Who Needs Cheeks?

Eyes and Mouths are Enough for Emotion Identification

and

Evidence for a "Face Superiority Effect"

Nila K Leigh

131 Ave B (Apt. 1B)

New York, NY 10009

Stuyvesant High School

345 Chambers St.

New York, NY 10282

N.K. Leigh (2000) Who needs cheeks? Eyes and mouths are enough for emotion identification and Evidence for a "Face Superiority Effect". Intel Science Talent Search. Semifinalist. http://psych.nyu.edu/pelli/#intel

SUMMARY

Behind this paper was the desire to discover whether the visual system concentrates on a specific area of the face when judging emotion. It was found that humans do not focus on just one zone. Depending on the situation, humans can identify an emotion by either looking at the eye-and-brow area or the mouth-and-chin area. A parallel was also drawn between "Word Superiority" for word recognition and findings of a "Face Superiority Effect" for emotion recognition.

TABLE OF CONTENTS

LITERATURE RESEARCH

Recognition of facial emotions is a basic life skill crucial to everyday interactions. Socially, it is extremely important for humans to be able to distinguish emotions accurately. Most people take this ability for granted. However, it is interesting to stop and wonder exactly how the human visual system processes an emotion. What do we really look at in the human face in order to accurately judge emotion?

Face Recognition:

A main question in assessing how facial expressions are recognized is whether faces are processed holistically (in one piece) or featurally (as representations of their component features). It has been found that the visual system relies on holistic representation for recognition of an individual's identity (Tanaka & Farah, 1993). However, this may be very different than the process that takes place when identifying specific emotions within an individual's face.

Emotion Recognition:

All normal faces have the same configuration. The brow, the eyes, the nose, the mouth, and the chin are always in the same fixed locations. Facial expressions of emotion are stereotypical configurations of movements of the facial muscles (Ekman, 1992). These expressions of emotion are the product of specific combinations of muscle movements. Most expressions of emotion are the product of several muscle movements occurring in various combinations; however, in some cases movements of individual muscles occur alone and can express an entire emotion by themselves. Knowing this, one can deduce that the visual system is accustomed to looking at faces in a certain way in determining the emotion being expressed.

One interesting question is whether there is an exact area of the face that the viewer looks at in order to identify the emotion being conveyed. One can think of facial areas in one of two ways. The first defines the term "area" as a zone of the face, such as the top or the bottom. The second defines "area" as a certain component, such as a raised brow, wrinkles in the forehead, or an open mouth. It has been found that because emotional expressions can utilize many muscles of the face at once or just one or two at a time, it is impossible for the visual system to only concentrate on a certain component, such as the eyes or lips. (Boucher & Ekman, 1975). A person looking to identify an emotional expression in this way would be so focused on one feature that he or she would miss many expressions that utilize muscle combinations not necessarily centralized around that one feature. Because of this, when researching on which area of the face an observer concentrates when judging an emotion, it is most useful to consider areas to be general regions of the face, instead of specific features. This gives the most precise reading of exactly where emotions are perceived.

In this paper, a "hotspot" is defined as the area most critical to the identification of emotion within the face. Schwartz, Bayer, and Pelli (1997) claimed that in identifying facial expressions the hotspot is the mouth-and-chin region. However, the set of facial expressions used as stimuli for this study did not portray emotions. (See **Fig 1**) The woman photographed for the set simply made twenty-one different faces, silly faces. When in comes to the basic universal emotions, it is generally agreed that there are either six or seven distinct emotions (Ekman, 1992). Due to the nature of the stimuli, it was impossible to view the results from their study as being applicable to emotions. Another problem with drawing conclusions from this earlier experiment was that facial emotions

utilize muscles from all over the face, not just those on the mouth area, as concentrated on by the woman in the set. (Bayer, Schwartz, and Pelli, 1997).

To look at perceptions of emotional expressions, a different face set had to be used. **Fig. 2** shows Professor Paul Ekman demonstrating the seven primary emotions. The eighth face in the set is neutral, showing no emotion at all. The seven primary emotions are, from left to right along the rows: happiness, surprise, anger, sadness, contempt, fear, and disgust. Professor Ekman utilizes all facial muscles, not just the mouth, and his emotional expressions are true to spontaneous, real-life emotions.

Word Superiority:

A phenomenon exists that is known as "Word Superiority". When an observer is shown a letter briefly, the task of identifying the letter displayed is more difficult when the letter is shown alone against a gray background than if it is shown within a word, even if the rest of the word provides no extra information. For example, when identifying a 'C' or a 'J', it is easier to recognize the letter when the observer is shown either the word 'COIN' or the word 'JOIN', despite the fact that the 'OIN' provides no extra information (Johnston & McClelland, 1974). When thinking about face recognition, it is often useful to compare the identification of faces to that of other types of stimuli. This paper addresses the question of whether there is “Face Superiority Effect”.

Techniques:

One of the most effective methods of analyzing how well an observer performs on visual recognition tests is by comparing contrast thresholds. Contrast threshold is the measure of how much contrast must be present in a stimulus for it to be identified accurately. Contrast threshold is a computer generated number measured after a subject performs a given task enough times for his or her performance to stabilize.

In order to discover exactly where a hotspot lies, it is useful to employ the curtain technique. In this procedure, a computer program automatically covers photographs with layers of varying densities of white gaussian noise. Subjects attempt to identify the stimulus behind the noise curtain. The noise covers varying amounts of the face from the top, bottom, left side, or right side. The computer then makes computations based on the subject's responses as to where performance on the task is most harmed by noise. Putting the results together, one can see where the critical range is on a face – the "hotspot". The curtain technique is good for identifying exactly where people look to detect which emotion is being displayed on the face they are observing.

Goals:

This paper attempts to answer two questions. The first is whether there is a specific portion of the face that observers concentrate on instinctively to identify emotions. The second question is whether there is a "Face Superiority Effect", in other words, whether it is easier to identify an isolated mouth when it is shown against the background of a neutral face than when it is presented alone against a gray background.

EXPERIMENT 1

Participants:

The author was the subject.

Stimuli:

Eight photographs of Professor Paul Ekman were loaded into the computer and cropped to an 8.01 by 8.01 degree size. The photographs showed Professor Ekman displaying the seven primary emotions (happiness, surprise, anger, sadness, contempt, fear, and disgust) and the neutral face. (See **Fig. 2**) The computer masked varying edge segments of each photograph with curtains of white gaussian noise. The curtain came from the left side, right side, top, or bottom of the stimulus. (See **Fig. 3**) The trials involving white gaussian noise curtains were alternated with trials of curtains consisting simply of varying contrasts. This method showed the difference between the two curtains and provided a threshold for the human observer.

Procedure:

The subject sits one meter from the computer screen. One of the eight facial emotions appears in the center of the computer screen for 200 ms. After another 200 ms, the entire set of 8 expressions comes up on the screen. This reply board is a fixed set that never changes in size, location or contrast. The subject uses the mouse to click on the emotion in the set that matches the one flashed 200 ms earlier. Correct answers are rewarded by a short tone. Each time the stimulus is presented, it is called a trial. A run

equals forty consecutive trials. For every run, the computer changes how far the masking noise or contrast curtain intrudes into the stimulus area. The computer calculates where the subject's performance is most damaged by the curtain. Putting together all the results for these trials shows where the "hotspot" is, the area that is most critical to recognizing what the particular emotion is. (See **Fig. 4**)

Results:

This study found two critical areas for emotion recognition. The hotspot actually shifts depending on where the noise curtain is. If the top half of the face is covered with noise, observers use the mouth-and-chin hotspot. If the bottom half of the face is covered with noise, observers concentrate on the eyes-and-brow hotspot. Calculations made from the positions of the right and left noise curtains show that the visual system does not have a preference for either side. This experiment provides concrete quantifiable dimensions for the critical area. The results show that the forehead, ears, cheeks, and nose provide relatively little information, too little for the visual system to concentrate on when identifying emotions. The process of identifying real emotions and the process of identifying faces that are the result of “making faces”, such as those used by Bayer, Schwartz, and Pelli (1997), are different operations.

EXPERIMENT 2

Participants:

As in Experiment 1, the author was the subject.

Stimuli:

The two sets of facial expressions used in Experiment 2 were based on the set used in Experiment 1. Adobe Photoshop alterations were performed on the Ekman full-face photographs to make a second set consisting of the eight different mouths set against the background of Professor Ekman's neutral face. (See **Fig. 5**) Each photograph in this hotspot set was 8.01 by 8.01 degrees. With this set, the neutral face background served the same purpose as the 'OIN' in COIN and JOIN (Johnston & McClelland, 1974), providing a non-informational familiar setting for the mouth expressions. In the same way, the mouths from the various emotions were designed to serve as the letters (J or C for example). The final set consisted of the eight mouths against a gray background. (See **Fig. 6**) The size of the mouth-alone set was adjusted to 1.98 degree by 1.98 degrees. The dimensions of the actual mouths were identical to those of the hotspot set.

Procedure:

Contrast thresholds were calculated for each of the two face sets. For each trial the subject sits one meter from the computer screen. One of the eight photographs appears in the center of the computer screen for 200 ms. After another 200ms, the entire set of 8 photographs comes up on the screen. Again, this reply board is a fixed set that

never changes in size, location or contrast. And again, the subject uses the mouse to click on the photograph in the set that matches the one flashed 200 ms earlier. This represents one trial with a run equaling forty consecutive trials. For each of the two sets, the subject performed approximately 50 runs in order to acquire a stable and reliable threshold.

The computer also automatically adjusts the contrast or noise from trial to trial within one run according to the subject's accuracy. For every run, the computer calculates the contrast energy threshold. Here, contrast threshold was measured with and without noise. Contrast threshold is an estimate of the contrast necessary for the subject to perform accurately 82% of the time. A lower contrast threshold indicates better performance on a task than does a higher contrast threshold.

One problem often run into with emotion experiments is that a great deal of ambiguity accompanies vocabulary and language. Making subjects verbalize their responses brings other factors into the results. (Dodson, Johnson & Schooler, 1997). These experiments eliminated language, thereby clarifying results.

Results:

The average contrast threshold for the mouth-only set was 0.4. The hotspot set, however, produced a contrast threshold of 0.2. (See **Fig 7**) These results show that there is in fact a “Face Superiority Effect” in play here. The visual system is used to looking at an entire face for emotion information and putting the mouth emotions in context improves performance. The familiarity of the neutral face background lowers the observer’s contrast threshold.

Future Research

These findings have further implications. These curtain results show that identification is as accurate with the eyes-and-brow hotspot as with the mouth-and-chin hotspot. Therefore, testing the eyes alone against the full neutral face with the eyes hotspot in an experiment duplicating Experiment 2 is likely to produce the same results as those for the mouth-and-chin hotspot. (See **Fig. 8a**, **8b**, and **8c**) It can be assumed that the eyes-and-brow area alone would yield a higher threshold than the neutral face with an emotion bearing eyes-and-brow area by approximately a factor of two, as was the case with the mouth-and-chin hotspot tested here.

FIGURES

Fig 1 - The Bayer, Schwartz, and Pelli face set.

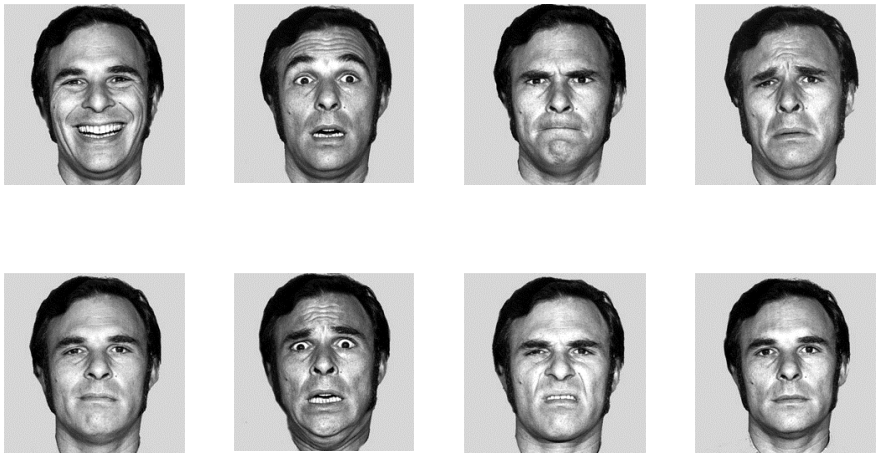

Fig 2 - The Paul Ekman full-face set of emotions.
The emotions, from left to right along the rows, are: happiness, surprise, anger, sadness, contempt, fear, disgust, and neutral.

Fig. 3 - Ekman face with bottom noise curtain.
Can you identify this emotion from the eight photographs in Fig. 2?

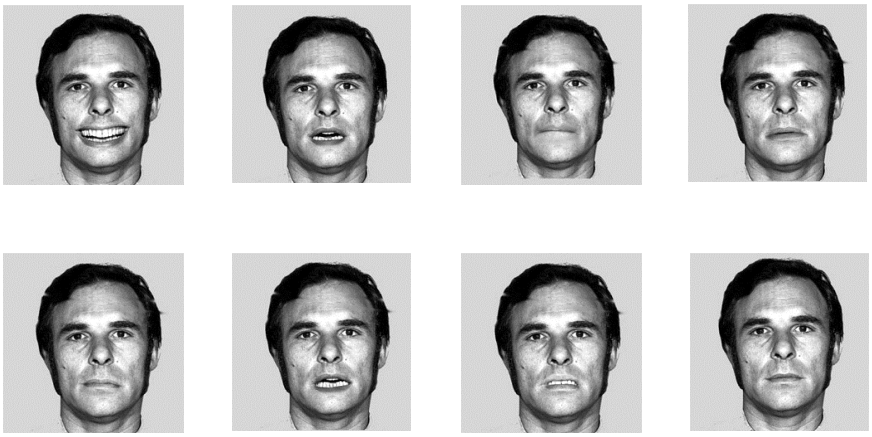

Fig.5 - Ekman hotspot face set. Mouth emotions on neutral face background.

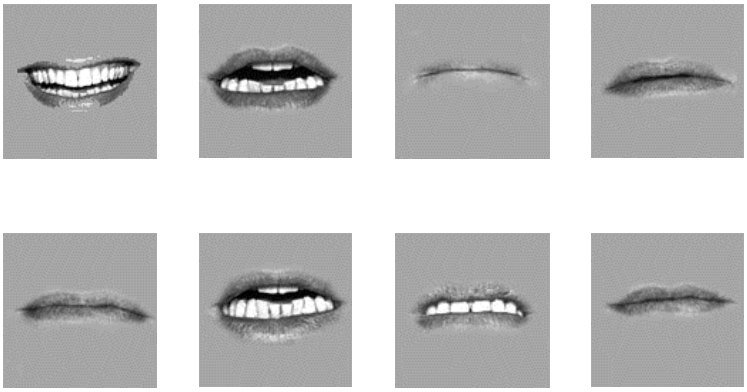

Fig. 6 - Ekman mouth-only set. Mouth emotions against gray background

Fig. 8a - Ekman full-face disgust with hotspot area shown.

Fig. 8b - Ekman disgust eye hotspot on background of neutral face.

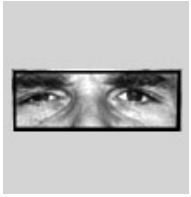

Fig. 8c - Ekman eye hotspot alone against gray background.

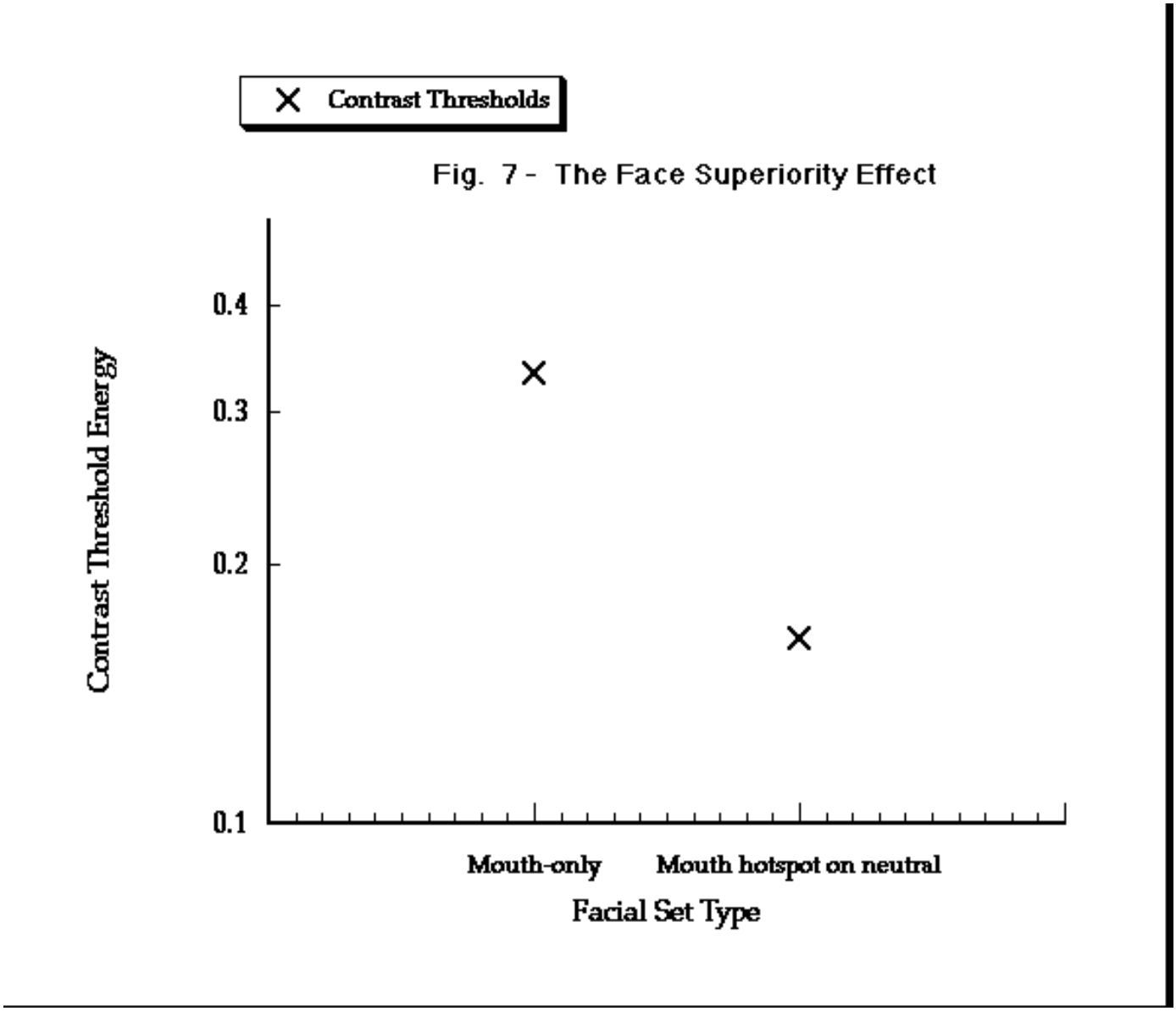

Fig. 7 - The Face Superiority Effect

DISCUSSION

As previously stated, two questions were asked when researching this paper. As to the first question, it was found that there is not only one hotspot for facial emotions. The area that subjects focus on varies depending on the circumstances. A person can identify emotion using either one of the hotspots or both at any time. In this way, one can see the value of expressing an emotion fully. Simply turning down the corners of one's mouth will not necessarily convey sadness. This also has implications when one thinks of the use of masks and cultural costumes. By covering his eye and brow area, Zorro forces those interacting with him to switch their focus to his mouth-and-chin area in order to judge his emotions accurately. Alternately, operating surgeons with their mouth-and-chin coverings force their patients and colleagues to judge emotion based on their eye and brow area only. Experiment 1 shows that it is possible to perceive emotions in each of the above scenarios despite limited information.

Experiment 2 illustrates that there is indeed a "Face Superiority Effect". As hypothesized, the principles of Word Superiority can be applied to emotion perception. The human visual system is accustomed to viewing an entire face. Although an emotion can be identified by one hotspot alone, the hotspot cannot be identified as easily when it is seen out of the context of a face. Placing a piece of the face showing its share of any emotion onto a full neutral face, helps an observer to identify the emotion displayed in that isolated piece. This finding implies that faces are perceived in a special way. The visual system is confused when it is shown partial emotional expressions away from the conventional background on which they are usually seen.

ACKNOWLEDGEMENTS

I would like to thank Professor Denis Pelli most sincerely for his excellent guidance and knowledge and for the generous use of his lab. A thank you is in order to Melanie C. Palomares whose constant support has been invaluable to me throughout the past year. I also would like to thank Professor Paul Ekman. I could not have completed this project without his guidance and the generous use of his facial images as the basis of my experiments. I want to thank Ms. Marianna L. Reep for her aid in preparing me for this Intel Report. Lastly, to all the students in Professor Pelli's lab who have encouraged me throughout the past year, thank you.

REFERENCES

Bayer, H., Schwartz, O., and Pelli, D. (1997). Recognizing Facial Expressions Efficiently. Poster presented at Center for Neural Science, New York University. September 17, 1997.

Boucher, J.D., and Ekman, P. (1975). Facial Areas and Emotional Information. *Journal of Communication. 25 (2),* 21-29.

Diamond, R, and Carey, S. (1986). Why Faces Are and Are Not Special: An Effect of Expertise. *Q J Exp Psychol: Gen, 115 (2),* 107-117

Dodson, C.S., Johnson, M.K., and Schooler, J.W. (1997). The Verbal Overshadowing Effect: Why Descriptions Impair Face Recognition. *Mem Cognit. 25 (2),* 129-139.

Ekman, P. (1992). Are There Basic Emotions? *Psychological Review, 99*, 550-553.

Johnston, J.C., and McClelland, J.L. (1974). Perception of Letters in Words: Seek Not and Ye Shall Find. *Science, 184*, 1192-1194.

Pelli, D., and Farell, B. (1999). Why Use Noise*? J Opt Soc Am A Opt Image Sci Vis, 16 (3),* 647-653.

Schwartz, O., Bayer, H. and Pelli, D. (1997) Features, Frequencies, and Facial Expressions. Poster presented at Center for Neural Science, New York University. September 17, 1997.

Tanaka J.W., and Farah, M.J. (1993). Parts and Wholes in Face Recognition. *Q J Exp Psychol [A], 46 (2)*, 225-245