Towards Automated Decision Making in Dating Apps Through Pupillary Responses

Jan Ehlers[1][a], Sebastian Laverde Alfonso[2] and Arup Mazumder[3]

[1] *Department of Computer Science, Bauhaus-Universität Weimar, Schwanseestr. 143, Weimar, Germany*

[2] *Unstructured.io, 7580 Horseshoe Bar Rd., Loomis, CA, U.S.A.*

[3] *University of Rhode Island, 45 Upper College Rd, Kingston, RI, U.S.A.*

Keywords: Cognitive Pupillometry, Affective Processing, Arousal, Decision Making, Automation.

Abstract: Decision making is a multi-stage process that involves a series of rational evaluations. Recently, bodily arousal has been identified as a factor that mediates individual decisions, particularly during partner selection. The current study investigates pupil size changes in response to facial images of the opposite sex from controlled eye-tracking data (Experiment 1) and by reading out signals from front-facing smartphone cameras in noisy environments (Experiment 2). The aim is to enable automated decision-making in dating apps using arousal-based information. The rating results showed a tendency towards moderate evaluations when coping with facial attractiveness, while pupil diameter did not clearly discriminate between all four rating categories. However, a ROCKET model was trained on the pupil data from Experiment 1 with a prediction accuracy of 77% for binary classification of clearly preferred and non-preferred images. Ambiguous responses will therefore continue to pose a problem for cognition-aware systems. Capturing pupil diameter from mobile phone cameras resulted in a high proportion of inadequate recordings, probably due to a lack of experimental control. However, an overly systematic approach should run contrary to the intended scenario of lifelike mobile dating app usage.

1 INTRODUCTION

Over the past decades, online dating has become increasingly popular with more than 200 million active users by the end of 2019 (Castro and Barrada, 2020). In recent years, however, regular use has shifted to mobile dating apps that feature real-time location based services and typically require users to decide for or against a candidate's photos by swiping left (rejection) or right (acceptation) across the screen (Sawyer et al., 2018; Wu and Trottier, 2022). This image-based approach is based on the view that physical appearance exerts a decisive influence on partner selection (Berscheid and Walster, 1974). Tallying with this, evolutionary psychology suggests that phenotypic features of the face (more than body characteristics) are particularly well suited to predicting overall attraction (Fink and Penton-Voak, 2002). And although the relevance of facial cues appears to differ between the sexes (Buss, 2016), both men and women will pay more attention to attractive faces compared to unattractive ones (Dai et al., 2010; Hahn and Perrett, 2014). Thus, today's dating apps meet our needs during partner search with the result that they are used frequently and in almost any environment. The binary selection mechanism (reject vs accept) features a straightforward design, but manual input can become tiring during long-term usage and even prone to input errors while on the go or in case of one-handed operation.

Decision making - in dating apps but also in general - constitutes a multistage process that involves a series of (rational) judgments. However, when the stakes are high and solutions have long-term consequences, individual choices are strongly influenced by visceral factors (due to anxiety) or drive states (e.g. sexual desire) (Damasio, 1996). Especially sexual arousal has recently gained recognition as a situational factor mediating individual mate choice (Skakoon-Sparling et al., 2016). The current work explores arousal-based changes in response to facial images of the opposite sex. To quantify arousal we aim for a peripheral measure that can be administered remotely and is applicable in different settings. As

[a] https://orcid.org/0000-0002-4475-2349

Ehlers, J., Alfonso, S. and Mazumder, A.
Towards Automated Decision Making in Dating Apps Through Pupillary Responses.
DOI: 10.5220/0012471100003660

In *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2024) - Volume 1: GRAPP, HUCAPP and IVAPP*, pages 522-529
ISBN: 978-989-758-679-8; ISSN: 2184-4321

indicated above, visual cognition is ubiquitous during partner selection, suggesting eye-based data to constitute a promising channel for incorporating bodily activation.

Pupil diameter primarily reflects the nervous system response to ambient luminance ('pupillary light reflex') (Kardon, 1995). However, in controlled lighting environments, pupil size changes are supposed to reflect processes underlying cognitive functions such as attention (Einhäuser et al., 2008) or working memory (Kucewicz et al., 2018; Grimmer et al., 2021). Affective pupillometry relies on innervations from the autonomic nervous system and indicates that pupil diameter increases with general bodily arousal (Ehlers et al., 2016; Ehlers et al., 2018). Liao et al. (Liao et al., 2021) as well as Pronk et al. (Pronk et al., 2021) report an increase in pupil diameter when viewing attractively rated faces. Laeng and Falkenberg (Laeng and Falkenberg, 2007) observe strong dilations in female participants to pictures of their sexual partners. Tallying with this, Rieger and Savin-Williams (Rieger and Savin-Williams, 2012) identify pupil diameter as an indicator of sexual orientation, with heterosexual men showing greater dilations to images of women than vice versa. Bernick et al. (Bernick et al., 1971) report pupil diameter to increase with sexual arousal and conclude that pupil size changes can be applied to distinguish between sexual arousal and more general states of activation. However, studies on affective pupillometry tend to apply cumulative results over a period of time after stimulus-onset and to ignore time-dependent changes in pupil diameter. And although research has focused almost exclusively on laboratory experiments, results are often confounded with stimulus luminance or contrast (Liao et al., 2021).

The present study builds on this research. The overall aim is to enable arousal-based decision making via pupil responses in order to automate the use of dating apps. Therefore, we investigate pupil size changes in response to facial images of the opposite sex in a fully controlled laboratory experiment (Experiment 1). A state-of-the-art machine learning model is applied to predict individual choices by classifying pupil responses. Experiment 2 is a field study that adopts the previous design but reads out pupil diameter from front-facing smartphone cameras. The results will show the extent to which the laboratory findings can be transferred to real-life conditions.

2 EXPERIMENT 1

Experiment 1 investigates changes in pupil diameter when viewing facial stimuli of the opposite sex. Environmental factors are kept constant in order to determine influences from cognitive and affective processing while coping with facial attractiveness.

2.1 Methods

The following subsections provide information on the experimental design and specify details on data processing and sample characteristics.

2.1.1 Design and Procedure

Experiment 1 was carried out in a laboratory room featuring a constant illumination of 300 lux. Participants received an introduction to the research subject and were then seated in an upright position in front of a computer monitor. A stationary Eye-tracker was attached to the lower edge of the screen at a distance of 90cm.

We applied a within-subject design by providing a set of 60 high-quality AI-generated images that was randomly taken from a pool of 400 stimuli (Fig. 1A) (Karras et al., 2019). Each image showed a human face that looked directly into the camera with no other object in the foreground or background. Our volunteers identified as heterosexual and were asked to passively view 30 images of the opposite sex. Each image was presented for five seconds against a light grey background and preceded by a one-second baseline recording. Baseline acquisition involved the presentation of a pixel scrambled version of the subsequent image to control for differences in luminance and contrast (Fig.1B).

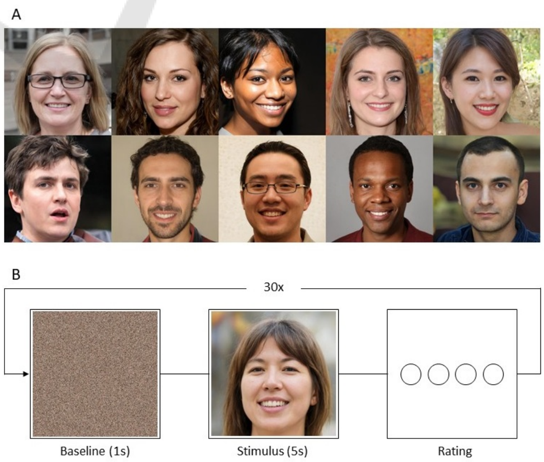

Figure 1: A: AI-generated images. B: Trial procedure during the image viewing task.

Subsequent to each image presentation, volunteers were asked to evaluate the corresponding face on a four-point scale ranging from (1) “Do not like at all”, (2) “Rather do not like” to (3) “Rather like” and (4) “Totally like”. After the individual assessment participants proceeded to the next trial via button press. Pupil diameter was recorded throughout the entire experiment (approx. five minutes).

2.1.2 Data Processing

Eye-tracking data was obtained binocular with a sampling rate of 60 Hz. According to (Mantiuk et al., 2012), data from the left eye were taken to determine changes in pupil diameter. Following (Leys et al., 2013), we applied the median absolute deviation (MAD) to correct for outliers. Recordings mainly included blink artefacts, i.e. segments of missing data as a result of eyelid closure. Exemplary raw data trials are depicted in Figure 2. Outliers were removed from the data set and replaced via linear interpolation. For purposes of presentation, eye-tracking data were smoothed using a moving window of 25 data points.

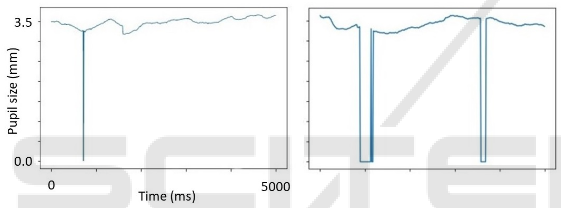

Figure 2: Exemplary raw data trials featuring pupil size changes during the image viewing task. Vertical lines indicate blink artefacts (segments of missing data).

2.1.3 Apparatus

Pupil size changes were measured with the Tobii Pro Nano Eye-tracker, Tobii Pro Lab Software (v. 1.123) was applied for data recording and storage. The Eye-tracker was attached to a 22-inch Dell P2213 monitor with a resolution of 1680 x 1050 pixels. Statistical analyses were carried out using the open-source software JASP (v.0.12.2) (Love et al., 2019).

2.1.4 Particiants

16 volunteers (five female, M: 28 years (SD: 3)), all of them students of computer science at Bauhaus-Universität, participated in Experiment 1. Information on regular medication was not collected, however, volunteers reported no history of head injury and no neurological or psychiatric disorders. Written informed consent was obtained prior to the experiment. All measurements were performed in accordance with the Declaration of Helsinki and approved by a local ethics committee (Ulm University).

2.2 Results

The following subsections summarize our findings on rating decisions, pupil size changes and data modelling.

2.2.1 Ratings

Participants were asked to evaluate 30 facial images of the opposite sex on basis of a four-point scale. 16% of all cases were rated as (1) “Do not like at all” whereas the majority of faces were either assessed as (2) “Rather do not like” (35%) or (3) “Rather like” (34%). 14 % were classified as (4) “Totally like”. Four participants allocated images to categories 2 and 3 only.

2.2.2 Pupillometry

Absolute pupil diameter differs between participants. To make individual responses comparable, we averaged across the last 30 data points of baseline recording and subtracted the result from each value during the five seconds of image viewing (Grimmer et al., 2021). P-values of the Shapiro-Wilk test indicated that pupillary data was well modelled by a normal distribution.

The first half second after stimulus-onset involved a pupil constriction that appeared comparably strong while viewing disliked facial images (category 1). Hereafter, pupil diameter increased and diverged according to feelings of affection after approx. two seconds. Averaged responses to preferred images (category 3 and 4) continued to increase whereas pupil diameter yielded a temporary constriction while viewing (rather) disliked images (category 1 and 2) (Fig. 3, left). Figure 4 depicts exemplary responses (individual averages) to images that were rated as ”Do not like at all” (category 1, left) and ”Totally like” (category 4, right). As can be seen, variability between individuals increased with increasing viewing duration, featuring similar courses only during the first second after stimulus-onset.

Descriptive statistics suggested a tendency towards larger diameters in response to attractively rated faces (Fig. 3, right). Facial images of category 1 (“Do not like at all”) involved an increase of 0.09mm (SE: 0.03) compared to baseline mean whereas category 2 stimuli (“Rather do not like”) induced diameter changes of 0.11mm (SE: 0.02). Averaged values during image viewing from category 3 (“Rather like”) proved to be similar (M: 0.11mm, SE: 0.02) whereas facial images from category 4 (“Totally like”) were associated with the strongest increase in pupil diameter (M: 0.13mm, SE: 0.02). However,

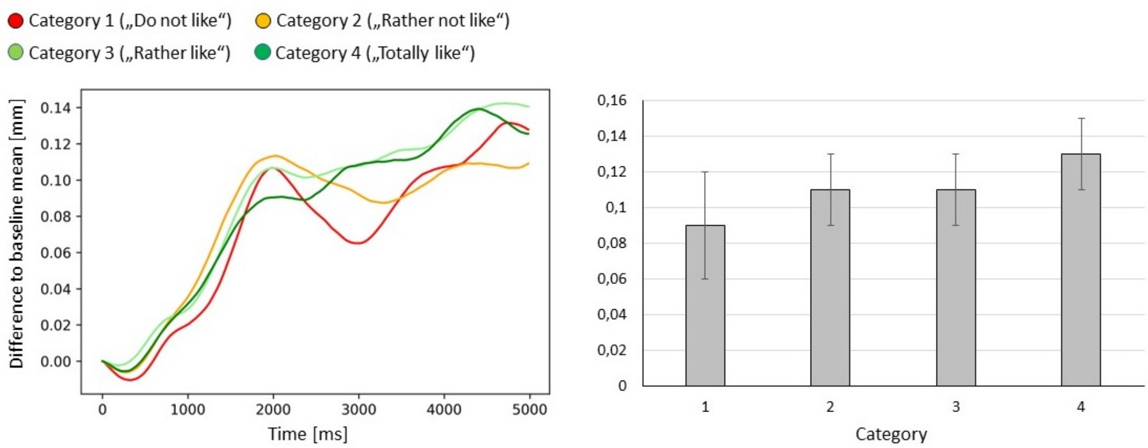

Figure 3: Total averages (n=16) of pupil diameter while viewing images of the opposite sex. Values depict differences from baseline mean and are grouped according to individual preferences. Left: Zero on the x-axis indicates stimulus-onset. For purposes of presentation, signal course changes are smoothed using a moving window of 25 data points.

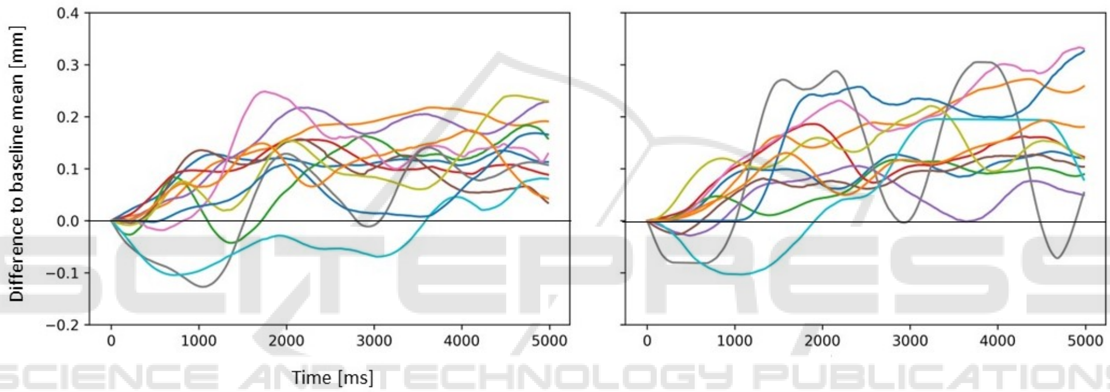

Figure 4: Individual averages (n=11) of pupil diameter while viewing images of the opposite sex that were rated as ”Do not like at all” (category 1, left) and ”Totally like” (category 4, right). Values depict differences from baseline mean; zero on the x-axis indicates stimulus-onset. For purposes of presentation, signal course changes are smoothed using a moving window of 25 data points.

a one-way repeated measures ANOVA (Greenhouse-Geisser corrected) indicated no considerable differences ($F(2.541, 27.950) = 1.149$, $p = 0.34$).

In a further step, we compared pupillary responses that involved clear preferences (category 1 vs category 4) via a paired samples t-test and observed a strong trend towards larger dilations while viewing attractively rated faces ($t(11) = -1.536$, $p = .08$, $d = -0.44$). However, this excludes trials that entail some degree of uncertainty (category 2 and 3). Following the selection technique from mobile dating apps, we regrouped pupillary data into binary samples by merging category 1 and 2 as well as category 3 and 4 and computed weighted averages due to different number of data sets. In a subsequent t-test, the effect of category was no longer visible ($t(11) = -0.184$, $p = .43$). Four participants were excluded from inferential statistics as they allocated images only to categories 2 and 3.

2.2.3 Time Series Classifciation

Machine learning (ML) methods enable to identify complex correlations in time series data by figuring out patterns and extracting useful features (Ruiz et al., 2021). We implemented an ML-model that was trained on the pupillary data from Experiment 1. To reduce ambiguity, we addressed data from only those trials that indicated clear preferences (category 1 and 4). Along with the outlier correction this resulted in a total of 79% data reduction. However, data quality was considerably increased and should be particularly suited for the model to learn from it. The remaining data were standardized by re-scaling the distribution (M:0, SD:1), series lengths were normalized to adjust the number of decimals.

We applied the state-of-the-art ROCKET model (Random Convolutional Kernel Transform) that features a high classification accuracy by transforming time series on basis of random convolutional kernels and training a linear classifier (Dempster et al., 2020). Data was split using Stratisfied Kfold (k=5) to ensure a balanced proportion of both labels resulting in 102 samples to train ROCKET and 26 samples to test it. Given the small data sets we selected the evaluation strategy LOO (LeaveOneOut) built in a Ridge Classifier. Training was carried out five times using different random seeds on top of Stratisfied Kfold to capture the best model (25 models trained in total). Table 1 indicates relevant metrics for the best obtained univariate ROCKET model featuring a prediction accuracy of at least 77% during binary classification (category 1 vs 4).

2.3 Discussion

Volunteers classified 30 facial images into four categories according to their individual preferences. Results suggest a tendency towards moderate ratings when dealing with facial attractiveness. Dislikes (category 1) as well as strong likes (category 4) account for only a small proportion of cases (around 30%), suggesting that ambiguous responses and a degree of uncertainty pose an ongoing challenge for cognition-aware systems. It could be argued that our five female volunteers saw different faces due to the presentation of images of the opposite-sex. This is a valid objection and further research may balance the sex ratio or focus on one sex only. On the other hand, recent studies have found no differences in the pupillary responses of heterosexual men and women to images of the opposite sex (Attard-Johnson et al., 2021).

Due to baseline referencing and controlled environmental conditions, current changes in pupil size cannot be explained through low-level factors such as luminance or contrast. During early processing stages (up to 1 second) pupil diameter was characterized by low variability between individuals and likely shaped by cognitive processes that underlying attention (Mathôt et al., 2014) or memory processes (Naber et al., 2013) during face recognition. Two seconds after stimulus onset, pupil responses diverged according to attractiveness ratings. While pupil reactions to unfavourable images (category 1 and 2) decreased, there was a strong and sustained autonomous response to favourable images (category 3 and 4). These findings replicate previous studies, indicating pupil diameter to increase when viewing preferred facial images of the opposite sex (Bernick et al., 1971; Laeng and Falkenberg, 2007; Liao et al., 2021; Rieger and Savin-Williams, 2012). However, arousal-related changes in pupil size did not clearly distinguish between all four categories. Rather, it appeared that pupil diameter was only sensitive to discriminating physical states of clear rejection from those of clear acceptance. More recently, Pronk et al. (Pronk et al., 2021) conducted a similar study with a larger number of participants and confirmed a positive association between pupil diameter and the acceptance of a hypothetical partner. Larger samples allow for more precise estimation of statistical properties and greater power; therefore it is reasonable to assume that a larger sample would yield similar results in the present case.

The ROCKET model yielded a promising performance. Although it's slightly biased towards the most positive category (4), the model predicted results for 77% of unseen data correctly. For comparison only: A dummy classifier randomly classifying samples or constantly predicting the same class provided an accuracy of no better than 57%.

3 EXPERIMENT 2

Experiment 2 adapts the design from Experiment 1. However, the setting is transferred from fully controlled laboratory conditions to various domestic locations. Moreover, video-based eye-tracking is replaced by recordings from unmodified smartphone cameras.

3.1 Methods

The following subsections provide information on the experimental design and specify details on technical implementation, data processing and sample characteristics.

Table 1: Classification metrics of the univariate ROCKET model for pupillary responses to facial images of category 1 (”Do not like at all”) and category 4 (”Totally like”).

	Precision	Recall	F1-score	Number of samples
1 (category 4)	71%	92%	80%	13
0 (category 1)	89%	92%	73%	13
Accuracy			77%	26
Macro avg	80%	92%	76%	26
Weighted avg	80%	92%	76%	26

3.1.1 Design and Procedure

Subsequent to Experiment 1 we collected another sample of volunteers in order to pilot a field study in domestic environments. A new set of images (30 male and 30 female faces) was selected from the same pool of 400 AI-generated pictures (Karras et al., 2019). An Android package (APK) including all relevant files was emailed to the participants. They were asked to install it on their smartphones and grant access to the mobile storage media in order to acquire video data from the front-facing camera. Experimental design and task procedure corresponded to the setting depicted in 2.1.1. Volunteers were asked to pick a quiet moment and sit comfortably while holding the phone steady in front of their face. However, task processing was not supervised and influencing factors like viewing angle, distance from screen or ambient lighting could not be controlled. Upon completion of the viewing task, all video files were uploaded to a server.

3.1.2 Technical Implementation

Pupillary data was captured from video files of the front-facing camera of mobile phones. To obtain the center of the pupil, we applied the deep learning framework GazeML which constitutes an established method to recognize landmarks of the eyes (Park et al., 2018). RGB values were employed for pupil detection and comparison with the closest region's color codes. To detect a similar colored pixel, we contrasted nearby regions until color values exceeded a predetermined range of black shading (RGB values 51,51,51).

3.1.3 Data Processing

Mobile devices featured varying frame rates and we normalized sampling to a constant value of 25 frames per second (fps). As during Experiment 1, the MAD algorithm was applied to detect outliers (Leys et al., 2013). However, while reading out pupillary data, relevant features could not entirely be captured frame to frame from video files which resulted in fragmentary data sets (Fig. 5). As a consequence, seven participants had to be removed from the data base due to insufficient recordings.

3.1.4 Apparaus

Pupil diameter was determined on basis of video files from the front-facing camera of latest Samsung smartphones. Android Studio, the development environment of Google's operating system, was used to design the front end of the mobile app. The Python flask

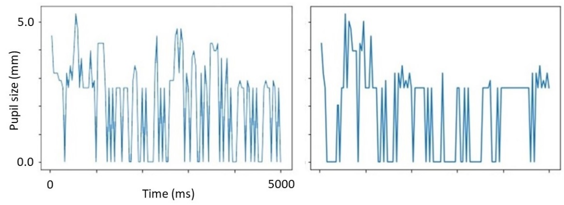

Figure 5: Exemplary raw data trials featuring pupil size changes during the image viewing task. Data were captured from video files of front-facing smartphone cameras. Vertical lines indicate segments of missing data).

framework was applied to create the backends' REST API.

3.1.5 Participants

12 male volunteers (M: 28 years, SD: 2) participated in Experiment 2. Seven data sets had to be excluded from pupillometric analyses due to unsuccessful recordings. Information on regular medication was not collected, however, volunteers reported no history of head injury and no neurological or psychiatric disorder. Written informed consent was obtained prior to the experiment. All measurements were performed in accordance with the Declaration of Helsinki.

3.2 Results

The following subsections summarize findings on rating decisions and pupil size changes.

3.2.1 Ratings

Participants were asked to evaluate 30 facial images of the opposite sex on basis of a four-point scale. 14% of all cases were rated as "Do not like at all" whereas the majority of faces were either assessed as "Rather do not like" (40%) or "Rather like" (29%). 17% of cases were classified as "Totally like".

3.2.2 Pupillometry

A large part of pupillary data from the video files was not analysable. Due to the small number of successful recordings (n=5), we forego the documentation of descriptive statistics and focus a qualitative evaluation. During the first second of task processing, averaged pupil size changes roughly corresponded to the eye-tracking data during Experiment 1, indicating pupil diameter to increase while viewing attractively rated faces (Fig. 6). Smartphone recordings then yielded strong responses to facial images from category 3 (approx. 0.3mm above baseline mean), whereas diameters fell below baseline level while viewing pictures from category 2. Pupillary reactions to non-preferred images (category 1) indicated a clear dilation whereas pictures from category 4 (”Totally like”) induced rather unspecific dynamics.

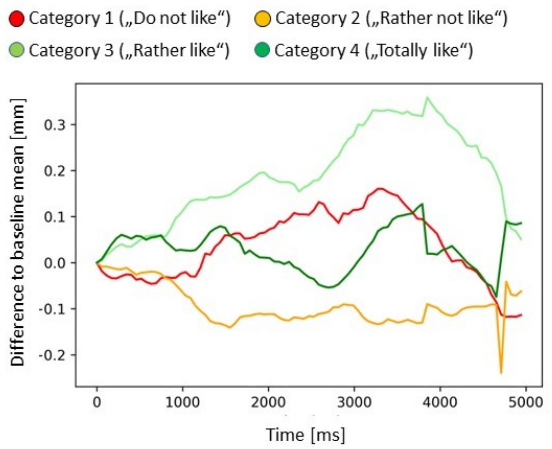

Figure 6: Total averages (n=5) of pupil diameter while viewing images of the opposite sex. Values depict differences from baseline mean and are grouped according to individual preferences; zero on the x-axis indicates stimulus-onset. Data were captured from video files of the mobile phone front-facing camera.

3.3 Discussion

We applied the viewing task from Experiment 1 in a noisy setting and during unsupervised task processing, while pupil data were captured by video files from front-facing smartphone cameras. Rating results from Experiment 1 were essentially replicated, indicating a tendency towards moderate ratings during attractiveness judgments. However, there was a large proportion of insuffient data for pupillometry. In addition, successful recordings (n=5) yielded inconsistent results, featuring both a sustained drop below baseline level (category 2) and a particularly strong response (category 3) when processing rather moderate stimuli.

Rafiqi et al. (Rafiqi et al., 2015) obtained pupil size changes from unmodified smartphone cameras in a fully controlled setting and reported high correlations with control measures from a certified Eye-tracker. Therefore, it can be assumed that our data loss was mainly due to processing issues, including low ambient lighting and unfavourable viewing distances or camera angles. In addition, the current sample consisted of volunteers from an international study program, including a large proportion of Indian students with rather dark eyes. This presents an additional challenge to the pupil detection algorithm, especially under unfavourable lighting conditions. The latter may be countered by environmental illumination compensation as recently introduced by Gollan et al. (Gollan and Ferscha, 2016). However, cognitive pupillometry based on digital cameras requires a certain degree of environmental control and a rather stationary camera position.

4 CONCLUSION

The present study takes a further step towards automated decision making in dating apps using arousal-based information. We applied pupillometry as an established method to map transient and long-lasting changes in underlying cognitive and affective processing. Under laboratory conditions, pupil diameter discriminated between states of clear rejection and clear acceptance. Accordingly, data modelling yielded promising results to enable proper automation of binary classifications. However, a controlled setting runs contrary to the intended scenario of life-like mobile dating app usage. Furthermore, the question of whether this type of automation is desirable in principle remains unanswered. A perceived loss of control during partner selection could lead to rejection of the application. Therefore, usability measures and qualitative data analyses should accompany further research to ensure a positive user experience.

REFERENCES

Attard-Johnson, J., Vasilev, M. R., Ó Ciardha, C., Bindemann, M., and Babchishin, K. M. (2021). Measurement of sexual interests with pupillary responses: A meta-analysis. *Archives of sexual behavior*, 50:3385–3411.

Bernick, N., Kling, A., and Borowitz, G. (1971). Physiologic differentiation of sexual arousal and anxiety. *Psychosomatic medicine*, 33(4):341–352.

Berscheid, E. and Walster, E. (1974). Physical attractiveness. In *Advances in experimental social psychology*, volume 7, pages 157–215. Elsevier.

Buss, D. M. (2016). *The evolution of desire: Strategies of human mating*. Hachette UK.

Castro, Á. and Barrada, J. R. (2020). Dating apps and their sociodemographic and psychosocial correlates: A systematic review. *International Journal of Environmental Research and Public Health*, 17(18):6500.

Dai, X., Brendl, C. M., and Ariely, D. (2010). Wanting, liking, and preference construction. *Emotion*, 10(3):324.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351(1346):1413–1420.

Dempster, A., Petitjean, F., and Webb, G. I. (2020). Rocket: exceptionally fast and accurate time series classification using random convolutional kernels. *Data Mining and Knowledge Discovery*, 34(5):1454–1495.

Ehlers, J., Strauch, C., Georgi, J., and Huckauf, A. (2016). Pupil size changes as an active information channel for biofeedback applications. *Applied psychophysiology and biofeedback*, 41:331–339.

Ehlers, J., Strauch, C., and Huckauf, A. (2018). A view to a click: Pupil size changes as input command in eyes-only human-computer interaction. *International journal of human-computer studies*, 119:28–34.

Einhäuser, W., Stout, J., Koch, C., and Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences*, 105(5):1704–1709.

Fink, B. and Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science*, 11(5):154–158.

Gollan, B. and Ferscha, A. (2016). Modeling pupil dilation as online input for estimation of cognitive load in non-laboratory attention-aware systems. *COGNITIVE*.

Grimmer, J., Simon, L., and Ehlers, J. (2021). The cognitive eye: Indexing oculomotor functions for mental workload assessment in cognition-aware systems. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–6.

Hahn, A. C. and Perrett, D. I. (2014). Neural and behavioral responses to attractiveness in adult and infant faces. *Neuroscience & Biobehavioral Reviews*, 46:591–603.

Kardon, R. (1995). Pupillary light reflex. *Current opinion in ophthalmology*, 6(6):20–26.

Karras, T., Laine, S., and Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4401–4410.

Kucewicz, M. T., Dolezal, J., Kremen, V., Berry, B. M., Miller, L. R., Magee, A. L., Fabian, V., and Worrell, G. A. (2018). Pupil size reflects successful encoding and recall of memory in humans. *Scientific reports*, 8(1):4949.

Laeng, B. and Falkenberg, L. (2007). Women's pupillary responses to sexually significant others during the hormonal cycle. *Hormones and behavior*, 52(4):520–530.

Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of experimental social psychology*, 49(4):764–766.

Liao, H.-I., Kashino, M., and Shimojo, S. (2021). Attractiveness in the eyes: A possibility of positive loop between transient pupil constriction and facial attraction. *Journal of Cognitive Neuroscience*, 33(2):315–340.

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Šmíra, M., Epskamp, S., et al. (2019). Jasp: Graphical statistical software for common statistical designs. *Journal of Statistical Software*, 88:1–17.

Mantiuk, R., Kowalik, M., Nowosielski, A., and Bazyluk, B. (2012). Do-it-yourself eye tracker: Low-cost pupil-based eye tracker for computer graphics applications. In *Advances in Multimedia Modeling: 18th International Conference, MMM 2012, Klagenfurt, Austria, January 4-6, 2012. Proceedings 18*, pages 115–125. Springer.

Mathôt, S., Dalmaijer, E., Grainger, J., and Van der Stigchel, S. (2014). The pupillary light response reflects exogenous attention and inhibition of return. *Journal of vision*, 14(14):7–7.

Naber, M., Frässle, S., Rutishauser, U., and Einhäuser, W. (2013). Pupil size signals novelty and predicts later retrieval success for declarative memories of natural scenes. *Journal of vision*, 13(2):11–11.

Park, S., Zhang, X., Bulling, A., and Hilliges, O. (2018). Learning to find eye region landmarks for remote gaze estimation in unconstrained settings. In *Proceedings of the 2018 ACM symposium on eye tracking research & applications*, pages 1–10.

Pronk, T. M., Bogaers, R. I., Verheijen, M. S., and Sleegers, W. W. (2021). Pupil size predicts partner choices in online dating. *Social Cognition*, 39(6):773–786.

Rafiqi, S., Wangwiwattana, C., Fernandez, E., Nair, S., and Larson, E. (2015). Work-in-progress, pupilware-m: Cognitive load estimation using unmodified smartphone cameras. In *2015 IEEE 12th International Conference on Mobile Ad Hoc and Sensor Systems*, pages 645–650. IEEE.

Rieger, G. and Savin-Williams, R. C. (2012). The eyes have it: Sex and sexual orientation differences in pupil dilation patterns. *PloS one*, 7(8):e40256.

Ruiz, A. P., Flynn, M., Large, J., Middlehurst, M., and Bagnall, A. (2021). The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data Mining and Knowledge Discovery*, 35(2):401–449.

Sawyer, A. N., Smith, E. R., and Benotsch, E. G. (2018). Dating application use and sexual risk behavior among young adults. *Sexuality Research and Social Policy*, 15:183–191.

Skakoon-Sparling, S., Cramer, K. M., and Shuper, P. A. (2016). The impact of sexual arousal on sexual risk-taking and decision-making in men and women. *Archives of sexual behavior*, 45:33–42.

Wu, S. and Trottier, D. (2022). Dating apps: a literature review. *Annals of the International Communication Association*, 46(2):91–115.