

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/331109585>

The danger of interpreting detection differences between image categories: A brief comment on "Mind the snake: Fear detection relies on low spatial frequencies" (Gomes, Soares, Sil...

Article in *Emotion* · February 2019

DOI: 10.1037/emo0000550

CITATION

1

READS

71

3 authors, including:



Surya Gayet

Radboud University

42 PUBLICATIONS 412 CITATIONS

[SEE PROFILE](#)



Timo Stein

University of Amsterdam

70 PUBLICATIONS 1,890 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Top-down modulations of conscious access [View project](#)



Visual working memory enhances neural representations of matching visual input [View project](#)

The Danger of Interpreting Detection Differences between Image Categories: A Brief
Comment on *Mind the Snake: Fear Detection Relies on Low Spatial Frequencies* (Gomes,
Soares, Silva, & Silva, 2018).

Surya Gayet¹, Timo Stein², & Marius V. Peelen¹

¹ Donders Institute for Brain, Cognition and Behaviour, Radboud University

Full address: Montessorilaan 3, 6525 HR Nijmegen, The Netherlands

² Amsterdam Brain and Cognition, University of Amsterdam

Full address: Nieuwe Achtergracht 129-B, 1018 WS Amsterdam, The Netherlands

Experiment files, raw data, and analyses

All experiment files, raw data, and analysis files used in the experiment described in this manuscript can be retrieved [here](#) (via the Open Science Framework).

Abstract

Using breaking continuous flash suppression (b-CFS; a perceptual suppression technique), Gomes and colleagues (2018) showed that human observers have an advantage in detecting images of snakes (constituting an evolutionarily old threat) over birds. In their study, images of snakes and birds were filtered to contain either coarse-scale or fine-grained information. The preferential detection of snakes relied on coarse-scale (rather than fine-grained) information, which was taken as support for the existence of an evolutionarily old subcortical pathway dedicated to snake detection. Here, we raise the concern that images of snakes and birds inherently differ in their visual characteristics, which can strongly affect detection times in b-CFS. Images of snakes, for instance, have a larger perimeter-to-surface ratio than images of birds. Importantly, these visual characteristics are not snake-specific, as they are shared with many non-threatening object categories. To illustrate this point, we compared detection times between images of bicycles and cars; non-threatening image categories that differ in visual characteristics but for which detection is unlikely to capitalize on an evolutionarily old dedicated subcortical pathway. Observers exhibited an advantage for detecting bicycles over cars. Mirroring the snake-bird differences reported in Gomes et al., this advantage was driven by the coarse-scale (rather than fine-grained) information in the images. Hence, differences in visual characteristics between two non-threatening, semantically matched stimulus categories suffice to produce the exact same pattern of findings as observed with snakes versus birds. We conclude that spatial frequency-specific detection differences in b-CFS cannot be unequivocally attributed to differences in processing pathways.

Keywords

Fear processing; Threat; Continuous flash suppression; Dual route model; Low road

The Danger of Interpreting Detection Differences between Image Categories

A key role of our visual system, from an evolutionary stance, is to detect threatening stimuli within our environment. Accordingly, there is ample evidence that stimuli signaling threat are preferentially processed over non-threatening stimuli (for reviews, see Mather, & Sutherland, 2011; Yiend, 2010). There is less consensus, however, regarding the neural mechanisms that enable this preferential processing of threatening stimuli. A long-standing debate in this regard concerns the role of a subcortical pathway in the fast and automatic detection of potentially threatening stimuli (Pessoa & Adolphs, 2010). This pathway connects the retina to the amygdala via the superior colliculus and the pulvinar, bypassing the cortex (LeDoux, 1998). Many studies investigating the involvement of this subcortical pathway in detecting threatening (or fear-inducing) stimuli capitalize on the fact that it transmits coarse-scale (i.e., low-spatial frequency) information, as opposed to the fine-grained information that is processed in the visual cortex (Vuilleumier, Armony, Driver, & Dolan, 2003; Pessoa & Adolphs, 2010).

Following this approach, Gomes, Soares, Silva, and Silva (2018) recently demonstrated that the preferential detection of snakes (threatening stimuli) over birds is driven mostly by low spatial frequency information. They presented images of snakes and birds to one eye, and suppressed these images from consciousness by presenting a high contrast dynamic pattern to the other eye (continuous flash suppression; Tsuchiya, & Koch, 2005). The time it takes participants to localize these initially suppressed images provides a measure of detectability, which can be compared between images of birds and snakes (i.e., the so-called breaking continuous flash suppression paradigm (b-CFS); Jiang, Costello, & He, 2007; Stein, Hebart, & Sterzer, 2011; for a review, see Gayet, Van der Stigchel, & Paffen, 2014). In line with their earlier work, the authors observed faster detection of snake images than bird images, thus showing preferential detection of snake images. Crucially, the spatial

frequency content of the images was also manipulated, such that images comprised only low spatial frequency (i.e., coarse scale) information, only high spatial frequency (fine-grained) information, or both (i.e., unfiltered, broadband images). The preferential detection of snakes over birds was observed for unfiltered images and for images containing only low spatial frequencies, but not for images containing only high spatial frequencies (see Figure 1A for a summary of their results). These results were interpreted as evidence for a specialized subcortical pathway dedicated to snake detection based on coarse-scale information.

In this Brief Commentary, we argue that the finding of a detection difference between different images in the b-CFS paradigm cannot be interpreted as evidence for or against the involvement of subcortical processing, even when these differences are driven selectively by low spatial frequency content. Please note that our commentary is neutral about the existence of a putative subcortical pathway for preferential processing of threatening stimuli. Rather, we present a more general limitation in interpreting detection times between different image categories in the b-CFS paradigm; one that similarly applies to other studies investigating the preferential detection of emotionally laden over emotionally neutral stimuli (e.g., Yang, Zald, & Blake, 2007; Sklar, Levy, Goldstein, Mandel, Maril, & Hassin, 2012; Gomes, Silva, Silva, & Soares, 2017).

It is well-known that low-level visual differences between image categories can bring about differences in detection times, and particularly so in the b-CFS paradigm (Tsuchiya, & Koch, 2005; Yang, & Blake, 2012; for an overview, see Gayet, et al., 2014). Accordingly, when previous studies included additional controls to assess the contribution of visual confounds to the difference in detection times between stimulus categories, the difference in detection times was often found to be at least partly driven by these visual confounds (e.g., Stein, Awad, Gayet, & Peelen, in press; Tsuchiya, Moradi, Felsen, Yamazaki, & Adolphs, 2009; Stein, Peelen, & Sterzer, 2011; Stein, & Sterzer, 2012; Chen, & Yeh, 2012; Gray,

Adams, Hedger, Newton, & Garner, 2013). To account for such differences in visual characteristics, Gomes and colleagues (2018) a priori equated the luminance and contrast between the two image categories, and asserted a posteriori that there was no significant difference in spatial frequency energy between the two image categories. Nevertheless, images of snakes and birds inherently differ in terms of visual characteristics. As such, it remains unclear whether the faster detection of snakes reflects a processing advantage for threatening stimuli, or a processing advantage for threat-unrelated visual properties that happen to differ between snakes and birds (e.g., curvature, elongation, or perimeter-to-surface ratio).

The addition of different spatial filtering conditions might appear to circumvent this problem, under the assumption that non-specific visual differences between images of snakes and birds remain constant across spatial filtering conditions. We make the case, however, that the addition of different spatial filtering conditions simply introduces new image categories, which differ idiosyncratically in their visual characteristics. That is, considering that the snake and bird image categories are comprised of different visual characteristics, these image categories will also be differently affected by spatial frequency filtering. For example, images with a higher perimeter-to-surface ratio (like snakes) will retain more troughs and peaks once blurred (i.e., low-pass filtered), whereas images with lower perimeter-to-surface ratio (like birds) will mostly comprise a single peak, with a surrounding trough (i.e., a blob). When these same image categories (of high and low perimeter-to-surface ratio) are high-pass filtered, both perimeter edges and within-stimulus edges will be retained, so that the images become more similar between categories (Supplementary Figure S3). Note that this is only one of many possible differences in visual characteristics between stimulus categories and spatial filtering conditions (other examples could include differences in elongation, center of mass, number of gaps, object curvature, etc.). Importantly, these differences are not

accounted for by equating variables such as contrast and luminance values, but they could similarly drive differences in detection times.

To illustrate this point, we ran an experiment comparing detection times between (luminance and contrast-equated) images of bicycles and cars; two image categories that arguably do not differ in terms of the fear that they induce, and for which detection surely does not rely on an evolutionarily old dedicated subcortical pathway. These image categories do differ, however, in the visual characteristics that they comprise, such as the perimeter-to-surface ratio (Supplementary Materials 3.4). Following the exact procedure of Gomes and colleagues (2018), we compared detection times to broadband (BSF), low-pass (LSF), and high-pass (HSF) images of bicycles and cars (see Figure 1B; the full Methods and Results can be found in the Supplementary Materials). For broadband (i.e., unfiltered) images, initially suppressed bicycles were detected 1.15 seconds faster than cars, $t(12) = 9.7$, $p < 0.001$, $d = 2.7$. For images containing only low spatial frequency content, bicycles were detected 1.19 seconds faster than cars, $t(12) = 7.7$, $p < 0.001$, $d = 2.1$. For images containing only high spatial frequency content, however, there was no difference in detection time between bicycles and cars (detection of cars was numerically 0.05 seconds faster), $t(12) = 0.2$, $p > 0.8$, $d < 0.1$. Note that absolute response times differed between filtering conditions (as also observed by Gomes and colleagues) because filtering inherently removes visual information from the image, thereby impacting absolute detection times. This is irrelevant to the current findings, in which we compare detection times between image categories, within filtering conditions. In sum, akin to the previously reported difference between snakes and birds, the detection difference between bicycles and cars (as observed with unfiltered images) appears to be fully driven by the low spatial frequency information in the images. In contrast to snakes, however, preferential detection of bicycles yields no clear evolutionary benefit.

Admittedly, the fact that the advantage of (A) detecting images of snakes relative to images of birds, and the advantage of (B) detecting images of bicycles relative to images of cars both rely on the low spatial frequency content of these image categories does not prove that both detection differences are underpinned by the same neural mechanism. As such, it is possible that a subcortical pathway to the amygdala causes the advantage for snake detection, and that a different (cortical) mechanism causes the advantage for bicycle detection. Considering that humans most probably did not evolve a specialized subcortical route for bicycle detection, our data at least demonstrate that a detection advantage in the low spatial frequency domain for some image category does not require the presence of a specialized subcortical route for processing said image category.

Given that the snake (or bicycle) detection advantage is only found for low spatial frequency content, does that not imply an involvement of the subcortical pathway? No, the reliance of an effect on low spatial frequency information does not provide evidence for subcortical processing. While cortical involvement is indeed required for processing detailed, high spatial frequency information (e.g., Stein, Seymour, Hebart, & Sterzer, 2014), this logic cannot be reversed, as the visual cortex processes a wide range of spatial frequencies. More generally, there is little evidence that stimuli composed of high and low spatial frequencies are effective in isolating magnocellular and parvocellular visual processing (Skottun, 2015).

One could argue that a specialized subcortical pathway that evolved for preferential snake detection incidentally causes preferential detection of all image categories that share visual similarities with snakes, including bicycles (for a similar argument in the comparison between manipulable and non-manipulable objects, see Almeida, Mahon, Zapater-Rabero, Dziuba, Cabaço, Marques, & Caramazza, 2014; Almeida, Fintzi, & Mahon, 2013). It is difficult to establish empirically whether the advantage for detecting bicycles is caused by a specialized pathway that was initially dedicated to detecting snakes. Nonetheless, if a

specialized subcortical pathway exists that favors snakes, and thereby generalizes to bicycles, we can conclude that this pathway capitalizes on *shape* detection (as argued here) rather than *fear* detection.

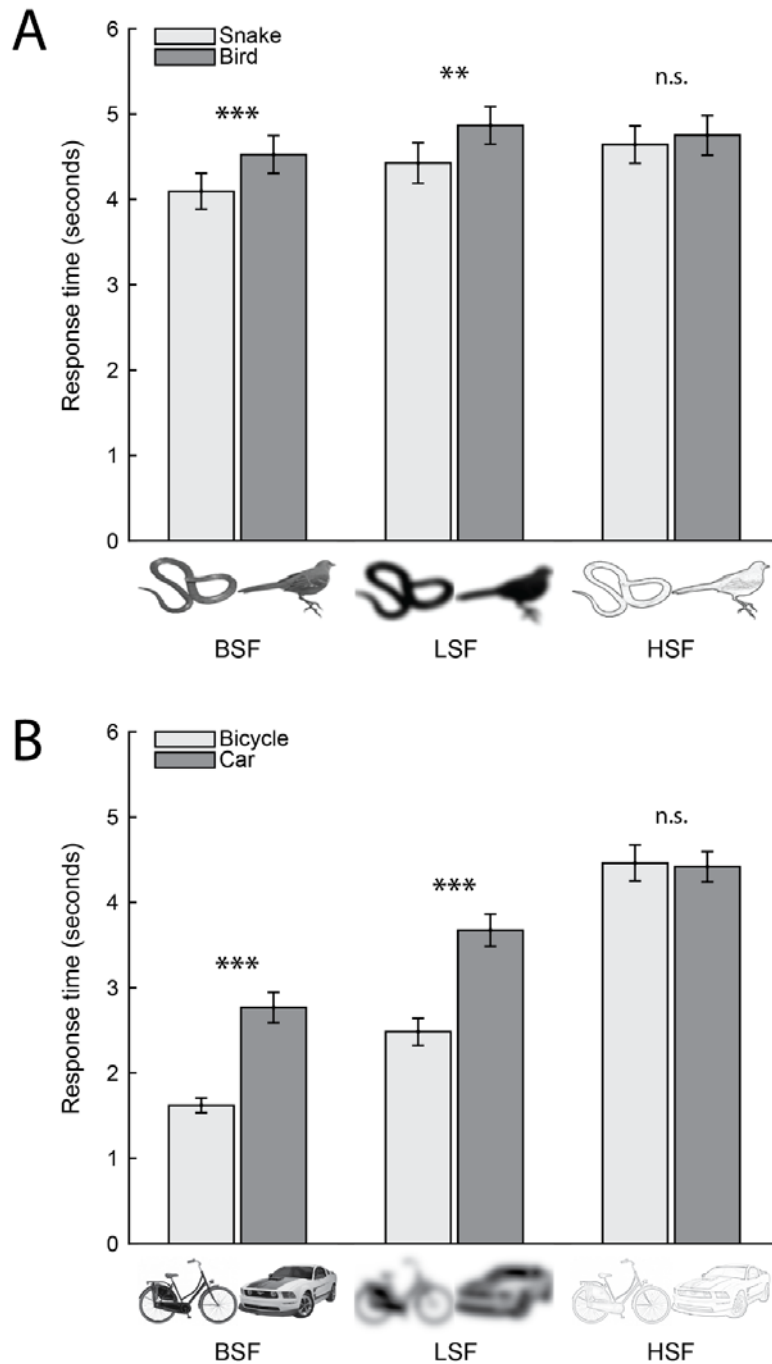


Figure 1. Panel A depicts the results of Gomes and colleagues' (2018) Experiment 1A, and is reconstructed on the basis of the means and standard deviations reported in their manuscript. Panel B depicts the results of our experiment (see the Supplemental Materials for a complete description of the Methods and Results). The bars depict average response times to initially suppressed snake and bird stimuli (panel A) or bicycle and car stimuli (panel B), which were either unfiltered (BSF), low-pass (LSF), or high-pass (HSF) filtered. Error bars depict the standard error of the mean. ** $p < 0.01$, *** $p < 0.001$

Discussion

It is well-established that certain types of threatening stimuli are preferentially processed over non-threatening stimuli (e.g., Öhman, & Soares, 1994; Schmack, Burk, Haynes, & Sterzer, 2016). While it is true that differences in visual characteristics between threatening and non-threatening stimuli could allow for the employment of an alternative (either cortical or subcortical) processing pathway that favors threatening over non-threatening stimuli, such conclusions cannot be inferred from the current experimental approach. This follows from the fact that comparing differences in detection times between threatening and non-threatening image categories does not allow for disentangling the causal role of ‘fear’ from that of non-specific visual characteristics that constitute the threatening stimuli. The addition of different spatial filtering conditions does not circumvent this problem, as it introduces new image categories with new image confounds and is thus equally uninformative vis-à-vis the role of fear in snake detection.

Nonetheless, we do not argue against the usage of the b-CFS paradigm to investigate the preferential detection of emotionally laden stimuli per se. The b-CFS paradigm can be valuable as long as the difference in emotional content between image categories is not confounded with a difference in visual characteristics. One promising approach to achieve this, is by capitalizing on image manipulations like inversion and/or polarity reversal, which partly disrupt the extraction of meaning while preserving most visual characteristics (Rock, 1974). Because these manipulation selectively disrupt the extraction of meaning, the difference in detection time between normal and (say) polarity inverted versions of an image is proportional to the contribution of non-visual factors to image detection (for a similar approach, see Hedger, Adams, & Garner, 2015; Stein et al., in press). Image inversion and polarity reversal are of particular interest for the present purpose, because they are orthogonal to spatial filtering (unlike the main differences in visual characteristics between snakes and

birds). Hence, if the advantage for detecting normal over polarity inverted snakes increases after applying low-spatial compared to high-spatial frequency filtering, one can conclude that there is something about the meaning (rather than the visual characteristics) of the snake that is preferentially extracted from low spatial frequency information. Another way to isolate the influence of emotional content when comparing detection times between image categories is to capitalize on between-subject differences. For instance, Schmack et al. (2016) showed that the preferential detection of spiders relative to flowers depends on the degree of spider phobia. Whereas the detection difference between images of spiders and flowers itself is confounded by differences in visual characteristics, these visual characteristics were identical for all participants despite their varying degrees of spider phobia. Hence, the authors could conclude that fear of spiders drives the preferential detection of spider images. Along similar lines, Tsuchiya et al. (2009) tested whether the advantage for detecting fearful over happy faces, as observed in a matched control group, was preserved in a patient with bilateral amygdala lesions, thus isolating the influence of visual characteristics on detection times. Finally, one can also consider Pavlovian conditioning as a tool to investigate the preferential detection of emotionally laden stimuli within the b-CFS paradigm. For instance, it was shown using classical fear conditioning that initially neutral stimuli are preferentially detected when they signal threat (i.e., when they were previously associated with an aversive event) compared to when they do not (Gayet, Paffen, Belopolsky, Theeuwes, & Van der Stigchel, 2016).

Conclusion

Because the b-CFS paradigm is extremely sensitive to differences in visual characteristics, it poses a challenge when investigating differences in non-visual attributes (such as threat or fear) between different image categories. Although the comparison of preferential detection of one image category over the other *between* spatial filtering conditions appears to circumvent this issue, this is not the case because spatial filtering (i.e., HSF and LSF) produces idiosyncratic differences in visual characteristics between image categories. As a result of this, the addition of spatial filtering conditions results merely in a larger number of image categories, that are all confounded with differences in visual characteristics. This argument is supported by our data, which show that strong differences in detection times between image categories (and across spatial filtering conditions) exist even for non-threatening, semantically-matched stimuli.

References

- Almeida, J., Fintzi, A. R., & Mahon, B. Z. (2013). Tool manipulation knowledge is retrieved by way of the ventral visual object processing pathway. *Cortex*, 49(9), 2334-2344.
- Almeida, J., Mahon, B. Z., Zapater-Rabero, V., Dziuba, A., Cabaco, T., Marques, J. F., & Caramazza, A. (2014). Grasping with the eyes: the role of elongation in visual recognition of manipulable objects. *Cognitive, Affective & Behavioral Neuroscience*, 14(1), 319-335.
- Chen, Y. C., & Yeh, S. L. (2012). Look into my eyes and I will see you: Unconscious processing of human gaze. *Consciousness and Cognition*, 21(4), 1703-1710.
- Gayet, S., Paffen, C. L. E., Belopolsky, A. V., Theeuwes, J., & Van der Stigchel, S. (2016). Visual input signaling threat gains preferential access to awareness in a breaking continuous flash suppression paradigm. *Cognition*, 149, 77-83.

- Gayet, S., Van der Stigchel, S., & Paffen, C. L. E. (2014). Breaking continuous flash suppression: competing for consciousness on the pre-semantic battlefield. *Frontiers in Psychology*, 5, 460.
- Gomes, N., Silva, S., Silva, C. F., & Soares, S. C. (2017). Beware the serpent: the advantage of ecologically-relevant stimuli in accessing visual awareness. *Evolution and Human Behavior*, 38(2), 227-234.
- Gomes, N., Soares, S. C., Silva, S., & Silva, C. F. (2018). Mind the snake: Fear detection relies on low spatial frequencies. *Emotion*, 18(6), 886-895.
- Gray, K. L., Adams, W. J., Hedger, N., Newton, K. E., & Garner, M. (2013). Faces and awareness: Low-level, not emotional factors determine perceptual dominance. *Emotion*, 13(3), 537-544.
- Hedger, N., Adams, W. J., & Garner, M. (2015). Fearful faces have a sensory advantage in the competition for awareness. *Journal of Experimental Psychology: Human Perception and Performance*, 41, 1748-1757.
- Jiang, Y., Costello, P., & He, S. (2007). Processing of invisible stimuli: Advantage of upright faces and recognizable words in overcoming interocular suppression. *Psychological Science*, 18(4), 349-355.
- LeDoux, J. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. New York City, NY: Simon and Schuster.
- Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science*, 6(2), 114-133.
- Öhman, A., & Soares, J. J. (1994). "Unconscious anxiety": phobic responses to masked stimuli. *Journal of abnormal psychology*, 103(2), 231.

- Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. *Nature Reviews Neuroscience*, *11*(11), 773.
- Rock, I. (1974). The perception of disoriented figures. *Scientific American*, *230*(1), 78-86.
- Schmack, K., Burk, J., Haynes, J. D., & Sterzer, P. (2016). Predicting subjective affective salience from cortical responses to invisible object stimuli. *Cerebral Cortex*, *26*(8), 3453-3460.
- Sklar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., & Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. *Proceedings of the National Academy of Sciences*, *109*(48), 19614-19619.
- Skottun, B. C. (2015). On the use of spatial frequency to isolate contributions from the magnocellular and parvocellular systems and the dorsal and ventral cortical streams. *Neuroscience & Biobehavioral Reviews*, *56*, 266-275.
- Stein, T., Awad, D., Gayet, S., & Peelen, M. V. (in press). Unconscious processing of facial dominance: The role of low-level factors in access to awareness. *Journal of Experimental Psychology: General*.
- Stein, T., Hebart, M. N., & Sterzer, P. (2011). Breaking continuous flash suppression: a new measure of unconscious processing during interocular suppression?. *Frontiers in Human Neuroscience*, *5*: 167.
- Stein, T., Peelen, M. V., & Sterzer, P. (2011). Adults' awareness of faces follows newborns' looking preferences. *PLoS One*, *6*(12), e29361.
- Stein, T., Seymour, K., Hebart, M. N., & Sterzer, P. (2014). Rapid fear detection relies on high spatial frequencies. *Psychological science*, *25*(2), 566-574.

- Stein, T., & Sterzer, P. (2012). Not just another face in the crowd: detecting emotional schematic faces during continuous flash suppression. *Emotion*, 12(5), 988-996.
- Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, 8(8), 1096-1101.
- Tsuchiya, N., Moradi, F., Felsen, C., Yamazaki, M., & Adolphs, R. (2009). Intact rapid detection of fearful faces in the absence of the amygdala. *Nature Neuroscience*, 12(10), 1224-1225.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, 6(6), 624-631.
- Yang, E., & Blake, R. (2012). Deconstructing continuous flash suppression. *Journal of Vision*, 12(3): 8.
- Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion*, 7(4), 882.
- Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition and Emotion*, 24(1), 3-47.

Supplementary materials

1. Article information

Authors: *Surya Gayet, Timo Stein, and Marius Peelen.*

Main article: *The Danger of Interpreting Detection Differences between Image Categories.*

Resources: *Download all experiment files, raw data, and analysis files [here](#).*

2. Background and research question

A recent study by Gomes, Silva, Silva, and Soares (in press, Emotion) showed that conscious detection of initially suppressed stimuli is faster for images of snakes than for images of birds, using the breaking continuous flash suppression paradigm (b-CFS; for a review, see Gayet, Van der Stigchel, & Paffen, 2014). The authors argue that this is because snakes are an evolutionarily old threat, and therefore detection of snakes is behaviorally relevant. The authors also demonstrate that the prioritized detection of snakes over birds is observed for low-pass filtered images (containing only coarse-scale information) but not for high-pass filtered images (containing only fine-grained information). The authors provide these findings as corroborating evidence that the prioritized detection of snakes relies on a subcortical ‘quick and dirty’ magnocellular pathway to the amygdala (bypassing the visual cortex) dedicated to the detection of snakes.

To investigate whether the pattern of results obtained by Gomes and colleagues (in press) can be caused by differences in visual characteristics between snakes and birds (rather than differences in threat, and the utilization of a dedicated subcortical pathway), we compared conscious detection of bicycles and cars. These image categories were chosen because (1) they obviously differ in terms of visual characteristics (even after equating a number of low-level visual characteristics such as luminance contrast), and (2a) they differed neither in their degree of threat, nor (2b) in their propensity to tap onto a dedicated subcortical pathway.

3. Methods

3.1. Participants and stopping rule.

Participants were recruited from the Radboud University subject pool, and participated for course credits or monetary reward. A total of 15 participants (4 males) with an average age of 22.3 years ($SD = 3.6$) took part in the experiment. A power analysis based on the results of a pilot study comprising only the unfiltered condition (with 10 different participants), revealed that 6 participants would suffice for detecting the response time difference between image categories with 95% power. Data collection for the current study was combined with data collection for an unrelated b-CFS study, which determined the eventual sample size of 15. Ethical approval was provided by the Ethical Committee Social Sciences (ECSW2017-2306-51).

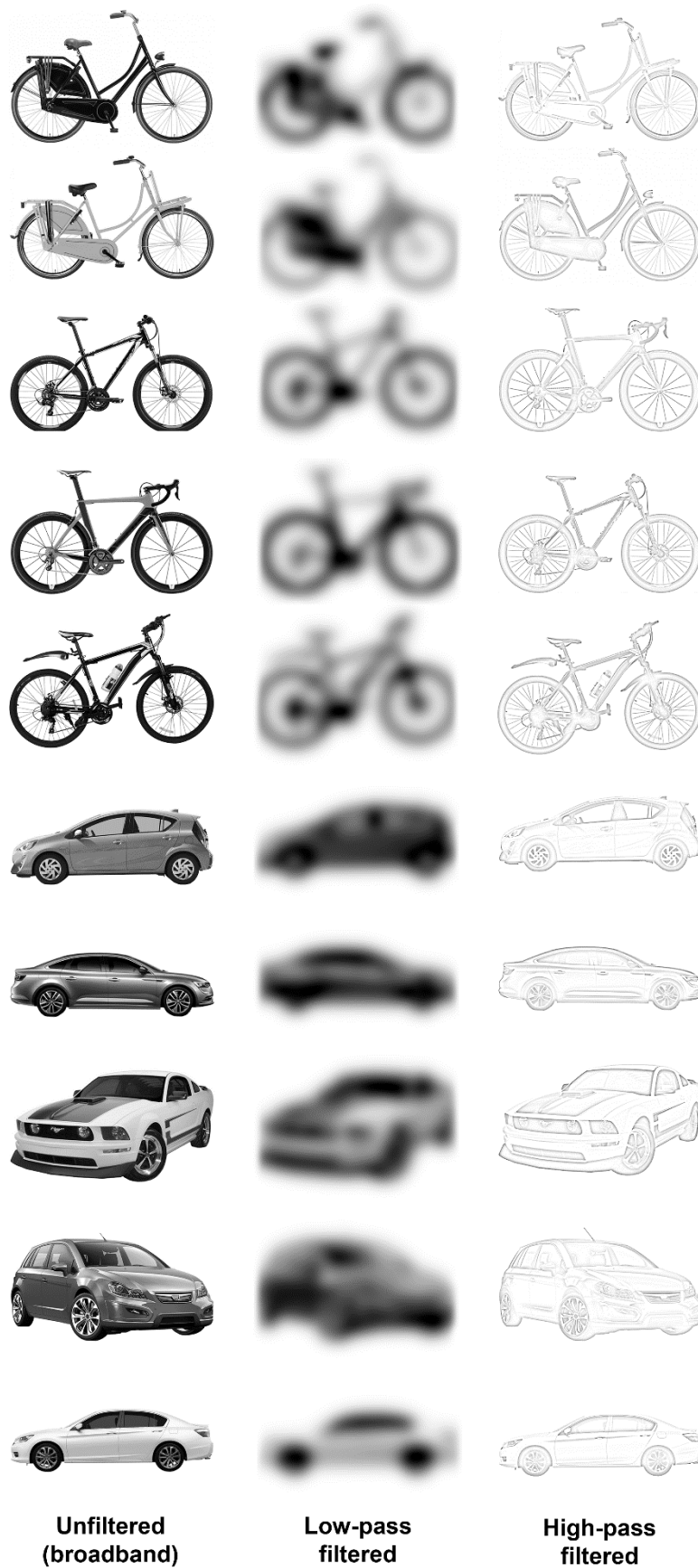


Figure S1. All target stimuli used in the current experiment: two image categories (bicycles and cars) with five exemplars each, and three filtering conditions (unfiltered, low-pass, and high-pass filtered).

3.2. Experimental Design.

The experimental design of our study was identical to that of Gomes and colleagues (in press). The factors-of-interest were manipulated in a 2 image categories (bicycle, or car) by 3 spatial filtering conditions (unfiltered, low-pass, or high-pass spatial filtering). Two other factors (of non-interest) were also manipulated within participant: 4 target locations (top-left, bottom-left, top-right, or bottom-right) by 5 image exemplars (the five different images of cars and bicycles). All factors were fully counterbalanced within-participant, which resulted in 120 unique trials that were presented in a randomized order to each participant.

3.3. Stimuli and apparatus.

Stimulus presentation and data collection were performed on a PC with an Intel Xeon CPU E5-1620 3.7 GHz processor, 8 Gb of RAM memory, running Windows 7 (64-bit). Responses were registered via a Windows keyboard (Corsair), and stimuli were presented on a 24" BenQ XL 2420Z LED monitor with a 144 Hz refresh rate, at the native resolution of 1920 by 1080 pixels. A chinrest ensured a fixed viewing distance to the screen of 95 cm.

To make our stimuli, we followed the same procedure as Gomes and colleagues (in press), with three notable exceptions. First, we used five images of bicycles and five images of cars (instead of snakes and birds) retrieved via Google Image Search. Second, our experiment was a combination of Experiments 1A and 1B of Gomes and colleagues. That is, we included all three filtering conditions, as in Experiment 1A, and we created the CFS masks (Tsuchiya & Koch, 2005) following the optimized procedure of Experiment 1B (more details below). Third, we presented the target images and masks to different eyes by means of a mirror stereoscope, rather than through the usage of red-blue anaglyph glasses (i.e., presenting the target image through the blue RGB channel of the monitor, and the masks through the red RGB channel). This was motivated by the fact that the usage of a mirror stereoscope guarantees separate stimulation of the two eyes, whereas presentation through anaglyph glasses is prone to stereoscopic crosstalk (Baker, Kaestner, & Gouws, 2015). To maximize the similarity with the approach of Gomes et al., however, we adjusted the luminance output of the masks and target images to the luminance output that is elicited by the red and blue RGB channels respectively.

All stimuli were presented on a uniform gray presentation area subtending 8 by 8 degrees of visual angle (dva), which was embedded within a 0.5 dva white noise frame to facilitate binocular fusion of the two complementary images presented to each eye. Target stimuli were trimmed to fit a square area of 3.8 by 3.8 dva, and could be presented in one of the four quadrants of the presentation area. Stimuli were either broadband (i.e., unfiltered), low-pass filtered (< 6 cycles per image width), or high-pass filtered (> 30 cycles per image width). The full set of target stimuli can be seen in Figure S1. Next, we equated the luminance and luminance (RMS) contrast of all 10 bicycle and car images, within each spatial filtering condition. To induce continuous flash suppression (CFS; Tsuchiya, & Koch, 2005), we created 40 different masks, comprised of superimposed grayscale circles of varying diameter (range: 0.39 to 1.4 dva). The eventual masks used in our experiment were obtained by linear averaging of the high-pass and low-pass filtered components of the original masks (see Figure S2).

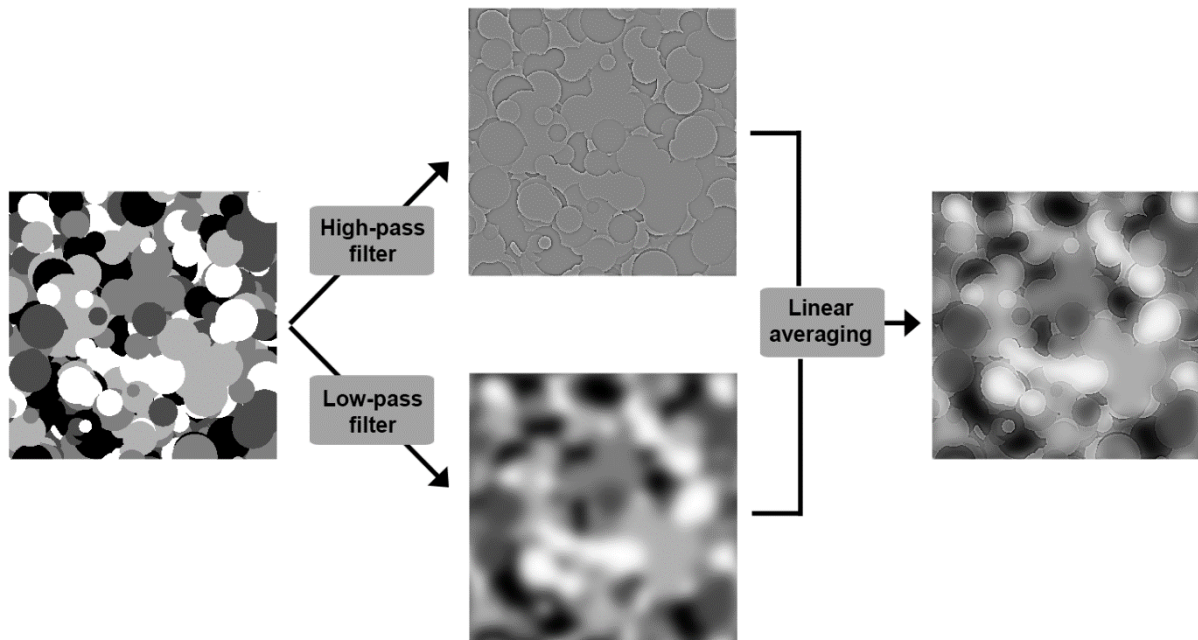


Figure S2. Masks for inducing continuous flash suppression were obtained (1) by creating broadband pattern masks, (2) by applying spatial filtering (as for the target stimuli depicted in Figure S1) to retrieve the low and high spatial frequency components, and (3) by computing their linear average.

3.4. *Perimeter-to-surface ratio*

Upon visual inspection of the ‘snake’ and ‘bird’ image categories (used in the study by Gomes et al., in press), we contemplated a number of visual characteristics that could underlie the faster detection of snake images compared to bird images. Among other differences, it was apparent that the snakes in the images comprised more edges that bordered the uniform image background than the birds (or: a higher perimeter-to-surface ratio). Such a difference could be of particular importance, because it is expected to be preserved in low-pass filtered images (in which the overall shape of the snakes/birds is preserved), but not in high-pass filtered images (in which surface information is discarded, but both within-figure and border edges are preserved). For the current study, we chose two image classes from the same semantic category, that were expected to differ in perimeter-to-surface ratio as well. Figure S3 shows that, indeed, unfiltered images of bicycles had a higher perimeter-to-surface ratio than cars (left panel), and that this difference was preserved after low-pass filtering (middle panel), but not after high-pass filtering (right panel).

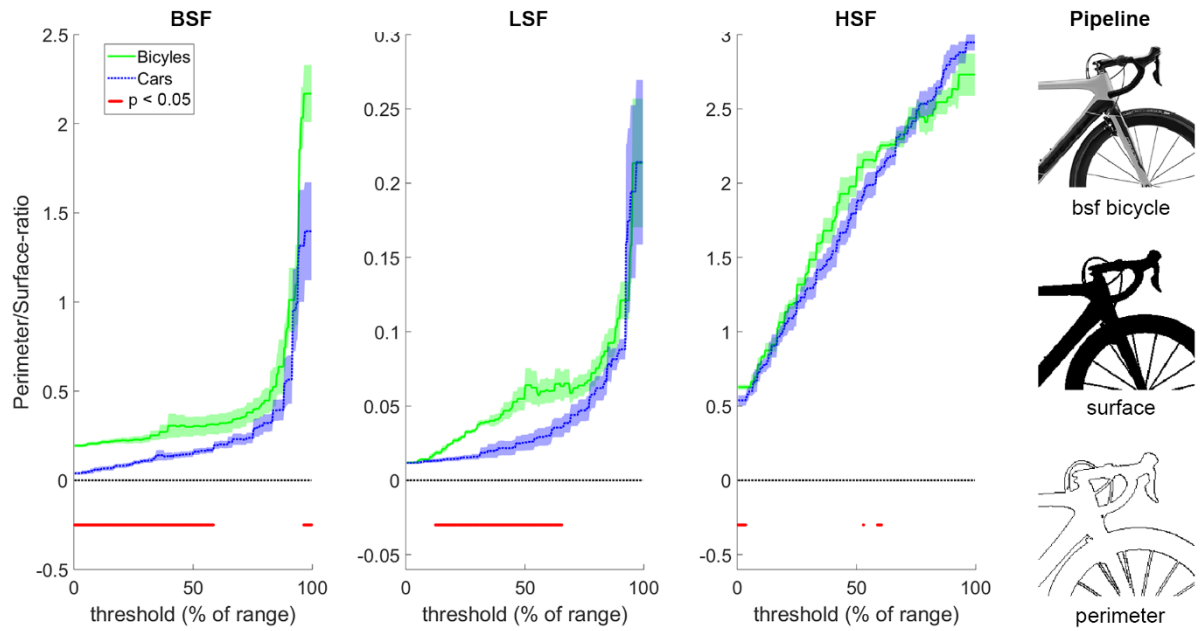


Figure S3. Perimeter-to-surface ratio, for the unfiltered (left panel), low-pass filtered (middle panel), and high pass filtered (right panel) images of bicycles and cars. Positive values on the Y-axes reflect a larger perimeter-to-surface ratio, the shaded area shows the standard error of the mean, and the red line depicts a significantly different (i.e., higher) ratio for bicycles than for cars. The perimeter-to-surface ratio was computed by (1) rounding the image to black-and-white to segregate the image from the background, using different thresholds depicted along the x-axis, then (2) computing the fraction of above-threshold pixels, reflecting the figure, (3) computing the fraction of above-threshold pixels adjacent to below-threshold pixels, reflecting the figure-edges, and (4) computing the ratio between the two. The right-most image shows a detail of an unfiltered bicycle images (top), with the surface area (middle) and perimeter (bottom) derived from an 8% pixel-intensity threshold. This illustrates the ability of our custom-made algorithm to retrieve surface and perimeter information from the images in our stimulus set.

3.5. Procedure.

The procedure of our study was identical to that of Gomes and colleagues (in press). Participants were seated in front of a dichoptic mirror set-up mounted on a chin-rest. First, we ensured proper binocular fusion of the complementary images presented to each eye of the participant. In case participants experienced diplopia, the distance between the two complementary images was adjusted, until binocular fusion was achieved. Before taking part in the current study, participants took part in an unrelated b-CFS study. We used the response times from this unrelated study (to targets presented in the left and right eyes) to derive participants' eye dominance. Next, the experimenter provided verbal instructions, and after a 30-trial practice block, participants took part in three experimental blocks of 40 trials each, separated by mandatory breaks.

The outline of a trial is depicted in Figure S4. On each trial, participants were required to report as fast as possible by means of the 'F', 'V', 'N', and 'J' keys of the keyboard whether (part of a) target stimulus appeared in (respectively) the top-left, bottom-left, bottom-right, or top-right quadrant of the presentation area. Each trial started with a one-second fixation screen, after which a rapid stream of masks (replaced at 10Hz) was presented to the non-dominant eye, and a target stimulus was presented to the dominant eye, in one of four quadrants of the presentation area. First, the target stimulus linearly increased from zero to

full contrast over 1.1 seconds. Next, the masks linearly decreased from full to zero contrast over 4 seconds. Eventually, the full contrast target stimulus remained on the screen until participants provided a response, or until 7 seconds had elapsed since target onset.

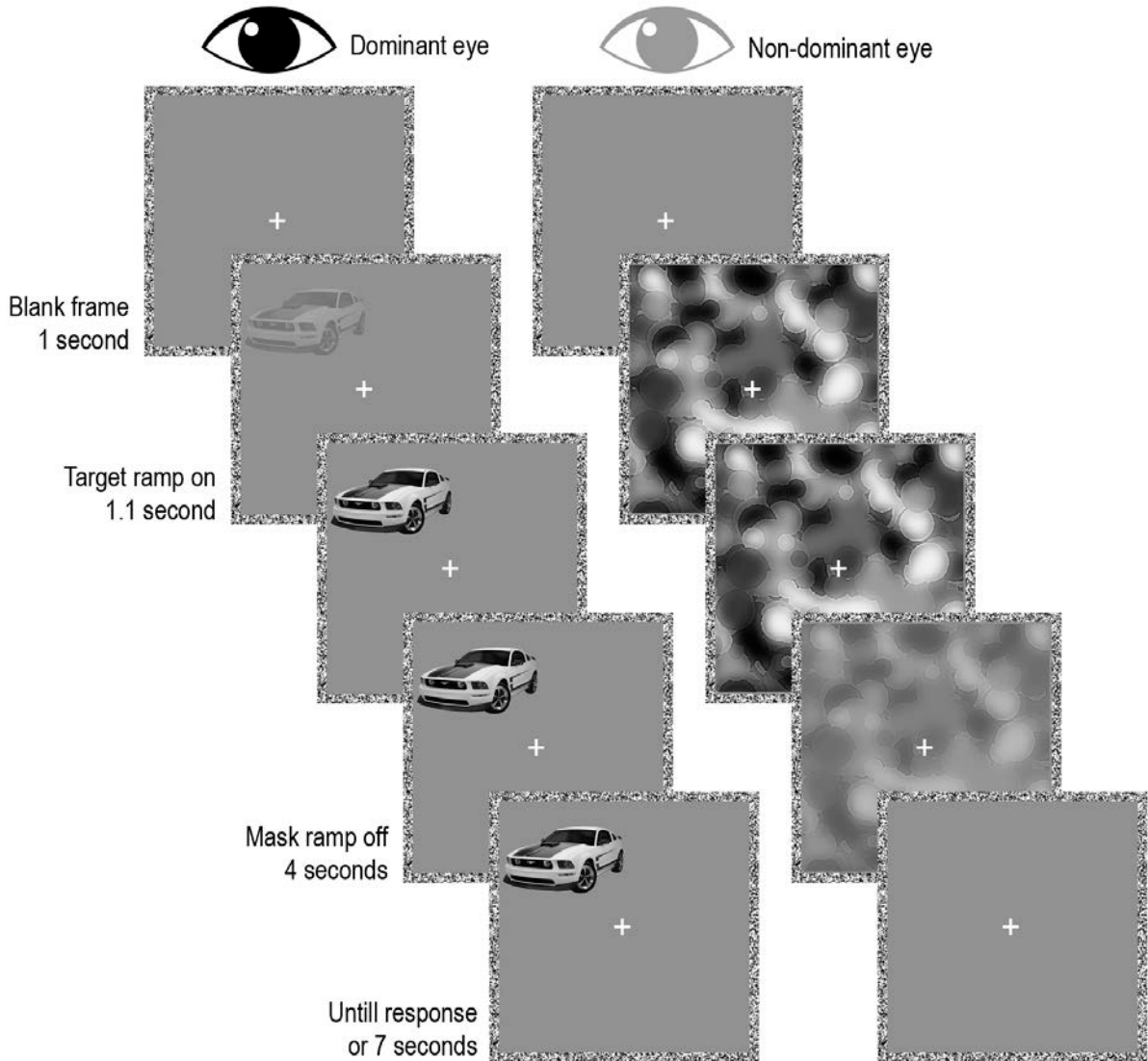


Figure S4. Schematic depiction of a trial. Participants were required to indicate as fast as possible in which of four quadrants a (part of a) target stimulus appeared. Each trial started with a one-second fixation interval. Next, a dynamic mask (10 Hz) was presented to the participants non-dominant eye, while a target stimulus was gradually increased in intensity from zero to full contrast in the participants dominant eye. Over the next four seconds, the contrast of the mask decreased from full to zero contrast. The full contrast target remained on the screen until the participant provided a response, or until seven seconds had elapsed since target onset.

3.6. Analyses.

3.6.1. Analyses of response times.

To investigate how image category and spatial filtering impacted response times in the b-CFS task, we first conducted a 2 by 3 Repeated-Measures ANOVA, with the within-subject factors *Image Category* (bicycle or car) and *Spatial Filtering* (unfiltered, low-pass, or high-

pass filtered) on the mean response times per participant. In case Mauchly's test of sphericity indicated that the assumption of sphericity is not met, the degrees of freedom were Huyhn-Feldt corrected. Effect sizes were reported as partial eta-squared (η_p^2). Subsequently, planned pairwise comparisons were conducted between the two image categories (bicycles versus cars), for each of the three spatial filtering conditions (unfiltered, low-pass, or high-pass filtered). To allow for optimal comparison between the three planned comparisons, all comparisons were conducted with the same test (t -test or Wilcoxon sign-ranked test) in case assumption violation is borderline for one comparison, and is clearly violated or not for the other two comparisons. Effect sizes are reported as Cohen's d where t -tests were employed, and as matched rank biserial correlations where Wilcoxon signed-rank tests were employed. Similarly, 95% confidence intervals were provided in the metric-of-interest (i.e., seconds or percentages) when t -tests were employed, and as the Hodges-Lehmann estimate where Wilcoxon signed-rank tests were employed. Trials with no response (25.8%, SD = 12.8%) and trials with incorrect responses (7.8%, SD = 6.3%) were excluded from these analyses. Additionally, two participants were excluded, because the high-pass filtered bicycle image condition contained no more data after exclusion of incorrect trials and no-response trials, leaving a total of 13 participants for the analysis of response times.

Analyses were conducted in MatLab and in JASP 0.8.2 (JASP Team, 2018). Analysis results obtained after applying the b-CFS normalization procedures described by Gayet and Stein (2017) were virtually identical to the results reported in the Results section below (i.e., on non-normalized RT data).

3.6.2. Analyses of localization accuracy.

To ensure that the differences in response times between conditions reflected genuine differences in detection speed, rather than differences in response criteria, we conducted the same analyses as described in Section 3.5.1. on the participants' localization accuracy (N = 13). That is, if participants would adopt a more liberal response criterion for (say) low-pass bicycle images than for low-pass car images, this would lead to faster response times, but also to more incorrect responses for low-pass bicycle images.

3.6.3. Analyses of target visibility.

Finally, because Gomes and colleagues (in press) had lower trial exclusion rates than we have (up to a maximum of 8%), we conducted a secondary analysis to corroborate our main analysis: we compared the proportion of targets that were (correctly) reported within the 7-second response deadline (i.e., the fraction of targets that eventually became visible to the participants) between the two image categories, for each of the three spatial filtering conditions. This analysis thus included all 15 participants. The analysis approach was identical to the one described above for the analysis of response times (3.5.1.) and accuracy (3.5.2).

4. Results

4.1. Response times.

Figure 1 (in the main manuscript) depicts the average response time per condition. A repeated-measures ANOVA on the response times revealed a main effect of Spatial Filtering, $F(1.4, 17.3) = 136.18, p < 0.001, \eta_p^2 = 0.919$ (Huyhn-Feldt corrected), a main effect of Image Category, $F(1, 12) = 85.83, p < 0.001, \eta_p^2 = 0.877$, and the crucial interaction between Spatial Filtering and Image Category, $F(2, 24) = 14.74, p < 0.001, \eta_p^2 = 0.551$. Subsequent paired-samples t -tests revealed that, for unfiltered images, initially suppressed bicycles ($M = 1.62$ seconds, $SD = 0.31$) were reported 1.15 seconds faster than cars ($M = 2.77$ seconds, $SD = 0.65$), $t(12) = 9.723, p < 0.001, d = 2.697$, 95% CI [1.41, 0.89]. Similarly, for low-pass filtered images, initially suppressed bicycles ($M = 2.48$ seconds, $SD = 0.58$) were reported 1.19 seconds faster than cars ($M = 3.67$ seconds, $SD = 0.68$), $t(12) = 7.698, p < 0.001, d = 2.135$, 95% CI [1.53, 0.85]. In contrast, for high-pass filtered images, there was no difference in response time between bicycles ($M = 4.46$ seconds, $SD = 0.76$) and cars ($M = 4.42$ seconds, $SD = 0.65$), $t(12) = -0.036, p = 0.972, d = -0.010$, 95% CI [0.46, -0.48]. Thus, we observe an advantage for detecting bicycle images over cars, for both unfiltered and low-pass filtered images, but not for high-pass filtered images.

4.2. Localization accuracy.

The repeated measures ANOVA on the localization accuracy revealed a main effect of Spatial Filtering, $F(1.1, 13.2) = 9.03, p = 0.009, \eta_p^2 = 0.429$ (Huyhn-Feldt corrected), but no main effect of Image Category, $F(1, 12) = 0.553, p = 0.471, \eta_p^2 = 0.044$, and no interaction between Spatial Filtering and Image Category, $F(1.18, 14.14) = 3.924, p = 0.062, \eta_p^2 = 0.246$ (Huyhn-Feldt corrected). Subsequent paired-samples t -tests revealed that, for unfiltered images, localization accuracy was similar for bicycles ($M = 97.7\%$, $SD = 3.3$) and cars ($M = 96.5\%$, $SD = 5.1$), $t(12) = 1.059, p = 0.310, d = 0.294$, 95% CI [-1.2, 3.6]. For low-pass filtered images, bicycles ($M = 95.6\%$, $SD = 6.9$) yielded marginally higher localization accuracies than cars ($M = 92.2\%$, $SD = 7.6$), $t(12) = 2.195, p = 0.049, d = 0.609$, 95% CI [0.0, 6.8]. For high-pass filtered images, there was no difference in localization accuracy between bicycles ($M = 71.1\%$, $SD = 29.6$) and cars ($M = 80.1\%$, $SD = 23.0$), $t(12) = -1.632, p = 0.129, d = -0.453$, 95% CI [-21.0, 3.0]. Note that, for each Spatial Filtering condition, image categories that yielded faster response times, also yielded (numerically) higher localization accuracies. Thus, the response time differences reported in Section 4.1. cannot be accounted for by a more liberal response criterion elicited by the image categories that yielded faster response times. Rather, these data suggest that the analyses of response times provided useful estimates of the difference in detection times between image categories.

4.3. Target visibility.

The repeated measures ANOVA on the target visibility revealed a main effect of Spatial Filtering, $F(1.3, 18.2) = 148.229, p < 0.001, \eta_p^2 = 0.914$ (Huyhn-Feldt corrected), no main effect of Image Category, $F(1, 14) = 0.043, p = 0.838, \eta_p^2 = 0.003$, but did reveal the crucial

interaction between Spatial Filtering and Image Category, $F(2, 28) = 60.924$, $p < 0.001$, $\eta_p^2 = 0.813$. Subsequent paired-samples t -tests revealed that, for unfiltered images, target visibility was higher for bicycles ($M = 97.3\%$, $SD = 3.2$) than for cars ($M = 89.7\%$, $SD = 15.1$), $W(14) = 42.00$, $p = 0.022$, $r = -0.300$, 95% CI [2.5, 25.0]. Similarly, for low-pass filtered images, target visibility was higher for bicycles ($M = 93.7\%$, $SD = 8.3$) than for cars ($M = 75.3\%$, $SD = 22.7$), $W(14) = 91.00$, $p = 0.002$, $r = 0.517$, 95% CI [12.5, 32.5]. In contrast, for high-pass filtered images, target visibility was lower for bicycles ($M = 19.3\%$, $SD = 15.0$) than for cars ($M = 43.0\%$, $SD = 26.6$), $W(14) = 1.500$, $p < 0.001$, $r = -0.975$, 95% CI [-35.0, -15.0]. Thus, in line with the analyses of response times reported in Section 4.1., an advantage for detecting images of bicycles over cars was observed with unfiltered images and low-pass filtered images, but not with high-pass filtered images. In fact, the analyses of target difficulty revealed that, with high-pass filtered images, an advantage was observed for detecting images of cars over images of bicycles (comparable to the advantage for detecting images of birds over snakes in study 1B of Gomes and Colleagues, in press).

5. Conclusion

Using the breaking continuous flash suppression paradigm, we investigated whether (1) observers would exhibit an advantage in detecting images of bicycles compared to images of cars (2) and if so, whether this advantage is driven by the high rather than low spatial frequency content of these image categories. We observed that (akin to images of snakes and birds; Gomes et al., in press) observers have an advantage in detecting images of bicycles over images of (contrast and luminance matched) cars, which is driven by the low spatial frequency content of the images. Cars and bicycles are similar in terms of (1) semantic category, (2) the amount of fear that they elicit, and (3) their propensity to rely on a dedicated subcortical processing pathway. Hence, we conclude that the specific pattern of results obtained for snake (compared to bird) detection, can be accounted for by differences in visual characteristics between the two image categories, and does not warrant the existence of a dedicated subcortical pathway for selective detection of one of those image categories.

5. Supplementary references

- Baker D. H., Kaestner M., Gouws A. (2015). Comparison of crosstalk for several methods of stereoscopic presentation. *Perception*, 44 (4), 455–455.
- Gayet, S., & Stein, T. (2017). Between-subject variability in the breaking continuous flash suppression paradigm: potential causes, consequences, and solutions. *Frontiers in Psychology*, 8, 437.
- Gayet, S., Van der Stigchel, S., & Paffen, C. L. E. (2014). Breaking continuous flash suppression: competing for consciousness on the pre-semantic battlefield. *Frontiers in Psychology*, 5: 460.
- Gomes, N., Soares, S. C., Silva, S., & Silva, C. F. (in press). Mind the snake: Fear detection relies on low spatial frequencies. *Emotion*.
- JASP Team (2018). JASP (Version 0.8.2)[Computer software].

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, 8(8), 1096-1101.