Emotion
2008, Vol. 8, No. 4, 445–452

1528-3542/08/$12.00 DOI: 10.1037/1528-3542.8.4.445

RESPONSE

Not All Negative Words Slow Down Lexical Decision and Naming Speed: Importance of Word Arousal

Randy J. Larsen, Kimberly A. Mercer, David A. Balota, and Michael J. Strube
Washington University in St. Louis

Previously the authors analyzed sets of words used in emotion Stroop experiments and found little evidence of automatic vigilance, for example, slower lexical decision time (LDT) or naming speed for negative words after controlling for lexical features. If there is a slowdown evoked by word negativity, most studies to date overestimate the effect because word negativity is often confounded with lexical features that promote slower word recognition. Estes and Adelman (this issue) analyze a new set of words, controlling for important lexical features, and find a small but significant effect for word negativity. Moreover, they conclude the effect is categorical. The authors analyze the same data set but include the arousal value of each word. The authors find nonlinear and interaction effects in predicting LDT and naming speed. Not all negative words produce the generic slowdown. Paradoxically, negative words that are moderate to low on arousal produce more LDT slowing than negative words higher on arousal. This finding presents a theoretical and empirical challenge to researchers wishing to understand the boundaries of the automatic vigilance effect.

Keywords: emotion words, automatic vigilance, emotion Stroop, lexical decision

In a previous study (Larsen, Mercer, & Balota, 2006), we investigated whether the emotional connotation of words influences word recognition speed above and beyond various lexical characteristics of the words. The rationale for this investigation was based on work suggesting that cognitive activity, including word recognition, may be momentarily disrupted when threatening information is detected in the perceptual stream (Algom, Chajut, & Lev, 2004). The idea is that an automatic vigilance mechanism monitors the sensory stream and causes a brief interruption when threat is detected (Pratto & John, 1991; Wentura, Rothermund, & Bak, 2000).

In this study (Larsen et al., 2006), we examined words used in 32 published emotion Stroop studies. We found that, in general, the threatening and control words used were confounded with word length and frequency of use. In particular, the threatening words were longer and more rare than the control words, making it ambiguous whether the observed generic slowdown in color naming is due to the negative words being more rare and longer, or whether the negative words capture attentional resources in a manner consistent with the automatic vigilance hypothesis.[1]

We (Larsen et al., 2006) *did* find a subset of words (coded as disorder-specific words) that were associated with a slowdown in lexical decision latency even after controlling for frequency of use and length. For example, words such as: *ache, bite, bleeds, bruises, cramp, defeat, disfigure, dishonesty, germ, hideous, illness, incest, infected, lying, rejection, repulsive, ridicule, tumor,* and *vomit* remained associated with generic lexical decision slowing even after controlling for lexical features. This leaves open the possibility that some words may, in fact, evoke the automatic vigilance effect.

Randy J. Larsen, Kimberly A. Mercer, David A. Balota, and Michael J. Strube, Department of Psychology, Washington University in St. Louis.

Preparation of this article was facilitated by Grant RO1-MH63732 from the National Institute of Mental Health to Randy J. Larsen. Support for the ELP database was provided by Grant BCS 0001801 from the National Science Foundation to David A. Balota. Support for the ANEW database was provided by a National Institute of Mental Health Center grant to Peter J. Lang. We thank Zachary Estes, James Adelman, and two anonymous reviewers for constructive and helpful reviews on a previous version of this article.

Correspondence concerning this article should be addressed to Randy J. Larsen, Department of Psychology, Campus Box 1125, One Brookings Drive, Washington University, St. Louis, MO 63130. E-mail: rlarsen@wustl.edu

[1] The emotional Stroop task consists of color naming of negative, neutral, and sometimes positive words. The underlying mechanism responsible for the emotion Stroop findings is *not* one of response competition, as is the case for the standard color Stroop effect (Burt, 2002). Instead, the underlying mechanism for the emotion Stroop is thought to be a generic interrupt system that acts early and in an automatic fashion when threatening information is detected in the perceptual stream (Algom et al., 2004). Algom et al. (2004), and others (e.g., Larsen et al., 2006), argue that, to the extent that negative words produce a slowdown in color naming, they do so through this mechanism of automatic threat vigilance, which should produce general slowing on all cognitive activity. Nevertheless, some researchers question whether the task of color naming is more sensitive to, or a better indicator of, this generic slowdown than other cognitive activities, such as making a lexical decision or simply naming a word. Algom et al., 2004 present the argument that threatening stimuli will temporarily disrupt *all* ongoing cognitive activity, including lexical decisions, word naming, and color naming. Indeed, Algom et al. (2004) demonstrate in several experiments that word negativity affects both color naming and word naming to a similar degree (Experiments 1-4) and word negativity affects lexical decision time (Experiment 5). As such, at least these three tasks (color naming, lexical decision time, and word naming) appear equivalent to each other as indicators of the generic slowdown associated with threatening stimuli.

The Larsen et al., 2006 study suffered from one major limitation. In that study, we used words from published emotion Stroop studies, and thus used the categorical codes provided by the original study authors to classify the words for our analyses. This is suboptimal for two reasons. First, word categories were generated by 32 different research groups. It would be better to have a single source of word ratings. Second, word valance was a categorical code, resulting in nominal or, at best, ordinal level of measurement. Interval level scaling of word valance would provide a more precise test. A stronger test of the automatic vigilance hypothesis would be to use interval scaling on negativity for each word, and to use a large and representative list of emotion words rated on valence by a single reliable source.

Estes and Adelman (2008) pursued this tack and obtained a large and representative list of emotion words that were interval scaled on word negativity by one research group (the Affective Norms for English Words (ANEW) word list, Bradley & Lang, 1999). They then obtained lexical decision time (LDT) and naming speed to these words and, after controlling for a number of important lexical features, found a small but significant effect of word negativity on LDT and naming speed. Moreover, Estes and Adelman (2008) concluded that the effect was categorical. That is, the category of negative words produced slower LDT and naming speed than the category of positive words, and this categorical effect was somewhat stronger than the effect obtained from using interval-level word negativity ratings.

Figure 1 from Estes and Adelman (2008) clearly shows the categorical effect of word negativity. However, in looking at their Figure 1, there is an obvious nonlinearity in the left half of the figure (i.e., among the negative words). It appears that, as words increase in negativity, there is actually a *decrease* in the slowdown effect. In fact, examining their values corrected for lexical features, there appears to be a very small difference, if any, between the most arousing negative words and the group of positive words. Despite this ocular evidence of nonlinearity in the category of negative words, the authors conclude that the automatic vigilance effect is categorical, implying that the effect is equivalent for all negative words.

The ANEW words have also been scaled for arousal value by the list originators, that is, Bradley and Lang, 1999. In Figure 1, we plot the arousal by valance value for the entire ANEW word list, which shows the strong U-shaped quadratic relation between valence and arousal. This figure, in combination with Figure 1 from Estes and Adelman (2008), has two important implications. First, arousal may play a role in predicting the automatic vigilance effect from negative words. Second, valence and arousal together may produce nonlinear or interactive effects in predicting automatic

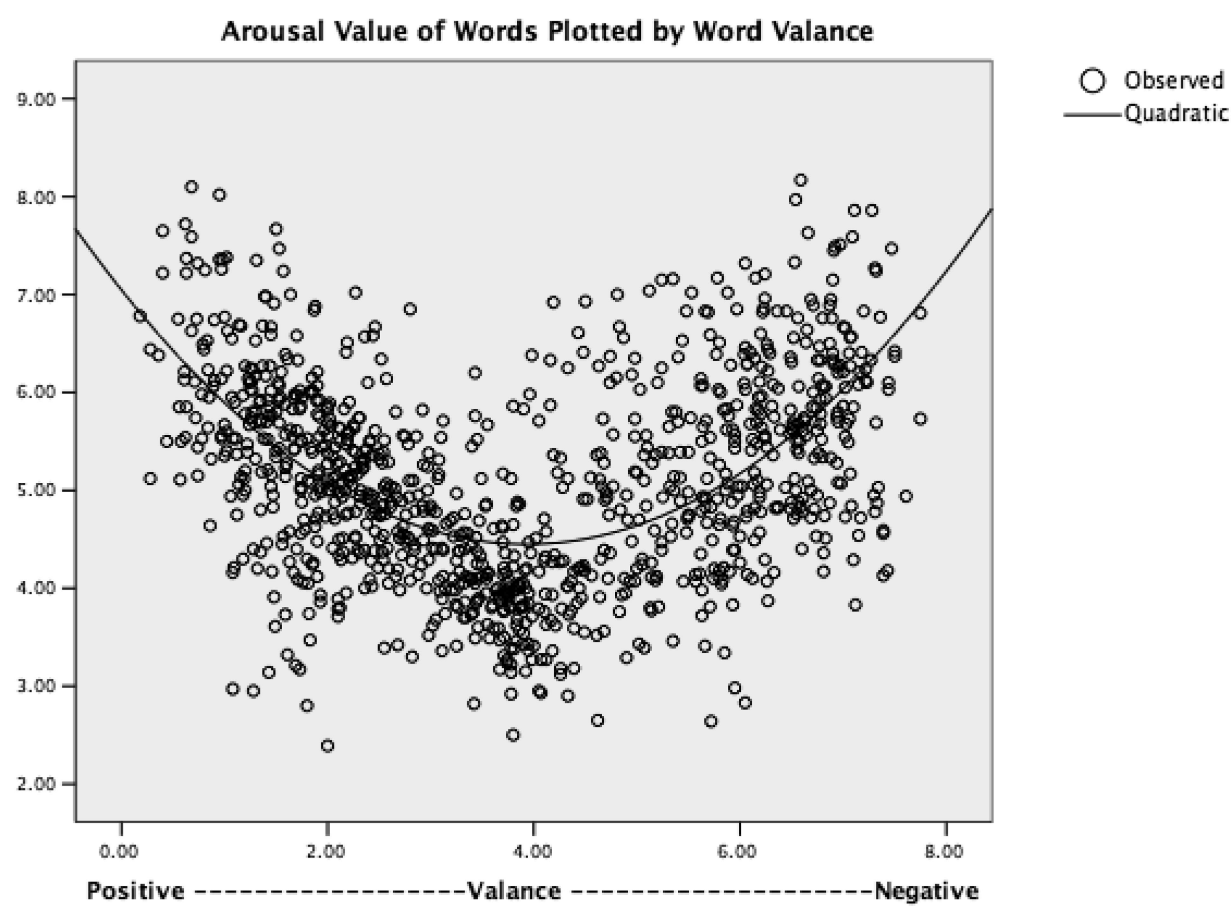

Figure 1. Relationship between word negativity and word arousal rating for entire sample of 1,021 words.

vigilance. In our response, we test whether these implications are true.

Method

Word Selection

Words were drawn from the list (ANEW; Bradley & Lang, 1999), the same list used by Estes and Adelman (2008). These words have been normed by a large group of college students on pleasantness and arousal, as described in Bradley and Lang (1999). The pleasantness dimension is a bipolar scale that runs from 1 to 9, with a rating of "1" indicating *extremely unpleasant*, a "5" indicating *neutral*, and "9" indicating *extremely pleasant*. For ease of interpretation, we reversed this scale so that larger numbers indicated more negativity. The arousal dimension is a unipolar scale that runs from 1 to 9, with a rating of "1" indicating *low arousal* and "9" indicating *high arousal*. Information for obtaining the ANEW words is available from the Center for the Study of Emotion and Attention at http://www.phhp.ufl.edu/csea/index.html.

Lexical and Behavioral Characteristics of the ANEW Words

The English Lexicon Project (ELP; Balota et al., 2002) is a searchable database containing lexical characteristics and naming and lexical decision times for over 40,000 words and is available online at: http://elexicon.wustl.edu/default.asp. From the ELP database we used the standardized[2] lexical decision time and naming speed variables on each word. While we report raw reaction time data in the tables below (so readers can see the millisecond metric), we used the standardized form of these measures in the correlational analyses. The ELP database also contains lexical characteristics on each of the words, which we extracted for the present analyses (see Larsen et al., 2006 for complete descriptions of these lexical features).

We submitted the 1,034 ANEW words to the ELP search engine, which found exact matches for 1,021 words. The valence ratings from the ANEW database were then merged with the lexical and behavioral data from the ELP database for each of these words. This list of 1,021 words forms the final data set used in our analyses.

Results

Descriptive information on the words is presented in Table 1. The words themselves were slightly positive (mean of 3.85 is slightly toward the positive direction from the midpoint of the bipolar 0–8 positive–negative scale), although there was a good deal of variability. The average arousal rating fell in the moderately arousing range, but again with a good deal of variability. The reaction time data are within the range of published lexical and naming speed values, with naming speed being faster than lexical decision speed, which is typical. The words also show a good deal of variability on all three lexical characteristics.

First-order Pearson correlations between all variables are presented in Table 2. We replicate the typical finding that both lexical decision time and naming speed are strongly related to word length, and are strongly inversely related to the frequency of use (log Hyperspace Analogue to Language [HAL] Index; Lund & Burgess, 1996), and are moderately inversely related to orthographic neighborhood size. For comparison purposes, we include the raw HAL frequency index as well as the Kucera and Francis (KF; 1967) frequency index in this table. In every case, the correlations with behavioral measures are stronger with the log (HAL) index than with the raw HAL or the KF frequency index. Consequently we use the log (HAL) index in the analyses reported below. First-order correlations with word negativity also suggest modest positive relations between negativity and lexical decision time and naming speed, with the more negative words taking longer in both of these word recognition tasks. However, word negativity also correlates inversely with the frequency of use index, implying that negative words are used less frequently in everyday linguistic behavior than positive words. This fact, while interesting in its own right, highlights the importance in controlling for lexical features of words when examining the automatic vigilance hypothesis in word recognition paradigms. Arousal showed no linear relationship to either lexical decision time or naming speed. Nevertheless, it is important to examine arousal effects in interaction with word negativity to determine if more or less arousing words produce longer LDT and naming times.

Of primary interest is to test whether word negativity is associated with slower lexical decision and naming latencies after controlling for lexical features of the words. We conduct these tests by running a series of general linear models, with the behav-

Table 1
Means and SDs *on Reaction Times, Lexical Characteristics, and Valance and Arousal on 1,021 Words From the Affective Norms for English Words (ANEW) Word List*

	M	SD
Lexical decision speed (in ms)	655	79.99
Naming speed (in ms)	633	58.14
Length (in letters)	6.16	1.82
Log (HAL frequency)	8.58	1.77
Orthographic neighborhood	2.81	4.16
Negativity [a]	3.85	1.99
Arousal [b]	5.11	1.05

[a] Rating made on a bipolar scale, ranging from 0 to 8, with 0 anchored as "*extremely positive*," 5 anchored as "*neutral*," and 8 anchored as "*extremely negative*."
[b] Rating made on a Likert scale, ranging from 1 to 9, with 1 representing "*not at all arousing*" and 9 representing "*extremely arousing*."

[2] Note that z-scores on RT data in the English Lexicon Project were not calculated as normal deviates. Rather, because the data were gathered on over 40,000 words, single subjects could not provide LDT and naming speed on each and every word. Instead, each subject was given a subset of approximately 2,500 words to respond to, depending on the task. Because subjects differ from each other in overall response latency and variability, responses were standardized within each subject (using that subject's mean and standard deviation across all the words he or she responded to) before combining into a composite standard score for each word. LDT and naming speed were gathered on each word from approximately 30 subjects. The mean z-scores for LDT and naming speed for each word in the English Lexicon Project thus control for individual differences in mean response latency and variability across the subjects who responded to that particular word.

Table 2
Pearson Correlations Between All Variables Used in the Regression Analyses, Calculated Across 1,021 Affective Norms for English Words (ANEW) Words

	Naming	Length	Log (HAL)	OrthoN	Negativity	Arousal	HAL	Kucera and Francis
LDT	$.69^{**}$	$.57^{**}$	$-.69^{**}$	$-.40^{**}$	$.24^{**}$	$-.01$	$-.31^{**}$	$-.35^{**}$
Naming Speed		$.52^{**}$	$-.53^{**}$	$-.36^{**}$	$.20^{*}$	$.04$	$-.25^{**}$	$-.25^{**}$
Length			$-.41^{**}$	$-.66^{**}$	$.05$	$.13^{*}$	$-.23^{**}$	$-.23^{**}$
Log (HAL)				$.32^{**}$	$-.28^{**}$	$.05$	$.58^{**}$	$.58^{**}$
OrthoN					$-.05$	$-.06$	$.22^{**}$	$.20^{**}$
Negativity						$.05$	$-.18^{**}$	$-.18^{**}$
Arousal							$.00$	$-.02$
HAL								$.83^{**}$

Note. LDT = lexical decision time.
** $p < .001$. * $p < .01$, two-tailed.

ioral data (lexical decision latency and naming speed) as the dependent variable in each regression. For independent variables, we entered the lexical characteristics (length, frequency, and orthographic neighborhood) as well as word negativity and arousal values, plus their interaction. Because the effects in Figure 1 of Estes and Adelman (2008) appear nonlinear, and the relation between word negativity and arousal is strongly curvilinear (see Figure 1), we also enter the squared and cubed terms for word negativity. The nonlinear components of negativity are necessary to produce normally distributed residuals when arousal is included as a predictor. To determine whether arousal moderates the linear and nonlinear effects of negativity, we also entered interactions (i.e., product variables) of arousal with negativity, negativity squared, and negativity cubed.

Predicting Lexical Decision Time

In predicting lexical decision latency, we applied the following multiple regression model:

$$\mathrm{LDT} = \mathrm{c} + \mathrm{length} + \mathrm{freq} + \mathrm{orthoN} + N + \mathrm{A} + (N{*}\mathrm{A}) + \mathrm{N}^2 + \mathrm{N}^3 + \mathrm{N}^2\mathrm{A} + \mathrm{N}^3\mathrm{A}$$

where *c* is a constant, length is word length in letters, *freq* is the log of the HAL frequency index, and *orthoN* is orthographic neighborhood size. *N* is word negativity, and *A* is the arousal value of the words. To determine the proportions of variance accounted for by models of increasing complexity, the predictors were entered in four stages. In Stage I, the linear predictors were entered (length, freq, orthoN, N, and A); in Stage II the two-way interactions were entered (N^{*}A and N^2); in Stage III the three-way interactions were entered (N^3, and N^2A). In the last stage, the four-way interaction (N^3A) was entered. We also centered all the independent variables prior to analysis to control for multicolinearity between predictor variables. The results for the first and last stages are presented in Table 3. The first stage results provide easy comparison to Estes and Adelman's findings; the last stage provides the full model.

The complete model accounted for a very large portion of variance in LDT, with an adjusted R^2 of .587 for the full model. Table 3 presents the parameter estimates for each of the terms in the model. In the first stage of the analysis, word length and frequency accounted for significant and large amounts of variability in LDT, at 7.3% and 20.0% of the variability respectively. Similar to Estes and Adelman (2008), we also find that word negativity is associated with a significant but relatively small

Table 3
Multiple Regression Predicting Lexical Decision Time Reaction Time (LDT RT; z-score) From Word Length, Frequency, Orthographic Neighborhood, Word Negativity, Arousal, the Interaction of Negativity and Arousal, and Cubic and Quadratic Interactions of Negativity and Arousal

	First Step				Last Step			
Parameter	*B*	SE_B	*t*	sr^2	*B*	SE_B	*t*	sr^2
Intercept	−0.472	0.005			−0.459	0.009		
Length	0.055	0.004	**13.350**	0.073	0.056	0.004	**13.517**	0.074
HAL	−0.078	0.004	**22.046**	0.200	−0.076	0.004	**21.311**	0.184
Ortho	0.001	0.002	0.634	0.000	0.001	0.002	0.707	0.000
Negativity (N)	0.011	0.003	**3.912**	0.006	0.028	0.007	**3.904**	0.006
Arousal (A)	−0.009	0.005	1.656	0.001	−0.003	0.009	0.359	0.000
N^2					−0.002	0.002	0.990	0.000
N × A					−0.025	0.007	**3.720**	0.006
N^3					−0.002	0.001	**2.199**	0.002
N^2 × A					−0.001	0.002	0.478	0.000
N^3 × A					0.003	0.001	**3.192**	0.004

Note. sr^2 = squared semi-partial correlation. All independent variables were centered prior to analysis. Boldfaced *t* values are significant at $p < .05$.

amount of slowing in LDT (0.6% of variability in LDT uniquely related to word negativity).

As implied by Figure 1 in Estes and Adelman (2008), we also found a significant cubic effect for negativity, which was further moderated by arousal. This highest order interaction accounted for about the same unique amount of variability as word negativity alone (0.4% for the negativity cubed by arousal interaction, compared to 0.6% for word negativity alone). Figure 2 portrays a three-dimensional surface plot of the complex relationship between arousal and valence (word negativity) in predicting LDT for these data. Predicted values for raw LDT (to provide easier interpretation and comparison to Estes and Adelman) were obtained from a full regression model for various values of word valence (range = −1.5 to + 1.5 *SD*) and arousal (range = −1.5 to + 1.5 *SD*). One can readily see that the greatest amount of LDT slowing is found for negative words that are low to moderate on the arousal dimension. The cubic relation between negativity and LDT is most apparent when arousal is low and largely disappears when arousal is moderate to high.

Predicting Word Naming Speed

Multiple regression results for word naming speed are presented in Table 4. In predicting naming speed, we applied the same model that we used for LDT. The complete model accounted for a large portion of variance in word naming speed, with an adjusted R^2 of .40 for the full model. Table 4 presents the parameter estimates for each term in the model. In Stage I, word length and frequency accounted for significant and large amounts of variability in word naming speed, at 7.4% and 9.5% of the variability respectively. Interesting to note, word frequency accounts for much more variability in lexical decision time (20%) than in naming speed (9.5%). We also find that word negativity is associated with a significant but relatively small amount of slowing in word naming speed (0.6% of variability in LDT uniquely related to word negativity). The complex interaction found for LDT was not found for naming speed, but a simpler cubic effect for negativity was significant. This relation is illustrated in Figure 3.

Discussion

We examined predictions drawn from the concept of automatic vigilance, which holds that, when a threatening stimulus is detected in the perceptual stream, cognitive resources are diverted to more thoroughly evaluate that stimulus, resulting in a generic slowdown in the cognitive processing of other attributes of that stimulus. Past studies on automatic vigilance using LDT and naming paradigms (Algom, Chajut, & Lev, 2004), and a color-naming paradigm (the emotion Stroop task; e.g., Pratto & John, 1991), are inconclusive because they did not control for important lexical features that influence word recognition speed. Indeed, we (Larsen et al., 2006) found that a high percentage of emotion Stroop studies use word lists that are not equivalent on important lexical features (Larsen et al., 2006), making their findings ambiguous with respect to the source of any slowdown observed in color naming.

Estes and Edelman (2008) obtained a large set of unique words normed for negativity (and arousal) by one laboratory then obtained the lexical features and LDT and naming speed for each word, obtained from another laboratory. After controlling for lexical features, they found that the relationship between word negativity and LDT and naming speed remained significant, albeit quite small. However, even though the effect is small, it is nevertheless theoretically important. Theories of word recognition typically give a miniscule role, if any, to the meaning of words in determining word recognition speed. No theory of word recognition identifies word negativity as playing a role in word recognition. Moreover, the effect is consistent with the automatic vigi-

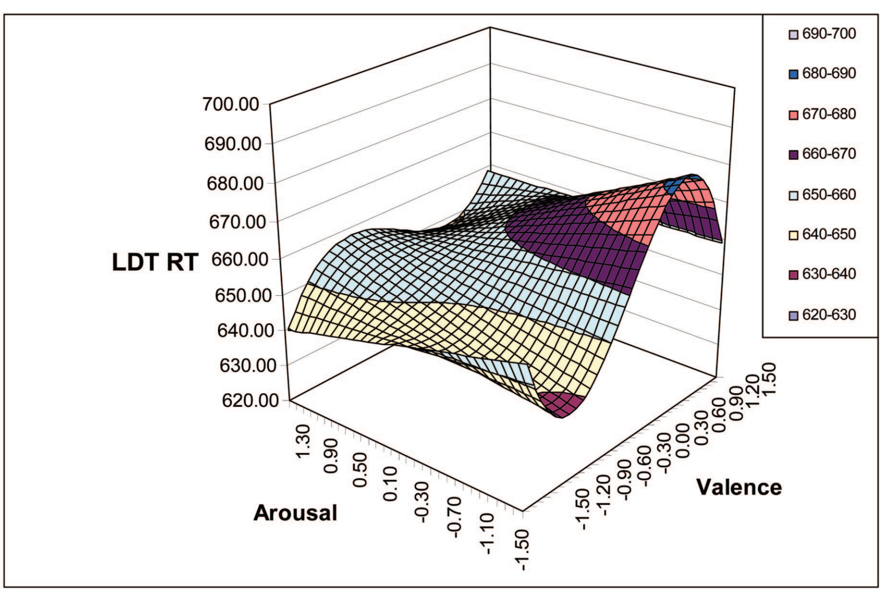

Figure 2. Three-dimensional surface plot of the relation between word valance (high numbers mean more negative), arousal rating, and lexical decision time (LDT) in milliseconds.

Table 4

Multiple Regression Predicting Word Naming Speed (z score) From Word Length, Frequency, Orthographic Neighborhood, Word Negativity, Arousal, the Interaction of Negativity and Arousal, and Cubic and Quadratic Interactions of Negativity and Arousal

	First Step				Last Step			
Parameter	*B*	SE_B	*t*	sr^2	*B*	SE_B	*t*	sr^2
Intercept	−0.420	0.006			−0.413	0.011		
Length	0.052	0.005	**11.208**	0.074	0.052	0.005	**11.293**	0.075
HAL	−0.050	0.004	**12.702**	0.095	−0.048	0.004	**12.125**	0.087
Ortho	0.000	0.002	0.160	0.000	0.000	0.002	0.216	0.000
Negativity (N)	0.010	0.003	**3.159**	0.006	0.028	0.008	**3.559**	0.007
Arousal (A)	0.000	0.006	0.046	0.000	0.003	0.010	0.347	0.000
N^2					0.000	0.002	0.139	0.000
N × A					−0.009	0.008	1.224	0.001
N^3					−0.003	0.001	**2.339**	0.003
N^2 × A					−0.001	0.002	0.496	0.000
N^3 × A					0.001	0.001	1.281	0.001

Note. sr^2 = squared semi-partial correlation. All independent variables were centered prior to analysis. Boldfaced *t* values are significant at $p < .05$.

lance notion, implying that negative information detected in the perceptual stream interferes with ongoing cognitive activity to produce the generic slowdown effect (Algom et al., 2004). We agree with Estes and Edelman (2008) on this conclusion.

Another conclusion of Estes and Edelman (2008), one that we disagree with, is that the effect of word negativity on LDT and naming speed is categorical. This implies that all negative words produce the effect to an equivalent degree. We evaluated this conclusion in the present paper by examining the role of arousal, in interaction with word negativity, in producing the slowdown in LDT and naming speed. As for LDT, in addition to finding a general word negativity effect, we also find similar sized effects for nonlinear components of word negativity, as well as interactions between negativity and arousal.

Our results clarify that the effect of word negativity on LDT is not categorical. That is, not all negative words produce the same level of automatic vigilance. Moreover, we identify a pattern of results related to arousal and nonlinear components of word negativity that predicts additional variability in LDT above and beyond word negativity. A particularly important finding is that the beta weight for the negativity × arousal interaction was itself negative, indicating that high arousal negative words produce *less* of a slowdown in LDT than low arousal negative words. This may be because many negative and highly threatening words, for example, *death, germs, rotten, stench, bereavement, urine, handicap, inferior, gloom, obesity, ache*, and *coffin* are low on arousal. It appears that categorical word negativity is not the only operative semantic variable that generates generic slowing in word recognition.

Our findings pose a challenge to emotion researchers to figure out what specific features of words are most predictive of the automatic vigilance effect. A likely next step would be a content analysis of the words that produce the largest slowdown in LDT. It may be that specific attributes of the words, beyond their mere

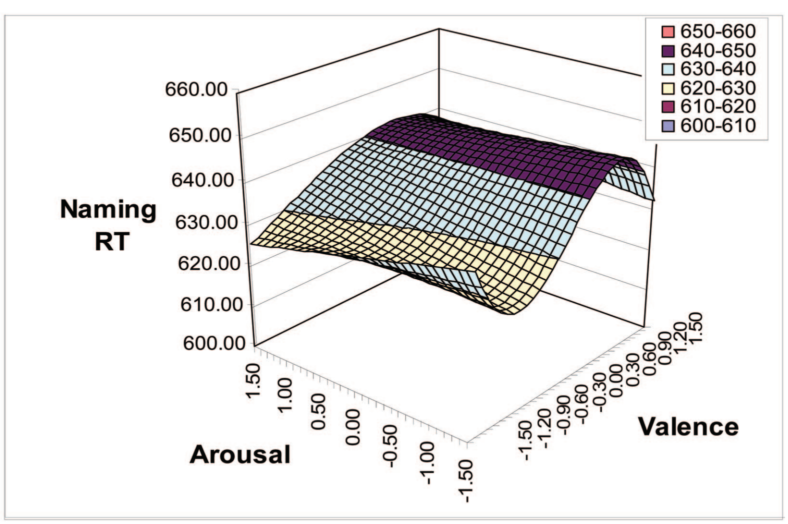

Figure 3. Three-dimensional surface plot of the relation between word valance (high numbers mean more negative), arousal rating, and naming speed in milliseconds.

negativity, can be found that predict which words produce the slowdown effect. For example, Wentura et al. (2000), using a color-naming paradigm, make a distinction between negative other-relevant traits (e.g., *cruel, vicious, violent, mean,* which are threatening for the people around the person who has these traits) and negative possessor-relevant traits (*lonely, depressed, frustrated, unhappy,* which are threatening for the person who has these traits). Wentura et al. (2000) reported larger automatic vigilance effects for other-relevant negative traits compared to possessor-relevant negative traits (even after controlling for lexical features of the trait words). An important line for future research would be to determine the specific attributes of words that most contribute to automatic vigilance effects.

Another neglected aspect of words that deserves mention concerns their standard deviations on the positive-negative dimension. Some words may have a negative mean rating yet also have a high standard deviation (e.g., *lesbian, tease, body, naked*), meaning that some people see the words as negative and others see them as positive. Other words are viewed as negative by almost everyone (e.g., *cancer, rape, grief, failure, rejected*), and thus might make better test words in studies of automatic vigilance. Similarly, some neutral words have high standard deviations (e.g., *hospital, obscene*), whereas others are seen as neutral by almost everyone (e.g., *pencil, table, icebox*). Clearly, neutral words with smaller standard deviations would be better control words. The point here, and one point of our previous article (Larsen et al., 2006), is that researchers should carefully evaluate the words they use, both in terms of their lexical features as well as their normative ratings, before applying them in a study of automatic vigilance.

For emotion researchers, our results show that even such quick and automatic cognitive processes as word recognition can be influenced by the emotional connotations of the stimuli. Human evolution most likely sculpted us in such a way that our perceptual and attentional systems are especially tuned to the threat value of objects in the perceptual stream. For example, perceptual search experiments on human facial expressions show that detection speed is faster for faces displaying an angry or a fearful expression than neutral or happy expressions (Tipples, Atkinson, & Yount, 2002). The survival value of such sensitivity to stimulus threat is obvious, and early humans without this sensitivity were most likely at a fitness disadvantage.

When threat is present in the perceptual stream, it is processed through a fast subcortical pathway that biases other lower-level cognitive processes, such as perception and attention (e.g., Ohman, Flykt, & Esteves, 2001). However, threat detection is followed by slower and more thorough higher-level cognitive process involving cortical structures (Koster Crombez, van Damme, Verschuere, & De Houwer, 2004). As such, threat is detected very quickly (around 100 ms; Smith, Cacioppo, Larsen, & Chartrand, 2003), yet slows secondary processing on more controlled tasks, such as color naming or lexical decision latency, an effect referred to as automatic vigilance. The concept of automatic vigilance is appearing in the cognitive literature (Algom et al., 2004), the emotion literature (Koster et al., 2004), and the social cognition literature (Wentura et al, 2000). Automatic vigilance effects may even produce bias in the cognitive system, resulting in a lower threshold for threat detection in the near future, such as is found with evaluative priming paradigms using negative primes (e.g., De Houwer & Randell, 2004; Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Hermans, De Houwer, & Eelen, 2003; Klauer, 2003).

In summary, we agree entirely with Estes and Adelman (2008) that the effects of automatic vigilance are real, albeit small. Moreover, because negative words are typically used less frequently, as shown in Table 2, it is especially important for researchers to carefully control for frequency of use in examining the cognitive processing of negative words. However, we disagree with Estes and Adelman (2008) that the effect of word negativity is categorical. Our analyses show that arousal interacts with word negativity in a counterintuitive manner, with the lower arousal negative words producing higher automatic vigilance effects than the highly arousing negative words. Our findings present a theoretical and empirical challenge to researchers wishing to understand the psychological processes that produce the automatic vigilance effect.

References

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.

Balota, D. A., Cortese, M. J., Hutchison, K. A., Neely, J. H., Nelson, D., Simpson, G. B., et al. (2002). The English Lexicon Project: A Web-based repository of descriptive and behavioral measures for 40,481 English words and non-words. Retrieved March 17, 2008, from http://elexicon.wustl.edu/default.asp

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW).* Gainesville, FL: The National Institute of Mental Health Center for the Study of Emotion and Attention, University of Florida.

Burt, J. S. (2002). Why do non-color words interfere with color naming? *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1019–1038.

De Houwer, J., & Randell, T. (2004). Robust affective priming effects in a conditional pronunciation task: Evidence for the semantic representation of evaluative information. *Cognition & Emotion, 18,* 251–264.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50,* 229–238.

Hermans, D., De Houwer, J., & Eelen, P. (2003). On the acquisition and activation of evaluative information in memory: The study of evaluative learning and affective priming combined. In J. Musch and K. C. Klauer (Eds.), *The psychology of evaluation: Affective process in cognition and emotion* (pp. 139–168). Mahwah, NJ: Erlbaum.

Klauer, K. C. (2003). Affective priming: Findings and theories. In J. Musch and K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7–50). Mahwah, NJ: Erlbaum.

Koster, E. H., Crombez, G., van Damme, S., Verschuere, B., & De Houwer, J. (2004). Does imminent threat capture and hold attention? *Emotion, 4,* 312–317.

Kucera, H., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

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

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavioral Research Methods, Instruments, & Computers, 28,* 203–208.

Ohman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General, 130,* 466–478.

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

Smith, K. N., Cacioppo, Larsen, J. T., & Chartrand, T. L. (2003). May I have your attention, please: Electrocortical responses to positive and negative stimuli. *Neuropsychologia,* 41, 171–183.

Tipples, J., Atkinson, A. P., & Young, A. W. (2002). The eyebrow frown: A salient social signal. *Emotion, 2,* 288–296.

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

Received September 29, 2006
Revision received June 6, 2007
Accepted July 19, 2007 ■