ACCEPTED MANUSCRIPT

This article may not exactly replicate the authoritative document published in the APA journal, *Journal of Experimental Psychology: Human Perception and Performance* (doi: 10.1037/xhp0000372). It is not the copy of record.

Individual differences in automatic semantic priming

Sally Andrews, Steson Lo and Violet Xia

School of Psychology

University of Sydney, Australia

Contact details
Sally Andrews
School of Psychology
University of Sydney
Sydney, NSW 2006
Australia
sally.andrews@sydney.edu.au
Ph: (612) 93518297

KEYWORDS
Semantic priming, lexical processing, reading ability

This research investigated whether automatic semantic priming is modulated by individual differences in lexical proficiency. A sample of 89 skilled readers, assessed on reading comprehension, vocabulary and spelling ability, were tested in a semantic categorisation task that required classification of words as animals or non-animals. Target words were preceded by brief (50 ms) masked semantic primes that were either congruent or incongruent with the category of the target. Congruent primes were also selected to be either high (e.g., *hawk EAGLE, pistol RIFLE*) or low (e.g., *mole EAGLE, boots RIFLE*) in feature overlap with the target. ‘Overall proficiency’, indexed by high performance on both a ‘semantic composite’ measure of reading comprehension and vocabulary and a ‘spelling composite’, predicted stronger congruence priming from both high and low feature overlap primes for animal exemplars, but only predicted priming from low overlap primes for non-exemplars. Classification of high frequency non-exemplars was also significantly modulated by an independent ‘spelling-meaning’ factor, indexed by differences between the semantic and spelling composites, which appeared to tap sensitivity to semantic relative to orthographic feature overlap between the prime and target. These findings show that higher lexical proficiency predicts stronger automatic semantic priming and suggest that individual differences in lexical quality modulate the division of labor between orthographic and semantic processing in early lexical retrieval.

The assumption that meaning is extracted rapidly and automatically from written words underpins interpretation of a variety of cognitive psychological phenomena across a range of domains. For example, variants of the Stroop effect are used to assess automatic cognitive biases relevant to psychopathology (e.g., Gorlin & Teachman, 2015); false memory effects are used to index automatic semantic processing relevant to forensic contexts (e.g., Marche, Brainerd & Reyna, 2010); and implicit association tests have become a standard diagnostic of automatically activated social attitudes and prejudices (e.g., Monteith, Woodcock & Gulker, 2013). Despite this widespread acceptance by the broader psychological community, questions about how early semantic information becomes available, and the extent to which it is activated automatically and independently of conscious awareness, remain a focus of debate in cognitive science (e.g., de Wit & Kinoshita, 2015; Kouider & Dehaene, 2007).

This debate has been fuelled by difficulties in defining the central constructs (e.g., Holender, 1986; Marcel, 1983) and developing effective methods for tapping them (e.g., Kouider & Dehaene, 2007). The operational definition of automatic semantic priming adopted here is that it occurs under conditions in which participants do not, and usually cannot, identify the prime and there is no incentive to do so because prime-target relatedness is irrelevant to the required response – for example, targets requiring 'yes' and 'no' responses are equally likely to be preceded by a semantically related prime. These conditions appear to be met by the masked priming paradigm (Forster & Davis, 1984), in which a brief lowercase prime is sandwiched between 500 ms presentations of a forward mask (usually nonlexical symbols e.g., #####) and the uppercase target word, as long the prime is presented for less than 60 ms (Forster, Mohan & Hector, 2003) and the stimulus conditions are arranged to ensure that primes do not predict the required response.

Under these conditions, orthographically similar primes consistently produce robust masked priming effects on lexical decision responses to word, but not nonword targets. Forster & Davis, (1984) interpreted this as evidence that masked primes pre-activate existing lexical representations. However, the selectivity of masked priming depends on the specific decision requirements of the task, as predicted by models that attribute masked priming to evidence accumulation and decision processes (e.g., Norris & Kinoshita, 2008; Kinoshita & Norris, 2009) rather than pre-activation of lexical representations.

Semantic influences on masked priming

In contrast to orthographic priming, masked semantic priming has proved relatively elusive (e.g., Gollan, Forster & Frost, 1997; Holender, 1986). This may be because any benefit of the semantic relationship between the prime and target is counteracted by the orthographic differences between them. The masked morphological priming paradigm circumvents this problem because morphologically related items usually share both orthographic and semantic features. It is therefore possible to assess whether semantic similarity boosts priming under conditions in which orthographic similarity is controlled.

Direct evidence of early semantic activation in this paradigm is provided by comparisons of *transparent* morphological pairs (e.g., *hunter-HUNT*), which are both morphologically and semantically related; with *pseudomorphemic* pairs (e.g., *corner-CORN*) that share the same apparent morphological structure but are not semantically related; and *control* pairs that are equally orthographically similar (e.g., *turnip TURN)* but do not have a morphological structure (e.g., -*ip* is not a suffix). A meta-analysis of the general pattern of data found in lexical decision tasks comparing these conditions revealed a numerically graded effect with stronger masked priming for transparent than opaque pairs, which both produced stronger priming than orthographically matched control pairs (Rastle & Davis, 2008). However, a subsequent, extended meta-analysis only reported significant priming for transparent pairs (Feldman, O'Connor, Moscoso del Prado Martin, 2009). A variety of methodological and statistical ambiguities have been proposed to account for such discrepancies, fuelling ongoing debate about the precise timecourse of orthographic and semantic processing implicated by these findings. Some researchers interpret the evidence as supporting *form-first* models which assume that orthographically-based segmentation into morphemic units precedes activation of shared semantic attributes (e.g., Davis & Rastle, 2010), while others just as adamantly argue that this view is challenged by evidence of early semantic influences (e.g., Feldman et al., 2009; Feldman, Milin, Moscoso del Prado and O'Connor, 2015).

Individual differences in masked priming

A possible resolution to these contradictions suggested by our recent data (Andrews & Lo, 2013) is that masked morphological priming effects in the average data for samples of skilled readers conflate two different patterns of priming that reflect systematic individual differences in relative sensitivity to orthographic and semantic priming. Andrews and Lo's (2013) average data replicated the graded effect of morphological similarity reported in many previous studies (e.g., Rastle & Davis, 2008). However, linear mixed analyses revealed that the pattern of morphological priming was significantly modulated by individual differences on a *'spelling-meaning' factor* that tapped discrepancies between performance on independent tests of vocabulary and spelling ability, partialling out effects of overall proficiency (Andrews, 2015). Individuals with a *semantic profile*, defined by relatively higher scores on tests of vocabulary than spelling, demonstrated robust priming for transparently related pairs but virtually no priming from equally orthographically similar pseudomorphemic and control primes. By contrast, the opposite *orthographic profile*, of higher spelling than vocabulary, showed strong and equivalent priming from both transparent and pseudomorphemic primes, the pattern predicted by form-first models. These findings suggest that skilled readers vary in the 'division of labor' between orthographic and semantic processing during early lexical retrieval (Harm & Seidenberg, 2004).

This conclusion is consistent with evidence of individual differences in masked orthographic priming. Andrews and Lo (2012) found that *overall proficiency*, assessed by the combination of higher reading comprehension, vocabulary and spelling ability predicted faster overall lexical classification and stronger sensitivity to both sublexical facilitation from nonword primes and lexical inhibition from word primes. However, over and above the effects of overall proficiency, a *spelling-meaning* factor indexed by the *difference* between spelling and comprehension/vocabulary ability also captured unique variance: the orthographic profile of higher spelling than comprehension was associated with inhibitory priming from masked transposed letter primes (e.g., *clam CALM; cilp CLIP*), suggesting that these readers were more sensitive to letter position or order. Additional analyses using the separate measures of reading comprehension, vocabulary and spelling as predictors showed that spelling was the only individual predictor that accounted for significant unique variance in priming: Good spellers showed inhibitory priming, particularly for TL primes, while poor spellers showed facilitatory priming from both TL and neighbor primes.

These differences among skilled readers are consistent with Perfetti's (2007) *lexical quality* framework (Andrews, 2015). According to this view, superior reading skill depends on precisely specified orthographic representations that are tightly interconnected with their associated phonological and semantic codes, so that written words synchronously and coherently activate a word's full lexical identity. Such high quality lexical representations support autonomous, perceptually-driven lexical retrieval, which is minimally reliant on context. This bottom-up reading strategy is thought to be optimally efficient because it allows attentional resources to be directed towards the higher-order processes required for comprehension (e.g., Stanovich, 2000). But it is not the only route to successful comprehension. Readers who fail to develop precise orthographic representations can compensate by relying more heavily on semantic context to identify words (Andrews & Bond, 2009). The selective priming for transparently related morphological pairs shown by individuals with the semantic profile of relatively higher vocabulary than spelling ability (Andrews & Lo, 2013) suggests that, although spelling ability is the most potent predictor of masked orthographic priming, semantically-focused measures such as reading comprehension and vocabulary may be the best predictors of early sensitivity to semantic attributes of masked primes.

Consistent with this view, higher vocabulary participants have been reported to show additive effects of frequency and semantic priming that contrasted with the stronger priming for low frequency words shown by those of lower vocabulary (Yap, Tse & Balota, 2009). These findings were interpreted as suggesting that "the lexical processing system of readers with greater vocabulary knowledge … [is] more modular in nature" (Yap et al., 2009, p. 321), as assumed by the lexical quality framework. However, this study used visible primes, which may encourage the adoption of post-lexical strategies, and did not include additional measures of written language proficiency, such as spelling, to assess whether the effects were specific to vocabulary.

A recent study of morphological priming that included measures of both vocabulary and spelling ability (Feldman et al., 2015, Exp. 3) failed to replicate Andrews and Lo's (2013) evidence that these variables modulate priming, using shorter (34 ms) prime durations than their 50 ms presentation. This may indicate that individual differences gradually emerge in the course of prime processing. This possibility is consistent with Andrews and Lo's RT distribution analyses which found that the differences in patterns of morphological priming between readers with the 'orthographic' and 'semantic' profiles were not evident in the fastest RTs but emerged across the RT distribution. Feldman et al.'s (2015, Exp 1) direct comparison of average morphological priming effects at different prime durations confirmed that the effects of 34 ms primes were very small but increased with prime duration. At earliest stages of prime processing tapped by very brief masked primes, individual differences may not yet be able to be reliably detected.

Using category congruence effects to probe automatic semantic priming

The present study was designed to confirm and extend previous evidence of individual differences in the balance between orthographic and semantic processing in early lexical retrieval by investigating sensitivity to category congruence in a semantic categorisation task.

In contrast to the weak semantic priming effects generally found in masked priming lexical decision tasks, brief masked primes yield robust *category congruence effects* (Dehaene & Naccache,

2001) in tasks requiring semantic categorisation, i.e., targets are classified as members of a pre-defined semantic category faster when they are preceded by masked primes from the same category rather than a different category. These effects appear to provide strong evidence that "subliminal stimuli are unconsciously processed up to a semantic level, in principle not different from conscious cognition" (Kunde, Kiesel & Hoffmann, 2003). However, most investigations of category congruence have used *narrow categories* composed of a small set of items that are typically presented many times during the experiment – most prominently, the small, finite set of single-digit numbers (Dehaene et al., 1998; Dehaene & Naccache, 2001). As elaborated below, such categories may yield different patterns of sensitivity to category congruence than *broad categories,* like 'animal' (Forster, 2004).

Investigations of broad categories raise questions about how semantic relatedness should be defined. Many words that are rated as being semantically related share an associative rather than semantic relationship (e.g., *mouse-cheese*) and the two are often confounded (e.g., *dog-cat*). Associative relationships may be represented in the lexical network due to their co-occurrence in spoken and written language (Witzel & Forster, 2014). Early findings that automatic semantic priming effects appeared to be limited to associatively related pairs were therefore interpreted as showing that apparent evidence of early semantic activation may actually arise at the lexical or word form level (Shelton & Martin, 1992).

However, non-associates have been found to produce significant semantic priming when they share many features. When semantic relatedness was defined by norms for degree of feature overlap derived from a feature-listing task, significant semantic priming occurred for prime-target pairs with high feature overlap (e.g., *hawk EAGLE; pistol RIFLE*) in both lexical decision (McRae & Boisvert, 1998) and semantic categorisation tasks (McRae, de Sa & Seidenberg, 1997), using short (200 ms) primes, which are assumed to tap automatic processes (Neely, 1991).

When primes are clearly visible, like the 200 ms primes used in the studies described above, it is difficult to rule out the possibility that semantic influences arise from conscious decision strategies based on prime-target relatedness (de Wit & Kinoshita, 2015, Neely, 1991). Stronger evidence that category congruence effects reflect automatic semantic activation is provided by studies using masked primes in semantic categorisation tasks. There is clear evidence for masked congruence effects on evaluative judgements (e.g., Greenwald, Draine & Abrahams, 1996) and number classification tasks (Dehaene et al., 1998) using narrow categories that require frequent repetitions of the same stimuli within the experiment. However, these effects appear to be at least partly due to intra-experimental learning of stimulus-response associations (Damian, 2001). Even when this is precluded by never presenting the masked primes as visible targets (e.g., Kunde et al., 2003), categorisation decisions about small categories such as single-digit numbers may be made by searching just the set of category members.

Consistent with this possibility, Forster (2004) found that *no* responses to non-exemplars of narrow categories (e.g., numbers, months, human body parts) did not show the word frequency effects typically used as a diagnostic of lexical access, while non-exemplar decisions for the broad category of *animal* produced a robust frequency effect. Forster also reported that *yes* responses to exemplars of narrow categories showed a category congruence effect (e.g., faster responses to *January AUGUST* than *machine AUGUST*) that was not evident for broad *animal* categorisation responses (e.g., *mole EAGLE vs boot EAGLE*; Forster et al., 2003). He interpreted this as evidence that semantic categorisation responses for narrow categories rely on "a special access mode (category search)…. without access to the lexical entry for the target word" while decisions about broad categories "are based on semantic properties retrieved from the lexical entry for the target word" (Forster, 2004, p. 276).

A direct comparison of broad and narrow categories by Quinn and Kinoshita (2008) found significant frequency effects for non-exemplars in both narrow and broad category decision tasks, contradicting Forster's (2004) findings. Also in contrast to Forster et al.'s (2003) data, they found category congruence effects for exemplars from the broad *animal* category but it was restricted to congruent primes that were high in feature overlap (e.g., *hawk EAGLE)* and did not occur for category congruent primes that shared few features with the target *(e.g., mole EAGLE*). In contrast, non-exemplars showed significant priming from both high and low overlap category congruent primes (e.g., *pistol RIFLE, boots RIFLE*). Quinn and Kinoshita argued that these data contradicted Forster's (2004) claim that different strategies are used to make decisions about narrow and broad categories and instead supported a 'selective feature monitoring' account in which semantic categorisation decisions are always made by "monitoring the semantic features activated following lexical access … [but] broad and narrow categories differ in the features that are selected to be monitored" (Quinn and Kinoshita, 2008, p. 300). Narrow categories allow a small set of features to be selected, while broad categories require monitoring of a large feature set. Judgements of both exemplars and non-exemplars for broad categories therefore benefit from a category congruent prime, but only if it shares category-relevant features.

The present study

The aim of the present study was to investigate whether category congruence effects in a semantic categorisation task for the broad category 'animal' are sensitive to individual differences amongst skilled readers. The design was modelled on Quinn and Kinoshita (2008, Exp. 1) which compared category congruence effects for high and low overlap congruent primes with category incongruent primes for exemplar and non-exemplar targets. To assess the contribution of lexical processing, frequency was also manipulated within the non-exemplars.

The critical items were identical to those used by Quinn and Kinoshita (2008) so the average data were expected to replicate their finding of significantly larger category congruence effects for high overlap than low overlap congruent primes, for both exemplars and non-exemplars, and significant frequency effects for non-exemplar classification.

Individual differences in the division of labor between orthographic and semantic processing in early lexical retrieval, paralleling those found in Andrews and Lo's (2013) morphological priming data, would be supported by evidence that congruence priming is predicted by semantic measures of comprehension/vocabulary rather than by the spelling ability measures that predict masked orthographic priming. Indeed, if the orthographic profile of higher spelling than comprehension/vocabulary is associated with a form-based access process involving limited activation of semantic information, as Andrews and Lo's (2013) morphological priming data implied, these individuals may be insensitive to category congruence, regardless of the degree of feature overlap.

Method

Participants. The participants were 95 undergraduate students (66 females; average age=21 years) from Sydney University, Australia, who participated in the experiment for course credit. All had normal or corrected-to-normal vision and had been speaking English by at least age 5.

Measures of individual differences. Four measures of written language proficiency were administered. Reading Comprehension and Vocabulary were assessed using a short version of the Nelson-Denny Reading Test1[1] (Brown, Fishco, & Hanna, 1993). Spelling ability was assessed using the measures of spelling dictation and spelling recognition described by Andrews and Hersch (2010), which have been shown to have high test-retest reliability (r>.89) and to predict performance in a range of tasks (Andrews, 2012). We report analyses based on two measures. A *'semantic'* composite score was calculated by averaging the standardized scores for reading comprehension and vocabulary, which were quite highly correlated (r=.64). This composite score is similar to the total score derived from the Nelson-Denny used in previous studies of individual differences among skilled readers (e.g., Ashby et al., 2005, Perfetti, 2007) and is therefore referred to as *ZTotal*. A composite spelling score (*ZSpell*) was created by averaging standard scores for the two highly correlated measures of spelling (r=.82). Presumably reflecting the restricted range of ability within this university sample, and consistent with our previous studies, these composite measures were only moderately correlated (r=0.38).

Experimental design and stimuli. The critical stimuli were those constructed by Quinn and Kinoshita (2008, Experiment 1). The critical target words were 45 animal names that served as exemplars and 90 non-animal non-exemplars. The exemplar targets had a low average frequency (mean= 4.8/million[2], range 0-40) and an average length of 6.6 letters. Half of the non-exemplars were high frequency (mean: 97.3, range 40-443) and half were low frequency (mean: 5.4, range 1-15). High and low frequency non-exemplars differed significantly in mean length (5.3 vs 6.5; t=3.81, p<.001).

Three primes were selected for each exemplar and nonexemplar target, none of which overlapped orthographically or phonologically with the target. Both High and Low overlap category congruent primes were from the same category (i.e., animal or non-animal) as the target, but differed in the average number of shared semantic features with the target assessed by McRae, Cree, Seidenberg, and McNorgan's (2005) semantic feature production norms. High overlap congruent primes (e.g., *hawk-EAGLE*, *pistol-RIFLE*) shared at least five features with the target word (mean= 8.4) while low overlap primes shared very few features. Category incongruent prime shared neither category membership nor semantic features with the target (*knee-EAGLE*, *camel-RIFLE*). The primes were matched with the target on length. To avoid response bias by presenting an equal number of exemplar and nonexemplar targets, 45 filler exemplar items were selected with similar properties to the critical exemplars (mean frequency and length of 4.4/million and 6.8 letters respectively) and equally divided between each of the prime types.

The ultimate test materials consisted of three counterbalanced lists of 180 prime-target stimuli each containing 50% exemplars and 50% non-exemplars. Each list contained only one instance of each target but, across lists, all targets occurred in all three prime conditions. Each list was presented to approximately equal numbers of participants.

Procedure Participants were tested individually in a single session lasting about 60 minutes. The animal categorization task and a separate lexical decision task not reported here were interleaved with the language proficiency tests.

Following the procedure used by Quinn & Kinoshita (2008), each semantic categorisation trial consisted of three successive displays presented in the centre of a computer screen: a forward mask (#########) for 500 ms, the prime in lower case for 50 ms and the target in upper case for 500 ms. Following the response (or a timeout of 1500 ms), there was a 300 ms blank screen until initiation of the next trial. All stimuli were presented in white 12-pt courier font on a black background. The stimulus presentation and timing of responses and errors was controlled by the DMDX system (Forster & Forster, 2003). Participants were instructed to classify the upper case targets as animal or nonanimal, by pressing a YES (with their dominant hand) or NO button on a button box. No mention was made of the primes. Participants were randomly allocated to one of the three lists and presented with an individually randomised sequence of items.

Results and Discussion

Six participants with error rates above 20% were excluded from analysis. The data for the remaining 89 participants were cleaned to remove incorrect responses and anticipatory responses of less than 200 ms (5.9% of total). The resulting average reaction time (RT) and percentage error rate (%ER) for the total sample in each condition are presented in Table 1.

The data were analysed by testing linear mixed effect (LME) models using the lme4 package (Bates, Maechler, Bolker & Walker, 2015). This analysis approach is ideal for the present context because it allows the effects of crossed subject and item factors to be considered within a single analysis. In addition to simultaneously evaluating the generality of fixed effects across subjects and items (Baayen, Davidson & Bates, 2008), this method allows assessment of interactions between stimulus factors and continuous individual difference measures while statistically controlling for a number of sources of extraneous variability.

The analyses treated subjects and items as crossed random factors and were conducted on inverse RT, because this best approximated normal distribution assumptions (Box & Cox, 1964) [3], but the results reported in tables and figures are converted back to raw RT to facilitate interpretation. To facilitate comparison with previous investigations of category congruence by allowing all pairwise comparisons between the three priming conditions to be tested, two separate models were tested for the exemplar and non-exemplar data. In the first set of models, the effects of prime type were assessed using normalised sum contrasts that separately compared the high and low feature overlap conditions with the category incongruent condition. The second set of models included the remaining pairwise comparison of the high feature overlap and low feature overlap congruent primes and an orthogonal contrast comparing the average of the two congruent prime conditions with the incongruent condition[4]. A generalised matrix inversion was conducted on both sets of contrast weights in order to yield interpretable main effects (Venables & Ripley, 2002).

Autocorrelation in the residuals was controlled by including participants' RT on the previous trial and trial sequence number as fixed effects in each model (e.g., Kinoshita, Mozer & Forster, 2011; Bates, Kliegl, Vasishth & Baayen, 2015). Analyses of the non-exemplar data also included a normalised sum contrast testing the difference between high and low frequency targets, and its interaction with the effects of prime condition.

The contribution of individual differences was assessed using the approach adopted by Andrews and Lo (2012, 2013) of conducting a principal components (PC) analysis on the composite individual differences measures of ZTotal and ZSpell to extract orthogonal estimates of their shared and unique variance. The first principal component (PC1) had high positive correlations with both ZTotal (r=0.88) and ZSpell (r=0.77), such that participants who were high on this dimension demonstrated high *overall proficiency* on both composite measures. The second principal component (PC2), which was orthogonal to PC1, captured the remaining 31% of variance between ZTotal and ZSpell. This component was positively correlated with ZTotal (r=0.47) and negatively correlated with ZSpell (r= -0.64), indicating that individuals with higher scores on PC2 performed better on the semantic composite of comprehension/vocabulary than on the spelling composite while low scores on this factor reflected the opposite asymmetry. Main effects and interactions involving these principal components were included as fixed effects in order to assess whether individual differences in written language proficiency systematically modulated category congruence effects for both exemplar and non-exemplar targets. [5]

Each model was estimated using the 'maximal' random effect structure as recommended by Barr, Levy, Scheepers and Tily (2013) [6]. Following established convention (e.g., Kliegl, Masson & Richter, 2010), effects were considered to be statistically significant if their estimated magnitude exceeded two standard errors (e.g., $|t|>2$). The full models are reported in Appendix A so the summary of results below only reports the t values for relevant effects.

TABLE 1 ABOUT HERE

Average data

The average exemplar and non-exemplar data presented in Table 1 replicated the pattern reported by Quinn and Kinoshita (2008). Analyses of 'yes' responses to exemplar targets revealed significant priming from high overlap congruent primes relative to both incongruent primes, t= -3.81, and low overlap congruent primes, t= 3.66, which did not significantly differ, t= -0.47.

'No' responses were significantly faster for high than low frequency non-exemplars, t= -4.86. In contrast to exemplars, responses to non-exemplars showed a graded priming effect with faster responses to targets primed by high overlap than low overlap congruent primes, t=4.37, and faster responses to targets preceded by low overlap congruent than incongruent primes, t= -4.95. Neither effect significantly interacted with target frequency, both $|t|< 1.43$.

Individual differences in congruence priming

The novel question focused on in the present research was whether these average effects were modulated by individual differences. Models of the exemplar data showed that the semantic priming effects in the averaged data interacted significantly with PC1. As depicted in the left panel of Figure 1, higher scores on PC1 were associated with faster average RT, t= -3.35, and stronger congruence priming from both high overlap, t= -2.3, and low overlap, t= -2.03, congruent primes. None of the main or interaction effects involving PC2 was significant, all |t|<1.63. Thus, the significant effect of congruence priming found for exemplars in the average data was best accounted for by the *combination* of superior comprehension/vocabulary *and* spelling ability.

FIGURE 1 ABOUT HERE

Similarly, models tested on the non-exemplar data showed that the PC1 dimension of overall proficiency significantly modulated the category congruence effects observed in the averaged results. As summarised in the right panel of Figure 1, higher scores on PC1 were associated with significantly faster average RT, t= -4.31, that was significantly more pronounced for targets preceded by low overlap congruent primes than incongruent primes, t= -3.45. PC1 did not significantly modulate the decrease in RT associated with high overlap congruent primes relative to either low overlap, t= -1.86, or incongruent primes, t= -1.61. The results therefore indicate that for non-exemplar targets, high overlap primes yielded statistically equivalent category congruence priming regardless of proficiency.

Although there was no significant main effect of PC2 on RT for the non-exemplars, t= -0.59, this factor significantly modulated the difference in RT between incongruent and high overlap congruent primes through its significant three-way interaction with target frequency, t= -2.14. To determine the source of this interaction, separate analyses were conducted of data for the high and low frequency non-exemplar targets, which are summarised in Figure 2. The analysis of low frequency targets showed that the significant category congruence effect for high overlap primes relative to incongruent primes, t= -5.09, did not significantly interact with PC2, t= 0.22. However, for high frequency targets, higher scores on PC2 were associated with significantly larger category congruence effects from high overlap relative to incongruent primes, t= -2.71 and marginally significantly more priming from high than low overlap congruent primes, t=1.97. Target frequency did not significantly qualify any of the interactions between congruence priming and PC1, all |t|<1.

FIGURE 2 ABOUT HERE

Thus, the significant congruence effects observed in the average data were principally due to participants high in overall proficiency on both the semantic and spelling composite measures. Participants below average on this overall proficiency dimension showed no congruence priming for exemplar targets. They did, however, show congruence priming for non-exemplars, but only from high overlap congruent primes, while higher proficiency participants showed congruence priming for both high and low overlap primes. The only significant effect of PC2, the discrepancy between the semantic and spelling composite scores, was on high frequency non-exemplar targets and occurred because the category congruence effect for high frequency targets was principally due to participants with higher comprehension/vocabulary than spelling ability.

Analysis using individual composite scores as predictors. Using principal component scores to index individual differences allowed us to tap independent dimensions of variation and to capture the effects of overall proficiency. However, these measures obscure the effects of the individual composite measures. A final set of analyses therefore used the continuous measures of ZTotal and ZSpell as predictors in the LME models. Since these measures were only moderately correlated their joint inclusion does not violate collinearity constraints.

FIGURE 3 ABOUT HERE

Models of the exemplar data showed that the significant semantic priming effect in the average data was entirely due to the ZTotal measure of comprehension/vocabulary. As depicted in left panel of Figure 3, this score was associated with faster average RT, t=-3.04, and stronger congruence priming from both high overlap, t=-2.61, and low overlap, t=-2.40, congruent primes. None of the main effects or interactions involving ZSpell was significant, all t<1. Thus, the significant effect of congruence priming found for exemplars in the average data was best accounted for by the semantic composite of reading comprehension and vocabulary.

The ZTotal measure was also the best predictor of average priming for non-exemplars, As summarised in the right panel of Figure 3, average RT was significantly faster for individuals with higher scores on the ZTotal composite, $t=-3.21$, but did not significantly vary with ZSpell, $t=-1.50$. Higher ZTotal was also associated with stronger congruence effects for both high overlap, $t=-2.42$, and low overlap, $t=-2.94$, congruent primes.

ZSpell did not yield significant interactions with the average congruence effects, $|t| < 1.62$, but its effects were qualified by target frequency: the difference between congruent high overlap primes and both incongruent primes, and low overlap congruent primes participated in significant 3-way interactions with spelling and target frequency, $t=2.07,-2.17$ respectively. Figure 4 shows that these interactions were driven by the large congruence priming effect shown by poor spellers for high frequency, but not low frequency, non-exemplar targets, which contrasted with the complete absence of difference between high and low overlap primes shown by good spellers. Frequency did not significantly modulate the interactions of comprehension/vocabulary with congruence effects, all $|t|<1.38$.

FIGURE 4 ABOUT HERE

These results complement the outcomes of the PC-based analyses by showing that the effects of overall proficiency on congruence priming were principally due to the semantic composite of reading comprehension and vocabulary, while the differential effect of PC2 on priming for high and low frequency exemplar targets was driven by spelling ability.

General Discussion

The average results replicated Quinn and Kinoshita's (2008) evidence that skilled readers' semantic categorisation decisions for the broad category 'animal' show category congruence effects that are enhanced when the masked prime shares many features with the target. Paralleling their findings, the average priming effects were somewhat different for exemplar and non-exemplar targets: ‘Yes’ responses to animal targets only benefited from high overlap primes, while ‘No’ responses to non-animals showed a graded priming effect that was significant even for low overlap primes. Also replicating Quinn & Kinoshita’s data, non-exemplar classifications showed a significant effect of word frequency that did not interact with priming, despite the greater power afforded by our larger sample size.

These finding converge with earlier evidence (Forster, 2004) showing that brief masked primes which were never presented in clear view or repeated in the experimental context are processed semantically, at least under the task demands of semantic categorisation. Like Quinn and Kinoshita’s results, the present data demonstrate that the effects are semantic in origin by showing significant effects of the extent of featural overlap between the prime and target. These average data support Quinn and Kinoshita’s claim that “the category congruence effect is a ‘semantic priming effect’….[implying that] semantic features are activated autonomously as a result of lexical access driven by perceptual input” (2008, p. 295-6) . However, significant interactions with our measures of individual differences revealed that skilled readers vary in their sensitivity to this index of automatic semantic priming.

Congruence effects for exemplars were limited to participants high in overall proficiency, defined by superior performance on the combined semantic and spelling composite measures (see Figure 1). Analyses using the individual composite predictors revealed that the semantic composite of comprehension/vocabulary was the most potent predictor of congruence priming (see Figure 3). This was also generally true for non-exemplars except that average priming from high overlap congruent primes for these items did not interact with overall proficiency. However, interactions with target frequency revealed that the *difference* between the semantic and spelling composite indexed by PC2 modulated the extent of priming from high overlap primes for high frequency targets (see Figure 2) and the composite-based analyses showed that this reflected selective effects of spelling ability on congruence priming in this condition (see Figure 4).

The mechanisms of masked congruence priming

Quinn and Kinoshita interpreted their data as showing that word meanings are represented as a network of distributed features (e.g., McRae & Boisvert, 1998) and semantic categorisation "is accomplished by monitoring semantic features activated following lexical access….but the selected features are not completely diagnostic of category membership, that is, there is no set of features corresponding to 'animalness' … that could be monitored" (2008, p. 301). These conclusions were based on the presence of frequency effects for non-exemplars, which showed that decisions were based on lexical retrieval rather than search of a single category (Forster, 2004), and the graded semantic effects observed for non-exemplars, which contradicted Carreiras, Perea and Grainger's (1997) claim that decisions are based on a single 'animalness' feature: if such a feature was available, non-exemplar responses should be insensitive to feature overlap because decisions could be made by simply monitoring 'animalness'.

The present results confirm both of these findings and extend them by showing that category congruence priming depends on overall proficiency and that the interactions with proficiency also differ between exemplars and non-exemplars: congruence priming of exemplars by both high and low overlap primes increased with higher proficiency but for non-exemplars proficiency only predicted priming for low overlap congruent primes. Somewhat surprisingly, this was because, on average, high overlap congruent primes produced priming across all levels of proficiency and, for high frequency targets, this effect was primarily due to poorer spellers.

This novel evidence of individual differences in congruence priming may provide insight into the rather puzzling differences between exemplars and non-exemplars observed in both Quinn and Kinoshita's (2008) data and our average results. Why does a category congruent prime that is low in semantic feature overlap yield significant benefit for non-exemplars but not non-exemplars –

e.g., why does *lunch BRIDGE* yield significant category priming when *horse TIGER* does not? From the perspective of Quinn and Kinoshita's feature monitoring account, a graded effect of feature overlap seems more likely for the category of 'animal', in which average feature overlap is relatively high, than the more diffuse category of 'non-animal'. However, as Quinn and Kinoshita note, it is difficult to make clear predictions for non-exemplars without specifying how 'no' decisions are made in binary semantic categorisation tasks.

The association between overall proficiency and congruence priming sheds some light on these issues. Congruence priming increased with overall proficiency, indicating that the 'gradedness' of congruence effects reflects variation between individuals as well as items. When both sources of variance are considered, congruence priming effects show the graded effects predicted by the semantic feature overlap account for both exemplars and non-exemplars, but in different ways. The increase in congruence priming for exemplar targets from both high and low overlap category primes suggests that higher overall proficiency is associated with faster access to semantic features, or with monitoring a broader range of semantic features. By contrast, non-exemplar targets showed a graded effect of semantic overlap in the average data, but the interactions with individual differences suggested that it reflects the combined effects of two independent factors.

Priming from low overlap primes increased with overall proficiency, as it did for exemplar targets. However the priming effects for high overlap primes were modulated by both target frequency and spelling ability. The PC-based analyses confirmed that this selective interaction of congruence with target frequency reflected an independent dimension of variability between individuals that was associated with PC2 rather than the overall proficiency captured by PC1. The composite-based analyses, which isolated unique variance associated with each individual composite score, showed that poorer spelling was associated with an unexpectedly large congruence priming effect that was restricted to high frequency targets and did not extend to low overlap primes, while better spellers showed no benefit for high relative to low overlap primes for high frequency targets. The independence of this contribution to congruence priming from the effects of overall proficiency is highlighted by the fact that it reflects the opposite relationship between priming and skill. Higher overall proficiency (which reflects higher scores on *both* the semantic and spelling composite) was associated with *stronger* congruence priming but an independent dimension of variance between individuals captured by the difference between semantic and spelling scores assessed by PC2, and by the unique effects of spelling in the composite-based analysis, showed that higher spelling ability was associated with *weaker* congruence priming for high frequency targets.

The determinants of variability in automatic semantic priming

This independent factor corresponds to the spelling-meaning factor identified in our previous investigations of individual differences in masked priming, reviewed in the Introduction. These studies (Andrews & Hersch, 2010; Andrews & Lo, 2012, 2013) used the lexical decision task which can, in principle, be successfully performed on the basis of orthographic access alone, without retrieving semantic information. In those data, comprehension ability predicted overall speed of lexical classification, but spelling was a more potent unique predictor of the masked orthographic priming assessed in these studies than comprehension/vocabulary. However, the animal categorisation task used in the present experiment requires semantic processing. Under these task demands, overall proficiency predicted both average speed and degree of congruence priming, and comprehension/vocabulary, rather than spelling ability, was the strongest unique predictor of priming. This is consistent with our previous finding that the 'semantic profile' of higher vocabulary than spelling predicted semantic influences on morphological priming (Andrews & Lo, 2013) and extends this evidence to show that individual differences in the division of labor between orthographic and semantic processing interact with task demands as well as stimulus factors to determine patterns of priming.

Although overall proficiency was the most potent predictor of congruence priming in the semantic categorisation task, the independent spelling-meaning factor also played a more subtle role, that was restricted to high frequency non-exemplar targets. For these items, the beneficial effects of better spelling when combined with superior comprehension/vocabulary (i.e., overall proficiency) were counteracted by independent effects of the *difference* between spelling and comprehension/vocabulary ability, indexed by PC2, that selectively enhanced congruence priming in poor spellers but reduced it in better spellers.

These finding converge with our previous evidence that the high orthographic precision indexed by better spelling is associated with stronger sensitivity to orthographic competition (Andrews & Hersch, 2010; Andrews &Lo, 2012, 2013). Both congruent and incongruent primes differed from the target orthographically and therefore had the potential to conflict with the orthographic features activated by the target word. The orthographic precision indexed by high spelling ability may be associated with greater sensitivity to orthographic discrepancies between primes and targets that potentially counteracts the benefits of shared semantic features.

When the orthographic and semantic representations of a word are tightly interconnected, as they are for individuals with high quality lexical representations, their coherent co-activation converges on the lexical representation for the target word and inhibits the residual orthographic activation of the prime. Decision time also benefits from the overlapping semantic features activated by congruent primes which support the strong congruence priming shown by individuals high in both comprehension/vocabulary and spelling. The discrepant profile of good spelling but poor comprehension/vocabulary may reflect orthographic precision in the absence of coherence (Andrews, 2015): weaker binding of precise orthographic representations to the complex of semantic and phonological attributes that define the word's full lexical representation (Perfetti, 2007). This impedes efficient resolution of the conflicting orthographic features of the prime and target, particularly for high frequency targets that are accessed more quickly, and dilutes the benefit of overlapping semantic features. In contrast, poor spellers with imprecise orthographic representations are insensitive to orthographic discrepancies between the prime and target and instead make rapid decisions based on the, convergent activation of category-relevant semantic features that occurs when high frequency targets are preceded by high overlap congruent primes, leading to strong, semantically-mediated congruence priming.

According to this interpretation, the differential patterns of priming observed for exemplars and non-exemplars in the average data reflect the combined impact of both individual and item variability. At the item level, the differences may be due to frequency rather than exemplar status *per se*. In the stimulus set constructed by Quinn and Kinoshita (2008, Exp. 1) that was used in the present experiment, the average frequency of the critical exemplar targets (M=4.8/million) was similar to the low frequency non-exemplars (M=5.4) and substantially lower than the high frequency non-exemplars (M=97.3). The frequency manipulation was limited to non-exemplars both because of constraints on the frequency range of animal targets and also because a frequency effect for non-exemplars provides the clearest evidence that categorisation decisions are based on lexical retrieval rather than category search (Forster, 2004), or expiry of a deadline for ‘no’ responses (Carreiras et al., 1997). Thus, the counteracting effects of orthographic precision on semantic priming from high overlap congruent primes appear to be restricted to high frequency targets which, within the present stimulus set, were all non-exemplars. This may be because high frequency words yield faster access to the convergent semantic features that underlie the large congruence effect demonstrated by poorer spellers and/or because better spellers are more likely to have established precise orthographic representations for high frequency words.

This account suggests that congruence priming for both exemplars and non-exemplars relies on the general mechanisms proposed by the semantic feature overlap hypothesis (Quinn & Kinoshita, 2008) and that these processes operate more efficiently in high proficiency readers, particularly those high in reading comprehension/ vocabulary. However, for high frequency words, categorisation decisions are also influenced by an independent dimension of variability between individuals that determines their relative sensitivity to semantic and orthographic overlap between the prime and target. Poor spellers benefit from prime-target pairs with high semantic overlap while good spellers suffer from the conflict between the orthographic features of such items.

Lexical quality and semantic priming

These data provide strong evidence for the multi-faceted nature of ‘lexical quality’. Perfetti’s (2007) definition of this construct emphasised both orthographic precision and semantic coherence (Andrews, 2015). Precision ensures rapid access to the word’s orthographic form but effective use of lexical information in reading requires that these orthographic representations are tightly interconnected with the word’s semantic and phonological features to ensure coherent, synchronous activation of these convergent codes to support comprehension.

Orthographic precision, indexed by spelling ability, appears to be necessary but not sufficient for optimally efficient semantic categorisation performance. The semantic composite of comprehension/vocabulary was a stronger predictor of congruence priming than spelling ability. However, the most efficient performance and the strongest congruence priming was associated with the high scores on both composite measures. This is consistent with the view that high lexical quality is defined by orthographic precision in combination with lexical coherence (Andrews, 2015) and that high quality representations support rapid, automatic semantic priming. By contrast, orthographic precision in the absence of coherence – indexed by higher spelling than comprehension/vocabulary – was associated with reduced congruence priming for high frequency words, suggesting sensitivity to orthographic discrepancy; while low spelling ability, partialling out reading/vocabulary, predicted strong semantic priming for the same high frequency words. These independent effect of the discrepancy between comprehension/vocabulary and spelling ability may reflect differences in the extent of binding between the different elements of a reader’s lexical representations that give rise to differences in the weighting of orthographic and semantic information in lexical retrieval and decision processes.

These findings converge with recent evidence that parafoveal semantic preview effects in eye movement studies of sentence reading are stronger in individuals with higher reading comprehension, but reduced in better spellers (Veldre & Andrews, 2016a). Given that better spelling was generally associated with deeper parafoveal processing, the absence of semantic preview effects was attributed to competition between the orthographic features of the preview and target. Subsequent evidence has shown that parafoveal semantic preview benefit is due to the contextual plausibility of the preview rather than its semantic relationship to the target word (Veldre & Andrews, 2016b). However, this form of semantic preview benefit was also reduced in higher proficiency readers, defined by high comprehension and spelling, who only benefited from identical previews, suggesting that they effectively discounted the information retrieved from non-identical previews because of its orthographic mismatch with the target word (Veldre & Andrews, 2016b). This evidence provides insight into how individual differences in the division of labor between orthographic and semantic processing in early lexical retrieval revealed by masked priming effects for isolated words contribute to the processes involved in 'real reading'.

In conclusion, the present results add to a growing body of evidence showing that at least part of the documented variability in semantic priming among skilled readers (Stolz, Besner & Carr, 2005) is accounted for by systematic individual differences in written language proficiency (e.g., Yap et al., 2009), even for the automatic processes tapped by masked semantic congruence effects . These individual differences in early semantic processing have important implications for future refinement of theories of lexical retrieval and reading.

ACKNOWLEDGEMENTS

This research was supported by an Australian Research Council Discovery grant to Sally Andrews.

Portions of these data were presented at the Psychonomic Society Conference, Long Beach, 2014.

FOOTNOTES

[1] The usual time limits for the Reading Comprehension and Vocabulary subtest were each halved – to 10 mins and 7.5 mins respectively. In both tests, the items are ordered randomly with respect to difficulty so there is no reason to believe that these shorter time limits should compromise the reliability and validity of this widely used test.

[2] All frequency statistics reported in this paper are per million values obtained from the CELEX database (Baayen, Piepencrock, & van Rijn, 1993) using Davis's (2005) N-Watch software.

[3] Parallel analyses conducted on raw RT using GLMM models that assumed an Inverse Gaussian distribution of residuals (Lo & Andrews, 2015) yielded virtually identical outcomes.

[4]The results associated with this contrast are subsumed by the separate pairwise comparisons of the High and Low Overlap primes with the Incongruent prime condition and are therefore not reported. [5] To address a reviewer's concern that this approach of testing full models containing all interaction terms is vulnerable to overfitting, additional analyses were conducted to confirm that the omission of 2-way interactions between prime type and PC1 in the models for both the exemplar and non-exemplar data yielded significantly poorer fits to the data following a backward stepping approach (χ^2=6.28, p<0.05; χ^2=11.93, p<0.05 respectively), and significantly better fits to the data when they were added following a forward stepping approach (χ^2=6.43, p<0.05; χ^2=11.92, p<0.05 respectively). We also confirmed that omitting effects which failed to improve or impair goodness-of-fit following a forward or backward stepping approach did not affect the statistical results in any way, supporting recent conclusions that "the presence of superfluous variance components has minute effects on the estimates of the variability of the fixed-effects estimates" (Bates, Kliegl, Vasishth & Baayen, 2015)

[6] Random correlations were not estimated because they yielded convergence errors for the non-exemplar data. According to Barr et al. (2013) dropping these extra parameters from the model does not affect assessment of the fixed effects and Bates et al. (2015) described these 'zero-correlation parameter models' as a "defensible" alternative to maximally specified linear mixed effect models.

APPENDIX

Results of the LME models tested on inverse RT for exemplar targets and non-exemplar targets. Two models were conducted on each dataset to allow all pairwise comparisons on prime type to be tested. Model 1 used normalised sum contrasts to test Incongruent primes separately against High and Low Overlap primes. Model 2 used a normalised sum contrast to test the difference between High and Low Overlap primes. This model also included and a final contrast that was mathematically orthogonal to this latter comparison (i.e., Incongruent primes with the average of High and Low Overlap primes) which is not reported because it does not provide any additional information beyond the three pair-wise comparisons. The tables below report the full set of fixed effects tested in Model 1 as well as the main and interaction effects for the additional pair-wise contrast tested in Model 2 *(in italics).* Significant effects are indicated in bold.

1. Analyses using PC1 and PC2 scores derived from principal component analysis as predictors

EXEMPLAR TARGETS

	Estimate	Std. Error	t value
(Intercept)	-1.8690	0.0232	-80.50
Target freq (log)	-0.1062	0.0377	**-2.82**
Previous trial /RT/	0.0004	0.0000	**9.98**
Previous trial response	-0.0006	0.0002	**-3.93**
C1: LowOverlap vs Incon	-0.0054	0.0117	-0.47
C2: HighOverlap vs Incon	-0.0510	0.0134	**-3.81**
C3: HighOverlap vs Low Overlap	*0.0458*	*0.0125*	***3.66***
PC1	-0.0574	0.0172	**-3.35**
PC2	-0.0303	0.0258	-1.17
C1 x PC1	0.0212	0.0105	**-2.03**
C1 x PC2	-0.0221	0.0157	-1.41
C2 x PC1	0.0241	0.0105	**-2.30**
C2: PC2	-0.0229	0.0157	-1.46
C3 x PC1	*0.0030*	*0.0104*	*0.29*
C3 x PC2	*0.0006*	*0.0154*	*0.04*

NON-EXEMPLAR TARGETS

	Estimate	Std. Error	t value
(Intercept)	-1.7090	0.0213	-75.94
Previous trial /RT/	0.0003	0.0000	**13.71**
Previous trial response	0.0000	0.0001	-0.45
Target freq [Tfreq]: High/Low	-0.0937	0.0193	**-4.86**
C1: LowOverlap vs Incon	-0.0402	0.0081	**-4.95**
C2: HighOverlap vs Incon	-0.0762	0.0087	**-8.75**
C3: HighOverlap vs Low Overlap	*0.0359*	*0.0087*	***4.37***
PC1	-0.0739	0.0171	**-4.31**
PC2	-0.0655	0.0258	-0.59
C1 x Tfreq	-0.0231	0.0162	-1.43
C2 x Tfreq	-0.0320	0.0174	-1.84

C3 x Tfreq	*-0.0091*	*0.0164*	*-1.62*
Tfreq x PC1	0.0013	0.0058	-0.23
Tfreq x PC2	-0.0158	0.0086	-1.83
C1 x PC1	-0.0238	0.0069	**-3.45**
C1 x PC2	-0.0095	0.0103	-0.92
C2 x PC1	-0.0110	0.0069	-1.61
C2 x PC2	-0.0182	0.0103	-1.78
C3 x PC1	*-0.0127*	*0.0092*	*-1.86*
C3 x PC2	*-0.0087*	*0.0084*	*0.85*
C1 x Tfreq x PC1	-0.0072	0.0137	-0.52
C1 x Tfreq x PC2	-0.0047	0.0205	-0.23
C2 x Tfreq x PC1	0.0059	0.0137	0.43
C2 x Tfreq x PC2	-0.0438	0.0205	**-2.14**
C3 x Tfreq x PC1	*-0.0136*	*0.0183*	*-1.00*
C3 x Tfreq x PC2	*0.0398*	*0.0167*	*1.95*

2. Analyses using composite ZSpell and ZTotal scores as predictors.

Only main effects and interactions involving ZSpell and ZTotal are reported because the contrast estimates and statistical outcome for all other effects are identical to the principal component-based analyses reported above.

EXEMPLAR TARGETS

	Estimate	Std. Error	t value
ZSpell	-0.0123	0.0228	-0.54
ZTotal	-0.0638	0.0210	**-3.04**
C1 x ZSpell	0.0040	0.0140	0.29
C1 x ZTotal	-0.0304	0.0126	**-2.40**
C2 x ZSpell	0.0027	0.0140	0.20
C2: ZTotal	-0.0332	0.0127	**-2.61**
C3 x ZSpell	*0.0014*	*0.0137*	*0.07*
C3 x ZTotal	*0.0027*	*0.0125*	*0.28*

NON-EXEMPLAR TARGETS

	Estimate	Std. Error	t value
ZSpell	-0.0342	0.0228	-1.50
ZTotal	-0.0673	0.0210	**-3.21**
Tfreq x ZSpell	0.0115	0.0077	1.56
Tfreq x ZTotal	-0.0109	0.0070	-1.56
C1 x ZSpell	-0.0075	0.0092	-0.81
C1 x ZTotal	-0.0245	0.0083	**-2.94**
C2 x ZSpell	0.0073	0.0092	0.80
C2 x ZTotal	-0.0200	0.0083	**-2.42**
C3 x ZSpell	*-0.0147*	*0.0091*	*-1.62*
C3 x ZTotal	*-0.0044*	*0.0083*	*-0.54*

C1 x Tfreq x ZSpell	-0.0009	0.0183	-0.05
C1 x Tfreq x ZTotal	-0.0085	0.0166	-0.51
C2 x Tfreq x ZSpell	0.0378	0.0183	**2.07**
C2 x Tfreq x ZTotal	-0.0228	0.0165	-1.38
C3 x Tfreq x ZSpell	*-0.0396*	*0.0182*	***-2.17***
C3 x Tfreq x ZTotal	*0.0143*	*0.0165*	*0.87*

REFERENCES

Andrews, S. (2012). Individual differences in skilled word recognition and reading: The role of lexical quality. In J. Adelman (Ed.,) *Visual word recognition* (Vol. 2). Hove: Psychology Press.

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

Andrews, S. & Bond, R. (2009). Lexical expertise and reading skill: Bottom-up and top-down processing of lexical ambiguity. *Reading & Writing: An Interdisciplinary Journal*, *22*, 687-711.

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. (2012). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning Memory and Cognition, 38*,152-163.

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.

Ashby, J., Rayner, K. & Clifton, C. (2005) Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *Quarterly Journal of Experimental Psychology, 58A*, 1065-1086.

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

Baayen, R.H., Piepencrock, R. & Gulikers, L. (1995). CELEX2. Philadelphia: Linguistic Data Consortium.

Barr, D. Levy, R., Scheepers, C. & Tily, H. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255-278.

Bates, D., Kliegl, R., Vasishth, S. & Baayen, H. (submitted). Parsimonious mixed models. http://arxiv.org/abs/1506.04967

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

Box, G. E. P, & Cox, D. R. (1964). An analysis of transformations (with discussion). Journal of the Royal Statistical Society B, 26, 211-252.

Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *Nelson-Denny reading test*. Rolling Meadows, IL: Riverside Publishing Company.

Carreiras, M., Perea, M. & Grainger, J. (1997). Effects of the orthographic neighborhoods in visual word recognition: cross-task comparisons. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 857-871.

Damian, M. (2001). Congruity effects evoked by subliminally presented primes: Automaticity rather than semantic processing. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 154-165.

Davis, C.J. (2005). N-Watch: A program for deriving neighbourhood size and other psycholinguistic statistics. *Behavior Research Methods, 37*, 65 – 70.

Davis, M. & Rastle, K. (2010). Form and meaning in early morphological processing: Comment on Feldman, O'Connor, and Moscoso del Prado Martin (2009). *Psychonomic Bulletin & Review, 17*, 749-755.

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E. Mueller, M., Dehaene-Lambertz, G., van de Moortele, P., & Le Biahn, D. (1998). Imaging unconscious semantic priming. *Nature, 395*, 597-600.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79*, 1-37.

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.

Feldman, L.B., Milin, P., Moscoso del Prado, F. & O'Connor, P.A. (2015). Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in Human Neuroscience, 9,* 111.

Feldman, L.B., O'Connor, P.A., & Moscoso del Prado Martín, F. (2009). Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review, 16*, 684–691.

Forster, K. (2004). Category size effects revisited: Frequency and masked priming effects in semantic categorization. Brain and Language, 90, 276–286.

Forster, K., & 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., Mohan, J. & Hector, J. (2003). The mechanics of masked priming. S. Kinoshita & S. Lupker (Eds.) (2003), *Masked Priming: The State of the Art*. (p. 3-37). New York: Psychology Press.

Gollan, T.H., Forster, K.I. & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and non-cognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1122-1139.

Gorlin, E. & Teachman, B. (2015). Inhibitory control as a moderator of threat-related interference biases in social anxiety. *Cognition and Emotion, 29*, 723-735.

Greenwald, A.G., Draine, S.C. & Abrahams, R.L. (1996). Three cognitive markers of unconscious semantic activation. *Science, 273*, 1699-1702.

Harm, M.W. & Seidenberg, M.S. (2004). Computing the meanings of words in reading: Co-operative division of labor between visual and phonological processes. *Psychological Review, 111*, 662-720.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision and visual masking: A survey and appraisal. *Behavioral and Brain Science, 9*, 1-23.

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*(4), 622-636.

Kinoshita, S. & Norris, D. (2009). Transposed-letter priming of prelexical orthographic representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 1-18.

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

Kouider, S. & Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society B, 362*, 857-875.

Kunde, W., Kiesel, A. & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition, 88*, 223-242.

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.*

Marcel, A.J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15,* 197-237.

Marche, T., Brainerd, C. & Reyna, V. (2010). Distinguishing true from false memories in forensic contexts: Can phenomenology tell us what is real? *Applied Cognitive Psychology, 24*, 1168-1182.

McRae, K. & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 558-572.

McRae, K., Cree, G., Seidenberg, M. & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods, 37,* 547-559.

McRae, K., de Sa, V. & Seidenberg, M. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126,* 99-130.

Monteith, M., Woodcock, A. & Gulker, J. (2013). Automaticity and control in stereotyping and prejudice: The revolutionary role of social cognition across three decades of research. In D. Carlston (Ed.), *The Oxford Handbook of Social Cognition*. (pp. 74-94). New York, NY, US: Oxford University Press.

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

Norris, D. & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology: General, 137,* 434-455.

Perfetti, C.A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*, 357-383.

Quinn, W. & Kinoshita, S. (2008). Congruence effects in semantic categorization with masked primes and broad and narrow categories. *Journal of Memory and Language, 58,* 286-306.

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes, 23*, 942–971.

Shelton, J. & Martin, R.C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1191-1210.

Stanovich, K.E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers.* NY: Guilford.

Stolz, J.A. & Besner, D. & Carr, T.H. (2010). Implications of measures of reliability for theories of priming: Activity in semantic memory is inherently noisy and uncoordinated. *Visual Cognition, 12*, 284-336.

Veldre, A., & Andrews, S. (2016a). Semantic preview benefit in English: Individual differences in the extraction and use of parafoveal semantic information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 837-854.

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

Venables, W. & Ripley, B. (2002). *Modern Applied Statistics with S.* New York: Springer

Witzel, J. & Forster, K.I. (2014) Lexical co-occurrence and ambiguity resolution. *Language, Cognition and Neuroscience, 29,* 2014, 158-185.

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.

Figure captions

Figure 1

LME model estimates of raw RT for exemplar targets (left panel) and non-exemplar targets (right panel) primed by Congruent/High Overlap, Congruent/Low Overlap and Incongruent masked primes as a function of a continuous measure of the first principal component (PC1) corresponding to the average of the semantic and spelling composite scores (see text for details). (Model estimates of inverse RT have been converted to raw RT for ease of interpretation)

Figure 2

LME model estimates of raw RT for Low frequency (left panel) and High frequency (right panel) non-exemplar target words primed by Congruent/High Overlap, Congruent/Low Overlap and Incongruent masked primes as a function of a continuous measure of the second principal component (PC2) corresponding to the difference between the semantic and spelling composite scores (see text for details). (Model estimates of inverse RT have been converted to raw RT for ease of interpretation)

Figure 3

LME model estimates of raw RT for exemplar targets (left panel) and non-exemplar targets (right panel) primed by Congruent/High Overlap, Congruent/Low Overlap and Incongruent masked primes as a function of a continuous measure of the semantic composite standard score from the Nelson-Denny tests of reading comprehension and vocabulary (ZTotal). (Model estimates of inverse RT have been converted to raw RT for ease of interpretation)

Figure 4

LME model estimates of raw RT for Low frequency (left panel) and High frequency (right panel) non-exemplar target words primed by Congruent/High Overlap, Congruent/Low Overlap and Incongruent masked primes as a function of a continuous measure of the composite spelling score (ZSpell). (Model estimates of inverse RT have been converted to raw RT for ease of interpretation)

Table 1: Mean RT (standard error in brackets) and percentage error rate (%ER) to exemplar and non-exemplar target words in the three masked priming conditions.

Target	**Prime type**					
	Congruent/ High overlap		Congruent/ Low overlap		Incongruent	
	RT	%ER	RT	%ER	RT	%ER
Exemplars	*hawk-EAGLE*		*mole-EAGLE*		*knee-EAGLE*	
	557 (10)	4.6	566 (9)	5.1	566 (8)	6.9
Non-Exemplars	*pistol-RIFLE*		*boots-RIFLE*		*camel-RIFLE*	
High frequency	586 (10)	3.1	598 (11)	3.6	612 (9)	4.3
Low frequency	620 (10)	8.0	632 (11)	9.1	643 (10)	8.9