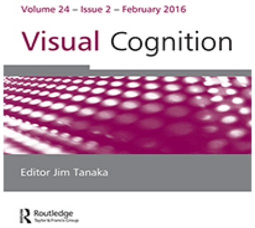
Routledge
Taylor & Francis Group



Visual Cognition

ISSN: 1350-6285 (Print) 1464-0716 (Online) Journal homepage: http://www.tandfonline.com/loi/pvis20

Are individual differences in masked repetition and semantic priming reliable?

Luuan Chin Tan & Melvin J. Yap

To cite this article: Luuan Chin Tan & Melvin J. Yap (2016) Are individual differences in masked repetition and semantic priming reliable?, Visual Cognition, 24:2, 182-200, DOI: 10.1080/13506285.2016.1214201

To link to this article: http://dx.doi.org/10.1080/13506285.2016.1214201

VISUAL COGNITION, 2016
VOL. 24, NO. 2, 182–200
http://dx.doi.org/10.1080/13506285.2016.1214201

Routledge
Taylor & Francis Group

Are individual differences in masked repetition and semantic priming reliable?

Luuan Chin Tan and Melvin J. Yap

Department of Psychology, National University of Singapore, Singapore

ABSTRACT

Despite the robustness of semantic priming (e.g., *cat–DOG*), the test-retest and internal reliabilities of semantic priming effects within individuals are surprisingly low. In contrast, repetition priming (e.g., *dog–DOG*) appears to be far more reliable across a range of conditions. While Stolz and colleagues attribute the low reliability in semantic priming to uncoordinated automatic processes in semantic memory, their use of unmasked priming paradigms makes it difficult to fully rule out the influence of strategic processes. In the present study, we explored the reliability of semantic and repetition priming when primes were heavily masked and cannot be consciously processed. We found that masked repetition, but not semantic, priming effects showed some degree of reliability. Interestingly, skilled lexical processors (as reflected by vocabulary knowledge and spelling ability) also produced larger masked repetition priming effects.

ARTICLE HISTORY
Received 25 November 2015
Revised 28 June 2016
Accepted 10 July 2016

KEYWORDS
Masked repetition priming; masked semantic priming; lexical decision; individual differences

Given that word recognition is a critical aspect of skilled reading, the processes that support the recognition of visually presented words have been extensively investigated. In addition to studies exploring the recognition of words presented in isolation, context effects on word recognition have also been widely studied via priming paradigms. In such paradigms, two letter strings are presented consecutively, and the nature of the relation between the two letter strings can be manipulated. For example, primes can be orthographically (*couch—TOUCH*), phonologically (*much—TOUCH*), morphologically (*touching—TOUCH*), or associatively/semantically (*feel—TOUCH*) related to the target word (see Yap & Balota, 2015, for a review); primes can also be identical to the target word (e.g., *touch—TOUCH*). In addition, primes can either be *unmasked* (i.e., available to conscious awareness) or *masked* (i.e., presented very briefly to minimize conscious processing). The crucial advantage of masked priming is that participants have little to no phenomenological awareness of the prime–target relation, thereby making it less likely that performance is contaminated by strategies (Forster, 1998; see also Kinoshita & Lupker, 2003). Masked priming can therefore potentially serve as a powerful tool for exploring early and relatively modular processes in word recognition (Kinoshita & Norris, 2012).

To preview, the present study investigates the psychometric reliability of masked repetition priming (*touch—TOUCH*) and masked semantic priming (*feel—TOUCH*), with the goal of providing further insights into the early processes that support these two forms of priming. The second goal is to explore the extent to which masked repetition and semantic priming effects are moderated by theoretically important individual differences such as vocabulary knowledge and spelling ability.

The bases of masked repetition and semantic priming

The most commonly used masked priming paradigm, by far, is the three-field paradigm introduced by Forster and Davis (1984) over three decades ago. In this paradigm, each trial comprises the following events: a forward mask (#####) for 500 ms, a briefly presented lowercase prime (for typically between 30–60 ms), followed by an uppercase target. Participants are required to respond to the target, either by classifying it as a word or nonword (lexical decision) or reading it aloud as accurately and quickly as possible (speeded pronunciation). Although participants were not able to see the prime, Forster and Davis (1984) reported that lexical decision times were

CONTACT Melvin J. Yap melvin@nus.edu.sg Department of Psychology, National University of Singapore, Block AS4, #02-07, 9 Arts Link, 117570, Singapore

reliably faster when targets (e.g., *APPLE*) were preceded by repetition (e.g., *apple*), compared to unrelated (e.g., *table*), primes. Furthermore, masked repetition priming effects were of comparable magnitude for low- and high-frequency targets (but see Kinoshita, 2006). To explain these findings, Forster and Davis (1984) suggested that masked repetition priming reflects an entry-opening mechanism, whereby a masked prime makes contact with the lexical entry for that item, and in so doing temporarily increases its accessibility.

It is worth noting that the entry-opening perspective rests on the idea that masked repetition priming implicates stable, abstract representations and automatic processes that are relatively insensitive to the experimental context and task demands (Davis & Kim, 2006). There has been some opposition to this view. Most notably, according to Bodner and Masson's (1997, 2001, 2003) memory recruitment account of priming, word recognition and memory phenomena are assumed to tap common underlying mechanisms. That is, the presentation of a masked prime creates an episodic memory trace which can then be subsequently recruited to facilitate target recognition. Importantly, the extent to which the episodic trace is relied upon depends on the prime's task relevance (Anderson & Milson, 1989). That is, in task contexts where the prime has greater utility, there should be greater retrospective reliance on the priming episode. Consistent with this perspective, Bodner and Masson (1997) reported that masked repetition priming effects became larger when the proportion of targets preceded by a repetition prime in an experiment was .8 instead of .2; the payoff for recruiting the masked prime in the .8, compared to the .2, proportion condition is higher. While this sort of flexibility is difficult to reconcile with the entry-opening account, it can be explained by the idea that there is more strategic recruitment of episodic prime information as the validity of the prime (as reflected by the proportion of repetition trials) increases (see Bowers, 2000; Tenpenny, 1995, for more discussion).

Masked associative/semantic priming is generally much smaller in magnitude than masked repetition priming (Bodner & Masson, 2003). In contrast to unmasked semantic priming, which is mediated by a mixture of automatic and controlled mechanisms (see McNamara, 2005; Neely, 1991, for reviews), masked semantic priming is typically assumed to reflect semantic processing of the prime via automatic spreading activation (Grossi, 2006). That is, the presentation of a prime (e.g., *cat*) activates its semantic representation, which in turn preactivates associatively or semantically related concepts (e.g., *dog*), thereby facilitating their subsequent identification (Posner & Snyder, 1975). The masked semantic priming literature is fairly contentious, and some researchers have questioned the premise that semantic priming can occur in the absence of conscious awareness. For example, De Wit and Kinoshita (2015) recently explored masked semantic priming in lexical decision and semantic categorization, and found reliable effects only in the latter task. They argued that this sort of task-dependency is difficult to reconcile with an automatic priming mechanism, which should yield effects that generalize across tasks. However, as a counterpoint to De Wit and Kinoshita's (2015) failure to observe masked semantic priming effects in lexical decision, an important meta-analysis of 46 published articles (Van den Bussche, Van den Noortgate, & Reynvoet, 2009) provides a preponderance of evidence that subliminally presented information *can* be processed semantically in visual word recognition, even when competing (non-semantic) explanations such as automatized stimulus–response mappings are ruled out. Additionally, although it is clear that masked semantic priming effects are indeed larger in semantic categorization (mean effect size = .80), they are nonetheless also significant in lexical decision and pronunciation (mean effect size = .47).

Interestingly, Bodner and Masson (2003) have proposed that masked semantic priming, like masked repetition priming, can also be explained by memory recruitment. That is, the masked semantic prime establishes an episodic trace that can be retrospectively recruited when participants are trying to recognize the word. In support of this, Bodner and Masson (2003) found larger masked semantic priming effects when the proportion of targets preceded by a semantic prime in an experiment was .8 instead of .2. However, it is worth noting that Bodner and Masson's (2003) claim has been undermined by other studies (e.g., Grossi, 2006; Perea & Rosa, 2002) that have failed to replicate this effect. More seriously, the general validity of the memory recruitment account has been widely criticized in recent years due to its underspecified nature and equivocal empirical support (see Bodner & Masson, 2014, for more discussion). For example, prime proportion effects in masked priming can also be explained by other mechanisms, such as an adaptive adjustment of the response deadline (Kinoshita, Forster, & Mozer, 2008; Kinoshita, Mozer, & Forster, 2011; see also Klapp, 2007).

Are semantic and repetition priming reliable?

For the most part, studies of word recognition have focused on group-level data that are averaged across participants. This emphasis on characterizing a "prototypical" reader seems at odds with evidence that there are not only substantial individual differences among readers, but that these differences have an impact on word recognition performance (Yap, Balota, Sibley, & Ratcliff, 2012). Central to the analysis of individual differences is the issue of reliability. Reliability refers to the degree to which one gets the same results each time one measures a construct. The consistency of such a construct can be evaluated across time (i.e., test-retest reliability) and across items within a test (inter-item or split-half reliability) (Nunnally & Bernstein, 1994). The reliability of a measure is an important psychometric property that greatly circumscribes its usefulness (Kopriva & Shaw, 1991). Specifically, the reliability of a measure places an upper limit on the extent to which priming effects might be expected to correlate with other measures (Hutchison, Balota, Cortese, & Watson, 2008). While one might intuitively expect effects that are very robust at the group level to be also psychometrically reliable, this intuition turns out to be incorrect. For example, despite the robustness of the classic Stroop effect (i.e., slower colour naming times for incongruent stimuli, e.g., *BLUE* printed in red), the test-retest reliability of the Stroop effect has been found to be unexpectedly low (Lowe & Rabbitt, 1998).

Are semantic and repetition priming effects reliable? Empirically, there is substantial variability in the magnitude of priming produced by different participants (Stolz, Besner, & Carr, 2005). However, it is unclear if this variability reflects systematic individual differences or random measurement noise. The reliability of priming has profound implications for any study that attempts to explore individual differences in priming (e.g., are semantic priming effects moderated by age?). Without first establishing the reliability of priming, one cannot tell if the absence of a correlation between priming and some other measure of interest truly reflects no real relation or low reliability on one or both measures. Should it turn out that repetition and semantic priming are indeed unreliable, this will qualify research that have examined individual differences in repetition priming in domains such as learning (e.g., Woltz & Shute, 1993) and lexical processing (e.g., Monahan, Florentino, & Poeppel, 2008), or individual differences in semantic priming in domains such as personality (e.g., Matthews & Harley, 1993), psychopathology (e.g., Morgan, Bedford, & Rossell, 2006), attentional control (Hutchison, 2007), and lexical processing (e.g., Laver & Burke, 1993; Plaut & Booth, 2000; Yap, Tse, & Balota, 2009).

To address this, Stolz et al. (2005) explored the reliability of semantic priming effects in visual word recognition across a range of experimental contexts. Specifically, the within-session reliability (assessed by split-half correlations) and between-session reliability (assessed by test-retest correlations) of semantic priming effects were evaluated across conditions where relatedness proportion (RP, i.e., the proportion of word targets preceded by a related prime) and prime–target stimulus onset asynchrony (SOA) were orthogonally manipulated. Importantly, although the prime–target pairs in the two halves or sessions were *not* identical, they were theoretically comparable, being similar on properties such as target (e.g., length, frequency) and prime–target (e.g., forward and backward associative strength) characteristics. The influence of controlled priming mechanisms should be minimized when RP is low and SOA is short, and maximized when RP is high and SOA is long (Stolz & Neely, 1995). Interestingly, Stolz et al. (2005) found that semantic priming was statistically unreliable ($r = -.06$) under conditions (i.e., SOA of 200 ms, RP of .25) which minimize the contributions of controlled priming mechanisms, suggesting that automatic priming reflects an "inherently noisy and uncoordinated" (p. 328) semantic system. Interestingly, as the task context increased the influence of controlled priming, priming became *more* reliable. This suggests that any observed reliability in semantic priming is due to priming performance gaining coherence through participants' reliance on controlled mechanisms such as expectancy generation (i.e., intentional generation of candidates for to-be-presented targets; Becker, 1980) and semantic matching (i.e., retrospective search for a target-to-prime relation; Neely, Keefe, & Ross, 1989).

Although Stolz et al.'s (2005) results are important and theoretically intriguing, it is possible that the absence of reliability in the short SOA, low RP condition may reflect uncoordinated *pre-semantic*, rather than semantic, processes. To illustrate this, consider the Interactive Activation model (see Figure 1) described by Stolz and Besner (1996). Within this framework, excitatory activation cascades forward from the feature level to the semantic level, by way of the letter and lexical levels. In this light, unreliability in semantic-level processing (Stolz et al., 2005) may thus implicate earlier pre-semantic processing at the lexical, letter, and/or feature levels. To address this, Waechter, Stolz, and Besner (2010) extended the study by Stolz et al. (2005) to examine the reliability of repetition priming. The rationale here is that semantic priming taps semantic processing whereas repetition priming taps pre-semantic processing. If repetition priming effects are also found to be unreliable, this would suggest that the lack of reliability seen in semantic priming may not be specific to semantic processing but is instead mediated by earlier processes.

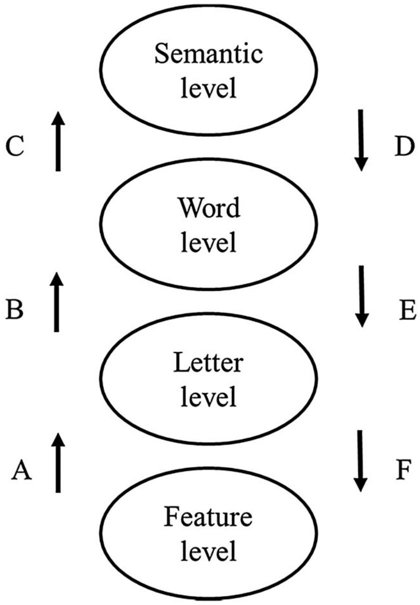


Figure 1. An Interactive Activation framework. Pathways A, B and C feed activation from the feature level to the higher levels. Pathways D, E and F feed activation from the semantic level to the lower levels. From "Role of Set in Visual Word Recognition: Activation and Activation Blocking as Nonautomatic Processes" by J. A. Stolz and D. Besner, 1996, *Journal of Experimental Psychology: Human Perception & Performance, 22*, p. 1168. Copyright 1996 by the American Psychological Association. Reproduced with permission.

For our purposes, the key finding from Waechter et al. (2010) was that even under conditions (i.e., 200 ms SOA, RP of .25) which minimize the contributions of controlled priming mechanisms, the repetition priming effect ($r = .33$) was moderate in size and statistically reliable. This contrasts strongly with the pattern in semantic priming. As a result, they concluded that the observed unreliability of semantic priming most likely arises from uncoordinated processes specific to semantic memory while automatic processes at earlier stages of visual word recognition unfold in a more coherent and consistent manner.

The present study

There are two related major objectives in the present study. The first is to follow up on the claim by Stolz and colleagues that automatic semantic priming is unreliable whereas automatic repetition priming is reliable. To recapitulate, Stolz et al. (2005) and Waechter et al. (2010) reported that semantic, but not repetition, priming was unreliable under experimental conditions which minimize reliance on controlled priming mechanisms. However, even at a very short SOA of 200 ms, there is conscious awareness of the primes (Forster, 1998), which may allow some participants to strategically engage controlled processes such as expectancy generation (Hutchison, 2007). To more fully rule out the influence of controlled processes, we will examine the test-retest and split-half reliabilities of masked repetition and semantic priming effects in the same set of participants. This will help determine if the results of Stolz and colleagues generalize when priming predominantly reflects relatively automatic mechanisms.

The second objective is to examine the extent to which masked repetition and semantic priming effects (assuming they are reliable) are predicted by theoretically important measures of individual differences such as vocabulary knowledge and spelling performance. There is surprisingly little work relating masked priming to individual differences. In the word recognition domain, Andrews and Hersch (2010) reported that masked orthographic priming effects (e.g., *jury—FURY*) were systematically related to individual differences in reading, spelling ability, and vocabulary knowledge (i.e., knowledge of word form and meaning). There is also evidence that masked semantic priming effects (e.g., *mole—EAGLE*) in the semantic categorization task (e.g., is this an animal?) are moderated by individual differences in spelling and vocabulary (Andrews, 2015). According to Perfetti and Hart's (2001) lexical quality hypothesis, readers vary in the precision and coherence of their orthographic representations, which can be selectively tapped by spelling ability and vocabulary knowledge respectively (Andrews, 2015). For highly skilled readers, word identification involves the precise activation of the corresponding underlying lexical presentation, with minimal activation of orthographically similar words. Furthermore, such readers are less dependent on the strategic use of context (e.g., prime information) to facilitate target identification (Yap et al., 2009). If fluent lexical processors are indeed less reliant on the prime context, one might expect skilled readers to show smaller effects of masked semantic and repetition priming, compared to readers who are less proficient.

Collectively, the results of these analyses will help shed more light on the relations between the quality of underlying lexical representations, masked priming phenomena, and measures of individual differences. To our knowledge, the present study is the first attempt to answer these intertwined questions in a unified manner.

Method

Participants

Two hundred and forty undergraduates from the National University of Singapore participated in this study in exchange for course credit. All participants reported English as their first language and had normal or corrected-to-normal vision.

Design

We examined the joint effects of Prime Relatedness (related vs. unrelated) and Session (first vs. second) as a function of Prime Type (repetition vs. semantic). Two 2 × 2 (Prime Relatedness × Session) designs were incorporated within the same study, with non-overlapping items used to examine the effects for each type of priming. All variables were manipulated within participants. The assignment of pairs to Session 1 or Session 2 was counterbalanced across participants, as was the order of prime type (e.g., repetition priming followed by semantic priming vs. semantic priming followed by repetition priming) within each session. In line with the studies by Stolz et al. (2005) and Waechter et al. (2010): (a) participants were presented with different but theoretically comparable prime–target pairs in the two sessions, (b) no target was repeated within an experiment, and (c) both word and nonword targets were preceded by word primes. A total of 16 counterbalancing versions were generated (see Figure 2 for one version); Session 1 is reflected by the two phases before the language test while Session 2 is reflected by the two phases after the test. Within each phase, there were 40 related and 40 unrelated word trials, as well as 80 nonword trials. The dependent variables were response time and accuracy.

Materials

Measures of individual differences

Participants were assessed on two measures of written language proficiency: spelling and vocabulary knowledge. In the spelling task, participants were presented with 88 letter strings, half of which was correctly spelled and the other half misspelled (Andrews & Hersch, 2010), and were required to identify words as being correctly spelled or not via a button press. In the vocabulary knowledge task, participants were asked to complete the 40-item vocabulary subscale test of the Shipley Institute of Living Scale (Shipley, 1940). Each target word was accompanied by four options, and participants had to select via a button press the number of the option that was closest in meaning to the target word. Spelling and vocabulary scores for each participant are reflected by the proportion of correct items on each task.

Stimuli

With respect to repetition priming, the same 160 targets used by Waechter et al. (2010) were used, as these have been shown to yield large, reliable repetition priming effects. An additional 160 words were obtained from the English Lexicon Project (ELP; Balota et al., 2007) and matched to related primes on word frequency (Brysbaert & New, 2009), number of letters, number of syllables, and number of orthographic neighbours (Coltheart, Davelaar, Jonasson, & Besner, 1977). Consistent with Waechter et al. (2010), these 160 new words were randomly paired with the targets to serve as unrelated primes.

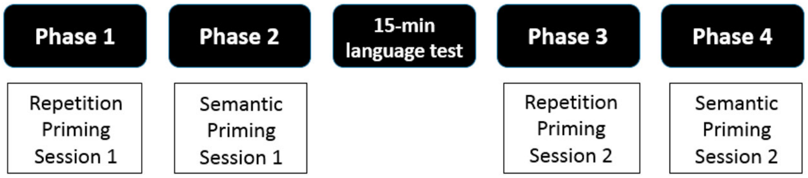


Figure 2. Schematic of the study.

Turning to semantic priming, we used the Nelson, McEvoy, and Schreiber (2004) norms to select 160 symmetric prime–target pairs that possessed relatively similar forward (prime-to-target) and backward (target-to-prime) association strengths. Symmetric prime–target pairs have been shown to produce much stronger priming than pairs which possess only a prime-to-target or target-to-prime relation (Thomas, Neely, & O'Connor, 2012). The close match between forward and backward associative strengths also makes it less likely that participants will be biased to rely on either prospective or retrospective priming mechanisms (Thomas et al., 2012). Across participants, stimuli were counterbalanced across the related and unrelated conditions; unrelated prime–target pairs were created by re-pairing the primes and targets within each set. The descriptive statistics of the stimuli are presented on Table 1.

Table 1. Descriptive statistics for prime and target stimuli used in the study.

	Repetition Priming		Semantic Priming	
	Mean	SD	Mean	SD
Target Concreteness[a]	3.97	0.98	3.65	1.02
Target Length	4.56	0.63	4.99	1.59
Target Word Frequency (LgSUBTLCD)[b]	2.87	0.60	3.10	0.65
Target Orthographic Neighbourhood Size	5.93	5.15	5.33	5.34
Target Phonological Neighbourhood Size	13.84	10.31	11.48	10.78
Target Orthographic Levenshtein Distance 20[c]	1.60	0.34	1.76	0.57
Target Phonological Levenshtein Distance 20[d]	1.38	0.38	1.59	0.66
Target Number of Syllables	1.21	0.44	1.42	0.58
Target Number of Morphemes	1.08	0.26	1.18	0.42
Prime Concreteness			3.60	1.05
Prime Length			4.99	1.65
Prime Word Frequency (LgSUBTLCD)			3.25	0.57
Prime Orthographic Neighbourhood Size			5.21	5.53
Prime Phonological Neighbourhood Size			11.81	11.13
Prime Orthographic Levenshtein Distance 20			1.78	0.58
Prime Phonological Levenshtein Distance 20			1.60	0.62
Prime Number of Syllables			1.47	0.65
Prime Number of Morphemes			1.16	0.39
Forward Association Strength[e]			0.49	0.14
Backward Association Strength[e]			0.57	0.16

Note: SD refers to standard deviation.
[a] Brysbaert, Warriner, and Kuperman (2014)
[b] Brysbaert and New (2009)
[c] Yarkoni, Balota, and Yap (2008)
[d] Yap and Balota (2009)
[e] Nelson, McEvoy, and Schreiber (2004)

For both priming paradigms, a total of 320 legal nonwords were created using Wuggy, a multilingual pseudoword generator (Keuleers & Brysbaert, 2010) which creates nonwords that conform to the phonotactic constraints of the English writing system, and are matched to words on number of letters, number of syllables, and orthographic neighbourhood size. An additional 320 words were selected from the ELP (Balota et al., 2007) to serve as nonword primes, which were matched to word primes on word frequency, number of letters, number of syllables, and orthographic neighbourhood size. At the start of each lexical decision task (LDT), 20 practice items (five related words, five unrelated words, 10 nonwords) were administered before the experimental trials began.

Procedure

PC-compatible computers with 17″ Viewsonic CRT monitors (75 Hz refresh rate) were used to present stimuli and collect data, using DMDX software (Forster & Forster, 2003). Participants were tested individually in a sound-attenuated cubicle in a single session lasting approximately 45 minutes. Each participant went through four LDTs in total, with a vocabulary and spelling test administered in the middle of the study. For the LDT, participants were instructed to decide whether the letter string presented after the string of hashes formed a word or nonword by pressing the appropriate button (right shift key for word and left shift key for nonword). They were encouraged to respond quickly, but not at the expense of accuracy. No mention was made of the masked primes. The 20 practice trials were then followed by two blocks of 80 experimental trials, with a break between the two blocks.

Each trial in the LDT comprised three successive displays: (1) a forward mask (##########) presented for

500 ms, (2) the lowercase prime for 40 ms, and (3) the uppercase target for up to 5000 ms (see Figure 3 for the trial structure). Stimuli were presented in white against a black background, using the 11-point Calibri font. The order of trials was randomized anew for each participant.

Results

Response times for trials that were correctly responded to were subjected to the non-recursive moving criterion procedure described in Van Selst and Jolicoeur (1994). In this procedure, the criterion cutoff for trimming RTs (in terms of standard deviations) is proportional to the number of observations within the different experimental cells for each participant; as the number of trials increases, a more conservative trimming criterion is used. For the RT analysis, data trimming removed 8.71% (5.88% errors; 2.83% RT outliers) of experimental trials.

In addition to priming effects based on raw RTs, we also computed z-score transformed RTs (zRTs) for our participants (Faust, Balota, Spieler, & Ferraro, 1999; Hutchison et al., 2008). That is, participants' RTs were converted to z-scores based on the mean and SD of their RTs. This allows us to express participant-level priming effects in standard deviation units. In general, as baseline RT (a proxy for processing speed) becomes slower, there is a tendency for effect sizes to become larger due to greater variance in the measure. Indeed, Hutchison et al. (2008) have cautioned that evaluating participant-level priming effects based on raw RTs can lead to misleading conclusions because baseline RT is not taken into account. Computing priming effects using zRTs helps to control for within-participant variation in baseline RTs across the different experimental blocks, and places priming effects across blocks on a common scale. Importantly, simulations by Bush, Hess, and Wolford (1993) have also showed that when analyzing skewed RT data, an approach that combines trimming and the Z transformation not only works well but yields greater statistical power than an analysis of raw means.

Analysis 1: group-level priming performance

Trials for each participant were first partitioned into Session 1 (S1) trials and Session 2 (S2) trials. Descriptive statistics as a function of Prime Type (repetition vs. semantic), Prime Relatedness (related vs. unrelated), and Session (first vs. second) are presented in Table 2, and the data were analyzed using repeated-measures ANOVA.

Masked repetition priming

With respect to RTs, the main effect of Prime Relatedness was significant, $F(1, 239) = 950.39$, $p < .001$, MSE $= 421.44$, $\eta_p^2 = .80$, with faster RTs for related ($M = 508$ ms), compared to unrelated ($M = 549$ ms), trials. The main effect of Session was significant, $F(1, 239) = 30.47$, $p < .001$, MSE $= 1544.86$, $\eta_p^2 = .11$, with faster RTs for Session 1 ($M = 521$ ms), compared to Session 2 ($M = 535$ ms), trials. The Prime Relatedness × Session interaction was significant, $F(1, 239) = 7.02$, $p = .009$, MSE $= 451.97$, $\eta_p^2 = .03$; the priming effect was slightly larger at Session 2 ($M = 44$ ms) than at Session 1 ($M = 38$ ms); tests of simple effects revealed that the masked repetition priming effect was statistically significant for both Session 1 ($p < .001$) and Session 2 ($p < .001$) trials.

Turning to the zRTs, the main effect of Prime Relatedness was significant, $F(1, 239) = 803.60$, $p < .001$, MSE $= .03$, $\eta_p^2 = .77$, with faster zRTs for related ($M = -.50$), compared to unrelated ($M = -.20$), trials. The main effect of Session was significant, $F(1, 239) = 28.74$, $p < .001$, MSE $= .06$, $\eta_p^2 = .11$, with faster zRTs for Session 1 ($M = -.39$), compared to Session 2 ($M = -.31$), trials. The Prime Relatedness × Session interaction was significant, $F(1, 239) = 9.94$, $p = .002$, MSE $= .02$, $\eta_p^2 = .04$; the priming effect was larger at Session 2 ($M = .32$) compared to Session 1 ($M = .27$); tests of simple effects revealed that the masked repetition priming effect was statistically significant across Session 1 ($p < .001$) and Session 2 ($p < .001$) trials.

Finally, for accuracy rates, the main effect of Prime Relatedness was significant, $F(1, 239) = 132.86$, $p < .001$, MSE $= .001$, $\eta_p^2 = .36$, with a higher accuracy rate for related ($M = .96$), compared to unrelated ($M = .93$), trials. The main effect of Session was significant, $F(1, 239) = 8.97$, $p = .003$, MSE $= .001$, $\eta_p^2 = .036$, with a higher accuracy rate for Session 1 ($M = .95$), compared to Session 2 ($M = .94$), trials. However, the Prime Relatedness × Session interaction was not significant, $F < 1$.

Masked semantic priming

With respect to raw RTs, the main effect of Prime Relatedness was significant, $F(1, 239) = 84.85$,

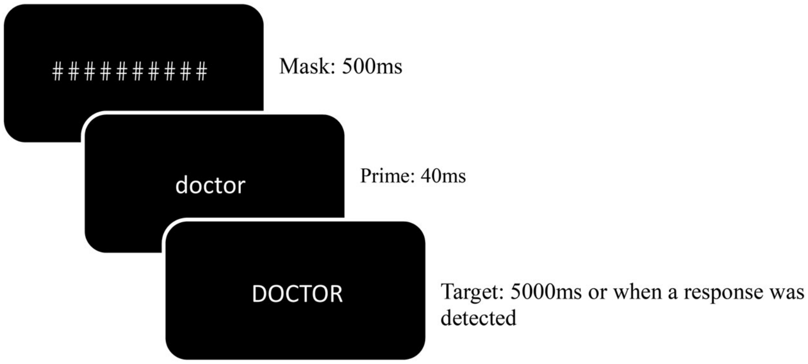


Figure 3. Sample trial structure.

$p < .001$, MSE = 365.20, $\eta_p^2 = .26$, with faster RTs for related ($M = 523$ ms), compared to unrelated ($M = 534$ ms), trials. The main effect of Session was significant, $F(1, 239) = 61.45$, $p < .001$, MSE = 1619.84, $\eta_p^2 = .20$, with faster RTs for Session 1 ($M = 518$ ms), compared to Session 2 ($M = 539$ ms), trials. The Prime Relatedness × Session interaction was not significant, $F < 1$.

Turning to zRTs, the main effect of Prime Relatedness was significant, $F(1, 239) = 107.37$, $p < .001$, MSE = .02, $\eta_p^2 = .31$, with faster zRTs for related ($M = -.39$), compared to unrelated ($M = -.30$), trials. The main effect of Session was significant, $F(1, 239) = 57.03$, $p < .001$, MSE = .07, $\eta_p^2 = .19$, with faster zRTs for Session 1 ($M = -.41$), compared to Session 2 ($M = -.28$), trials. The Prime Relatedness × Session interaction was not significant, $F < 1$.

Finally, for accuracy rates, the main effect of Prime Relatedness was significant, $F(1, 239) = 27.63$, $p < .001$, MSE = .001, $\eta_p^2 = .10$, with a higher accuracy rate for related ($M = .96$), compared to unrelated ($M = .95$), trials. The main effect of Session was significant, $F(1, 239) = 5.53$, $p = .02$, MSE = .001, $\eta_p^2 = .02$, with a higher accuracy rate for Session 1 ($M = .954$), compared to Session 2 ($M = .949$), trials. The Prime Relatedness × Session interaction was not significant, $F < 1$.

In summary, at the group level, we observed statistically reliable effects for both masked repetition and masked semantic priming. The magnitude of masked repetition priming effects ($M = 41$ ms) was much larger than for masked semantic priming ($M = 11$ ms), which is consistent with the extant literature (e.g., Gomez, Perea, & Ratcliff, 2013). It is also interesting and noteworthy that the magnitude of masked repetition priming is almost equivalent to the absolute value of the SOA (40 ms), consistent with the idea that priming reflects a savings or head-start (Forster, Mohan, & Hector, 2003).

Analysis 2: reliability of masked priming effects

Trials for each participant were partitioned into Session 1 (S1), Session 2 (S2), odd-numbered, and even-numbered trials; trial number refers to the order in which the trial was presented to the participant. For each participant, priming effects based on zRTs were computed for all trials, S1 trials, S2 trials, odd-numbered trials, and even-numbered trials, for both masked repetition and semantic priming. As discussed earlier, using z-score transformed priming effects helps control for within-participant variability in baseline RTs across blocks (Hutchison et al., 2008).

Table 2. Means and standard deviations of priming performance across priming conditions as a function of dependent variables.

			Raw RT (ms)		*z*-score RT		Accuracy	
			M	SD	M	SD	M	SD
Masked Repetition Priming	Session 1	Related	502	60	−0.53	0.20	0.96	0.03
		Unrelated	540	59	−0.26	0.18	0.93	0.05
	Session 2	Related	513	70	−0.47	0.22	0.96	0.04
		Unrelated	557	66	−0.14	0.19	0.93	0.06
Masked Semantic Priming	Session 1	Related	513	60	−0.45	0.19	0.96	0.04
		Unrelated	524	58	−0.37	0.19	0.95	0.04
	Session 2	Related	533	71	−0.33	0.21	0.96	0.04
		Unrelated	544	68	−0.24	0.20	0.94	0.05

In addition, the distinguishing of S1 from S2 trials, and of odd- from even-numbered trials, are necessary for computing test-retest and split-half reliabilities respectively. Table 3 presents the means and standard deviations of masked priming effects by experimental conditions and trial types (all trials, odd-numbered trials, even-numbered trials, S1 trials, and S2 trials).

Figure 4 presents the distributions of zRT priming effects as a function of prime type. These boxplots show clear and substantial between-participants variability in the magnitude of masked semantic and repetition priming effects. It is also clear that priming is not a universal phenomenon (Stolz et al., 2005), particularly for semantic priming: 24% (57/240) of the participants did not show facilitatory masked semantic priming, whereas only 1.7% (4/240) did not show facilitatory masked repetition priming. Interestingly, about 20% of Stolz et al.'s (2005) participants also produced no semantic priming effect in the low RP, short SOA condition.

Turning to the reliability analyses, robust correlation techniques were used to estimate the correlations between S1 and S2 trials (test-retest reliability) and between odd- and even-numbered trials (split-half reliability) as a function of prime type. Robust correlation coefficients are less susceptible to violations of parametric assumptions and the presence of outliers (Larson-Hall, 2016). Using R (R Development Core Team, 2014), we ran the corr.plot function in the mvoutlier package (Filzmoser & Gschwandtner, 2015) to obtain the classical (i.e., Pearson) and robust correlation ellipsoids for each scatterplot, along with the respective correlation coefficients for each ellipsoid (see Figure 5).

Let us first consider masked repetition priming. For classical correlations, both the test-retest ($r = .21$, $p = .001$) and split-half ($r = .36$, $p < .001$) correlations were significant. For robust correlations, the test-retest ($r = .29$, $p < .001$) and split-half ($r = .43$, $p < .001$) correlations were also significant. In contrast, the results were more equivocal and less stable for masked semantic priming. Specifically, while the robust correlations were weakly significant (test-retest $r = .15$, $p = .02$; split-half $r = .17$, $p = .008$), the classical correlations were respectively .07 and .05, *ns*. In sum, when one considers the collective evidence, test-retest and split-half correlations were generally moderate in magnitude and statistically significant for masked repetition priming, but this was not the case for masked semantic priming.

Table 3. Means and standard deviations of z-score transformed masked priming effects as a function of prime type and trial type.

	Overall	Odd	Even	Session 1	Session 2
Masked Repetition Priming (zRTs)					
M	0.30	0.29	0.29	0.27	0.32
SD	0.16	0.20	0.19	0.19	0.22
Masked Semantic Priming (zRTs)					
M	0.08	0.08	0.09	0.08	0.09
SD	0.12	0.18	0.16	0.15	0.19

Analysis 3: individual differences in masked priming

To investigate whether masked priming effects were moderated by theoretically important individual differences in spelling ability and vocabulary knowledge, we analyzed our data using linear mixed effects (LME) models (Baayen, Davidson, & Bates, 2008). Using R (R Development Core Team, 2014), we fitted our RT data using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015); *p*-values for fixed effects were obtained using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2016). The main and interactive effects of the factors of interest were treated as fixed effects; effect coding was used for Prime Relatedness, whereby related trials were coded as −.5 and unrelated trials were coded as .5. Random intercepts for participants and targets, along with by-participant and by-target random slopes for relatedness, were also included in each

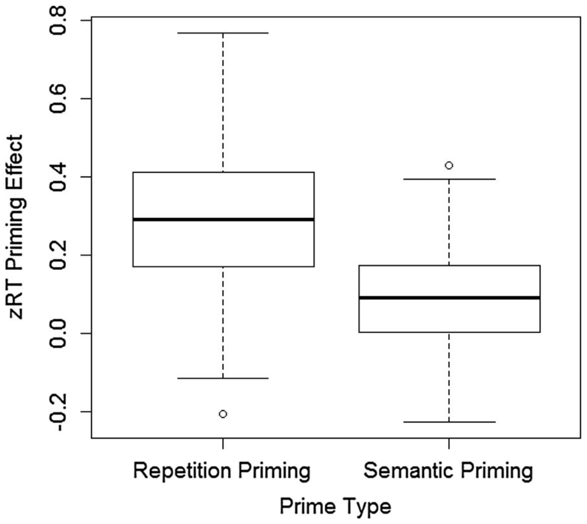


Figure 4. Distributions of masked repetition and semantic priming effects.

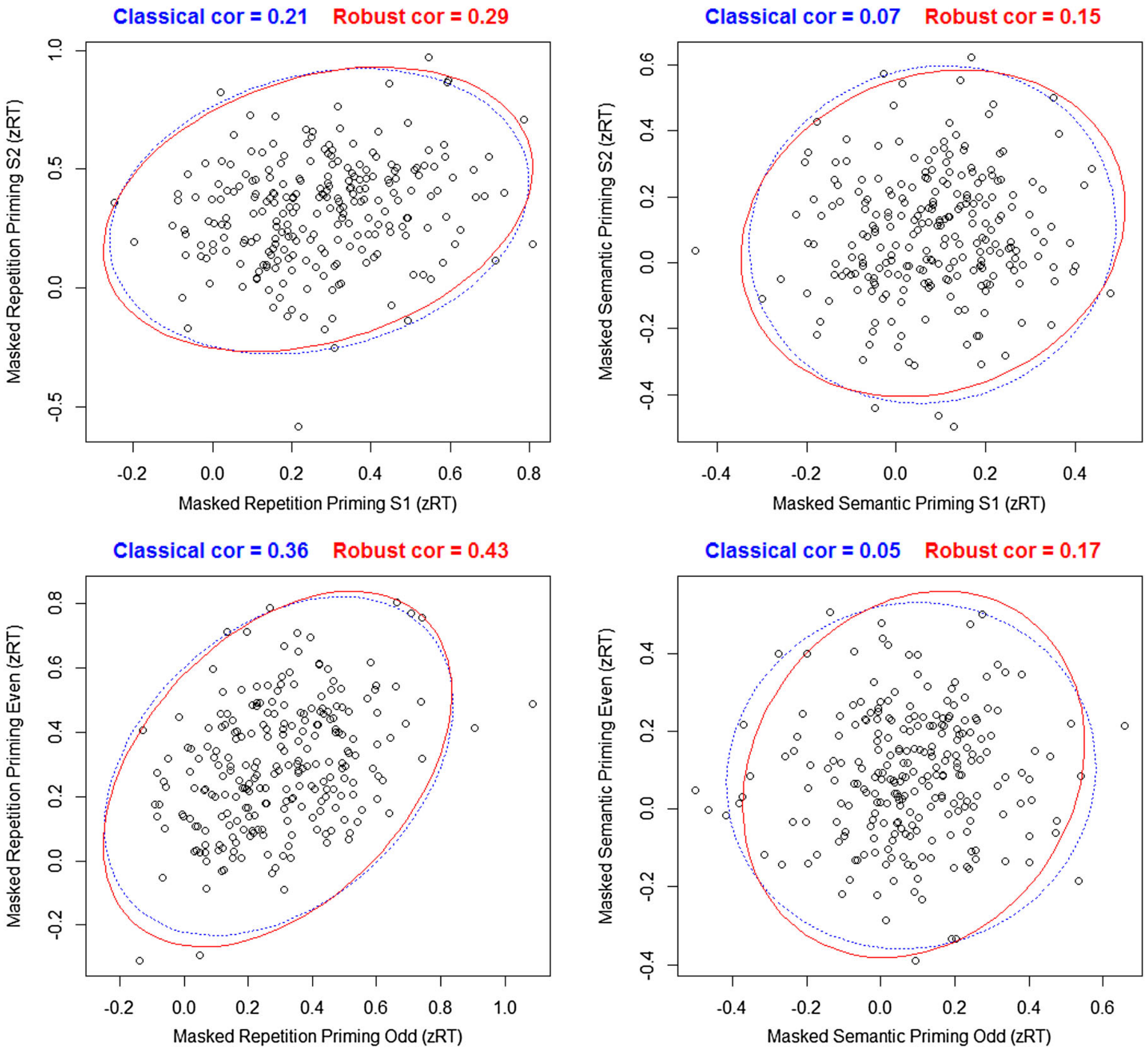


Figure 5. Scatterplots reflecting test-retest and split-half correlations for masked repetition and semantic priming.

model. In mixed-effects modelling, the standard recommendation is to use reciprocally transformed RTs (i.e., −1/RT; Masson & Kliegl, 2013; see also Kliegl, Masson, & Richter, 2010, for more discussion). RT data are typically positively skewed in cognitive tasks, and the non-linear reciprocal transformation helps to normalize the residuals.

There are a couple of problems with using reciprocally transformed RTs in the present analyses. First, there is mounting evidence that the indiscriminate use of reciprocally transformed RT data in LME models is associated with serious drawbacks (Balota, Aschenbrenner, & Yap, 2013; Lo & Andrews, 2015). Briefly, when the same data are analyzed using LME analyses of transformed RTs or ANOVAs of mean RT data, there are meaningful and systematic discrepancies between the two sets of analyses (Balota et al., 2013; Lo & Andrews, 2015). Specifically, overadditive effects (i.e., larger effects of Factor A at higher levels of Factor B) become more additive, and additive effects (i.e., main effects of both factors but no interaction) become more underadditive (i.e., smaller effects of Factor A at higher levels of Factor B). Furthermore, some statisticians (e.g., Gelman & Hill, 2007) have pointed out that violating the normality of residual assumption has virtually no impact on estimating regression slopes, and specifically recommend *against* including this as a regression diagnostic. This suggests that the modest benefits of residual normality may be more than offset by the spurious inferences non-linear transformations may yield. Second, and of even greater concern, reciprocal transformation, unlike the z-score transformation, does *not* control for baseline RT. Given our interest in the extent to which participant-level priming is moderated by individual differences, it is important that the priming effects for different participants are placed on a common scale. In light of the foregoing considerations, the analyses reported in this section are based on zRTs.

Table 4 (masked repetition priming) and Table 5 (masked semantic priming) present the results for the joint effects of Prime Relatedness with Spelling Ability (top panel) and Vocabulary Knowledge (bottom panel). Participants with vocabulary (*M* = .75, SD = .08) or spelling (*M* = .86, SD = .06) scores more than 3 interquartile ranges below the first quartile were classified as extreme outliers. We decided to exclude the single participant (out of 240 participants) who met this criterion, given the possibility that his or her unusually low score on the spelling test (.61) is likely to reflect a lack of motivation or a fundamental misunderstanding of the task instructions.

The results of the foregoing analyses are relatively easy to summarize. Both masked repetition (Table 4) and masked semantic (Table 5) priming effects were statistically significant, consistent with the results of Analysis 1. Masked semantic priming did not interact with either spelling or vocabulary (Table 5). Interestingly though, the interaction between masked *repetition* priming and vocabulary was significant (Table 4). Plotting the simple slope for this interaction (see Figure 6) revealed that masked repetition priming effects were larger for participants with more vocabulary knowledge.

Given that Analysis 2 established that masked semantic priming effects are psychometrically unreliable, it is unsurprising that masked semantic priming effects were *not* related to either the vocabulary or spelling measure. That being said, the analyses involving masked repetition priming are less conclusive. Specifically, while masked repetition priming interacted significantly with vocabulary knowledge, its moderation by spelling ability was not significant. However, although the present study included 240 participants, they were recruited from a relatively homogenous participant pool, which may restrict the range of spelling and vocabulary scores. To establish the generalizability and robustness of the interactions involving masked repetition priming, analysis of archival data from the masked Form Priming Project (FPP; Adelman et al., 2014) was conducted. The FPP

Table 4. LME model estimates (based on zRT) for fixed and random effects for the joint effects of masked repetition priming with spelling ability (top panel) and vocabulary knowledge (bottom panel).

Random Effects	Variance	SD	*r*
Items			
Intercept	.065	.255	
Prime Relatedness	.016	.127	.350
Participants			
Intercept	.011	.106	
Prime Relatedness	.015	.124	–.060

Fixed Effects	Coefficient	Standard Error	*p*-value
Intercept	–.328	.022	<.001
Prime Relatedness	.305	.014	<.001
Spelling Ability (centred)	–.009	.133	NS
Vocabulary Knowledge (centred)	–.134	.098	NS
Prime Relatedness × Spelling Ability	.279	.172	NS

Random Effects	Variance	SD	*r*
Items			
Intercept	.065	.255	
Prime Relatedness	.016	.127	.350
Participants			
Intercept	.011	.106	
Prime Relatedness	.015	.124	–.060

Fixed Effects	Coefficient	Standard Error	*p*-value
Intercept	–.328	.022	<.001
Prime Relatedness	.305	.014	<.001
Spelling Ability (centred)	–.001	.133	NS
Vocabulary Knowledge (centred)	–.141	.098	NS
Prime Relatedness × Vocabulary Knowledge	.265	.126	.037

Table 5. LME model estimates (based on zRT) for fixed and random effects for the joint effects of masked semantic priming with spelling ability (top panel) and vocabulary knowledge (bottom panel).

Random Effects	Variance	SD	*r*
Items			
Intercept	.062	.250	
Prime Relatedness	.014	.117	.330
Participants			
Intercept	.013	.115	
Prime Relatedness	.004	.061	–.230

Fixed Effects	Coefficient	Standard Error	*p*-value
Intercept	–.327	.021	<.001
Prime Relatedness	.091	.012	<.001
Spelling Ability (centred)	.082	.142	NS
Vocabulary Knowledge (centred)	.142	.104	NS
Prime Relatedness × Spelling Ability	.102	.125	NS

Random Effects	Variance	SD	*r*
Items			
Intercept	.062	.250	
Prime Relatedness	.014	.117	.330
Participants			
Intercept	.013	.115	
Prime Relatedness	.004	.062	–.230

Fixed Effects	Coefficient	Standard Error	*p*-value
Intercept	–.327	.021	<.001
Prime Relatedness	.091	.012	<.001
Spelling Ability (centred)	.094	.141	NS
Vocabulary Knowledge (centred)	.140	.105	NS
Prime Relatedness × Vocabulary Knowledge	.019	.092	NS

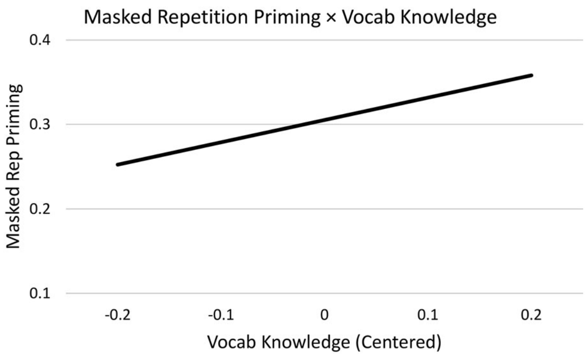


Figure 6. Masked repetition priming as a function of spelling ability (top panel) and vocabulary knowledge (bottom panel). Masked repetition priming is computed by zRT (unrelated)—zRT (related).

contains behavioural data for over 800,000 lexical decision trials across 28 types of form priming (including repetition priming) from 1015 participants over 14 testing universities. Importantly, the participants in the FPP went through the same vocabulary and spelling tasks used in the present study. Given that the FPP is based on a much larger, more diverse sample, range restriction should be less of a problem, and the data should also afford more statistical power to detect subtle relations between masked repetition priming and the individual difference measures. Indeed, if we look at Figure 7, it is clear that there is more variability in the FPP dataset, particularly for the spelling scores.

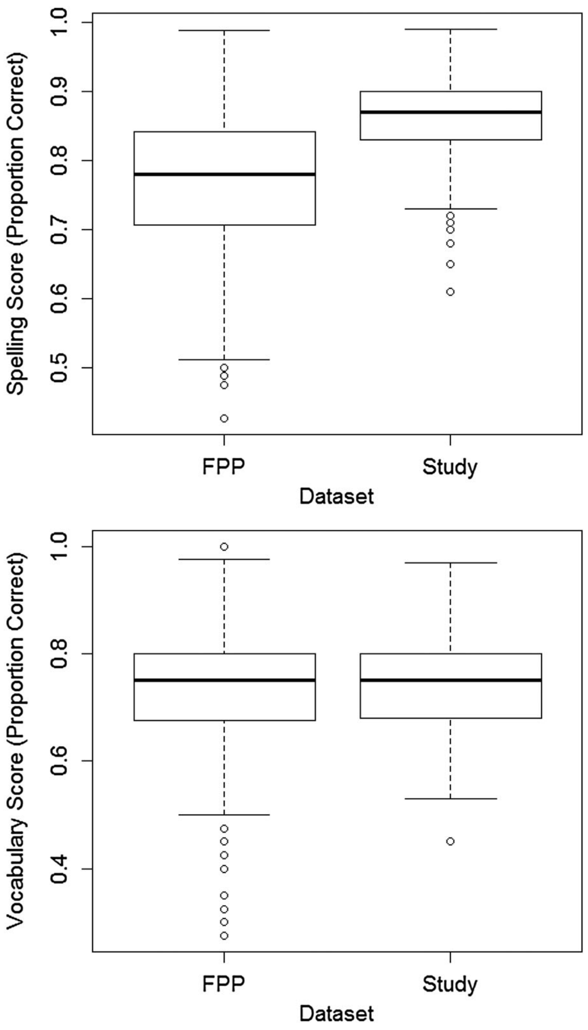


Figure 7. Distributions of spelling (top panel) and vocabulary (bottom panel) scores for the two datasets.

Using the FPP data, we examined the extent to which masked repetition priming was moderated by spelling and vocabulary. Of the original 1015 participants, a total of six participants were excluded, four because they did not possess spelling and vocabulary scores, and two because their scores were more than 3 interquartile ranges below the first quartile. We should also note that in the FPP dataset, the unrelated condition was a pronounceable nonword (e.g., *voctal—DESIGN*), whereas words served as unrelated primes in the present study. Despite this methodological difference, we found converging patterns in the FPP. Specifically, as can be seen in Table 6 (see also Figure 8), both spelling ability and vocabulary

Table 6. LME model estimates (based on zRT) for fixed and random effects for the joint effects of masked repetition priming with spelling ability (top panel) and vocabulary knowledge (bottom panel), from Form Priming Project data.

Random Effects	Variance	SD	*r*
Items			
Intercept	.095	.309	
Prime Relatedness	.008	.088	.370
Participants			
Intercept	.008	.089	
Prime Relatedness	.004	.062	−.120
Fixed Effects	Coefficient	Standard Error	*p*-value
Intercept	−.199	.016	<.001
Prime Relatedness	.256	.012	<.001
Spelling Ability (centred)	−.039	.075	NS
Vocabulary Knowledge (centred)	−.071	.067	NS
Prime Relatedness × Spelling Ability	.260	.121	.032
Random Effects	Variance	SD	*r*
Items			
Intercept	.095	.309	
Prime Relatedness	.008	.087	.370
Participants			
Intercept	.008	.089	
Prime Relatedness	.004	.060	−.130
Fixed Effects	Coefficient	Standard Error	*p*-value
Intercept	−.199	.016	<.001
Prime Relatedness	.256	.012	<.001
Spelling Ability (centred)	−.039	.075	NS
Vocabulary Knowledge (centred)	−.071	.067	NS
Prime Relatedness × Vocabulary Knowledge	.264	.109	.015

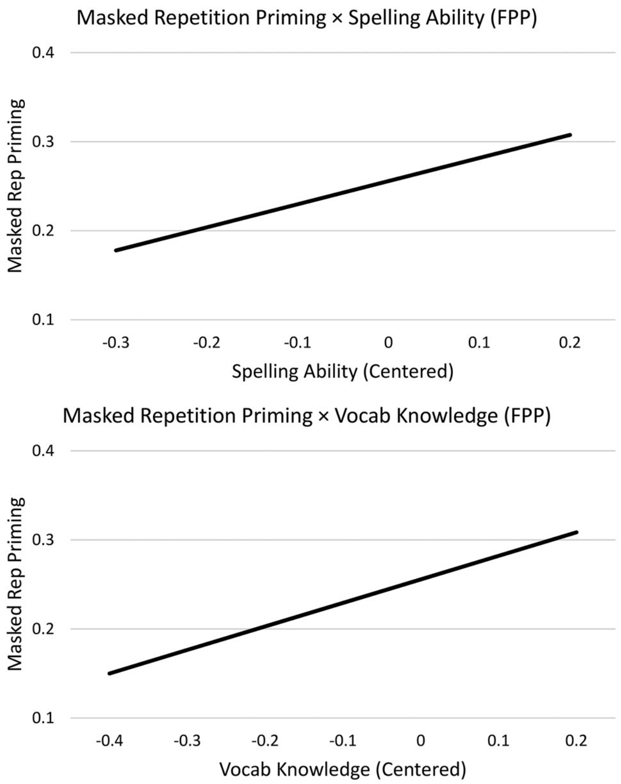


Figure 8. Masked repetition priming as a function of spelling ability (top panel) and vocabulary knowledge (bottom panel). Masked repetition priming is computed by zRT (unrelated)—zRT (related).

knowledge interacted significantly with masked repetition priming; masked repetition priming effects were larger for participants with higher vocabulary and spelling scores.[1]

In summary, the two independent sets of analyses converge on the conclusion that vocabulary knowledge and spelling ability are able to predict the magnitude of masked repetition, but not masked semantic priming, wherein more skilled readers show greater priming than less skilled readers.

Discussion

In the present study, we examined the reliability of the early processes underlying repetition and semantic priming, as well as the extent to which they are moderated by theoretically important individual differences. This study extends earlier work by Stolz et al. (2005) and Waechter et al. (2010), who respectively examined the reliability of semantic and repetition priming when visible primes were used. Our results are straightforward and easy to summarize. One, even though statistically significant effects of masked repetition and semantic priming were obtained at the level of the group (Analysis 1), only masked repetition was associated with moderate within-session and between-session reliability (Analysis 2). These results nicely corroborate and extend the claims made by Stolz and colleagues by showing that their basic findings are replicated even when the influence of controlled processing is minimized. Two, we observed that skilled readers, as reflected by vocabulary knowledge and spelling ability, were associated with larger masked repetition, but not semantic, priming effects. We will now turn to a discussion of these findings.

Masked semantic priming is psychometrically unreliable

The semantic priming effect is one of the most well-known and important effects in the experimental psychology literature. The ubiquity of semantic priming suggests that it implicates fundamental mechanisms of retrieval from memory (McNamara, 1992; Ratcliff & McKoon, 1988). Indeed, the semantic priming effect has played a major role in helping researchers better understand lexical and semantic processing (Collins & Loftus, 1975; Masson, 1995), the distinctions between conscious and unconscious processing (Dagenbach, Carr, & Wilhelmesen, 1989), the nature of automaticity (Neely, 1977; Posner & Snyder, 1975), the basis for language processing deficits in aphasia (Bushell, 1996), and psychopathology (Minzenberg, Ober, & Vinogradov, 2002; Scott, Mogg, & Bradley, 2001).

Importantly, despite the robustness of semantic priming at the group level, this stability does not appear to extend to individual-level performance when priming predominantly reflects automatic

---

[1] Our results may appear inconsistent with aspects of the individual differences analyses reported by Adelman et al. (2014). Specifically, they observed correlations of −.038 and −.020 between repetition priming and spelling, and between repetition priming and vocabulary respectively (see their Table 6). Collectively, these negative correlations suggest that repetition priming effects are *smaller* for more skilled readers. However, neither correlation was statistically significant, and more importantly, the analyses used unpronounceable nonwords (e.g., *cbhaux—DESIGN*) as the unrelated baseline to compute priming. We computed the zero-order correlations between repetition priming, spelling, and vocabulary in the FPP dataset, using both pronounceable and unpronounceable nonword baselines. While correlations were indeed negative when unpronounceable nonwords were used, they became positive when *pronounceable* words were used; furthermore, all corrections were not statistically significant.

processes (Stolz et al., 2005). However, because Stolz et al.'s (2005) primes could be consciously processed, it is difficult to fully rule out the influence of controlled priming mechanisms such as expectancy generation or backward semantic matching. The present study is the first to demonstrate that the instability of semantic priming persists at the individual level when heavily masked primes are used, even when the group-level effect is itself reliable. This further reinforces Stolz et al.'s (2005) claim that the automatic mechanisms (e.g., spreading activation) supporting semantic priming are inherently noisy and uncoordinated. That is, just because *cat* strongly primes *DOG* at some point in time for a participant, there is no assurance that it will do so to the same extent at a later time.

Collectively, these results have important implications for researchers who are interested in using an individual's semantic priming effect to predict other outcomes of interest. For example, individual differences in semantic priming have been used as a tool to investigate associative network dysfunction in thought-disordered schizophrenic patients (Spitzer, Braun, Hermle, & Maier, 1993). However, given the demonstrable unreliability of automatic semantic priming, it is perhaps unsurprising that the literature on semantic priming in schizophrenia has been mixed. Indeed, in their review, Minzenberg et al. (2002) pointed out that the studies which have reported the most consistent findings have tended to feature experimental conditions which emphasize *controlled*, rather than automatic, processes in semantic priming. To recapitulate, the unreliability of automatic semantic priming makes it a poor candidate for studying individual differences in other domains, as it is unclear how much of the between-participants variability is simply measurement error.

That being said, we are not suggesting that individual differences in semantic priming are always meaningless. Stolz et al. (2005) reported that the reliability of semantic priming increased under experimental conditions (e.g., high RP, long SOA) which encouraged greater reliance on controlled mechanisms such as expectancy generation. That is, a participant who strategically relies on expectancy generation in Session 1 is also likely to rely on expectancy generation in Session 2, thereby increasing the correlation in performance between the two sessions. It is probable that the reliability of these controlled processes is accounting for some of the extant findings in the semantic priming literature. For example, Yap et al. (2009) reported that the joint effects of unmasked semantic priming and word-frequency were reliably moderated by vocabulary knowledge. Specifically, participants with more vocabulary knowledge yielded comparable priming for low- and high-frequency words, whereas participants with less vocabulary knowledge produced more priming for low- compared to high-frequency words. Yap et al. (2009) argued that these differences were due to the fact that participants with *less* vocabulary knowledge were relying more on controlled priming mechanisms (e.g., backward semantic matching) that recruit more prime information as target processing becomes more difficult. Likewise, Yap, Hutchison, and Tan (in press) also reported that (unmasked) semantic priming effects were reliably larger for participants who had better reading comprehension and more vocabulary knowledge. In addition, they suggested that highly skilled lexical processors could identify prime words more rapidly, thereby increasing the efficiency of priming mechanisms. In light of the present findings, it seems increasingly likely that semantic priming mechanisms moderated by individual differences are controlled, rather than automatic, in nature.

Masked repetition priming is weakly reliable

Because visual word recognition involves interactive processing between multiple levels of representation (Stolz & Besner, 1996; see Figure 1), it is logically possible for the unreliability in semantic priming (Stolz et al., 2005) to be mediated by earlier, lower-level processes. To test this, Waechter et al. (2010) examined the reliability of repetition priming (which is assumed to tap pre-semantic processing) and found that, unlike semantic priming, repetition priming remained reliable even under experimental contexts which strongly facilitate automatic processing. Although this seems to confirm that the locus of semantic priming's unreliability is indeed limited to semantic-level processing, the primes in Waechter et al.'s (2010) study were not masked, and the influence of controlled processing cannot be fully excluded. The present study is the first to demonstrate that the stability of repetition priming persists at the individual level when heavily masked primes are used. This lends additional support to the idea that automatic processes at the pre-lexical and lexical levels unfold in both a "coordinated" and "consistent" manner (Waechter et al., 2010, p. 553).

We should acknowledge here that the test-retest ($r = .29$) and split-half ($r = .43$) reliability coefficients for masked repetition priming are somewhat low (see Figure 5), relative to traditional recommendations that instruments used in basic research should have a reliability of at least .70 (Nunnally, 1978). However, the magnitude of our correlations, modest though they are, are not out of line with estimates from related studies in the literature. For example, Waechter et al. (2010) obtained a test-retest reliability of .33 for unmasked repetition priming, in their .25 RP, 200 ms SOA condition. When words are presented in isolation, the test-retest reliabilities of participants' sensitivity to lexical characteristics (e.g., word frequency) are also not that high, ranging between .38 and .71 (Yap et al., 2012). Furthermore, consistent with Stolz and colleagues' methodology, we computed test-retest reliability by correlating priming effects for different (but theoretically comparable) prime–target pairs across two sessions. It is possible that presenting *identical* prime–target pairs across both sessions will yield higher estimates of test-retest reliability. This is an interesting question that can be addressed in future research.

With this caveat out of the way, we believe that these results provide some reassurance for researchers who are interested in studying individual differences in masked repetition priming. For example, masked repetition priming has been used to examine differences in implicit memory between Alzheimer's patients and healthy controls (Schnyer, Allen, Kaszniak, & Forster, 1999). Masked repetition priming performance is generally assumed to reflect early orthographic input coding processes, which serve to code visually presented letter strings and to map these representations onto abstract lexical representations (Davis, 2010); our results suggest that these early processes are relatively stable. We think our results also mesh well with findings from the isolated word recognition literature. For example, the study by Yap et al. (2012) reported that participants' sensitivity to the underlying lexical characteristics (frequency, number of letters, neighbourhood density) of words also show low to moderate test-retest reliability (Yap et al., 2012), reinforcing the basic claim that the incoherence of spreading activation processes does not appear to extend to the processing of a word's structural and lexical (i.e., orthographic and phonological) properties.

Individual differences in masked priming

The majority of lexical processing studies continues to rely on measures of group-level performance, despite the implausibility of the uniformity assumption that skilled readers have developed the same cognitive architecture and read the same way (Andrews, 2012). Fortunately, an increasing number of researchers have begun to consider the influence of individual differences on word recognition performance (see Andrews, 2015, for a review). For example, as discussed earlier, Yap et al. (2009) demonstrated that the joint effects of priming and frequency were moderated by vocabulary knowledge. Related to this, Yap et al. (2012) also showed that participants with more vocabulary knowledge were less sensitive to the influence of lexical characteristics (Yap et al., 2012). Such findings can be seen as broadly consistent with Perfetti and Hart's (2001) lexical quality hypothesis, which claims that as readers acquire more experience with written language, the quality of their lexical representations increases. In other words, the more skilled the reader is, the more precise, coherent, and stable their lexical representations are. These skilled readers also become more reliant on relatively automatic lexical processing mechanisms and are consequently less influenced by word characteristics or contextual influences.

In the present study, we investigated the relations between lexical quality (as reflected by vocabulary knowledge and spelling ability) and masked repetition and semantic priming. Unsurprisingly, there was no relation between masked semantic priming and individual differences. Given the unreliability of masked semantic priming, one would not expect it to correlate with any other measure. However, masked repetition priming was associated weakly and positively with spelling ability and vocabulary knowledge. Why are masked repetition priming effects larger for highly skilled readers? We suggest that higher-quality lexical representations afford more precise and quicker identification of the prime stimulus, allowing related primes to provide a greater head-start (Hutchison, Heap, Neely, & Thomas, 2014; Yap et al., in press); this increases the magnitude of repetition priming.

The present results can also be seen as consistent with a study by Andrews and Hersch (2010), which reported that spelling ability moderated masked neighbour priming effects for four-letter targets from low- or high-density orthographic neighbourhoods. A neighbour prime is a word that is one letter different from the target (e.g., *jury*—*FURY*). The most intriguing finding from Andrews and Hersch (2010) is from the experimental condition where participants make lexical decisions to targets with many neighbours. Here, the presentation of a neighbour prime is facilitatory (i.e., faster RTs relative to an unrelated control) for poor spellers but inhibitory (i.e., slower RTs relative to the control) for good spellers. For good spellers, lexical representations are very precise. Therefore, the prime (e.g., *jury*) is able to facilitate activation of only *jury*, while suppressing activation of competing alternatives (e.g., *fury*). For poorer spellers, the prime activates both the target word as well as orthographically similar neighbours, which explains the facilitatory masked neighbourhood priming effect for such participants. Tying this back to the present study, good spellers are associated with larger masked repetition priming effects because the prime mainly activates its lexical representation. In contrast, for poorer spellers, the activation is less precise and is diffused over multiple competing representations. Of course, this is speculative and awaits empirical verification in future work.

Limitations and future directions

Although masked semantic priming was found to be less reliable than masked repetition priming, one could argue that because masked semantic priming is generally so much weaker (~10 ms) than masked repetition priming (40–60 ms), range restriction places an upper limit on the former's reliability. While this is undoubtedly a valid point, it is unlikely that our results can be fully explained by this. Specifically, in Stolz et al. (2005), the test-retest reliability of unmasked semantic priming was also not significant ($r = -.06$), even though the priming effect ($M = 28$ ms) they observed was considerably larger than the present study's. To address this question in a more fine-grained manner, one could examine the reliability of masked form priming effects which are associated with smaller effect sizes. For example, when the medial letters of a target are substituted (e.g., *dewvgn*—*DESIGN*), priming effects are only about a quarter as large as when identical primes are used (i.e., *design*—*DESIGN*) (Adelman et al., 2014). The prediction is that even when masked form priming effects are greatly attenuated, they should still show psychometric reliability.

We also acknowledge that because our experimental design was closely modelled on the two studies by Stolz and colleagues, the nonword targets in the repetition priming condition were preceded by words, which were necessarily unrelated to the nonwords. This implies that repetition trials with related prime–target pairs were always associated with a word response. Future work can address this methodological limitation by ensuring that both words and nonwords are repeated. Nonetheless, it is reassuring that the individual differences analyses in the present study converged with the results from the FPP (Adelman et al., 2014), in which there were repetition conditions for both words and nonwords.

It is worth noting that we have focused on priming performance at the level of the mean. There is some intriguing recent evidence in the literature that individual differences may selectively influence different portions of the RT distribution in masked morphological priming (e.g., Andrews & Lo, 2013) and unmasked semantic priming (Yap et al., 2009). RT distributional analyses could be used in future research to provide additional insights into individual differences in masked priming. Apart from spelling and vocabulary knowledge, researchers could also consider how repetition priming mechanisms are modulated by individual differences in attentional control (AC), where AC refers to the coordination of attention and memory so as to optimize task performance by enhancing task-relevant information (Hutchison, 2007). Finally, repetition priming is only one of many types of form priming (see Adelman et al., 2014, for a comprehensive list), and future research can be directed at exploring individual differences in form priming, when other types of form primes (e.g., transposition, insertion, deletion, substitution) are used.

In conclusion, we found that under masked conditions, repetition priming is reliable, but semantic priming is not, replicating and extending earlier work by Stolz et al. (2005) and Waechter et al. (2010). We also found that skilled readers were associated with larger masked repetition priming effects; this is an intriguing finding that merits more sustained investigation in the future.

Acknowledgements

We thank Jeff Bowers, Jennifer Stolz, and two anonymous reviewers for their very helpful comments on an earlier version of this article.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was carried out as an Master's thesis by L. C. T. under the direction of M. J. Y., and was supported in part by an ODPRT Grant for Research Excellence [grant number R-581-000-144-646] to M. J. Y.

ORCiD

Melvin J. Yap http://orcid.org/0000-0002-5372-4323

References

Adelman, J. S., Johnson, R. L., McCormick, S. E., McKague, M., Kinoshita, S., Bowers, J., … Davis, C. J. (2014). A behavioral database for masked form priming. *Behavior Research Methods*, *46*, 1052–1067.

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, *96*, 703–719.

Andrews, S. (2012). Individual differences in skilled visual word recognition and reading: The role of lexical quality. In J. S. Adelman (Ed.), *Visual word recognition volume 2: Meaning and context, individuals, and development* (pp. 151–172). Hove: Psychology Press.

Andrews, S. (2015). Individual differences among skilled readers: The role of lexical quality. In A. Pollatsek & R. Treiman (Eds.), *Oxford handbook of reading* (pp. 129–148). New York: Oxford University Press.

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General*, *139*, 299–318.

Andrews, S., & Lo, S. (2013). Is morphological priming stronger for transparent than opaque words? It depends on individual differences in spelling and vocabulary. *Journal of Memory and Language*, *68*, 279–296.

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

Balota, D. A., Aschenbrenner, A., & Yap, M. J. (2013). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1563–1571.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., … Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*, 445–459.

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

Becker, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & Cognition*, *8*, 493–512.

Bodner, G. E., & Masson, M. E. J. (1997). Masked repetition priming of words and nonwords: Evidence for a nonlexical basis for priming. *Journal of Memory and Language*, *37*, 268–293.

Bodner, G. E., & Masson, M. E. J. (2001). Prime validity affects masked repetition priming: Evidence for an episodic resource account of priming. *Journal of Memory and Language*, *45*, 616–647.

Bodner, G. E., & Masson, M. E. J. (2003). Beyond spreading activation: An influence of relatedness proportion on masked semantic priming. *Psychonomic Bulletin & Review*, *10*, 645–652.

Bodner, G. E., & Masson, M. E. J. (2014). Memory recruitment: A backward idea about masked priming. In B. Ross (Ed.), *The psychology of learning and motivation, Volume 61* (pp. 179–213). San Diego: Academic Press.

Bowers, J. S. (2000). In defense of abstractionist theories of repetition priming and word identification. *Psychonomic Bulletin & Review*, *7*, 83–99.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977–990.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*, 904–911.

Bush, L. K., Hess, U., & Wolford, G. (1993). Transformations for within-subject designs: A Monte Carlo investigation. *Psychological Bulletin*, *113*, 566–579.

Bushell, C. M. (1996). Dissociated identity and semantic priming in Broca's aphasia: How controlled processing produces inhibitory semantic priming. *Brain and Language*, *55*, 264–288.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

Coltheart, M., Davelaar, E., Jonasson, J. F., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention & Performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

Dagenbach, D., Carr, T. H., & Wilhelmesen, A. (1989). Task-induced strategies and near-threshold priming: Conscious influences on unconscious perception. *Journal of Memory and Language*, *28*, 412–443.

Davis, C. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*, 713–758.

Davis, C., & Kim, J. (2006). Changing circumstance: How flexible is lexical access? In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 207–226). New York: Psychology Press.

De Wit, B., & Kinoshita, S. (2015). The masked semantic priming effect is task dependent: Reconsidering the automatic spreading activation process. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1062–1075.

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

Filzmoser, P., & Gschwandtner, M. (2015). *mvoutlier: Multivariate outlier detection based on robust methods*. R Package Version 2.0.6. [Software]. Retrieved from https://cran.r-project.org/web/packages/mvoutlier/index.html

Forster, K. I. (1998). The pros and cos of masked priming. *Journal of Psycholinguistic Research*, *27*, 203–233.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 680–698.

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

Forster, K. I., Mohan, K., & Hector, J. (2003). The mechanics of masked priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: State of the art* (pp. 3–37). New York: Psychology Press.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.

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*, 1731–1740.

Grossi, G. (2006). Relatedness proportion effects on masked associative priming: An ERP study. *Psychophysiology*, *43*, 21–30.

Hutchison, K. A. (2007). Attentional control and the relatedness proportion effect in semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 645–662.

Hutchison, K. A., Balota, D. A., Cortese, M. J., & Watson, J. M. (2008). Predicting semantic priming at the item level. *The Quarterly Journal of Experimental Psychology*, *61*, 1036–1066.

Hutchison, K. A., Heap, S. J., Neely, J. H., & Thomas, M. A. (2014). Attentional control and asymmetric associative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 844–856.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*, 627–633.

Kinoshita, S. (2006). Additive and interactive effects of word frequency and masked repetition in the lexical decision task. *Psychonomic Bulletin & Review*, *13*, 668–673.

Kinoshita, S., Forster, K. I., & Mozer, M. C. (2008). Unconscious cognition isn't that smart: Modulation of masked repetition priming effect in the word naming task. *Cognition*, *107*, 623–649.

Kinoshita, S., & Lupker, S. J. (2003). *Masked priming: State of the art*. New York: Psychology Press.

Kinoshita, S., Mozer, M. C., & Forster, K. I. (2011). Dynamic adaptation to history of trial difficulty explains the effect of congruency proportion on masked priming. *Journal of Experimental Psychology: General*, *140*, 622–636.

Kinoshita, S., & Norris, D. (2012). Task-dependent masked priming effects in visual word recognition. *Frontiers in Psychology*, *3*, 178. doi: 10.3389/fpsyg.2012.00178

Klapp, S. T. (2007). Nonconscious control mimics a purposeful strategy: Strength of Strooplike interference is automatically modulated by proportion of compatible trials. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1366–1376.

Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition*, *18*, 655–681.

Kopriva, R. J., & Shaw, D. G. (1991). Power estimates: The effect of dependent variable reliability on the power of one-factor ANOVAs. *Educational and Psychological Measurement*, *51*, 585–595.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: Tests in linear mixed effects models*. R Package Version 2.0-30. [Software]. Retrieved from https://cran.r-project.org/web/packages/lmerTest/index.html

Larson-Hall, J. (2016). *A guide to doing statistics in second language research using SPSS and R*. New York: Routledge.

Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. *Psychology and Aging*, *8*, 34–43.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171. doi:10.3389/fpsyg.2015.01171

Lowe, C., & Rabbitt, P. (1998). Test\re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: Theoretical and practical issues. *Neuropsychologia*, *36*, 915–923.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 3–23.

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

Matthews, G., & Harley, T. A. (1993). Effects of extraversion and self-report arousal on semantic priming: A connectionist approach. *Journal of Personality and Social Psychology*, *65*, 735–756.

McNamara, T. P. (1992). Priming and constraints it places on theories of memory and retrieval. *Psychological Review*, *99*, 650–662.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. Hove, UK: Psychology Press.

Minzenberg, M. J., Ober, B. A., & Vinogradov, S. (2002). Semantic priming in schizophrenia: A review and synthesis. *Journal of the International Neuropsychological Society*, *8*, 699–720.

Monahan, P. J., Florentino, R., & Poeppel, D. (2008). Masked repetition priming using magnetoencephalography. *Brain and Language*, *106*, 65–71.

Morgan, C. J. A., Bedford, N. J., & Rossell, S. L. (2006). Evidence of semantic disorganisation using semantic priming in individuals with high schizotypy. *Schizophrenia Research*, *84*, 272–280.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic Priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Memory & Language*, *15*, 1003–1019.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*, 402–407.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Perea, M., & Rosa, E. (2002). Does the proportion of associatively related pairs modulate the associative priming effect at very brief stimulus-onset asynchronies? *Acta Psychologica*, *110*, 103–124.

Perfetti, C. A., & Hart, L. (2001). The lexical bases of comprehension skill. In D. S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). Washington, DC: American Psychological Association.

Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review*, *107*, 786–823.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*, 385–408.

R Development Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Schnyer, D. M., Allen, J. J. B., Kaszniak, A. W., & Forster, K. I. (1999). An event-related potential examination of masked and unmasked repetition priming in Alzheimer's Disease: Implications for theories of implicit memory. *Neuropsychology*, *13*, 323–337.

Scott, K. M., Mogg, K., & Bradley, B. P. (2001). Masked semantic priming of emotional information in subclinical depression. *Cognitive Therapy and Research*, *25*, 505–524.

Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *The Journal of Psychology*, *9*, 371–377.

Spitzer, M., Braun, U., Hermle, L., & Maier, S. (1993). Associative semantic network dysfunction in thought-disordered schizophrenic patients: Direct evidence from indirect semantic priming. *Biological Psychiatry*, *34*, 864–877.

Stolz, J. A., & Besner, D. (1996). Role of set in visual word recognition: Activation and activation blocking as nonautomatic processes. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1166–1177.

Stolz, J. A., Besner, D., & Carr, T. H. (2005). Implications of measures of reliability for theories of priming: Activity in semantic memory is inherently noisy and uncoordinated. *Visual Cognition*, *12*, 284–336.

Stolz, J. A., & Neely, J. H. (1995). When target degradation does and does not enhance semantic context effects in word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 596–611.

Tenpenny, P. L. (1995). Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin & Review*, *2*, 339–363.

Thomas, M. A., Neely, J. H., & O'Connor, P. (2012). When word identification gets tough, retrospective semantic processing comes to the rescue. *Journal of Memory and Language*, *66*, 623–643.

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

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology Section A*, *47*, 631–650.

Waechter, S., Stolz, J. A., & Besner, D. (2010). Visual word recognition: On the reliability of repetition priming. *Visual Cognition*, *18*, 537–558.

Woltz, D. J., & Shute, V. J. (1993). Individual difference in repetition priming and its relationship to declarative knowledge acquisition. *Intelligence*, *17*, 333–359.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, *60*, 502–529.

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

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

Yap, M. J., Hutchison, K. A., & Tan, L. C. (in press). Individual differences in semantic priming performance: Insights from the Semantic Priming Project. In M. N. Jones (Ed.), *Big data in cognitive science: From methods to insights*. New York: Psychology Press.

Yap, M. J., Tse, C.-S., & Balota, D. A. (2009). Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language*, *61*, 303–325.

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