preview was plausible compared to implausible (e.g., Schotter & Jia, 2016; Veldre & Andrews, 2016). In contrast, another account that Schotter and Leinenger (2016) proposed can explain these effects because their account posits that the durations of (at least some) fixations are determined by information obtained during parafoveal preview, regardless of the subsequently fixated target information.

Schotter and Leinenger (2016) suggested that previewing a word that is easy to process (i.e., plausible and/or high frequency) can lead the reading system to not only skip over it (Rayner, 2009), but also to make a *forced fixation* on it. Forced fixations are short single fixations on the word before moving forward, even when the subsequently fixated target word is completely unrelated to the preview. More precisely, forced fixations are cases in which the system would otherwise skip the word but the skipping decision was made too late, at a point when the saccade program toward the to-be-skipped word could not be cancelled. During forced fixations, the reader is assumed to be insensitive to properties of the fixated target because attention has already shifted to the subsequent word (Morrison, 1984; Schotter, Leinenger, & von der Malsburg, 2018; see the Discussion and Figure 5).

Forced fixations are more likely when the parafoveal preview is easy to process (e.g., high frequency and/or plausible) than when it is not. Importantly, in contrast to purely oculomotor theories of reading (e.g., McConkie & Yang, 2003; Vitu's, 2003 comment on Reichle, Rayner, & Pollatsek, 2003), the duration of forced fixations is determined by lexical properties, specifically the ease of lexically processing the parafoveal preview.

A Variety of Parafoveal Preview Effects: A Hybrid Account

Forced fixations only account for fixations that are relatively short in duration (i.e., those that would otherwise be skips). Consequently, they cannot account for standard preview benefit effects observed for parafoveal nonword previews that are orthographically similar (Rayner, 1975) or phonologically similar (Miellet & Sparrow, 2004) to the target because nonword previews are not easy to recognize and therefore would not lead to skips or forced fixations. To find evidence for forced fixations, it is necessary to use stimuli in which the previews are plausible words that are unrelated to the targets (Schotter & Leinenger, 2016). However, as discussed above, only forced fixations can account for preview effects that do not depend on similarity between the preview and target (i.e., reversed preview benefit and plausibility preview effects).

It seems that there may be multiple ways in which parafoveal information is used in the reading process, either independent of the target (i.e., via forced fixations) or in comparison to it (i.e., via trans-saccadic integration). In fact, using a quantile regression analysis and complementary Vincentile plots, Schotter and Leinenger (2016) demonstrated that short single fixations showed only an influence of the preview word frequency but not an influence of the target word frequency. However, long single fixations showed only an effect of the target word frequency and not the preview word frequency. Lastly, intermediate fixations showed an interaction between preview and target frequency (i.e., an effect of display change; see also Risse & Kliegl, 2014). These data suggest that, even within a single distribution of fixations obtained within a single study, there is evidence for different populations of fixations that are influenced by different properties.

This variety of preview effects may be explained by a hybrid mechanism of saccade triggering that includes both forced fixations and trans-saccadic integration. Reingold et al. (2012) first suggested a hybrid mechanism of saccadic control, which Schotter and Leinenger (2016) also suggested in relation to their forced fixations account. On this hybrid account, forced fixations reflect one saccade triggering mechanism that is engaged when the preview is easy to process whereas preview costs reflect another mechanism that intervenes when trans-saccadic integration failure occurs (i.e., if the system detects a lack of similarity between the preview and target). Phenomenologically, forced fixations are short fixations followed by progressive saccades whereas fixations that are influenced by trans-saccadic integration failure are relatively longer and/or followed by refixations or regressions. Because forced fixations are relatively short there is little to no time for the foveal target information to enter the system (e.g., due to the retina–brain lag; see Reichle & Reingold, 2013) and therefore little influence of the target on the duration or likelihood of these progressive saccades. When forced fixations do not occur (i.e.,

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

when the preview is difficult to process), the fixation on the target word will be longer and there will be more opportunity to obtain information from the foveal target word. Therefore, the fixation duration may be influenced exclusively by the frequency of the target,[1] as shown by Schotter and Leinenger's (2016) longest quantiles, or by the lack of similarity between the preview and target, as shown by Schotter and Leinenger's (2016) middle quantiles.

In addition to investigating oculomotor behavior, there is a question as to the reader's ultimate linguistic interpretation of the sentence. That is, do they encode the preview or the target word? After all, on the forced fixations account, the preview word determines subsequent eye movement behavior and also predicts that, under specific circumstances, the preview word—not the foveally fixated target word—enters the linguistic processing stream. Testing this prediction, will therefore inform us about which of the three accounts outlined above best explains the interplay of oculomotor control and word recognition.

Schotter et al. (2018) explicitly probed the reader's encoding of the text with two-alternative forced choice questions for which the preview and target word were the response options (see also Schotter & Jia, 2016). They found that readers reported reading the target word the majority of the time, except when they skipped over it or fixated it for less than 100 ms (i.e., made a forced fixation on it) and did not return to reread the target word. Similar to the dissociation between early and late processing suggested by Schotter and Leinenger's (2016) quantile regressions, very short forced fixations led to a different type of linguistic encoding than longer fixations. The duration of forced fixations and subsequent linguistic encoding was based on the preview information whereas the duration of longer fixations and the subsequent linguistic encoding was influenced by the target information. Unsurprisingly, if the reader reread the target they almost exclusively reported it rather than the preview, presumably because the reencountered representation of the target replaced whatever the initially encoded representation was (see Booth & Weger, 2013). Thus, comprehension questions at the end of a sentence are only partially informative because online processing (i.e., subsequent reading behavior after initial forced fixations) may contaminate the measurement of the relationship between online processing and comprehension by providing an additional chance to encode the target word.

It is unclear what caused rereading in the Schotter et al. (2018) study in the first place; regressions to reread prior text generally occur when a sentence suddenly stops making sense (e.g., when ambiguity is disambiguated) or to correct oculomotor error (see Bicknell & Levy, 2011; Rayner, 2009; Schotter, Tran, & Rayner, 2014 for discussions of regressions). By design, in studies of forced fixations both the preview and target words made sense in the sentence and therefore comprehension failure is an unlikely explanation. In display change studies where the identity of a word changes during a saccade, readers most often do not notice these changes because of saccadic suppression (Matin, 1974). However, they are sometimes aware of these changes (Angele, Slattery, & Rayner, 2016; Slattery, Angele, & Rayner, 2011), and regressions may increase when they notice them, or perhaps when trans-saccadic integration fails even if readers are not explicitly aware of the change.

The Present Study

In the present study, we investigated evidence for a hybrid mechanism (i.e., both forced fixations and dissociations between downstream linguistic processing behaviors that might indicate trans-saccadic integration failure or only foveal word encoding). We modified sentences from Schotter & Leinenger (2016; Schotter et al., 2018) so that the end of the sentence (the *probe region*) ultimately made the preview and/or the target word implausible. This manipulation allows us to compare rereading behavior between trials in which only the preview eventually becomes implausible and trials in which only the target eventually becomes implausible. Comparisons between the rate of regressions in these two trials indicate which word meaning (i.e., preview or target) had been encoded. If readers had encoded the preview word they should be more likely to make a regression out of the probe region when the preview word becomes implausible than if they had encoded the (plausible) target. Conversely, if readers had encoded the target word, they should be more likely to make a regression out of the probe region when the target word becomes implausible than if they had encoded the (plausible) preview.

To dissociate regressions caused by comprehension processes from regressions made in response to the display changes (i.e., trans-saccadic integration failure), we inserted a few words between the target region and the probe region (which created the implausibility). In this intervening *buffer region*, both words continued to be plausible so any regressions made out of this region could only be due to trans-saccadic integration failure, rather than comprehension processes. Lastly, on this hybrid account, trans-saccadic integration failure should not happen (or should be delayed) during forced fixations. Therefore, we compare the rate of regressions in response to display changes for high-frequency and low-frequency previews. On a forced fixations account, readers should be less likely to notice and immediately respond to display changes when the preview was high frequency (and forced fixations are more likely) than when it was low frequency (and forced fixations are less likely).

Method

Participants

Twenty-four undergraduates from the University of California, San Diego, participated in this experiment for course credit. All were native English speakers, had normal vision, and were naïve to the purpose of the experiment. The study was approved by the University of California, San Diego institutional review board.

Apparatus

Eye movements were recorded with an SR Research Ltd. Eyelink 1000 eye tracker (sampling rate of 1,000 Hz) in a tower setup that restrained head movements with forehead and chin rests. Viewing was binocular, but only the eye movements of the right eye were recorded. Subjects were seated approximately 60 cm

[1] Note, however, that the likelihood of being part of this population of fixations is still determined by lexical properties of the preview word, which determines whether the fixation is forced or not.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

away from an HP p1230 CRT monitor with a screen resolution of 1024 × 768 pixels and a refresh rate of 150 Hz. Text was displayed in black, 12-point, fixed-width Courier New font on a white background. Sentences were always displayed in the vertical center of the screen in one line of text, and 2.65 characters subtended 1° of visual angle. Display changes were completed, on average, within 4 ms of the tracker detecting a saccade crossing the invisible boundary, which was located at the beginning of the space preceding the target word.

Materials

Seventy-five high- and low-frequency noun pairs that were matched in length and with limited orthographic, phonological, and semantic overlap were taken from Schotter et al. (2018; see Table 1).

The sentence frames from Schotter et al. (2018) were modified so that the beginning of the sentence and at least three words following the target remained neutral (i.e., each member of the target pair fit equally well), but the end of the sentence rendered one member of the pair implausible. One frame made the high frequency word implausible and the other made the low frequency word implausible, for a total of 150 experimental sentences (see Figure 1).

For each sentence frame, the two words of the pair could be the preview and/or target word and the orthogonal crossing of these variables created four conditions: (a) both plausible (the non-display-change condition where the preview/target remains plausible at the probe region), (b) target implausible (the display change condition where only the target is implausible at the probe region), (c) preview implausible (the display change condition where only the preview is implausible at the probe region), and (d) both implausible (the non-display-change condition where the preview/target becomes implausible at the probe region). Thus, for each pair of target words, there were eight possible conditions created by a 2 (preview word frequency: high vs. low) × 2 (display type: identical vs. change) × 2 (target word plausibility: plausible vs. implausible) design. The eight conditions were counterbalanced across participants and items in a Latin square design and presented in a randomized order, intermixed with 62 filler sentences (45 of which were followed by comprehension questions) that were plausible and of similar length to the experimental stimuli.

Normative data. Ten native English speakers from the United States, who did not participate in the reading experiment, participated in online norming through Amazon's Mechanical Turk service for monetary compensation. They completed a sentence acceptability-rating task to ensure that each version of the sentence supported one member of the target pair while rendering the other member implausible. The average acceptability score (on a 1–7 scale, with 7 being plausible) for sentences with the compatible target was $M = 6.05$ ($SD = .59$, range = 4.40–7.00) and for the incompatible target was $M = 3.60$ ($SD = 1.24$, range = 1.20–6.20). Below, we report two analyses for regressions out of the probe region: one on the entire stimulus set and one only including items where the implausible items were rated lower than the lowest-rated plausible item (106 items). The results of both analyses were the same.

An additional set of 10 participants, who did not participate in the reading experiment or the sentence acceptability norming, were recruited via Amazon's Mechanical Turk service to complete a cloze norming task for monetary compensation. This norming revealed that both the high and low frequency preview/target words were on average not predictable given the prior sentence context. Average cloze scores were .05 and .02 for the high and low frequency preview/target words respectively.

Procedure

Subjects were instructed to read the sentences for comprehension and to respond to occasional comprehension questions. At the start of the experiment, the eye-tracker was calibrated with a 3-point calibration scheme; calibration error was required to be below .3 degrees of visual angle for every point to proceed to the experiment. Each trial began with a fixation point in the center of the screen, which the subject was required to fixate until the experimenter initiated the trial (or paused to redo the calibration if error was too high). A fixation box then appeared on the left side of the screen at the location of the beginning of the sentence. Once a stable fixation was detected within the box, the box disappeared and was replaced by the sentence, which remained on the screen until the subject pressed a button signaling that they were done reading.

The experiment consisted of 150 experimental sentences in which an invisible boundary was located at the end of the pretarget word (i.e., to the left of the space preceding the target word). While a subject's eyes were to the left of the boundary, the preview word was either the high frequency word (e.g., *phone*) or the low frequency word (e.g., *scarf*). When the tracker detected that the reader's eyes crossed the boundary, either an identical target or the higher- or lower-frequency member of the pair replaced the preview word. There were also 62 filler sentences that were plausible and did not contain display changes, 45 of which (21% of the total number of trials) were followed by comprehension questions (an-

Table 1
Summary Statistics for Target/Preview Words

	High-frequency				Low-frequency			
Word property	*M*	*SD*	Min.	Max.	*M*	*SD*	Min.	Max.
Log HAL frequency/400 mil	10.41	1.01	8.21	12.66	7	1.1	4.61	8.9
Raw frequency per million	136	150	9	786	5	5	0	18
Length	5.83	.9	5	8	5.83	.9	5	8
Cloze predictability	.05	.12	0	.9	.02	.07	0	.5

Note. HAL = Hyperspace Analogue to Language.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

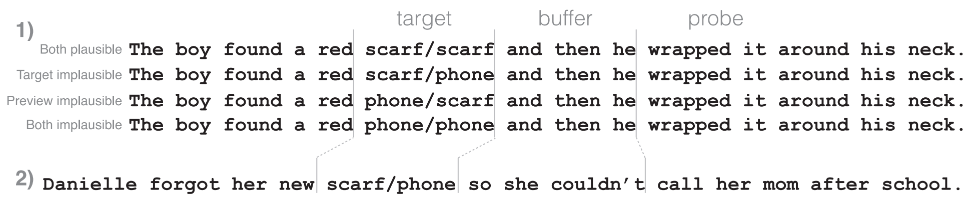

Figure 1. Example sentences with experimental conditions for Sentence 1 labeled to the left (Sentence 2 shows only the preview implausible condition) and analysis regions labeled above. The first word in the target region is the preview and second word is the target; *scarf* is low frequency and *phone* is high frequency.

swered with "yes" or "no"). Responses to comprehension questions were accurate ($M = 91\%$, range $= 81$–97%), suggesting the subjects were reading closely. The experiment lasted approximately 45 min.

Results

Following the data processing procedure used in Schotter and Leinenger (2016), sequential fixations shorter than 81 ms were combined (i.e., summed) with an adjacent fixation if they were within one character space, or were included in the dataset if they were further than one character space from an adjacent fixation. All fixations longer than 800 ms were eliminated. Trials in which there was a blink or track loss on the target word during first pass reading were excluded, as were trials in which the display change was triggered by a saccade that landed to the left of the boundary (i.e., j hooks) or trials in which the display change completed after the reader had started fixating a word. These procedures left 2,951 trials available for analysis (82% of the original data), of which 67% contained single fixations on the target (the critical measure for testing for reversed preview benefit), leaving 1,981 single fixations available for analysis.

We report inferential statistics based on linear mixed-effects models (LMMs) for fixation duration data and generalized linear mixed-effects models (GLMMs) with a logit link for fixation probability data. To fit the (G)LMMs, we used the lmer function from the lme4 package (Version 1.1–12; Bates, Maechler, Bolker, & Walker, 2015) within the R Environment for Statistical Computing (Version 3.3.2; R Development Core Team, 2016). For single fixation duration, because fixation time measures are skewed we ran LMMs on log-transformed data (Schotter & Leinenger, 2016 argued that this better demonstrates the effect on forced fixations at the short end of the distribution[2]). For all models, subjects and items were entered as crossed random effects (see Baayen, Davidson, & Bates, 2008) with the maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013) except when overparameterization required a reduced random effects structure, noted below. We first report single fixation duration to establish a replication of the reversed preview benefit effect, then turn to analyses of downstream regression behavior.

Single Fixation Duration

To test for a reversed preview benefit (i.e., evidence for forced fixations), we analyzed the single fixation duration data with the same LMM model structure used by Schotter and Leinenger (2016; Schotter et al., 2018). We conducted the analyses two ways (i.e., used two models) with custom contrasts that directly test for the magnitude of the preview effects (i.e., a centered [sum-coded] contrast for display change vs. identical) for each target frequency separately (i.e., an uncentered [dummy coded] contrast with either the high- or low-frequency target word as the baseline). This approach is preferable in these studies because Schotter and Leinenger reported a crossover interaction between target type and display type and running two models allows us to estimate the target condition-specific effect of display type directly.

In contrast to Schotter and Leinenger (2016, Experiment 1; Schotter et al., 2018), the main effect of target word frequency was not significant ($t < .01$), likely because of the significant crossover interaction between display type and target frequency ($t = 3.78$; Table 2, Figure 2). Importantly, the pattern of data for the effect of display type for each of the two target frequency conditions is the same as the previous experiments; thus, we report the effects of display type separately for the two baseline conditions.

In the high-frequency target condition, we observed a standard preview benefit (i.e., evidence for trans-saccadic integration failure); longer reading times when the display changed than when it was identical ($t = 3.40$). In the low-frequency target condition, we observed a reversed preview benefit (i.e., evidence for forced fixations) that was on the border of statistical significance ($t = -1.95$). The reversed preview benefit likely failed to reach significance because of larger standard error in the analysis (due to a smaller stimulus set and slightly weaker frequency manipulation than the previous study[3]) relative to studies in which it was fully significant (Schotter & Leinenger, 2016; Schotter et al., 2018), but the consistent pattern of results across the three studies is strong evidence for reversed preview benefit (see the Appendix).

Overview of Analyses of Regressions Out

We first present data for regressions out of the probe region, which tests which word meaning the reader had encoded by the end of the sentence in cases in which they had not already made a regression. To assess what causes regressions to happen prior to

[2] The difference in the *t*-values between the raw and log-transformed data is due to the log-transform shrinking the variance estimates, particularly in the long duration end of the distribution, relative to the mean. Given that forced fixations relate to changes in the short duration end of the distribution, this can lead to large discrepancies between the raw and log-transformed LMM results.

[3] In Schotter and Leinenger (2016) there were 288 items (144 high–low frequency pairs) where the average raw counts per million were 159 and 4, respectively. In the current study, there were 150 items (75 pairs) where the corresponding raw counts per million were 136 and 5.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 2
Results of Linear Mixed Effects Models for Log-Transformed Single Fixation Duration From Models With the High Frequency Target (Left Columns) or Low Frequency Target as the Baseline (Right Columns)

	Model with high frequency baseline			Model with low frequency baseline		
Contrast	b	*SE*	|t|	b	*SE*	|t|
Intercept	**5.452**	**.0293**	**186.10**	**5.452**	**.0300**	**181.69**
Target frequency effect	.0004	.0167	.00	−.0004	.0167	.00
Display type for baseline	**.0976**	**.0287**	**3.40**	−.0484	.0248	1.95†
Display Type × Frequency	**−.1460**	**.0386**	**3.78**	**.1460**	**.0386**	**3.78**

Note. The intercept represents the mean duration for the baseline target frequency averaged across display type and the effect of target frequency is the difference between target frequency conditions averaged across display type. The effects of target frequency and the interaction are identical (with reversed sign) between the two models. Significant effects indicated by boldface; marginally significant effect indicated with a dagger.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

the probe region, we then report analyses of regressions out of the target region and the intervening buffer region. To test for the effects of experimental manipulations on regressions out of each of the regions of interest, we conducted models with only the fixed effects that were relevant or of theoretical interest at that point in the sentence (noted in the beginning of each section below). All contrasts were centered (sum coded) so that the main effects and interactions can be interpreted the same way as an analysis of variance (i.e., main effects and interactions), but we used LMMs with fully crossed random effects structures.

Probe region. As discussed in the introduction, one of the goals of this study is to investigate readers' downstream reading behavior based on linguistic encoding of the preview/target word. Through the plausibility manipulation in the probe region (i.e., in display change trials only the preview or target word was plausible) we can investigate which of the words the reader had likely encoded by comparing the rate of regressions between experimental conditions. We analyzed regressions out of the probe region as a function of target plausibility, display type, and the interaction between them (Figure 3, right panel). Note that the main effect of preview plausibility is statistically equivalent to the interaction between target plausibility and display type (Risse & Kliegl, 2014). For this we only included trials with a single fixation on the target region, and no regression out of the preceding regions (i.e., either the target or buffer region), which would provide a chance to reread the target word before reaching the probe region, reducing the strength of our manipulation; 1,721 trials, 58%).

There was a significant effect of target plausibility; readers made more regressions out of the probe region when the target became implausible than when it remained plausible (b = .64, z = 3.37, $p < .001$). Neither the main effect of display type (b = 0.33, z = 1.71, p = .09) nor the interaction (i.e., the effect of preview plausibility) were statistically significant (b = −0.11, z = 0.30, p = .77). In addition, none of the effects involving preview frequency were statistically significant (all ps > .60). If readers had not already made a regression due to noticing the display change (i.e., from the target or buffer region; see below) any regressions made due to implausibility were triggered by the reader's encoding of the directly fixated target word, rather than the parafoveal preview. These data are in line with the results of Schotter et al. (2018) and suggest that the majority of word encoding occurs in foveal vision, despite the fact that parafoveal information influences eye movement planning as demonstrated by the reversed preview benefit reported above.

Not all sentences had an equally strong plausibility manipulation. Although all sentences with the plausible word were rated as highly plausible (min = 4.4, max = 7.0 on a 7-point scale), there was a larger range in the plausibility ratings for the implausible word (min = 1.2, max = 6.2). For a stronger test of the effect of plausibility, we redid the analysis with only the items for which the implausible item was rated lower than the lowest-rated plausible item (i.e., lower than 4.4 on the 7-point scale). This led to the inclusion of 106 of the 150 original items (71%) for which the new descriptive statistics for the implausible word are (M = 2.93, SD = 0.74, range = 1.2–4.2). The results of this analysis were similar to the analysis on the full stimulus set; there was a main effect of target plausibility (b = 0.77, z = 3.65, $p < .001$), no effect of display type (b = 0.21, z = 0.98, p = .33), no interaction (b = 0.004, z = 0.01, p = .99), and no effects involving preview frequency (all ps > .12).

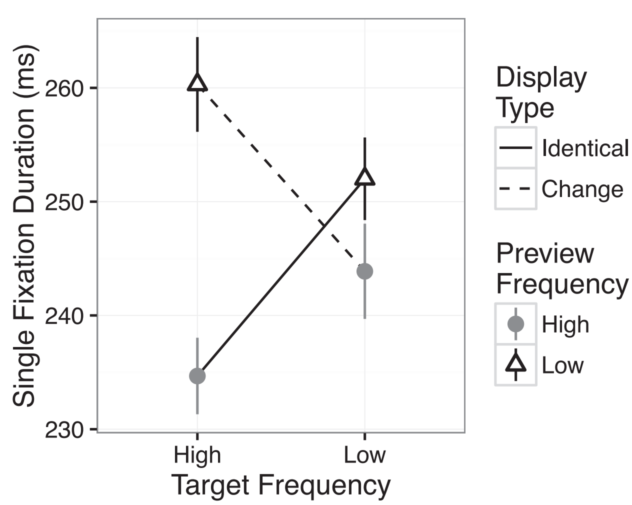

Figure 2. Single fixation duration on the target word as a function of target frequency, preview frequency, and display type. Error bars represent ± 1 SEM.

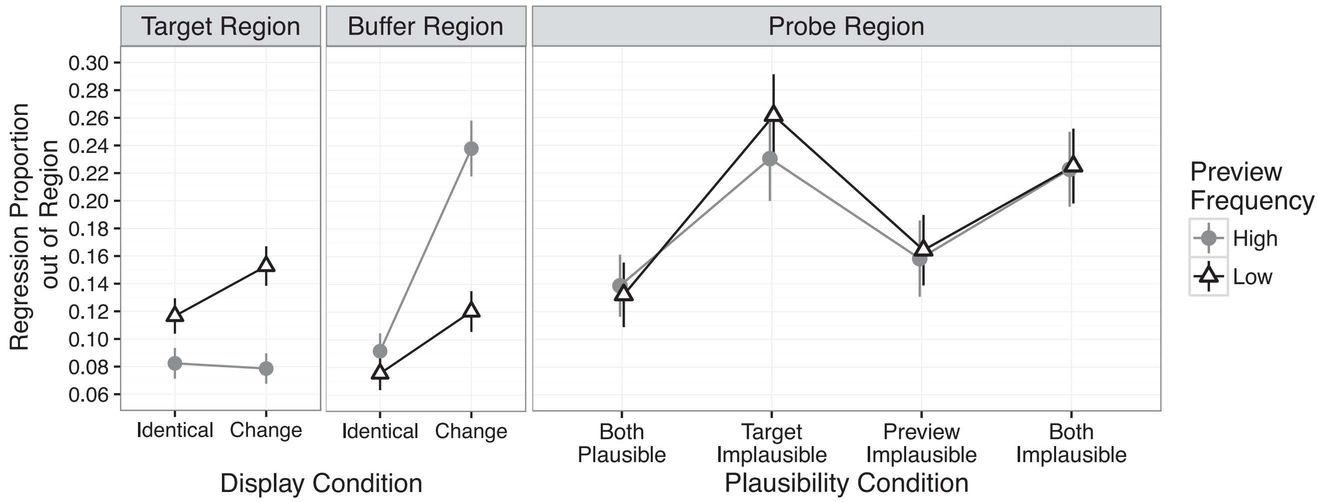

Figure 3. Probability of making a regression out of each region of the sentence (panels) as a function of preview frequency (shape/color), display type (x-axis for the target and buffer region panels) and plausibility condition (x-axis; probe region panel only; the inner two conditions of the probe region panel are display change conditions). Data are only included in the buffer region if there was no regression out of the target region, and were only included in the probe region if there was neither a regression out of the target or buffer regions. Error bars represent ± 1 SEM.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Regressions prior to the probe region. Readers made regressions out of regions before the probe region, and these trials were excluded from the analysis of the probe region. Therefore, in the following sections we investigate which of the experimental manipulations influenced these regressions and what these patterns suggest about a hybrid mechanism involving forced fixations and trans-saccadic integration. On a hybrid mechanism account, we would expect that trans-saccadic integration failure might increase regressions, but not in cases where readers were likely to make a forced fixation. When the preview is low frequency, regressions out of the target region might be higher when the display changed than when it did not because forced fixations should be relatively rare. In contrast, when the preview is high frequency, forced fixations should be relatively common and should lead to no difference in regressions out of the target region between display change and non-display-change conditions. If the different target information is eventually encoded on display change trials for high frequency previews, this may lead to an increase in regressions that is observed after the forced fixation is executed (i.e., from the buffer region).

Target region. We analyzed regressions out of the target region (when it was not skipped; 2,466 trials, 84%) as a function of preview frequency and display type. There was a significant preview frequency effect; readers made more regressions out of target words when the preview was low frequency ($b = 0.83$, $z = 2.73$, $p < .01$). As described above, this is predicted by a hybrid mechanism; when the preview is high frequency readers would be more likely to make a forced fixation—preprogram a forward saccade. The main effect of display type was not statistically significant ($b = -0.24$, $z = 0.75$, $p = .46$), but there was a significant interaction ($b = 1.35$, $z = 2.17$, $p < .05$; Figure 3, left panel). There were more regressions out of the target when the display had changed than when it was identical only when the preview was low frequency, which also aligns with the idea of a hybrid mechanism. When the preview is high frequency, there is no opportunity for the dissimilar target information to influence the preinitiated saccades that are part of the forced fixations phenomenon. However, in the absence of forced fixations (i.e., when the preview is low frequency) there is more opportunity for the different target information to lead to trans-saccadic integration failure during the fixation on the target and cause a relative increase in regressions for display change relative to nondisplay-change conditions.

Buffer region. We analyzed regressions out of the buffer region with the same model structure as regressions out of the target region (Figure 3, middle panel). We included only trials in which readers made a single forward fixation on the target region (i.e., did not skip it and only fixated it once, which is the cleanest test of the forced fixations account; see Schotter & Leinenger, 2016) and did not make a regression out of target region (i.e., had not already responded to the display change; 1,981 trials, 67%). There was a marginally significant preview frequency effect; readers made more regressions out of the buffer region when the preview was high frequency ($b = -.48$, $z = 1.75$, $p = .08$). This pattern also fits with the hybrid mechanism; when the preview was high frequency readers were more likely to have made forced fixations on the target and therefore did not immediately respond to the display change in the target region so that the response showed up later, once they landed in the buffer region. The main effect of display type was statistically significant ($b = 0.99$, $z = 3.56$, $p < .001$); there were more regressions out of the buffer region when the display changed. While numerically the effect of display change was larger when the preview was high frequency, the interaction was not statistically significant ($b = -0.91$, $z = 1.65$, $p = .10$). We found evidence for trans-saccadic integration failure in regressions out of the buffer region and a numerical trend

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

suggesting that the process is qualified by the presence of forced fixations. These data further support a hybrid mechanism in which both forced fixations and trans-saccadic integration failure influence reading behavior.

The Time Course of Regression Behavior

To visualize the time course of regression behavior across the sentence, we plotted the cumulative distribution of regressions as a function of character position in the sentence separately for each of the plausibility and display change condition combinations (see Figure 4). Only trials in which there was a single regression are included, and we plotted the average locations of the boundaries between regions for reference.[4] This figure shows the pattern suggested by the sequence of analyses by region reported above. Regressions out of the buffer region show an effect of display change (i.e., the increase in regressions in the red and pink lines relative to the black and gray lines in the central region), whereas regressions out of the probe region show an effect of target plausibility (i.e., the increase in regressions in the red and black lines relative to the pink and gray lines in the right-hand region).

Discussion

The current study reports data suggesting that preview and target word properties can have different effects on reading behavior and that eye movements during reading are controlled by a hybrid mechanism of saccade triggering. We found evidence for reversed preview benefit when the preview word was higher frequency than the target word, which can be explained by forced fixations but not trans-saccadic integration. We also found that the completion of linguistic encoding occurs primarily in foveal vision because regressions out of a region at the end of the sentence that made one of the preview–target pair implausible showed increased regressions only based on the target word plausibility. In addition, we found evidence for a hybrid mechanism of saccadic control because regressions in response to display changes (i.e., trans-saccadic integration failure) increased relative to nondisplay change conditions in different regions, depending on the frequency

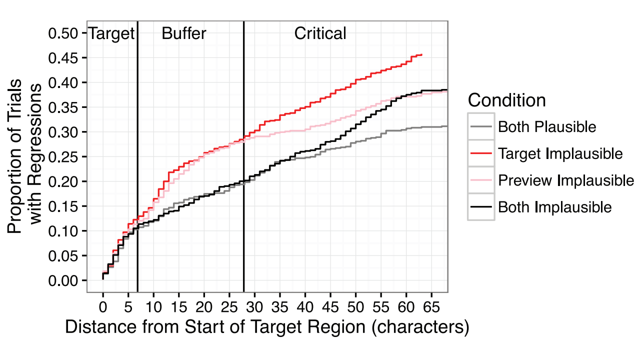

Figure 4. Cumulative plots of regression proportion as a function of distance from the beginning of the target region (in characters) and plausibility condition (colored lines, which are higher on the y-axis, are display change conditions and greyscale lines, which are lower on the y-axis are nondisplay change conditions). Average locations of the boundaries between regions are marked and regions are labeled. See the online article for the color version of this figure.

of the preview word (i.e., the likelihood of forced fixations). Regressions caused by display-changes were more likely from the target region when the preview was low frequency, but such regressions were more likely from the buffer region when the preview was high frequency. We hypothesize that this is because the reader was relatively more likely to have preinitiated the progressive saccade based on the preview information (i.e., make a forced fixation), before the different target information had registered. In contrast, when the preview was low frequency, the reader was relatively less likely to preinitiate a progressive saccade and there was enough time to register that the preview and target were different, leading to a failure of trans-saccadic integration and an increased likelihood of making a regression out of the target region (Figure 3, left panel).

The dissociation between the patterns of regressions from the target region and buffer region lend further evidence for the idea of a hybrid mechanism of saccadic control (Reingold et al., 2012). Trans-saccadic integration failure influenced reading behavior immediately (i.e., in the target region) when forced fixations did not occur, and influenced reading behavior later (i.e., in the buffer region) after forced fixations occurred. These data add to research suggesting that, in the gaze-contingent boundary paradigm, the reading system sometimes detects a display change either explicitly (Angele et al., 2016; Slattery et al., 2011) or implicitly via trans-saccadic integration failure, causing a regression either from the target or immediately following region, depending on how far processing of the preview word had progressed. However, of the 1,233 display change trials in which the reader fixated the target (thereby providing an opportunity to detect that the display had changed) only 318 trials (26%) included a regression from either the target or buffer region. This suggests that this does not happen the majority of the time, especially when considering that the percent of trials with regressions from these regions in nondisplay change trials is not negligible either (17% of trials).

Although the forced fixations account suggests that progressive saccade triggering is directly controlled by lexical identification, the fact that regressions out of the probe region were mostly determined by target word implausibility suggests that it is partial, not complete recognition, that influences these saccades. In fact, the data from Schotter et al. (2018) using probe questions after the sentence suggest that there is some proportion of cases in which a reader directly fixates a word but actually had encoded the preview. Clearly, the relationship between higher-order language and oculomotor behavior is complex; other studies show that the plausibility of a word (determined by the preceding context) can have an immediate influence on saccade decisions when that word is encountered (e.g., gaze duration; Rayner, Warren, Juhasz, & Liversedge, 2004; Reichle, Warren, & McConnell, 2009; Warren & McConnell, 2007; Warren, McConnell, & Rayner, 2008), even based only on a parafoveal preview (Schotter & Jia, 2016; Veldre & Andrews, 2016). It will be important for future research to delineate what lexical properties are used to plan progressive saccades, and the degree to which those influences are dissociable from complete word recognition during reading.

[4] Locations plotted are average across all stimuli. Across sentences the character position of the boundaries of regions varied, but the location was always in the same place for all four conditions of a given sentence.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Accounting for These Effects in Models of Oculomotor Control in Reading

As mentioned, these data suggest that progressive saccade planning and complete word recognition may be dissociable during reading. In fact, this suggestion is part of the architecture of current models of oculomotor control in reading, which posit that the decision to start planning an eye movement forward from a word is made before complete recognition of it. For example, the E-Z Reader model (Pollatsek, Reichle, & Rayner, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998) posits two stages of word recognition; the first stage is a *familiarity check* (L_1), which initiates both saccade programming and a second stage of *lexical access* (L_2), in which word recognition is completed. Although the process that determines fixation durations is only partial recognition of words, such a mechanism still assumes that oculomotor decisions are directly controlled by cognitive (i.e., lexical) processing (i.e., as opposed to oculomotor accounts, which deny the involvement of cognitive-lexical processing; McConkie & Yang, 2003; Vitu, 2003).

Given this architecture, forced fixations can already be explained within the E-Z Reader model (Reichle et al., 1998) and potentially any model that has two-stage lexical processing and two-stage saccade program assumptions (e.g., Engbert, Nuthmann, Richter, & Kliegl, 2005). In fact, the forced fixations account that Schotter and Leinenger (2016) proposed was based on the modeling approach of Schotter, Reichle, and Rayner (2014), who used the E-Z Reader model to estimate how far into lexical processing the model has progressed based on the preview. They found that there were times in which the model reached the second stage of word recognition (i.e., L_2, which is initiated by the completion of the partial word recognition stage, L_1, which also initiates saccade programming).

Forced fixations in oculomotor models of reading. Within the architecture of the E-Z Reader model, fixation behavior on the upcoming word depends on the timing of word recognition of the upcoming word relative to the timing of saccade planning toward that word. In addition to the two stages of word recognition mentioned above, E-Z Reader posits two stages of saccade planning; during the first stage (M_1), the current saccade plan can be canceled, but during the second stage (M_2) the current saccade cannot be cancelled. If L_1 for the upcoming word completes during the M_1 stage of the saccade program toward it, that saccade would be cancelled and replaced with a skip. However, if L_1 for the upcoming word completes during the M_2 stage, the saccade toward that word could not be cancelled and the system would instead preinitiate the subsequent saccade program forward from the upcoming word because saccades can be programmed in parallel (Becker & Jürgens, 1979; Morrison, 1984). The preinitiation of subsequent saccade programs due to an inability to execute a skip leads to the intervening fixation on that word being relatively short—these are forced fixations (see Figure 5). When skips or forced fixations do not happen, the word will be fixated for a relatively longer amount of time and foveal information is needed to initiate the progressive saccade. If the display had changed in these cases, the system must deal with the fact that the new information obtained in the fovea differs from the information that has been previously obtained from the parafovea.

Trans-saccadic integration in oculomotor models of reading. Other modeling work has focused on explaining standard preview

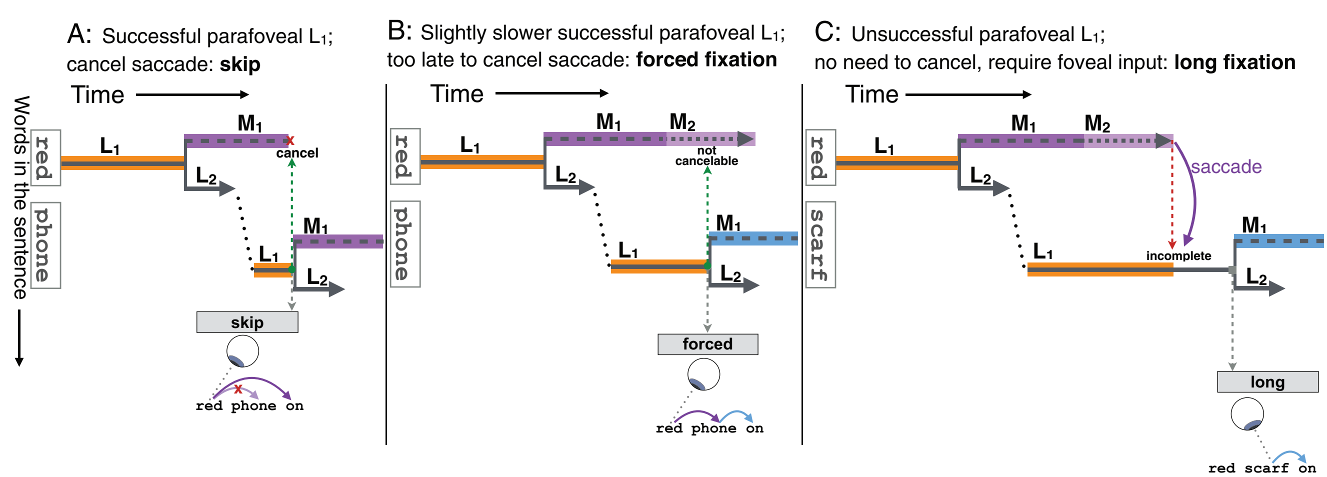

Figure 5. Schematic of the forced fixation account within the architecture of the E-Z Reader model. The type of fixation behavior (represented in three separate panels) depends on when the L_1 stage of word identification for word $n + 1$ completes relative to the saccade planning stages for the saccade from word n to word $n + 1$ (see Schotter & Leinenger, 2016 for details). A: Skips: the first saccade plan, represented in purple and with the label M_1, is replaced with a second saccade plan that moves from n to $n + 2$; B: Forced fixations: the first saccade plan cannot be canceled so the subsequent saccade plan, represented below in blue and also labeled M_1, is initiated while the first saccade plan is being implemented. This is more likely when word $n + 1$ is easy to identify (e.g., is a high frequency word like *phone*); C: Long fixations: the first saccade plan is completely executed before the second saccade plan is ever initiated because word recognition in the parafovea never completes by the time the word is fixated. This is more likely when word $n + 1$ is difficult to identify (e.g., a low frequency word like *scarf* or a nonword). See the online article for the color version of this figure.

effects by assuming that lexical processing does not start until fixation on the target (e.g., Pollatsek et al., 2006), analogous to the idea that there is only an influence of the target word, or have reset lexical processing after a display change (e.g., Risse, Hohenstein, Kliegl, & Engbert, 2014), analogous to the idea of trans-saccadic integration failure. It may be the case that delaying lexical processing is what happens when the parafoveal preview is completely unrecognizable (i.e., a nonword), but such a process could not explain why nonword previews that are orthographically or phonologically related to the target provide a benefit relative to unrelated previews (see Schotter et al., 2012). Likewise, completely resetting lexical processing after a display change would not be able to account for these effects, either. Therefore, any model simulations that set out to explain standard preview benefit effects that align with the idea of trans-saccadic integration must, to some degree, take into account how the words are represented as they are being processed. That is, presumably these standard preview benefit effects are due to the "resetting" of processing being less severe or costly when the preview and target are similar compared to when they are dissimilar. This is a very interesting area for future research because adequately simulating these effects might give us a better idea of the ongoing process of word representation during reading.

Together, the data reported above suggest a fairly sophisticated division of labor within the reading system that may help to optimize the trade-off between speed and accuracy. When ongoing understanding of the sentence breaks down (e.g., due to implausibility), the comprehension system can intervene and trigger regressions. The time course with which these regressions are observed depends on how far through the oculomotor programming process the reading system had progressed. Thus, reading can be considered a two-stage process in which partial lexical processing of parafoveal information is used to "hedge a bet" that word recognition will be successful and to program saccades to maintain reading speed (i.e., by skipping or preinitiating subsequent eye movements). But such an architecture may ultimately not be too risky, as it has relatively small negative consequences on the reader's ultimate understanding of most of the words in sentences (i.e., those that are not skipped or do not receive a forced fixation). The reading system takes advantage of high acuity foveal information and the opportunity to make regressions under uncertainty or confusion to maintain reading comprehension accuracy (Bicknell & Levy, 2011; Booth & Weger, 2013; Schotter et al., 2014). If readers become less confident in their interpretation of the text they can make regressions and reread to reencode the words through direct foveal vision.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

References

Angele, B., Slattery, T. J., & Rayner, K. (2016). Two stages of parafoveal processing during reading: Evidence from a display change detection task. *Psychonomic Bulletin & Review, 23,* 1241–1249.

Baayen, R. H., Davidson, D. H., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412.

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,* 255–278.

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

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research, 19,* 967–983.

Bicknell, K., & Levy, R. (2011). Why readers regress to previous words: A statistical analysis. In L. Carlson, C. Holscher, & T. Shipley (Eds.), *Proceedings of the Cognitive Science Society* (pp. 931–936). Austin, TX: Cognitive Science Society.

Booth, R. W., & Weger, U. W. (2013). The function of regressions in reading: Backward eye movements allow rereading. *Memory & Cognition, 41,* 82–97.

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2015). How is information integrated across fixations in reading? In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 245–260). Oxford, UK: Oxford University Press.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112,* 777–813.

Kliegl, R., Hohenstein, S., Yan, M., & McDonald, S. A. (2013). How preview space/time translates into preview cost/benefit for fixation durations during reading. *The Quarterly Journal of Experimental Psychology, 66,* 581–600.

Marx, C., Hawelka, S., Schuster, S., & Hutzler, F. (2015). An incremental boundary study on parafoveal preprocessing in children reading aloud: Parafoveal masks overestimate the preview benefit. *Journal of Cognitive Psychology, 27,* 549–561.

Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin, 81,* 899–917.

McConkie, G. W., & Yang, S. N. (2003). How cognition affects eye movements during reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 413–428). Oxford, UK: Elsevier.

Miellet, S., & Sparrow, L. (2004). Phonological codes are assembled before word fixation: Evidence from boundary paradigm in sentence reading. *Brain and Language, 90,* 299–310.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 667–682.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 148–162.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52,* 1–56.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81.

Rayner, K., Schotter, E. R., Masson, M., Potter, M. C., & Treiman, R. (2016). So much to read, so little time: How do we read, and can speed reading help? *Psychological Science in the Public Interest, 17,* 4–34.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1290–1301.

Rayner, K. (2009). The Thirty Fifth Sir Frederick Bartlett Lecture: Eye movements and attention during reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62,* 1457–1506.

R Development Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125–157.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26,* 445–476.

Reichle, E. D., & Reingold, E. M. (2013). Neurophysiological constraints on the eye-mind link. *Frontiers in Human Neuroscience, 7,* 361.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using EZ Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16,* 1–21.

Risse, S., Hohenstein, S., Kliegl, R., & Engbert, R. (2014). A theoretical analysis of the perceptual span based on SWIFT simulations of the n+ 2 boundary paradigm. *Visual Cognition, 22,* 283–308.

Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive Psychology, 65,* 177–206.

Risse, S., & Kliegl, R. (2014). Dissociating preview validity and preview difficulty in parafoveal processing of word n + 1 during reading. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 653–668.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception & Psychophysics, 74,* 5–35. http://dx.doi.org/10.3758/s13414-011-0219-2

Schotter, E. R., & Jia, A. (2016). Semantic and plausibility preview benefit effects in English: Evidence from eye movements. *Journal Experimental Psychology: Learning, Memory & Cognition, 42,* 1839–1866.

Schotter, E. R., & Leinenger, M. (2016). Reversed preview benefit effects: Forced fixations emphasize the importance of parafoveal vision for efficient reading. *Journal of Experimental Psychology: Human Perception and Performance, 42,* 2039–2067. http://dx.doi.org/10.1037/xhp0000270

Schotter, E. R., Leinenger, M., & von der Malsburg, T. (2018). When your mind skips what your eyes fixate: How forced fixations lead to comprehension illusions in reading. *Psychonomic Bulletin & Review*. Advance online publication. http://dx.doi.org/10.3758/s13423-017-1356-y

Schotter, E. R., Reichle, E. D., & Rayner, K. (2014). Rethinking parafoveal processing in reading: Serial attention models can account for semantic preview benefit and N+2 preview effects. *Visual Cognition, 22,* 309–333.

Schotter, E. R., Tran, R., & Rayner, K. (2014). Don't believe what you read (only once): Comprehension is supported by regressions during reading. *Psychological Science, 25,* 1218–1226.

Slattery, T. J., Angele, B., & Rayner, K. (2011). Eye movements and display change detection during reading. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1924–1938.

Vasilev, M. R., & Angele, B. (2017). Parafoveal preview effects from word N+ 1 and word N+ 2 during reading: A critical review and Bayesian meta-analysis. *Psychonomic Bulletin & Review, 24,* 666–689.

Veldre, A., & Andrews, S. (2016). Is semantic preview benefit due to relatedness or plausibility? *Journal of Experimental Psychology: Human Perception and Performance, 42,* 939–952.

Vitu, F. (2003). The basic assumptions of EZ Reader are not well-founded. *Behavioral and Brain Sciences, 26,* 506–507.

Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review, 14,* 770–775.

Warren, T., McConnell, K., & Rayner, K. (2008). Effects of context on eye movements when reading about possible and impossible events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1001–1010.

(*Appendix follows*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix

Meta-Analysis of Reversed Preview Benefit

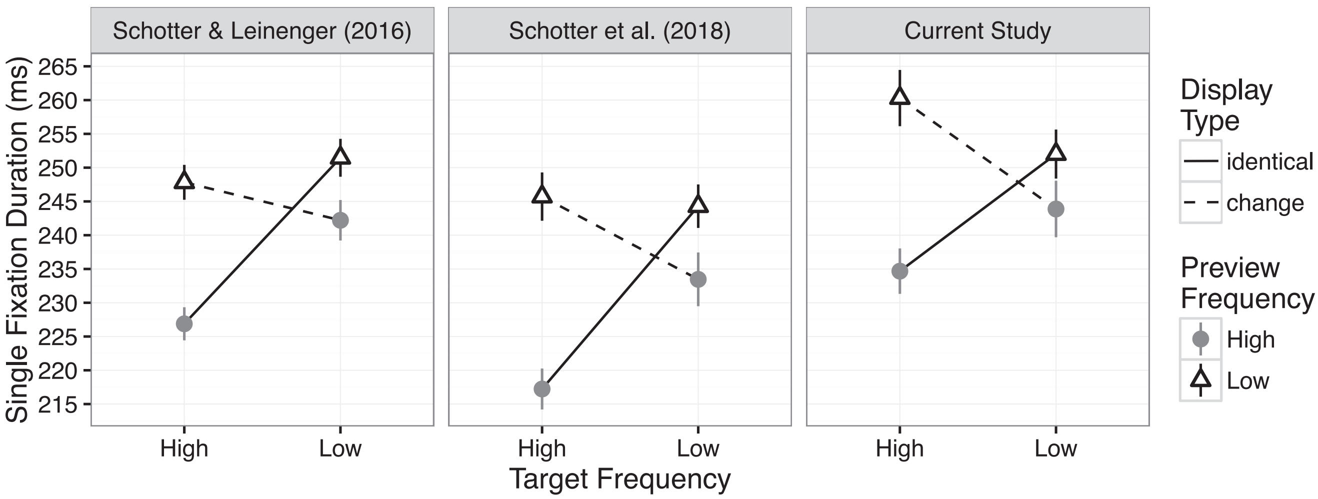

Figure A1. Raw single fixation duration on the target word as a function of target frequency, preview frequency, and display type in three experiments: Schotter and Leinenger (2016, Experiment 1), Schotter et al. (2018), and the current study. Error bars represent ± 1 SEM.

We conducted a meta-analysis on the log-transformed single fixation data (the model with the low frequency baseline), including experiment as an additional factor with three levels with the original study as the baseline and separate contrasts for each of the comparisons between the replication studies to the baseline (e.g., Schotter et al., 2018 vs. Schotter & Leinenger, 2016 and the current study vs. Schotter & Leinenger, 2016) and the interactions between those contrasts and the other fixed effects in the model.[5] There was a significant effect of target frequency, a significant reversed preview benefit, and a significant interaction (all ts $>$ 2.55). None of the main effects or interactions involving experiment were statistically significant (all ts $<$ 1.59), suggesting that these effects were robust to different groups of subjects and the cutting of power associated with choosing a smaller set of items from Schotter and Leinenger (2016) for Schotter et al. (2018) and the current study (see Figure A1).

[5] Because subjects did not participate in more than one experiment, the random effects for subjects did not include experiment. Due to overparameterization, we reduced the random effects for items to include the intercept and slopes for display, experiment, and their interaction.

Received July 5, 2017
Revision received March 13, 2018
Accepted March 19, 2018 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.