BRIEF REPORT

Cross-Modal Evaluative Priming: Emotional Sounds Influence the Processing of Emotion Words

Laura D. Scherer and Randy J. Larsen
Washington University in St. Louis

Cross-modal priming occurs when a prime presented in one sensory modality influences responses to a target in a different sensory modality. Currently, demonstrations of cross-modal evaluative priming have been sparse and limited. In the present study, we seek to partially rectify this state of affairs by examining cross-modal evaluative priming from auditory primes to visual targets. Significant cross-modal priming effects were found, but only for negative primes. Results are discussed in terms of the negativity bias, and several suggestions are provided for using cross-modal evaluative priming to address theoretically important questions about emotion and cognition.

Keywords: evaluative priming, cross-modal priming, automatic vigilance

Fazio, Sanbonmatsu, Powell, and Kardes (1986) developed an evaluative priming paradigm by adapting procedures from the semantic priming literature (e.g., Neely, 1977) and generalizing them to evaluative stimuli. In their first demonstration of evaluative priming, Fazio et al. (1986) presented participants with prime-target word pairs that were either congruent or incongruent on valence. Then participants were instructed to categorize target words as either “good” or “bad” as quickly as possible. This study, and many others on evaluative priming that followed (e.g., Bargh, Chaiken, Govender, & Pratto, 1992; Murphy & Zajonc, 1993), showed that responses to evaluative target stimuli are facilitated (i.e., are faster and/or more accurate) when those targets are preceded by evaluatively congruent primes compared to incongruent primes. Evaluative priming effects are theoretically important because (1) the processing of evaluative connotations occurs spontaneously, and (2) once activated, evaluations automatically activate responses to other cognitive content sharing the same valence (Klauer & Musch, 2003).

Cross-Modal Semantic Priming

The importation of priming paradigms from the semantic priming literature into emotion research occurred about two decades ago. Since that time, important developments continued to occur in the semantic priming literature. One relatively recent and important development in the semantic priming literature has been investigations of multisensory, or cross-modal, semantic priming. Cross-modal semantic priming occurs when the prime and the target are processed by different sensory modalities, such as an auditory prime and a visual target, or vice versa. Cross-modal semantic priming has proved useful in testing theories in several different cognitive domains. For example, Lukatela, Eaton, Moreno, and Turvey (2007) asked whether a common lexicon underlies words seen and words heard, and argued for a single lexicon based on strong cross-modal semantic priming. Furthermore, one classic view of word recognition is that activation flows from orthography (written) to phonology (sounds). However, Whatmough, Arguin, and Bub (1999) present a strong case, based on cross-modal semantic priming, that activation can also go from phonology to orthography (see also Lamy, Mudrik, & Deouell, 2007; Schneider, Engel, & Debener, 2008).

Cross-Modal Evaluative Priming

Semantic information can be conveyed visually and through sound, but evaluations can be conveyed in virtually any sensory channel. For example, touch can be unambiguously unpleasant, in the form of pain. Tastes and smells can be pleasurable or noxious. Similarly, sounds can communicate negativity or positivity: a scream can be experienced as “hair-raising,” whereas a child’s laughter can be pleasant.

Yet in spite of the sensory-universal nature of evaluations, evaluative priming has been largely concerned with priming within one sensory modality, namely, vision. In the vast majority of studies, participants see a prime that is either a picture or a word, and then they see a target that is also a picture or a word. This amounts to a unimodal priming paradigm, in which brief presentations of stimuli in one sensory modality influence the perception of other stimuli presented in that same sensory modality.
In contrast, the present article is concerned with cross-modal evaluative priming. To date, there have been only a handful of instances (to our knowledge) in which cross-modal evaluative priming effects have been observed. In two experiments, odors were found to influence the processing of faces and words (Hermans, Baeyens, & Eelen, 1998; Li, Moollem, Puller, & Gottfried, 2007). In another experiment, auditory primes were found to influence face and word processing (Carroll & Young, 2005). However, out of the three existing experiments that examine cross-modal evaluative priming, two (Carroll & Young, 2005; Hermans et al., 1998) possess an important limitation, insofar as they provided analyses of only congruent versus incongruent trials and collapsed over effects for positive and negative stimuli.1

This failure to report valence-specific effects is potentially important because certain theoretical perspectives (e.g., automatic vigilance) imply that cognitive priming might be adaptive primarily for negative stimuli (Larsen, 2004). Furthermore, much of the emotion and attitude literature has shown that effects for negative stimuli are often stronger than those for positive stimuli (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Hence, one prediction might be that evaluative asymmetries are also pertinent to the influence of stimuli across multiple sensory modalities. For example, emotionally aversive sounds (such as a growl) may signal the need for immediate action that requires other senses (e.g., running). In contrast, positive sounds generally require no such action, and, as a result, such sounds may have a relatively small influence on behavior and action tendencies involving other sensory modalities. In sum, the current literature on cross-modal evaluative priming is both small and limited. The purpose of the present article is to expand this literature by examining the automatic effects that emotional sounds (both positive and negative) have on the identification of emotion words.

The Present Research

In the following experiment, we ask participants to complete a typical evaluative priming task. However, instead of using visual stimuli (e.g., words or pictures) as primes, we employ auditory stimuli that are normatively positive or negative. Moreover, the targets are visually presented words that are either positive or negative. Hence, we predict that the affective connotations elicited by sounds will influence participants’ accuracy in responding to the evaluative meaning of the visually presented words.

Method

Participants and Design

A total of 45 (16 male and 29 female) college students participated in this experiment in return for partial course credit. Two participants made a large number of errors (greater than 2 SD above the mean), and one participant made an exceptionally large number of responses outside of the response window (greater than 3 SD above the mean). These participants were removed from the analyses, leaving a total of 42 participants. All participants completed the same experimental procedure, in which both the valence of the prime sounds (pleasant vs. unpleasant) and the valence of the target words (positive vs. negative) were manipulated.

Stimuli

The primes were 16 sounds selected from the International Affective Digitized Sounds (IADS-2; Bradley & Lang, 2007) based on their valence norms. In norming the sounds, Bradley and Lang (2007) had at least 100 male and female participants report how each sound made them feel on three separate dimensions (Valence, Arousal, Dominance) using the Self-Assessment Manikin (SAM), an affective rating system devised by Bradley and Lang (1994). In this system, a graphic manikin is manipulated to portray values along each of the three dimensions on a continuously varying scale and is used by participants to report how the sounds made him or her feel. For example, on the valence dimension, the manikin ranges from a smiling, happy figure to a frowning, unhappy figure. There are nine distinct gradations between endpoints on each dimension, and thus the ratings are converted to a 9-point Likert scale for analysis (see Bradley & Lang, 2007, for a more thorough description of normative rating methods). In the present study, we were primarily interested in selecting auditory stimuli from the endpoints of the valence dimension. The mean ratings for all IADS-2 sounds on the valence dimension ranged from 1.57 (highly negative) to 7.90 (highly positive). We chose 8 sounds that were affectively negative (mean rating = 2.0) and 8 sounds that were affectively positive (mean rating = 6.93). See Table 1 for a description of the sounds as well as the IADS-2 sound number and ratings.

Design and Procedure

Participants were seated in front of a desktop computer in a private room and were given a headphone set, which they wore for the duration of the experiment. In the first task, participants were presented with each of the 16 sounds and were asked to rate the pleasantness of each one using a −3 to +3 Likert-type scale. The order in which the sounds were presented was randomized. This rating task served two distinct purposes. The first was to obtain an explicit rating of the sounds from each of our participants so that we could determine whether participants’ evaluative reactions to the sounds corresponded to their overt ratings. Second, and perhaps more importantly, this initial rating task allowed participants to listen to each sound in its entirety, whereas in the subsequent priming task, participants were exposed to just a 1-s clip of each sound. Past research (as well as our own pilot testing) has shown that priming task effects decrease as the duration of the prime increases (e.g., Klauer, Teige-Mocigemba, & Spruyt, 2009). This presents a problem for the presentation of sounds in the priming

1 These experiments also report findings that are somewhat inconsistent. For example, Hermans et al. (1998) found priming from odor to visual stimuli for consciously presented odors. However, Li et al. (2007) found similar priming effects only when odors were presented below detection thresholds. More work is needed to resolve these inconsistencies.
task, because sounds unfold over time. Hence, playing the sounds in anticipation of the priming task served a very important purpose, namely, to introduce participants to sounds that they would only be “reminded” of in the later task by partial sound clips.

After listening to and rating the 16 sounds, all participants performed the priming task. They were informed that they would be performing a task that was meant to examine “how listening affects reading” and that the task would include a presentation of the sounds that they previously rated, followed by a word. The participants’ job was to ignore the sound and to indicate whether the target word was a “good” or “bad” word as quickly as possible after they appeared on the screen.

During the priming task, each sound was presented for 1 s, which meant that only a portion of each sound served as a prime. The sound was immediately followed by the presentation of a target word, which remained on the screen for 150 ms and was immediately replaced by a mask (a row of asterisks). The mask remained on the screen for 400 ms, leaving a total of 550 ms in which the participant could respond. If participants responded outside of the 550 ms window, the words “Faster Please!” appeared in red type on the screen and remained for 500 ms. If they responded within the window, the next trial would begin immediately. Before beginning the critical trials, participants were given four practice trials that included four of the primes (2 pleasant and 2 unpleasant) paired with four of the target words (2 positive and 2 negative). In the critical phase of the priming task, each prime was followed by each target exactly once, which resulted in 256 total trials (16 primes x 16 targets). The target-prime pairs were presented in random order. After completing the priming task, participants were debriefed and dismissed.

### Results

Pretest pleasantness ratings of the sounds provided by our subjects correlated strongly with the normative pleasantness values, $r = .88$, $p < .001$, implying our sample rated the sounds similarly to the normative sample. On the priming trials, for responses made within the 550 ms window, errors on each prime-type (pleasant, unpleasant) and word-type (positive, negative) combination were summed and divided by the number of on-time trials for that prime-target combination. This resulted in a total of two error indices for the pleasant and unpleasant primes, one index for good words and another for bad words. In this paradigm, positive spontaneous evaluations are assumed when the prime causes participants to make more errors on bad words than good words. This is because, for example, an error on a bad word indicates that the prime has caused the participant to press the “good” key in error.

### Priming Task Results

The error indices were submitted to a 2 (Word-type: positive vs. negative) x 2 (Prime: pleasant vs. unpleasant) within-subjects ANOVA. As predicted, a 2-way Prime X Word-type interaction emerged, $F(1, 41) = 4.20$, $p < .05$ (see Figure 1). No main effects were significant ($p > .20$). An inspection of Figure 1 reveals that the unpleasant sounds elicited significantly more errors on good words relative to bad words, $F(1, 41) = 4.67$, $p < .05$. For the pleasant sounds, the difference in errors on good versus bad words

<table>
<thead>
<tr>
<th>Description</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>POSITIVE SOUNDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seagull</td>
<td>6.95</td>
<td>1.65</td>
<td>1.07</td>
<td>1.38</td>
</tr>
<tr>
<td>Robin</td>
<td>7.12</td>
<td>1.56</td>
<td>2.13</td>
<td>.93</td>
</tr>
<tr>
<td>Kids</td>
<td>6.11</td>
<td>1.90</td>
<td>0.52</td>
<td>1.41</td>
</tr>
<tr>
<td>Applause</td>
<td>7.32</td>
<td>1.62</td>
<td>1.91</td>
<td>.86</td>
</tr>
<tr>
<td>Baseball</td>
<td>7.38</td>
<td>1.53</td>
<td>1.74</td>
<td>.92</td>
</tr>
<tr>
<td>Native Song</td>
<td>6.17</td>
<td>1.99</td>
<td>1.67</td>
<td>.94</td>
</tr>
<tr>
<td>Harp</td>
<td>7.44</td>
<td>1.41</td>
<td>2.24</td>
<td>.94</td>
</tr>
<tr>
<td>Guitar</td>
<td>6.98</td>
<td>1.90</td>
<td>2.04</td>
<td>.89</td>
</tr>
<tr>
<td>NEGATIVE SOUNDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vomiting</td>
<td>2.08</td>
<td>1.78</td>
<td>−2.54</td>
<td>.62</td>
</tr>
<tr>
<td>Female Scream</td>
<td>1.93</td>
<td>1.63</td>
<td>−2.70</td>
<td>.55</td>
</tr>
<tr>
<td>Attack1</td>
<td>1.68</td>
<td>1.31</td>
<td>−2.87</td>
<td>.34</td>
</tr>
<tr>
<td>Attack3</td>
<td>2.01</td>
<td>1.48</td>
<td>−2.63</td>
<td>.74</td>
</tr>
<tr>
<td>Attack2</td>
<td>1.80</td>
<td>1.56</td>
<td>−2.93</td>
<td>.25</td>
</tr>
<tr>
<td>Fight1</td>
<td>1.65</td>
<td>1.27</td>
<td>−2.87</td>
<td>.40</td>
</tr>
<tr>
<td>Car Wreck</td>
<td>2.04</td>
<td>1.52</td>
<td>−2.04</td>
<td>.98</td>
</tr>
<tr>
<td>Dentist Drill</td>
<td>2.89</td>
<td>1.67</td>
<td>−1.09</td>
<td>1.05</td>
</tr>
</tbody>
</table>
predicted priming, $r = .37$, $p = .18$ across the 16 sounds. When controlling for unpleasant-pleasantness ratings, arousal did not predicting priming, $r = -.17$, $p = .54$. These data suggest that both valence and arousal act in tandem to produce priming effects, with valence perhaps explaining a bit more of the variance in priming effect than arousal. Nevertheless, with sounds, valence and arousal are strongly colinear.

**Discussion**

Our findings demonstrate that the evaluative system is likely independent of sensory modality, and that emotions that are activated by sounds (especially negative evaluations) can quickly and significantly influence responses to evaluative visual stimuli. Importantly, the present results extended past research by showing that negative auditory primes have a greater effect on visual processing than do positive auditory stimuli. This finding is consistent with an expanding literature suggesting that negative stimuli have a stronger psychological impact than positive stimuli (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Furthermore, these data suggest that valence and arousal levels of the primes may act together to produce priming. In the IADS-2 list, negative sounds are, on average, more arousing than pleasant sounds, and so it is possible that arousal and pleasantness are, to some degree, inextricably confounded for auditory stimuli. This negative correlation between affect and arousal is consistent with current theories that suggest that negative stimuli signal the need for change and action. For example, Baumeister et al. (2001) theorize that "it may be that humans and animals show heightened awareness of and responded more quickly to negative information because it signals a need for change" (p. 357). If negative sounds indeed signal the need for change and action, it follows that such stimuli will command greater attention and be experienced as relatively more arousing than sounds that do not require action.²

**Potential Mechanisms for Cross-Modal Evaluative Priming**

Currently, cross-modal priming effects can be explained by at least three different theoretical approaches. The first draws from a spread-

---

² One additional explanation of these asymmetrical effects is that the negative stimuli were more easily identifiable than the positive stimuli. We address this issue in three ways. First, we analyzed the average time taken to explicitly rate each of the emotional sounds and found that these reaction times were unrelated to either arousal or valence levels of the stimuli (both $r < .08$, ns). Assuming that participants must recognize the sounds before they can rate them, these results indicate that the negative sounds were not more quickly recognizable than the positive sounds. Second, if there is something inherent about negative stimuli that make them more easily recognizable, then participants should be faster to name negative than positive stimuli. To explore this hypothesis, we analyzed the positive and negative target words using the English Lexicon Project (ELP) database (Balota et al., 2002; available online at http://elexicon.wustl.edu/default.asp). Using the ELP database, we found that the naming times for the negative and positive target words used in this study were approximately equal (642 ms and 645 ms, respectively). Finally, we believe that the present experimental design makes it unlikely that the positive and negative words had different recognition rates, because participants were asked to listen to and rate the sounds just prior to the priming task, making all of the primes familiar.
Cross-modal priming could also be explained by response competition models of evaluative priming (for a review, see Klauer & Musch, 2003), and by the response mapping model recently proposed by Scherer and Lambert (2009). While those two models are different in many respects, both assume that primes activate a given response and that priming occurs when that response either facilitates or interferes with the response required by the target. Hence, cross-modal evaluative priming could occur if an emotion sound had the effect of activating a good or bad response, which then interfered with responding to a subsequent evaluatively incongruent target.

Finally, cross-modal priming can also be understood using an embodied cognition approach (e.g., Niedenthal, Barsalou, Winkelmann, Krauth-Gruber, & Ric, 2005). According to this perspective, emotions are not amodal representations but rather are represented by multimodal bodily states. For example, research has shown that participants who sit upright in a chair report feeling more pride than those who are instructed to slump (Stepper & Strack, 1993). Such influences of bodily states on emotion could potentially explain cross-modal priming by assuming that emotions are processed not by one particular mental system but by the interaction of many bodily systems together. Hence, the activation of an emotion via sound could potentially activate a cascade of changes in multiple bodily states, which could in turn influence processing in all of the body’s sensory modalities.

**Directions for Future Research**

We hope that the cross-modal priming effects demonstrated here will be generative, and hence we see a number of avenues for future research. First, it would be interesting to examine the evaluative connotations of the human voice, for example, how the timbre, tone, pitch, and accent of spoken language activates automatic evaluations. Second, we have demonstrated that auditory stimuli can influence evaluative reactions to visually presented words, but it might be important to show this effect in reverse (from sight to sound) and also demonstrate these effects using different paradigms, such as a lexical decision task (Neely, 1991) or the affective Simon task (De Houwer, Crombez, Baeyens, & Hermans, 2001). Third, it will be important for future research to investigate the asymmetries that are often observed across negative and positive stimuli to see if those asymmetries are equal in size and frequency in cross-modal versus unimodal paradigms. Finally, in this paper, we examined just one additional sensory modality, namely, sound. Yet all of our five senses contribute to our overall daily emotional experiences. Each sense probably influences behaviors that are both within and outside of our control, and each is worthy of empirical investigation.

**References**


Scherer, L. D., & Lambert, A. J. (2009). Contrast effects in evaluative...


Received October 8, 2009
Revision received September 23, 2010
Accepted October 25, 2010