# The Catcher in the Eye: Recognizing Users by their Blinks

Ryo Iijima
ryo.iijima@aist.go.jp
National Institute of Advanced Industrial Science and Technology
Koto, Tokyo, Japan

Tatsuya Takehisa
t.takehisa@nict.go.jp
National Institute of Information and Communications Technology
Koganei, Tokyo, Japan

Tetsushi Ohki
ohki@sec.inf.shizuoka.ac.jp
Shizuoka University
Shizuoka, Tokyo, Japan

Tatsuya Mori
mori@nsl.cs.waseda.ac.jp
Waseda University
Shinjuku, Tokyo, Japan

## ABSTRACT

In this paper, we develop a novel behavioral biometric recognition framework, BlinkAuth, that takes advantage of a user's blinking. BlinkAuth utilizes electrooculogram (EOG) data, (i.e., the electric potential difference between the corneal and retinal sides of the eye), and applies a machine-learning model to achieve user recognition. BlinkAuth works with devices like smart glasses and VR headsets and can be used simultaneously in activities such as driving or cooking. Using JINS MEME, a glasses-type wearable device that can measure EOG, we collected EOG data from 31 participants under various conditions and evaluated the recognition accuracy of BlinkAuth. The results demonstrate that BlinkAuth can achieve high accuracy as a behavioral biometric recognition with an average AUC of 95.8% and an average EER of 9.28%. We developed a system for implementing BlinkAuth for real-time recognition and evaluated the time required for the recognition process and the system's usability with the System Usability Scale (SUS). The results show an overall processing time of approximately 0.6 seconds, including the data measurement time, and an average SUS score of 82.50, which indicates high usability equivalent to rank A in the standard criteria for interpreting SUS scores. Six extensive user experiments and 17 evaluation perspectives reveal that BlinkAuth is highly robust to environmental changes, such as skin moisture and makeup, participant actions, and eye strain conditions, as well as to attacks that imitate the target's blinking.

## CCS CONCEPTS

• **Security and privacy → Usability in security and privacy**; **Biometrics**.

## KEYWORDS

Bio-signals, ElectroOculoGram, User recognition, User study


*Conference acronym 'XX, June 03–05, 2018, Woodstock, NY*

ACM ISBN 978-1-4503-XXXX-X/18/06
https://doi.org/XXXXXXX.XXXXXXX

**ACM Reference Format:**
Ryo Iijima, Tatsuya Takehisa, Tetsushi Ohki, and Tatsuya Mori. 2018. The Catcher in the Eye: Recognizing Users by their Blinks. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX).* ACM, New York, NY, USA, 14 pages. https://doi.org/XXXXXXX.XXXXXXX

## 1 INTRODUCTION

Recent wearable devices, such as smart glasses and smartwatches, are equipped with sensors that measure bioelectrical potentials such as electrocardiogram (ECG), electroencephalogram (EEG), electromyogram (EMG), and electrooculogram (EOG), providing users the ability to track their health status and daily activities. Traditionally, these bioelectrical potentials have been collected by expensive medical devices. However, as sensors have become affordable, it is now possible to measure bioelectrical potentials using commercially available wearable devices.

Biopotentials are electric potentials generated by the activity of nerve cells and muscles. It is expected to be applied to personal recognition technology because there are individual differences in the signal waveform of biopotentials [6, 23, 29]. Although sensors in wearable devices are now capable of measuring bioelectric potentials, recognition technology using bioelectric potentials has not become widespread in these devices. Currently, most wearable devices, such as smart glasses and smartwatches, adopt passwords and/or PINs as recognition functions. These recognition methods do not have high usability owing to the long time required for recognition [27]. This study aimed to develop a usable biopotential-based recognition system that operates with commercially available wearable devices based on the aforementioned background.

In this study, we focused on EOG as a promising biopotential for usage in a behavioral biometric recognition system. EOG is a method for measuring eye movement using the electric potential difference between the retinal and corneal sides of the eye. EOG is used for eye tracking, concentration analysis, and electrophysiological testing in medicine [31, 33]. Blinking involves eye movements and changes in eyelid resistance, generating a characteristic EOG signal pattern [8]. We developed a biometric recognition framework, BlinkAuth, that exploits the unique pattern of EOG when a user blinks. We expect BlinkAuth to be mounted on smart glasses or VR headsets. BlinkAuth can be categorized as a type of behavioral biometric technology based on conscious or unconscious blinking behavior. In general, the intensity of EOG is on the order of 100 $\mu$V,

**Table 1: Comparison table of related works. "Simul." means "Simultaneously" and indicates whether something can be performed simultaneously with actions other than recognition.**

| | Performance | | Evaluation | | | Other differences | |
|---|---|---|---|---|---|---|---|
| Ref. | EER | Time | Usability | Robust | Security | Stimuli | Cost (USD) |
| Zhang2017 [53] | 6.90% | N/A | ad-hoc | - | - | Visual | - |
| OcuLock [30] | 3.55% | 3–10 s | ad-hoc | Time | Closed | Visual | 50K |
| Blink to Get In [16] | 3.90% | 3–10 s | ad-hoc | - | - | Visual | 50K |
| Seha2019 [41] | 9.90% | N/A | - | - | - | Visual | - |
| SoundLock [54] | 3.90% | 10 s– | ad-hoc | - | Closed | Audio | 2K |
| Abbas2017 [1–3] | 4.40% | N/A | - | - | - | **Not Required** | 1.5K |
| BlinkAuth(ours) | **1.83%** (driving) | **0.6 s** | **SUS [11] =82.5** | **Time Skin Action** | Closed **Open** [36] | **Not Required** | **150** |

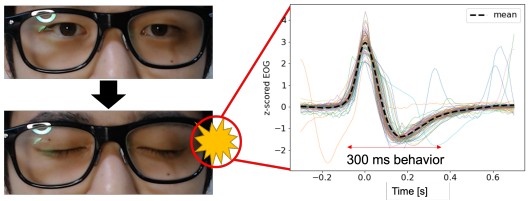


**Figure 1: Measurement of EOG signal patterns generated by blinking. The center graph presents the EOG waveforms of blinks.**

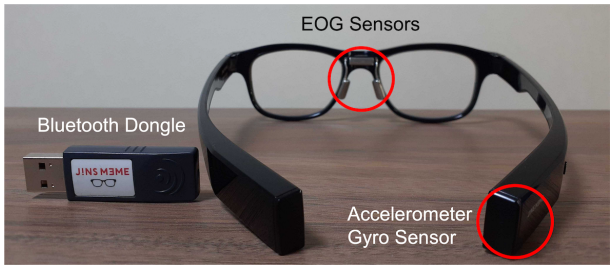


**Figure 2: JINS MEME smart glasses. The nose pads and the bridge are made of metal and are used as sensors to acquire EOG voltage values.**

and the signal intensity is high compared to EEG on the order of 10 $\mu$V. Fig. 1 illustrates an overview of a biometric recognition scheme based on EOG.

The advantages of this study, compared to other related research (refer to Table 1 and Sec. 2.3), are as follows:

- **High Accuracy.** By implementing the BlinkAuth Framework (Sec 3), we achieved the lowest Equal Error Rate (EER) of 1.83% among the related studies (Sec 5.4).
- **High Speed.** In general, behavioral biometrics tends to take a long time to complete the recognition [6, 30, 32] (see Table 1). BlinkAuth is the fastest recognition method among the related studies, with the recognition process being completed within 600 ms. (Sec 4.3) In practical use-case, the recognition time significantly impacts usability. BlinkAuth has achieved the highest Rank A in usability according to the System Usability Scale (SUS) (refer to Sec 4.3).
- **Robustness.** BlinkAuth is robust against environmental changes such as the time effect (Sec 5.1), the presence of the attacker (Sec 5.2), makeup around the eyes (Sec 5.3), eye strain (Sec 5.4), and the presence of other actions such as driving a car. (Sec 5.4)
- **Impersonation Attack Evaluation.** In prior work, impersonation attacks were studied with closed-set evaluations, including attackers in the training data. In this study, along with closed-set evaluations, we also conducted an open-set evaluation, thereby excluding the attacker from the training data.
- **Stimulus-Free.** Generally, related recognition methods often involve presenting stimuli such as audio or visual cues for the recognition process. In this study, we achieved recognition without those stimulus presentations (Table 1). Thus, the user need not pay attention to the sensor when measuring data, and the recognition can be performed simultaneously with other actions. (Sec 5.4)
- **Cost.** Several related works on recognition systems using bioelectrical signals have relied on costly medical equipment [16, 30] and do not reflect commercially available wearable devices. We developed BlinkAuth using a commercially available device that is priced at approximately 150 USD.

## 2 BACKGROUND

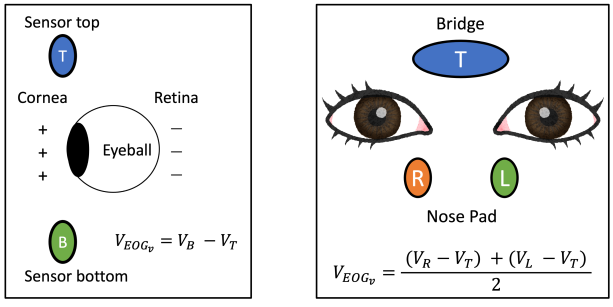


**Figure 3: Left: Diagram of EOG measurement. Right: diagram of JINS MEME measurement [24].**

## 2.1 ElectroOculoGraphy (EOG)

EOG [12] is a method of measuring eye movement using the electric potential difference between the cornea and retina. The cornea of the eye has a positive potential and the retina has a negative potential, and changes in eye movements and eyelid resistance cause changes in the measured voltage. EOG measurement is performed by attaching metal electrodes to the skin around the eyes. Left side of Figure 3 shows an image of the EOG measurement. In general, vertical EOG ($EOG_v$) can be obtained by adding electrodes to the upper and lower lids of the eye, and horizontal EOG ($EOG_h$) can be obtained by adding electrodes to the right and left edges of the eye [12]. In this study, we will focus on $EOG_v$. We simply refer to $EOG_v$ as EOG. One example of a commercial product that can acquire EOG is BITalino [39]. Because of its unnatural appearance and it takes time to wear the sensor, developers improve the sensor shape to wear the EOG sensor with a wearable device. For example, in the case of JINS MEME used in this research, as shown in Figure 2, the nose pads and bridge are metal, which are called 3-point electrooculography (EOG) sensors [24]. The measurement image of the 3-point EOG sensor is shown in right side of Figure 3. The 3-point EOG sensor is based on the sleep analysis research [5]. EOG can be measured by attaching metal electrodes to the skin around the eyes [5]. The left side of Fig. 3 presents an image of the EOG measurement. In general, vertical EOG ($EOG_v$) can be measured by attaching electrodes to the upper and lower eye lids, and horizontal EOG ($EOG_h$) can be measured by attaching electrodes to the right and left edges of the eye [12]. In this study, we focused on the vertical EOG, simply referred to as EOG.

In the case of unconscious blinking, the range of changing voltage is known to appear between 10–200 $\mu$V and frequency between 0.5–15 Hz [8]. Fig. 6 presents an example of measured EOG waveform. The EOG waveform is characterized by a positive change in potential from the start of blinking until the eyes are completely closed and a negative change in potential from the beginning of the opening of the eye until the eyes are completely opened. The EOG signal that appears during blinking has a higher voltage value than the EEG signal.

Researchers and medical professionals who perform EEG analysis have treated eye potentials as noise and have performed preprocessing to remove the eye potential component before analysis [4, 42]. EOG, on the other hand, has promising applications such as analysis of concentration and fatigue levels and eye tracking while driving [47]. The commoditization of eye potential measurement sensors is progressing, and wearable devices capable of measuring EOG, such as JINS MEME, are gaining popularity.

Conventional commercial EOG measurement devices such as BITalino [39] are unnatural in shape and time consuming to wear. For the practical application of EOG measurement, developers have improved EOG measurement devices to make them lightweight and easy to use. For example, JINS MEME, which is adopted in this study, is an easy-to-use EOG measurement device.

## 2.2 Wearable device that can be equipped with EOG measurement functionality

BlinkAuth is a framework for biometric recognition based on EOG. We outline wearable devices that are already equipped or capable of being equipped with EOG measurement functions as follows.

**Smart glasses.** Smart glasses are wearable devices with displays and sensors mounted on glasses. They can project maps and information about the surrounding area as augmented reality onto the lenses and perform activity tracking based on sensor readings. Examples of commercially available smart glasses include Google Glass, Ray-ban Stories [35], Nreal Air [18], Neon [19], SMI ETG [17] and JINS MEME [24]. JINS MEME is equipped with an acceleration sensor and gyro sensor and capable of activity tracking. JINS MEMS is also equipped with a sensor capable of measuring EOG, which enables the analysis of concentration and eye movement. In this study, we use JINS MEME for the experiment. Google Glass and Nreal Air are equipped with transparent displays that display AR objects and screens of other devices. Neon is an eye tracking device, which is equipped with a special camera for eye tracking. SMI Eye Tracking Glasses are smart glasses that use two infrared cameras for eye tracking [17]. Berkovsky et al. used the device to measure eye-gaze behavior and analyze personality traits [9]. As such, at the time of writing this paper, many of smart glasses, except JINS MEMS, are not equipped with EOG sensors. Since EOG sensors are small and consume little power, we expect that smart glasses with EOG sensors installed may become more popular in future.

**VR headset.** A VR headset is a device that is worn on the head to experience virtual reality. VR headsets are widely used for gaming, live viewing, education, simulators, trainers, etc. VR headsets are equipped with accelerometers, gyroscopes, and magnetic sensors capable of tracking head motion in addition to a stereoscopic head-mounted display. The three authentication methods used in typical VR headsets today are pattern lock, password, and PIN. Pattern locks are used by Meta Quest [34], PINs are used by Vive Focus Plus [50], Vive Focus 3 [49], and other VR devices from Vive, and passwords are used by PlayStation VR [44] and applications within Meta Quest. Currently, there are no VR headsets that can measure EOG, but since VR headsets are shaped to cover the area around the eyes and the upper part of the nose, it is easy to attach sensors that measure EOG in the same way as smart glasses. In fact, there exist VR headsets that integrate eye tracking sensors.

## 2.3 Related works

This chapter summarizes research that has proposed authentication methods using eye movements. A summary of related work features can be found in Table 1. Authentication methods based on eye movements can be broadly categorized into those using EOG sensors as described in Sec. 2.1, and those that employ cameras, known as VideoOculoGraphy (VOG) based authentication. OcuLock, proposed by Luo et al. [30], and research conducted by Gupta et al.[16] and Abo et al. [1–3], have developed authentication systems utilizing EOG generated by eye movements. Authentication methods based on eye movements can be broadly categorized into those using EOG sensors as described in Sec. 2.1, and those that employ cameras, known as VideoOculoGraphy (VOG) based

authentication. Gupta et al. [16] authenticate by providing visual stimuli to the eyes. Most previous research utilized the BIOPAC MP36R [10] for EOG measurements, which is extremely expensive at over 50,000 USD, rendering it unsuitable for practical application.

Related research such as that of Zhang et al. [53], Seha et al. [41], and SoundLock by Zhu et al. [54], all focus on VOG-based authentication using cameras. In VOG, infrared and camera sensors are used to determine the position of the pupil. However, VOG utilization comes with challenges, including the necessity for calibration at startup, resulting in reduced user-friendliness. Additionally, noise removal and accuracy improvement are difficult. One innovation aiming to improve accuracy is SoundLock [54], which detects unconscious pupil movements caused by auditory stimuli using VOG and achieved a high detection accuracy of 3.90%.

Compared to the aforementioned related research, our study achieved the lowest Equal Error Rate (EER) of 1.83%, whereas behavioral biometrics require long time to complete recognition [6, 30, 32] (see Table 1), BlinkAuth is the fastest recognition method among the related research with process completion requiring only 600 ms. For a practical use-case, recognition time significantly impacts usability. Additionally, BlinkAuth is robust against environmental changes such as the time effect, presence of the attacker, eye makeup and strain, and the presence of other actions such as driving a car. Moreover, recognition was achieved without those stimulus presentations (Table 1), whereas related recognition methods frequently involve presenting stimuli such as audio or visual cues for the recognition process. Thus, using BlinkAuth, the user is not required to pay attention to the sensor when measuring data, and recognition can be simultaneously performed with other actions. Finally, Several related publications focus on recognition systems using bioelectrical signals that rely on costly medical equipment [16, 30], thus precluding their use for commercially available wearable devices. For example, Ref [16, 30] used the BIOPAC [10], which costs expensive medical devices over 500,000 USD, to measure the EOG. In contrast, we developed BlinkAuth using a commercially available device that is priced at approximately 150 USD.

# 3 DESCRIPTIONS OF THE BLINKAUTH FRAMEWORK

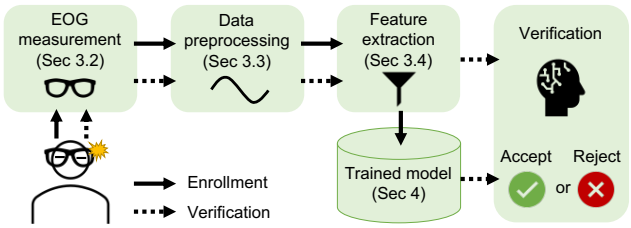


**Figure 4: Workflow of the BlinkAuth framework.**

In this section, we describe the overview of BlinkAuth and the technical modules that comprise the framework. After presenting an overview of the framework (Sec. 3.1), we describe the method of EOG data collection (Sec. 4.1), processing the measured EOG data (Sec. 3.3), and extraction of features for application to machine learning algorithms that perform recognition (Sec. 3.4).

## 3.1 Overview of the framework

Fig. 4 illustrates the workflow of BlinkAuth. A user either blinks actively (single recognition mode) or unconsciously (continuous recognition mode). Devices equipped with EOG sensors, such as smart glasses, are used to measure changes in EOG generated by blinking (Sec. 3.2). BlinkAuth performs preprocessing on the measured EOG data (Sec. 3.3). The preprocessing can be performed in real-time. Next, BlinkAuth extracts features for the machine learning model that performs recognition (Sec. 3.4).

To use BlinkAuth for the behavioral biometric recognition task, we first train the machine learning model using the data of the user to be verified (the trained model in the figure). After the training is completed, we input the features extracted from the EOG measurement data generated by the user's blink into the machine learning model. Recognition of the user can be performed based on the output of the machine learning model. The details of the machine learning models employed in this study will be described in Section 4.

## 3.2 Measuring EOG

BlinkAuth leverages wearable devices equipped with EOG sensors to measure the unique EOG patterns generated by blinking. JINS MEME [24] is an example of such a device. In this study, we implement and evaluate BlinkAuth using JINS MEME Academic Pack. Fig. 2 is a photograph showing the appearance of JINS MEME. JINS MEME is smart glasses equipped with an EOG sensor that can acquire left and right eye potentials. JINS MEME has electrodes installed at three points on the two nose pads and the bridge, which act as EOG measurement sensors [24]. Fig. 3 to the right shows an image of EOG measurement with the three-point electrode sensor of JINS MEME. It can measure the EOG signals with dry electrodes between the eyebrows and on the nose piece. As shown Fig. 2, JINS MEME has electrodes installed at three points on the two nose pads and the bridge, which act as EOG measurement sensors [24]. In addition to the EOG sensors, JINS MEME is equipped with an accelerometer and a gyroscope to estimate the posture of the person wearing the glasses. In this study, the sampling frequency of JINS MEME was set to $f_s = 100$ Hz.

## 3.3 Data Preprocessing

This section shows the preprocessing steps performed on the measured raw EOG data. Fig. 5 presents the preprocessing pipeline. First, noise reduction and drift correction are applied to the waveform. Next, the waveform corresponding to a single blink is extracted. Finally, outliers are removed. The details of each process are shown below.

**Filtration and drift correction**

A bandpass filter is applied to the raw EOG data to perform noise filtering. A 6th-order Butterworth filter was employed as a filter to remove waveforms not included in the range of 0.25–45 Hz. [1].

[1]We determined the lower frequency limit based on the characteristics of the EOG. For the upper frequency limit, we set the sampling frequency of JINS MEME to 100

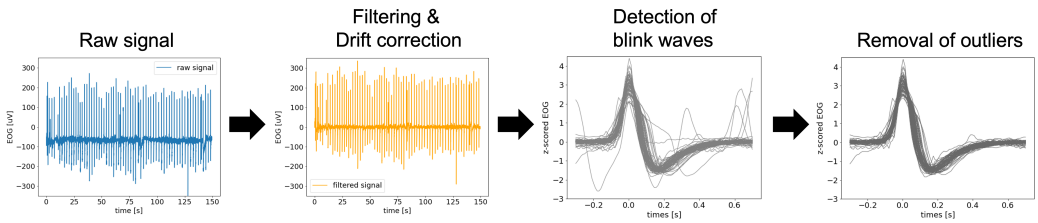


**Figure 5: EOG data preprocessing pipeline.**

The waveform at 0 Hz can be regarded as the drift from the point of 0 V (DC offset). As described above, the waveform can be converted to a waveform with 0 drift correction by cutting the frequency components, including 0 Hz. This filtering process can remove non-zero-based components caused by skin conditions, perspiration, and other influences.

**Extraction of blink waveforms**

As shown in Fig. 5, the EOG waveform shows a sharp peak with blinking. BlinkAuth extracts waveforms corresponding to blinks by applying a peak detection algorithm to the filtered EOG data. We define the peak voltage of the EOG generated by a blink as the voltage positive peak $V_p$. We extract a single blink waveform from 0.3 seconds before to 0.7 seconds after the time when $V_p$ occurs ($D_p = 0.3$ and $D_n = 0.7$ in Fig. 6). Since this method includes outliers in some cases, they are removed using the method presented below. Note that in this study, SciPy's `find_peaks` function was employed to implement the peak detection algorithm.

**Removal of outliers**

The BlinkAuth framework takes advantage of waveform features that are unique to each user to remove outliers. To this end, we normalize the waveforms using the Z-scoring. Note that the normalization process is applied only for outlier removal and not for feature extraction. We compute the average of the waveforms for several blink waveforms collected during user registration and use the resulting one as a template. We compute the Root mean square error (RMSE) of the template and inspected waveforms and remove the inspected waveform as an outlier if the RMSE is greater than 0.5. This threshold was empirically determined using actual measured data so that the percentage of error waveforms would be around 5–10% of the total. We accurately recorded the start time of the blink waveform in our experiments by asking the user to press the button at the same time as the blink.

**Data Augmentation**

Owing to the limited amount of data obtained from the user study, we employed data augmentation techniques for the training data. Based on audio data augmentation methods [21], we implemented three methods: (1) adding gaussian noise, (2) phase shifting, and (3) linear stretching. Through the three types of data augmentation mentioned above, the size of the training dataset becomes four times the original size.

Hz, so based on the sampling theorem, we set the upper frequency limit to 45 Hz, a frequency slightly lower than half of 100 Hz.

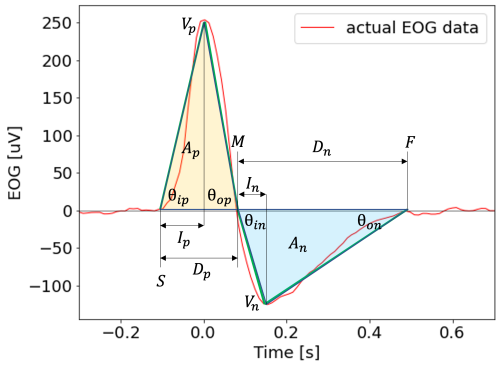


**Figure 6: Extraction of features based on the graphic characteristics of a blink waveform.**

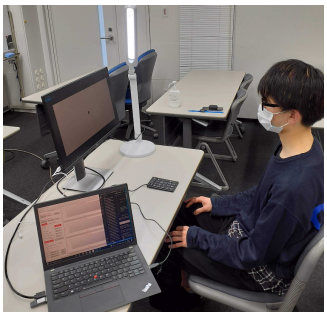

**Figure 7: Experimental setup: measuring blink EOG data using JINS MEME.**

## 3.4 Feature Extraction

BlinkAuth extracts features based on (1) EOG waveforms and potential values, and (2) frequency analysis of EOG potential values for the preprocessed blink waveforms. We present those methods as follows.

### 3.4.1 Features based on waveform and potential values

We extract features based on (i) graphic characteristics of the waveform, (ii) autocorrelation coefficients of the potential values, and (iii) moments of the potential values. Each of these feature extraction methods is shown below.

**(i) Graphic characteristics of the waveform**

The following feature extraction approach is based on the one proposed in Ref [2]. The graphic characteristics of the measured

blink waveform are extracted, as shown in Fig. 6. The actual waveform (red line) is simplified by connecting the vertices representing positive and negative peak values and the intersection points of the waveform and the X axis with straight lines (green line). The variables shown in Fig. 6 are extracted as features. Where $\theta$ denotes angle, $A$ denotes area, $V$ denotes voltage, and $I, D$ denotes time, respectively. The subscripts $p$ and $n$ denote positive and negative voltages, respectively. Unsubscripted variables represent the position (time) on the horizontal axis. Note that $t = 0$ is the time when the voltage value peaks.

**(ii) Autocorrelation coefficient of potential values**

Parameter estimation of the Autoregressive (AR) model is performed on the time series data of potential values constituting a single blink waveform. The AR model is represented by the following equation

$$V_t = c + \Sigma_{i=1}^{m} \phi_i V_{t-i} + \epsilon_t \quad (1)$$

where $V_t$ is the potential of the blinking waveform at some discrete time $t$, $c$ is the intercept, $\phi_i$ is the $i$th order autocorrelation coefficient, and $\epsilon_t$ is the disturbance term. We extract the autocorrelation coefficients resulting from the parameter estimation of the AR model as features. We adopt $m = 20$ as the dimension of the AR model based on previous studies [3] that have applied the AR model to the analysis of EEG.

**(iii) Moment of potential values**

As the basic time-series statistic used for biological signals [27], moments such as variance and skewness are calculated for the potential value $V_t$ of each blink waveform and used as feature values.

#### 3.4.2 Features based on frequency analysis

We employ the power spectral density (PSD) to extract features based on frequency analysis, computed by applying the fourier transform to the EOG potential values of a blink and normalizing them by the sampling interval $\Delta f$. With reference to the evaluation of each frequency band used for EEG measurements [45, 52], i.e., $\delta$ (0.5–3 Hz), $\theta$(4–7 Hz), $\alpha$ (8–13 Hz), $\beta$ (14–30 Hz), and $\gamma$ (30 Hz–), the integrated PSD values corresponding to the respective frequency bands are adopted as features. In addition, the mean frequency, variance frequency, mean power, and total power were adopted as the frequency-domain features [38]. We note that these features are also used in myopotential measurement and motion estimation.

## 4 RECOGNITION ACCURACY OF THE BLINKAUTH FRAMEWORK

In this section, we report the results of our experimental evaluation of the recognition accuracy of BlinkAuth. We evaluated the impact of the machine learning model on recognition accuracy and the impact of blinking with both eyes or one eye on recognition accuracy. We also evaluated the effectiveness of continuous recognition in addition to single-time recognition.

**Ethical Considerations.** We carefully considered the concerns regarding protecting the participants' health and privacy because we used bioelectrical potentials to identify individuals. We submitted our research plan to the IRB of our organization and received formal approval to experiment before its commencement. In this study, we focused only on the bioelectric potentials that could be acquired by the EOG sensor. We did not involve any invasive procedure or intervention on the participants in the experiment. Since EOG is data that can identify individuals, we received informed consent from participants prior to the experiment. Specifically, we carefully explained the purpose of the experiment, the experimental procedure, how the data would be used, and that the participants would not suffer any disadvantages by participating in the experiment. We conducted the study only with participants who signed a consent form. We respected the wishes of the experimental participants to the greatest extent possible and informed them that they could stop immediately if they felt sick or uncomfortable during the experiment. We checked the participants' health conditions before starting the experiment and ensured they were not physically burdened. For all of the experiments, we recruited participants by posting on the student jobs section of the author's university. Participants received compensation at a rate of approximately $15 per hour.

### 4.1 Data Collection

Here, we describe the procedure for acquiring blink EOG data that we used to evaluate the performance of BlinkAuth. For the experiment, we recruited participants by posting on the student jobs section of the author's university. We recruited 31 participants, comprised of university students and faculty members, with 17 males and 14 females, ranging in age from 18 to 48.

**Experimental setup**

Fig. 7 presents the experimental setup for measuring the EOG data of blinking. We closed the blinds in our room and set up a stand light with adjustable brightness during the data measurements. Such a setup allows us to maintain a constant brightness of the experimental environment. We adjusted the light to maintain a brightness level of 480-520 lux near the participant's eyes. The participants were seated such that they can see the black dot in the middle of the black circle displayed on the monitor and keep their attention on the dot. The evaluation for instances when the users are not focusing on the black dot was conducted in Sec 4.3 and after. Participants were seated in a chair and kept a constant distance from the screen. The purpose of these setups was to eliminate the effects of gaze blurring caused by unstable posture.

We have collected 180 blinks (60 blinks × 3 sessions) for four types: unconscious blinks, conscious blinks, dominant-eye blinks, and non-dominant eye blinks. Each session takes 3 minutes. When measuring active blinking, we sounded a tone once every 3 seconds and instructed participants to blink in response to the tone. The reason for the 3-second interval was to prevent successive blinks from interfering with each other. Because it is difficult to execute a blink precisely in time with the sound played at 3-second intervals, participants were instructed to blink with timing to the extent possible. We also instructed the participants to press the button at the timing when they performed the blinking motion. The timestamp information of the button press was recorded as an artifact. By referring to the artifact, the EOG waveform of the blink could be accurately obtained. We present detailed experimental procedures and design in Appendix A.

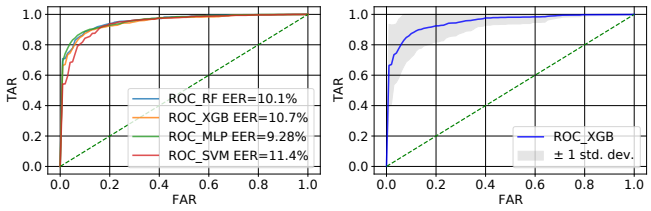


**Figure 8: Left: Mean ROC for the four recognition models. Right: Mean and Std ROC for XGB model.**

## 4.2 Evaluation of the Recognition Accuracy

In this section, we evaluate the recognition accuracy of BlinkAuth.

**Creating a dataset for evaluation**

First, after acquiring the participant's EOG data using the procedure described in Section 4.1, the preprocessing and feature extraction described in Section 3 are applied to generate a feature vector of blink waveforms. As features, all the features based on the waveform and potential values, as well as those based on frequency analysis as described in Sec 3.4, were employed. We then assign a label of 1 to the feature vector of the user to be authenticated and a label of 0 to the feature vectors of other users. After creating the labeled feature vectors, we perform undersampling and adjust the number of labels 0 and 1 to be equal, respectively [46], in order to prevent the label bias.

**Metrics used for accuracy evaluation**

To evaluate recognition accuracy using machine learning models, we employ AUC (Area Under the Curve) score, EER (Equal Error Rate), F1-score, and ROC (Radio Operating Curves). These metrics are widely used to evaluate the performance of authentication systems [46] 1:1 verification models were created for each of the 31 participants, and the AUC score, EER, F1, and ROC were calculated for each model.

**Machine learning model**

The machine learning models used for recognition are support vector machine (SVM) as a nonlinear model, Random Forest (RF) and XGBoost (XGB) as models using decision trees, and neural networks (NN) with relatively simple structures. We completely separated the obtained data into training and test data at a 8:2 ratio.

**Evaluation of the recognition accuracy**

Table 2 shows the average AUC, EER, and Accuracy for each machine learning model. The left side of Fig. 8 shows the mean ROC curve for each model and the right of Fig. 8 shows the mean ROC and variance for the 31 participants in the case of XGboost. BlinkAuth achieved high accuracy with AUC scores exceeding 95% for all machine learning models. The EER was approximately 9–10 %, which is a very high level of accuracy for behavioral biometric recognition. Significantly, the lowest Equal Error Rate (EER) of 9.28% was achieved by MLP. In the following evaluations, we will conduct assessments using XGBoost.

**Recognition Accuracy of Cases with One-eye Wink.**

We targeted both eye blinks in our evaluations discussed in the previous section. In addition to the two-eye blink, BlinkAuth allows for variations on the eyelid-closing motion, such as a wink with one eye or two blinking motions. In this section, we study the recognition accuracy when the user uses the one-eye wink. Here, we adopt the "Dominant eye test" by the Ref [15] to distinguish between dominant and non-dominant winks. In the section of appendix B, we present the procedure of the dominant eye test. As a result of the dominant eye test, 22 out of 31 (71%) were determined to be right-eye dominant and 9 out of 31 (29%) were determined to be left-eye dominant. Table 3 shows the results of the evaluation of recognition accuracy when one eye was winked. For comparison, the results for both eye blinks are also shown. We employed XGB as the machine learning model in both cases. The wink with the dominant eye achieved EER=12.1%, and the wink with the non-dominant eye achieved EER=12.5%, both lower than 15%, even though the recognition error for the one-eye wink was somewhat higher than that for the two-eye blink. These results suggest that recognition using one-eye winking is practical in BlinkAuth.

**Continuous Recognition**

So far, we have targeted single-time recognition using actively executed blinks or winks. BlinkAuth can realize continuous recognition using unconscious blinking, as mentioned in Section 1. Promising use cases for single-time recognition and continuous recognition are discussed in Section 6.1. In the following, we show the results of evaluating the accuracy of continuous recognition. Table 4 presents the results. The results of single-time recognition are also shown for comparison. The accuracy of continuous recognition can achieve relatively high accuracy of more than 86% for F1 and 92% for AUC, even though it is slightly lower than that of single-time recognition. The reasons for the lower accuracy of continuous recognition compared to single-time recognition are that (1) the extraction accuracy of blink waveforms is lower because artifacts cannot be added to unconscious blinks during measurement and (2) the peak potential of unconscious blinking is smaller than that of active blinking, making it more susceptible to noise. It is necessary to improve the accuracy of blink detection to improve the accuracy of continuous recognition, which is an issue to be addressed in the future.

## 4.3 Processing Time & Usability: Real-time Recognition Test

We demonstrate how the BlinkAuth framework can be implemented as a real-time recognition system. We have developed a real-time recognition system that implements a series of workflows shown in the top of the Fig. 9. The objectives of this experiment are (1) to evaluate the end-to-end performance when users attempt to recognize themselves using BlinkAuth, and (2) to evaluate the usability of BlinkAuth as the recognition system. In the following, we evaluate the performance of the implemented system. We also evaluate the usability of the system through user study experiments.

**Experimental Setup**

The top of Fig. 9 illustrates the overview of the implementation of the real-time recognition system. The appearance of the real-time recognition setup is shown in Fig. 9 to the bottom. The recognition system first receives the raw EOG data from JINS MEME via the Bluetooth channel. It then performs pre-processing, feature extraction, and recognition, using the trained model. All processes, except

**Table 2: Results of evaluating recognition accuracy using different machine learning models.**

| ML models | AUC | EER | F1 |
|---|---|---|---|
| RF | 0.990 | 0.101 | 0.956 |
| XGB | 0.951 | 0.107 | 0.886 |
| SVM | 0.936 | 0.114 | 0.881 |
| MLP | 0.958 | 0.0928 | 0.862 |

**Table 3: Comparison of recognition accuracy: winking one eye vs. blinking both eyes.**

| Eyes | AUC | EER | F1 |
|---|---|---|---|
| Dominant | 0.946 | 0.121 | 0.880 |
| Non-dominant | 0.948 | 0.125 | 0.898 |
| Both | 0.951 | 0.107 | 0.886 |

**Table 4: Recognition accuracy: single-time recognition vs. continuous recognition.**

| Recognition mode | AUC | EER | F1 |
|---|---|---|---|
| Continuous | 0.925 | 0.124 | 0.867 |
| Single-time | 0.951 | 0.107 | 0.886 |

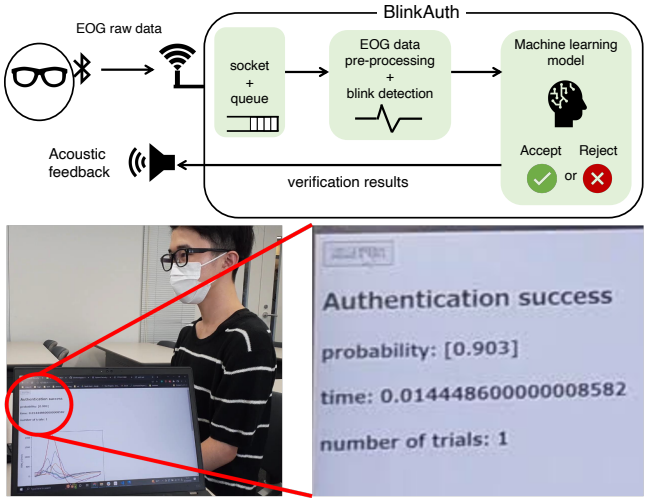


**Figure 9: PoC implementation of the real-time recognition system with the BlinkAuth framework. Top: Design, Bottom: Examples of realtime evaluation.**

the training model, are performed in background processing. The system provides an acoustic tone when notifying the user to start the recognition. If user recognition is accepted, the system provides acoustic feedback with a different sound.

To validate the performance of the real-time system, we measured the recognition time through user study experiments by recruiting 12 participants (7 women and 5 men, ranging in aged from 18–28). Note that the following experiments are independent of the experiments discussed in the previous section, and the number of participants in the experiments are different.

The processing time was defined as the duration from which the notification sound indicates that authentication has started until the feedback confirming authentication success is sent. Note that if authentication fails, the process continues until it succeeds and is included in the measured duration.

We followed the best practices developed in the HCI research community and conducted user studies with 12–15 experimental participants in the following experiments. In the experiment, we collected 180 active blinks (60 blinks × 3 sets) for each participant. For each participant, we train the machine learning model using the blink data collected. We install the trained model on the recognition system and let the participant test the recognition process and record the results. Finally, we ask participants to work on the SUS questionnaire [11]. In addition, with the goals of developing a qualitative understanding of the user experience and exploring the potential of BlinkAuth, we asked the participants two open-ended questions. **Q1**: "*What methods do you think exist to further improve BlinkAuth?*" and **Q2**: "*What are the possible use cases of BlinkAuth for you?*"

### Performance Evaluation

Fig. 10 to the left shows box plots of processing time and SUS final score. The median processing time of the recognition system was 0.3 seconds. Since the time required from the start of blinking to the end of blinking is about 0.3 seconds, overall, the entire recognition process can be completed in about 0.6 seconds. As shown above, the BlinkAuth framework can be used practically as a real-time recognition system.

### Results of System Usability Scale (SUS)

To evaluate the usability of the BlinkAuth framework, we adopt the System Usability Scale, SUS [11], which aims to measure the usability of a system after a user completes the tasks on the system. The SUS consists of a simple questionnaire that asks users to agree or disagree with a series of questions. Respondents are asked to indicate their level of agreement with each question according to a Likert scale of 1 to 5, where 5 means complete agreement and 1 means complete disagreement. In SUS, the questions are categorized into positive and negative questions, and Based on the scores given to each question according to a Likert scale of 1 to 5, the final score is converted so that 100 indicates the most usable and 0 indicates the least usable.

The average SUS score for all users was 82.50, the median was 81.25, and the standard deviation was 10.28. For reference, detailed results for each SUS question are shown in Appendix (Fig. 14). The SUS score of 80.3 or higher is known to correspond to usability within the top 10%, according to Sauro et al. [22]. Thus, BlinkAuth was shown to have high usability.

### Summary of answers to open-ended questions

**Q1**: *What methods do you think exist to further improve BlinkAuth?*

We obtained comments such as "I would expect the recognition to be successful even with varying blink intensities." (User 59) and "I found the task of initiating a blink as soon as the sound was made difficult." (User 70). These comments suggest registering more data or more time for the user to acquire sufficient familiarity with the proposed system as needed to improve the accuracy of recognition.

**Q2**: *What are the possible use cases of BlinkAuth for you?*

For this question, we received several interesting use case scenarios for situations where the user's hands are full; e.g., "I want to pass through ticket gates with a blink behavior instead of holding up a smartphone or card (User 58)," or "I want to unlock auto locks

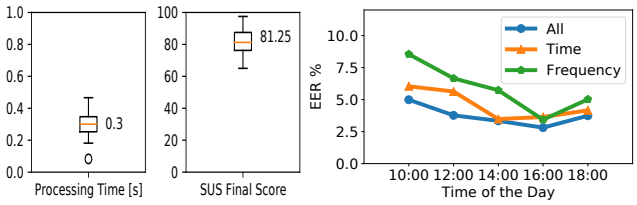


**Figure 10: Left: The processing time and usability scores in the realtime recognition. The orange line shows the median result. Right: EER of time effect (short period.)**

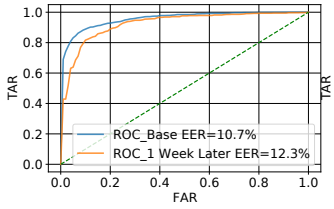


**Figure 11: The ROC for the time effect (long period.)**

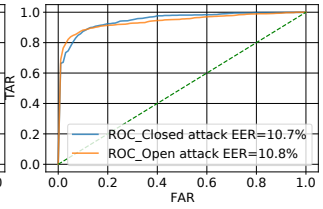


**Figure 12: The ROC for the attack evaluation on BlinkAuth.**

of houses and hotels when my hands are full. (User 44)." We will further discuss these ideas in Section 6.1.

# 5 ROBUSTNESS OF THE BLINKAUTH FRAMEWORK

In this section, we examine whether the BlinkAuth framework is robust against external and internal factors when used in the real world. Specifically, we evaluated (1) the effect of waveform changes over time, (2) the effect of the presence of an attacker, (3) the impact of environmental changes, such as skin conditions, on the accuracy of the recognition as external factors, and (4) the impact of the behavior of the participants and the state of fatigue on the accuracy of the recognition as internal factors. Based on the experimental results of this section, we present discussions regarding improvements for robustness in Section 6.2.

## 5.1 Robustness against Time Passage

The objective of this experiment is to evaluate whether the passage of time affects the accuracy of the authentication results. Evaluations examining variations within the same day are categorized as "Short Period," while evaluations over the course of a week are classified as "Long Period."

**Short period.** This experiment aimed to investigate whether a user's condition within the same day could result in variations in accuracy. We collected data every two hours, from morning through evening. A total of 13 individuals, comprising 6 women and 7 men, ranging in age from 18–28, participated in the study. Measurements were performed at five times throughout the day: 10:00, 12:00, 14:00, 16:00, and 18:00. During each 10-minute measurement period, 180 conscious blinking behaviors were collected. Between each measurement, all participants were in a state of eye strain resulting from operation of either a laptop or a smartphone. Fig. 10 to the right presents the EER results from the recognition, segmented by time and compared across different types of features. As users became accustomed to the recognition actions, the results showed a gradual improvement in accuracy until 16:00. However, by 18:00, the error rate had risen by approximately 0.5-1%. This result might be attributed to minor changes in blink patterns owing to the participants' level of sleepiness or fatigue. This issue is discussed in Sec 6.2.

**Long period.** In this experiment, we collected the same participants from Section 4.1 again and reported the data collected 7–10 days after the measurements from Section 4.1. This is aimed at investigating temporal changes over multiple days. The experimental setup is the same as in Section 4.1. Fig. 11 presents the ROC generated by XGBoost, as reported in Section 4, and the ROC generated by XGBoost from this experiment. The EER increased by approximately 1.6% due to the week-long evaluation. This increase might be caused by the users unfamiliarity with the authentication actions, as they were not performed during the week. Although this error rate is within the acceptable range, it is slightly greater than that of other evaluations. We discuss potential improvements to address this in Section 6.2.

## 5.2 Impersonation Attack

The National Institute of Standards and Technology (NIST) defines closed-set and open-set verifications in the context of face-biometric security: the former involves registered users conducting impersonation attacks, and the latter involves unregistered users [36]. Most academic recognition papers, including OcuLock [30] and our work in Sec 4.2, employ closed-set evaluations. In this study, we implemented an open-set security evaluation, considering impersonation attacks from unregistered users [36].

**Experimental Setup**

For the open-set security evaluation, we reuse the data gathered from 31 individuals in Sec 4.1, comprised of university students and faculty members, with 17 males and 14 females, ranging in age from 18 to 48. We followed the open-set evaluation defined by NIST for face recognition [36] and adapted the evaluation to fit the EOG verification model. The following points describe how to build models for open-set security evaluation. First, from the dataset, we select one individual as a legitimate user (L) and another as not a legitimate user, instead as an attacker (A). Before training begins, we excluded an attacker A's data from the dataset of 31 participants. The dataset was divided, excluding the attacker's data, into training and test sets at an 8:2 ratio. The training set was subsequenatly used to train XGBoost. After XGBoost model training was conducted, we combined the excluded attacker A's dataset with the legitimate user L's test dataset at a 1:1 ratio to populate the test dataset. Note that Attacker A's data is not included in the training process. Using this mixed test data, the EER and ROC were evaluated, followed by an analysis of results. This procedure is conducted for every combination of legitimate users and attackers. Namely, we create ${}_{31}C_2 = LA = 31 \times 30 = 930$ models according to the aforementioned procedures.

**Table 5: Recognition accuracy under different skin conditions.**

| Skin conditions | AUC | EER | F1 |
|---|---|---|---|
| raw | 0.940 | 0.201 | 0.860 |
| skincare | 0.997 | 0.0758 | 0.895 |
| makeup | 0.999 | 0.0356 | 0.914 |

**Results**

Fig. 12 presents the mean ROC and EER generated following the aforementioned procedure for open-set attack evaluation. For comparison, we include the results from the closed-set attack evaluation conducted in Section 4.2. The EER in the presence of an attacker exhibits a marginal increase of 0.1%, which is within the acceptable range. This suggests the model's robustness even under security attacks. Although beyond the scope of this research, creating a model that performs presentation attack detection (PAD) or separate attack detection prior to recognition could be an effective countermeasure.

In the aforementioned evaluation, we assessed the system using test data created by mixing the legitimate user with the attacker dataset at a 1:1 ratio to compare closed- and open-sets under the same conditions. To further the evaluation, we composed the test data solely of attacker samples. Based on the probability of the legitimate user, which can get the value from the function of predict_proba(), we made the models which set with thresholds when FAR=0.01 and 0.05, and we attributed the proportion of successful attack samples as the Successful Attack Rate (SAR). SAR is a metric defined in ISO/IEC 30136: Performance testing of biometric template protection schemes [20]. The SAR evaluation results showed that for FAR=0.01, the average SAR was 0%, while for FAR=0.05, the average SAR was 1.52%. The results above explain that the SAR can be adjusted based on the chosen threshold and tailored to the specific use case. For instance, if the primary goal is to prevent all potential attacks, then setting the threshold to FAR=0.01 would be considered ideal. However, one might opt for the threshold of FAR=0.05 if usability is a priority. Other measures against attacks, besides threshold adjustments, are discussed in Sec. 6.2.

## 5.3 Robustness to Skin Conditions

We conducted user study experiments to validate that the BlinkAuth framework is robust against external factors that could affect EOG measurement through changes in skin conditions. We assumed that changes in skin condition are due to the attachment of moisture and makeup products because moisture and / or makeup products affect the contact resistance of skin and electrodermal activity (EDA) [14]. Specifically, we measured EOG signals under three conditions: raw skin, skin with a skincare product, and skin with a makeup product.

**Experimental Setup**

The purpose of this experiment is to study the effect of changing skin conditions by applying skincare and makeup products on the EOG measurement and the BlinkAuth recognition. Fifteen participants (5 males, and 10 females), ranging in the age from 19–24, participated in this experiment. Participants performed active blinking exercises with both eyes under raw skin, skincare, and makeup conditions, completing 60 blinks in 3-minute sessions for a total of 3 sessions per condition. Table 8 in the appendix lists the names and Fig. 15 in Appendix shows figures of the products used in the experiment. We instructed the participants on the standard usage of the products based on Ref. [37]. Finally, the participants applied skincare and makeup products around their eyes. For reference, examples of each skin condition are shown in Appendix Fig. 16.


**Figure 13: Setup of the driving simulation.**

**Table 6: Recognition accuracy under the different conditions.**

| Conditions | AUC | EER | F1 |
|---|---|---|---|
| Before driving | 0.944 | 0.109 | 0.895 |
| During driving | 0.991 | 0.0246 | 0.972 |
| After driving | 0.997 | 0.0183 | 0.978 |

**Results**

Table 5 presents the results. The AUC and F1 were all high under the skincare-applied condition, and were almost the same as those under the raw skin condition. The result suggests that the effect of moisture adhering to the skin on the recognition accuracy is extremely small. The slightly higher EER in the raw state may be due to participants' unfamiliarity with the blinking action during initial measurements, similar to that observed in the Fig. 10 to the right. These results demonstrate that the BlinkAuth framework is robust against factors of skin conditions (moisture and makeup) and can perform recognition with high accuracy.

## 5.4 Robustness to Human Activities and Eye Strain

In this section, we focus on the internal factors that may affect the EOG measurement results, such as the actions that an user may perform during the recognition process and their eye strain condition, to clarify their influence on recognition accuracy. We adopt driving as a possible action that a user performs simultaneously with recognition. Driving involves a wide range of factors that could influence the accuracy of the recognition, including increased eye movement to check for oncoming vehicles, pedestrians and traffic signals, and increased tension to avoid traffic accidents. For safety reasons, this study used a driving simulation instead of driving a real car.

**Experimental Setup**

As a driving simulator, we adopt Euro Truck Simulator 2 (ETS2) [43]. We adopt a steering wheel controller and brake / acceleration pedals as shown in Fig. 13 to make driving experiences realistic. The simulation environment was configured in a sunny condition and the stand-light on the right side was adjusted to provide the same level of light as in the sunny condition.

We limited the participants to those who had a driver's license. We recruited 14 participants, comprised of university students and faculty members, with 6 males and 8 females, ranging in age from 20 to 29. To ensure that participants' eyes were not fatigued at the start of the experiment, we requested that they had slept well (say, 7 to 8 hours) before the experiment and that they had refrained from using PCs and smartphones too much on the day of the experiment. We checked their conditions prior to the experiment using a questionnaire. The actual sleep and health status, as well as the frequency of smartphone and screen use, were investigated by means of a questionnaire before the experiment.

We first checked the health status of the participants with a brief questionnaire. Next, we gave the participants an overview of the experiment. The participants then completed three sets of 60 active blinking tasks with both eyes. After a 5-minute test drive on the ETS2, they then performed a 60-minute task in which they attempted recognition by blinking while driving the car on ETS2, wherein, we sounded a tone once every 30 seconds to signal the user to initiate recognition by blinking. Finally, after completing the driving simulation, the participants completed three sets of a task in which they performed 60 active blinks with both eyes. The last task was designed to simulate recognition in an eye-strain condition.

**Results**

Table 6 presents the results of the evaluation of recognition accuracy under different conditions. The results demonstrate that authentication can be performed with high accuracy, with EER below 11%, in all verification conditions: before driving, during driving, and after driving. The lower accuracy in the "Before Driving" state can be attributed to the initial instability of blink movements, as seen in Fig. 10 to the right. The "While Driving" results show that high accuracy authentication can be achieved even in environments where other activities, such as driving, are mixed in.

# 6 DISCUSSION

We provide a table summarizing all the experimental conditions, evaluation perspectives, and main results for discussing the promising use cases (Sec. 6.1) and limitations as well as future study (Sec. 6.2) in Table 7. According to Table 7, in this research, we conducted a total of 6 user studies and evaluated the effectiveness of BlinkAuth from 17 perspectives. A table comparing the results with related works is shown in Table 1, and its advantages are summarized in the end of Sec. 1 and Sec. 2.3.

## 6.1 Promising Use Cases of BlinkAuth

BlinkAuth can realize single-time and continuous recognition using blinking. In the following, we discuss promising use cases for each recognition mode.

**Single-time recognition**

BlinkAuth can perform a single-time recognition using the user's intentional blinking. For example, BlinkAuth's single-time recognition mode is suitable for cases where recognition must be performed when the user's hand is not available, such as when unlocking a car, approving an automatic reply to a message received while driving, or unlocking a smartphone screen while cooking. In addition, as mentioned in the user comments discussed in Section 4.3, BlinkAuth can be used in situations wherein our hands are occupied and we cannot retrieve our smartphones, e.g., opening the auto-lock of a house with one's hands full or going through a ticketing gate while holding luggage. Moreover, BlinkAuth can be used for all smart glasses or VR headset applications that require recognition. As mentioned in Section 1, most VR headsets currently use a password or PIN-based recognition, which takes approximately 7.5 s on a PC and 12.8–13.2 s on a smartphone or tablet to complete recognition with an 8-character password [51]. VR headsets without keyboards generally require additional time for authentication, as characters must be entered using a controller with only a simple interface, such as buttons or triggers.

**Continuous recognition**

BlinkAuth can continuously verify the user's identity using their unconscious blink. In the continuous recognition mode, it is possible to check whether the same person continues to use the device or whether the device has been stolen and is being used by someone else. Examples of daily use of EOG include use cases where quadriplegics and severely paralyzed patients control a hospital alarm system by blinks or elderly people leverage eye movements and blinking to operate a wheelchair [26, 48]. In such situations, we expect BlinkAuth to be a promising approach for verifying user identity.

## 6.2 Limitations and Future Study

**User adaptation design for improving EER**

Evaluations in Sec 5.1 suggest that as time passes, user fatigue or sleepiness during later hours, along with inactivity over a week, may lead to changes in the biometric signal patterns necessary for recognition. These challenges can be mitigated through a model that uses incremental learning and online training to sequentially learn from authentication attempt data collected under these conditions. Security concerns arising from learning model updates, such as "biometric backdoors" [28], remains a future work.

Additionally, second measurements (skincare and while driving) showed improved accuracy than the first measurements (raw and before driving) in skin condition (Sec 5.3) and driving (Sec 5.4) evaluations. As indicated in Fig. 10 to the right (Sec 5.1), the accuracy can increase as users acclimate to the authentication actions. This highlights the importance of a well-designed registration interface for quick user adaptation, in addition to incremental learning and online training: a key aspect of future work for practical use.

**Changes in Features Due to Eye Injuries**

In this research, we conducted a user study primarily involving healthy individuals. Individuals with injuries, particularly around the eyes, might result in changes in their blinking features. Owing to ethical considerations, evaluating changes before and after an injury is difficult and out of the scope of this research. Therefore, it

**Table 7: Summary of experiments and main results.**

| Sec.# Experiment | Sec. 4.2 Accuracy | Sec. 4.3 Realtime | Sec. 5.1 Time Effect | Sec. 5.2 Attack | Sec. 5.3 Skin | Sec. 5.4 Action |
|---|---|---|---|---|---|---|
| # of Participants (M:F) | 31 (17:14) | 12 (5:7) | 13 (7:6) | 31 (17:14) | 15 (5:10) | 14 (6:8) |
| Age | 18–48 | 18–28 | 18–28 | 18–48 | 19–24 | 20–29 |
| Evaluation perspectives | Model<br>Wink<br>Continuous | Processing Time<br>SUS<br>Open-ended Q | Short period<br>Long period | Closed-set<br>Open-set<br>SAR | Raw<br>Skincare<br>Makeup | Before<br>While<br>After |
| Main Result (Condition) | EER=9.28% | Time=0.6 s<br>SUS=81.25 | EER=2.81%<br>(Short 16:00) | SAR=1.52%<br>(Open) | EER=3.56%<br>(Makeup) | EER=1.83%<br>(After) |

is essential to assess with special precautions in a medical setting in the future.

**User interface**

In this study, we developed a real-time recognition system as an implementation of BlinkAuth and evaluated its performance and usability. The user interface for processing user registrations was not implemented as a system, and we manually processed the data. Developing a usable user interface that systematically measures the EOG generated by conscious/unconscious blinking and trains a machine-learning model to register a user is a challenge to be addressed by future research.

**Need for the field experiment**

We conducted our user study in a controlled environment, aiming to minimize the effects of environmental noises. The experiment was carried out during the day in a room with blinds to block the sunlight from outside the windows, and the room's brightness was adjusted using a desk light. Since the user's simultaneous driving and recognition was evaluated using a driving simulator, we could not evaluate the effects of acceleration and vibration changes that occur when a real car is driven. Conducting field studies in more natural, diverse, and dynamic environments will be the subject of future work.

**Presentation Attacks by applying voltages**

In this research, we conducted an attack evaluation by presenting another person's data (See Sec.5.2). In real-world scenarios, applying voltage to the electrodes may realize a presentation attack on biometric signals. Although not specific to EOG studies, Eberz et al. discussed a method that involved presenting an electrocardiogram (ECG) waveform to ECG sensors to attack user recognition [13]. To realize a presentation attack through voltage application, it would be necessary to either directly modify the victim's hardware [13], attach electrodes directly to the victim, or secretly install electrodes in devices or accessories such as hats worn by users. Although these assumptions appear too strong for an attacker and may not be considered realistic, voltage application attacks could serve as an additional security assessment for EOG authentication. We note that conducting such tests would require further ethical considerations than this research (See 4), which did not involve invasive procedures or interventions on experiment participants.

Furthermore, to execute a presentation attack, the attacker must acquire a target user's EOG voltages. This would involve a data leakage or the direct application of electrodes to the victim, both of which are unlikely scenarios. Some interesting research in the biometrics field includes attacking the PPG authentication model using rPPG (remote PPG), which obtains Photoplethysmogram (PPG) signals by estimating from videos, without sensors [25, 40]. In future, it is possible that studies will be identified that estimate signals other than PPG, such as EOG, from videos, predicting the target's signals. Evaluating these attacks remains a challenge for future research.

**Realizing multimodal behavioral biometric recognition**

In this study, we showed that BlinkAuth users can be verified successfully while participating in other activities (see Sec 5.4). This fact suggests that BlinkAuth can be combined with other behavioral biometric recognition techniques to achieve multimodal recognition. Since BlinkAuth is a highly usable method with fast recognition, it is expected to be a promising candidate for multimodal recognition in combination with other behavioral biometric recognition methodologies [7]. The development and evaluation of a multimodal behavioral biometric recognition based on BlinkAuth is a challenge to be addressed by future research.

## 7 CONCLUSION

We developed BlinkAuth, a fundamental framework for behavioral biometric recognition using EOG signals generated during blinking. We implemented BlinkAuth and conducted a user study experiment with 31 participants. The results demonstrated that the average AUC and EER were 95.8% and 9.28%, respectively, showing a high accuracy for a behavioral biometric recognition technology. Furthermore, we found that BlinkAuth is robust against environmental variations, such as the time effect, the presence of the attacker, makeup around the eyes, the presence of other actions, such as driving a car, and eye strain. Blinking is a simple, quick action with the advantage of high usability and a large range of applicability, including compatibility with other actions. Field studies of BlinkAuth in diverse and dynamic environments, the development of a usable user interface required for user registration, and the development of multimodal recognition in combination with other recognition techniques are future research challenges.

## ACKNOWLEDGMENTS

We thank the anonymous reviewers for their helpful comments. This research was supported in part by JSPS KAKENHI Grant Number 22K17890, 22K19782 and Tateishi Grant Number 2221002.

## REFERENCES


[1] Sherif N. Abbas and M. Abo-Zahhad. *Eye Blinking EOG Signals as Biometrics*, pages 121–140. Springer International Publishing, Cham, 2017.

[2] M. Abo-Zahhad, Sabah M. Ahmed, and Sherif N. Abbas. A novel biometric approach for human identification and verification using eye blinking signal. *IEEE Signal Processing Letters*, 22(7):876–880, 2015.

[3] M. Abo-Zahhad, Sabah M. Ahmed, and Sherif N. Abbas. A new multi-level approach to eeg based human authentication using eye blinking. *Pattern Recognition Letter*, 82:216–225, 2016.

[4] Mohit Agarwal and Raghupathy Sivakumar. Blink: A fully automated unsupervised algorithm for eye-blink detection in eeg signals. In *57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1113–1121, 2019.

[5] R. Agarwal, T. Takeuchi, S. Laroche, and J. Gotman. Detection of rapid-eye movements in sleep studies. *IEEE Transactions on Biomedical Engineering*, 52(8):1390–1396, 2005.

[6] Patricia Arias-Cabarcos, Thilo Habrich, Karen Becker, Christian Becker, and Thorsten Strufe. Inexpensive brainwave authentication: New techniques and insights on user acceptance. In *30th USENIX Security Symposium*, pages 55–72, 2021.

[7] Patricia Arias-Cabarcos, Christian Krupitzer, and Christian Becker. A survey on adaptive authentication. *ACM Comput. Surv.*, 52(4), sep 2019.

[8] Patrick Berg and Michael Scherg. Dipole models of eye movements and blinks. *Electroencephalography and Clinical Neurophysiology*, 79(1):36–44, 1991.

[9] Shlomo Berkovsky, Ronnie Taib, Irena Koprinska, Eileen Wang, Yucheng Zeng, Jingjie Li, and Sabina Kleitman. Detecting personality traits using eye-tracking data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–12, New York, NY, USA, 2019. Association for Computing Machinery.

[10] BIOPAC. Mp36 system, 2023. https://www.biopac.com/product/upgrade-to-mp36-system/.

[11] John Brooke. *SUS-A quick and dirty usability scale." Usability evaluation in industry*. CRC Press, 1996.

[12] Donnell J. Creel. Chapter 33 - the electrooculogram. In Kerry H. Levin and Patrick Chauvel, editors, *Clinical Neurophysiology: Basis and Technical Aspects*, volume 160 of *Handbook of Clinical Neurology*, pages 495–499. Elsevier, 2019.

[13] Simon Eberz, Nicola Paoletti, Marc Roeschlin, Andrea Patané, Marta Kwiatkowska, and Ivan Martinovic. Broken hearted: How to attack ECG biometrics. In *24th Annual Network and Distributed System Security Symposium, NDSS*. The Internet Society, 2017.

[14] Raymond Fish and Leslie Geddes. Conduction of electrical current to and through the human body: A review. *Eplasty*, 9:e44, 10 2009.

[15] Gary Heiting, OD, All about vision. https://www.allaboutvision.com/resources/dominant-eye-test.htm, 2023.

[16] Ekansh Gupta, Mohit Agarwal, and Raghupathy Sivakumar. Blink to get in: Biometric authentication for mobile devices using eeg signals. In *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, pages 1–6, 2020.

[17] iMotions. Smi eye tracking glasses, 2023. https://imotions.com/products/hardware/smi-eye-tracking-glasses/.

[18] Nreal Japan Inc. Nreal air, 2022. https://www.nreal.jp/.

[19] Pupil Labs Inc. Neon, 2023. https://pupil-labs.com/products/neon/.

[20] ISO. Iso/iec 30136: Performance testing of biometric template protection schemes, 2018. https://www.iso.org/standard/53256.html.

[21] iver56. https://github.com/iver56/audiomentations, 2023.

[22] Sauro Jeff. *A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices*. Measuring Usability LLC, 2011.

[23] Xinyu Jiang, Xiangyu Liu, Jiahao Fan, Xinming Ye, Chenyun Dai, Edward A. Clancy, Dario Farina, and Wei Chen. Enhancing iot security via cancelable hd-semg-based biometric authentication password, encoded by gesture. *IEEE Internet of Things Journal*, 8(22):16535–16547, 2021.

[24] JINS. Jins meme, 2022. https://jins-meme.github.io/apdoc/en/.

[25] Lin Li, Chao Chen, Lei Pan, Jun Zhang, and Yang Xiang. Video is all you need: Attacking ppg-based biometric authentication. In *Proceedings of the 15th ACM Workshop on Artificial Intelligence and Security*, AISec'22, page 57–66, New York, NY, USA, 2022. Association for Computing Machinery.

[26] Yuanqing Li, Shenghong He, Qiyun Huang, Zhenghui Gu, and Zhu Yu. A eog-based switch and its application for "start/stop" control of a wheelchair. *Neurocomputing*, 275, 10 2017.

[27] Shuqi Liu, Wei Shao, Tan Li, Weitao Xu, and Linqi Song. Recent advances in biometrics-based user authentication for wearable devices: A contemporary survey. *Digital Signal Processing*, 125:103–120, 2022.

[28] Giulio Lovisotto, Simon Eberz, and Ivan Martinovic. Biometric backdoors: A poisoning attack against unsupervised template updating. In *IEEE European Symposium on Security and Privacy, EuroS&P 2020*, pages 184–197. IEEE, 2020.

[29] Giulio Lovisotto, Henry Turner, Simon Eberz, and Ivan Martinovic. Seeing red: Ppg biometrics using smartphone cameras. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 3565–3574, 2020.

[30] Shiqing Luo, Anh Nguyen, Chen Song, Feng Lin, Wenyao Xu, and Zhisheng Yan. Oculock: Exploring human visual system for authentication in virtual reality head-mounted display. In *27th Annual Network and Distributed System Security Symposium, NDSS*, 2020.

[31] A. López, F.J. Ferrero, M. Valledor, J.C. Campo, and O. Postolache. A study on electrode placement in eog systems for medical applications. In *2016 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pages 1–5, 2016.

[32] Ahmed Mahfouz, Tarek M. Mahmoud, and Ahmed Sharaf Eldin. A survey on behavioral biometric authentication on smartphones. *Journal of Information Security and Applications*, 37:28–37, 2017.

[33] Michael F. Marmor and Eberhart Zrenner. Standard for Clinical Electro-oculography. *Archives of Ophthalmology*, 111(5):601–604, 05 1993.

[34] Meta. https://www.meta.com/jp/en/quest/, 2022.

[35] Meta. Ray-ban stories smart glasses, 2022. https://www.ray-ban.com/usa/ray-ban-stories.

[36] NIST. Face recognition vendor test (frvt) part 2: Identification, nist ir 8271, 2023. https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8271.pdf.

[37] Personalized Beauty Discovery, Inc. https://www.colorescience.com/blogs/learn/how-to-apply-makeup, 2022.

[38] Angkoon Phinyomark, Sirinee Thongpanja, Huosheng Hu, Pornchai Phukpattaranont, and Chusak Limsakul. The usefulness of mean and median frequencies in electromyography analysis. In Ganesh R. Naik, editor, *Computational Intelligence in Electromyography Analysis*, chapter 8. IntechOpen, Rijeka, 2012.

[39] PLUX Biosignals. https://www.pluxbiosignals.com/collections/bitalino, 2023.

[40] Robert Mark Seepers, Wenjin Wang, Gerard de Haan, Ioannis Sourdis, and Christos Strydis. Attacks on heartbeat-based security using remote photoplethysmography. *IEEE Journal of Biomedical and Health Informatics*, 22(3):714–721, 2018.

[41] Sherif Seha, Georgios Papangelakis, Dimitrios Hatzinakos, Ali Shahidi Zandi, and Felix JE Comeau. Improving eye movement biometrics using remote registration of eye blinking patterns. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2562–2566, 2019.

[42] Mohammad Shahbakhti, Vahidreza Khalili, and Golnoosh Kamaee. Removal of blink from eeg by empirical mode decomposition (emd). In *The 5th 2012 Biomedical Engineering International Conference*, pages 1–5, 2012.

[43] SCS Software. Euro truck simulator 2, 2022. https://eurotrucksimulator2.com/.

[44] SONY. Playstation vr, 2022. https://www.playstation.com/en-us/ps-vr/.

[45] Jinani Sooriyaarachchi, Suranga Seneviratne, Kanchana Thilakarathna, and Albert Y. Zomaya. Musicid: A brainwave-based user authentication system for internet of things. *IEEE Internet of Things Journal*, 8(10):8304–8313, 2021.

[46] Shridatt Sugrim, Can Liu, Meghan McLean, and Janne Lindqvist. Robust performance metrics for authentication systems. In *26th Annual Network and Distributed System Security Symposium, NDSS 2019*. The Internet Society, 2019.

[47] Yuanyuan Tian and Jingyu Cao. Fatigue driving detection based on electrooculography: a review. *EURASIP J. Image Video Process.*, 2021(1):33, 2021.

[48] S. Venkataramanan, P. Prabhat, S.R. Choudhury, H.B. Nemade, and J.S. Sahambi. Biomedical instrumentation based on electrooculogram (eog) signal processing and application to a hospital alarm system. In *Proceedings of 2005 International Conference on Intelligent Sensing and Information Processing, 2005.*, pages 535–540, 2005.

[49] VIVE. Vive focus 3, 2022. https://www.vive.com/us/product/vive-focus3/.

[50] VIVE. Vive focus plus, 2022. https://business.vive.com/eu/product/focus-plus/.

[51] Emanuel von Zezschwitz, Alexander De Luca, and Heinrich Hussmann. Honey, i shrunk the keys: Influences of mobile devices on password composition and authentication performance. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, NordiCHI '14, page 461–470. Association for Computing Machinery, 2014.

[52] Jonathan Wolpaw and Elizabeth Winter Wolpaw. *Brain–Computer Interfaces: Principles and Practice*. Oxford University Press, 01 2012.

[53] Youming Zhang and Martti Juhola. On biometrics with eye movements. *IEEE Journal of Biomedical and Health Informatics*, 21(5):1360–1366, 2017.

[54] Huadi Zhu, Mingyan Xiao, Demoria Sherman, and Ming Li. Soundlock: A novel user authentication scheme for VR devices using auditory-pupillary response. In *30th Annual Network and Distributed System Security Symposium, NDSS*, 2023.


## A DETAILS OF THE EXPERIMENT IN SEC. 4.

We designed the experiment such that each participant completes it within approximately 60 minutes, including 10-minute breaks to minimize the burden on the participants. The flow of the experimental task for one participant is as follows:

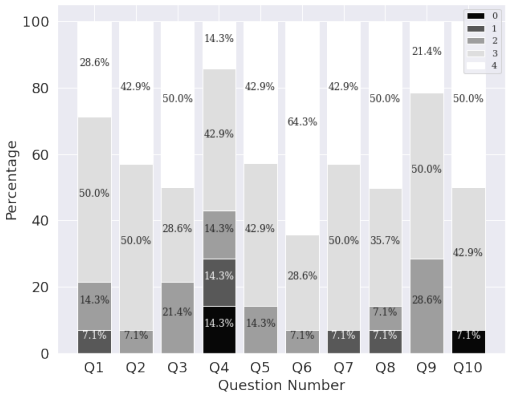


**Figure 14: Percentage of scores to each SUS question.**

1) Provide an overview of the experiment, including the informed consent agreement. (3 min)
2) Identification of the dominant eye. (2 min) (Sec 4.4)
3) Unconscious 60 blinks of both eyes (continuous mode) 3 min × 3 times
4) Active 60 blinks of both eyes (single-time mode) 3 min × 3 times + break (5 min)
5) Active 60 blinks of right eye (single-time mode) 3 min × 3 times + break (5 min)
6) Active 60 blinks of left eye (single-time mode) 3 min × 3 times

## B THE PROCEDURE OF THE DOMINANT EYE TEST.

Here, we distinguish between dominant and non-dominant winking. The following procedure is used to verify the dominant eye.

1) The participant stands, arms outstretched, and makes a triangle with both hands.
2) The participant lets an object that exists at the same height as their eyes be positioned exactly in the middle of the triangle made with both hands. In our experiment, we used a wall clock as the object.
3) Participants close one eye alternately. We determine that the closed eye is the dominant eye when the amount by which the object is displaced from the center of the triangle is greater.

## C DETAILS OF SUS RESULTS

Figure 14 to the left shows the percentage of scores for each question.


**Figure 15: Products used in the experiment. Skincare products: (S1) toner lotion, (S2) lotion, and (S3) sunscreen. Makeup products: (M1) primer, (M2) liquid foundation, and (M3) powder foundation.**

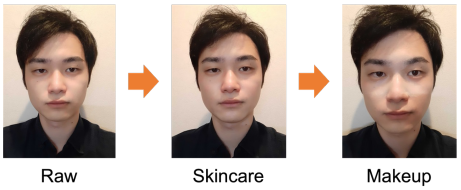


**Figure 16: Example of each skin condition in the robustness to skin conditions study. In the make-up state, all make-up products are attached around the area where the sensor touches.**

**Table 8: List of skincare / makeup products used in the experiment.**

| Group | ID | Types | Product name |
|---|---|---|---|
| Skincare | S1 | toner lotion | CEZANNE, Deep Moisture skin conditioner |
| | S2 | lotion | CEZANNE, Moisture Rich Essence Milk |
| | S3 | sunscreen | NIVEA, UV Super Water 50 Gel |
| Makeup | M1 | primer | CEZANNE Make Keep Base |
| | M2 | foundation | CEZANNE, Lasting Cover Foundation (Liquid) |
| | M3 | foundation | CANMAKE, Foundation (powder) |

## D DETAILS OF ROBUSTNESS TO SKIN CONDITIONS

To carefully select the products to be used in the experiment, we had discussions with five people who use both products on a daily basis. The criteria for selecting the products were: products that are considered effective in verifying the impact of moisture, common and inexpensive brands sold in most drugstores, and products that are skin-friendly so that they do not affect the skin of the participants. After discussion, we selected the products shown in Figure 15. Detailed product names are given in the Appendix Table 8. We also purchased two types of makeup remover, one liquid and one wet wipe type, so that participants could remove their makeup after the experiment. We asked participants who were in the habit of wearing makeup on a daily basis to select the type of product they normally use and apply it to the skin around their eyes. We also asked them to ensure that the amount of product they applied to their skin was the same as the amount they normally use.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009