In press, *Emotion*

Automatic vigilance for negative words in lexical decision and naming: Comment on Larsen, Mercer, and Balota (2006)

Zachary Estes & James S. Adelman

University of Warwick

Word count: 2134

Correspondence to:

Zachary Estes

Department of Psychology

University of Warwick

Coventry CV4 7AL U.K.

Email: z.estes@warwick.ac.uk

Phone: +44 (0) 247 652 3183

Abstract

An *automatic vigilance* hypothesis states that humans preferentially attend to negative stimuli, and that this attention to negative valence disrupts the processing of other stimulus properties. Thus negative words typically elicit slower color naming, word naming, and lexical decisions than neutral or positive words. Larsen, Mercer, and Balota (2006) analyzed the stimuli from 32 published studies, and they found that word valence was confounded with several lexical factors known to affect word recognition. Indeed, with these lexical factors covaried out, Larsen et al. found no evidence of automatic vigilance. We report a more sensitive analysis of 1,011 words. Results revealed a small but reliable valence effect, such that negative words (e.g., “shark”) elicit slower lexical decisions and naming than positive words (e.g., “beach”). Moreover, the relation between valence and recognition was categorical rather than linear; the extremity of a word’s valence did not affect its recognition. This valence effect was not attributable to word length, frequency, orthographic neighborhood size, contextual diversity, first phoneme, or arousal. Thus, the present analysis provides the most powerful demonstration of automatic vigilance to date.

KEYWORDS: affective valence; automatic vigilance; emotional Stroop; lexical decision; word naming; word recognition

An *automatic vigilance* hypothesis states that humans preferentially attend to negative stimuli (Ohman & Mineka, 2001; Pratto & John, 1991; Wentura, Rothermund, & Bak, 2000). Specifically, attention is disengaged from negative stimuli more slowly than from neutral or positive stimuli. One functional consequence of automatic vigilance is that, following the presentation of a negative stimulus, responses to subsequent stimuli are hindered (e.g., Fox, Russo, Bowles, & Dutton, 2001; McKenna & Sharma, 2004; Most, Chun, Widders, & Zald, 2005). In the emotional Stroop task, for instance, neutral words elicit slower color naming when preceded by a negative word than when preceded by a neutral or positive word (McKenna & Sharma, 2004). A more immediate consequence of automatic vigilance is that this sustained attention to negative valence may also hinder responding to the negative stimulus itself. Indeed, negative words typically elicit slower color naming (see Williams, Mathews, & MacLeod, 1996), lexical decisions (Wentura et al., 2000), and word naming (Algom, Chajut, & Lev, 2004) than do neutral or positive words. Apparently, this attention to negative valence diverts processing resources away from other properties of the stimulus, such as its color, its lexical status, and its pronunciation.

Unfortunately, previous studies of automatic vigilance have suffered a number of methodological shortcomings. For instance, many studies have used small samples of items (e.g., five items per condition), and the same items have often been used across different studies, thereby rendering equivocal the generality of the result. Moreover, many studies have used only negative and neutral words. To provide a complete test of automatic vigilance, one must demonstrate that negative stimuli also elicit slower responses than positive stimuli. Otherwise, the effect could be attributed to valence in general rather than negative valence in particular. Equally problematic is the observation that many studies have failed to control important lexical variables. Larsen, Mercer, and Balota (2006) analyzed 1033 stimulus words from 32 emotional Stroop studies, and they found that word valence was confounded with word length, word frequency, and orthographic neighborhood size, all of which are known to affect lexical processing. Specifically, across those studies the negative words were longer, were less frequent, and had fewer orthographic neighbors than the neutral words. Any or all of these confounds could explain the slower responses to negative words (see Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004).

Given these methodological shortcomings, one might reasonably conclude that the evidential status of automatic vigilance is equivocal at best. Indeed, when Larsen and colleagues (2006) covaried the spurious lexical factors out, they found no evidence that negative words elicit slower responses (see also McKenna & Sharma, 2004). However, Larsen and colleagues were cautious not to reject the validity of automatic vigilance. They concluded instead that automatic vigilance may be a real phenomenon, but if so, its apparent magnitude has been grossly inflated by the confounding of word valence with other lexical factors. In fact, Larsen et al. cited the methodology of Wentura and colleagues (2000) as a noteworthy exception to these criticisms. Those researchers used 50 negative words and 50 positive words, and after length and frequency were covaried out, negative words still elicited slower lexical decisions. Thus, despite the methodological problems exhibited by many studies, some evidence of automatic vigilance appears immune to those criticisms.

If negative words do elicit slower lexical decisions than positive words, as suggested by Wentura et al. (2000), then one may wonder why the large-scale analysis of Larsen and colleagues (2006) provided little evidence of this. For each word, Larsen et al. adopted the valence designation (i.e., negative, neutral, positive) of the original study from which it was sampled. Because that analysis included stimuli from various studies, there was little consistency in the methods, measures, and criteria that were used to determine valence, nor in the participant populations who provided the ratings (e.g., some were clinical populations). Consequently, the sensitivity of the valence factor may have been relatively low. Larsen et al.

obtained lexical decision and word naming latencies and accuracies from the publicly available database of Balota et al. (2002). Although those data were collected at several universities by different researchers, the methodology was standardized. Thus, of the data used by Larsen and colleagues, the dependent variables (i.e., latencies and accuracies) appear reliable but the independent variable (i.e., valence) may be relatively insensitive.

We therefore undertook an alternative analysis that was conceptually similar to that of Larsen et al. (2006), with one critical methodological deviation: The present study used a set of items for which valence ratings were collected via a uniform procedure across all stimuli. Specifically, valence ratings were obtained from the Affective Norms for English Words (ANEW; Bradley & Lang, 1999), which includes ratings of 1034 words on a scale from 1 (unpleasant) to 9 (pleasant). And following Larsen et al., word recognition data were obtained from the English Lexicon Project (ELP; Balota et al., 2002), which includes latencies and accuracies for both lexical decisions and word naming. We merged these two datasets in order to test whether lexical decisions and word naming are slower for negative words than for positive words, as predicted by automatic vigilance. This constancy in the collection of valence ratings should lend greater statistical power to detect the presumed relationship between word valence and recognition.

A number of known predictors of lexical retrieval (see Balota et al., 2004) were included as covariates. Two measures of word length—letters and syllables—were included. Word frequencies were calculated from the TASA12 corpus (Landauer, Foltz, & Laham, 1998), which consists of 8.26 million tokens. Contextual diversity—the number of distinct documents in which a word occurs (Adelman, Brown, & Quesada, 2006)—was also calculated from TASA12. Word frequency and contextual diversity were log transformed. To avoid problems with zero counts, a constant of one was added prior to transformation. Orthographic neighborhood size (orthographic *N*; see Andrews, 1997) and first phoneme were collected from the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995). First phoneme was only included as a covariate in the analysis of naming data. We also included arousal ratings from ANEW, since arousal tends to correlate with valence (Jennings, McGinnis, Lovejoy, & Stirling, 2000). Of the 1034 words in ANEW, 23 were absent from either ELP or CELEX. Thus, analyses were calculated on 1011 words.

Initial analyses confirmed significant correlations between word valence and recognition latency in both lexical decision [$r = -.233$, $t(1009) = -7.63$, $p < .001$] and naming [$r = -.194$, $t(1009) = -6.29$, $p < .001$]. To test whether valence provides a unique contribution to word recognition, we also conducted multiple regressions with the aforementioned covariates factored out. Valence exhibited no unique contribution to accuracy in either task, but did account for a significant amount of unique variance in both lexical decision [$t(1003) = -3.47$, $p < .001$] and naming latencies [$t(968) = -2.40$, $p = .017$]. In support of the automatic vigilance hypothesis (Algom et al., 2004; Pratto & John, 1991; Wentura et al., 2000), lexical decisions and naming were slower for negative words (e.g., “shark”) than for positive words (e.g., “beach”).

Figure 1 illustrates this relationship between valence ratings and recognition latencies. Before the covariates were factored out (solid lines in Figure 1), response latencies were clearly longer for words on the negative side of the valence scale than for words on the positive side. But when those covariates were removed (dashed lines), the slopes were attenuated. This change in slopes indicates that much of the difference in response latencies between positive and negative words is attributable to spurious covariates such as word length, word frequency, and arousal, rather than to word valence per se. This is effectively the observation made by Larsen and colleagues (2006). Unlike their original analysis, however, the present analysis did reveal a significant relationship between valence and latencies even after the critical covariates were statistically controlled.

Interestingly, the relation between valence and recognition was categorical rather than linear; the extremity of a word's valence did not affect its recognition. This observation was confirmed by multiple regression analyses that included two measures of valence: In addition to the continuous valence factor used in the analysis above, we also created a categorical valence factor. All items with a mean valence rating of less than 5 (the scalar midpoint) were labeled "negative", and all items with a mean of greater than 5 were labeled "positive". (This led to the exclusion of one word, "taxi", for which the mean valence rating was 5.00.) The categorical factor accounted for unique variance over and above the continuous factor in both lexical decision [$t(1001) = -2.227$, $p = .026$] and naming latencies [$t(966) = -2.701$, $p = .007$]. In contrast, the continuous factor did not account for any unique variance in either lexical decision [$t(1001) = .204$, $p = .829$] or naming [$t(966) = 1.11$, $p = .267$]. Thus, valence was categorically related to recognition latency. Note that the words in the ANEW database span the entire range of valence, with many words on or near the midpoint of the valence scale. In fact, 15% of the items fall within the range of 4.50 to 5.50. (12.5% would be expected in that range if the items were uniformly distributed across the scale.) Examples include "fur" (4.51), "square" (4.74), "foot" (5.02), "pencil" (5.22), and "glacier" (5.50). So given that the valence ratings were continuously distributed across the scale, the categorical nature of the response latencies is striking.

Table 1 provides a comparison of one set of regressions in which valence was treated as a continuous predictor, and another set of regressions in which valence was treated as a categorical predictor. The magnitude of each predictor's effect is reported in milliseconds. When valence was treated as a categorical predictor, naming latencies were about 10 ms slower and lexical decisions were about 15 ms slower for negative words than for positive words. We also calculated the amount of unique variance in response latencies explained by word valence. For this analysis, we conducted one regression with the aforementioned covariates only (i.e., letters, syllables, frequency, contextual diversity, orthographic *N*, first phoneme, and arousal), and another regression with the categorical valence factor included. The difference in those R^2 values (i.e., ΔR^2) provides a measure of effect size. Although the preceding analyses confirmed that the effect of valence was significant, this effect was small for both lexical decisions (ΔR^2 = 0.79%) and word naming (ΔR^2 = 0.58%).

The present analysis yields three important observations: (1) Word valence does predict lexical decision and word naming latencies, (2) this effect of word valence is categorical rather than linear, and (3) the effect is small. We speculate that the latter two observations might explain why the relationship between valence and latency was not evident in Larsen and colleagues' (2006) analysis. Their study included 322 negative words, 393 neutral words, and 240 positive words (as well as 78 disorder-specific words). Assuming that the neutral words would cluster around the midpoint of the valence scale (that is, if all the items were to be rated on a single scale), that group of items would consist of a heterogeneous mixture of slightly negative and slightly positive words. Given the categorical relation between valence and recognition latency, these slightly negative and slightly positive words would exhibit an inordinate amount of variance in latencies when considered as a group. That increase in variance could mask the relationship between valence and latency, especially since the neutral words were over-represented in the sample, and since the effect appears to be quite small to begin with. In contrast, the present analysis used a different set of stimuli with a more sensitive measure of valence. This methodological difference proved critical for revealing the subtle relationship between word valence and recognition latency.

Despite the small magnitude of this effect—approximately 10 to 15 ms, or less than 1% of the variance in word recognition latencies—its theoretical significance looms large. Larsen and colleagues (2006) provided a critical and valid observation, namely, that much of the prior evidence of automatic vigilance has confounded word valence with other lexical factors known to affect word recognition. In contrast, the presently observed delay in responding to negative stimuli was not attributable to word length, word frequency, orthographic neighborhood size, contextual diversity, first phoneme, or arousal. Thus, the present analysis provides the most powerful demonstration of automatic vigilance to date. The categorical nature of this effect suggests that affective evaluation is fast but crude (Pratto & John, 1991). Evidently, one immediately categorizes a stimulus as negative or positive (Fazio, 2001; Lazarus, 1982; Zajonc, 1980), and this categorical evaluation affects one's response. If the stimulus is positive, then responding generally proceeds rapidly. But if the stimulus is negative, then responses to other aspects of the stimulus (such as its color, its pronunciation, or its lexical status) are delayed.

References

Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word naming and lexical decision times. *Psychological Science, 17,* 814-823.

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional Stroop phenomenon: A generic slowdown, not a Stroop effect. *Journal of Experimental Psychology: General, 133,* 323-338.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4,* 439-461.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database (Release 2)* [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Balota, D. A., Cortese, M. J., Hutchison, K. A., Neely, J. H., Nelson, D., Simpson, G. B., & Treiman, R. (2002). The English Lexicon Project: A Web-based repository of descriptive and behavioral measures for 40,481 English words and non-words. Available at http://elexicon.wustl.edu.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283-316.

Bradley, M. M. & Lang, P. J. (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition & Emotion, 15,* 115-141.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General, 130,* 681-700.

Jennings, P. D., McGinnis, D., Lovejoy, S., & Stirling, J. (2000). Valence and arousal ratings for Velten mood induction statements. *Motivation and Emotion, 24,* 285-297.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25,* 259-284.

Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion, 6,* 62-72.

Lazarus, R. S. (1982). Thoughts on the relationship between emotion and cognition. *American Psychologist, 37,* 1019-1024.

McKenna, F. P., & Sharma, D. (2004). Reversing the emotional Stroop effect reveals that it is not what it seems: The role of fast and slow components. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 382-392.

Most, S. B., Chun, M. M., Widders, D. M., & Zald, D. H. (2005). Attentional rubbernecking: Cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin & Review, 12,* 654-661.

Ohman, A. & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108,* 483-522.

Pratto, F. & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology, 61,* 380-391.

Wentura, D., Rothermund, K., & Bak, P. (2000). Automatic vigilance: The attention-grabbing power of approach- and avoidance-related social information. *Journal of Personality and Social Psychology, 78,* 1024-1037.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120,* 3-24.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151-175.

Author Notes

This research was supported by ESRC grant RES 000221558 and Grant F/215/AY from the Leverhulme Trust (UK). We thank Randy J. Larsen, Kimberly A. Mercer, and David A. Balota for an insightful and constructive review. We also thank Jose Quesada for calculating word frequency and contextual diversity from the TASA12 corpus, and we thank Michelle Verges for comments on an earlier draft of this manuscript.

Table 1. Regression analyses for word naming and lexical decision latencies as a function of valence (treated continuously and categorically in separate analyses).

	Word Naming				Lexical Decision			
	Continuous		Categorical		Continuous		Categorical	
Variable	ms	t	ms	t	ms	t	ms	t
Letters	11.289	8.740 ***	11.252	8.713 ***	14.336	8.730 ***	14.302	8.706 ***
Syllables	6.122	2.989 *	6.508	2.182 *	5.898	1.641	6.624	1.744 †
Log WF	0.242	0.040	1.370	0.224	-23.481	-2.930 ***	-21.451	-2.680 **
Orth N	0.344	0.926	0.351	0.948	0.435	0.901	0.445	0.924
Log CD	-12.013	-1.874 †	-12.828	-2.013 *	1.859	0.222	0.114	0.014
Arousal	-2.791	-2.207 *	-3.332	-2.619 **	-8.736	-5.221 ***	-9.430	-5.600 ***
Valence	-1.701	-2.402 *	-9.942	-3.443 ***	-3.186	-3.466 ***	-15.490	-4.122 ***
Total R^2	52.26%		52.58%		53.01%		53.24%	

Note. These regressions were conducted separately with valence treated as either continuous or categorical. Effect magnitudes are in milliseconds (ms). WF = word frequency, Orth N = orthographic neighborhood size, CD = contextual diversity. In the word naming analyses, the effect of first phoneme was also removed and was highly significant. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Fig. 1. Recognition latency ($M \pm SE$) as a function of word valence.