Attention, Perception and Psychophysics: Brief Report

The role of spatial frequency in emotional face classification

Ben J. Jennings[1], Yinan Yu[2] & Frederick A. A. Kingdom[1]

[1]Mcgill Vision Research, Department of Ophthalmology, McGill University, Montreal, Canada

[2]Department of Psychology, McGill University, Montreal, Canada

Abstract

Previous studies with emotional face stimuli have revealed that our ability to identify different emotional states is dependent on the faces' spatial frequency content. However, these studies typically only tested a limited number of emotional states. In the present study we measure the consistency with which twenty-four different emotional states are classified, when the faces are either unfiltered, high- or low-pass filtered, using a novel rating method that simultaneously measures perceived arousal (high to low) and valence (pleasant to unpleasant). The data reveal that consistent ratings are made for every emotional state independent of spatial frequency content. We conclude that emotional faces must possess both high and low frequency information that can be relied upon to facilitate classification.

Introduction

The ability to recognize the emotions in others' faces is of considerable importance (Darwin, 1872). A quick glance can reveal whether someone is sad, excited or worried, and their facial emotion is often indicative of their future behavior (Ekman, 1982; Izard, 1972). Moreover, the detectability and saliency of a face is dependent on its emotional content. For example, using a dynamic flash suppression paradigm (Tsuchiya & Koch, 2005), faces exhibiting fearful expressions emerge from suppression faster than faces containing neutral or happy expressions (Yang et al., 2007). In addition, visual search times for faces are influenced by their emotional expressions (Frischen, Eastwood & Smilek, 2008), for example angry faces are detected more quickly and more accurately than happy faces (Pitica *et al*., 2012). The conclusion from these studies is that ecologically relevant emotional expressions enjoy preferential treatment by the visual system, giving them a fast route into awareness. The rapid and accurate detection of a fearful or angry face may provide information about a potential local danger or threat, signaling that appropriate action may be required to maintain survival.

The well-known contrast sensitivity function, or CSF, which describes how sensitivity to contrast varies with spatial frequency, is inverse U-shaped with a peak around 2-6 cycles/deg under normal viewing conditions (Campbell & Robson, 1968; Graham & Nachmias, 1971). The CSF is believed to be the envelope of a number of underlying spatial frequency channels, each narrowly selective for a given range of spatial frequency (Sachs, Nachmias & Robson, 1971). Numerous studies have examined how the detection and perception of faces, including their emotional content, is influenced by spatial frequency content. The faces in these studies are typically high- and low-pass

filtered. In general high spatial frequency content provides the fine details of the face and low spatial frequency content its overall structure; however this belies a more nuanced involvement of spatial frequency in the perception of facial emotion. Low spatial frequencies appear to be important for the detection of threats (Bar *et al*., 2006) and play a prominent role in identifying faces that express pain (Wang, Eccleston & Keogh, 2015), happiness (Kumar and Srinivasan, 2011) and fear (for example, Holmes, Winston & Eimer, 2005), although this has been challenged for fearful faces (Morawetz et al., 2012). High spatial frequencies on the other hand have been shown to play a prominent role in identifying sad, happy, or again fearful emotional faces (Wang, Eccleston & Keogh, 2015; Goren & Wilson, 2006).

The above studies typically employ only small numbers of emotional states, for example sad vs. happy vs. neutral. Moreover, different studies have tested different emotions, used different faces (image databases), different high- and low-spatial-frequency cut-off points and used different methods, e.g. identification, detection in noise, and reaction times. All these factors make cross-study comparisons problematic. In particular, the differences found in previous studies may have been due to the task employed, for example a detection task may favor low frequency faces.

In the present study we examine how high and low spatial frequency information is made use of in a classification task that employs a wide variety (twenty-four) of emotional expressions. The faces contained either the full range of spatial frequencies, or the high or the low spatial frequency content in isolation - see Fig. 1a for examples. Our data reveals that observers rate emotions identically irrespective of their spatial frequency content.

Hence, features specific to each emotion must be represented by both high and low frequency filtered faces.

General Methods

Participants

Eighteen observers took part in the experiment (10 female), with mean age 22.3±2.1. All observers had 6/6 vision, in some cases achieved through optical correction.

Equipment

The experiment was performed using a MacBook Pro (Apple Inc.) installed with a 2.9 GHz i7 processor and 4 GB of DDR3 memory running OS X (El Capitan, version: 10.11.6). The gamma corrected display had a resolution of 1280 x 800 pixels and a frame rate of 60 Hz. MatLab (Mathworks Inc.) running PsychToolbox V3 (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) was used to present the stimuli, while observer responses were submitted via the built-in trackpad. During data collection observers were positioned 60 cm from the display.

Stimuli, experimental procedure and data analysis

Twenty-four emotional expressions were selected from the McGill Face Database (Schmidtmann et al., 2016). They were: affectionate, alarmed, amused, baffled, comforting, contented, convinced, depressed, entertained, fantasizing, fearful, flirtatious, friendly,

hateful, hostile, joking, panicked, playful, puzzled, reflective, relaxed, satisfied, terrified and threatening. The stimuli used to isolate the high and low spatial frequency ranges were generated using custom-written MatLab software, in which the images were filtered using a bank of log-Gabor filters, with four orientations (0, 45, 90 and 135°) and with the DC component being restored in the high frequency stimuli. The frequency ranges selected were based on previous studies (Vuilleumier et al., 2003; Bannerman et al., 2012) and correspond to cut-off frequencies of ~20 cycles per face for the high frequency condition and ~6 cycles per face for the low frequency condition in the horizontal direction, this removed spatial frequencies known to be important for normal face perception.

The different conditions were interleaved in random order in a novel "point-and-click" computer based task. One trial consisted of the presentation of a single image for a duration of 200 ms, then replaced with an image of an Arousal-Valence emotion space (Russell, 1980), as illustrated in Fig. 1b - the red labels in the figure give the reader the gist of the emotions represented by each region but were not shown during testing. The space defined two dimensions of emotion along (i) the arousal level, for example a panicked or annoyed face would convey a high arousal level, whereas a contented or relaxed face would convey a low arousal level; (ii) the valence, i.e. pleasant vs. unpleasant, for example an unpleasant, horrified or disgusted expression would be placed towards the left-hand side of the space (negative valence) and a pleasant expression, for example a happy or amused face would be placed towards the right-hand side of the space (positive valence). Faces that are perceived with neutral emotions would be positioned towards the center of the space. Each face each subtended ~5° horizontally and the square response area subtended ~12.3 x 12.3° of visual angle.

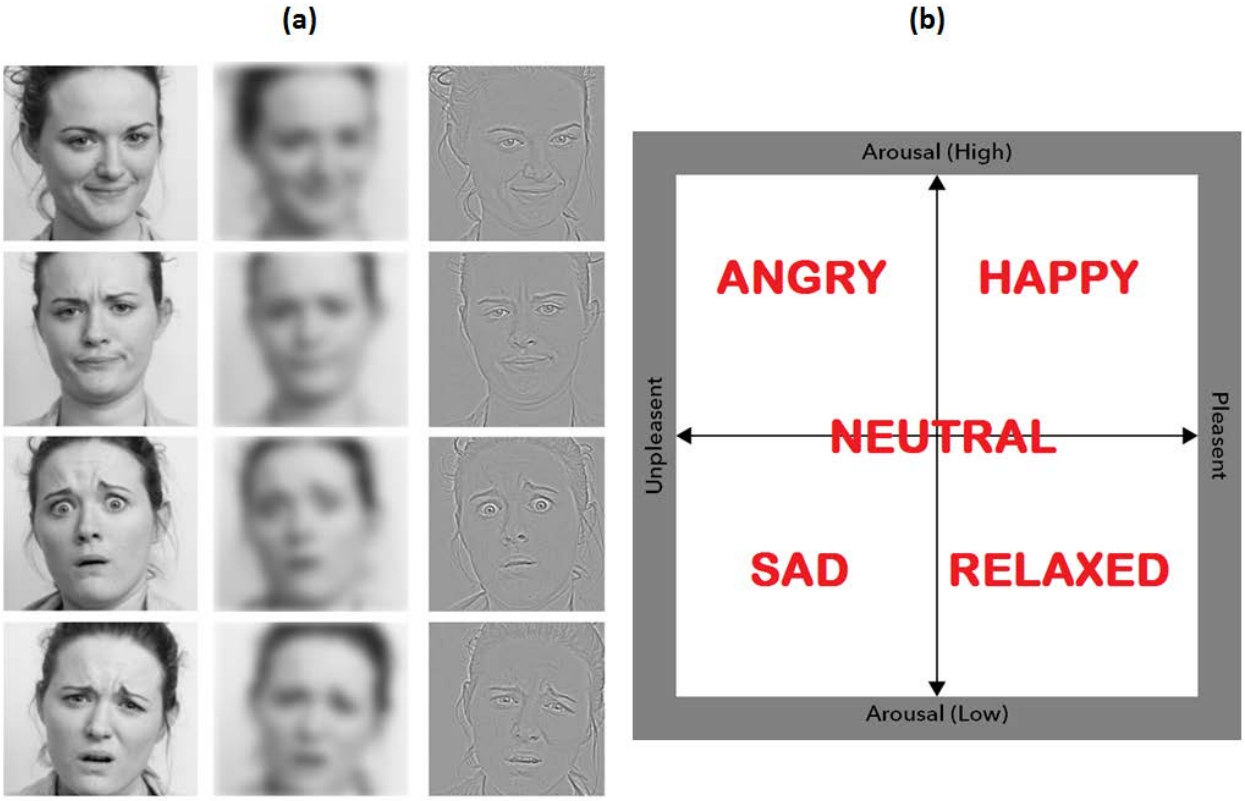

Fig.1a-b. *(a) Illustrates four exemplar emotional faces, with unfiltered faces (left), low frequency isolating (central) and high frequency isolating (right). (b) shows the valence vs. arousal emotion space within which observers placed responses. Example emotional states are displayed in red text (not present during testing).*

The task for each observer was to position the on-screen cursor using a computer trackpad, and click the location in the emotion space that corresponded to the emotion being expressed, i.e., the emotion of the face and not the emotion, if any, induced in the observer.

Data for each observer was collapsed per emotion and the mean emotion coordinates were calculated for each of the three spatial frequency conditions. In order to reveal any shift in location of a given emotion resulting from the frequency content manipulation, spatial statistics were performed and corrected p-values were obtained.

Results and discussion

Mean data (n=18) for the unfiltered stimuli are plotted in Fig. 2a along with the emotion name as specified by the McGill Face Database. All data, i.e. all frequency conditions, are plotted in Fig. 1b, with black circles for the unfiltered faces, red circles for the high frequency faces, and green circles for the low frequency faces. For clarity, each tested emotion is plotted separately (Fig. 2c), using the same colour code as Fig. 2b.

Each emotion was analysed in turn by performing multiple (Bonferroni corrected) 2-tailed t-tests for each combination of spatial frequency (unfiltered, high frequency and low frequency) to test for differences in classification location in both the arousal and valence directions. The statistics reveal no differences in the arousal direction for any combination of unfiltered, high frequency or low frequency stimuli for a given emotion type. The lowest and highest p-values in the arousal direction were for the differences in location between respectively the unfiltered and high-frequency conditions for the fearful expression ($t(17)=-3.10$, $p=.47$), and the unfiltered and high-frequency conditions for the relaxed expression ($t(17)=0.03$, $p=1$). In the valence direction the statistics again revealed no differences in the locations between any combination of unfiltered, high frequency or low frequency stimuli for each emotion type. The lowest and highest p-values in the valence direction were for the differences in location between respectively the unfiltered and high-frequency conditions for the depressed expression ($t(17)=-1.76$, $p=.29$), and the low and high-frequency conditions for the threatening expression ($t(17)=-0.025$, $p=1$).

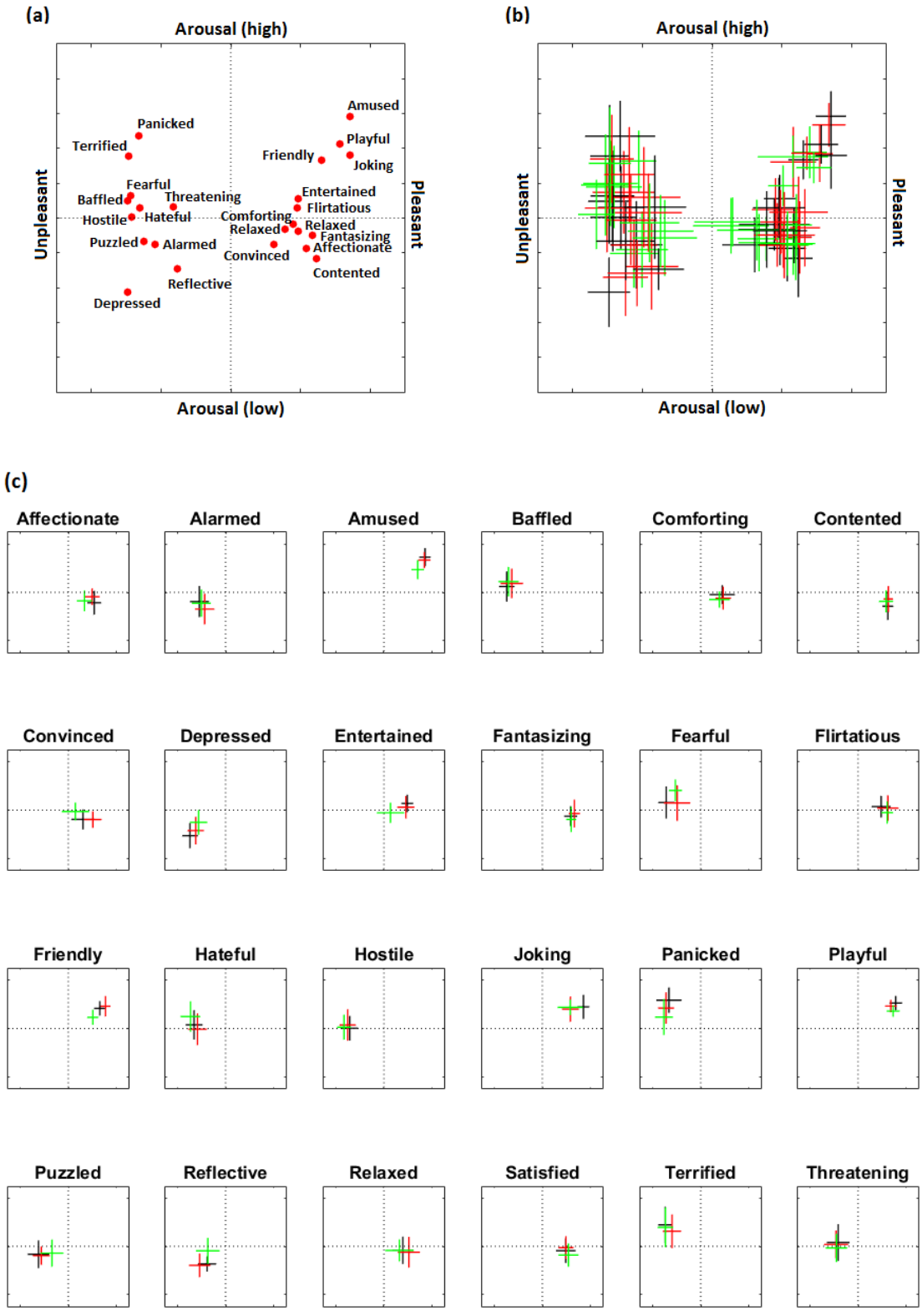

Fig. 2a-c. *(a) plots mean data for the unfiltered faces in the arousal vs. valence emotion space - the data point labels are taken from the McGill Face Database. The data form a U-shaped*

distribution within the space. (b) plots all spatial frequency data: unfiltered (black), high frequency (red) and low frequency (green). Error bars are ±2SE. (c) Each panel plots the three spatial frequency conditions for a given emotion using the same colour coding as (b).

As our conclusion is based upon a statistically null result we calculated confidence intervals (CIs) for effect sizes, separately for both the arousal and valence directions, using a bootstrapping resampling method (Banjanovic and Osborn, 2016). Taking each emotion in turn subsets of data were randomly selected (1000 simulations were performed per emotion) and CI for effect sizes were calculated and collapsed across spatial frequency. Hence for each of the arousal and valence directions the effect size and corresponding CIs were obtained for each combination of spatial frequency condition, i.e. between (i) the unfiltered and high-frequency conditions, (ii) the unfiltered and low-frequency conditions and finally (iii) the high and low-frequency conditions. These effect sizes, along with their 95% bootstrapped CIs, in both arousal and valence directions, are plotted in Fig. 3. The red horizontal lines indicate the v*ery small* (0.01), *small* (0.2) and *medium* (0.5) effect size ranges as termed by both Cohen (1988) and Sawilowsky (2009). It can be seen that the mean values based on the above bootstrapping method fall in the range 0.14 to 0.22.

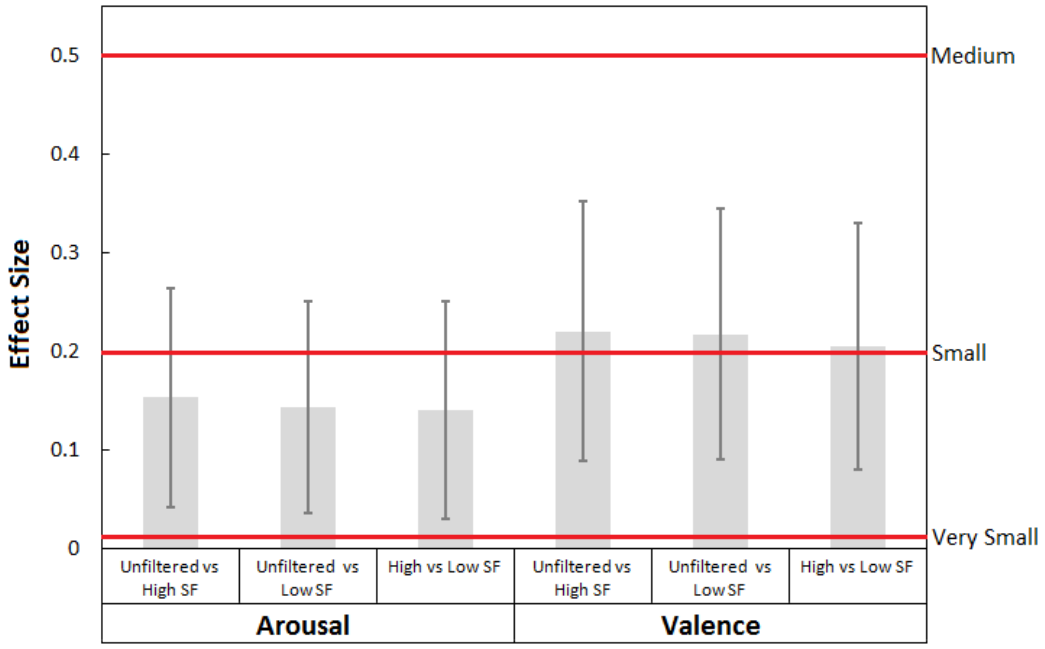

Fig. 3. *Mean effect sizes are plotted with 95% bootstrapped confidence limits for all combinations of spatial frequency conditions, in both the arousal and valence directions.*

The data obtained using this new psychophysical method for classifying perceived emotions demonstrates that emotional faces can be perceived and hence classified consistently independent of high or low spatial frequency content. This is the case for ratings in both the arousal and valence dimensions.

Some emotions can be recognised from just the mouth (Guarnera et al., 2015). However by employing faces conveying a wide range of emotions Schmidtmann et al. (2016) showed that accuracy for selecting an emotionally descriptive word in a 4AFC task was equal when either the entire face or a restricted stimulus showing just the eyes was employed. It is thus plausible that arousal-valence ratings are based on eye information only.. Our classification data however cannot confirm this as our stimuli were spatial-frequency filtered, and the salient face feature might be spatial-frequency dependent.

Our data forms a U-shaped curve within the emotion space, leaving some regions of the space unused. This is partly to be expected as it is unlikely that a face with a 'neutral' valence could be perceived as highly arousing. The obtained U-shape distribution is similar to that obtained using the International Affective Picture System (Lang *et al*., 1988; Libkuman *et al*., 2007), where observers made ratings using just a coarse scale, a result that has since been replicated cross culturally (Silva, 2011). The task employed by Libkuman *et al*. was for observers to indicate the intensity of the emotional response they experienced. Interestingly, between the three aforementioned and the present study, in which observers were required to classify the perceived emotion in a face, similar results were found.

Conclusion

Although different information is contained in images of emotional faces filtered to isolate either their high or low spatial frequency content, human observers can utilize either to identify emotional states.

Acknowledgments

This work was funded by the Canadian Institute of Health Research grant #MOP 123349 given to FAAK.

References

Bannerman, R. L., Hibbard, P. B., Chalmers, K. & Sahraie, A. (2012). Saccadic latency is modulated by emotional content of spatially filtered face stimuli. Emotion, 12(6), 1384-92.

Banjanovic, E. S. & Osborn, J.W. (2016). Confidence Intervals for Effect Sizes: Applying Bootstrap Resampling. Practical Assessment, Research & Evaluation. 21:5.

Brainard, D. H. (1997). The Psychophysics Toolbox, Spatial Vision 10*:433-436.*

Campbell, F. W. & Robson, G. J. (1968). Application of fourier analysis to the visibility of gratings. Journal of Physiology, 197(3), 551–566.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. Routledge. ISBN 1-134-74270-3.

Darwin, C. R. (1872). The Expression of the Emotions in Man and Animals. John Murray, London.

Ekman, P. (1982). emotion in the human face (Second Edition). Cambridge University Press.

Fiorentini, A., Maffei, L. & Sandini, G. (1983). The role of high spatial frequencies in face perception, Perception. 12(2), 195-201.

Frischen, A, Eastwood, J. D. & Smilek, D. (2008). Visual search for faces with emotional expressions. Psychological Bulletin, 134(5), 662-76.

Goren, D. & Wilson, H.R. (2006). Quantifying facial expression recognition across viewing conditions. Vision Research, 46, 1253-1262.

Holmes, A., Winston, J.S. & Eimer, M. (2005). The role of spatial frequency information for ERP components sensitive to faces and emotional facial expression. Cognitive Brain Research, 25(2), 508.

Izard, C. E. (1972). Patterns of emotion: a new analysis of anxiety and depression. Academic Press. New York.

Kleiner, M., Brainard, D., Pelli, D. (2007). What's new in Psychtoolbox-3? Perception 36 ECVP Abstract Supplement.

Lang, P. J., Öhman, A., &Vaitl, D. (1988). The International Affective Picture System [Photographic slides]. Gainesville: University of Florida, Center for Research in Psychophysiology.

Libkuman, T. M., Otani, H., Kern, R., Viger, S. G. & Novak, N. (2007). Multidimensional normative ratings for the International Affective Picture System. Behavior Research Methods, 39(2), 326-334.

Morawetz, C., Baudewig, J., Treue, S. & Dechent, P. (2011). Effects of spatial frequency and location of fearful faces on human amygdala activity. Brain Research, 1371, 87-99.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies, Spatial Vision 10:437-442.

Pitica, I., Susa, G., Benga, O. & Miclez, M. (2012). Visual search for real emotional faces: the advantage of anger. Procedia - Social and Behavioral Sciences, 33, 632-636.

Russell, J. A. (1980). A circumplex model of affect. Journal of Personality and Social Psychology, 39(6), 1161-1178.

Sachs, M.B., Nachmias, J., & Robson, J.G. (1971). Spatial frequency channels in human vision. Journal of the Optical Society of America A, 61, 1176–1186.

Sawilowsky, S. (2009). New effect size rules of thumb. Journal of Modern Applied Statistical Methods, 8 (2): 467–474.

Schmidtmann, G., Sleiman, D., Pollack, J. & Gold, I. (2016). Reading the Mind in the Blink of an Eye - A novel database for facial expressions. Perception, 45, 238-239.

Silva, J. (2011). International Affective Picture System (IAPS) in Chile: A cross-cultural adaptation and validation study. Terapia Psicologica, 29(2):251-258

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

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), 624-31.

Wang. S., Eccleston, C. & Keogh E. (2015). The role of spatial frequency information in the recognition of facial expressions of pain. Pain, 156(9), 1670-82.

Yang, E., Zald, D. H. & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. Emotion, 7(4), 882-886.