How spatial attention and attentional resources influence the processing of emotional visual scenes

Aida Gutiérrez García and Manuel G. Calvo
Universidad de La Laguna

Pleasant, neutral, or unpleasant pictures appeared for 150 ms in either peripheral or foveal vision, with or without a concurrent foveal load task. Participants indicated whether the visual scene in the picture was or was not pleasant, or was or was not unpleasant. The manipulation of picture location (foveal vs. peripheral) aimed to tap spatial attention, while the perceptual load task was used to manipulate the availability of attentional resources. Results showed that emotional valence was discriminated above the chance level even in the attentionally-constrained conditions (peripheral presentation combined with perceptual load). Nevertheless, valence encoding depended on both attentional mechanisms, as indicated by reductions in accuracy and by slowed reaction times in valence identification when attention was allocated elsewhere, relative to when the scene appeared at fixation and when there was no concurrent task. This indicates that emotional processing requires attention.

Cómo la atención espacial y los recursos atencionales influyen en el procesamiento de escenas visuales emocionales. Se presentaron imágenes con valencia emocional positiva, neutra o negativa durante 150 milisegundos en visión periférica o foveal. Los participantes respondían si eran agradables o no, o si eran desagradables o no. Mediante la ubicación periférica o foveal de las imágenes se manipuló su acceso a la atención espacial; mediante la presencia de una tarea concurrente se manipuló la disponibilidad de recursos atencionales. Los resultados indicaron que la valencia emocional de las imágenes se discriminó incluso en las condiciones de menor restricción de la atención. No obstante, se produjeron decrementos en la probabilidad de identificar la valencia correcta y también aumentos en los tiempos de identificación tanto en la condición de presentación periférica como en la de tarea concurrente. Esto indica que el procesamiento de la valencia emocional no es automático, sino que depende de la atención espacial y de los recursos atencionales.

Emotional processing serves the purpose of assessing how events are related to our well-being. Emotions arise in response to stimuli that are related to threat or harm and to benefit or pleasure. For adaptive purposes, the cognitive system is expected to analyze with minimal attentional demands whether stimuli are good or bad. This would ensure that emotional stimuli are readily detected even when our limited attentional capacity is engaged in the processing of other concurrent stimuli.

In accordance with this view of the adaptive importance of emotional stimulus processing, several lines of research have examined automaticity in the affective assessment of visual scenes (see Calvo & Nummenmaa, 2007). For example, first, the visualization of affective pictures potentiates or inhibits automatic reactions such as the cardiac defense response or the startle reflex (Sánchez et al., 2002). Second, affective valence of pictorial stimuli can be extracted rapidly (within 200 ms or less) as indicated by intracranial recordings of the amygdala and ventral prefrontal cortex (Kawasaki et al., 2001). Third, emotional meaning is also perceived unconsciously, as shown by studies using subliminal presentation or masking paradigms (Banse, 2001). Fourth, emotional processing occurs involuntarily, even when it is task-irrelevant (Ferré, 2002).

In the current study, we have investigated an additional issue related to automaticity, namely, the requirement of attentional resources. To address this issue, we have made a distinction between attentional capacity, as a general-purpose resource for cognitive operations, and spatial attention, which involves selective processing of specific locations or stimuli (see Ruthruff, Allen, Lien, & Grabbe, 2008). Then we have applied this distinction to the investigation of the independent and the interactive role of each attentional function in emotional scene processing. This issue has not been satisfactorily addressed in prior research because the two attentional functions have been investigated separately rather than orthogonally combined.

In previous research, the role of spatial attention has been examined by presenting emotional and neutral scenes at extrafoveal vs. foveal locations in the visual field. When scenes were presented to the parafovea (2.5° away from a fixation point), Calvo and Lang.
(2005) found that semantic content of pleasant and unpleasant rather than neutral scenes was especially likely to be recognized. Calvo and Avero (2008) demonstrated affective priming, namely, faster encoding of probe scenes when they were preceded by prime scenes with congruent relative to incongruent emotional valence. This priming occurred when the prime scenes were presented under foveal gaze-contingent masking, and therefore could not be fixed. This suggests that affective scene content is analyzed outside the focus of spatial attention.

The role of available attentional capacity has been examined by presenting a non-emotional stimulus concurrently with an emotional distracter. Erthal et al. (2005) and Okon-Singer, Tzelgov, and Henik (2007) presented an emotional scene at fixation while a letter or two bars appeared peripherally; the participants made a speeded discrimination response to the letter/bars while ignoring the scene. Attentional load was varied by increasing the number of letters or the degree of similarity between the bars. These authors found that, in comparison with neutral scenes, unpleasant scenes interfered with task performance under low-load, but not under high-load conditions. When the attentional demands of the non-emotional task increased—thus reducing the available resources for affective processing—the interference caused by emotional stimuli decreased. This suggests that the processing of emotional stimuli is dependent on the availability of attentional resources.

In sum, there are discrepancies in prior research regarding the role of attention in emotional scene processing: While affective processing is not dependent on spatial attention, it is dependent on the availability of attentional capacity. In the current study, we attempted to clarify this issue and extend prior research in three respects. First, as indicated above, all the previous studies have considered the effect of spatial attention and attentional resources separately. In no study were spatial eccentricity and perceptual load varied simultaneously. This is necessary to estimate the combined effects and the potential interactions of the two attentional functions. To this end, we manipulated both factors orthogonally. Second, it is unclear whether both unpleasant and pleasant scenes are processed in a similar fashion. Whereas in some studies both unpleasant and pleasant scenes were presented (Calvo & Lang, 2005; Calvo & Avero, 2008), in others only unpleasant scenes were used (Erthal et al., 2005; Okon-Singer et al., 2007). To examine the potentially automatic processing of both aversive and appetitive stimuli, we included stimuli of both categories. And, third, in previous studies the scenes appeared in either an extrafoveal or a foveal condition. To this end, we manipulated both factors orthogonally. Second, it is unclear whether both unpleasant and pleasant scenes are processed in a similar fashion. Whereas in some studies both unpleasant and pleasant scenes were presented (Calvo & Lang, 2005; Calvo & Avero, 2008), in others only unpleasant scenes were used (Erthal et al., 2005; Okon-Singer et al., 2007). To examine the potentially automatic processing of both aversive and appetitive stimuli, we included stimuli of both categories. And, third, in previous studies the scenes appeared in either an extrafoveal or a foveal condition (Calvo & Lang, 2005; Calvo & Avero, 2008; Erthal et al., 2005).

To estimate the contribution of spatial attention to emotional processing, it is imperative to determine how much affective processing is reduced when the stimulus is presented extrafoveally in comparison with foveal presentation. This comparison was made in the present study.

We presented pictures (unpleasant, neutral, or pleasant in affective valence) for 150 ms either in foveal (i.e., at fixation) or peripheral (a minimum of 5.2° to the right or left of a central fixation cross) vision, either concurrently with or without a foveal letter-identification task (i.e., the load task). The experimental task involved the evaluation of whether the scene content was pleasant or not, or whether it was unpleasant not. We made the following predictions. If emotional processing is contingent on attentional resources, the load task will impair affect discrimination. If emotion perception is contingent on the focus of selective spatial attention, better valence discrimination is predicted under the foveal than the peripheral display conditions. In addition, perceptual load and scene location are predicted to interact, such that the load task will be especially detrimental on emotional processing in the peripheral scene condition.

Method

Participants

A total of 192 psychology undergraduates (144 female; 48 male) participated for course credit. Twenty-four of them (18 female) were randomly assigned to each of eight between-subjects experimental conditions (see Design). Although the absolute proportion of females and males was different within each condition, their relative proportion was the same across all the experimental conditions, and so any potential differences as a function of gender were controlled.

Stimuli and apparatus

Sixty-four target pictures portraying unpleasant (32) or pleasant (32) scenes, and 64 pictures portraying neutral scenes were used (see Calvo, Nummenmaa, & Hyönä, 2008). All scenes involved people. They were selected from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert 2005). In a 9-point scale, the mean valence scores of the selected pictures were 2.36 (unpleasant), 5.10 (neutral), and 7.67 (pleasant). A main effect of valence category (unpleasant vs. neutral vs. pleasant picture) emerged in a one-way ANOVA, $F(2, 125)= 812.00, p<.0001$. Post hoc HSD Tukey tests for multiple comparisons showed statistically significant differences between all three categories (all $p<.0001$).

Basic image statistics such as luminance, contrast, kurtosis, skewness, and energy were computed with Matlab 7.0 (The Mathworks, Natick, MA). There were no significant differences in these image characteristics between valence categories (see Calvo et al., 2008). The pictures subtended a visual angle of 13.3° (width) by 11.1° (height) at a viewing distance of 50 cm. Participants had their heads positioned on a chin rest. Pictures were presented in their original colors against a dark background on a 17” SVGA monitor connected to a Pentium-IV computer. The E-Prime experimental software controlled stimulus presentation and response collection.

Procedure

Figure 1 shows the sequence of events on each trial. A trial started with a central fixation cross for 500 ms. This was followed by one target picture and a central letter (A or O, subtending 1° × 0.8°) for 150 ms. In the peripheral location condition, the target picture appeared either to the left or right of fixation, at the same time as a meaningless picture (a random combination of colors; same size as the target picture) on the opposite side. In the foveal condition, the target picture appeared with the central letter superimposed. After a 500-ms mask display, a prompt to respond whether the target scene was unpleasant or not (or pleasant or not) was presented. In the perceptual load condition, a prompt to respond whether the letter was an A or an O was presented after the participant had classified the scene valence. The inter-trial interval was 1,500 ms.

The participants were told that they would be presented with a photograph and a letter (A or O, each 50% of the times) on each
The participant’s task was to attempt to identify the affective valence of the scene and also to identify the letter (only in the load condition). For both tasks, the participants were to respond with a keypress on keys labeled as YES or NO as soon as possible. Response accuracy and reaction times were collected. There were 26 practice trials and 128 experimental trials.

**Design**

The experimental conditions were combined in a mixed factorial design, with valence (pleasant vs. neutral vs. unpleasant) of the target pictures as a within-participants factor, and picture location (foveal vs. peripheral), concurrent load (yes vs. no), and evaluation instructions (unpleasantness vs. pleasantness), as between-participants factors. Each of the 128 target pictures was presented once to each subject in random order.

In the peripheral condition, the distance between the centre of the letter that served as the fixation point and the inner edge of the lateralized pictures was 5.2° of visual angle. In the foveal condition, the target scene appeared at fixation, with the letter in the center. In the load condition, the participants had to identify the central letter appearing at the same time as the picture and keep the letter in working memory during the scene valence evaluation. In the no load condition, the participants were told that the letter would appear simply as a fixation point. In the unpleasantness-evaluation condition, the participants responded whether the scene was unpleasant or not; in the pleasantness-evaluation condition, the participants responded whether the scene was unpleasant or not.

The same exposure and eccentricity parameters used in the current study (i.e., 150-ms picture display at a 5.2° distance from the fixation point) in the peripheral location condition have been found to prevent fixations on the pictures in previous research. In Calvo et al. (2008), the mean latency of the first saccade to the picture was 175 ms (hence above the 150-ms display); the probability that the peripheral picture was fixated (within the 150-ms display) was less than 1%; and the mean fixation time in these few cases was only 4 ms. This implies that, in the peripheral condition of the current study, the pictures were also very unlikely to be fixated, and therefore remained in peripheral vision.

**Data analysis**

By means of the SPSS 17.0 for Windows, the following analyses were performed. First, response accuracy and reaction times on the letter-identification task were analyzed in mixed-model factorial ANOVA (evaluation instructions by location by valence), with repeated measures for valence. Second, a repeated-measures one-way (valence) ANOVA was conducted on the probability of responding «unpleasant» or «pleasant» to the scenes, followed by post hoc Bonferroni-corrected multiple contrasts, to determine that there was discrimination of affective content. Third, one-sample t tests were computed to examine whether the probability of correct responses in the affective evaluation task exceeded the chance level (i.e., the empirical probability scores were compared to the .50 test value). Finally, the probability that scene valence was correctly identified and the response latencies were analyzed by means of a mixed-model factorial ANOVA (evaluation instructions by location by load by valence; the neutral scene category was excluded as these scenes are not emotional), with repeated measures for the valence factor.

**Results**

**Performance Accuracy and Reaction Times on the Concurrent Letter-identification Task**

The 2 (instructions) × 2 (location) × 3 (valence) ANOVA yielded no statistically significant effects on letter identification. All the Fs were less than 1.0 (and p > .38), except for the main effects of valence on reaction times, F(2, 184) = 1.86, p = .16, ns (M = 404 vs. 417 vs. 405 ms, for unpleasant vs. neutral vs. pleasant scenes, respectively). The mean probabilities of correct responses were
near ceiling in all the experimental conditions (M accuracy=.935). The high accuracy and the lack of differences between conditions rule out the possibility of trade-offs that might affect the valence identification task.

**Performance on the Affective Evaluation Task: Discrimination between Valence Categories**

As indicated by one-way (valence) ANOVAs, in the unpleasantness evaluation condition, the probability that participants responded as «unpleasant» to the unpleasant scenes was greater than for the neutral scenes, which was greater than for the pleasant scenes, F(2, 190)=1.080.38, p<.0001 (M=.86 vs .21 s .09, respectively; all ps<.0001, after Bonferroni corrections). Conversely, in the unpleasantness evaluation condition, the probability that participants responded as «pleasant» to the pleasant scenes was greater than for the neutral scenes, which was greater than for the unpleasant scenes, F(2, 190)=525.25, p<.0001 (M=.92 vs .50 s .15, respectively; all ps<.0001). This revealed that participants discriminated between scenes as a function of affective content. Details of these comparisons are presented in Table 1 for each location and load condition.

**Performance on the Affective Evaluation Task: Response Accuracy above the Chance Level**

One-sample t tests showed that, for every combination of location and load (see Table 1), unpleasant scenes were responded to as «unpleasant» above the chance level (all ts(23)>7.60, p<.0001). In contrast, neutral scenes (all ts(23)>7.50, p<.0001) and pleasant scenes (all ts(23)>18.00, p<.0001) were responded to as «unpleasant» below the chance level, and thus they were perceived as not being unpleasant. Conversely, pleasant scenes were responded to as «pleasant» above the chance level in all the experimental conditions (all ts(23)>13.00, p<.0001). In contrast, neutral scenes were responded to as «pleasant» at chance (all ts(23)<1, p>.60), and unpleasant scenes were responded to as «pleasant» below the chance level (all ts(23)>4.30, p<.0001); thus neutral scenes were not perceived as pleasant, and unpleasant scenes were perceived as not being pleasant.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Unpleasantness Evaluation</th>
<th>Pleasantness Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unpleasant</td>
<td>Neutral</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foveal-No Load</td>
<td>.929</td>
<td>.163</td>
</tr>
<tr>
<td>Foveal-Load</td>
<td>.907</td>
<td>.159</td>
</tr>
<tr>
<td>Peripheral-No Load</td>
<td>.869</td>
<td>.220</td>
</tr>
<tr>
<td>Peripheral-Load</td>
<td>.753</td>
<td>.280</td>
</tr>
</tbody>
</table>

**Note:** Scores with different superscripts (horizontally) are significantly different (after Bonferroni corrections for multiple comparisons); * above/below chance level (after one-sample t tests against the .5 probability value). For reaction times for correct responses, the ANOV A revealed main effects of valence, location, and load, F(1, 184)=707.80, p<.0001, and load, F(1, 184)=8.89, p<.01, with no interactions. Response latencies were shorter for pleasant than for unpleasant scenes (M=725 vs. 746 ms), in the foveal than in the peripheral location (M=460 vs. 1,012 ms), and under no-load than under load conditions (M=705 vs. 767 ms). There were no significant interactions, with load slowing down responses both in the peripheral, F(1, 94)=3.98, p<.05, and the foveal, F(1, 94)=10.12, p<.01, conditions. See figure 2.

**Discussion**

The results of the current study showed that the affective valence of visual scenes can be discriminated to a significant extent even when spatial attention is directed elsewhere such that the scene cannot be overtly attended to and attentional resources are used by a concurrent non-emotional task. Nevertheless, the encoding of emotional valence is additively and interactively impaired by both types of attentional constraints, thus suggesting that affective processing is contingent on the availability of attentional resources. These findings are relevant to examine the role automaticity in emotional processing.
Emotional Content Can to Some Extent Be Reliably Processed with Reduced Attention

Prior studies have provided some evidence that the affective valence of pictures can be assessed when these appear in peripheral vision. First, in eye-movement studies using simultaneous presentation of emotional and neutral pictures matched in low-level physical image properties, the probability of first fixation was higher for emotional than for neutral scenes (Calvo & Lang, 2004; Calvo et al., 2008; Nummenmaa, Hyönä, & Calvo, 2006). As attentional orienting was initiated before a picture was fixated, this implies that the viewers had perceived something of the emotional scene content while it was still in peripheral vision, which then selectively captured overt attention. Second, in recognition studies using paired emotional and neutral prime scenes, an increased false alarm rate for emotional vs. neutral scenes has been found in display conditions that prevent fixations (Calvo, 2006; Calvo et al., 2008). This suggests that the information acquired peripherally is relatively vague, such as a general affective impression of the scene (whether something «good» or «bad» occurs). False alarms to conceptually similar but visually different probe pictures have been interpreted as an indication of meaningful, although coarse, processing of prime pictures (Potter, Staub, & O’Connor, 2004). Coarse processing would thus lead to erroneous stimulus identification, as revealed by false alarms to scenes that are related in affective valence, and this coarse processing would be sufficient to induce selective overt orienting of attention.

Nevertheless, in the Okon-Singer et al. (2007; Experiment 1) study, peripheral emotional scenes did not impair the performance of a concurrent neutral task presented at fixation. This suggests that affective content is not always extracted from extraveival scenes. It is, however, possible that the interference paradigm used by Okon-Singer et al. is not sensitive enough, as it involves measurement of the processing of emotional stimuli through their influence on an unrelated task. We have argued that emotional valence can be processed coarsely in the visual periphery, albeit lacking in detail about the specific actions depicted in the scene. The scene details (e.g., high-spatial frequencies) are not required for processing of its «affective gist» (or a coarse impression) (see Carretié, Hinojosa, López-Martín, & Tapia, 2007). If so, in the absence of a vivid, detailed representation of the scene content, affective valence may not be sufficiently salient to cause interference with other ongoing cognitive tasks. Yet such a coarse representation would be sufficient to produce effects on attentional orienting and recognition. Actually, in the current study, performance on the concurrent letter-identification task did not differ (and so was not impaired) as a function of scene valence, yet emotional valence was discriminated. This shows that experimental paradigms building on interference with concurrent task performance may not provide sensitive measures of emotional processing.

All three measures (attentional orienting, false alarms, and interference) used in previous studies are indirect indices of whether the emotional valence of «unattended» scenes is processed. In the current study we used a straightforward index, by asking participants to explicitly evaluate the scene valence. Two findings corroborate the view that emotional significance can be reliably identified both under constrained spatial attention and reduced attentional resources, i.e., in the visual periphery and under perceptual load. First, response accuracy was well above chance level for pleasant and unpleasant scenes in all the experimental conditions, even the one that combined peripheral presentation with perceptual load. Second, comparisons across scene valence categories (unpleasant, neutral, and pleasant) clearly indicated that these were discriminated from each other. Consistently, Meseguer et al. (2007) found that, relative to neutral pictures, pleasant and unpleasant IAPS pictures not only activated some common brain areas (e.g., amygdala), but also some specific areas (e.g., orbitofrontal cortex: pleasant; e.g., dorsolateral prefrontal cortex: unpleasant) under passive picture viewing conditions, while participants were involved in a letter-discrimination task, thus showing valence discrimination when attention was allocated to a concurrent task.

…Yet Emotional Processing Depends on Spatial Attention and the Availability of Attentional Resources

We have seen that emotional scene valence can be identified under reduced spatial attention and attentional resources. The next step involves exploring whether and how much the processing of emotional valence depends on attention. To assess the role of spatial attention, we compared a foveal and a peripheral presentations condition. To assess whether emotional content might have access to the cognitive system even when attentional resources are used by a concurrent task, we compared emotional processing under perceptual load vs. no load conditions. Our results indicated that the processing of emotional valence is dependent on spatial attention and is also resource-limited. Accuracy in the identification of emotional valence was impaired and slowed down in the peripheral vs. the foveal condition, and in the load vs. the no-load condition.

Some prior studies have provided relevant data concerning the dependence of emotional scene processing on attention. Okon-Singer et al. (2007; Experiment 1) found that emotional processing—as assessed by interference with a concurrent non-emotional task — was eliminated when unpleasant scenes were presented peripherally. In addition, Okon-Singer et al. (2007; Experiment 2), as well as Erthal et al. (2005), found that unpleasant scenes interfered with performance on a concurrent non-emotional task under low-load, but not under high-load conditions. Consistently, in our study, emotional processing was reduced by constraints on both spatial attention and attentional load. In addition, beyond prior studies in which only unpleasant pictures were used (Erthal et al., 2005; Okon-Singer et al., 2007), we included both unpleasant and pleasant scenes. It has been frequently assumed that, for adaptive reasons, aversive stimuli are more relevant than appetitive stimuli, and therefore an automatic processing advantage should exist for negatively-valenced content. The present data show that encoding of both unpleasant and pleasant scene content processing is attention-dependent.

Another novel contribution of the current study concerns the relative roles of spatial attention and attentional resources. This issue was addressed by an orthogonal combination of the scene spatial eccentricity and the concurrent task. In prior studies these two factors have been varied independently (Calvo & Lang, 2005; Calvo & Nummenmaa, 2007; Erthal et al., 2005; Okon-Singer et al., 2007). The interaction between scene location and concurrent load reveals that attentional demands affect valence identification when the scenes appear peripherally more than when they appear at fixation. Presumably, in the load condition combined with peripheral scene presentation, viewers need to distribute
their attentional resources by overtly attending to the letter and covertly attending to the scene. As attentional capacity is required to process the emotional valence of scenes, this condition leads to more pronounced decrements in scene valence encoding. In the peripheral scene condition, viewers need to widen the scope of spatial attention to incorporate in it both the foveal letter and the peripheral scene, but this is not needed when the letter and the picture appear superimposed at fixation (i.e., in foveal vision). As the required scope of spatial attention is smaller in the foveal presentation condition, less disruption is caused by the load task.

In conclusion, the affective significance of visual scenes can be encoded even when attention is reduced, either by preventing fixations on the scenes or by allocating cognitive resources to a concurrent task. The findings of affective encoding when attention is allocated elsewhere suggest that there is parallel processing of emotional content. Nevertheless, affective processing is not totally automatic. Both when spatial attention to the emotional scene is constrained and when attentional load is imposed, accuracy in valence identification is reduced and the identification speed is slowed down. Accordingly, even though emotional processing can be automatic in the sense of being fast, involuntary, and performed in parallel with other unrelated tasks, it is dependent on attention. The current study has shown the combined effects and the interactive roles of spatial attention and attentional resources on emotional scene processing.

Acknowledgements

This research was supported by Grant PSI2009-07245 from the Spanish Ministry of Science and Innovation.

References


