Journal of Experimental Psychology: General
2010, Vol. 139, No. 2, 222–246

0096-3445/10/$12.00 DOI: 10.1037/a0018858

Semantic Categorization Precedes Affective Evaluation of Visual Scenes

Lauri Nummenmaa
Aalto University School of Science and Technology and Turku PET Centre

Jukka Hyönä
University of Turku

Manuel G. Calvo
University of La Laguna

We compared the primacy of affective versus semantic categorization by using forced-choice saccadic and manual response tasks. Participants viewed paired emotional and neutral scenes involving humans or animals flashed rapidly in extrafoveal vision. Participants were instructed to categorize the targets by saccading toward the location occupied by a predefined target scene. The affective task involved saccading toward an unpleasant or pleasant scene, and the semantic task involved saccading toward a scene containing an animal. Both affective and semantic target scenes could be reliably categorized in less than 220 ms, but semantic categorization was always faster than affective categorization. This finding was replicated with singly, foveally presented scenes and manual responses. In comparison with foveal presentation, extrafoveal presentation slowed down the categorization of affective targets more than that of semantic targets. Exposure threshold for accurate categorization was lower for semantic information than for affective information. Superordinate-, basic-, and subordinate-level semantic categorizations were faster than affective evaluation. We conclude that affective analysis of scenes cannot bypass object recognition. Rather, semantic categorization precedes and is required for affective evaluation.

Keywords: emotion, recognition, eye movements, attention, scene perception

When you stroll around in the garden and suddenly notice a green, tube-shaped, and curvy object moving around the grass, what do you notice first? Do you recognize the object as a snake, or do you instead detect that there is something potentially threatening in your vicinity? According to the *affective primacy hypothesis*, you will first notice the potential danger associated with the event and then semantically recognize the object as a snake (LeDoux, 1996; S. T. Murphy & Zajonc, 1993; Stapel, Koomen, & Ruys, 2002; Zajonc, 1980). In contrast, on the basis of the view that affective evaluation is dependent on cognition, we may argue that you first need to recognize the object as a snake before proceeding with the affective evaluation (Calvo & Nummenmaa, 2007; Lazarus, 1984; Rolls, 1999; Storbeck & Clore, 2007; Storbeck, Robinson, & McCourt, 2006).

Emotional processes assess the importance of sensory events to our well-being and adjust our physiological, behavioral, and cognitive responses to cope with challenges (Lang, 1995). Emotions occur when we observe potentially harmful or beneficial events, and they prepare us to engage in appropriate approach–avoidance behavior. This requires that the cognitive systems in the brain effectively decide whether an object or an event is good or bad for the organism. An important issue is whether this emotional evaluation can occur prior to cognitive assessment of the nature of the object or event, that is, whether it can be classified as good or bad prior to knowing what it is. The relationship between affect and cognition—more specifically, the temporal primacy of one over the other—has been an enduring question in the field of cognitive and affective science (Calvo & Nummenmaa, 2007; Lazarus, 1984; Storbeck & Clore, 2007; Zajonc, 1980).[1] It is highly relevant to the issue of whether and how quickly affective processes can modulate the information processing priorities of the brain and whether affective analysis is itself contingent on basic perceptual and cognitive analysis.

In the present study, we assessed the temporal primacy of affect versus cognition in the initial evaluation of visual sensory input by measuring the minimum as well as the typical latency required for affective and semantic categorization of complex pictorial scenes. In a series of seven experiments, we used brief presentations of paired (Experiments 1–3 and 7) or singly presented (Experiments

Lauri Nummenmaa, Low Temperature Laboratory, Brain Research Unit, and Department of Biomedical Engineering and Computational Science, Aalto University School of Science and Technology, Espoo, Finland, and Turku PET Centre, Turku, Finland; Jukka Hyönä, Department of Psychology, University of Turku, Turku, Finland; Manuel G. Calvo, Department of Cognitive Psychology, Universidad de La Laguna, La Laguna, Spain.

This research was supported by Academy of Finland Grant 121031 to Lauri Nummenmaa, by the aivoAALTO Grant from Aalto University, and by Spanish Ministry of Science and Innovation Grant PSI2009-07245 to Manuel G. Calvo.

Correspondence concerning this article should be addressed to Lauri Nummenmaa, Low Temperature Laboratory, Brain Research Unit, Aalto University School of Science and Technology, 00076 Aalto, Finland. E-mail: nummenmaa@neuro.hut.fi

[1] It is evident that there is no universal answer to the primacy of affect versus cognition in the most general sense. Some cognitive operations may occur earlier (c.f., this study) and some are bound to occur later (such as abstract reasoning) than affective processing. Accordingly, in the context of the present study, the primacy of affective versus semantic processing refers to the initial processing of novel sensory input in the visual domain.

4–6) target and nontarget scenes and measured the latency of eye movements or manual responses in a two-alternative forced-choice (2AFC) paradigm, in which the viewers must discriminate between a target scene and a nontarget scene with respect to semantic category (e.g., animal vs. human: semantic task) or emotional content (e.g., unpleasant or pleasant vs. nonemotional: affective task). We demonstrate that both affective and semantic recognition can be performed accurately within 180–220 ms, with an average latency of 300 ms, but that semantic categorization practically always precedes affective evaluation.

Fast and Automated Processing of Affect

In models of emotional processing, it is assumed that affective processing is quick, effortless, and automatic (Bargh, 1997). Some theorists (S. T. Murphy & Zajonc, 1993; Zajonc, 1980) have taken this argument even further by claiming that affective analysis of an event may even precede the semantic recognition or identification of the event. A critical assumption of such affective primacy models is that emotional processing does not require detailed perceptual processing or semantic identification of the stimulus to be accomplished. In other words, it is postulated that recognition of the emotional meaning of visual events occurs prior to or independently of semantic recognition of the perceived objects or their parts. In several neurophysiological theories of emotion, independence of affective and semantic recognition systems in the brain is also assumed, and several recent reviews of the neurophysiology of emotions have established that specialized neural circuits are devoted to the detection of affect and generating emotions (Kober et al., 2008; LeDoux, 1995; F. C. Murphy, Nimmo-Smith, & Lawrence, 2003). For example, one view (LeDoux, 1995; Morris, Ohman, & Dolan, 1998b; Vuilleumier, 2005) posits that the affective significance of sensory inputs could be crudely appraised via an extrageniculostriate pathway projecting from the pulvinar thalamus to the amygdala. Some forms of affective evaluation could thus occur even before visual information is transmitted to the striate cortex and the temporal regions involved in object recognition. On the contrary, object recognition (i.e., semantic recognition) is thought to rely rather strictly on the ventral visual pathway through neurons with increasingly large receptive fields responding to increasingly complex stimulus features (Goodale & Milner, 1992). The stream starts with the early visual areas extracting physical image properties (V1, V2, V3) and ends with inferior temporal (IT) cortex. Cells in the IT display object-selective, size, and rotation invariant response properties (Desimone, Albright, Gross, & Bruce, 1984; Perrett, Hietanen, Oram, & Benson, 1992; Tanaka, 1996) and probably underlie perceptual recognition.

Three lines of empirical studies support or are at least consistent with the affective primacy hypothesis. First, psychophysiological studies have shown that affective evaluations may occur even when the stimuli are presented outside the observer's awareness (i.e., subliminally). Affective valence of briefly presented and visually masked stimuli evokes detectable changes in the viewers' peripheral physiology (facial electromyographic responses: Dimberg, Thunberg, & Elmehed, 2000; electrodermal responses: Glascher & Adolphs, 2003; Ohman & Soares, 1998) and in the brain regions associated with affective processing (Morris et al., 1998b), thus suggesting affective processing in the absence of conscious recognition. Second, studies with the affective priming paradigm have implied that affective processing of stimuli is automatic and mandatory (see a review in Klauer & Musch, 2003). These studies have shown that affective valence of to-be-ignored prime scenes influences the affective categorization speed of subsequently presented probe scenes. Furthermore, priming effects may even occur when the primes are presented outside of awareness (Hermans, Spruyt, De Houwer, & Eelen, 2003) and when affective congruency between prime and probe scenes is manipulated in a concurrent semantic priming task (i.e., when the semantic category of the probe is task-relevant and the affective category is task-irrelevant; Calvo & Nummenmaa, 2007). Third, the emotional valence of events is automatically assessed from eccentric and unattended locations of the visual field (Calvo & Nummenmaa, 2007), which implies that detailed, attentive processing is not necessary for affective evaluation. Also, visual search studies with extrafoveally presented unpleasant faces (Öhman, Lundqvist, & Esteves, 2001) and scenes (Öhman, Flykt, & Esteves, 2001) suggest that affective valence may be processed in the absence of attention, as indexed by near-flat search slopes when searching for affective targets (although see Calvo & Nummenmaa, 2008; Horstmann, 2007, for alternative explanations).

Does Affective Processing Precede Semantic Recognition?

Different models of object recognition have different assumptions regarding the hierarchy of categorization processes. For example, Rosch and colleagues (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) argued that the visual stimuli are first classified at the basic level (i.e., the most abstract level in which visually similar objects still belong to same category), after which the more exclusive (subordinate) and inclusive (superordinate) levels are accessed. On the contrary, the parallel distributed processing theory (McClelland & Rogers, 2003) posits that object representations are activated from broad to fine, so that broad categories are accessed first and narrow categories are accessed last. Nevertheless, a common assumption of most models of visual recognition is that one category level always serves as the entry level that is accessed first. Against this background, it is interesting to ask (a) whether the affective valence of scenes would actually be the entry level category of visual scenes (i.e., whether valence would be the first representation to be activated) and, if not, (b) when does affective processing occur along the hierarchy of the levels of object categorizations?

The studies reviewed above suggest that affective information may have special access to awareness. But does affective categorization occur prior to semantic recognition? Recent evidence stemming from patient studies and behavioral studies, as well as from electromagnetic experiments, seems to challenge this view. First, studies on affective blind sight (i.e., affective categorization in the absence of striate cortex) suggest that cortical processing similar to that involved in object recognition may be necessary for affective judgments. Although certain patients show somewhat accurate classification of facial expressions presented to their blind visual field (de Gelder, Vroomen, Pourtois, & Weiskrantz, 1999), it seems that this effect is restricted to facial expressions and does not extend to complex affective scenes (de Gelder, Pourtois, &

Weiskrantz, 2002) whose processing seems to require inputs to striate cortex.

Second, recent affective priming studies also imply that affective analysis requires recognition of the scene objects. Priming effects for emotional scenes are dependent on prior exposure to the scenes (Calvo & Nummenmaa, 2007), thus suggesting that object or scene recognition must occur prior to emotion perception. In a similar vein, in a study investigating affective priming of angry and happy facial expressions with liked versus disliked facial identities, it was found that conscious recognition of facial identities was necessary for the priming effects to occur (Nummenmaa, Peets, & Salmivalli, 2008). Third, recent studies on the visual search of faces and abstract fear-conditioned shapes have failed to establish pop-out effects for emotional targets, as revealed both by steep search slopes (Batty, Cave, & Pauli, 2005; Fox et al., 2000; Horstmann & Bauland, 2006) and the need to foveally examine the discrepant faces prior to detecting the target (Calvo, Nummenmaa, & Avero, 2008). Taken together, these data imply that affective information cannot be processed preattentively (or fully in parallel) and independently of cognition. Instead, serial attentive processing may be necessary.

The primacy of semantic over affective processing is also supported by electrophysiological recordings, although studies directly comparing the processing latency of affective and semantic categorization are relatively sparse. Electromagnetic studies of visual categorization in humans (Liu, Harris, & Kanwisher, 2002; Meeren, Hadjikhani, Ahlfors, Hämäläinen, & de Gelder, 2008; Mouchetant-Rostaing, Giard, Delpuech, Echallier, & Pernier, 2000) have revealed that processing of single objects may start as early as 100 ms poststimulus, as indexed by category-selective effects on inverted versus upright presentation of the stimuli. Intracranial field potentials recorded from visual cortices (including IT) have demonstrated that object-selective responses to object transformations occur already within 100 ms from stimulus onset (Liu, Agam, Madsen, & Kreiman, 2009). Categorization of complex scenes (Rousselet, Fabre-Thorpe, & Thorpe, 2002) may begin slightly later, with differential event related potentials (ERPs) emerging around 150–170 ms from stimulus onset, although behavioral studies have suggested that conscious recognition of complex scenes may occur as early as 120 ms poststimulus (Kirchner & Thorpe, 2006).

In contrast with the very early categorization of objects, the earliest reliable signs of facial affect processing are typically observed around 170 ms poststimulus, indexed by the face-selective N170 potential being reliably modulated by facial expressions (Batty & Taylor, 2003; Eimer & Holmes, 2002; Pizzagalli, Lehmann, Koenig, Regard, & Pascual-Marqui, 2000). For complex emotional scenes, the results are less conclusive. A limited number of intracranial recordings of the amygdala (Oya, Kawasaki, Howard, & Adolphs, 2002) and the medial prefrontal cortex (Kawasaki et al., 2001) suggest that elementary affective classification of visual scenes could be accomplished in less than 150 ms. A number of extracranial studies (Carretié, Hinojosa, López-Martín, & Tapia, 2007; Carretié, Hinojosa, Martín-Loeches, Mercado, & Tapia, 2004) also report an early affective ERP modulation around the P1 latency (>100 ms poststimulus) range. However, such results are difficult to interpret, as these early potentials are also sensitive to physical image properties that are devoid of any meaning. It has actually been shown (Bradley, Hamby, Low, & Lang, 2007) that image composition influences the early, but not the late, posterior (and frontal) components, such that no early differences between affective and neutral pictures are found when image complexity is controlled for. It is thus possible that stimulus selection may explain the great variability (for a review, see Olofsson, Nordin, Sequeira, & Polich, 2008) in results within the early latency range of ERPs evoked by complex pictorial scenes. For ERPs with longer latencies (200 to over 300 ms poststimulus) the findings are more consistent, with emotional (i.e., arousing) rather than neutral pictures resulting in early posterior negativity (EPN) 200–300 ms poststimulus (e.g., Cuthbert, Schupp, Bradley, Birbaumer, & Lang, 2000; Junghöfer, Bradley, Elbert, & Lang, 2001; Schupp, Junghöfer, Weike, & Hamm, 2003; Schupp, Junghöfer, Weike, & Hamm, 2004).

The Current Study

Taken together, the electrophysiological studies suggest that systematic ERP modulation by affective content emerges later than does the earliest nonemotional category-specific responses, although the possibility for early (~100 ms) affective evaluation cannot be ruled out. Nevertheless, the evidence is rather mixed and indirect, as none of these studies has directly compared emotional and semantic processing in the same experiment. Moreover, we are not aware of any studies that would have systematically compared affective categorization with semantic categorization occurring at different levels (superordinate, basic, and subordinate), which would allow one to systematically assess the latency of different affective–semantic categorization operations. It should also be noted that results from electrophysiological studies on visual processing speed are difficult to interpret, as it is not completely clear whether the observed early ERP latency and amplitude differences reliably index conscious recognition of the scene's emotional content or valence. In contrast, behavioral measurements have an advantage over psychophysiological recordings because the former yield an estimate of conscious recognition of the scene content. Although manual responses are slow (taking over 250 ms to execute) and thus distort the latency estimates, eye movement responses are well suited for studying visual processing speed, as they are much faster than are manual responses, with minimum latencies of visually guided saccades below 100 ms (Fischer & Weber, 1993). In fact, one previous study (Kirchner & Thorpe, 2006) used simultaneous rapid presentation of target and distracter scenes, combined with a task requiring participants to saccade to the target scene as fast as possible. This study established that conscious semantic classification of complex visual scenes (i.e., deciding whether the scenes involve animals or not) may occur in less than 120 ms, which is considerably faster than the lower-bound latency estimates of affective classification arising from most ERP studies. We thus decided to take advantage of this methodology to address the issue of processing speed for affective versus semantic information.

We used five methodological developments. First, we directly compared with each other the latency of affective and semantic categorization of visual scenes. To that end, we used saccadic latencies in a 2AFC classification task of semantic and affective target scenes. Measurement of saccade latencies enabled us to pinpoint more accurately the speed of affective (i.e., discrimination between pleasant or unpleasant and neutral content) and semantic (i.e., discrimination between animal and human scenes) recognition. Second, we compared both the typical and the minimum latencies of correct saccades toward predefined targets (e.g., unpleasant or animal) and examined the response latency distributions by using both by participants (F1) and by items (F2) analyses. The F2 analyses enabled us to assess whether a subset of affective stimuli would be faster to categorize than the respective semantic stimuli and whether some stimulus types would enjoy a more significant processing advantage than would others. Third, given the proposed significance (Morris, Ohman, & Dolan, 1999; Vuilleumier, 2005; Vuilleumier, Armony, Driver, & Dolan, 2003) of the magnocellular visual system (projecting from the peripheral retina) in emotion perception, we contrasted the recognition speed and accuracy of affective versus semantic target scenes presented in either foveal or extrafoveal vision. Fourth, we compared the minimum exposure thresholds for affective and semantic recognition in order to assess how much visual input is required for each type of task. Fifth, we compared different levels of semantic recognition against affective processing to estimate whether affective recognition can bypass all, some, or no object recognition stages.

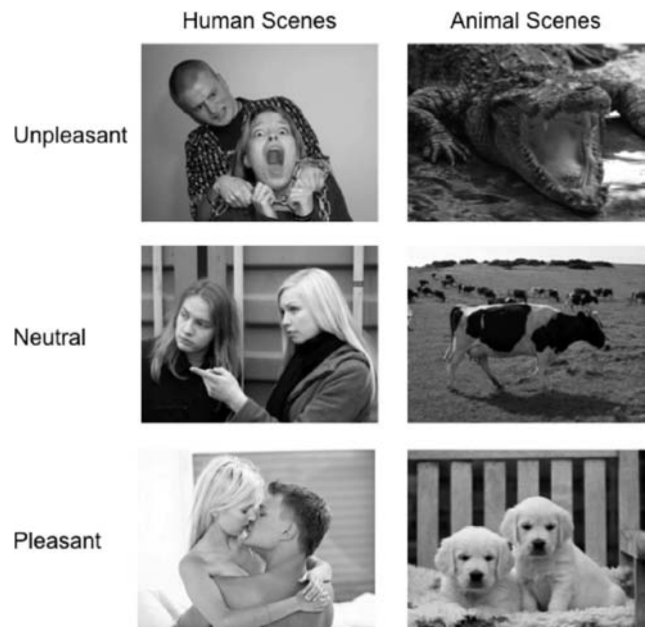

Figure 1. Illustration of unpleasant, neutral, and pleasant scenes used in Experiments 1–7. Note that these example pictures were not among the experimental stimuli.

Experiment 1: Estimating the Speed of Emotion Recognition From Scenes

In Experiment 1, we benchmarked the 2AFC saccadic response paradigm for categorization of emotional scenes and estimated the lower bound latency of affective classification of complex emotional targets. On each trial, two scenes were briefly flashed to participants: one emotional scene and one neutral scene that appeared simultaneously. The participants' task was to perform as quickly as possible a saccade to the location where the emotional scene (unpleasant or pleasant, depending on the counterbalancing condition) was flashed. We measured both the median latency of correct saccades and the lowest latency bin at which correctly directed saccades were more frequent than erroneously directed ones. This way we could reveal both the typical and the minimum processing speed required for affective evaluation.

Method

Participants. A sample of 24 undergraduate psychology students (17 female; 7 male; with a mean age of 21 years) from the University of La Laguna (La Laguna, Spain) participated for course credit. In this and the following experiments, all the participants gave informed consent and had normal or corrected-to-normal vision.

Apparatus. Stimuli were presented on a 21 in. (53.34 cm) monitor (120 Hz refresh rate) with a 3.2 GHz Pentium IV computer. Participants' eye movements were recorded with an EyeLink II eyetracker (SR Research, Mississauga, Ontario, Canada) connected to a 2.8 GHz Pentium IV computer. The sampling rate of the eyetracker was 500 Hz, and the spatial accuracy was better than 0.5°, with a 0.01° resolution in the pupil-tracking mode.

Materials. The stimuli (see Figure 1 for illustrations) were 128 pictures selected from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2005; Center for the Study of Emotion and Attention, 2005). There were 32 unpleasant, 32 pleasant, and 64 neutral pictures (see Appendix). The unpleasant pictures represented people suffering serious threat or harm. The pleasant pictures portrayed people showing or experiencing positive affect. The neutral pictures depicted people in daily nonemotional activities. Valence and arousal ratings (ranging from 1 to 9, see Table 1) for each picture have been obtained in norming studies (Lang et al., 2005). Valence (unpleasantness vs. pleasantness) reflects the dominant motive system activated (avoidance or approach). Arousal reflects the intensity of the motive system activation, from calm to tension. For the stimuli used in the current study, a one-way analysis of variance (ANOVA) yielded an effect of stimulus type on valence ratings, $F(2, 127) = 492.68$, $p < .001$, $\eta_p^2 = .93$, with significant differences among all three stimulus categories (for all post hoc comparisons, *p*s $< .001$; Bonferroni corrected). There was also a stimulus type effect in arousal ratings, $F(2, 127) = 91.11$, $p < .001$, $\eta_p^2 = .80$, with significantly higher ratings for the unpleasant and pleasant stimuli than for the neutral stimuli (*p*s $< .001$), but no difference between the pleasant stimuli and the unpleasant stimuli.

As eye movements during initial stages of scene perception are mainly driven by visual features rather than semantic features (Henderson, 2003; Itti & Koch, 2001), we controlled for potential confounds due to low-level visual image properties by computing with Matlab 7.0 (The MathWorks, Natick, Massachusetts) basic image statistics (see Table 1), including mean luminosity, standard deviation of luminosity, contrast density (root mean square [RMS] contrast; Peli, 1990), skewness, kurtosis, and energy. The one-way ANOVAs showed no significant differences between the valence categories in any of the image characteristics (*F*s ≤ 1.60, $p \geq .10$; for energy), $F(2, 125) = 3.90$, $p \geq .072$, *ns*, after Bonferroni corrections for multiple contrasts.

Stimulus displays. See Figure 2 for illustrations of stimulus displays. The initial fixation target was a white circle with a black

Table 1
Means and Standard Deviations of Stimulus Characteristics of the Unpleasant, Neutral, and Pleasant Stimuli Involving Humans Used in Experiments 1–4

	Unpleasant		Neutral		Pleasant	
Characteristic	*M*	*SD*	*M*	*SD*	*M*	*SD*
Valence rating	2.60_a	0.64	5.27_b	0.59	7.20_c	0.53
Arousal rating	5.59_a	1.00	3.58_b	0.55	5.45_a	1.03
Luminance (average)	99.35	19.66	101.06	19.55	109.22	22.46
Luminance (*SD*)	73.72	12.49	67.22	13.52	73.17	13.57
RMS contrast	0.76	0.14	0.68	0.17	0.70	0.21
Skewness	0.57	0.48	0.54	0.50	0.39	0.58
Kurtosis	2.36	0.67	2.57	1.17	2.27	0.86
Energy ($\times 10^{-7}$)	7,226	2,930	8,394	3,079	6,767	1,669

Note. A different subscript indicates significant differences between scene categories. If two scores share an identical subscript, they are equivalent. RMS = root mean square.

center and a diameter of 1.5° presented at the center of the screen. The size of the target and the distracter pictures was 15.36° (width) × 9.32° (height) at a 60-cm viewing distance. The stimulus displays consisted of an emotional target (either pleasant or unpleasant) and a neutral distracter picture. The pictures were centered on the central horizontal axis, and the distance between the innermost edges of the picture areas from the center of the fixation circle was 2.5°. The pictures were followed by two white circles, which served as the saccade targets (see below for an explanation) and were presented at the center of the location where the target scenes had appeared (i.e., 10.18° away from the center of the screen aligned on the horizontal axis). To obtain a sufficient number of trials per condition (128) for the response latency distribution analysis, each stimulus scene was presented four times during the experiment.

Procedure. The participants were told that on each trial they were going to see two pictures (one emotional and one neutral) briefly appearing at the visual periphery. Their task was (depending on the block) to execute, as quickly as possible, a saccade to the side where an unpleasant or pleasant scene was flashed. Next, the eyetracker was calibrated. The calibration was accepted if the average error was less than 0.5°.

Each trial (see Figure 2 for a description of the sequence of trial events) began with a drift correction. A fixation circle appeared at the center of the screen, and the participants had to focus their gaze at the center of the circle. When the participant's eye was fixated on the circle, the experimenter initiated the trial. A random delay of 0–100 ms was appended at the beginning of all trials to prevent anticipatory saccades. After a 200 ms gap period (introduced to accelerate saccade initiation, see Fischer & Weber, 1993), the target scenes were displayed for 30 ms and were subsequently replaced with the saccade target circles. Participants were instructed to make a saccade as fast as possible to the circle that replaced the pleasant scene or the unpleasant scene. After an intertrial interval of 1,000 ms, the central fixation point reappeared, and the next trial was initiated.

Each participant performed one block with pleasant target scenes and one block with unpleasant target scenes, with the block order counterbalanced across participants. At the beginning of each block, the participants performed 10 practice trials representing the forthcoming experimental condition. Each block consisted of 128 trials in random order. The target scenes were equiprobable in both visual fields. Each target scene was presented four times (twice in each visual field), each time paired with a different distracter scene.

Eye movement analysis. Three different analysis strategies were used. In the first approach, we removed anticipatory responses (i.e., latencies < 0 ms) and undershoots (saccades with amplitudes below 2°) and then computed participantwise accuracy scores (%) and median reaction times (RT) for correctly directed saccades in the unpleasant and pleasant target conditions. These data thus represent the typical accuracy and time taken to encode the emotional scene content. In the second approach, we divided participantwise saccade latency distributions into 20-ms time bins and searched for the first bin containing significantly more (i.e., $p < .05$ in a paired samples t test) correctly directed saccades than erroneously directed saccades, followed by at least four more consecutive bins with more correctly directed saccades than erro-

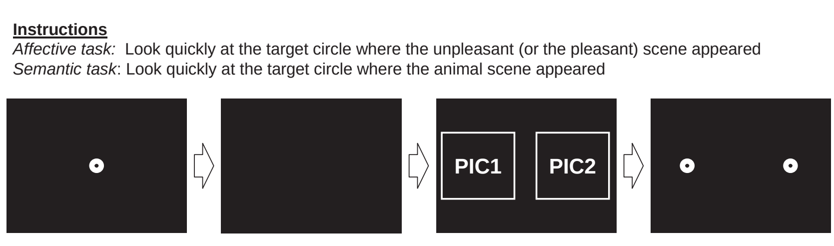

Figure 2. Trial sequence in Experiments 1–3. PIC = picture.

neously directed saccades (see Kirchner & Thorpe, 2006). These data thus represent the lower boundary of reliable classification speed (i.e., the earliest time at which identification of emotional valence occurred). In the third approach (F2 analysis), we computed stimuluswise median saccadic RTs and frequencies of correct responses and correlated these data with each other (i.e., RTs and accuracy), as well as with the low-level image features, to control for possible visual confounds.

Results

See Figure 3 for a summary of results. There were no differences in recognition accuracy for unpleasant (M = 72%) scenes versus pleasant (M = 74%) scenes, $t(23) = 1.4$, $p > .1$. One-sample t tests were used to examine whether the probability of correct responses exceeded the chance level (i.e., 50%). It was found that the hit rate exceeded the chance level for both unpleasant scenes and pleasant scenes, $ts(19) > 9.8$, $ps < .001$, thus confirming that the scene valence was reliably recognized. Median saccadic RTs indicated that pleasant scenes were recognized faster than were unpleasant scenes, $t(23) = 2.85$, $p = .01$, $\eta_p^2 = .26$ (means were 311 ms and 330 ms, respectively). The minimum latency, that is, the first 20-ms bin to contain more correct responses than errors, was observed to be at the 220–240-ms bin for both unpleasant, $t(23) = 3.01$, $p < .05$, and pleasant, $t(23) = 4.30$, $p < .001$, scenes. Itemwise hit rates and median RTs correlated negatively with each other for both unpleasant, $r(23) = -.37$, $p < .05$, and pleasant, $r(23) = -.74$, $p < .001$, scenes. However, neither RTs nor accuracies correlated with valence, arousal, or any of the low-level image features ($rs < .2$, $ps > .1$).

Discussion

Experiment 1 established that participants were able to consciously classify the pictures as pleasant or unpleasant in around 220–240 ms. If we assume that it takes on average 25 ms for the target-guided saccades to be programmed in the brainstem saccade generator (Schiller & Kendall, 2004), this leaves around 195–215 ms for the affective discrimination of the scenes. This conscious visual categorization speed thus seems surprisingly fast, given that the latency of reflexive saccades (i.e., those made automatically toward meaningless singleton targets) is rarely below 150 ms (Rayner, 1998). Although some of the extremely fast saccadic responses are probably anticipatory responses, the criterion used for determining the lower bound of encoding latency (i.e., first latency bin with more correct responses than erroneous responses followed by at least four bins with the same pattern) provides a reasonable safeguard against erroneous labeling of anticipatory albeit correct responses as recognition. It is also important to note that for the bins prior to the first identified bin, the proportion of erroneous and correctly directed saccades was essentially identical; moreover, after the proportion of correct responses begins to exceed that of errors, their proportion increases substantially. Hence, it seems reasonable to assume that the method used here is reliable for establishing the earliest time-point at which conscious recognition of affect starts.

The minimum latency of affective processing observed here is in line with estimates with single-neuron intracranial recordings of the amygdala (Oya et al., 2002) and the prefrontal cortex (Kawasaki et al., 2001). However, our lower bound latency estimates are lower than the EPN modulations (around 300 ms poststimulus) by highly arousing images (e.g., Cuthbert et al., 2000; Junghöfer, et al., 2001; Schupp et al., 2003; Schupp et al., 2004). Our findings suggest that the emotional valence of a scene is consciously accessible as early as around 200 ms poststimulus; thus, rudimentary affect recognition processes must be completed before 200 ms after stimulus onset. The median classification times (an average of 300 ms assuming a 20 ms delay in saccadic programming) are more in line with the EPN latency range,

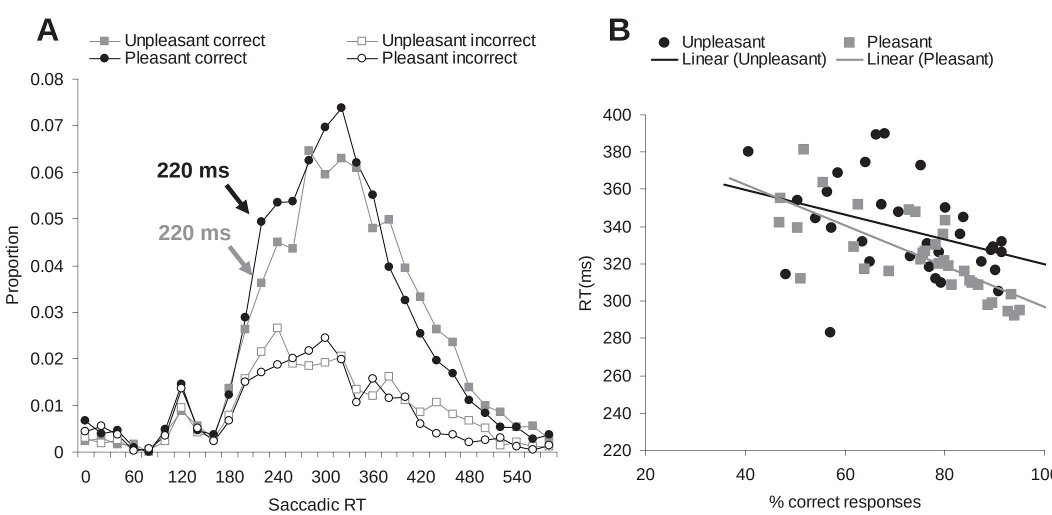

Figure 3. Results of Experiment 1 with unpleasant and pleasant human targets (and neutral human distracters). A: Saccadic latency distributions. The x-axis shows saccadic reaction times (across 20-ms bins), and the y-axis shows the proportion of trials with correct and incorrect saccadic responses for pleasant and unpleasant scenes. Arrows indicate the earliest point in time when the proportion of correct saccades toward the target significantly exceeded that of erroneous saccades toward the distracter. B: Itemwise median reaction times plotted against the proportion of correct responses, with a regression line, separately for the unpleasant and the pleasant targets. RT = reaction time.

although it must be noted that EPN effects are typically found for arousal rather than for valence (see Olofsson et al., 2008). Our findings thus add to the EPN literature by showing that conscious recognition and discrimination of emotional valence also occurs within the 200–300-ms EPN range.

The present data also demonstrate that very limited visual input is sufficient for affective recognition, as reliable classification was observed when a pair of scenes was flashed only for 30 ms. Shifts of covert attention are estimated to take around 25–50 ms per item (Itti & Koch, 2001); hence, it is plausible to assume that on most trials the participants did not have time to attend either of the scenes. Instead, they were able to accomplish categorization by relying on visual information acquired simultaneously from two unattended sources. This fits with previous findings (Glascher & Adolphs, 2003; Vuilleumier, Armony, Driver, & Dolan, 2001) that showed that affective processing requires only minimal attentional resources (but see Pessoa, 2005).

Although researchers in some studies have argued for a detection advantage of unpleasant over pleasant objects in emotional pictures (Fox et al., 2000; Öhman, Lundqvist, et al., 2001), in the present study, we found no clear evidence for such an advantage. Although lower bound latency for correct classification was similar for both unpleasant and pleasant scenes, the median RTs (reflecting the typical time for encoding affect from the scenes) were faster for pleasant than for unpleasant scenes. This latter finding is consistent with prior studies showing preferential processing of positively valenced information rather than negatively valenced information (see review in N. A. Murphy & Isaacowitz, 2008).

In the context of face recognition (Calvo & Nummenmaa, 2008), we have argued that the facilitated detection of pleasant facial affect (happy faces) is based on effective detection of single salient visual features (such as the smile), which is used as a shortcut for encoding the affective content. We thus turned to F2 analyses to examine whether specific scene features could be responsible for facilitated detection of pleasant scenes as well. No speed–accuracy tradeoffs were observed in the F2 analysis. Quite the opposite, the stimuli that were classified most reliably were also classified fastest. Nevertheless, the F2 plot shows that there is great variability in the speed with which different scenes can be classified as pleasant or unpleasant. Whereas for some scenes the average classification speed is around 280 ms, for others it is close to 400 ms. Eyeballing the contents of the fastest and slowest scenes suggests that the fastest pleasant scenes univocally contained either erotica or romantic couples, whereas the slowest pleasant scenes contained biologically less salient events such as recreational activities (e.g., sports, amusement, etc.). Accordingly, the color of (naked) human bodies (Jiang, Costello, Fang, Huang, & He, 2006) or the smile of the happy faces (Calvo & Nummenmaa, 2008) could be used as shortcuts for detecting pleasant affect. For unpleasant scenes, the contents of the fast and slow scenes were not clearly distinctive in content, and the lack of clear distinctive diagnostic information may have rendered their processing slower than that of pleasant scenes.

Experiment 2: Contrasting Affective and Semantic Categorization Speed

Experiment 1 established the minimum and typical speed of visual affective processing and confirmed that the eye-movement methodology is well suited for investigating this issue. We next moved to directly comparing affective and semantic processing. To this end, we conducted Experiments 2–5 in which we contrasted the speed of affective evaluation of pleasantness or unpleasantness with that of semantic recognition of animals versus humans. We had two reasons for selecting animals as the semantic targets. First, it allows for the comparison of superordinate-level semantic categorization (animal or human) of a natural category with superordinate-level affective categorization (unpleasant or pleasant). Second, it allows us to compare affective and semantic recognition latency estimates directly with those obtained by Kirchner and Thorpe (2006), who used animal targets in a similar saccadic response paradigm involving only semantic recognition. In Experiment 2, we compared affective classification speed of the scenes used in Experiment 1 with semantic recognition of scenes involving neutral animals. If affective processing precedes semantic recognition, affective scene recognition should enjoy a significant advantage over semantic recognition.

Method

Participants. Forty undergraduate psychology students (31 female, 9 male; with a mean age of 19.7 years) from the University of La Laguna participated for course credit.

Procedure and design. The basic design, stimulus layout, and timing were similar to those in Experiment 1. On each trial, a pair of images were flashed for 30 ms, followed by two target circles replacing the images, and participants were asked to saccade as quickly as possible to the circle replacing the prespecified target. In Experiment 2, participants performed two different tasks. In the affective task they were asked to saccade toward an emotional (either pleasant or unpleasant, depending on the counterbalancing condition) scene, whereas in the semantic task they were asked to saccade toward a scene containing an animal. Neutral scenes of human beings served as distracter scenes.

The 32 unpleasant and 32 pleasant target scenes and the 64 neutral distracter scenes were the same as in Experiment 1. Additionally, a set of 32 neutral images depicting nondangerous animals (e.g., elks, birds, hedgehogs) in various natural environments were added as semantic targets. The images were retrieved from various sources (most of them were freely available in the Internet), and their basic image statistics (see Table 2 for mean scores) were computed similarly to those of Experiment 1. Comparison of the neutral animal scenes with the neutral and affective scenes involving humans (mean scores presented in Table 1) showed no significant differences in the low-level visual features. One-way ANOVAs with image category (unpleasant with humans vs. neutral with humans vs. pleasant with humans vs. neutral animals) revealed a significant difference only in the standard deviation of luminance, $F(3, 159) = 2.91$, $p = .04$, $\eta_p^2 = .05$, and energy, $F(3, 159) = 4.45$, $p = .005$, $\eta_p^2 = .08$, but none of the planned contrasts proved significant after correcting for multiple comparisons.

For the experimental stimulus displays, the animal scenes as well as the pleasant scenes and the unpleasant scenes were paired with the neutral scenes involving humans. The experimental conditions were combined into a mixed factorial design, with task (affective vs. semantic recognition) as a within-participants factor and valence of the affective targets (unpleasant vs. pleasant) as a between-participants factor. Twenty participants performed the affective, pleasant-stimulus task and the semantic task, and another

Table 2
Means and Standard Deviations of Stimulus Characteristics of the Unpleasant, Neutral, and Pleasant Animal Scenes Used in Experiments 2–5

Characteristic	Unpleasant animals *M*	Unpleasant animals *SD*	Neutral animals *M*	Neutral animals *SD*	Pleasant animals *M*	Pleasant animals *SD*
Luminance (average)	95.05	35.76	104.43	28.45	108.85	33.55
Luminance (*SD*)	62.10	11.18	71.33	10.41	63.79	11.92
RMS contrast	0.77	0.38	0.73	0.25	0.66	0.30
Skewness	0.64	0.77	0.46	0.56	0.31	0.64
Kurtosis	3.18	2.44	2.41	1.16	2.63	0.97
Energy ($\times 10^{-7}$)	4,997.61	1,491.53	9,144.63	3,603.24	4,320.59	1,614.93

Note. RMS = root mean square.

twenty performed the affective, unpleasant-stimulus task and the semantic task. Each participant saw 128 trials with semantic targets in one block and 128 trials with affective targets in another block, with block order counterbalanced between participants. Each target scene was repeated four times. Data were analyzed and processed similarly to those of Experiment 1.

Results

See Figure 4 for a summary of results. First, one-sample *t* tests were used to verify that the probability of correct responses in all the tasks (detect unpleasant/detect pleasant/detect animal) exceeded the chance level (i.e., 50%). This was confirmed to be the case in all conditions, $ts(19) > 15.34$, $ps < .001$, implying that both affective valence and semantic category membership were reliably recognized from the scenes. Next, accuracy scores and median saccadic RTs were subjected to 2 (categorization type: semantic vs. affective) $\times$ 2 (valence of affective scenes: unpleasant vs. pleasant) mixed ANOVAs, with valence of affective scenes as a between-participants factor. For accuracy scores, a main effect of categorization type emerged, $F(1, 38) = 300.00$, $p < .001$, $\eta_p^2 = .89$, with higher accuracy in the semantic task than in the affective categorization task (92% vs. 74%, respectively), as well as an interaction between categorization type and valence of affective scenes, $F(1, 38) = 9.94$, $p = .003$, $\eta_p^2 = .21$. The interaction resulted from semantic categorization having a larger advantage over classification of unpleasant (a difference of 21%) scenes than over classification of pleasant (a difference of 14%) scenes.

For saccadic median RTs, the analysis yielded a main effect of categorization type, $F(1, 38) = 62.78$, $p < .001$, $\eta_p^2 = .62$, with faster responses in the semantic task than in the affective task (274 ms vs. 330 ms, respectively). The interaction between categorization type and valence of affective scenes was also significant, $F(1, 38) = 9.49$, $p = .005$, $\eta_p^2 = .20$, resulting from semantic categorization enjoying a larger advantage over categorization of unpleasant (78 ms) scenes than over that of pleasant (34 ms) scenes. The minimum latency for the accurate semantic classification of scenes was 180–200 ms, $t(39) = 3.44$, $p < .001$; it was 220–240 ms for unpleasant scenes, $t(19) = 3.01$, $p = .03$, and 200–220 ms for pleasant scenes, $t(19) = 3.10$, $p = .025$. Itemwise hit rates and median RTs were negatively correlated in the semantic categorization task, $r(31) = -.55$, $p = .01$, and in the affective categorization of pleasant scenes, $r(31) = -.76$, $p < .001$, but not in that of unpleasant scenes, $r(31) = -.04$.

Discussion

Experiment 2 successfully replicated the lower bound processing speed of affective categorization. Experiment 2 also established that semantic categorization occurs at least 30 ms earlier than does affective categorization. The minimum latency for semantic classification was 180–200 ms, whereas the respective value for affective classification (averaged over unpleasant and pleasant) was 210–230 ms. The lower bound latency of semantic recognition was slightly longer than that (150 ms) reported by Kirchner and Thorpe (2006), who used a similar paradigm, but can probably be accounted for by differences in the selection of target scenes, as well as display parameters such as target scene size, distance from the center of the screen and so forth. The recognition advantage for semantic scenes became even more pronounced (58 ms) when median latencies were considered, for which a semantic recognition primacy over affective identification was found for both pleasant and unpleasant scenes.

Given that pictorial affective scenes attract attention in a reflexive, involuntary manner when presented together with visually matched, neutral distracter scenes (Calvo & Lang, 2004; Nummenmaa, Hyönä, & Calvo, 2006), it seems striking that this attentional bias was not reflected in the goal-directed saccadic RTs in the current paradigm (i.e., that saccades to affective targets were actually slower than those to semantic targets). A possible explanation for this is that semantic analysis indeed precedes affective analysis of the scenes and that the output of semantic recognition may be passed on to the eye-movement control system before the attentional bias due to emotional content occurs. All in all, the current data support the temporal primacy of semantic over affective processing (Calvo & Nummenmaa, 2007; Lazarus, 1984; Rolls, 1999; Storbeck & Clore, 2007; Storbeck et al., 2006).

Given that accuracy scores were higher in the semantic task than in the affective task, it could be argued that maybe the greater difficulty of the affective discrimination task (due to a subset of scenes being very difficult to classify) resulted in slower median latencies. However, inspection of the F2 data (Figures 4C–4D) suggests that this was not the case. Most of the semantic targets were indeed classified rapidly, whereas the same holds true for only a small subset of affective target scenes. Additionally, the advantage for semantic over affective classification was clear for almost all targets. Of unpleasant scenes, only 6% could be categorized faster than the slowest semantic scene (overlap between RT distributions being 12%), whereas the corresponding overlap

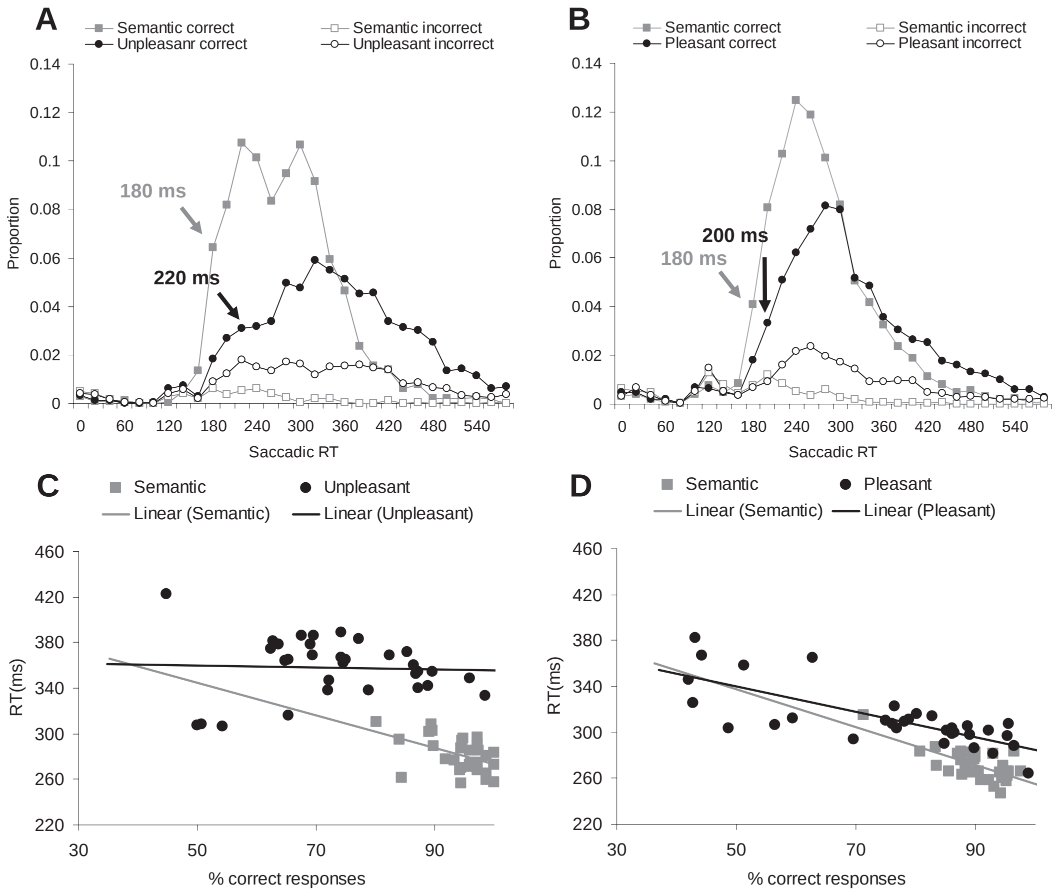

Figure 4. Results from Experiment 2 with unpleasant and pleasant human targets and neutral animal targets. A and B: Saccadic latency distributions. The *x*-axis shows saccadic reaction times (across 20-ms bins), and the *y*-axis shows the proportion of trials with correct and incorrect saccadic responses in the semantic and the affective classification task. Arrows indicate the earliest point in time when the proportion of correct saccades toward the target significantly exceeded that of erroneous saccades toward the distracter. C and D: Itemwise median reaction times plotted against the proportion of correct responses, with regression line, separately for the affective and the semantic classification task. RT = reaction time.

for the pleasant versus semantic scenes was 15%. It thus seems clear that semantic recognition holds a genuine and generalizable latency advantage over affective recognition, and the effect is not due to a small subset of semantic targets.

Experiment 3: Recognizing Semantic Category and Affective Valence of Dangerous Animals

Experiment 2 suggests that semantic categorization is faster than affective categorization. However, this finding is not fully conclusive because different images were used as affective and semantic targets. Although we matched the affective and semantic targets with respect to low-level visual features, it is possible that the emotional scenes were more complex, for example, in terms of the number of objects present in the scene or the composition of scene elements. Although detecting an animal among bushes only requires recognition of one scene element, encoding the affective valence of a scene that depicts a man strangling a girl requires (at least) an assessment of the nature of the depicted event and the roles of the depicted persons, which could result in slower processing.

To control for such confounds, we conducted Experiment 3, in which exactly the same scenes were used as targets in the semantic and the affective classification trials. Accordingly, any potential speed or accuracy differences between the tasks could not be explained by stimulus factors. Predatory animals and snakes are considered canonical examples of stimuli to which the affect system or the threat module (Ohman & Mineka, 2001) rapidly reacts. Accordingly, we chose a sample of 32 scenes with dangerous animals (snakes, wolves, alligators, etc.) from the IAPS and other sources. These scenes were paired with neutral distracter scenes involving humans. We could thus use exactly the same target—distracter pairs in both the affective (detect an unpleasant scene) and the semantic classification (detect an animal scene) tasks. This enabled us to compare the affective and semantic processing speed of the same visual input: If the semantic recognition advantage observed in Experiment 2 is genuine, participants should be quicker in classifying the picture of a snake as an animal rather than as unpleasant. On the contrary, if the semantic processing advantage in Experiment 2 was due to visual confounds (e.g., scene complexity) not captured by the available image statistics, no such processing advantage should occur in Experiment 3.

Additionally, this approach enabled us to correlate the itemwise RTs in the affective and semantic classification tasks in order to assess the temporal relationship between affective and semantic recognition. If recognition of the semantic category of the animal targets is necessary for their affective evaluation, RTs in the affective and semantic classification tasks should be positively correlated, as affective processing time would sum up with the semantic processing time. However, if affective and semantic processing is undertaken by independent neurocognitive systems, no correlation should be observed.

Method

Twenty undergraduate psychology students (16 female; 4 male; with a mean age 21.1 years) from the University of Turku (Turku, Finland) took part in the experiment for course credit. The basic design was identical to that in Experiment 2, with the following exceptions. Only unpleasant affective targets were used. In both affective and semantic categorization tasks, the target scenes were 32 unpleasant animal scenes that were paired with neutral scenes involving humans. Additionally, 50% of the trials were fillers in both tasks. In the affective categorization task, the filler trials were 32 unpleasant scenes depicting people paired with neutral scenes depicting humans. In the semantic categorization task, these were 32 scenes involving neutral animals paired with neutral humans. By using the filler items, we ensured that the affective and semantic recognition tasks were orthogonal: Semantic category of the target could not be used as a recognition cue in the affective task (because half of the targets were humans and half were animals), and affective valence could not be used as a cue in the semantic task (because half of the targets were unpleasant and half were neutral). All participants performed the affective task (unpleasant targets) and the semantic recognition task (animal targets) as two separate blocks, with the order of blocks counterbalanced across participants. Each block had 64 target trials and 64 filler trials and began with 10 practice trials representing the forthcoming experimental condition.

Results

The results are summarized in Figure 5. Again, one-sample t tests were used to verify that the probability of correct response in both tasks (detect unpleasant/detect animal) exceeded the chance level. The difference between the observed hit rate and chance level was significant for both tasks, $ts(19) > 7.02$, $ps < .001$, showing that participants were accurate in identifying category membership as well as affective valence. Next, accuracy scores and median saccadic RTs were analyzed with paired samples t tests. These revealed that accuracy was higher (94% vs. 71%), $t(19) = 8.39$, $p < .001$, and RTs were faster (311 ms vs. 375 ms), $t(19) = 5.30$, $p < .001$, in the semantic task than in the affective task. The analysis of the minimum saccade latency replicated the finding of Experiment 2. Faster lower bound classification was observed in the semantic (180–200 ms), $t(19) = 2.56$, $p < .05$, task than in the affective (200–220 ms), $t(19) = 2.36$, $p < .05$, task. F2-analysis revealed that hit rates and latencies were negatively correlated in the semantic task, $r(31) = -.34$, $p < .05$, but not in the affective task. It is important to note that semantic categorization speed was faster than or equally as fast as affective categorization speed for each item (i.e., target scene), with differences ($M_{\text{semantic}} - M_{\text{affective}}$) ranging from 0 to 130 ms. Further, there was a significant positive relationship between RTs of correct responses in the semantic task and in the affective task, $r(31) = .43$, $p = .01$ (see Figure 6).

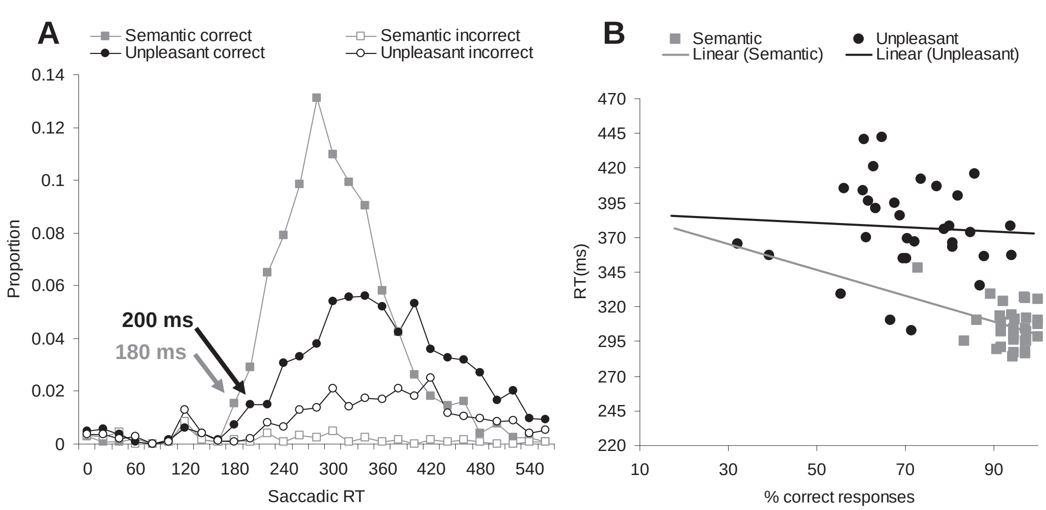

Figure 5. Results from Experiment 3 with unpleasant animal targets. A: Saccadic latency distributions. The x-axis shows saccadic reaction times (20-ms bins), and the y-axis shows the proportion of trials separately for correct and incorrect responses for the semantic and the affective classification task. Arrows indicate the earliest point in time when the proportion of correct saccades toward the target significantly exceeded that of erroneous saccades toward the distracter. B: Itemwise median reaction times plotted against the proportion of correct responses with regression line, separately for the affective and semantic classification task. RT = reaction time.

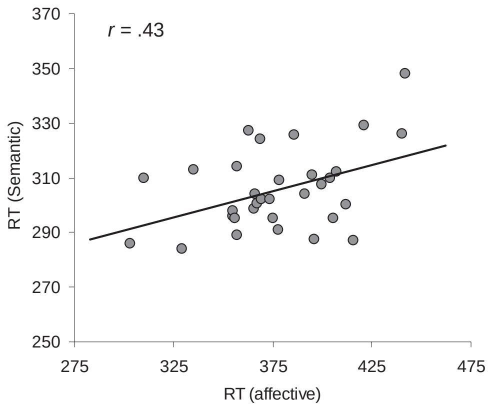

Figure 6. Association of itemwise response latencies in the affective and semantic recognition tasks in Experiment 3. RT = reaction time.

Discussion

Experiment 3 provides direct evidence for the priority of semantic over affective processing. Three findings support this view. First, when participants classified exactly the same images depicting phylogenetically significant, threatening animals with respect to either the affective (unpleasant or not) dimension or the semantic (animal or not) dimension, median RTs were 64 ms faster in the semantic task than in the affective task, and the lower bound latency for correct discrimination was 20 ms faster for the semantic task. These results unambiguously demonstrate that semantic recognition is accomplished earlier than affective evaluation. Second, the F2-analysis revealed that affective classification never preceded semantic analysis, as semantic recognition speed exceeded or equaled that of affective recognition for all items. In other words, not a single target scene could be classified faster according to its affective dimension than according to its semantic dimension. This suggests that affective processing depends on semantic recognition, in that affective analysis cannot bypass object recognition. Third, itemwise median RTs showed a significant positive correlation between the semantic tasks and the affective tasks. Together with the finding of faster RTs for semantic categorization than for affective categorization, this finding suggests that semantic recognition occurs prior to and is necessary for affective analysis. It is important to note that the positive association between affective and semantic categorization RTs shown in Figure 6 reveals that semantic and affective processing of images is additive in nature, thus supporting serial processing of semantic and affective information (see the General Discussion). In sum, Experiment 3 provides evidence for the view that semantic recognition of objects or their parts in the scene is necessary for subsequent affective evaluation (Calvo & Nummenmaa, 2007; Storbeck et al., 2006).

Experiment 4: Contrasting Foveal and Extrafoveal Semantic and Affective Categorization

In Experiments 1–3, extrafoveal presentation of target scenes was used. It is known that stimulus recognition becomes progressively worse the further away from the fovea the stimulus appears (e.g., Thorpe, Gegenfurtner, Fabre-Thorpe, & Bülthoff, 2001). Yet, it is possible that extrafoveally classifying a snake as an animal, versus a nonanimal, could be accomplished with more coarse-grained visual information (i.e., so that it can be processed by the magnocellular outputs originating from the peripheral retina) than categorizing a snake as harmful, versus harmless. Hence, the peripheral presentation of scenes could inherently favor semantic discrimination, as evidenced by the higher hit rates in all the semantic tasks, whereas an advantage for affective discrimination may perhaps occur when scenes are perceived foveally.

To test this prediction, we conducted Experiment 4, in which we compared the speed of affective and semantic categorization under foveal and extrafoveal (2.5 degrees away from fixation, as in Experiments 1–3) presentation conditions. As only a single picture could be presented foveally, the extrafoveal condition included a target picture paired with a meaningless image made up from random colors, to mimic the paired stimulus presentation of Experiments 1–3. Furthermore, as the saccadic response paradigm cannot be applied to the foveal condition, we resorted to a manual 2AFC task in which participants were asked to categorize the stimuli as affective (unpleasant vs. not unpleasant or pleasant vs. not pleasant), as containing animals or as not containing animals. If the results of Experiments 2–3 are due to the extrafoveal presentation favoring semantic recognition, no differences in hit rates and median RTs should be observed under the foveal condition of Experiment 4. However, if there is a genuine advantage of semantic over affective recognition, the RT difference should remain even when the scenes are seen foveally.

Experiment 4 also had a secondary aim. Namely, Experiments 2–3 demonstrating a semantic processing advantage used a simultaneous presentation of target scenes and distracters. Although parallel processing of semantic features of scenes has been established (Rousselet et al., 2002), parallel processing of affective scene information might not be possible (although, see Haberman & Whitney, 2007, for recent evidence supporting parallel processing of facial affect), which could explain the semantic processing advantage observed in Experiments 2–3. If so, the semantic processing advantage should be abolished in Experiment 4 when the pictures are presented singly, either foveally or extrafoveally.

Method

Participants. Forty undergraduate psychology students (30 female; 10 male; with a mean age of 19.7 years) from the University of La Laguna participated in the experiment for course credit.

Stimuli and apparatus. The stimuli were 128 target pictures, of which 64 portrayed people (32 unpleasant, 32 pleasant) and another 64 portrayed animals (32 unpleasant, 32 pleasant). The unpleasant and pleasant pictures portraying people were those used in Experiment 2, and those portraying unpleasant animals were used in Experiment 3. The additional 32 scenes with pleasant animals (pups, kittens, etc.) were selected from the IAPS and other sources. All the images were presented both in the affective and semantic classification task. Participants had their heads positioned on a chin and forehead rest. The pictures subtended a visual angle of 13.3° by 11.1° at a constant viewing distance of 60 cm. The stimuli were presented against a black background on a 17 in. (43.18 cm) super video graphics array (SVGA) monitor (with a 100-Hz refresh rate) connected to a Pentium IV 2.8-GHz computer. The E-Prime 2.0 experimental software controlled stimulus presentation and response collection.

Procedure. The participants were told that they would be presented with photographs of people or animals, which could be either unpleasant or not unpleasant. The participant's task was to attempt to identify the affective valence of the scene in one block and to detect the presence of an animal in the other block. For both tasks, the participants were to respond with a key press of the letter *D* or the letter *L* (labeled as *YES* or *NO*) as soon as possible, using the right and left index finger. Figure 7 shows the sequence of events during a trial. A trial started with a central fixation cross displayed for 500–1,000 ms. This was followed by a 200 ms gap period and a target picture for 30 ms. In the extrafoveal condition, the target picture appeared either to the left or to the right of the fixation cross, and a meaningless picture (a random combination of colors; same size and luminance as the target picture) appeared on the opposite side. The distance from the central cross to the inner edge of the lateralized target scene (and the scrambled image) was 2.5°. In the foveal condition, the target picture appeared at fixation, in the center of the screen (with no simultaneous meaningless picture). When the scene (and the meaningless picture) disappeared, it was replaced by a question mark that served as a prompt to respond as to whether the target scene was unpleasant (Block 1 or Block 2) or whether it portrayed an animal (Block 2 or Block 1). The intertrial interval was 1,500 ms. Response accuracy and RTs were collected. In each block, there were 16 practice trials followed by 128 experimental trials.

Design. The experimental conditions were combined in a mixed factorial design, with evaluation task (affective valence or semantic category) as a within-participant factor and target picture location (foveal vs. extrafoveal) as a between-participants factor, with 20 participants in each location condition. The order of blocks was counterbalanced, such that half of the participants received the valence evaluation task first and the other half received the semantic categorization task first. Within each block, each of the 128 target pictures was presented once to each participant in a random order.

Results

One-sample *t* tests were first computed to examine whether the proportion of correct responses in both the valence evaluation and the semantic task exceeded the chance level (i.e., .50). For all combinations of affective valence and location as well as animal category and location, the difference between the observed hit rate and the chance level was highly significant, $ts(19) > 8.0$, $p < .001$. This implies that both valence and semantic category membership were reliably recognized. In other words, unpleasant scenes were generally perceived as unpleasant, animal scenes were perceived as animals, pleasant scenes were correctly perceived as not being unpleasant, and people were perceived as not being animals.

The proportions of correct responses and the corresponding response latencies were analyzed with a 2 (task) × 2 (location) ANOVA. The mean accuracy and latency scores are shown in Figure 8. For response accuracy, the ANOVA yielded a main effect of task, $F(1, 38) = 61.80$, $p < .001$, $\eta_p^2 = .62$, and location, $F(1, 38) = 21.76$, $p < .001$, $\eta_p^2 = .36$. Semantic recognition was performed more accurately than was affective evaluation ($M = .95$ vs. .82), and the hit rate was higher for foveal than for extrafoveal scenes ($M = .92$ vs. .85). A Task × Location interaction, $F(1, 38) = 7.64$, $p < .01$, $\eta_p^2 = .17$, revealed that the impairment due to extrafoveal—relative to foveal—stimulus presentation was significant for the valence evaluation task, $t(38) = 4.06$, $p < .001$, but not for the semantic recognition task, $t(38) = 1.90$, $p = .07$, and that the advantage for semantic categorization over affective evaluation was greater in the extrafoveal, $t(19) = 6.43$, $p < .001$, than in the foveal condition, $t(19) = 4.52$, $p < .001$, although the advantage was significant for both presentation conditions.

RTs of correct responses were affected by task, $F(1, 38) = 54.79$, $p < .001$, $\eta_p^2 = .590$, and location, $F(1, 38) = 14.59$, $p < .001$, $\eta_p^2 = .28$. Response latencies were shorter in the semantic task than in the affective task ($M = 380$ ms vs. $M = 502$ ms), and they were shorter in the foveal condition than in the extrafoveal condition ($M = 401$ ms vs. $M = 481$ ms). A Task × Location interaction, $F(1, 38) = 6.97$, $p < .025$, $\eta_p^2 = .16$, revealed that the impairment due to peripheral stimulus presentation was significant for the valence evaluation task, $t(38) = 4.70$, $p < .001$, but not for the semantic evaluation task, $t(38) = 1.37$, $p = .18$, and that the advantage for semantic recognition over affective evaluation was greater in the extrafoveal condition, $t(19) = 6.24$, $p < .001$, than in the foveal condition, $t(19) = 4.01$, $p < .001$, although the advantage was significant for both presentation conditions.

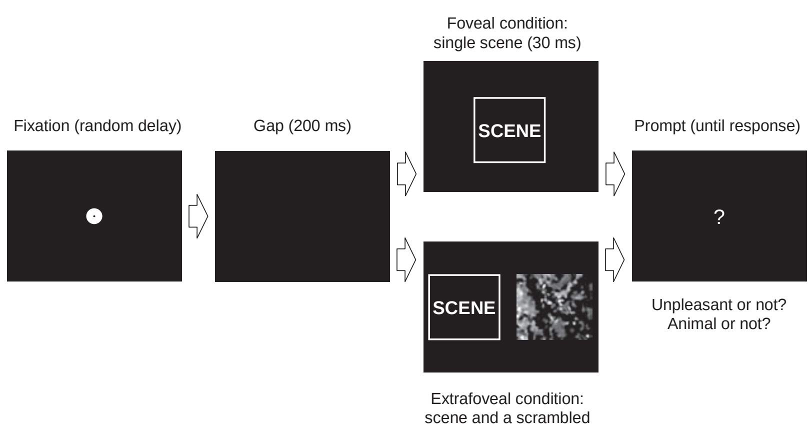

Figure 7. Trial sequence in Experiment 4.

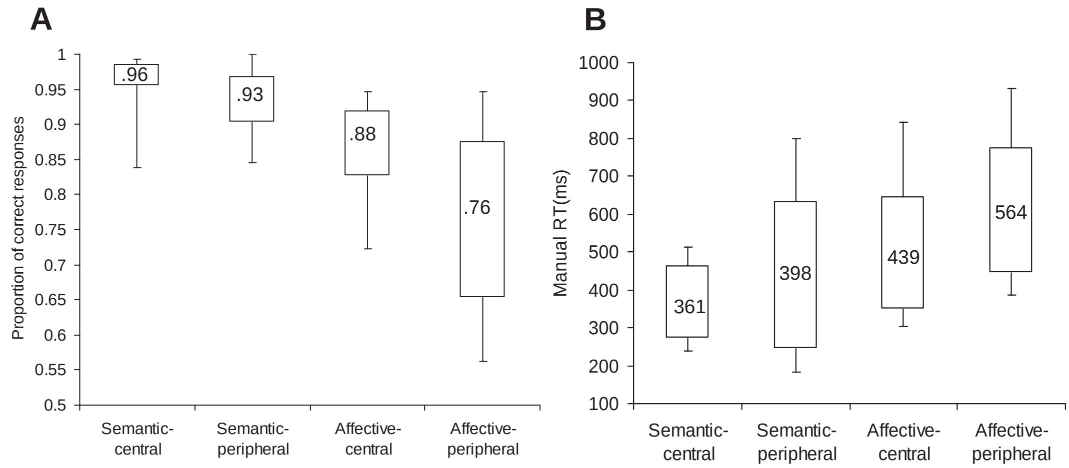

Figure 8. Distributions of accuracy scores (A) and median reaction times (B), as a function of categorization task (affective vs. semantic) and eccentricity (foveal vs. extrafoveal) in Experiment 4. Bars and box represent 95% of the observations. RT = reaction time.

Discussion

Experiment 4 confirmed that semantic categorization speed is superior to affective categorization speed both in foveal and extrafoveal vision. Although manual responses were over 100 ms slower than the respective eye-movement responses in the previous experiments, the pattern of results remained essentially the same. These findings have three important implications. First, as the semantic recognition advantage was observed under conditions involving only a single target, these data confirm that the results of Experiments 1–3 were not confounded by the simultaneous presentation of target images. Hence, the results of Experiments 1–3 may be taken as evidence for parallel affective processing of multiple scenes, similarly to parallel processing of semantic content of scenes (Rousselet et al., 2002). Second, the data of Experiment 4 strengthen the conclusions of the previous experiments by demonstrating a semantic classification advantage in another response modality (manual responses). This is particularly important, as manual responses to foveally and singly presented scenes are probably less influenced by variations in low-level visual features of the target scenes than are saccadic responses when two extrafoveal scenes are presented simultaneously.

Third, and most important, Experiment 4 indicates that the semantic superiority effect observed in Experiments 2–3 is not an artifact of presenting the scenes in peripheral vision because the advantage for semantic classification held even under foveal presentation conditions. As revealed by the Task $\times$ Location interaction, affective evaluation was more impaired than semantic categorization by the extrafoveal presentation of scenes. Hence, the extrafoveal stimulus presentation used in Experiments 1–3 may have made the semantic processing advantage more pronounced, but the results from the foveal condition of Experiment 4 show that the semantic categorization advantage (a difference of 78 ms) holds even when scenes are processed with the high-resolution foveal vision.

The significant impairment of affective processing in extrafoveal vision versus foveal vision is somewhat at odds with the claims that emerged from studies with spatial frequency filtering manipulations (mimicking extrafoveal stimulus presentation) of affective images. ERP studies have suggested that the magnocellular pathway projecting particularly from the extrafoveal–peripheral (but also from the foveal) retina might be more crucially involved in the affective evaluation of facial expressions (Pourtois, Dan, Grandjean, Sander, & Vuilleumier, 2005) and complex emotional scenes (Carretié, et al., 2007) than is the parvocellular pathway projecting from the fovea. Although our data do not undermine the role of the magnocellular pathway in extracting affective valence, they nevertheless suggest (a) that this pathway may be even more efficient in semantic recognition and (b) that the parvocellular pathway is more important for affective than for semantic recognition. Nevertheless, it must be noted that in addition to spatial frequency, processing of many other visual characteristics (e.g., color, motion, contrast and so forth) also change as a function of the eccentricity of the stimulus from the fovea; hence, the observed pattern of results is not likely to be only related to spatial frequency.

Finally, in Experiment 4 we controlled for one potential confounding factor that was present in the affective categorization task in Experiments 1–3: As those experiments involved the presentation of paired emotional target and nonemotional distracter scenes, it could be argued that the participants could perform the task by merely relying on the arousal (rather than valence) dimension. But, as Experiment 4 yielded an essentially identical pattern of response latencies with a task necessitating valence recognition, it can be assumed that also the results of Experiments 1–3 reflect valence processing.

Experiment 5: Comparing the Exposure Threshold for Affective and Semantic Recognition

Despite the evidence reported above, the primacy of affect over cognition may still be defended by arguing that the visual threshold is lower for affective information than for semantic information. A bulk of psychophysiological studies has indeed shown that unpleasant emotional information presented subliminally (typically for 20–30 ms and backwardly masked) can evoke affect-dependent electrodermal (Esteves, Dimbert, & Ohman, 1994; Glascher & Adolphs, 2003) and electromyographic responses (Dimberg et al., 2000), as well as hemodynamic changes in the brain regions involved in emotional processing (Morris et al., 1999). This evidence suggests that there is rapid processing of emotional information outside of awareness, which is consistent with the affective primacy hypothesis. Nevertheless, this important issue warrants further investigation: Although in these experiments it was verified that affective processing occurred in the absence of conscious perception, it was not tested whether semantic processing would have occurred as well. Indeed, if semantic processing could also be accomplished, this would cast doubt on the privileged access of emotion to awareness. This is a relevant concern, given that a recent study (Grill-Spector & Kanwisher, 2005) has established that semantic recognition of backwardly masked scenes presented for 20 ms (which is also the typical display duration in 'subliminal' affect processing studies) is indeed possible.

There is an additional concern regarding the paradigms we have used so far. Namely, it is possible that in our previous experiments affective processing could have been accomplished earlier than semantic recognition, but consciously accessing the affective representation would simply have taken longer, which would then be reflected in longer response latencies. To estimate the visual processing speed of affective and semantic recognition without resorting to speeded categorization, we asked participants to perform a semantic categorization (animal/human) and an affective evaluation (unpleasant/pleasant) task for foveally presented scenes under conditions in which the time available for visual processing was varied. The scenes were displayed for 20 ms, 40 ms, and 80 ms, followed by a mask. As masking interrupts visual processing (Breitmeyer & Ogmen, 2000), performance should be better for a given stimulus duration in the task (affective or. semantic) that can be accomplished earlier. Accordingly, if affective content has privileged access to awareness and is visually processed earlier than semantic category, performance should be more accurate in the affective task than in the semantic recognition task for the shortest (20 ms) exposure duration, with differences decreasing at longer exposure durations. On the contrary, if semantic information related to scene objects can (and must) be accessed prior to affective content, an advantage is predicted for the semantic recognition task over the affective evaluation task at the shortest stimulus displays.

Method

Participants. Twelve undergraduate students (9 female; 3 male; with a mean age of 22 years) from the University of Turku participated in the experiment for course credit.

Stimuli and apparatus. The stimuli were the 64 pleasant scenes (32 involving animals, 32 involving humans) and 64 unpleasant scenes (32 involving animals, 32 involving humans) used in Experiment 4. They were presented against a black background on a 20 in. (50.8 cm) SVGA monitor (with a 100-Hz refresh rate) connected to a Pentium IV 2.8-GHz computer. The E-Prime experimental software controlled stimulus presentation and response collection.

Procedure. The participants were told that they would be presented with photographs of people and animals, which could be either unpleasant or pleasant. The participant's task was to attempt to identify the affective valence (pleasant/unpleasant) of the scene in one block and the presence of animals versus humans in the other block. For both tasks, the participants were to respond with a key press of the letter *Z* or the letter *M* as soon as possible, using the right and left index finger. Figure 9 shows the sequence of events during a trial. Each trial started with a short (50 ms) beep played to the earphones. Next, the target picture was displayed for 20 ms, 40 ms, or 80 ms, followed by a mask (random combination of colors, generated randomly for each trial) displayed for 1,250 ms. The mask also served as a prompt to respond as to whether the target scene was unpleasant or pleasant (Block 1 or Block 2) or whether it portrayed an animal or a human being (Block 2 or Block 1). The mask remained on the screen until the next target scene was presented; a 50 ms auditory signal was played 200–250 ms (randomly selected) prior to the target scene. Response accuracy and RTs were collected. In each block, there were 16 practice trials followed by 384 experimental trials. The participants were encouraged to respond to each photograph even if they felt that they had not perceived the scene at all (in which case they were asked to guess).

Design. The experiment involved a 2 (evaluation task: affective valence or semantic category) $\times$ 3 (exposure duration: 20, 40, or 80 ms) within-participants design. The order of blocks was counterbalanced, such that half of the participants received the valence evaluation task first and the other half received the semantic categorization task first. Within each block and display duration, each of the 128 target pictures was presented once to each participant in a random order; thus, the total number of trials was 768.

Results

Participantwise median RTs and response accuracy were computed for all the cells in the design. One-sample *t* tests revealed that the classification accuracy exceeded the chance level in all experimental conditions, $ts(11) > 2.89$, $ps < .01$, except for affective classification in the 20-ms display condition. Next, accuracy scores and RTs were analyzed with 2 (evaluation task) $\times$ 3 (exposure duration) repeated-measures ANOVAs. The results are summarized in Figure 10. For accuracy, the ANOVA yielded a main effect of task, $F(1, 11) = 4.82$, $p < .05$, $\eta_p^2 = .31$, with more

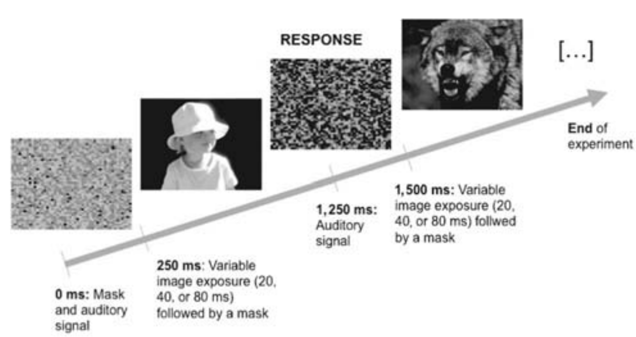

Figure 9. Trial sequence in Experiment 5.

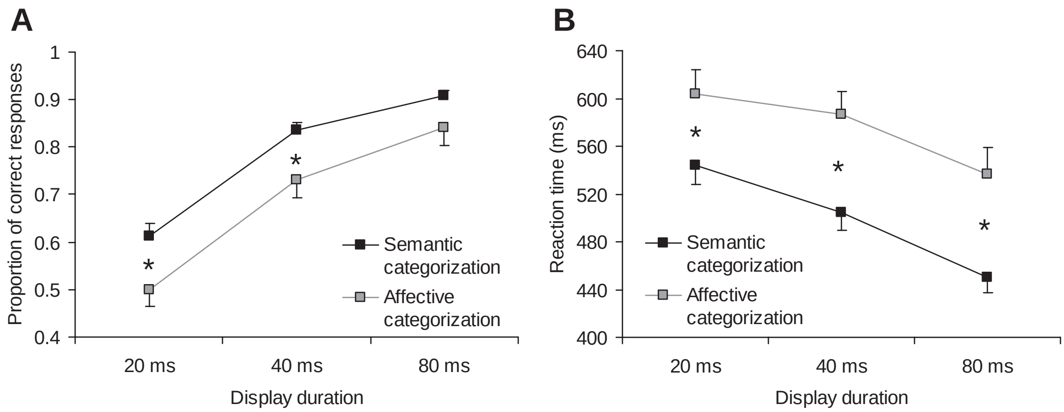

Figure 10. Means and standard errors of accuracy scores (A) and median reaction times and their standard errors (B), as a function of categorization task (affective vs. semantic) and exposure duration (20 ms, 40 ms, or 80 ms) in Experiment 5.

accurate classifications made in the semantic task than in the affective task (M = .78 and .69, respectively). The main effect of exposure duration was also significant, $F(2, 22) = 236.55$, $p <$ .001, $\eta_p^2 = .96$, and multiple comparisons show that accuracy became progressively better with each increase in exposure duration (Fs(1, 11) > 136.24, ps < .001). For RTs, the ANOVA yielded a main effect of task, $F(1, 11) = 12.90$, $p < .004$, $\eta_p^2 =$ .54, which resulted from RTs being faster for the semantic task than the affective task (M = 499 ms and 575 ms, respectively). There was also a main effect of exposure duration, $F(2, 22) =$ 56.70, $p < .001$, $\eta_p^2 = .84$, reflecting the fact that RTs became progressively faster as the exposure duration increased (Fs(1, 11) > 15.19, ps < .001).

As each image was presented once with each exposure duration and the order of the exposure durations was fully randomized, we were also able to test the effect of image repetition on recognition performance at the critical 20-ms exposure time, at which semantic but not affective recognition was reliable. In other words, we tested whether recognition of the scenes at the 20-ms display duration was enhanced by seeing them previously with the 40 ms and 80 ms exposures. A 2 (evaluation task: affective valence vs. semantic category) × 3 (repetition: 1st, 2nd, 3rd) repeated-measures ANOVA revealed a main effect of task, $F(1, 11) = 5.50$, $p < .05$, $\eta_p^2 = .33$, and repetition. $F(2, 22) = 3.95$, $p = .04$, $\eta_p^2 =$ 26, with recognition performance being better in the semantic task than in the affective task (means .61 and .50, respectively), and with each repetition enhancing performance (Ms = .52, .57, and .59, respectively). Although the task by repetition interaction did not reach significance, we conducted planned comparisons as we had an a priori prediction regarding a repetition effect specifically in the affective task. Specifically, we have previously found that processing of affect from briefly displayed pictorial scenes is contingent on prior foveal or repeated parafoveal exposure to the scenes (Calvo & Nummenmaa, 2007), so we wanted to assess whether similar preexposure is specific to affective processing or whether it is required for semantic recognition as well. As predicted, repeated stimulus exposure enhanced performance in the affective task, $ts(11) > 1.91$, ps < .04; however, this was not the case in the semantic task, $ts(11) < 1.10$, ps > .15. One-sample t tests revealed that semantic recognition performance was above chance level at all repetitions (1st = .59; 2nd = .63; 3rd = .63), $ts(11) > 1.80$, ps > .05, whereas affective recognition never exceeded the chance level (1st = .44; 2nd = .51; 3rd = .56), $ts(11) < 1.34$, ps > .10.

Discussion

The key finding of Experiment 5 was that affective evaluation required a longer exposure time than did semantic categorization to be performed above the chance level. Whereas semantic categorization was already well above chance at the 20-ms exposure duration, affective evaluation of the same scenes could not be reliably performed. Affective evaluation of the scenes required at least a 40-ms (i.e., 20 ms longer) exposure time for reliable classification; at the 80-ms display duration, the accuracy for affective and semantic classification became similar in magnitude. The RT data from Experiment 5 are also in line with those obtained in the previous experiments. Across exposure durations, semantic recognition was constantly faster than affective evaluation, with a mean difference of 67 ms. The data can also be interpreted as an indication that the visual information available within a 20-ms display duration was sufficient for semantic recognition but not for affective evaluation. Accordingly, different amounts (or types) of visual information are required for superordinate-level semantic categorization and affective processing. It is thus possible that the diagnostic information (within a single scene) differs for affective and semantic processing (Schyns, 1998), in that semantic classifications could be performed with more sparse (and easily attainable) information. Hence, the RT results of Experiments 2–4 would not be directly related to the processing speed of the affective and semantic recognition systems only. Rather, they would tell us how rapidly these systems could receive and process the relevant visual information. However, it must be noted that Experiment 3 established that RTs for classifying a set of images as affective and classifying them as containing animals were positively correlated, suggesting that at least partially similar diagnostic information was used in both tasks. Additionally, even if different diagnostic information was used in the affective or semantic tasks, this would not rebut our key argument that semantic categorization operations are accomplished prior to affective evaluations.

These data strongly suggest that semantic information processing begins and is completed prior to affective evaluation. Affective information does not have privileged access to awareness, at least when the timing of the formation of a consciously accessible representation is considered. Although unconscious affective processing (i.e., without conscious access to the emotional meaning of the visual input) may occur with subliminal presentation (i.e., with a 20-ms exposure duration of the stimulus followed by a mask), as revealed by peripheral physiological responses (Dimberg et al., 2000; Esteves et al., 1994; Glascher & Adolphs, 2003), it is noteworthy that in the present experiment a conscious representation of the semantic category membership of scene objects was also already available at this time (see also Grill-Spector & Kanwisher, 2005). This casts doubt on the claims regarding affective primacy in visual perception.

Combined with studies showing affective priming with subliminal stimulus presentation (Hermans et al., 2003; S. T. Murphy & Zajonc, 1993), the present data can be interpreted as demonstrating that affective information may actually have restricted access to awareness. Although some sort of affective information may be extracted during a subliminal 20-ms masked presentation (Dimberg et al., 2000; Esteves et al., 1994; Glascher & Adolphs, 2003), unlike semantic information, affective information does not become consciously accessible during this time (as shown in Experiment 5). Further, our data point out that the operational definition of subliminal perception is strictly linked to the visual task at hand, such that in the current experiment the 20-ms exposure duration was subliminal for affective categorization but not for semantic categorization. In sum, Experiment 5 demonstrated that during scene perception the cognitive system initially has access to the semantic content of the scene (as indexed by the minimum exposure time required for accurate recognition), whereas affective analysis requires a longer stimulus exposure time and is performed significantly later than semantic analysis (as indexed by RTs).

Experiment 6: Hierarchy of Semantic and Affective Categorization Operations

Semantic categorization has been proposed to operate hierarchically, so that one of the levels in the semantic taxonomy (e.g., superordinate, basic, and subordinate level) is always accessed first (Rosch et al., 1976). Although there is some controversy about what that entry level might be, recent studies with rapid visual classification tasks suggest that the superordinate level (such as animal vs. human categories in Experiments 1–5) can be accessed faster than the basic level (such as snake vs. bird; Macè, Joubert, Nespoulous, & Fabre-Thorpe, 2009). However, as opposite results have also been reported (Rosch et al., 1976), it is important to test whether both superordinate-level and basic-level semantic categorization or only superordinate-level semantic categorization holds primacy over affective categorization. Such data would be revealing as to whether affective recognition can bypass the initial stages of object recognition. Furthermore, it would be theoretically important to establish where affective categorization should be located in the hierarchy of categorization: Does affective evaluation have similar perceptual and cognitive constraints as basic-level categorization (thus being equally fast), or does affective evaluation require basic-level categorization to be completed first (thus being slower than basic-level categorization)?

In Experiment 6, we explored this issue by contrasting affective categorization speed with both a superordinate-level and a basic-level categorization task. We presented participants with unpleasant and pleasant scenes, some of which involved animals (including snakes) and some of which did not. We extended the design of Experiment 5; The participants performed three different 2AFC tasks: a superordinate-level categorization task (animal detection), a basic-level categorization task (snake detection), and an affective categorization task (valence detection). Only one scene was presented on each trial. It is critical that the snake pictures served as targets in all the tasks because they could be classified at all three levels. Accordingly, the classification task was not confounded by visual differences between target scenes.

By comparing the speed with which observers can classify snake scenes as containing an animal, a snake, or an unpleasant event, we aimed to assess the hierarchy of semantic and affective categorization and to compare serial and parallel models of semantic and affective processing. Specifically, if semantic recognition strictly precedes affective evaluation (i.e., processing is fully serial), recognizing a snake as an animal (superordinate level) would not be sufficient for subsequently classifying the snake as unpleasant, for which the serial system would also need to compute the basic-level category (i.e., to recognize the stimulus as a snake) prior to proceeding with affective evaluation. Accordingly, a serial model predicts that superordinate-level recognition is fastest, followed by basic-level categorization and, subsequently, by affective evaluation. However, it is also possible that affective processing and semantic processing are parallel (or only partially serial), in that the initial superordinate-level processing would precede affective categorization, but subsequent basic-level and affective categorization would be undertaken in parallel. Hence, a parallel model predicts that superordinate-level processing is fastest, followed by concurrent basic-level and affective categorization.

Method

Participants. Eighteen psychology undergraduates from the University of Turku (11 female, 5 male; with a mean age of 23 years) participated for course credit. Data from 2 persons had to be discarded because they failed to obey the task instructions in one or more task conditions.

Stimuli. The target pictures were 32 unpleasant and 32 pleasant scenes involving humans (same as in Experiment 5), 32 unpleasant scenes involving snakes (all new), and 32 unpleasant scenes involving animals other than snakes (mostly the same as in Experiment 5). Stimulus size and apparatus were similar to those in Experiment 5. In the superordinate and affective categorization tasks, the stimuli were 32 pleasant and 32 unpleasant scenes involving humans, 32 pleasant scenes involving animals, and 32 unpleasant scenes involving snakes. In the basic-level categorization task, the stimuli were 32 unpleasant animal scenes with snakes and 32 unpleasant animal scenes with other animals (e.g., sharks, spiders, etc.). Such design allowed the use of the same critical snake targets in all tasks and yielded a balanced (50%/50%) distribution of target and nontarget trials in all tasks. The trials not involving snakes were fillers that were not comparable across all conditions and were thus not analyzed.

Procedure. The participants were told that they would be presented with photographs of people or animals, which could be either unpleasant or pleasant and which could contain snakes or other animals. The participant's task was to make a superordinate-level (animals vs. humans), a basic-level (snakes vs. other animals) and an affective valence (pleasant vs. unpleasant) categorization, all in separate blocks. In all tasks, the participants were to respond with a key press of the letter *Z* or the letter *M* as soon as possible, using the right or left index finger.

Each trial began with a fixation cross displayed for a random time (500 to 1,000 ms), after which the target scene was flashed at the center of the screen for 30 ms. The target scene was followed by a 1-s blank screen. Limited response time was used so that performance would rely only on perceptual representations with no interference from linguistic representations (c.f. Macè, et al., 2009). The participants were encouraged to respond to each photograph even if they felt that they had not perceived the scene at all (in which case they were asked to guess). Response accuracy and RTs were collected. In each block, all the target scenes were presented twice; each block was preceded by 12 practice trials representing the forthcoming experimental condition. The order of blocks was fully counterbalanced across participants. The experiment involved a three-level categorization task (superordinate vs. basic level vs. affective) fully within-participants design.

Results

Participantwise median RTs of correct responses and response accuracy were computed. One-sample *t* tests revealed that classification accuracy exceeded the chance level in all the experimental conditions ($ts > 9.20$, $ps < .001$). For the proportion of correct responses, there were no differences among the superordinate level, the basic level, and the affective tasks, $F(2, 30) = 1.90$, $p = .17$ ($M_{\text{superordinate level}} = .97$, $M_{\text{basic level}} = .95$, $M_{\text{affective}} = .91$). For response latencies, there was a significant main effect of task, $F(2, 30) = 23.82$, $p < .001$, $\eta_p^2 = .614$, with superordinate-level classification being faster than basic level, $t(15) = 2.80$, $p = .01$, or affective categorization, $t(15) = 7.37$, $p < .001$, and basic-level categorization being faster than affective categorization, $t(15) = 3.72$, $p = .002$. See Figure 11 for a summary of the results. We also analyzed itemwise (F2) data to assess how consistently semantic information was processed prior to affect. Of the 32 snake scenes, all were assessed faster as animals (superordinate level) than as unpleasant, and 81% (26 out of the 32 scenes) were assessed faster as snakes (basic level) than as unpleasant, suggesting that affective analysis could practically never precede initial semantic recognition and only very infrequently bypass a secondary step (here basic-level categorization) of semantic categorization. Additionally, it was found that 81% of the scenes were classified faster at superordinate level than at basic level.

Discussion

In line with Macè and colleagues (Macè et al., 2009), we found that superodinate-level categorization was faster than basic-level categorization, although a small number of stimuli (19%) were recognized faster at the basic level. However, although our affective stimuli were chosen to represent a canonical and biologically relevant threat signal (i.e., snakes, see Ohman & Mineka, 2001), we again failed to observe any evidence for affective primacy over semantic recognition in visual processing. Quite the contrary, we observed that affective processing was slower than both superordinate-level and basic-level processing. Furthermore, response accuracy was well above chance level and practically identical in the superordinate, basic-level, and affective categorization tasks (mean accuracies were .97, .95, and .91, respectively). This provides an important control for the prior experiments by showing that even when the task difficulty is matched for semantic and affective processing, the processing speed for semantic information exceeded that for affective information. It is interesting that there was almost a linear ($R^2 = .98$; see Figure 11) increase in RTs from superordinate via basic level to affective categorization, which implies that affective processing is a similar extra step (in terms of time needed) in visual categorization as basic-level categorization is to superordinate-level categorization.

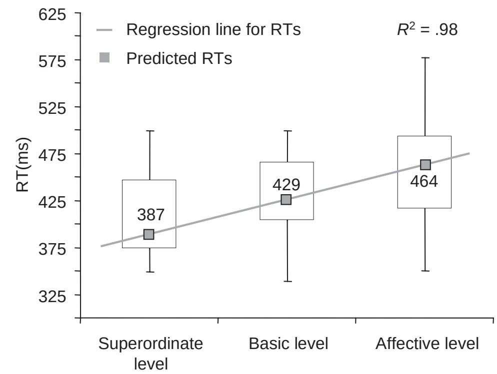

Figure 11. Distributions of median reaction times and their standard errors, as a function of categorization task (superordinate vs. basic level vs. affective), and regression line and predicted values of the RTs in Experiment 6. RT = reaction time.

Our findings are in line with a serial model of semantic and affective processing, according to which processing begins with global semantic categorization and proceeds then to finer-grained semantic discriminations, which in turn leads to affective analysis. However, as indicated by the F2 analyses, later stages of semantic processing and affect processing do not take place strictly serially, as on infrequent occasions affective associations could be retrieved prior to basic-level category information. It is possible that on some occasions individuals have established strong links between certain salient visual features and their affective valence; hence a rapid detection of these features could serve as a shortcut for retrieving affective associations (Calvo & Nummenmaa, 2008). To sum up, we argue that the earliest, superordinate-level recognition always precedes affective processing but that affective processing can be initiated as soon as sufficient information about the scene contents is available; on most (but not all) occasions this requires subordinate-level information about the object category (see the General Discussion).

Experiment 7: Are Animals a Special Case?

All the previous experiments have relied on the comparison of affective processing with animal detection. However, one may ask whether the animals constitute a special category due to their evolutionary significance, thus their processing advantage over affective content could be an exception rather than a rule (New, Cosmides, & Tooby, 2007). Although a similar argument could be put forward to support the special affective role of predatory animals, snakes, and spiders (see Ohman & Mineka, 2001, yet their affective processing was slower than their semantic categorization) used in Experiments 3–6, we considered it important to replicate our findings with a nonanimal stimulus set and a different task. This approach allows us to determine whether the semantic primacy in visual categorization is a genuine, generalizable effect and not restricted to animal stimuli.

In Experiment 7, we used the 2AFC saccade task (see Experiments 1–3) but somewhat changed the semantic task. Participants were briefly and simultaneously presented with two visual scenes that differed in affective valence (pleasant or unpleasant vs. neutral) and in the gender of the people depicted in the scenes. In the semantic task, viewers were to saccade as quickly as possibly toward a scene involving a female target rather than toward the paired male distracter scene.

We reasoned that a gender categorization task would be an appropriate (and a more demanding) control for the affective task for two reasons. First, gender is a subordinate-level category of human, and it is not likely to be the entry-level category for human figures. Hence, recognition of gender should require even more processing than the animal (or snake) detection examined in Experiment 6. Second, as both the target and distracter scenes now involve humans and the gender target stimuli involve one or more women (potentially embedded with men), no single diagnostic criterion can be used for female detection. Accordingly, if the semantic processing superiority holds even under such conditions, it would be strong evidence in support of the semantic superiority hypothesis.

Method

Participants. Forty psychology undergraduates at La Laguna University (34 female, 6 male; with a mean age of 21 years old) participated for course credit.

Stimuli and procedure. Sixty-four target pictures portraying unpleasant (32) or pleasant (32) scenes, and 64 pictures portraying neutral scenes were used. All scenes involved people. In half of the unpleasant and pleasant scenes, there was at least 1 female participant (either woman or girl, with or without male participants), whereas in the other half of the unpleasant and pleasant scenes there was at least 1 male participant, but no female participants. All the unpleasant and pleasant stimuli and most of the neutral stimuli were selected from the IAPS (Lang et al., 2005).

The general layout of the stimulus displays was analogous to that used in Experiments 1–3. Two pictures were presented simultaneously on each trial: a target picture (either unpleasant or pleasant in the affective task, or depicting a scene with a female in the semantic task) and a distracter picture (a neutral scene in the affective task, or a scene depicting male subjects only in the semantic task). Each participant was presented with two blocks of 128 experimental trials (and 16 practice trials), one with the female detection instructions and another with the instructions to detect affective valence. The order of blocks was counterbalanced across participants. Twenty participants were randomly assigned to each of the two affective valence conditions. In the pleasant scene condition, a pleasant and a neutral scene were presented simultaneously, one of which portrayed a woman (or women; 50% pleasant; 50% neutral) and one of which portrayed only men. In the unpleasant scene condition, an unpleasant scene and a neutral scene were presented, one of which portrayed a woman (or women; 50% unpleasant; 50% neutral) and the other portrayed only men. In the semantic task, the same stimulus pairings were used as in the affective task. Each target scene was presented four times in each task, twice in the left visual field and twice in the left field, each time paired with a different distracter scene.

Design and data analysis. Two experimental factors were orthogonally combined: task (semantic vs. affective), as a within-participants variable, and affective valence (unpleasant vs. pleasant) as a between-participants variable. Data were analyzed similarly to those of Experiments 1–3. Only the critical trials involving affective (unpleasant or pleasant) scenes with women in one of the picture pairs were included in the analyses, as these scenes served as the targets in both the affective and the semantic tasks. The trials involving affective scenes with men and neutral scenes with women were fillers, which ensured that the affective and semantic recognition tasks were orthogonal (see Experiment 3).

Results

A 2 (task: semantic vs. affective) $\times$ 2 (affective valence: unpleasant vs. pleasant) ANOVA was conducted on the dependent measures. For saccade response accuracy, no significant effects appeared (Fs < 1, for task; M semantic $= 69.7\%$; M affective $= 69.4\%$), and for the interaction, the main effect of valence approached significance, $F(1, 38) = 3.51$, $p = .07$, $\eta_p^2 = .08$ (M unpleasant $= 67.6\%$; M pleasant $= 71.5\%$). One-sample t tests were used to examine whether the probability of correct saccades exceeded the chance level (i.e., 50%). It was found that the hit rate was above chance level on all four experimental conditions (M semantic task/unpleasant $= 67.6\%$; M semantic task/pleasant $= 71.8\%$; M affective task/unpleasant $= 67.5\%$; M affective task/pleasant $= 71.3\%$), ts$(19) > 35$, ps $< .0001$, thus confirming that gender and valence were reliably recognized in the scenes. Median saccade latencies were affected by task, $F(1, 38) = 7.26$, $p < .01$, $\eta_p^2 = .16$, but not by valence, ($F < 1$; M unpleasant $= 350$ ms; M pleasant $= 354$ ms); the interaction was also nonsignificant ($F < 1$). Latencies were shorter in the semantic (343 ms) than in the affective task (361 ms).

The minimum latency, that is, the first 20-ms bin that contained more correct responses (i.e., saccades to the target) than errors (saccades to the distracter), was 200 ms for unpleasant scenes in the semantic task, $t(19) = 2.59$, $p < .025$. In contrast, in the affective task the minimum latency was delayed up to 240 ms, $t(19) = 2.26$, $p < .05$. For pleasant scenes in the semantic task, the minimum latency was 180 ms, $t(19) = 2.50$, $p < .025$. In contrast, in the affective task the minimum latency was again delayed up to 240 ms, $t(19) = 2.81$, $p < .025$. The saccadic latency distributions are presented in Figure 12.

Discussion

Using novel stimuli and task instructions, in Experiment 7 we replicated the semantic primacy effect observed in all the previous

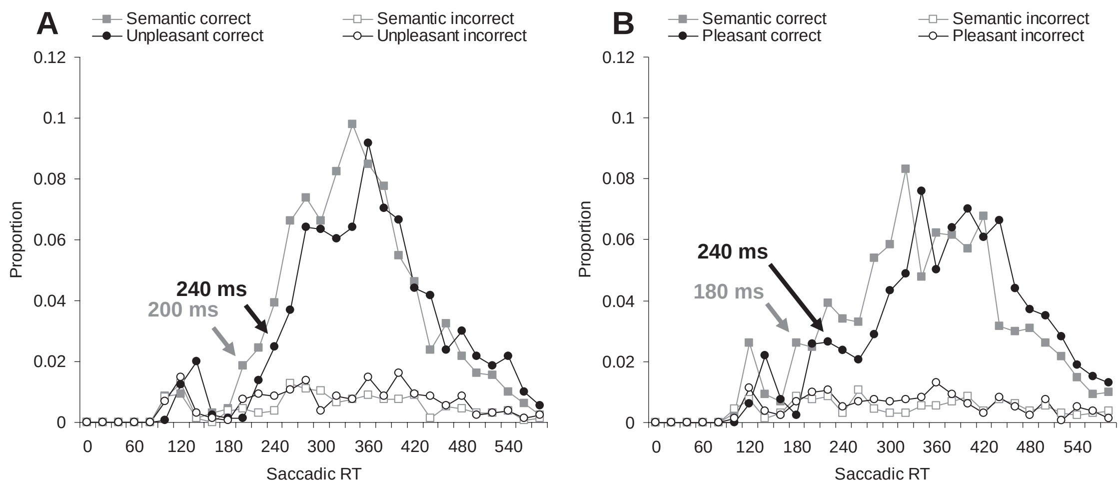

Figure 12. Saccadic latency distributions in Experiment 7 with unpleasant (A) and pleasant (B) female targets. The *x*-axis shows saccadic reaction times (20-ms bins), and the *y*-axis shows the proportion of trials separately for correct and incorrect responses for the semantic (gender) and the affective classification task. Arrows indicate the earliest point in time when the proportion of correct saccades toward the target significantly exceeded that of erroneous saccades toward the distracter. RT = reaction time.

experiments. Participants were faster to detect women embedded in the scenes than to evaluate the valence of the same scenes. When typical (median) saccade latencies were considered, there was a significant 18-ms advantage for semantic over affective recognition; when the minimum latencies were considered, the advantage ranged between 40 ms and 60 ms. Neither finding is consistent with the affective primacy hypothesis (S. T. Murphy & Zajonc, 1993; Stapel et al., 2002; Zajonc, 1980) but instead favor a semantic primacy hypothesis. Moreover, these results confirm that the primacy of animal detection is not a special case of fast semantic categorization. Rather, semantic primacy extends to other natural stimulus categories (here gender) as well. As the same target and distracter stimuli were used for the semantic (gender categorization) and the affective (valence evaluation) tasks, any explanation in terms of low-level visual image factors can be ruled out.

Experiment 7 also revealed that even subordinate-level semantic categorization is faster than affective evaluation. This seems quite striking, given that the subordinate-level information about people's gender was irrelevant to the affective evaluation of scenes. Combining the results of Experiment 6 and 7, we can conclude that the traditional levels of object categorization (Rosch et al., 1976) are accomplished faster than affective categorization. However, we do not know whether the later semantic categorization levels (specifically, the subordinate level) need to be fully completed prior to affective analysis. It may well be that affective analysis begins as soon as sufficiently detailed semantic information (e.g., basic level) is available, but that the later stages (i.e., subordinate level) of object categorization are still performed faster (see General Discussion and Figure 13).

General Discussion

Previous research has demonstrated that the affective significance of emotional scenes is successfully processed with very brief display durations and outside the focus of foveal attention (Calvo & Nummenmaa, 2007). Semantic and gist-level processing is enhanced for affective—relative to neutral—scenes in extrafoveal vision (Calvo & Lang, 2005; Calvo, Nummenmaa, & Hyönä, 2008), and emotional scenes attract attention in a reflexive manner (Kissler & Keil, 2008; Nummenmaa, Hyönä, & Calvo, 2009; Öhman, Flykt, et al., 2001). There is thus no doubt that sensory processing of affective content is enhanced in comparison with nonaffective content. But does this mean that affective content has privileged access to awareness, such that it would be processed prior to recognition of the objects or the semantic content of the scene? The results from the present study provide a negative answer to this question. Across a series of seven experiments, we systematically observed a speed advantage for semantic over affective recognition, and this advantage held (a) when exactly the same scenes were to be classified with respect to affective valence versus semantic category membership, (b) for both foveal and extrafoveal presentation conditions, (c) for superordinate-, basic-, and subordinate-level semantic processing, and (d) with both saccadic and manual responses. Affective valence is thus clearly not the entry-level category of visual scenes. These results have important implications for the neural timing of affective processing as well as for the temporal relationship between affective and semantic recognition, which is next discussed in more detail.

How Fast Is Affective Processing?

As reviewed in the introduction, ERP studies have yielded variable estimates of lower bound latency for emotional processing of complex visual scenes, ranging from 100 ms (Carretié, et al., 2007; Carretié, et al., 2004; Keil et al., 2001; Stolarova, Keil, & Moratti, 2006) to 200–300 ms (Cuthbert et al., 2000; Junghöfer, et al., 2001; Schupp et al., 2003; Schupp et al., 2004). Using a multiexperiment approach with a large stimulus set controlled for low-level visual features, we established that conscious affective processing may take as little as 200–220 ms to be accomplished. This lower bound latency was derived from the saccadic latency distributions as the earliest time when the proportion of correct responses exceeded that of errors and remained higher for at least the next 100 ms. Although the saccadic latency distributions show slight bimodality, with the first peak around 100–120 ms potentially reflecting express saccades (see Fischer & Weber, 1993), the saccades within this latency range were not accurately directed toward the correct target. Thus, express saccades cannot account for the findings. Rather, correct saccades were generally executed within a normal saccade latency range (i.e., between 175 ms and 200 ms, see Rayner, 1998).

It must, however, be noted that 200 ms is the lower bound of the conscious processing speed. The median latency for affective categorization is over 100 ms longer (~300 ms). Accordingly, there is considerable variability in the processing speed of different scenes (see e.g., Figures 3–5): Although the quickest pictures may be classified on average in 250 ms, the slowest scenes take over 100 ms longer to be categorized. It is interesting to note that our median estimates for affective processing speed fall within the typical EPN latency range (~300 ms). Although EPN modulations are typically observed for arousal rather than valence (e.g., Junghöfer et al., 2001; Schupp et al., 2003; Schupp et al., 2004), the present results suggest that conscious perception of scene pleasantness or unpleasantness (rather than mere arousal level) may already be available within the EPN latency range. Nevertheless, it is likely that some stages of the valence (or arousal) recognition process begin earlier, similarly to what has been reported for object recognition (see e.g., Liu et al., 2002; Meeren et al., 2008, for early category-selective responses). The analysis of the earliest time bin for reliable classification suggests that some form of valence processing begins to unfold in less than 200 ms (i.e., after removing a 25-ms saccadic programming time) poststimulus. Accordingly, this suggests that the early affect-sensitive responses recorded intracranially (Kawasaki et al., 2001; Oya et al., 2002) or extracranially (Carretié, et al., 2004) may already reflect selective affect-category processing.

Semantic Recognition of Complex Scenes Precedes Their Affective Evaluation

The major contribution of the present series of experiments is that they show that semantic recognition of objects in scenes is faster than their affective evaluation. With respect to the minimum recognition latency, semantic processing was found to be on average 20 ms faster than affective processing. When typical (median) latencies were considered, the advantage of semantic over affective encoding was more than 60 ms, and it was found to be larger in the visual periphery than in foveal vision (Experiment 4). As regards semantic recognition, our latency estimates fit reasonably well with those obtained in tasks involving semantic categorization of complex pictorial scenes (Kirchner & Thorpe, 2006; Li, VanRullen, Koch, & Perona, 2002; Thorpe et al., 2001). We extend those findings by showing that the visual processing speed of semantic information cannot be exceeded by the speed of affective information processing and that affective processing cannot bypass the elementary stages of object recognition. Rather, our findings support the view that semantic recognition of scenes and objects precedes their affective analysis.

It has recently been suggested that although superordinate-level object categorization (i.e., animal vs. car) does not require more visual processing than mere object detection (i.e., object vs. no object), subordinate-level categorization or object identification (i.e., Porsche 911 vs. Volkswagen Beetle) requires substantially more processing time (Grill-Spector & Kanwisher, 2005). In Experiment 5, we extended these findings by showing that relative to superordinate-level object categorization, affective recognition requires substantially longer visual exposure times (at least 40 ms) and more encoding time (76 ms). Furthermore, when both basic- and superordinate-level categorization were contrasted with affective processing, it was found that basic-level categorization took ~40 ms longer to complete than superordinate-level categorization and that affective evaluation took an additional ~40 ms (i.e., 80 ms longer than superordinate-level categorization). This suggests that object identification and affective evaluation are undertaken in a hierarchical fashion and that they share similar processing constraints, although they are not necessarily undertaken by similar neurocognitive systems. Altogether, these data show that recognizing the object's basic category is necessary but not sufficient for its affective evaluation: The visual input that is sufficient for superordinate-level categorization is not sufficient for affective recognition. Affect recognition can only proceed after basic-level category information of the target object has been acquired, which clearly contrasts with the affective primacy view.

The semantic processing advantage observed in Experiments 2, 3, and 7 involving paired target–nontarget stimulus presentation can be related to the well-established role of emotional pictorial scenes in guiding attentional deployment in a reflexive manner (Nummenmaa et al., 2006). The current findings suggest that automatic attentional shifts toward emotional scenes must be preceded by semantic categorization of scene elements, followed by an access to the affective valence of (some of) the elements. In other words, if affect had been processed prior to semantic content, the attentional bias toward emotional scenes should have resulted in faster saccadic responses in the affective task than in the semantic task. However, an opposite pattern of results was obtained. This is in line with the results of a study with a parafoveal affective priming paradigm (Calvo & Nummenmaa, 2007), which showed that prior exposure to the prime scenes was required for affective priming to occur, thus suggesting that object recognition is required for affective evaluation. In other words, these data are consistent with the proposal that the visual system does not directly detect affect but instead takes care of object recognition, which then leads to rapid retrieval of object-relevant affective associations (Storbeck et al., 2006). Affective analysis is thus contingent on basic perceptual and cognitive analysis and is unlikely to be accomplished preattentively or in parallel with object recognition.

It should be acknowledged that the present series of experiments did not allow us to assess whether there are differences in the affective versus semantic task as to what scene information was diagnostic in accomplishing the task (see e.g., Schyns, 1998). Even though the same target images were used in Experiments 3–7, differences in the diagnostic information could have influenced the results. Because we used a large sample (>250) of naturalistic images of which practically all were categorized faster in the semantic domain than in the affective domain, we are inclined to interpret potential differences in the diagnostic information as an inherent property of the visual arrays one typically encounters in daily visual environments rather than as a serious confounding factor. Nevertheless, in future studies it would be interesting to use, for example, the bubbles task (Gosselin & Schyns, 2001) or other reverse correlation techniques for estimating the diagnostic information used for affective and different levels of semantic recognition.

Semantic Recognition Is Necessary for Affective Evaluations

Two alternative models could account for the semantic recognition superiority obtained in the present experiments. First, it is possible that affective and semantic recognition are processed by fully independent, parallel systems and that emotional processing simply takes longer to accomplish. Alternatively, emotional and semantic information may be processed by interacting albeit serial systems, in which the semantic recognition system feeds its output to the affective evaluation system (an extra step). Hence, an increased latency in the semantic recognition stage is also reflected in the timing of affective processing.[2] The data from Experiment 3 that are critical to this question are consistent with the latter view. When participants were asked to classify exactly the same scenes with respect to category membership or affective valence, the semantic classification of a scene was practically always (i.e., for all scenes) faster than its affective evaluation. It is important to note that itemwise affective and semantic recognition times were positively correlated. This indicates that the time taken for affective evaluation sums up to the time taken by semantic categorization, thus suggesting that affective processing is an additional processing step in a serial system: The visual recognition system thus first encodes object category and then passes this information onto the emotion system for affective analysis.

The serial processing view proposed above and the critical role of IT in visual affective processing are supported by intracranial recordings in primates, as well as human patient studies. In monkeys, it has been established that synaptic cooling of neurons in IT (whose cells are insensitive to affective valence, see e.g., Rolls, 1999) attenuates the responses of food-reward selective cells in the amygdala and also changes response profiles of some food-selective cells to nonselective (Fukuda, Ono, & Nakamura, 1987). In a similar vein, a prosopagnosic patient, LF, for whom connections between occipitotemporal and limbic areas were lacking, lost the ability to become emotionally aroused by visual stimuli but not by auditory stimuli (Bauer, 1982), thus suggesting that output from the temporal area to the limbic areas is required for affect processing. Combined with the finding that affective blind sight (i.e., affective evaluation in the absence of visual cortex) does not extend to complex pictorial scenes (de Gelder et al., 2002), it seems reasonable to argue that affective evaluation of complex pictorial scenes must be preceded by at least rudimentary object recognition in the ventral visual stream. In the future, transcranial magnetic stimulation (TMS) studies targeting the IT cortices during affective and semantic categorization would be particularly useful in addressing this issue in detail.

Figure 13 provides a summary of the temporal relationship between semantic and affective processing. After visual attention selects the relevant visual input to be processed, semantic processing of scene contents begins in a hierarchical fashion (starting from superordinate level, leading to basic-level and subordinate-level categorization). This is followed by affective evaluation of scene contents, which on most occasions requires basic-level category information of the scene objects. This processing stream is indexed by the thick gray arrows. However, affective analysis may proceed as soon as sufficiently detailed information of the scene contents is obtained, which may sometimes occur prior to basic-level categorization. Hence, earlier stages of the object recognition system may also output to the affective system, although these links are much weaker than those from the basic-level processing stage, as indicated by the thin dotted arrows.

The processing of semantics and affect is thus mostly serial. When categorization is completed, the information is passed onto cognitive systems responsible for attentional deployment, response selection, and so forth. Our results thus extend the predictions of the parallel distributed processing theory (McClelland & Rogers, 2003) of visual recognition: Broadest categories (e.g., animal) are activated first, followed by more specified categories (e.g., snake) and, ultimately, the affective categorization of the recognized object by the emotion system. Accordingly, the attentional and perceptual biases resulting from emotional scene content are contingent on semantic and perceptual analyses of scene content (Calvo & Nummenmaa, 2007) and must thus be explained by strong links between attentional circuits and emotional processing mechanisms, rather than by earlier or privileged processing of affect.

Can Affective Processing Sometimes Precede Semantic Recognition?

It needs to be stressed that the primacy of cognition over affect observed in the present study may not hold for all types of information processing or for all sensory domains. First, it is evident that some cognitive operations such as complex reasoning are bound to be slower than affective evaluation, so it is clear that cognitive operations do not always precede affect. Our argument is thus that when novel visual arrays are encountered, semantic recognition must precede affective evaluation of the scene contents. However, it is interesting to ask whether affective processing could sometimes precede semantic recognition of scene objects. We feel that under some circumstances this may be possible. Namely, the present Experiment 5 as well as data from our affective priming studies (Calvo & Nummenmaa, 2007) suggest that learning facilitates extraction of affect from visual scenes. Accordingly, repeated exposures to specific emotional scenes or procedures such as fear conditioning (Morris, Ohman, & Dolan, 1998) could tune the response properties of the affect system in such way that recognition of the emotional meaning of that specific event is facilitated. If the affect system has very narrowly tuned representations of certain emotional events, these could lead to rapid recognition of the affective valence of these events, which may occur prior to semantic recognition. Via learning, the initial semantic recognition primacy could thus sometimes be flipped around into an affective primacy effect.

Additionally, it must be kept in mind that our data are strictly confined to the visual sensory domain. Other sensory modalities

[2] The data could also be accounted for by a single-system processing object category and affect. However, given the well-established role of specialized neural circuits for affective processing (Kober et al., 2008; LeDoux, 1995; F. C. Murphy et al., 2003), this alternative does not seem likely and is not be considered here.

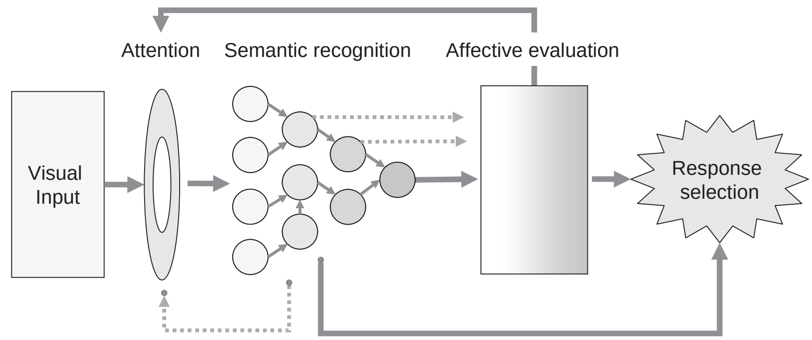

Figure 13. Schematic model of the semantic recognition advantage demonstrated in the present experiments. The relative thickness of the arrows depicts the strength of the connections between processing units. Visual attention selects features from the visual array for further scrutiny. This information is passed onto the parallel distributed mechanism for object recognition, whose output is then fed to the emotion evaluation system. Depending on scene complexity and content, different levels of semantic recognition (superordinate/basic/subordinate) are required for affective evaluation to start. Affective biases observed, for example, in the domain of visual attention are contingent on prior semantic recognition of scene content and are explained by the strong reciprocal links between the emotion recognition and attention systems rather than rapid processing of affect.

have different constraints for semantic and affective processing, which might give rise to affective primacy. For example, during speech perception, individuals may be able to grasp the affective valence of the voice quickly from nonlinguistic cues (see e.g., Brosch, Grandjean, Sander, & Scherer, 2008), whereas semantic processing of the message will require processing of longer speech segments, which renders semantic processing slower than affective processing.

Can Visual Factors and Task Effects Confound the Results?

One potential criticism of the present results is that differences in processing speed could be accounted for by low-level visual factors (c.f. Bradley et al., 2007). This is a serious issue, given that we used eye-movement responses toward laterally presented visual scenes, and the role of low-level visual features in guiding eye movements and visual attention is well-known (Henderson, 2003; Itti & Koch, 2001). However, there are four arguments against a low-level visual explanation. First, because our experimental tasks involved categorization of scenes in which the overall composition, as well as the size and shape of the targets, was greatly variable, it seems unlikely that any single scene feature would be sufficient for explaining the semantic processing advantage (see Thorpe et al., 2001, for similar conclusions). Second, we controlled for various low-level visual differences (see Tables 1 and 2) between the target and distracter scenes. Third, F2-analyses did not reveal significant associations among classification latencies, accuracies, and low-level visual features. Fourth, and most important, Experiments 3–7 used exactly the same scenes as affective and semantic targets. Any potential low-level differences between the target and the distracter scenes were thus the same for the semantic and the affective task. Yet, semantic processing of practically all the target scenes was faster than their affective classification. Finally, the same semantic versus affective categorization advantage appeared when the scenes were presented at fixation—hence no selective eye movements (that might be guided by low-level factors) were necessary—as when they were presented parafoveally. This rules out any explanation based on visual (dis) similarity or low-level image statistics.

It must also be noted that the use of behavioral responses is not fully unproblematic. For example, one prior study (Rousselet, Mace, Thorpe, & Fabre-Thorpe, 2007) showed that even when manual response latencies for animal and human face targets are similar in a detection task, simultaneous ERP recordings show that visual processing of faces begins earlier than that of animals. In a similar vein, it could be argued that affective processing might begin earlier than semantic processing. We admit this possibility, but even if the above reasoning was true, it does not rebut our argument that the cognitive system can both complete processing (categorize the target) and access (use the outcome of the categorization for decision making) and react to (by performing an endogenous saccade or a manual keypress response) semantic information earlier than affective information.

Conclusions

We conclude that semantic processing of visual scenes is faster than their affective processing and that semantic categorization precedes affective evaluation. Although affective valence can be extracted from minimal visual input, its encoding does not hold primacy over semantic recognition. The biases that emotional content exerts over cognitive processes (for example, on selective visual attention) would occur after semantic recognition of scene objects. Thus, affective analysis is contingent on attentive object identification. Nevertheless, the present results do not undermine the claims that affective information is prioritized over nonemotional information by the cognitive system. Rather, the present study qualifies this claim by suggesting that the prioritization of affective information must occur after semantic processing of visual information, which can lead to a perceptual and attentional bias toward affective scenes. In conclusion, when you notice a curvy and tube-shaped object moving around in the garden, you had better recognize it as a snake. Otherwise, you would fail to notice the threat you may soon be facing.

References

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer (Ed.), *Advances in social cognition: The automaticity of everyday life* (Vol. 10, pp. 1–61). Mahwah, NJ: Erlbaum.

Batty, M. J., Cave, K. R., & Pauli, P. (2005). Abstract stimuli associated with threat through conditioning cannot be detected preattentively. *Emotion, 5,* 418–430.

Batty, M., & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research, 17,* 613–620.

Bauer, R. M. (1982). Visual hypoemotionality as a symptom of visual-limbic disconnection in man. *Archives of Neurology, 39,* 702–708.

Bradley, M. M., Hamby, S., Low, A., & Lang, P. J. (2007). Brain potentials in perception: Picture complexity and emotional arousal. *Psychophysiology, 44,* 364–373.

Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics, 62,* 1572–1595.

Brosch, T., Grandjean, D., Sander, D., & Scherer, K. R. (2008). Behold the voice of wrath: Cross-modal modulation of visual attention by anger prosody. *Cognition, 106,* 1497–1503.

Calvo, M. G., & Lang, P. J. (2004). Gaze patterns when looking at emotional pictures: Motivationally biased attention. *Motivation and Emotion, 28,* 221–243.

Calvo, M. G., & Lang, P. J. (2005). Parafoveal semantic processing of emotional visual scenes. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 502–519.

Calvo, M. G., & Nummenmaa, L. (2007). Processing of unattended emotional visual scenes. *Journal of Experimental Psychology: General, 136,* 347–369.

Calvo, M. G., & Nummenmaa, L. (2008). Detection of emotional faces: Salient physical features guide effective visual search. *Journal of Experimental Psychology: General, 137,* 471–494.

Calvo, M. G., Nummenmaa, L., & Avero, P. (2008). Visual search of emotional faces: Eye-movement assessment of component processes. *Experimental Psychology, 55,* 359–370.

Calvo, M. G., Nummenmaa, L., & Hyönä, J. (2008). Emotional scenes in peripheral vision: Selective orienting and gist processing but not content identification. *Emotion, 8,* 68–80.

Carretié, L., Hinojosa, J. A., López-Martín, S., & Tapia, M. (2007). An electrophysiological study on the interaction between emotional content and spatial frequency of visual stimuli. *Neuropsychologia, 45,* 1187–1195.

Carretié, L., Hinojosa, J. A., Martín-Loeches, M., Mercado, F., & Tapia, M. (2004). Automatic attention to emotional stimuli: Neural correlates. *Human Brain Mapping, 22,* 290–299.

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology, 52,* 95–111.

de Gelder, B., Pourtois, G., & Weiskrantz, L. (2002). Fear recognition in the voice is modulated by unconsciously recognized facial expressions but not by unconsciously recognized affective pictures. *Proceedings of the National Academy of Sciences of the United States of America, 99,* 4121–4126.

de Gelder, B., Vroomen, J., Pourtois, G., & Weiskrantz, L. (1999). Non-conscious recognition of affect in the absence of striate cortex. *NeuroReport, 10,* 3759–3763.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. J. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience, 4,* 2051–2062.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science, 11,* 86–89.

Eimer, M., & Holmes, A. (2002). An ERP study on the time course of emotional face processing. *NeuroReport, 13,* 427–431.

Esteves, F., Dimberg, U., & Ohman, A. (1994). Automatically elicited fear— Conditioned skin-conductance responses to masked facial expressions. *Cognition & Emotion, 8,* 393–413.

Fischer, B., & Weber, H. (1993). Express saccades and visual-attention. *Behavioral and Brain Sciences, 16,* 553–567.

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition & Emotion, 14,* 61–92.

Fukuda, M., Ono, T., & Nakamura, K. (1987). Functional relations among inferotemporal cortex, amygdala, and lateral hypothalamus in monkey operant feeding-behavior. *Journal of Neurophysiology, 57,* 1060–1077.

Glascher, J., & Adolphs, R. (2003). Processing of the arousal of subliminal and supraliminal emotional stimuli by the human amygdala. *Journal of Neuroscience, 23,* 10274–10282.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15,* 20–25.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41,* 2261–2271.

Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science, 16,* 152–160.

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology, 17,* R751–R753.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7,* 498–504.

Hermans, D., Spruyt, A., De Houwer, J., & Eelen, P. (2003). Affective priming with subliminally presented pictures. *Canadian Journal of Experimental Psychology, 57,* 97–114.

Horstmann, G. (2007). Preattentive face processing: What do visual search experiments with schematic faces tell us? *Visual Cognition, 15,* 799–833.

Horstmann, G., & Bauland, A. (2006). Search asymmetries with real faces: Testing the anger-superiority effect. *Emotion, 6,* 193–207.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2,* 194–203.

Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences, 103,* 17048–17052.

Junghöfer, M., Bradley, M. M., Elbert, T. R., & Lang, P. J. (2001). Fleeting images: A new look at early emotion discrimination. *Psychophysiology, 38,* 175–178.

Kawasaki, H., Kaufman, O., Damasio, H., Damasio, A. R., Granner, M., Bakken, H., . . . Howard, M. A. (2001). Single-neuron responses to emotional visual stimuli recorded in human ventral prefrontal cortex. *Nature Neuroscience, 4,* 15–16.

Keil, A., Muller, M. M., Gruber, T., Wienbruch, C., Stolarova, M., & Elbert, T. (2001). Effects of emotional arousal in the cerebral hemispheres: A study of oscillatory brain activity and event-related potentials. *Clinical Neurophysiology, 112,* 2057–2068.

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research, 46,* 1762–1776.

Kissler, J., & Keil, A. (2008). Look—don't look! How emotional pictures affect pro- and antisaccades. *Experimental Brain Research, 188,* 215–222.

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7–49). Mahwah, NJ: Erlbaum.

Kober, H., Barrett, L. F., Joseph, J., Bliss-Moreau, E., Lindquist, K., & Wager, T. D. (2008). Functional grouping and cortical–subcortical interactions in emotion: A meta-analysis of neuroimaging studies. *NeuroImage, 42,* 998–1031.

Lang, P. J. (1995). The emotion probe—Studies of motivation and attention. *American Psychologist, 50,* 372–385.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International Affective Picture System (IAPS): Digitized photographs, instruction manual, and affective ratings* (Report No. A-6). Gainesville: University of Florida.

Lazarus, R. S. (1984). On the primacy of cognition. *American Psychologist, 39,* 124–129.

LeDoux, J. E. (1995). Emotion: Clues from the brain. *Annual Review of Psychology, 1995,* 209–235.

LeDoux, J. (1996). *The emotional brain.* New York, NY: Simon & Schuster.

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America, 99,* 9596–9601.

Liu, H., Agam, Y., Madsen, J. R., & Kreiman, G. (2009). Timing, timing, timing: Fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron, 62,* 281–290.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience, 5,* 910–916.

Macè, M. J. M., Joubert, O. R., Nespoulous, J.-L., & Fabre-Thorpe, M. (2009). The time-course of visual categorizations: You spot the animal faster than the bird. *PLoS ONE, 4,* e5927.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience, 4,* 310–322.

Meeren, H. K. M., Hadjikhani, N., Ahlfors, S. P., Hämäläinen, M. S., & de Gelder, B. (2008). Early category-specific cortical activation revealed by visual stimulus inversion. *PLoS ONE, 3,* e3503.

Morris, J. S., Ohman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature, 393,* 467–470.

Morris, J. S., Ohman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Sciences of the United States of America, 96,* 1680–1685.

Mouchetant-Rostaing, Y., Giard, M. H., Delpuech, C., Echallier, J. F., & Pernier, J. (2000). Early signs of visual categorization for biological and non-biological stimuli in humans. *NeuroReport, 11,* 2521–2525.

Murphy, F. C., Nimmo-Smith, I., & Lawrence, A. D. (2003). Functional neuroanatomy of emotion: A meta-analysis. *Cognitive, Behavioral, and Affective Neuroscience, 3,* 207–233.

Murphy, N. A., & Isaacowitz, D. M. (2008). Preferences for emotional information in older and younger adults: A meta-analysis of memory and attention tasks. *Psychology and Aging, 23,* 263–286.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness—Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64,* 723–739.

New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences, 104,* 16598–16603.

Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. *Emotion, 6,* 257–268.

Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2009). Emotional scene content drives the saccade generation system reflexively. *Journal of Experimental Psychology-Human Perception and Performance, 35,* 305–323.

Nummenmaa, L., Peets, K., & Salmivalli, C. (2008). Automatic activation of adolescents' peer-relational schemas: Evidence from priming with facial identity. *Child Development, 79,* 1659–1675.

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

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology, 80,* 381–396.

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

Ohman, A., & Soares, J. J. F. (1998). Emotional conditioning to masked stimuli: Expectancies for aversive outcomes following nonrecognized fear-relevant stimuli. *Journal of Experimental Psychology: General, 127,* 69–82.

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology, 77,* 247–265.

Oya, H., Kawasaki, H., Howard, M. A., III, & Adolphs, R. (2002). Electrophysiological responses in the human amygdala discriminate emotion categories of complex visual stimuli. *Journal of Neuroscience, 22,* 9502–9512.

Peli, E. (1990). Contrast in complex images. *Journal of the Optical Society of America A—Optics Image Science and Vision, 7,* 2032–2040.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, B335,* 23–30.

Pessoa, L. (2005). To what extent are emotional visual stimuli processed without attention and awareness? *Current Opinion in Neurobiology, 16,* 1–9.

Pizzagalli, D., Lehmann, D., Koenig, T., Regard, M., & Pascual-Marqui, R. D. (2000). Face-elicited ERPs and affective attitude: Brain electric microstate and tomography analyses. *Clinical Neurophysiology, 111,* 521–531.

Pourtois, G., Dan, E. S., Grandjean, D., Sander, D., & Vuilleumier, P. (2005). Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: Time course and topographic evoked-potentials mapping. *Human Brain Mapping, 26,* 65–79.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422.

Rolls, E. T. (1999). *The brain and emotion.* Oxford, England: Oxford University Press.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439.

Rousselet, G. A., Fabre-Thorpe, M., & Thorpe, S. J. (2002). Parallel processing in high-level categorization of natural images. *Nature Neuroscience, 5,* 629–630.

Rousselet, G. A., Mace, M. J. M., Thorpe, S. J., & Fabre-Thorpe, M. (2007). Limits of event-related potential differences in tracking object processing speed. *Journal of Cognitive Neuroscience, 19,* 1241–1258.

Schiller, P. H., & Kendall, J. (2004). Temporal factors in target selection with saccadic eye movements. *Experimental Brain Research, 154,* 154–159.

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003). Emotional facilitation of sensory processing in the visual cortex. *Psychological Science, 14,* 7–13.

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2004). The selective processing of briefly presented affective pictures: An ERP analysis. *Psychophysiology, 41,* 441–449.

Schyns, P. G. (1998). Diagnostic recognition: Task constraints, object information, and their interactions. *Cognition, 67,* 147–179.

Stapel, D. A., Koomen, W., & Ruys, K. I. (2002). The effects of diffuse and distinct affect. *Journal of Personality and Social Psychology, 83,* 60–74.

Stolarova, M., Keil, A., & Moratti, S. (2006). Modulation of the C1 visual event-related component by conditioned stimuli: Evidence for sensory plasticity in early affective perception. *Cerebral Cortex, 16,* 876–887.

Storbeck, J., & Clore, G. L. (2007). On the interdependence of cognition and emotion. *Cognition & Emotion, 21,* 1212–1237.
Storbeck, J., Robinson, M. D., & McCourt, M. E. (2006). Semantic processing precedes affect retrieval: The neurological case for cognitive primacy in visual processing. *Review of General Psychology, 10,* 41–55.
Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience, 19,* 109–139.
Thorpe, S. J., Gegenfurtner, K. R., Fabre-Thorpe, M., & Bülthoff, H. H. (2001). Detection of animals in natural images using far peripheral vision. *European Journal of Neuroscience, 14,* 869–876.
Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences, 9,* 585–594.
Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron, 30,* 829–841.
Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience, 6,* 624–631.
Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151–175.

Appendix

International Affective Picture System Numbers for the Stimulus Pictures

Neutral pictures: 2,037; 2,102; 2,190; 2,191; 2,191.1; 2,200; 2,220; 2,221; 2,270; 2,272; 2,272.1; 2,305; 2,312; 2,312.1; 2,357; 2,372; 2,383; 2,389; 2,393; 2,393.1; 2,394; 2,396; 2,397; 2,397.1; 2,410; 2,491; 2,493; 2,512; 2,513; 2,513.1; 2,515; 2,560; 2,560.1; 2,575; 2,575.1; 2,579; 2,593; 2,593.1; 2,594; 2,594.1; 2,595; 2,595.1; 2,598; 2,598.1; 2,635; 2,635.1; 2,745.1; 2,745.2; 2,749; 2,749.1; 2,840; 2,850; 2,870; 54,10; 7,493; 7,496; 7,496.1; 7,550; 7,550.1; 7,620; 7,620.1; 9,070; 9,210; and 9,210.1; Unpleasant pictures: 2,399; 2,399.1; 2,683; 2,691; 2,703; 2,716; 2,718; 2,722; 2,799; 2,800; 2,811; 2,900; 3,051; 3,180; 3,181; 3,225; 3,300; 3,350; 6,010; 6,250; 6,313; 6,315; 6,550; 6,560; 8,480; 8,485; 9,250; 9,254; 9,410; 9,415; 9,423; and 9,435; Pleasant pictures: 2,040; 2,070; 2,160; 2,165; 2,311; 2,332; 2,352; 2,540; 2,550; 4,599; 4,610; 4,624; 4,647; 4,658; 4,660; 4,669; 4,676; 4,680; 4,687; 4,694; 4,700; 5,621; 5,831; 5,836; 7,325; 8,021; 8,080; 8,161; 8,186; 8,200; 8,490; and 8,499. Unpleasant pictures with women: 2,141; 2,799; 3,180; 3,181; 3,225; 6,312; 6,313; 6,315; 6,550; 6,560; 6,838; 9,253; 9,254; 9,249; 9,921; and 2,399. Unpleasant pictures with men: 2,490; 2,703; 2,810; 2,811; 2,900; 3,530; 6,010; 6,242; 6,250; 6,821; 6,840; 8,231; 8,485; 9,400; 9,410; and 9,421. Neutral pictures with women: 2,037; 2,104; 2,272; 2,305; 2,312; 2,372; 2,383; 2,389; 2,396; 2,435; 2,513; 2,515; 2,560; 2,579; 2,594; 2,595; 2,598; 2,745.1; 2,850; 7,506; and 9,210. Neutral pictures with men: 2,102; 2,190; 2,191; 2,221; 2,235; 2,270; 2,357; 2,397; 2,410; 2,491; 2,493; 2,520; 2,575; 2,593; 2,635; 2,749; 2,840; 2,870; 2,890; 5,410; 5,875; 7,493; 7,550; and 9,070. Pleasant pictures with women: 2,070; 2,332; 2,340; 2,352; 2,360; 2,540; 2,550; 4,599; 4,641; 4,687; 4,695; 4,700; 5,836; 7,325; 8,032; and 8,461. Pleasant pictures with men: 2,057; 2,154; 2,160; 2,165; 2,260; 2,339; 2,655; 4,572; 4,614; 5,831; 8,021; 8,050; 8,161; 8,185; 8,186; and 8,200.

Received June 12, 2009
Revision received December 7, 2009
Accepted December 8, 2009 ■