2015

# The Electrooculogram and a New Blink Detection Algorithm

Samantha L. Epling

Matt Middendorf

Michael Hoepf

Christina Gruenwald

Lucas Stork

*See next page for additional authors*

THE ELECTROOCULOGRAM AND A NEW BLINK DETECTION ALGORITHM

SAMANTHA L. EPLING, Ball Aerospace and Technologies Corporation
MATT MIDDENDORF, Middendorf Scientific Services
MICHAEL HOEPF, Oak Ridge Institute for Science and Education
CHRISTINA GRUENWALD, Southwestern Ohio Council for Higher Education
LUCAS STORK, Southwestern Ohio Council for Higher Education
SCOTT GALSTER, Air Force Research Laboratory

Accurate and efficient real-time cognitive workload assessment has many important applications, and physiological monitoring has proven quite helpful with this assessment. One such physiological signal, the electrooculogram (EOG), can provide blink rate and blink duration measures. In a recent study, we developed and validated a robust blink detection algorithm based on the vertical EOG (VEOG). This algorithm does not require baseline data and is adaptive in the sense that it works for a wide variety of individuals without any experimenter adjustments. The performance of the algorithm is quantified using truth data based on video recordings. The algorithm produced blink rate and blink duration data for participants in a simulated remotely piloted aircraft experiment. Although this paper focuses on the blink detection algorithm, some results from the study will be included. Specifically, it was found that participants blinked fewer times and with a shorter duration in the more difficult experimental conditions.

A warfighter's dynamic cognitive workload is a ubiquitous concern, due in part to the close relationship between workload and performance (Cain, 2007). Methods of high workload detection and mitigation are therefore important areas of research. Cognitive workload is a complex and dynamic construct - it is dependent on many variables such as the operator's level of training, expertise, experience, motivation, time on task, and the difficulty of the task itself. The physiological measures used to assess workload are also affected by various factors such as fatigue, stress, engagement, and the environment (Wang & Zhou, 2013).

Physiological measures have proven quite useful in real-time cognitive workload assessment, in both laboratory and real-world settings. Wilson and Russell (2007) demonstrated that physiologically-determined adaptive automation could be used to significantly improve performance. Blink measures based on EOG can shed light on mental workload, and because of their noninvasive nature, they are well suited for real-time use.

Currently, despite the widespread use of eye blinks in research, a singular method of detection does not exist. Counting blinks from videotaped recordings is traditionally rejected due to the inordinate amount of time it takes. The use of cameras with complex eye tracking packages are an improvement technologically, but complexity and practicality remain a concern for field use. Blink detection using EOG has advantages over both of the approaches, and there is always a need to explore new detection algorithms.

The purpose of the current research is to present a blink detection algorithm that has a high accuracy level, is adaptive, dynamic, works in real-time, and does not require experimenter adjustments or calibration. The details of this algorithm and its validation using truth data will be presented. Results from a recent study employing this algorithm will be shown.

## Background

Various eye metrics, including blinks, have been shown to be useful indicators of cognitive state (Wang & Zhou, 2013). Blinks are classified in three ways. Voluntary blinks are those that occur with a conscious decision to shortly close the eye. Involuntary blinks involve both reflexive (startle) blinks, which occur to protect the eye in reaction to an external impetus, and spontaneous (endogenous) blinks, which are also reflexive but serve to maintain corneal moisture (Andreassi, 2007). Any blink mentioned henceforth refers to a spontaneous blink.

According to Andreassi (2007), blinks in a relaxed state occur at an average of 15-20 times per minute, have average amplitude of 380µV, and average duration of 120ms. Blink rate varies widely within an individual depending on the type of task, environment in which it is performed, and information processing demands. There are also large differences between individuals. For example, Sforza and colleagues (2008) showed that women spontaneously blink more frequently than men, and that younger people blink with more eyelid displacement than older people. Kong and Wilson (1998) claimed that such variability substantiates the need for blink detection algorithms, using the EOG signal, that are robust to noise, artifacts, and intra- and inter- individual variations.

Because blink rate and duration have been shown to relate to workload in environments with visual task demands (Wang & Zhou, 2013), the accurate detection of blinks holds great promise for continuous workload monitoring during many common human-computer interactions.

## Blink Detection Algorithms

While a blink in the VEOG is visually distinct to the human expert, the varied parameters and noisy signal make blink detecting algorithms quite difficult to create. Kong and Wilson (1998) filtered the signal four different ways before processing it with their algorithm. Blinks were then determined by finding a negative peak followed by a positive peak within a specified time window, along with other features used to give each potential blink a composite score. Another method, the workload assessment monitoring (WAM) system, also detects blinks using an algorithm that finds consecutive negative and positive slopes within a specified range (Wilson, 1994). However, unlike Kong and Wilson's approach, WAM criteria must be adjusted for every participant.

In order to detect blinks in the VEOG, a reliable algorithm is essential. Without accurately detecting the blinks, the extracted features (blink rate and durations) are less useful for assessing cognitive workload. Many researchers in the human factors, psychophysiology, ophthalmology, and human computer interaction domains have attempted blink detection in different ways, for different purposes. However, most literature available on these various approaches lack specific detail. Therefore, the present research fills that gap by presenting an algorithm that performs with a high level of accuracy and has a well-documented methodology.

## The New Blink Detection Algorithm

**Blink characteristics.** The basic shape of a blink in the VEOG signal has distinctive features. Andreassi (2007) describes the waveform as a sharp rise immediately followed by a sharp fall. The duration is short, the peak is rounded, and there is a noticeable overshoot before the signal returns to zero (Figure 1). Each time the VEOG signal goes above and below a threshold value is referred to as a bump. Data extraction software was written to compute a simple threshold, and use it to extract features from all bumps in the VEOG signal. The bumps in the signal include blinks and non-blinks (i.e., eye movements and noise). The features extracted by the software include the slope up at the midpoint, the slope down at the midpoint, the peak amplitude, and the duration at the midpoint.

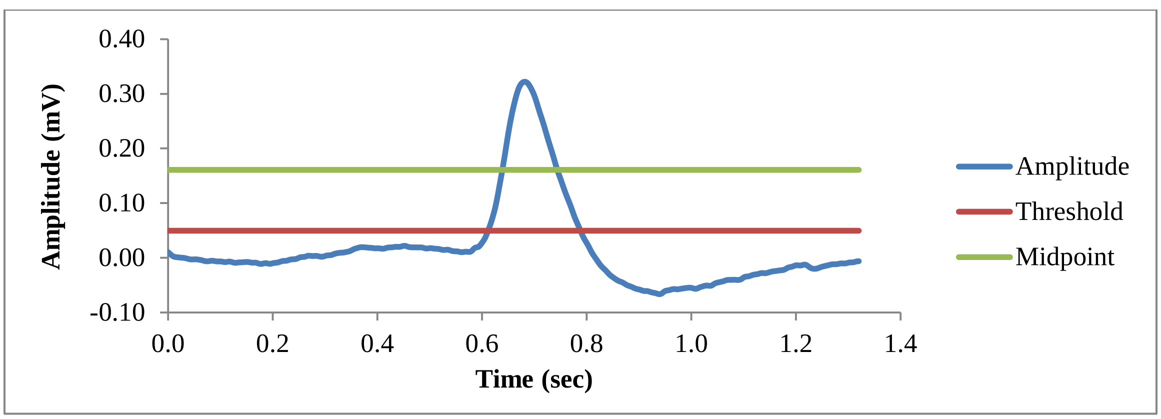


*Figure 1.* The basic shape of a blink.

**Primary criteria.** Not all excursions in the VEOG signal that go above and below threshold (i.e., bumps) are blinks, so criteria values needed to be established to distinguish blinks from non-blinks. To accomplish this, the data extraction software was used on a large database of existing VEOG data to extract the four features described in the above paragraph. Raters were trained to recognize the basic shape of a blink. The raters then visually observed each VEOG signal and coded each bump as a blink or non-blink. This data was used to determine eight criteria values needed to develop the initial blink detection algorithm. These eight criteria are the minimum and maximum values for slope up, slope down, peak amplitude, and duration at the midpoint. The extracted values for blink amplitude and duration were sorted and plotted for visual inspection (see Figure 2). Histograms were also created and the data was found to be normally distributed. The data for the two slope features were also examined and found to be normally distributed. The initial values for the primary criteria were determined using the 98$^{th}$ percentile of the distributions. The initial values were slightly tweaked following some testing. The final values of the primary criteria used in the detection algorithm are shown in Table 1.

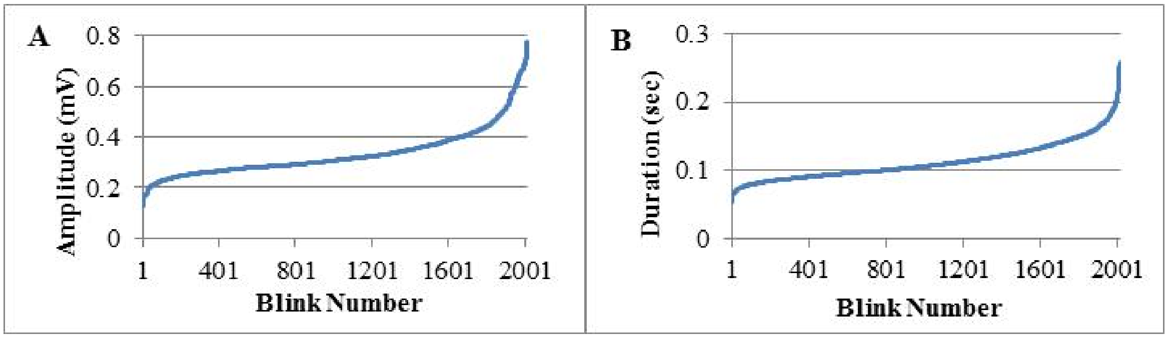


*Figure 2.* Blink amplitude (A) and duration (B) extracted from over 2000 blinks.

Table 1.
*Criteria Values for the Primary Features*

| Blink Feature | Minimum Criteria Value | Maximum Criteria Value |
|---|---|---|
| Amplitude (mV) | 0.1211 | 0.6483 |
| Blink Duration at the Midpoint (s) | 0.06 | 0.198 |
| Slope Up at the Midpoint (mV/s) | 1.5 | 13.41 |
| Slope Down at the Midpoint (mV/s) | -10.0 | -1.25 |

The existing database of VEOG data used above is from 12 participants performing four different tasks. This resulted in a total of 3102 bumps that went above and below threshold. The raters manually coded 2020 as blinks and 1082 as non-blinks. Note this is not absolute truth data because the raters were observing an electrical recording (VEOG) rather than a video recording. Therefore it is possible for a rater to occasionally miscode a bump. The use of video recording to generate actual truth data is discussed in the Algorithm Use and Validation section.

**Secondary criteria.** The initial blink detection algorithm was written to perform blink classification using the eight primary criteria identified above. New VEOG data was collected to test the classification logic. Each extracted feature from the new VEOG data had to fall within the range of the corresponding primary criteria values. For example, the amplitude must be between 0.1211 and 0.6483, otherwise the bump is not classified as a blink. The same logic is applied to the other three main features (slope up, slope down, and duration at the midpoint). For a bump to be classified as a blink, all four extracted features must be within their associated ranges. This classification logic was tested with additional new VEOG data and a few false positives were occurring.

To remedy this problem, five secondary features were extracted from each bump. In a manner similar to the primary features, criteria values were established for the secondary features. These features were used to provide a confidence assessment to refine classification accuracy. Specifically, all four of the primary features must fall within their associated ranges and three of five secondary features must meet their criteria values. Additional detail is provided in the scoring and classification section.

The five secondary features are the closure duration, the two $R^2$ values for linear fits at the midpoint and two additional duration measures (see Figure 3). The distance between the two linear fits at the peak is referred to as the closure duration. The distance between the two linear fits at the zero crossing is the blink duration at the zero crossing due to midpoint extrapolation. A similar duration is measured using linear fits about the threshold. The blink classification code was enhanced to incorporate the five secondary features. The number of false positives produced by the algorithm was substantially reduced. The actual values of the secondary criteria used in the detection algorithm are shown in Table 2.

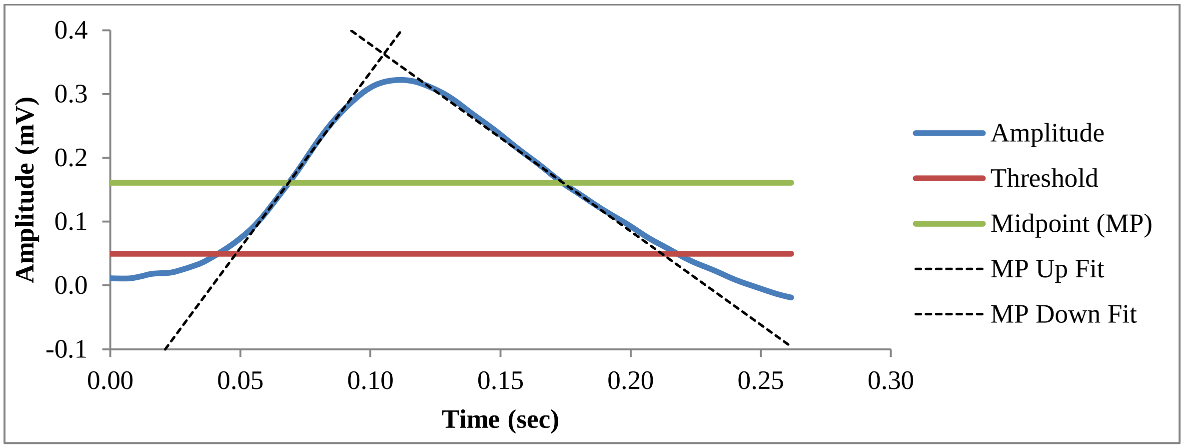


*Figure 3.* A typical blink with linear fits about the midpoints.

Table 2.
*Criteria Values for the Secondary Features*

| Blink Feature | Minimum Criteria Value | Maximum Criteria Value |
|---|---|---|
| Closure Duration (s) | 0.01 | 0.10 |
| Slope Up at the Midpoint $R^2$ | 0.996 | N/A |
| Slope Down at the Midpoint $R^2$ | 0.995 | N/A |
| Blink Duration ZCMP (s) | 0.1162 | 0.3 |
| Blink Duration ZCT (s) | 0.1 | 0.35 |

*Note.* ZCMP is duration at the zero crossing due to midpoint extrapolation. ZCT is duration at the zero crossing due to threshold extrapolation.

**How the Algorithm Works**

The major components of the blink detection algorithm are threshold generation, feature extraction state machine, scoring and classification, and blink save and false positive detection logic.

**Threshold generation.** The threshold generation approach uses a sliding five second window of raw VEOG data. To minimize the effects of blinks and eye movement on the threshold, the data is high pass filtered using a first order Butterworth filter with a break frequency of 10 Hz. This essentially leaves in the “noise” from which the threshold is calculated. The filtered signal is then rectified and the median is taken for the *raw* threshold value. The median is used because the data in the five second window can be highly skewed when there is a blink in the window.

The second stage of threshold generation imposes limits on the *raw* threshold and adds in a threshold reduction value to accommodate double and multiple blinks. Initially the threshold limits are static, but after ten blinks have been detected, the limits are dynamic based on the mean amplitude of the recorded blinks. The threshold reduction value is necessary due to the high pass filter in the signal acquisition hardware, which causes the signal to overshoot zero on the down slope of the blink. If the blink is immediately followed by another blink the subsequent blink starts below zero (Figure 4). If the threshold reduction value is not applied, the subsequent blink(s) may be easily missed. The threshold reduction value is based on the amount of overshoot of the previous blink. The threshold returns to its normal (non-reduced) value using a function that is the inverse of the high pass filter implemented in the signal acquisition hardware.

**Feature extraction state machine.** This state machine uses the threshold to monitor the VEOG signal. The state machine has four values (0, 1, 2, and 3). In state zero the logic waits for the signal to be below threshold. In state one it waits for the signal to go above threshold, at which time upward threshold crossing data is captured and the threshold is frozen. In state two the logic is waiting for the signal to go back below threshold. During this time peak data and downward threshold crossing data are captured and the threshold is unfrozen. In state three the signal overshoot value is captured and the extracted features are scored to see if the signal excursion above and below threshold is a blink. The state machine then returns to state zero.

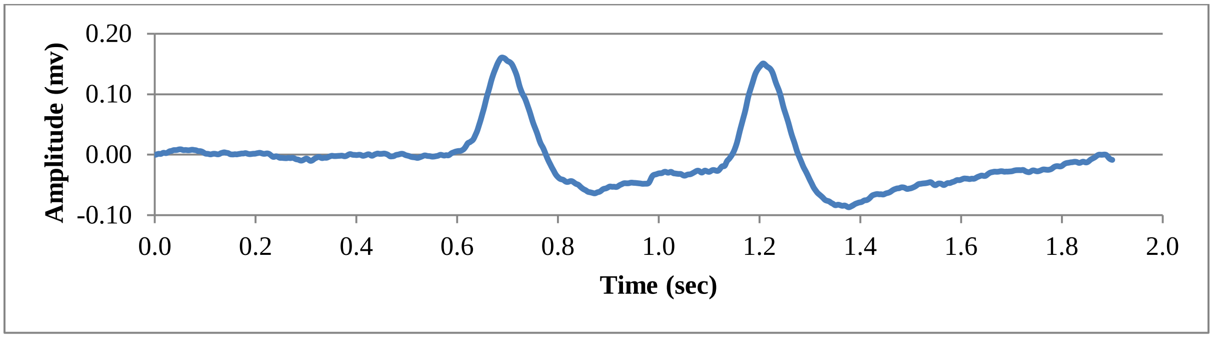


*Figure 4.* The threshold reduction logic is needed when multiple blinks occur in a short time frame. Because of the overshoot following a blink, the next blink starts from a lower value.

**Scoring and classification.** The VEOG bump that goes above and below threshold is scored using criteria values described in the previous sections. One point is awarded when each of the four primary criteria are met and one tenth of a point is awarded when each of the five secondary criteria are met. Therefore, the maximum score for a VEOG bump is 4.5 points. Bumps that score 4.3 points or higher are reliably classified as blinks. This requires that all four of the main features be met, and at least three of the secondary features be met. Requiring scores higher than 4.3 points results in some blinks being missed. Allowing scores lower than 4.3 results in some false positives.

**Blink save and false detection logic.** This logic applies to a very small number VEOG bumps. Bumps that fail only one of the four main criteria, but otherwise have a nearly perfect score (3.4 and 3.5), are given a second look. When a bump fails the maximum amplitude criterion, the criterion can be adjusted upward using amplitude data from previous blinks (minimum of 10 required). A likewise, adaptive test is applied when the minimum amplitude fails or the slope down at the midpoint fails. Currently only one false positive test is performed. Bumps that have two peaks are rejected.

## Algorithm Use and Validation

**Experimental results.** In a recent study participants were asked to track targets using remotely piloted aircraft. Workload was experimentally manipulated and physiological measures were collected. VEOG data was processed using the blink detection algorithm discussed in this paper. For the sake of brevity, only the blink rate and duration results are discussed here. For a full discussion of the experiment, see Hoepf, Middendorf, Epling, and Galster, this volume. In the study, high workload had a statistically significant effect on blink rate and duration. Blink rate was slower and blink duration was shorter. It was encouraging that the algorithm was sensitive to small changes in the blink measures due to the workload manipulation.

**Truth data validation.** Video recordings were used to generate truth data to help validate and quantify the blink detection algorithm. Eight participants were video recorded while performing trials in a recent experiment. Two trials were recorded for each participant using a Basler high speed camera. The output of the blink detection algorithm was evaluated by two separate individuals using the video recordings. Both individuals watched the recording of each trial and noted each time the participant blinked. If a participant blinked and the algorithm did not detect the blink, a “miss” would be counted. If the algorithm detected a blink when there was not one observed, it was classified as a “false positive.” Only 2.5 percent of blinks were missed, whereas 1.0 percent of blinks were falsely detected. Overall, the blink detection algorithm had an accuracy rating of 96.7 percent

## Discussion

In our recent study we collected eye activity data using both EOG and a camera-based eye tracking system. Two advantages of the camera-based system are its completely off-body and it produces position measures. For example, its eye lid opening measure is in meters. The EOG signal is an electrical measure and cannot be directly related to position. Due to drift in the EOG signals, a high pass filter is commonly used in the signal acquisition hardware. Therefore EOG is good for detecting rapid eye movements, but is not good for measuring gaze angles.

An advantage of EOG is that it does not have restricted field-of-view. The camera-based system can lose its lock on the eye if the participant slouches, changes seating position, or turns their head. Conversely, the EOG signal remains continuous regardless of participant movement. In our experiment participants needed to occasionally look down at the keyboard to press a key. When they did this the camera-based system stopped producing data, whereas the EOG approach did not.

The blink detection algorithm discussed here is adaptive in the sense that it works well for a wide selection of individuals. In addition, after the algorithm has compiled statistics on a few blinks, it can adapt some on the criteria to improve its classification accuracy. The algorithm is also dynamic in the sense that the detection threshold will change in real-time in response to changes in the VEOG signal.

A positive aspect of the blink detection algorithm is that it does not require baseline data or calibration. There is no need (or mechanism) for experimenter adjustments. This algorithm produces measures in real-time, which is an advantage over *post hoc* approaches.

## Conclusion

The new blink detection algorithm discussed in this paper works extremely well, in regard to both misses and false positives. It has served well as a tool to support the analysis of electroencephalogram (EEG) data using artifact separation (Credlebaugh, Middendorf, Hoepf, & Galster, this volume). The algorithm produced blink rate and blink duration measures that are sensitive to changes in cognitive workload.

## Acknowledgements

The authors would like to thank Chelsey Credlebaugh and Jonathan Mead for their assistance in data reduction and Chuck Goodyear for his help with statistical analysis. We would also like to thank Kevin Durkee, Noah DePriest, and Mark Squire for their technical support. The views expressed in this report are solely those of the authors and do not necessarily reflect the views of the employers or granting organizations.

## References

Andreassi. Psychophysiology: Human Behavior and Physiological Response 5th ed. NY, Taylor & Francis Group, LLC. 2007.

Cain, B. (2007). A review of the mental workload literature, Technical Report. Defense Research and Development Canada Toronto.

Credlebaugh, C., Middendorf, M., Hoepf, M., & Galster, S. (this volume). EEG data analysis using artifact separation. In *Proceedings of the Eighteenth International Symposium on Aviation Psychology*, Wright State University.

Hoepf, M., Middendorf, M., Epling, S. & Galster, S. (this volume). Physiological indicators of workload in a remotely piloted aircraft simulation. In *Proceedings of the Eighteenth International Symposium on Aviation Psychology,* Wright State University.

Kong, X. & Wilson, G. (1998). A new EOG-based eyeblink detection algorithm. *Behavior Research Methods, Instruments, & Computers. 30*(4), 713-719.

Sforza, C., Rango, M., Galante, D., Bresolin, N., & Ferrario, V. (2008). Spontaneous blinking in healthy persons: An optoelectronic study of eyelid motion. *Opthal. Physiol. Opt. 28*, 345-353.

Wang, Y. & Zhou, J. (2013). Literature review on physiological measures of cognitive workload. Machine learning research group – NICTA

Wilson, G. & Russell, C. (2007). Performance enhancement in an uninhabited air vehicle task using psychophysiological determined adaptive aiding. *Human Factors, 49(6),* 1005-1018.

Wilson, G. F. (1994). Workload assessment monitor (WAM). In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 38* (15), pp. 944-944). SAGE Publications.