Pupillary Responses of Asian Observers in Discriminating Real from Fake Smiles: a Preliminary Study

M.Z. Hossain[1], T.D. Gedeon[1], R. Sankaranarayana[1], D. Apthorp[2], A. Dawel[2]

[1]Research School of Computer Science, Australian National University, Canberra, Australia

[2]Research School of Psychology, Australian National University, Canberra, Australia

Abstract

This work is a part of a research effort on detecting the differences between human responses to real and fake smiles, by watching observers of different smiling video stimuli (displayer's smiles). Pupil diameter was recorded from 10 Asian observers (6 males and 4 females), while watching 9 real smile stimuli and 10 fake smile stimuli. The preliminary analysis on pupil data revealed that the pupil size increased more for fake smile stimuli as compared to real smile stimuli in the case of male observers, and vice versa in the case of female observers. To the best of our knowledge, most of the work on discriminating real from fake smiles was analyzed from the smile displayer's point of view. Our results were found from observers' pupillary responses. If we consider fake smiles as negative emotions and real smiles as positive emotions, then this outcome may be comparable with previous findings on pupil dilation. Our results show that analyzing an observer's pupillary responses may help to understand the displayer's actual state of mind.

Key Words: Real Smile, Fake Smile, Pupil Size Variation, Asian Observers, Video Stimuli

Introduction

Human smiling would appear to be one of the most complex, but seemingly straightforward facial displays. Although smiles reflect positive affect [1], this can be from a variety of emotions [2]. People do not smile only for happiness, but also when they are socially anxious, embarrassed, depressed, surprised and even more [3]. In many cultures, cheerfulness is mandatory and workers are required to smile as part of their jobs. Furthermore, it has been noted that smiling is less a sign of original emotion than a social display meant for others [4]. In this paper, we refer to posed, acted and non-happy smiles as fake smiles and only consider smiles grown from happiness as real smiles. In this context, understanding the meaning of a smile can help us to know a displayer's actual state of mind.

In the past, Valstar et. al. [5] distinguished real from fake smiles by analyzing 202 smiling videos directly. The smiling characteristics of virtual agents were studied in [6]. The displayers' perceived meanings regarding smiles were found to be related to specific characteristics [7]. The dynamics and morphological characteristics of displayers' smiles were analyzed to disambiguate real from fake smiles in face to face interaction [3]. They addressed displayers' smiling characteristics to discriminate real from fake smiles. In our work we try to analyze an observer's pupillary responses while they watch a (video of a) displayer's smile.

The human pupil may dilate for various reasons, including memory load, stress, pain, different emotional stimuli and so on [8]. Chapman et al. [9] reported that pupil dilation changes significantly with the increase of pain intensity. Kang et al. [10] observed that pupil dilation was lager when people were more curious about the answer. Wang et al. [11] found that pupil dilation was proportional to the 'size' of a lie. The pupil size variation also reflects the judgments of males on females and vice versa. For example, the pupil size was increased when observers were shown facial stimuli of opposite sex, and larger pupils of females evoked more positive feelings in males [12]. Thus measuring pupillary responses would offer a good method for smile detection, because no sensors need to be attached to the observer, and certainly not to the displayer.

The above mentioned findings on pupillary responses suggest the measurement of pupil size may be useful to discriminate real from fake smiles. Our study using observers' pupillary responses actually reflect some of how and what a person is thinking about another's smiling behavior. This preliminary study offered us a good outcome in discriminating real from fake smiles with an indication of gender differences. We found evidence that Asian males differ from Asian females according to their pupillary judgments.

Materials and Methods

Observers

Ten healthy, Asian-background right-handed graduate students were participants as observers of video stimuli, with a mean age of 29.8 ± 4.8 for males, and 34.7 ± 4.8 for females (mean ± SD). Each observer had normal or corrected to normal vision, and provided their written consent prior to his/her voluntarily participation. An approval from the Australian National University's Human Research Ethics Committee was granted for this study.

Stimuli

It is well known that video is more effective and easier than static images for human recognition purposes [17]. For that reason, twenty-five video stimuli were collected from three benchmark databases: ten from AFEW [13], five from MMI [14] and, ten from MAHNOB (five from HCI and five from Laughter) [15-16]. The collected video stimuli were processed using MATLAB R2014b to convert to grey scale, mp4 format and each lasting 10s from smile onset. This paper is based on the analysis of collected pupillary responses, while observing 10 AFEW and 9 MAHNOB video stimuli. Luminance of each stimulus over the frames in the videos was computed using the MATLAB SHINE toolbox [18]. The ten stimuli (mean luminance = 70.2 and $Standard\ deviation$ = 4.6) of AFEW database were classified as fake smiles, because these were acted smiles by professional actors. On the other hand, participants' positive emotions (happiness) were elicited in the case of MAHNOB stimuli (mean luminance = 75.8 and $standard\ deviation$ = 2.1) and were treated as real smiles. In the case of other 6 videos: we could find no clear specification or explanation in the source for one stimulus collected from the MAHNOB Laughter database so we could not classify it as fake or real; and the luminance of the MMI videos were much higher than the rest. For this reason, they were not considered for analysis purposes. The remaining videos had luminance which varied within 6 ALU (Arbitrary Linear Unit) for each stimulus. This consistent range allows us to confidently remove the effect of luminance from the pupil dilation.

Procedure

The observer was introduced to the laboratory and comfortably seated in a static chair in front of a 17.1 inch computer monitor. Their chair was moved forward or backwards to adjust the distance between the chair and eye tracker. The participant (observer) was told that their eye movements would be recorded during video stimuli presentation and instructed to limit their body movements to reduce unwanted artefacts in the signals.

The observer was instructed to track nine fixation points, which were displayed in the monitor for calibration purposes. The experiment was started with a short introduction of fake versus real smiles and finished with feedback through a web-based questionnaire. The pupil size variation of the observers was recorded from both eyes with a facelab (Seeing Machines) remote eye-tracker system. The sampling rate of the system is 60 Hz. Every stimulus presentation was followed by a five point (-2 to +2) Likert scale (see Figure 1) where the observer rated his/her experience as evoked by the stimuli using a computer mouse. The negative and positive ends of the scale represented fake and real smile stimuli respectively. The centre of the scale represented confusion. The ends and centre of the scales were labelled appropriately. Each viewing stimulus to the observer was rated by this scale and documented in a text file.

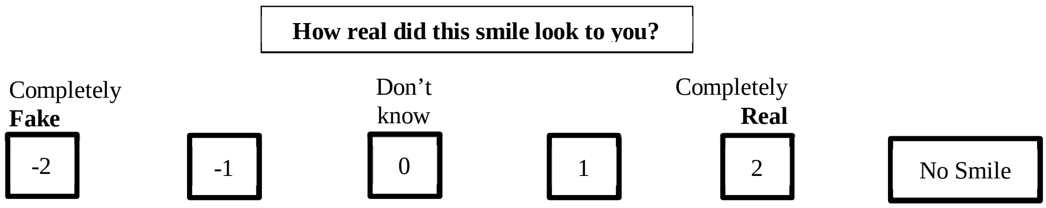

Figure 1. Five point Likert scale to accumulate the observer's self-report.

Data Processing

Samples where the pupil was obscured due to blinking were measured as zero by the eye tracking system, and cubic spline interpolation was applied to reconstruct the pupil size. Then, 10-point Hann moving window average was used to smooth the pupil signal. The minimum pupil diameter of each observer was subtracted from each pupil diameter of that observer, to reduce the effect of age among observers and to normalise for luminance in the videos which all had a consistent 6 ALU luminance range. The final pupil signal consisted of baseline corrected data averaged from both eyes over the whole 10s video stimulus interval.

All the data was analysed with paired sample two-tailed permutation test to report on significance levels. This test is performed to assess the statistical significance on waveform differences and produces p values for any number of time points [22]. Two-way analysis of variance (ANOVA) also introduced to check the effect of smile type, effect of sex, and their interactions.

Results

The timeline analysis of the pupil data revealed a common trend for each stimulus. Figure 2 illustrates the time point average of pupil diameter over observers when viewing all video stimuli. The pupil constricted from stimulus onset and reached a minimum within 1-2s, after which a sharp dilation started and continued till 3-4s. Then, either a smooth dilation or constriction started and continued, which is sustained in a consistent range, until the end of our analysis window. It is worth noting that the trends were separated according to real and fake smiling stimuli from about the 3s of stimuli onset. The paired sample permutation tests showed that pupil dilation differed significantly for real smile stimuli as compared to fake smile stimuli ($t = 4.56, df = 9, p = 0.051$), while mostly lower p values ($p < 0.1$) are found between 8.62s and 8.67s.

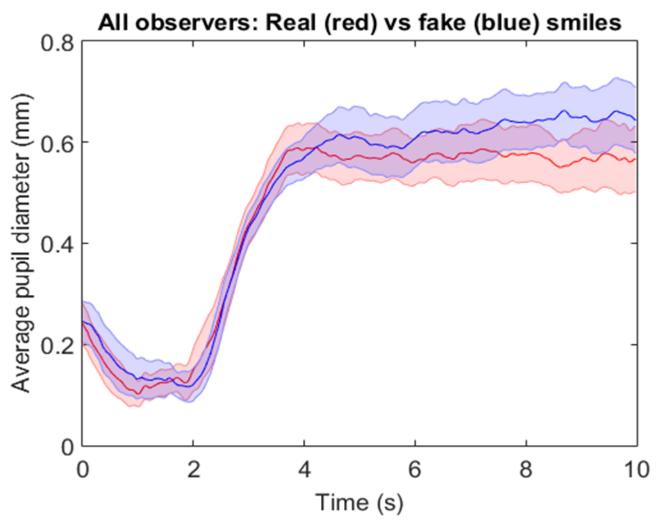

Figure 2. Average pupil diameter timelines for all stimuli over all observers.

The timeline of the pupil diameter changes was somewhat different for female and male observers (see Figures. 3(a) and 3(b)). The fake smile stimuli provoked the strongest pupil dilations for male observers, where the real smile stimuli provoked the strongest dilations for female observers. In the case of female observers, the difference of real from fake smiles started from 3s of stimulus onset and ended within 8s to 10s from stimulus onset (see Figure 3(a)). Paired Sample permutation test comparisons showed that the pupil diameter was not significantly smaller for fake smile stimuli than real smile stimuli ($t = 11.30, df = 3, p = 0.125$). In the case of male observers, the pupil dilates continuously from 4s of stimulus onset for fake smile stimuli, but no dilation or contraction for real smile stimuli (see Figure 3(b)). The pupil diameter was significantly larger for fake smile stimuli than real smile stimuli ($t = 6.41, df = 5, p = 0.063$), mostly lower p values ($p < 0.1$) are found from 7.75s to 7.93s and from 9.87s to 9.89s. The greater significant results may be found, if the number of observers is increased.

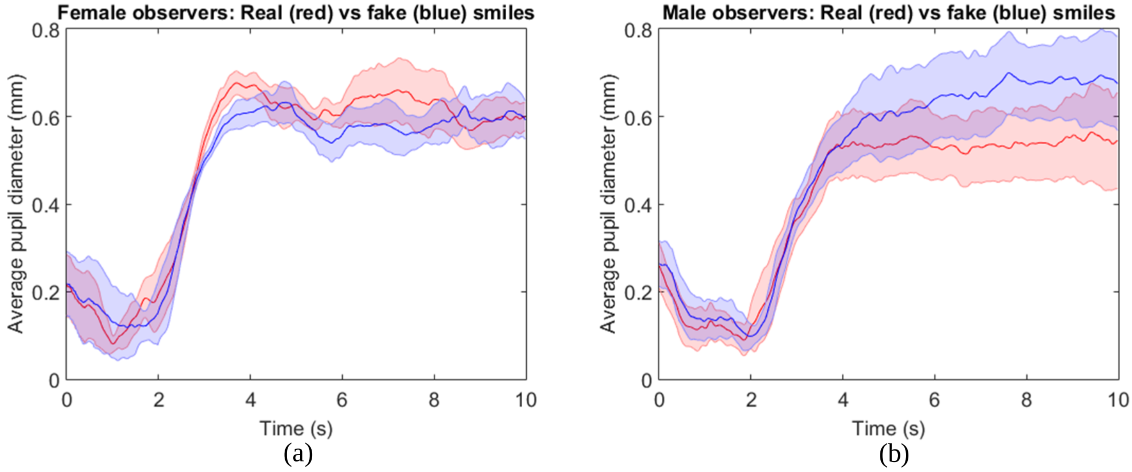

Figure 3. Average pupil diameter timelines for all stimuli over (a) female and (b) male observers.

We have also calculated the average value of individual observers' pupil size, while watching fake and real smile stimuli. Individual analysis on the whole presenting stimulus interval verified that larger pupil dilations were detected by male observers in the case of fake smile stimuli as compared to real smile stimuli and vice versa in the case of female observers (see Figure 4). In this Figure, error bars indicate standard deviations.

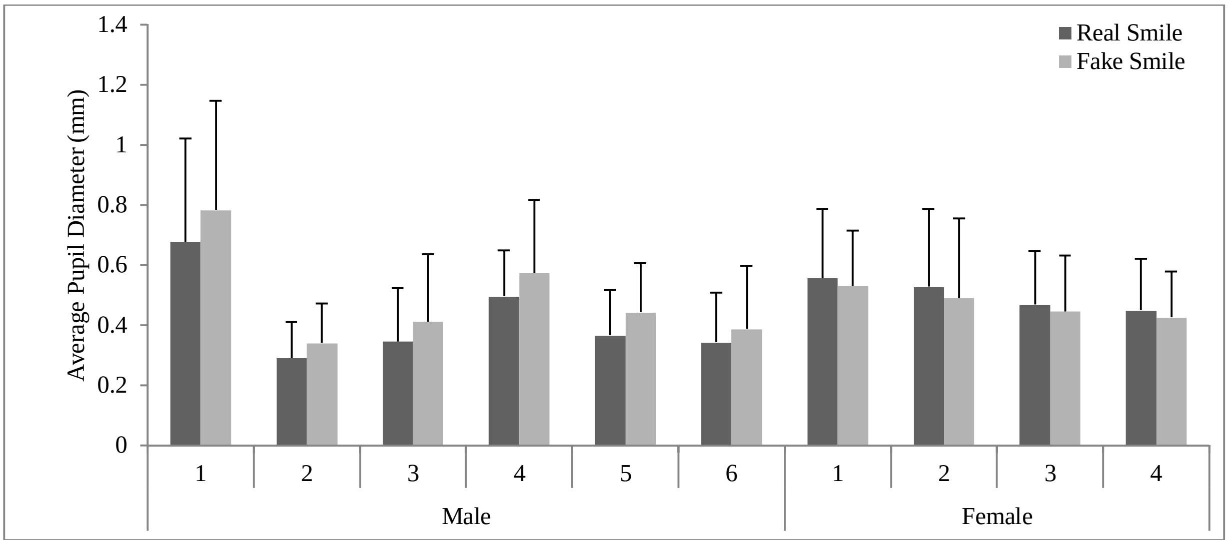

Figure 4. Average pupil diameter variation for individual observers.

Additional calculations on pupil size variation revealed that male observers have lower average values compared to female observers while watching real smile stimuli, and vice versa in the case of fake smile stimuli (see Figure 5). Here, error bars also denote standard deviations as in Figure 4.

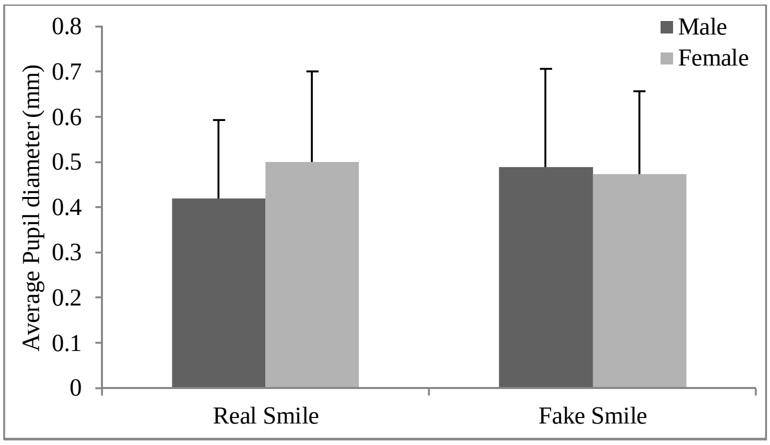

Figure 5. Average pupil diameter variation between female and male observers.

We also employed ANOVA test to check the effects of sex (male vs. female), smile type (real vs. fake), and their interactions on pupil data. The computed $F\ scores$ were 45.81, 20.42 and 104.17 for the effects of sex, smile type and their interactions respectively, that have significant impact ($p < 0.001$ for each case) on pupil data. We have also observed the $R\ squared$ value that explained 79% variations in the pupil data by the combinations of these effects and their interactions.

Discussion

Our results showed larger pupil size dilation in response to fake smile stimuli in the case of Asian male observers as compared to the pupil dilation in response to real smile stimuli and vice versa in case of Asian female observers. These differences were found during smiling video stimuli observation. The analyses on individual observers' pupil dilation also revealed that real smile stimuli caused different pupillary responses as compared to fake smile stimuli.

The visual inspection of the timeline curves discovered that the individual observers' average pupil dilation over stimuli coherently followed a similar type of curve. A sudden decrease in pupil size from the stimulus onset was followed by a sharp increase from about 2s to about 4s. Then, the observers' pupil responses were separated according to displayer's real and fake smile stimuli, from about 4s after the stimulus onset. These differences were observable clearly for male observers and averagely for female observers till the end of the viewing stimuli.

The outcomes of this experiment show that we can access the observer's different recognition and responses to real smile compared to fake smile stimuli via their pupil diameter. This is because the size of pupil is controlled by their autonomic nervous system which is known to respond to a person's emotional state [19]. According to evolutionary hypotheses of sex differences in emotional processing, the anatomy as well as function of the brain of the male is different (statistically) from the female [20]. Our results support the notion that pupil responses varied according to different emotional stimuli and have some similarity with recent findings that females respond more strongly to positive sounds and males respond more strongly to negative ones [21]. This is also similar to the average pupil size measured in [19] to discriminate negative from positive emotion. Furthermore, they found significant differences between emotions which are consistent with our outcomes considering fake versus real smile stimuli. We can also construct socio-biological explanations why males would respond more strongly to acted smiles while females would respond more strongly to real smiles.

The measurement of pupil size has important advantages compared to considering other physiological signal measurements. No sensors need be attached to the user, and pupil size may not be as sensitive to artefacts caused by various bodily movements. An important benefit of pupil size measurement is that it is not easy to control pupil size variation voluntarily and, thus, provides spontaneous and non-conscious outcomes. On the other hand, visually observable changes in facial expression can be masked, inhibited, exaggerated, and faked [2]. In this respect, pupil size measurement avoids these problems that are inherent in monitoring with video cameras. In this paper, we wanted to keep the video stimuli in a luminance range to make these as natural as possible. But careful attention is required to design a system in human-computer interaction for affective computing, because pupil size can vary according to light reflex, different stimulus parameters (e.g. visual and chemical), information-processing load and so on [12]. This control environment will be considered in our future research to compare with the present study and natural environments.

In conclusion, we analysed average pupil diameter timelines over 19 stimuli with 10 young adults from Asian backgrounds. These provide us an indication that the observer's pupil size varied differently for our fake smile stimuli as compared to real smile stimuli. Our outcome also provides evidence for gender differences. These findings suggest that by extracting and analysing an observer's pupillary responses, we can disambiguate real from fake smiles, based on the observer's innate and non-conscious recognition of the stimuli as representing fake or real smiles.

References

1. Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology: II. *Journal of Personality and Social Psychology*, **58**, 342–353.

2. Ekman, P., Friesen, W.V. (1982). Felt, false, and miserable smiles'. *Journal of Nonverbal Behavior*, **6**, 238–252.

3. Hoque, M., Morency, L. P., Picard, R. W., (2011). Are You Friendly or Just Polite? – Analysis of Smiles in Spontaneous Face-to-Face Interactions, Lecture Notes in Computer Science, 6974, 135-144.

4. Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin*, **115**, 268–287.

5. Valstar, M. F., Gunes H., Pantic, M. (2007). How to Distinguish Posed from Spontaneous Smiles using Geometric Features, Proceedings *of ACM International Conference on Multimodal Interfaces* (Nagoya, November 2007), 38-45.

6. Ochs, M., Niewiadmoski, R., Pelachaud, C. (2010). How a virtual agent should smile? Morphological and Dynamic Characteristics of virtual agent's smiles, *Proceedings of 10th International Conference on Intelligent Virtual Agents*, (Philadelphia, 20-22 September 2010), 427-440.

7. Ambadar, Z., Cohn, J. F., Reed, L. I. (2009). All Smiles are Not Created Equal: Morphology and Timing of Smiles Perceived as Amused, Polite, and Embarrassed/Nervous, *J Nonverbal Behav*. 33(1), 17–34.

8. Beatty, J. (1982). Task-Evoked Pupillary Responses, Processing Load, and the Structure of Processing Resources. *Psychological Bulletin*, **91(2)**, 276-292.

9. Chapman, C. R., Oka, S., Bradshaw, D. H., Jacobson, R. C., Donaldson, G. W. (1999). Phasic pupil dilation response to noxious stimulation in normal volunteers: Relationship to brain evoked potentials and pain report. *Psychophysiology*, **36(1),** 44-52.

10. Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T.-y., et al. (2009). The Wick in the Candle of Learning: Epistemic Curiosity Activates Reward Circuitry and Enhances Memory, *Psychological Science*, **20**, 963-973.

11. Wang, J. T.-y., Spezio, M., Camerer, C. F. (2009). Pinocchio's Pupil: Using Eye tracking and Pupil Dilation To Understand Truth-telling and Deception in Sender-Receiver Games, *American Economic Review*, **100(3),** 984-1007.

12. Hess, E.H., Petrovich, S.B., (1987). Pupillary behaviour in communication, In: *Siegman*, A.W., Feldstein, S. (Eds.), Nonverbal Behaviour and Communication. Erlbaum, Hillsdale, NJ, 327–348.

13. Dhall, A., Goecke, R., Joshi, J., Sikka K., Gedeon, T. (2014). The Second Emotion Recognition In The Wild Challenge, *16th ACM Int'l Conf. on Multimodal Interaction*, (Istanbul, September 2014).

14. Pantic, M., Valstar, M., Rademaker, R., Maat, L. (2005). Web-Based Database for Facial Expression Analysis, *Proc. IEEE Int'l Conf. Multimedia and Expo*, (Netherlands, 6-8 July 2005), 5.

15. Soleymani, M., Lichtenauer, J., Pun, T., Pantic, M. (2012). A multimodal database for affect recognition and implicit tagging, *IEEE Transactions on Affective Computing*, **3(1),** 42 – 55.

16. Petridisa, S., Martineza, B., Pantica, M. (2013). The MAHNOB Laughter database, *Image and Vision Computing*, **31(2),** 186–202.

17. Picard, R. W. (2000). Affective Computing, the MIT Press, Cambridge, Massachusetts, London, England, 175

18. Willenbockel, V., Sadr, J., Fiset, D., Horne, GO., Gosselin, F., Tanaka, JW. (2010). controlling low-level image properties: the SHINE toolbox, *Behav Res Methods*, **42(3),** 671-684.

19. Aracena, C., Basterrech, S., Snášel V., Velásquez, J. (2015). Neural Networks for emotion recognition based on eye tracking data, Proceedings of IEEE SMC Conference, (Hong Kong, 9-12 October 2015).

20. Cahill, L. (2005). His Brain, her brain, *Scientific American*, **292**, 40-47.

21. Partala T., Surakka, V. (2003). Pupil size variation as an indication of affective processing, *International Journal of Human-Computer Studies*, **59(1-2),** 185-198.

22. Blair, R. C., Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials, *Psychophysiology*, 30, 518-524.