

Tracking Early Differences in Tetris Performance using Eye Aspect Ratio Extracted Blinks

Gianluca Guglielmo, Michal Klincewicz, Elisabeth Huis in 't Veld, and Pieter Spronck

Abstract— This study aimed to evaluate if eye blinks can be used to discriminate players with different performance in a session of Nintendo Entertainment System (NES) Tetris. To that end, we developed a state-of-the-art method for blink extraction from EAR measures, which is robust enough to be used with data collected by a low-grade webcam such as the ones widely available on laptop computers. Our results show a significant decrease in blink rate per minute (blinks/m) during the first minute of playing Tetris. After having defined 3 groups of proficiency based on in-game performance (Novices, Intermediates, and Experts) we found out that expert players display a significantly lower decrease in blinks/m compared to novices during the first minute of gameplay, which shows that Tetris players' proficiency can be detected by looking at eye blinks/m variations during the early phase of a game session. This difference in blinks/m is observed throughout the entire game session, which supports the general conclusion that proficient Tetris players have a lower decrease in blinks/m, even when playing more difficult levels. Finally, we offer some interpretations of this effect and the relationship that our results may have with the visual cognitive workload experienced during the gameplay

Index Terms—Eye blinks, Expertise, Machine learning, Performance, Video games

I. INTRODUCTION

PLAYING a video game for hours typically leads to mastery of its mechanics. Unsurprisingly, experienced players behave differently than novices when they play a video game [1], thus showcasing this mastery and developing strategies for their effective use to achieve game-related goals. This study aimed to assess whether a player's blinks rate per minute (blinks/m) during the first minute of a game session of Tetris already provides enough information to discriminate between participants with different levels of proficiency in the game. In Tetris, the players have to successfully place Tetrazoids (henceforth zoids), the tile falling from the upper part of the game environment, in order to clean lines and avoid reaching the upper edge of the game environment. To do so, the player has to successfully clean lines placing zoids correctly avoiding leaving unoccupied spaces. The game increases its difficulty as the player clears more lines; more specifically the level increases every time the player clears a number of lines given by the current level $\times 10 + 10$. The speed of the zoids' fall increases as the player moves to harder levels. Given the specifications of Tetris, we expect that proficient players cope better, in terms of cognitive load, with harder levels than less proficient players. We speculated that expert participants, defined by their in-game performance [2, 3, 4, 5], may experience a lower cognitive workload compared to less experienced ones and that this would be reflected in their blinks. In typical circumstances, humans blink on average about 15 to 20 times per minute, with each blink lasting between 50 and 500 milliseconds (ms) [6, 7]. A typical spontaneous blink rate in general ranges between 2.8 and 48 per minute [8] with

inter-blink intervals lasting between 2 and 10 seconds [9] although they can be as short as 100 ms [8]. Neuroimaging studies have pointed towards a relationship between blinks and dopaminergic activity, suggesting that blinks may be an appropriate proxy to measure cognitive performance and goal-oriented behavior [9,10,11]. For example, a higher blinking frequency has been found to be positively associated with the ability to ignore specific stimuli in a Go/Nogo task, an inhibition task where the subject has to respond to the word "Go" by pressing a button and simply ignore the "No go" word when is presented [12]. On the other hand, blinking rates have been found to decrease during tasks requiring visual attention such as reading [13] or playing video games [14]. The rate of blinks/m also has an inverse relationship with cognitive load [10] during a task and its level of difficulty [15]. This has been observed in expert video game players during a Hearthstone tournament, suggesting that gameplay is less cognitively demanding for them [16]. Based on this related literature, we hypothesize that proficient Tetris players may also have lower variations of blinks/m. Such a difference may manifest not only throughout the entire game session but already in the early phase of the gameplay, after having corrected for their baseline [17,18]. Such correction is applied since the raw blinks/m during a task may be affected by not accounting for the individual blinks/m at rest.

Previous studies have already shown that some early behaviors are predictive of performance in longer Tetris sessions finding results based on in-game behaviors such as zoids positioning or keyloggers [3, 4]. These behaviors, extracted during the early phases of a game session, may be enough to detect expert and proficient players in longer sessions. For example, previous studies suggest that the ability of experts to clear a higher number of minimum lines and manage piles of zoids already emerges at level 0 [2]. Another independent study showed that the keys pressed by the players during the first 45 seconds of a game session, with specific reference to the down key and left key, are sufficient to discriminate proficient players from less proficient ones based on performance in longer sessions [4]. We, on the other hand, take note that making complex decisions quickly and with a desired result is cognitively demanding, but this can improve with practice and expertise reducing the cognitive demands required from the player. Blinks can reflect this given their indirect connection to cognitive load [10].

However, up to date, no study investigated the connection between early variations in blinks/m and the performance obtained in Tetris. For this reason, in this study, we analyze the baseline corrected blinks/m that players have during the first

minute of gameplay. We decided to use the first minute since, during this time frame, all the players still played at level 0, and there was no effect due to the level of difficulty. As a consequence, this study aims to investigate if the baseline corrected blinks/m during the first minute of gameplay can be used to discriminate players obtaining a different performance on a full Tetris session.

II. EXPERIMENT 1

A. Methods

A total of $n = 160$ participants were recruited to participate in this experiment (55% female, Mage= 21.81, SDage = 4.99; and 2 participants did not declare their sex). All participants filled out an informed consent form and were informed that their faces would be recorded with the laptop cameras (720 pixels, 16:9 aspect ratio, and at a rate of 30 frames per second). The use of cameras characterized by such a frame per second recording should suffice to capture blinks since their duration spans between 50 ms and 500 ms [6,7] (each frame of a 30 fps recording = 33.34 ms) and generally, blink rate is evaluated on minutes [9, 10]. This study was approved by the Ethics Committee of Tilburg University under code REDC 2021.35a.

After providing their consent, each participant was asked to watch a looped video for three minutes to establish their baseline blinks/m [19]. The video was selected from the OpenLAV library dataset [20] on account of it being emotionally neutral¹. Each video recorded using the Open Broadcaster Software (OBS) contained both the participant's face and the screen recording. The first author of this work manually annotated the blinks of 160 participants during the recording and counted the number of blinks to obtain the ground truth

The videos collected with OBS were cropped so they contained just the participants' faces while playing the game. Then, the Eye Aspect Ratio (EAR), a measurement that calculates the distance between the upper eyelid and the lower eyelid, was extracted [16, 21]. The code to extract the EAR was implemented in Python and more specifically using the cv2 library (version = 4.5.5) and cvzone (version = 1.5.6) libraries. The cv2 library was used to detect the participants' faces while the landmarks of interest to measure the EAR were detected using the FaceMeshDetector available in the cvzone library [16], as in Figure 1. The number of landmarks needed to detect the EAR and their coordinate position as specified in the FaceMeshDetector is shown in Figure 2.

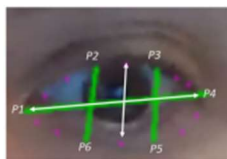


Fig. 1. The FaceMeshDetector's landmarks and their position on the eyes.

After having plotted the points using the previously mentioned method, the EAR is extracted using the following formula $(|P2-P6|+|P3-P5|)/(2|P1-P4|)$. The final EAR is an average calculated for both eyes since blinks occur when both eyes are closed, with a value ranging between 0.20 and 0.45. The EAR per each frame and the landmarks plotted on the participants' faces are visualized in Figure 2; please note that EAR is multiplied by 100 [16, 21].

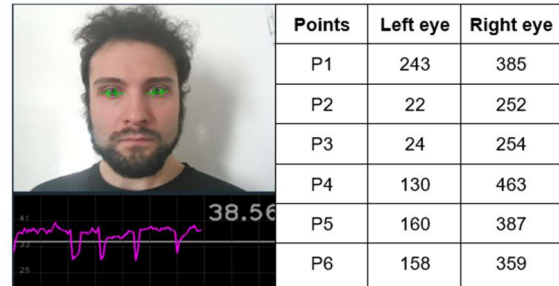


Fig. 2. On the left, an example of FaceMeshDetector landmarks plotted on the participant's face and the online EAR recording. The peaks represent moments in which the participant blinked. On the right, the FaceMeshDetector landmarks' numbers to identify the left and right eye

There are several methods for automatic blink detection. To assess which method performed best, we applied and compared four methods: the rule-of-thumb [35], rule-of-thumb with differentiated thresholds [16], isolation forest, and Isolation forest with filtering. The first rule-of-thumb method uses a threshold of at least 3 consecutive frames of approximately 100 ms being below EAR 30 [35]. This may be an unreliable measure since the EAR signal is influenced by the distance of the face from the screen, the position of the face, the shape of the participant's eyes, and a 100 ms duration does not fit with a possible blink duration between 50 ms and 500 ms [7]. For this reason, other studies used a threshold-based algorithm with filtering, which used three distinct thresholds (EAR 25, 30, and 35) [16]. This second method of extracting blinks from EAR looks for a sequence of frames by selecting sequences that are between 50 ms and 500 ms (approximately 2 and 15 frames long), which corresponds to the assumed lower and upper bounds of a typical blink duration. Then it looks for the minimum interval between blinks by comparing the distance between two continuous sequences and if this difference is lower than 9 frames (approx. 300ms) then the two sequences are concatenated to form one sequence corresponding to a blink [16]. The third method uses an isolation forest; an outlier detection algorithm that identifies sequences of outlier EAR values and treats them as belonging to blinks [22]. Specifically, this algorithm uses 100 trees and a contamination parameter, which is the expected percentage of outliers in the data, based on 3 standard deviations from the mean using a 100 frames-wide rolling window [22]. Afterward, the Isolation Forest can be set to detect a blink every time an outlier is detected for more

¹ the video is available at the following link https://www.youtube.com/watch?v=dHG_eKFJHtM and it is authored by Adrian Soare

than 2 consecutive frames in the sequence of EAR measures. However, this approach still has one of the problems of the rule-of-thumb approach [21], since it does not take into account the potential interval between blinks and the wide range (50ms – 500ms) of blink durations.

The last method and the one developed for this study adapts the isolation forest method based on a rolling window of 100 frames by adding a new definition of the contamination parameter, which uses the median with a median absolute deviation of 2.5. This specific value for the median absolute deviation is suggested by previous studies that focus on outlier detection [23, 24]. We also defined an upper and a lower threshold (to account for blink duration) of respectively 2 and 15 frames, approximately equal to 50 - 500 milliseconds, rather than the 2 frames used in the previous method with isolation forest [22]. At this stage of the process, the isolation forest could detect groups of outlier frames with a length between 2 and 15 frames that may be candidates to be parts of blinks. However, as suggested in previous studies, blinks have at least 100 ms between one and another [14]. For this reason, the new isolation forest method takes all intervals shorter than 4 frames (132 milliseconds considering that 3 frames still represent a duration lower than 100 milliseconds) occurring between two groups of frames detected as outliers and marks them as outliers. This was done to filter potential noise or frames belonging to blinks that the algorithm failed to detect. Afterward, all outlier sequences with a length between 2 and 15 frames were labeled as blinks.

Finally, we compared the mean number of blinks found across the 3 minutes recording of the rule-of-thumb method, the rule-of-thumb method with differentiated thresholds (20, 25, and 35), the isolation forest method, and our new isolation forest with filtering method using the mean absolute percentage error (MAPE) and Pearson's r coefficient (r).

B. Results

The results demonstrate that the isolation forest with filtering method developed for this study performs the best (see Table 1). Furthermore, statistical analysis using a Mann-Whitney U test accounting for the non-normality of the residuals shows that only for this method, the extracted average blinks/m on 3 minutes across all the participants (M) does not statistically differ from the ground truth blinks/m on 3 minutes (U).

TABLE I.
COMPARISON WITH 4 METHODS WITH CORRELATION COEFFICIENT (PEARSON'S R) AND MAPE.

	r	MAPE (%)	Blinks/m of 3 min (M, SD)	U	p
Ground truth			19.66 (12.17)		
Rule of Thumb [21]	0.58	60.91	14.02 (16.01)	17343	$p < .001$
Rule of Thumb with diff. thres. [16]	0.61	70.90	25.98 (17.00)	9877	$p < .001$
Isolation Forest [22]	0.50	45.58	9.65 (7.02)	19911	$p < .001$
Isolation Forest with filtering	<u>0.88</u>	<u>27.98</u>	19.76 (10.56)	12278.5	$p = .53$

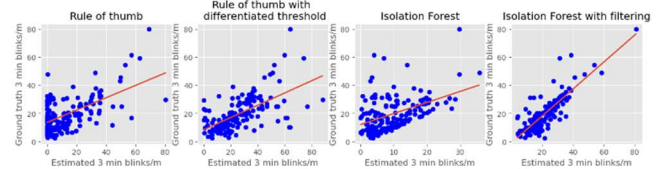


Fig 3. Correlation plot for the 4 methods adopted with the best fitting line. The Y axis represents the ground truth blinks/m on 3 minutes while the X represents the blinks/m on 3 minutes detected by the 4 methods compared in this study.

Moreover, our isolation forest with filtering method makes it possible to extract other features from EAR measures; an overview of these features is in Appendix A. The code for the Isolation Forest with filtering implemented for this work can be found at the following GitHub link: <https://github.com/G-Guglielmo/Blinks-tracker>.

III. EXPERIMENT 2

A. Methods

Eighty participants (49.37 % female, $Age = 22.27$, $SD = 5.92$) were recruited through the SONA participant recruitment platform to participate in this study. All participants filled out an informed consent form and were informed that their faces would be recorded with the laptop camera. One participant declined to answer the question regarding biological sex. Participants recruited at Tilburg University received a

formative credit required by their program of studies. This study was approved by the Ethics Committee of Tilburg University under code REDC2021.35.

A full experimental session (introduction, test session, game session, debrief) typically lasted approximately 40 minutes. At the beginning of the session, after having given informed consent, participants were asked to self-assess their experience with the Tetris game on a Likert scale between 1 (not experienced at all) and 5 (really experienced) and how many hours they spent playing video games, different than Tetris, every week. At this point, participants were asked to watch a 3-minute neutral video (see Methods of Experiment 1) to obtain their baseline blinks/m. Afterward, the experimenter introduced the participants to NES Tetris demonstrating how the 5 keys of a laptop computer keyboard can be used to play the game (Down for a forced drop, Left for a left translation, Right for a right translation, X for a clockwise rotation, and Z for a counter-clockwise rotation). Videos of participants' faces were taken during each session with the laptops' cameras (720 pixels, 16:9 aspect ratio, and at a rate of 30 frames per second). The actual gameplay was also recorded with Open Broadcaster Software (OBS).

After being introduced to the game, participants were given 2 minutes to practice with the NES Tetris. During this phase, the experimenter explained that players could use the information provided by the “Next” zoid area to plan where they will place the next zoid and that the “Statistics” square on the gameboard provided real-time information about all the zoids placed during the match so far. An overview of the information provided by the game to the players can be found in Figure 4

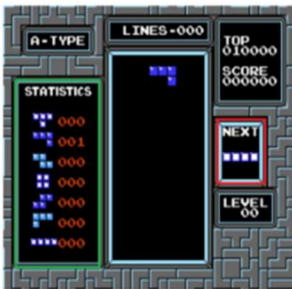


Fig. 4. The classic NES Tetris layout where the red square shows the “Next” zoid and the green square provides “Statistics” about the zoids placed

After the 2-minute practice session, participants played NES Tetris for 13 minutes on the same laptop. In case of a lost match (zoids filling up the game board), participants were asked to restart the game at level 0.

Previous studies defined Tetris expertise using the average score obtained across completed matches [3, 4, 5] and/or the number of matches played in a limited amount of time [2,4]. According to these approaches, the most expert players typically play fewer matches and obtain a higher score. We clustered the number of matches played by the players and the average score obtained across completed matches to define expertise. The number of matches that each participant played

had a limited skewness of 0.34 [25], however, the average score was severely skewed (skewness = 1.09) [25]. Therefore, the average score was log-transformed [2,6]. Participants were then clustered into a Novice, Intermediate, or Expert group using a k-means algorithm and the elbow method [2,6,5] similar to other studies [2,6,5]. Our participants had an average score of 4291.93 ($SD = 6771$) across completed matches and played an average of 2.76 matches ($SD = 1.88$). The only exceptions were 2 players who did not lose across the 13 minutes of gameplay; for this reason, we included their score obtained at minute 13 [2]. The clusters (proficiency groups) used for this study, representing the 3 groups with different levels of performance, are the same as the ones used in a previous study that used early keystrokes to predict levels of performance in Tetris [4]. It is standard to group game players into brackets, or broad categories, such as expert or novice, and we followed this approach here.

For what concerns blinks' extraction, first, EAR was extracted from video recordings of baseline (neutral video watching) and the 13 minutes of gameplay for each participant using the cvzone library. Second, an isolation forest with filtering (best method conveyed in experiment 1) was trained on both the baseline EAR values and complete gameplay EAR values for all participants. Third, using this model, the number of blinks from the 3-minute baseline EAR values for each participant and from the first minute of gameplay were extracted. At this point, a baseline corrected value of the blinks/m of the first minute of gameplay was calculated by subtracting from the baseline value for each participant, to account for the individual differences in blinks/m at rest [17, 18].

Finally, a two-way ANOVA was run with the baseline corrected blinks/m during the game as the dependent variable and biological sex (control variable) and the level of expertise (novice, intermediate, expert) as independent variables. Before modeling, assumptions concerning the normality of the residuals were checked by comparing the data against a normal distribution using the Kolmogorov-Smirnov test [26] and concerning the homogeneity of variance using the Bartlett's test [27].

B. Results

The number of matches, average scores, and self-assessed Tetris experience are visualized in Table 2. A significant main effect was found for self-assessed experience ($F(2,77) = 9.32$, $p < .001$), and posthoc analyses run using an Holm correction, revealed that experts had higher self-assessed Tetris experience than Intermediates ($p < .05$), who in turn had higher self-assessed Tetris experience than Novices ($p < .001$). Across all the matches played, the Expert group reached, on average, level 5.76 ($SD = 2.13$) while the Intermediate group and the Novice group respectively reached level 0.99 ($SD = 0.54$) and level 0.11 ($SD = 0.16$). These differences, after having run a Kruskal-Wallis test accounting for the non-normality of the residuals, resulted to be significant ($H(2) = 65$, $p < .001$). More specifically, running a Dunn post-hoc correction for non-parametric analysis, Experts reached a higher level than

Intermediates ($p < .001$) and Novices ($p < .001$) while Novices reached a lower level than intermediates ($p < .001$). No significant effect of hours playing video games every week was found when running a Kruskal-Wallis test, across the 3 groups (Experts: $M = 2.21$, $SD = 2.37$; Intermediates: $M = 1.80$, $SD = 1.94$; Novices: $M = 1.77$, $SD = 2.76$; $H(2) = 2.50$, $p = 0.29$). The 3 groups had a balanced number of females and males; the Novices were composed of 11 males and 11 females, the Intermediates of 20 females and 17 males, and the Experts of 12 males and 8 females. Females reported overall to have a lower self-assessed experience in Tetris ($M = 1.95$, $SD = 0.81$) than males ($M = 2.5$, $SD = 0.87$; $t(77) = 2.88$, $p < .01$). However, running a Mann-Whitney U test accounting for the non-normality of the residuals, this did not result in a significant difference in terms of the average score when comparing males ($M = 4904.5$, $SD = 7312.22$) and females (females: $M = 3720.85$, $SD = 6296.75$) throughout the 13 minutes game session ($U = 575$, $p = .07$) and in the average number of matches played (males: $M = 2.83$, $SD = 1.69$; females: $M = 2.84$, $SD = 2.01$; $U = 725.5$, $p = .73$). These results show no statistically significant differences between the biological sexes when looking at the variables used to cluster the participants in the 3 levels of performance.

TABLE II.

RESULTS OF CLUSTERING ALGORITHM AND CORRESPONDING AVERAGE SELF-ASSESSED TETRIS EXPERIENCE PER GROUP

	Novices	Intermediates	Experts
Matches played	5.22 (1.41)	2.32 (0.86)	1.1 (0.30)
Average score	305.69 (147.07)	1483.26 (971.14)	14013.45 (7556.00)
Tetris experience	1.73 (0.69)	2.21 (0.87)	2.80 (0.75)

In our sample, participants had an average blinks/m of 19.47 ($SD = 10.80$) during baseline and an average of 13.22 ($SD = 7.92$) during the first minute of gameplay. To investigate the difference between baseline and the first minute of gameplay between males and females, and the effect of self-assessed experience on blinks/m (given the difference between males and females), we run a repeated measures ANOVA with biological sex and Tetris self-assessed experience as a between-subject factor and time (baseline, first minute of the session) as a within-subject factor. Indeed, we found that females ($M = 21.64$, $SD = 11.79$) have higher blinks/m than males ($M = 17.42$, $SD = 9.30$; $F(1,74) = 5.81$, $p = .02$) and that blinks/m are lower during the Tetris session than during the baseline ($F(1,76) = 28.85$, $p < .001$). Tetris self-assessed experience had no effect on the blinks/m decrease during the first minute of gameplay ($F(1,74) = 0.01$, $p = 0.97$). No interaction effect of biological sex and time was found ($F(1,76) = 0.06$, $p = .95$). Given the

different self-assessed experience of males and females in our sample, we also checked the interaction effect of biological sex and Tetris experience. Such interaction term resulted not to be significant ($F(1,74) = 0.39$, $p = .54$).

Finally, a two-way ANOVA (after having confirmed the normality of the data and the homogeneity of variance), to evaluate if groups with different performance have different baseline corrected blinks/m, showed a significant effect of group ($F(2,74) = 3.17$, $p = .048$) but not for biological sex ($F(1,74) = 0.10$, $p = .75$) suggesting that the baseline corrected blinks/m is not affected by the biological sex of the player. A Post-hoc Holm correction showed that Experts ($M = 2.04$, $SD = 8.68$) experienced a lower variation in blinks when compared to Novices ($M = 9.83$, $SD = 8.70$; $p = .024$). However, no significant differences were observed when comparing Intermediates ($M = 6.24$, $SD = 10.73$) with Experts ($p = .30$) and Novices ($p = .30$) (see Figure 6).

In order to evaluate if the patterns found in one minute of gameplay can be found on the entire recording the same methods were used to analyse the full 13 minutes of Tetris session. Similarly to what occurs in one minute, a repeated measure ANOVA, shows that there is a significant decrease in blinks between the baseline ($M = 19.47$, $SD = 10.80$) and the average blinks/m across the 13 minutes ($M = 8.68$, $SD = 8.97$); both biological sex ($F(1,74) = 5.32$, $p = .03$) and the baseline blinks/m ($F(1,76) = 71.36$, $p < .001$) are significant predictors of the blinks/m during the 13 minutes of Tetris session while Tetris self-assessed experience is not a significant predictor ($F(1,74) = 0.004$, $p = .95$). Even in this case, no interaction effect between biological sex and time ($F(1,76) = 0.33$, $p = .57$) and between Tetris experience and biological sex ($F(1,74) = 1.41$, $p = .24$) was observed. A white-corrected two-way ANOVA, accounting for the non-homogeneity of the variance in the sample, showed a significant effect for group ($F(2,74) = 4.89$, $p = .03$) but not for biological sex ($F(1,74) = 0.16$, $p = .69$). A Games-Howell post-hoc correction, used in case of unequal variance in the sample, showed that Experts ($M = 5.01$, $SD = 5.03$), similar to the previous analysis, had a lower variation in blinks when compared to Novices ($M = 10.39$, $SD = 8.34$; $p = .046$) but not with Intermediates ($M = 9.54$, $SD = 10.30$; $p = .085$). No significant difference was found when comparing Intermediates and Novices ($p = .94$).

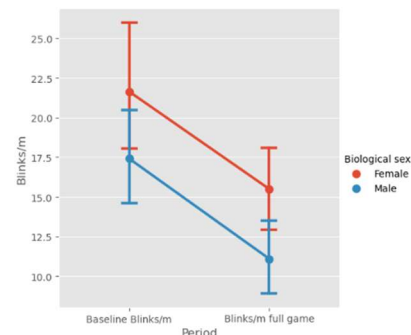


Fig. 5. Lines representing the decrease of blinks for males and females comparing the baseline and the first minute of gameplay

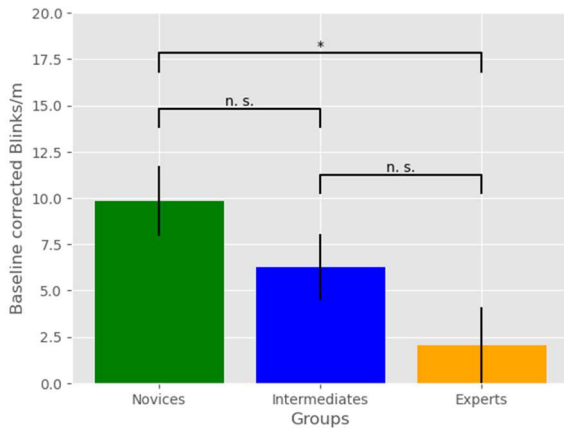


Fig 6. Comparison between the 3 groups showing the baseline corrected blinks/m.

IV. DISCUSSION

This study aimed to test the hypothesis about using blinks to discriminate groups with different Tetris performances during the early stages of a game session. In order to achieve that goal, a method that can track blinks non-invasively using a widely available computer webcam was developed and validated against other available methods. In our sample, we found that the hours spent every week playing video games were not significantly different across the 3 groups. This suggests that there was no transfer between the number of hours spent playing video games and the performance obtained in Tetris in our experiment. We found that males self-assessed themselves as being more experienced in Tetris, however, this factor did not significantly impact the blinks/m variation between the baseline and the gameplay blinks/m for the first minute and for the complete game session. For what concerns the baseline-corrected blinks/m variations, the results show a significant difference between novice and expert Tetris players in the first minute of gameplay. Furthermore, the lack of significant effect of biological sex suggests that baseline corrected blinks/m are a robust measure that is not affected by biological sex. Previous studies with similar aims, but assessing behavioral factors, demonstrated that expert players can be differentiated from less experienced ones just by looking at behavior during level 0 and level 1, such as their positioning of zoids or minimum lines cleared [3,4]. Furthermore, these differences may manifest in the use of the down and left keys in the first 45 seconds of gameplay [2]. Similarly, Guglielmo and colleagues found that more experienced players in Hearthstone have a higher blinks/m than novices during gameplay [2]. Our results show that during the early phases of a game session, group differences also occur on a physiological level. More interesting is that the significant patterns found in the first minute of gameplay seem to be present when looking at the entire 13 minutes of the game session. This suggests that experts, independently from the level reached and played, undergo a lower variation in blinks/m compared to novices. This interpretation is in line with previous studies, which demonstrate that blinks/m decreases with an increase in cognitive load [10] and that proficiency in a task relates to

higher blinks/m [15]. To sum up: proficient Tetris players display a lower decrease in blinks/m as early as one minute and across a full session of complex tasks, in this case playing Tetris.

Despite our results establishing a connection between blinks/m variation and performance in a video game, our study presents several limitations that may be worth mentioning. First, experiment 2 used a sample of participants that is smaller than those used in other studies using Tetris and early detection of expertise [3, 4]. Having more participants may result in wider variations in obtained scores and in different levels of proficiency across the players. Second, we did not directly control for other cognitive processes that may impact blinks/m such as attention [30]. The effect of attention on blinks/m variation may be further investigated in a future study involving eye-tracking measures [31] or EEG measures [32]. Third, the presence of an error in estimating the blinks using our Isolation Forest with filtering may also have played a role in the results obtained. Fourth, in this study, we applied the subtractive method to perform a baseline correction [17, 18]; however, applying another correction method, such as a divisive one, may lead to different results. This issue reflects a gap in the literature where the best method to perform baseline correction for blinks has not been investigated yet.

Additionally, this study makes a methodological contribution by introducing and validating a novel method—Isolation Forest with filtering—to detect blinks even using a low-quality webcam. This method is more noise-resistant compared to any other in existent literature while allowing for the extraction of additional features, such as blink durations and blink intervals, among others (see: Appendix A). Arguably, the Isolation Forest with filtering could be used with higher-quality cameras or an eye-tracker [22], which would further increase accuracy. This provides an additional, easy-to-use psychophysiological measurement option for other researchers, which only requires a webcam. The use of blinking behavior has been associated with dopaminergic activity and receptors [12, 10] but also with the detection of fatigue [28], drowsiness [15], and cognitive load [11] just to mention a few. This suggests that the method introduced here may be effective in investigating these phenomena. However, despite its potential applications, Isolation Forest with filtering still suffers from some limitations, which should be mentioned. First, the EAR measure is sensitive to lighting conditions. Second, its accuracy may be influenced by the sample size of the time series used to train the algorithm, as with any other machine-learning method. Furthermore, according to our results, the method seems effective when using a camera with 30 fps recording but such a method has not been tested yet with lower-quality cameras. Future studies may investigate if the method currently used in this study applies successfully to lower-quality cameras as well. Given this, there is still room to reduce the error between the ground truth blinks and the blinks detected using EAR-based methods. Future studies should aim to further fine-tune the algorithm.

V. CONCLUSION

In this work we provide evidence that players with different levels of performance may exhibit different blinks/m variations during the early phase of the game session. To investigate this we developed a new method that effectively tracks blinks using the EAR extracted from a low-grade camera. Future study may apply our methods to other games to detect early performance and use the method year introduced in other fields of research.

ACKNOWLEDGEMENTS

The research reported in this study is funded by the MasterMinds and Data2Game projects, part of the RegionDeal Midland WestBrabant, and is co-funded by the Ministry of Economic Affairs, Region Hart van Brabant, REWIN, Region West-Brabant, Midpoint Brabant, Municipality of Breda, Netherlands Research Organisation (NWO), and Municipality of Tilburg awarded to MML

APPENDIX A

A list of general features extractable from the EAR using the algorithm introduced in this work and their description.

General Features

Blinks/m: this measure returns to the number of blinks performed on average per minute.

Mean_blinks_interval: It returns the calculation of the average interval between one blink and another.

Std_int: it is the standard deviation between one blink interval and another (in seconds)

Min_int: it is the minimum interval occurring between one blink and another (in seconds)

Max_int: it is the maximum interval occurring between one blink and another (in seconds)

Blinks per recording: the total number of blinks recorded during the video.

Mean blink duration: This refers to the mean duration of the blinks across the entire recording. It is based on the number of frames that are considered as part of blinks (in milliseconds).

Std_duration: it refers to the standard deviation for the blinks' durations across the recording (in milliseconds)

Min_duration: the duration for the shortest blink in the recording (in milliseconds)

Max_duration: the duration for the longest blink in the recording (in milliseconds)

RMSSD: the root mean square of the successive difference between one blink and another [29]. This measure is able to capture changes in the blink intervals that may be caused by specific phenomena (e.g. visual workload) [29].

Besides the features above mentioned, it is possible to extract additional features when binning the recording (for example considering variations across minutes of recording).

Additional Features

Std_blinks: returns the standard deviation for the number of blinks occurring across the bins defined.

Min_blinks: this feature extracts the minimum number of blinks that occur during the recording according to the bins defined.

Max_blinks: this feature extracts the maximum number of blinks that occur during the recording according to the bins defined.

REFERENCES

- [1] J.-C. Hong and M.-C. Liu, "A study on thinking strategy between experts and novices of computer games," *Computers in Human Behavior*, vol. 19, no. 2, pp. 245–258, Mar. 2003, doi: 10.1016/s0747-5632(02)00013-4.
- [2] G. Guglielmo, M. Klincewicz, E. M. J. H. in 't Veld, and P. Spronck, "Predicting Tetris Performance Using Early Keystrokes." In *Foundation of Digital Games 2023*. 2023. doi: 10.1145/3582437.3587184.
- [3] J. K. Lindstedt and W. D. Gray, "Distinguishing experts from novices by the Mind's Hand and Mind's Eye," *Cognitive Psychology*, vol. 109, pp. 1–25, Mar. 2019, doi: 10.1016/j.cogpsych.2018.11.003.
- [4] W. D. Gray and S. Banerjee, "Constructing Expertise: Surmounting Performance Plateaus by Tasks, by Tools, and by Techniques," *Topics in Cognitive Science*, vol. 13, no. 4, pp. 610–665, Oct. 2021, doi: 10.1111/tops.12575.
- [5] Paul P. Maglio, 1995. *The Computational Basis of Interactive Skill*. University of California, San Diego. (1995).
- [6] C. C. Liu, R. C. N. D'Arcy, T. Cheung, X. Song, and R. C. N. D'Arcy, "Spontaneous Blinks Activate the Precuneus: Characterizing Blink-Related Oscillations Using Magnetoencephalography.," *Frontiers in Human Neuroscience*, vol. 11, p. 489, Oct. 2017, doi: 10.3389/fnhum.2017.00489.
- [7] Y. Wang, S. Toor, R. Gautam, and D. B. Henson, "Blink Frequency and Duration during Perimetry and Their Relationship to Test-Retest Threshold Variability," *Investigative Ophthalmology & Visual Science*, vol. 52, no. 7, p. 4546, Jun. 2011, doi: 10.1167/iovs.10-6553.
- [8] M. J. Doughty and T. Naase, "Further Analysis of the Human Spontaneous Eye Blink Rate by a Cluster Analysis-Based Approach to Categorize Individuals With 'Normal' Versus 'Frequent' Eye Blink Activity," *Eye & Contact Lens-science and Clinical Practice*, vol. 32, no. 6, pp. 294–299, Dec. 2006, doi: 10.1097/01.icl.0000224359.32709.4d.
- [9] A. Abusharha, "Changes in blink rate and ocular symptoms during different reading tasks," *Clinical Optometry*, vol. Volume 9, pp. 133–138, Nov. 2017, doi: 10.2147/opto.s142718.
- [10] A. C. S. Costa, A. Peppe, I. D. Mazzù, M. Longarzo, C. Caltagirone, and G. A. Carlesimo, "Dopamine Treatment and Cognitive Functioning in Individuals with Parkinson's Disease: The 'Cognitive Flexibility' Hypothesis Seems to Work," *Behavioural Neurology*, vol. 2014, pp. 1–11, Jan. 2014, doi: 10.1155/2014/260896.
- [11] B. J. Jongkees and B. Hommel, "Spontaneous eye blink rate as predictor of dopamine-related cognitive function—A review," *Neuroscience & Biobehavioral Reviews*, vol. 71, pp. 58–82, Dec. 2016, doi: 10.1016/j.neubiorev.2016.08.020.
- [12] T. Zhang, C. Wang, F. Tan, D. Mou, L. Zheng, and A. Chen, "Different relationships between central dopamine system and sub-processes of inhibition: Spontaneous eye blink rate relates with N2 but not P3 in a Go/Nogo task," *PubMed*, vol. 105, pp. 95–103, Jun. 2016, doi: 10.1016/j.bandc.2016.04.003.
- [13] C. Talens-Estareles, J. J. Esteve-Taboada, V. Sanchis-Jurado, Á. Pons, and S. García-Lázaro, "Blinking kinematics characterization during digital displays use," *Graefes Archive for Clinical and Experimental Ophthalmology*, vol. 260, no. 4, pp. 1183–1193, Nov. 2021, doi: 10.1007/s00417-021-05490-9.
- [14] G. Cardona, C. F. García, C. Serés, M. Vilaseca, and J. Gispets, "Blink Rate, Blink Amplitude, and Tear Film Integrity during Dynamic Visual Display Terminal Tasks," *Current Eye Research*, vol. 36, no. 3, pp. 190–197, Feb. 2011, doi: 10.3109/02713683.2010.544442.
- [15] A. Rahman, M. Sirshar, and A. Khan, *Real time drowsiness detection using eye blink monitoring*. 2015. doi: 10.1109/nsec.2015.7396336.

- [16] G. Guglielmo, P. M. Blom, M. Klineciewicz, E. M. J. H. in 't Veld, and P. Spronck, Blink To Win. *In Foundation of Digital Games 2022*. 2022. doi: 10.1145/3555858.3555864.
- [17] M. Brych, S. Murali, and B. Händel, "How the motor aspect of speaking influences the blink rate," *PLOS ONE*, vol. 16, no. 10, p. e0258322, Jan. 2021, doi: 10.1371/journal.pone.0258322.
- [18] S. Leal and A. Vrij, "Blinking During and After Lying," *Journal of Nonverbal Behavior*, vol. 32, no. 4, pp. 187–194, Jul. 2008, doi: 10.1007/s10919-008-0051-0.
- [19] C. N. Karson, K. F. Berman, E. F. Donnelly, W. B. Mendelson, J. E. Kleinman, and R. J. Wyatt, "Speaking, thinking, and blinking," *Psychiatry Research* 5.3, vol. 5, no. 3, pp. 243–246, Dec. 1981, doi: 10.1016/0165-1781(81)90070-6.
- [20] L. Israel, P. Paukner, L. Schiestel, K. Diepold, and F. SchÖNbrodt, "Data for: Open Library for Affective Videos (OpenLAV)," *PsychArchives*, Aug. 12, 2021. <https://www.psycharchives.org/en/item/009953b8-a55a-4771-b25a-f6235bb159a2>
- [21] D. N. Nishanth and G. M. Rao, "Liveness Detection Based on Human eye Blinking for Photo Attacks," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 1, pp. 4074–4077, Oct. 2019, doi: 10.35940/ijeat.a1314.109119. HERE
- [22] B. A. Fernando, A. Sridhar, S. Talebi, J. Waczak, and D. J. Lary, "Unsupervised Blink Detection Using Eye Aspect Ratio Values," *Preprints.org*, Mar. 2022, doi: 10.20944/preprints202203.0200.v1.
- [23] S. Mehrang, E. Helander, M. Pavel, A. Chieh, and I. Korhonen, *Outlier detection in weight time series of connected scales*. 2015. doi: 10.1109/bibm.2015.7359896.
- [24] P. J. Rousseeuw and C. Croux, "Alternatives to the Median Absolute Deviation," *Journal of the American Statistical Association*, vol. 88, no. 424, p. 1273, Dec. 1993, doi: 10.2307/2291267.
- [25] Barbara G. Tabachnick and Linda S. Fidell. 2013. *Using Multivariate Statistics* (sixth edition). Pearson College. New Jersey, Usa
- [26] P. Mishra, C. K. Pandey, U. Singh, A. Gupta, C. Sahu, and A. Keshri, "Descriptive statistics and normality tests for statistical data," *Annals of Cardiac Anaesthesia*, vol. 22, no. 1, p. 67, Jan. 2019, doi: 10.4103/aca.aca_157_18.
- [27] C. Viwatwongkasem, "A comparison of type I error and power of Bartlett's test, Levene's test and Cochran's test under violation of assumptions," 2004.
- [28] R. Schleicher, N. Galley, S. Briest, and L. Galley, "Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired?," *Ergonomics*, vol. 51, no. 7, pp. 982–1010, Jun. 2008, doi: 10.1080/00140130701817062.
- [29] A. Lenskiy and R. Paprocki, *Blink rate variability during resting and reading sessions*. 2016. doi: 10.1109/norbert.2016.7547466.
- [30] M. Antonio and A. Angrilli. "Spontaneous eye blink rate: An index of dopaminergic component of sustained attention and fatigue." *International Journal of Psychophysiology* 123 (2018): 58-63. Jan 2018.
- [31] M. R. Blair, M. R. Watson, R. C. Walshe & F. Maj. "Extremely selective attention: eye-tracking studies of the dynamic allocation of attention to stimulus features in categorization." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35.5: 1196. (2009)
- [32] P. Mundy, J. Card, & N. Fox. EEG correlates of the development of infant joint attention skills. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 2000, 36.4: 325-338.