Are Coarse Scales Sufficient for Fast Detection of Visual Threat?
Martial Mermillod, Sylvie Droit-Volet, Damien Devaux, Alexandre Schaefer and Nicolas Vermeulen

Psychological Science 2010 21: 1429 originally published online 3 September 2010
DOI: 10.1177/0956797610381503

The online version of this article can be found at:
http://pss.sagepub.com/content/21/10/1429
The ability to rapidly detect threatening stimuli in the environment is one of the most basic survival needs. For instance, quickly recognizing a predator is probably one of the most vital requirements for every species, from invertebrates to reptiles to humans. Detection of threatening information has been linked to the amygdala (Ledoux, 1996; Phelps et al., 2001). In addition, recent research has suggested that rapid detection of threats in the environment may rely on the ability of the human brain to use perceptually crude, low-spatial-frequency (LSF) information to evaluate visual cues of potential danger. This evidence was obtained with neuroimaging methods (Alorda, Serrano-Pedraza, Campos-Bueno, Sierra-Vazquez, & Montoya, 2007; Holmes, Winston, & Eimer, 2005; Morris, Öhman, & Dolan, 1999; Pourtois, Dan, Grandjean, Sander, & Vuilleumier, 2005; Vuilleumier, Armony, Driver, & Dolan, 2003), neural-network modeling (Mermillod, Bonin, Mondillon, Alleysson, & Vermeulen, in press; Mermillod, Vuilleumier, Peyrin, Alleysson, & Marendaz, 2009), and behavioral experiments (Bocanegra & Zeelenberg, 2009; Holmes, Green, & Vuilleumier, 2005). Figure 1 gives an example of an image presented in broad spatial frequencies (BSFs), LSF bands, and high-spatial-frequency (HSF) bands.

LSF and HSF images are processed by two different visual streams at the level of the lateral geniculate nucleus: The magnocellular neurons primarily provide rapid LSF cues that encode configurual features as well as brightness and motion; the parvocellular neurons provide slower HSF information—and, therefore, finer visual details than LSF information—about local shape features, color, and texture (Livingstone & Hubel, 1988). Surprisingly, few behavioral studies have addressed the role of LSF information in fear processing. Many experiments have shown the involvement of the amygdala in fear conditioning and in the processing of potential danger in general (Bechara et al., 1995; Phelps et al., 2001). However, these studies have not examined the specific visual processes involved in behavioral responses to fearful events (for a review, see Delgado, Olsson, & Phelps, 2006). In addition, different psychophysical studies on visual perception...
have investigated the role of spatial-frequency channels in cognitive processes, but without regard to fear processes (Bullier, 2001; Guyader, Chauvin, Peyrin, Hérault, & Marendaz, 2004; Hughes, Nozawa, & Kitterle, 1996; Peyrin, Chokron, et al., 2006; Peyrin, Mermillod, Chokron, & Marendaz, 2006; Schyns & Oliva, 1994; M.L. Smith, Cottrell, Gosselin, & Schyns, 2005).

In terms of fear processing, the rapidity of the detection of dangerous stimuli has been shown to vary also as a function of the meaning of the stimulus for specific species. According to Öhman and Soares (1993), phylogenetic development has resulted in a predisposition to process specific threatening stimuli, such as those that have been identified as dangerous to the survival of the species, in a high-priority mode. From an evolutionary perspective, very basic organisms, such as reptiles or even invertebrate animals, are able to recognize a predator, whereas they are not always able to recognize nonliving threats (e.g., a trap). However, this effect may also be related to lower perceptual and cognitive complexity of recognizing threats posed by living organisms than threats posed by nonliving objects. Therefore, it is possible that the visual detection of danger could be faster for LSF stimuli than for HSF stimuli, and specifically for a living danger compared with a nonliving one. Thus, the effect of LSF information on rapid recognition of threatening stimuli might not be independent of the nature of the stimuli perceived.

To address the question of LSF pathways for fast visual recognition of potential danger, Öhman, Lundqvist, and Esteves (2001) used indirect methods, such as measuring responses to schematic faces. However, this method is limited with respect to the perceptual properties of real stimuli (Mermillod, Vermeulen, Lundqvist, & Niedenthal, 2009). In the case of faces, Schyns and Oliva (1999) have shown that HSF information is important in determining whether a face is expressive, whereas LSF information seems more efficient in determining the specific content of emotional facial expressions (EFs), such as determining whether they are happy or angry. Moreover, these diagnostic HSF and LSF scales were recently specified for each EF (F.W. Smith & Schyns, 2009). Two previous studies have shown the importance of emotion (fearful faces) in directing attention to different parts of the visual field, for instance, when detecting oriented bars (Holmes, Green, & Vuilleumier, 2005) or oriented LSF Gabor stimuli (Bocanegra & Zeelenberg, 2009). Holmes, Green, and Vuilleumier (2005) have shown the greater importance of LSF information than HSF information for quickly detecting the orientation of probe targets, constituted of either horizontal or vertical light-gray rectangular bars, after priming with LSF fearful EFs. Bocanegra and Zeelenberg (2009) have shown that rapid presentation of fearful faces is sufficient to enhance sensitivity for the orientation of subsequently presented LSF Gabor patches while diminishing orientation sensitivity for HSF Gabor patches. In this article, we directly address the question of the importance of LSF information for detecting a threat in the visual environment.

Neuroimaging studies on the detection of fearful stimuli (Alorda et al., 2007; Vuilleumier et al., 2003) have so far failed to obtain evidence at a behavioral level that LSF channels play a critical role in determining whether information is dangerous or neutral. This could be due to the fact that participants did not experience fear per se in these studies. In other words, these studies used paradigms in which the failure to rapidly recognize danger was not associated with an objective aversive consequence. In the present study, we tested the effects that varying spatial-frequency channels played in the detection of visual stimuli containing threat-related information in a context of fear expectation. Fear expectation was operationalized as the occurrence of an aversive auditory stimulus contingent on the failure to rapidly detect visual danger.

The aim of our study was to provide behavioral evidence about the involvement of different spatial-frequency channels in the fast detection of visual danger in association with an objectively threatening aversive event. More precisely, we investigated the role of LSF and HSF information in behavioral responses to pictures of dangerous and neutral stimuli from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2005). The aim of this experiment was not to investigate the simple visual detection of danger across a wide range of stimuli, but to investigate the associative process, at different scales, between auditory threats and visual cues. This investigation consisted of a visual danger-detection task conducted with a limited but carefully controlled pattern of neutral versus dangerous stimuli. Note that we do not make any categorical assumptions about the neural stream involved in visual detection of a threat. LSF information spreads to a wide range of cortical areas, including frontoparietal cortical regions (Bar, 2004; Bullier, 2001), temporal cortex regions (Livingstone & Hubel, 1988), and subcortical tegato-pulvinar regions (Schiller, Malpeli, & Schein, 1979).

On the basis of the aforementioned literature, we assumed that LSF stimuli could be sufficient for fast detection of visual danger under threat pressure. This behavioral effect should occur despite the drastic loss of information in LSF channels compared with BSF channels or HSF channels. Indeed, LSF stimuli should be more difficult to recognize because of their

![Fig. 1. Examples of spatial-frequency channels. The top image is presented in unfiltered broad spatial frequencies. The bottom images, starting on the left, show the same stimulus presented in filtered spatial-frequency bands increasing from low to high (cpi = cycles per image).](Image)
rough level of detail, compared with BSF or HSF stimuli; this difficulty is generally observed in the literature for nonemotional stimuli that are perceived for more than 50 ms (Parker, Lishman, & Hughes, 1997; Schyns & Oliva, 1994). Therefore, because of the reversion of the coarse-to-fine bias for supraliminal perception range, we predicted a perceptual bias toward better recognition of HSF stimuli compared with LSF stimuli. However, we also predicted that when there is an expectation of punishment, the global loss of performance for LSF stimuli compared with HSF stimuli should be reduced for threat-related stimuli, relative to neutral stimuli, and this should be especially true in the case of living dangerous stimuli.

Method

Participants

Thirty-four undergraduate students (28 women and 6 men) of the Université Blaise Pascal at Clermont-Ferrand, France, participated to fulfill a course requirement. All had corrected-to-normal vision and audition. The mean age of the participants was 21.5 years ($SD = 4.98$).

Stimuli and material

Participants were tested individually in a shielded room at a distance of 90 cm from a computer connected to a 17-in. CRT monitor (ViewSonic, Walnut, CA). The experimental events were controlled by E-Prime software (Version 1.4.1; Psychology Software Tools, Inc., Pittsburgh, PA). Sounds were transmitted through Sennheiser headphones (Old Lyme, CT).

The stimuli consisted of pictures chosen from the IAPS according to their arousal and valence qualities. There were 10 pictures depicting a potential danger (e.g., tornado, snake) and 10 pictures depicting a neutral stimulus (e.g., clock, cup, rabbit, poppies). In each category, half of the pictures depicted a living stimulus, and half of the pictures depicted a nonliving stimulus. We used a limited but highly controlled pattern of stimuli in order to avoid fatiguing participants. BSF stimuli were unfiltered, and LSF and HSF stimuli were created by filtering the original images with MATLAB software (MathWorks, Natick, MA) in five bands: less than 8 cycles per image (cpi), 8–16 cpi, 16–32 cpi, 32–64 cpi, and more than 64 cpi. The goal was to provide a tuning curve covering the entire spatial-frequency spectrum. Average luminance was normalized, and average contrast energy did not differ significantly between dangerous and neutral stimuli.

To ensure that the dangerous stimuli were more arousing and more aversive than the neutral stimuli, we conducted an analysis of variance (ANOVA) on the arousal and the valence of the stimuli (Lang et al., 2005), with visual danger (dangerous vs. neutral) and stimulus category (living vs. nonliving) as between-subject variables. The data analysis showed that the mean arousal value was significantly higher for the dangerous stimuli ($M = 5.90, SD = 1.03$) than for the neutral stimuli ($M = 4.03, SD = 1.32$), $F(1, 16) = 14.4, p < .003$. The main effect of stimulus category on arousal was not significant, $F(1, 16) = 2.67$. The mean valence was significantly lower for the dangerous stimuli ($M = 3.62, SD = 0.698$) than for the neutral stimuli ($M = 6.85, SD = 1.22$), $F(1, 16) = 54.3, p < .001$. No difference was observed for valence as a function of stimulus category, $F(1, 16) < 1, n.s.$ In addition, there was no two-way interaction (Visual Danger × Stimulus Category), either for arousal, $F(1, 16) = 1.98, n.s.$, or for valence, $F(1, 16) = 2.24, n.s.$

Procedure

We used a paradigm of expectation of punishment, in which a threat cue (a colored triangle) predictive of a conditional aversive stimulus was presented prior to the IAPS stimuli. As illustrated in Figure 2, we used a response-window procedure (Greenwald, Draine, & Abrams, 1996), in which participants had to detect a potential danger in a picture as quickly as possible (<600 ms).

The experiment began with a familiarization session, in which all 20 BSF stimuli were presented. The familiarization session was immediately followed by a training session, during which each participant was presented with each of the 20 images in each of the six spatial-frequencies ranges. Each stimulus picture was preceded by a fixation cross with a random duration between 500 and 1,000 ms; the cross was followed by a brief presentation (30 ms) of either a punishment triangle (PT) or a no-punishment triangle (NPT). Each type of triangle was associated with a color (blue or red), and the punishment-color association was counterbalanced across subjects. The blue or red triangles were postmasked by a violet triangle (250 ms).

The target stimulus was presented in the center of the screen after the violet triangle. Participants pressed the “D” key or the “F” key on an AZERTY keyboard to indicate whether the stimulus was dangerous or nondangerous. Assignment of key to response was counterbalanced across subjects. Then, a blank screen was displayed for a random duration between 500 and 1,000 ms. This screen was followed by feedback, which consisted of either the word “correct,” for correct and fast responses (<600 ms), or “wrong,” for incorrect or slow responses (>600 ms). In trials during which PT appeared, a false response or a slow response to a dangerous stimulus was followed by an aversive white noise (50 ms, 90 dB). In trials during which the NPT appeared, all responses were followed by a 50-ms noise at 0 dB.

After completing the training session, participants rested for 5 min. Then they completed an extinction session. The extinction session was identical to the training session, except without the 90-dB white noises.

The order of stimulus presentation was randomized in both the training and the extinction sessions. The two sessions together consisted of a total of 240 trials, during which all
stimuli were presented twice (once during the training session and once during the extinction session). The BSF stimuli were presented three times: once during the training session, once during the extinction session, and once during the familiarization session.

**Results**

**Reaction times**

We conducted an ANOVA on reaction times (RTs), with experimental session (training vs. extinction), spatial frequency (BSF, < 8 cpi, 8–16 cpi, 16–32 cpi, 32–64 cpi, and > 64 cpi), visual danger (dangerous pictures vs. neutral pictures), and stimulus category (living vs. nonliving) as within-subjects variables. The ANOVA showed a significant main effect for visual danger, $F(1, 30) = 9.33, MSE = 22,225, p < .01$; stimulus category, $F(1, 30) = 34.62, MSE = 4,666, p < .001$; and spatial frequency, $F(5, 150) = 34.6, MSE = 2,620, p < .001$. (Fig. 3 shows these main effects; however, we do not report details of them because our focus of interest is the interaction of these factors).

The top panel of Figure 3 shows the mean RTs for dangerous pictures and neutral pictures as a function of spatial-frequency channel, for both the living and nonliving categories. The first significant two-way interaction was between visual danger and stimulus category; it indicated that the difference in the RTs between neutral pictures and dangerous pictures was higher for the living category (577.5 ms vs. 535.6 ms) than for the nonliving category (579.4 ms vs. 574.8 ms), $F(1, 30) = 19.4, MSE = 6,895, p < .001$. The second significant two-way interaction was between visual danger and spatial frequency, $F(5, 150) = 4.72, MSE = 2,732, p < .001$; the difference between neutral pictures and dangerous pictures decreased monotonically from LSF to HSF channels. Recognition tended to be faster for dangerous pictures than for neutral pictures in LSF channels, whereas performance tended to be similar for dangerous pictures and neutral pictures in HSF channels.

Pair-wise comparisons confirmed that this latter effect was significant for four channels—BSF: $F(1, 30) = 11.5, MSE = 6,868,
Fig. 3. Participants’ mean reaction times (upper panel) and mean accuracy (proportion of correct responses; lower panel) in detecting whether stimuli were dangerous or neutral, as a function of the spatial-frequency information presented. Each image was presented in a broad-spatial-frequency (BSF) version and in five spatial-frequency channels (measured in cycles per image, cpi). Results are shown separately for stimuli from the living and nonliving categories. Error bars represent 95% confidence intervals.
dangerous pictures and neutral pictures showed that dangerous pictures were, or tended to be, better recognized than neutral pictures in three spatial-frequencies—BSF: \( F(1, 32) = 3.76, \ MSE = 0.062, p = .06; \) less than 8 cpi: \( F(1, 32) = 18.9, MSE = 0.058, p < .001; \) and 8–16 cpi stimuli: \( F(1, 32) = 4.48, MSE = 0.079, p < .05 \) (all other spatial frequencies; \( ps > .10 \)).

As in the RT data, a significant interaction was observed between stimulus category and visual danger, \( F(1, 32) = 6, MSE = 0.07, p < .05 \). This interaction reflected a higher difference in accuracy between neutral pictures and dangerous pictures for the living stimuli (.62 vs. .71) compared with the nonliving stimuli (.63 vs. .6). All other interactions were not significant.

Moreover, as shown in Figure 4, we also observed a significant main effect of experimental session, \( F(1, 32) = 47.13, MSE = 0.16, p = .001 \); accuracy was greater during the extinction session (\( M = .7, SE = .14 \)) compared with the training session (\( M = .57, SE = .19 \)).

**Discussion**

The basic goal of our study was to investigate the possibility that coarse scales might represent a sufficient signal for fast detection of visual danger when that danger is associated with an expectation of punishment. Our results showed that the LSF component of visual stimuli was a very crude source of information resulting in overall lower performance (decreased accuracy and increased RTs) when compared with the HSF component of visual stimuli; this was especially the case for the neutral stimuli and during the training session. On the basis of our visual presentation onset (400–600 ms, on average), this finding is consistent with the literature on the coarse-to-fine bias showing a superiority of HSF information (at the level of accuracy and RTs) for supraliminal presentation durations (Hughes et al., 1996; Parker et al., 1997; Schyns & Oliva, 1994).

However, despite the roughness of LSF stimuli, participants were able to reliably use LSF information after a brief training session to quickly identify the dangerous stimuli in the visual environment. Moreover, the RT results suggest that LSF channels could be as efficient as finer but slower HSF channels (provided only by parvocellular layers) for fast detection of visual danger. This experiment confirms that, after fear conditioning, the difference in accuracy and RTs between the dangerous stimuli and the neutral stimuli decreases monotonically from LSF channels to HSF channels (Fig. 4). These results suggest that, under supraliminal perception durations, viewing dangerous stimuli in association with fear expectations is able to disrupt the well-established reversion of the coarse-to-fine bias. Normally, this reversion results in a better and faster recognition of HSF stimuli than LSF stimuli (Hughes et al., 1996; Schyns & Oliva, 1994). Therefore, fear-associative processes, possibly occurring at the level of the amygdala, seem to be sufficient to suppress this powerful perceptual bias.
Perceptual Processes of Threat Detection

Training Session

Extinction Session

Fig. 4. Participants’ mean reaction times (upper panel) and accuracy (proportion of correct responses; lower panel) in detecting whether stimuli were dangerous or neutral, as a function of the spatial-frequency information presented. Each image was presented in a broad-spatial-frequency (BSF) version and in five spatial-frequency channels (measured in cycles per image, cpi). Results are shown separately for the training and extinction sessions. Error bars represent 95% confidence intervals.
Moreover, the results of the present study show that this effect is specific to living stimuli. This is an original and important finding, and further research is required to understand its exact nature. We propose three hypotheses. The first is an evolutionary hypothesis in line with Öhman and Soares (1993). Predators often constitute a more direct and immediate danger for the organism than nonliving objects do. Therefore, a fear system dedicated to faster recognition of threatening stimuli that are animate, living entities may increase chances of survival. Second, an alternative hypothesis lies at a perceptual level. Organisms may require fewer perceptual details to recognize living dangers than nonliving dangers. For example, predators often have the same perceptual structure (elongated body, two eyes, open mouth, etc.) and could be easier to identify on the basis of LSF components only. A third hypothesis lies at a cognitive level. Dangerous situations involving nonliving objects could require a more complex cognitive analysis, probably involving high-level cortical areas, for the viewer to understand the dangerous content of the situation (e.g., a tornado or a harmful device). If the amygdala plays a crucial role in fear conditioning, this subcortical structure alone might not be sufficient for a complex understanding of nonliving threats, and the involvement of high-level associative areas could be necessary to access the dangerous meaning of these stimuli.

Conclusion

Previous studies involving neuroimaging (Alorda et al., 2007; Pourtois et al., 2005; Vuilleumier et al., 2003), computational modeling (Mermillod, Vuilleumier, et al., 2009; Mermillod et al., in press), and behavior (Bocanegra & Zeelenberg, 2009; Holmes, Green, & Vuilleumier, 2005) have suggested that LSF information might be sufficient for expedient detection of visual danger in the perceptual environment. This result should occur despite a strong perceptual bias toward better and faster recognition of HSF stimuli than LSF stimuli under supraliminal presentation durations when the stimulus is not associated with a fearful event. Our study aimed to test this hypothesis, and it therefore fills a gap between the literature on fear processing (Delgado et al., 2006; Löw, Lang, Smith, & Bradley, 2008; Vermeulen, Godefroid, & Mermillod, 2009) and psychophysical studies on visual processes (Parker et al., 1997; Schyns & Oliva, 1994).

To our knowledge, these findings are the first behavioral evidence that coarse scales can produce fast and reliable detection of visual threats under supraliminal perception durations. Further experiments are required to determine the visual pathways responsible for this effect (magnocellular vs. parvocellular layers). Moreover, we found that this effect was limited to basic visual threats (i.e., living categories such as predators). This property could be based on the phylogenetic development of the human species. However, we also propose that dangerous situations involving nonliving objects are perceptually and conceptually more complex to understand than dangerous situations involving living organisms; therefore, they could require the use of more complex cortical associative areas. Further neuroimaging and behavioral experiments are required to test these alternative hypotheses.

Acknowledgments

We thank Claire Pottage and Amy Bland for the English revision.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This work was supported by Grant ANR-06-BLAN-0360-01 from the French National Research Agency (ANR) to Martial Mermillod and Sylvie Droit-Volet, and by Grants ANR-06-CORP-019 and ANR-BLAN08-1_353820 from the ANR to Martial Mermillod.

Notes

1. The IAPS identification numbers of the chosen stimuli were 1050, 1300, 1321, 1450, 1590, 1710, 1930, 2080, 5030, 5891, 5971, 6610, 6940, 7054, 7175, 7200, 7211, 8502, 9635.1, and 9635.2.
2. The effect of PT color was not significant for RTs, $F(1, 30) = 0.27$, $MSE = 231,894, p = .61$, or for accuracy, $F(1, 32) = 0.63$, $MSE = 1.67, p = .36$. Therefore, this factor was removed from the analysis.

References


