# Gender Differences in Pupillary Responses of Asian Observers while Discriminating Real from Fake Stimuli

Damanvir Singh

**Abstract.** This work builds upon a previously completed study on detecting the pupillary response of humans while observing real and fake smiles. The previous study was done on 10 Asian observers (6 male and 4 female) and found that there were differences in the pupillary response of males and females depending on the stimuli (real/fake smile). They found that Asian females respond stronger to real stimuli and Asian males respond stronger to fake stimuli. This work further investigates the gender differences in pupillary responses of Asian males and females, to try and understand the unique relationship between gender and pupillary response to stimuli.

**Keywords:** Real Smile, Fake Smile, Pupil Size Variation, Asian Observers, Male vs. Female, Real Stimuli, Fake Stimuli

## Introduction

A study named "*Pupillary Responses of Asian Observers in Discriminating Real from Fake Smiles: a Preliminary Study*" was conducted to detect the differences between human responses to real and fake smiles (real/fake stimuli) [1]. The study was conducted on 10 Asian observers (6 males and 4 females) where all observers were shown 10-second-long videos (stimuli) of people smiling. The stimuli consisted of 9 real smiles and 10 fake smiles. The study showed that regardless of the gender, the pupil size of observers (given real or fake stimuli) always increased. The results from the study suggested that male observers' pupillary responses were greater than female observers' pupillary responses when exposed to fake stimuli, and female observers' pupillary responses were greater than male observers' when exposed to real stimuli. These findings support the findings from another study which found that "females respond more strongly to positive sounds and males respond more strongly to negative ones" [1].

In our work, the findings from the original study about pupillary responses being different in genders (given real/fake stimuli) are further investigated in an attempt to prove/disapprove this discovery. The results further suggested towards clear distinctions between genders in terms of pupillary responses to real/fake stimuli (smiles).

## Method

### Original Datasets

The datasets available from this study include "Asian_Observers_Pupil_Data" (AOPD), "LEPD" and "REPD". AOPD is the main dataset which contains the mean pupillary responses of observers to all real/fake stimuli. LEPD contains the pupillary responses of the observers' left eye for each real/fake stimulus (9 real and 10 fake). REPD contains the same data as LEPD for the same stimuli (9 real and 10 fake) but for the right eye. Due to the nature of the experiment, all the AOPD dataset is in the form of a time-series with records of pupillary responses over the span of ~10 seconds. LEPD and REPD are also in the form of time-series, however, they span for longer than 10 seconds. Since the original study only focused on the first 10 seconds of an observer's pupillary response, we will focus on the first 10 seconds of all datasets as well.

### Curated Dataset

The original AOPD dataset is extremely small with averaged pupillary observations for real and fake stimuli, for 10 observers. Such small amount of data is not sufficient for any classification or regression task as the model would be prone to overfitting. Therefore, a new dataset was created with LEPD and REPD combined. All observations from the two datasets were first transposed and then added to the new dataset, with some extra features. The LEPD and REPD datasets were transposed during the migration because by doing so, every single value of pupillary response could be used as a feature. The following is what the structure of the new dataset looks like:

| Name | Description |
|---|---|
| **person** | The observer.<br>(p2, p4, p5, p6, p7, p8, p9, p10, p11, p12) |
| **sex** | Gender of observer.<br>(0 = male, 1 = female) |
| **eye** | Eye of the observer.<br>(0 = left, 1 = right) |
| **dataset** | Which stimulus the observer watched.<br>(L1, L2 – L5, H1 – H5, A1 – A10) |
| **stimuli** | Real or fake stimuli.<br>(0 = real, 1 = fake) |
| **0** | Observer's pupillary response starting at 0 and ending at 540 (~10 seconds).<br>(0 – 1) |
| **…** | … |
| **540** | (0 – 1) |

## Results

The results from the original study were first reproduced using the AOPD dataset and further analysed. The initial analysis revealed that a highly influencing outlier is present in the dataset for males. Figure 1 shows that the male outlier's pupillary response (black line) is much greater than the average male pupillary response (blue line) for both real and fake stimuli.

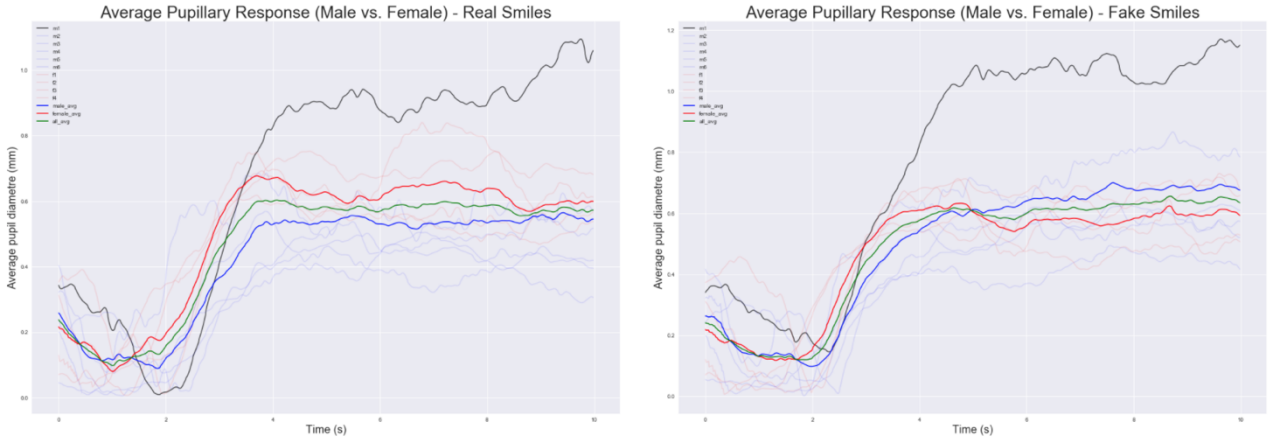


**Fig. 1.** Average pupillary response of all observers for real (left) and fake (right) stimuli. Male outlier highlighted in black.

It is visible that the male outlier is bringing-up the average of the male pupillary response by a considerable margin. In the case of average pupillary response of observers given fake stimuli (see Figure 1 right), the male and female averages are fairly close to each other, even though the male average is higher. However, once the male outlier is removed from the dataset (see Figure 2), the male average drops significantly, especially in the case of fake stimuli (see Figure 2 right). The female average is still higher than the male average for real smiles (see Figure 2 left).

In the results from Figure 1, the male average clearly overtakes the female average for fake stimuli (see Figure 1 right) just before the 5s mark. However, the new calculated average (after removing male outlier) for males is mostly lower or similar than the female average, apart from the small bump from ~7s - ~8.5s. The new male average for fake stimuli makes it extremely difficult to distinguish if there is a clear distinction between the average pupillary response of Asian males and females given fake stimuli (fake smiles).

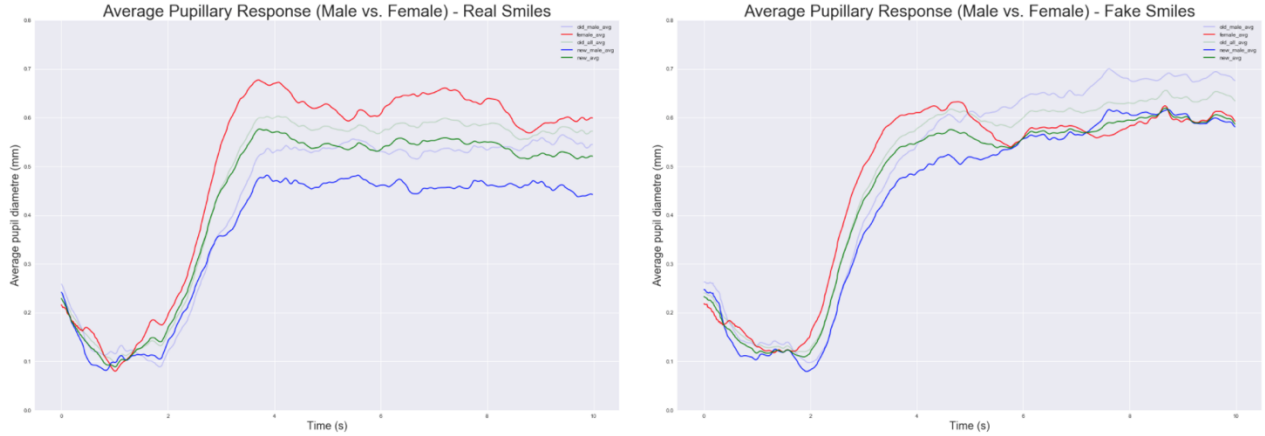


**Fig. 2.** Average pupillary response of all observers (without male outlier) for real (left) and fake (right) stimuli.

Figure 3 illustrates that the average pupillary response of Asian males is just lower than the Asian females'. The average pupillary response of Asian males is 0.43 (rounded to 2 decimal places) and Asian females is 0.47 (rounded to 2 decimal places). However, it must be noted how small the difference is, especially when the AOPD dataset was used for these results (Figures 1, 2 and 3). AOPD is a small dataset which contains only the averages of pupillary observations for all stimuli, for 10 observers.

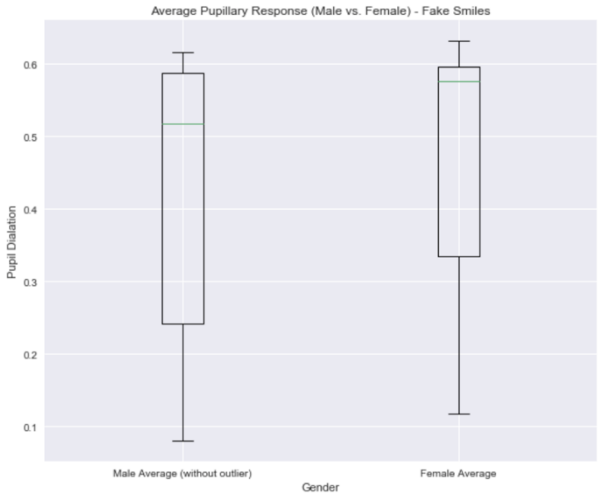


**Fig. 3.** Average Pupillary response (Male vs. Female) – Fake Stimuli.

Therefore, from this analysis, we have two findings a) the pupillary response of Asian females is greater than Asian males while exposed to real smiles (real stimuli) and b) we cannot sufficiently prove if the pupillary response of Asian males is greater or lower than Asian females while exposed to fake smiles (fake stimuli), since there is only a difference of 0.04 and the dataset used contains a small sample of observers (10).

To further examine if there is any clear distinction between the pupillary responses of Asian males and females, an attempt at classifying the gender of the observer based on their pupillary response to stimuli (real or fake smiles) was made, using a transformed AOPD dataset. The dataset was transformed by first, transposing it and then adding "sex" and "stimuli" as binary features. If any classification model can achieve a training and testing score close to 0.7, then that would indicate that there is a distinguishable difference in the pupillary responses of Asian males and females.

The dataset was first split into a training set (80%) and a testing test (20%). A basic Decision Tree Classifier (DTC) was then trained on the training data. The model returned a perfect training score of 1.0 with a testing score of 0.5 (rounded to 2 decimal place). This is a poor result as evident from the perfect training score and much lower testing score; the model is overfitting. To confirm if the model was overfitting, a Logistic Regression Classifier (LRC) was trained on the

same training data. The LRC also returned a perfect training score of 1.0 and a testing score of 0.5 which confirmed that this model was overfitting as well.

To avoid overfitting, the LRC was then trained on the curated AOPD dataset which contains pupillary responses of all observer for each stimulus. The model returned a training score of 0.61 and a testing score of 0.59. The model did not overfit, however, it is not accurate either.

## Discussion

The results from this study support the original study's findings that Asian females do respond stronger than Asian males in term of pupillary responses to real stimuli (smiles). However, the results also suggest the opposite to one of the other findings from the original study. It is appearing that Asian males do not respond stronger than Asian females in terms of pupillary responses to fake stimuli (smiles). The original study talked about the possibility of this relation, however, a more in-depth analysis of the dataset revealed that an outlier in the male observers was heavily influencing the male pupillary average. This 1 male observer out of a total of 10 observers had the largest increase in pupil size while observing each stimulus. Due to the dataset only having 10 observers, 1 outlier can easily influence the results.

The LRC model trained on the curated AOPD dataset did not return accurate results. Its testing score was only 0.59 where a score of 0.7 would have at least suggested a possibility of there being some clear distinctions in the pupillary responses of Asian males and females. However, since it is evident that Asian females do respond stronger than males to real stimuli (smiles), this does suggest that the same or the opposite could be true if given fake stimuli (stimuli).

## Future Work

Due to the limited data of 10 observers, none of the findings can be conclusively proved. Therefore, the study should be redone with a much larger dataset of observers. The new study should focus on other demographics and details including country of origin, ethnicity, condition of eyesight and etc. A detailed dataset would allow future studies to go further in-depth with analysis or machine learning tasks.

Both DTC and LRC were used on the AOPD dataset but only the LRC model was used for the curated AOPD dataset. The choice was made due to logistic regression being useful for two-class classification and help map pupillary response data and stimuli type (real/fake) to the gender of the observer. The accuracy of the final LRC model could potentially be improved through Bimodal Distribution Removal as it is a technique used for cleaning up noisy datasets to improve generalization [2].

## References


1. Hossain, M.Z., Gedeon, T.D., Sankaranarayana, R., Apthorp, D., Dawel. A.:Pupillary Responses of Asian Observers in Discriminating Real from Fake Smiles: a Preliminary Study**. (**2016)
2. Partala, T., Surakka, V.:Pupil size variation as an indication of affective processing. International Journal of Human-Computer Studies. 59(1-2), 185-198. (2003)