Eyes on the prize - Investigating Cognitive Load in Pac-Man gameplay through Eye Tracking

Akaran Sivakumar, AU720104

202208344@post.au.dk

Perception and Action, Cognitive Science, Aarhus University

Jens Chr. Skous Vej 2, 8000 Aarhus, Denmark

04/01/2024

Current attention levels in society have caused a shift in the media consumed and games played. Everything must be shorter, and less taxing cognitively. This study aims to investigate the major cultural trend in the 1980s, the arcade game Ms. Pac-Man, to try and deduce what made the game a global sensation. We investigated cognitive load dynamics through player gaze behaviour. The experiment consists of 4 participants undergoing a total of 20 trials. Players played Ms. Pac-Man for 15 minutes, while the experimenters monitored player action, gaze behaviour and visual input from the game. We uncovered a task-specific learning curve as novice players transitioned from an egocentric perspective to an allocentric perspective. Correlation tests showed a strong spatial association ($r_{s\ x-axis} = 0.73, r_{s\ y-axis} = 0.77$ between gaze location and Pac-Man's positions, supporting the initial egocentric perspective. Regression modelling resulted in a significant relationship between fixation duration and time since the start of the trial, as well as the distance to Pac-Man, though caution is warranted due to a low R-squared value ($R_2 = 0.05$). Despite limitations in the saccade data, the analysis revealed an unexpected negative logarithmic function relationship between saccade amplitude and duration. This relationship is theorized to have occurred due to either a natural optimization from evolution or because of the nature of the task and the task-specific schema created by the participant. Future research can build upon this study by incorporating pupil dilation data, recording a better dataset, incorporating self-report elements, and controlling for the difficulty of the task. The methodology used in this study can be applied in other domains such as driver attentiveness and better interpersonal understanding.

Table of Content

1. Introduction

Human and similar mammal eye movements follow a fixate-saccade-fixate pattern, characterized by brief periods of stability (fixation) alternated with rapid gaze shifts (saccade). The duration of fixation for humans is typically 0.2 - 0.3 seconds, influenced by factors like task complexity. Saccadic eye movements, occurring in the order of 0.01 - 0.1 seconds, depend systematically on movement amplitude (Rizzo et al., 2022). Fixations and saccades are the main forms of eye movement, which are thought to be a combination of bottom-up and top-down processes. In scenarios involving moving objects or observer locomotion, maintaining gaze stability necessitates 'tracking fixations,' involving smooth pursuit eye movements. These various eye movement types coexist in natural behaviour, challenging differentiation based on oculomotor properties or neurophysiology. A significant portion of our understanding of oculomotor control circuits stems from controlled laboratory experiments where participants' heads are stabilized using a chin rest or bite bar, and stimuli and tasks are constrained to elicit the specific types of eye movements the experimenter wishes. However, to understand the utilization of gaze control in natural behaviour, it is important to compare the observed oculomotor behaviour in laboratory settings with gaze recordings in uncontrolled, real-world conditions (Pekkanen & Lappi, 2017).

2. Cognitive load

John Sweller was the first to distinguish cognitive load into three different types. The first kind is called an *intrinsic load*. This is directly connected to the task at hand. This is the perceived difficulty of the task by the participant. As this is related to the ability of the participant, this is not something that can be manipulated by the designer of a system. The second kind, *extraneous load*, refers to the load that is imposed by the format and how information is presented (e.g. the structure of an interface). The last kind, *germane load*, is caused by the effort that has to be made to understand and process the materials. A successful visualization is characterized by keeping the *extraneous* load to a minimum, which can be done by providing smooth system interaction. A low *extraneous* load leads to more cognitive capacities for *intrinsic* and *germane* load that both help to solve the task at hand. Previous research has also shown that novices use significantly more cognitive workload in schema acquisition rather than problem-solving, resulting in an overlap in the cognitive processes required for both and prioritization of one over the other (Sweller, 1988).

Traditionally, brain activity has been used to identify cognitive load. Activity has been measured through techniques such as electroencephalography (EEG) and magnetoencephalography (MEG) whose objective is to capture changes in magnetic fields at the scalp caused by changing electrical currents in brain neurons. The main strength of these techniques is their millisecond-level time precision. However, this measurement is quite intrusive to the user, requires a time-consuming setup and its analysis is often very complex. There are luckily many other techniques that are related to brain activity and cognitive load. Cortical activity causes a small nervous response that is translated into variations in heart rate, blood pressure, electrodermal activity, electrical activity in facial muscles, eye movements and small dilations of the pupil. Several prior studies have investigated the relation of voluntary eye movements like fixations and saccades on cognitive processes, as well as involuntary eye movements like blink rate and pupil dilation. These eye movements are grouped as behavioural (voluntary) and physiological (involuntary).

(Rudmann et al., 2003) found that the gaze direction indicates the AOI (area of interest), that is relevant for the current cognitive activity. This can be interpreted as a repeated interest in an area. One could also use the fixation duration, the period of time that a user looked at a certain AOI. The fixation duration has been related to the level of cognitive processing with a high fixation duration indicating an increased strain on the working memory. This is accompanied by a decrease in fixation rate. Results from (Rizzo et al., 2022) show that the attention level influences gaze behaviour but, only for what concerns fixations-related variables. (Rizzo et al., 2022) also hypothesize that backwards saccades are connected to more complex types of reasoning, typical of attentive processes, which are led by a need to re-analyse or re-sample already visited portions of the scene. They found that subjects, among different tasks, implement task-specific schemes to regulate their gaze dynamics. One could therefore use the fixation duration and fixation rate as indicators of a change in the attention that is needed as the complexity of the task increases.

Another eye-tracking event, saccades, has also shown some influence on cognitive load, according to prior research. The most common visualization for saccades is scan paths. One can measure the velocity and length of saccades and observe the patterns of a scan path. (Chen et al., 2011) found that saccade velocity and length are highly discriminatory parameters, both related to achieving high performance. Based on this, one could assume that saccade length and saccade velocity influence cognitive load. Lastly, the involuntary response of pupils is theorized to be influenced by cognitive load. The pupil can range in diameter from 1.5 mm to more than 8 mm. Psychologists have argued for more than twenty years that changes in pupil dilation accompany effortful cognitive processing. Previous research has shown that users' pupils dilate when the difficulty of the task and their cognitive effort to solve it increase (Chen et al., 2011; Pomplun & Sunkara, n.d.). For example, task-invoked pupillary response has been found to vary linearly with the amount of information processed in short-term, and long-term memory tasks and task difficulty levels (Kramer, 1990). Prior research has also confirmed this dilation effect for tasks such as mental arithmetic, sentence comprehension, and letter matching (Kahneman, 1973). Studies such as (Porta et al., 2012) have even proved a decrease in pupil size as tasks progress, suggesting fatigue. As more working memory and attentional resources are required to achieve high task performance, participants tend to increase pupil size and fixation duration, and at the same time decrease their pupil size deviation, fixation rate, saccade speed and saccade size (Chen et al., 2011).

3. Data collection and analysis

The data used in this project is part of a large-scale dataset, created to conduct reinforcement and imitation learning. The dataset is named Atari-HEAD (**Atari** **H**uman **E**ye-tracking **A**nd **D**emonstration). The precursor to this dataset is the Atari Grand Challenge, a large-scale public dataset of human demonstration collected through online crowdsourcing with players of diverse skill levels. The dataset was created hoping to allow researchers to study the relation between attention and decision. The data was collected using the Arcade Learning Environment (ALE) (Bellemare et al., 2013). This structure allows for capturing of many interesting aspects of the natural visuomotor tasks while allowing better experimental control than real-world tasks. The use of ALE is deterministic given the same game seed. The seed was however randomly generated to introduce stochasticity for gameplay. The Arcade Learning Environment (ALE) was created to evaluate general, domain-independent AI technology. ALE offers the opportunity for models, machine learning and reinforcement learning to be tested on Atari 2600 games, which are seen as challenging and interesting even for human players. ALE allows for the development and benchmarking of domain-independent agents on over 55 different games, showcasing the potential of established AI techniques in the realm of perception and action (Bellemare et al., 2013). Ms. Pac-Man, a classic maze-chase game released in 1981, serves as an influential case study in perception and action. Developed by Namco as a sequel to the original Pac-Man, Ms. Pac-Man introduced dynamic improvements, including faster gameplay and intricate ghost movement patterns. This game was a cultural phenomenon, that contributed significantly to the 1980s arcade gaming scene. We have chosen to investigate Ms. Pac-Man as it offers a dynamic maze environment, perfect for exploring decision-making, attentional shifts, and cognitive load.

3.1. Data Collection

For every game image frame *i,* we recorded its corresponding image frame I_i, human keystroke action a_i, human decision time t_i, gaze positions $g_{i1} \ldots g_{in}$, and the immediate reward r_i returned by the environment. The game screen was 64.6 × 40.0 cm (or 1280 × 840 in pixels), and the distance to the subjects' eyes was 78.7 cm. The human subjects were amateur players who were familiar with the games. The data contains 4 subjects playing 20 different Atari games. This report will only focus on the gameplay of the Atari 2600 game, Ms. Pac-Man. The total game time is 4.87 hours, with 353,428 usable gaze samples. The subjects were only allowed to play for 15 minutes and were required to rest for at least 15 minutes before the next trial. The trials are all 15 minutes as the current literature does not yet propose any AIs that reach human performance by 15 minutes. The gaze data was recorded using an EyeLink 1000 eye tracker at 1000 Hz. The EyeLink 1000 tracker was calibrated using a 16-point calibration procedure at the beginning of each trial, and the same 16 points were used at the end of the trial to estimate the gaze positional error. The average end-of-trial gaze positional error across 471 trials was 0.4 cm (2.1 pixels), less than 1% of the stimulus size. Such high tracking accuracy is necessary when dealing with Atari games since many OOI (objects of interest) are small and hard to track without high-quality equipment.

To optimize the dataset for imitation learning (IL), the Arcade Learning Environment (ALE) default setting, challenging for expert players at 60 Hz, was adjusted. In the new setup, the game pauses at each frame until a keyboard action is taken, allowing subjects to hold a key for continuous play at a more comfortable 20Hz. This change resolves issues such as state-action mismatch, aligning actions with states at each time step and enhancing compatibility with supervised learning algorithms. The semi-frame-by-frame mode also aims to relax gameplay, reduce fatigue, and minimize suboptimal decisions due to inattentive blindness. By recording human decision time and eye movements at every frame, the dataset ensures capturing states requiring sophisticated planning, contributing to effective learning algorithms.

3.2. Data analysis

The data was downloaded from the Arxiv Library (Zhang et al., 2019) on 29th November 2023. The data was then pre-processed which included creating a temporally ordered .csv file and downsampling the data from 1000 Hz to 50 Hz. This was chosen, as Ms. Pac-Man in the ALE can at the highest speed run at 20 Hz, which makes using eye tracking data at 1000 Hz seem redundant. The python package CatEyes (Gütlin, 2021/2021) was then used to classify fixations, saccades, smooth pursuits and PSOs (Post-saccadic oscillations). The method of eye-movement signal segmentation and event classification used is NSLR-HMM (Naïve Segmented Linear Regression - Hidden Markov Models). Unlike traditional workflows, NSLR integrates denoising into segmentation, making it the initial step in the analysis. Classification is then performed on denoised segments. This versatile approach identifies fixations, saccades, smooth pursuits, and post-saccadic oscillations, accommodating experiments with complex gaze behaviour. This allows it to be directly applied to noisy data, yielding robust gaze position and velocity estimates for both high-quality lab data and challenging mobile data on natural gaze behaviour, requiring minimal manual parameter setting as it autonomously estimates signal noise levels and gaze feature parameters from human classification examples (Pekkanen & Lappi, 2017). The received output from the method is displayed in Figure 1.

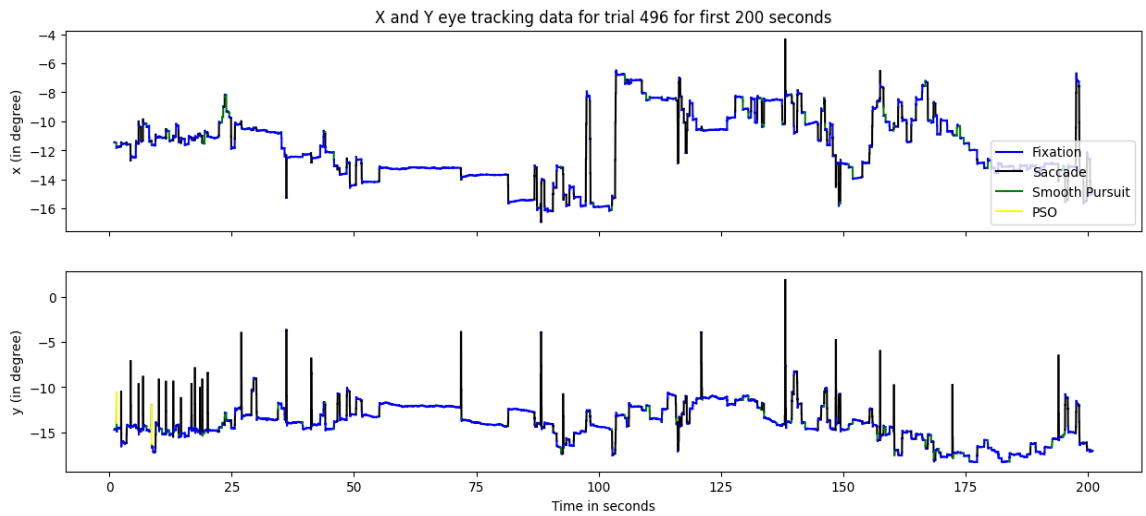

Figure 1. This figure shows an excerpt of the event classification done by the Python package CatEyes.

During the data analysis phase, the obtained game frames from the trials underwent processing using the OpenCV Python package (Bradski, 2000), to precisely localize ghosts and Pac-Man in each frame. This was feasible due to the distinct colour palette of retro game consoles like Atari, where colours for Objects of Interest (OOI) are intentionally different, as the console follows a 128-color palette. The localized positions were then utilized to calculate the distances between ghosts and Pac-Man, as well as between gaze location and Pac-Man, and gaze location and ghosts. This made it possible to create a novel variable named 'gaze_location', that specifies whether the participant is looking at Pac-Man, a ghost, or neither. This 'gaze_location' variable was then investigated temporally. Proportions of time spent looking at each value in 'gaze_location' were calculated for 10-second intervals, and the temporal dynamics of gaze location were visually inspected using the Matplotlib Python package (Hunter, 2007).

Next, the average fixation durations were plotted in 5-second intervals to address the hypothesis related to fixation duration and cognitive load. An ordinary least-squares linear regression model was hypothesized, treating fixation duration as the dependent variable, with time since the start of the trial and distance to Pac-Man as predictor variables. This model, created using the Statsmodels Python package (Seabold & Perktold, 2010), revealed that the distance to Pac-Man and the time since the start of the trial both had a significant impact on fixation duration. Finally, we explored the correlation between gaze location and Pac-Man's location temporally, due to the discoveries made in the prior step. The temporal dynamics were then plotted over one-minute intervals.

4. Results

4.1. Visual inspections

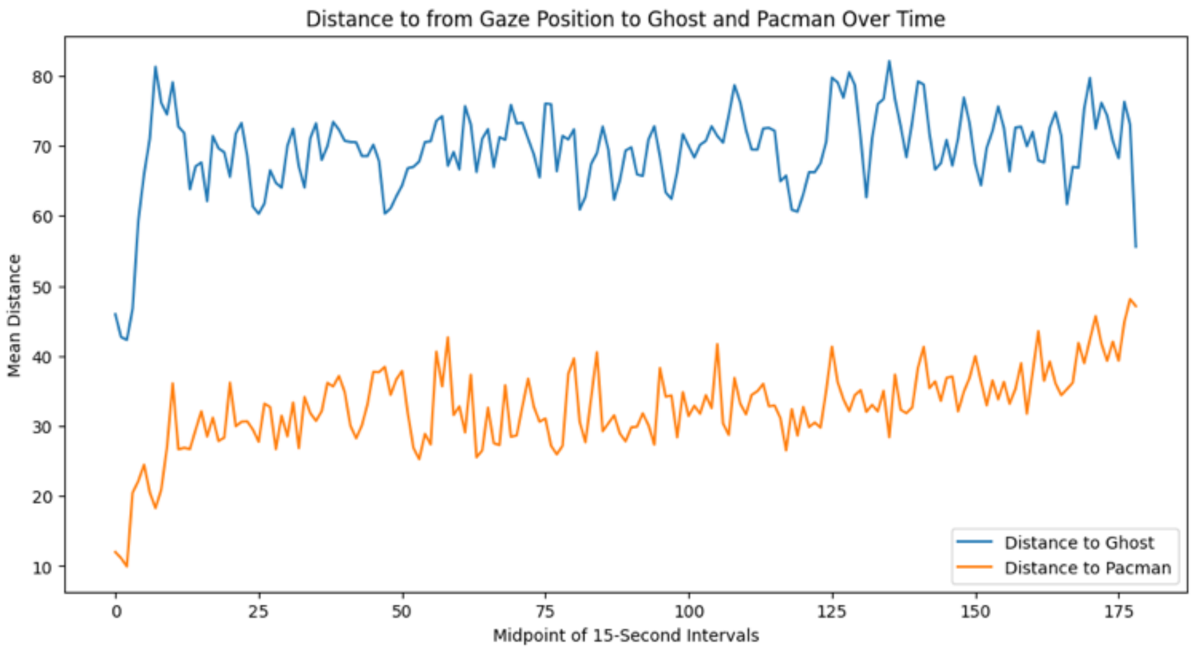

Figure 2. This figure shows the average distance from gaze position to Pac-Man and ghost over 15-second intervals.

Figure 2 shows the average distance per 15 seconds from the gaze position to Pac-Man and Ghosts for all participants. Here there is a clear positive trendline throughout the trial as the distance from the gaze position to Pac-Man is increasing over time. This is hypothesized to be due to the participants developing a task-specific schema for regulating their gaze dynamics. So, this is in actuality a learning curve which has been visualised, as the participant throughout the trial learns that their gaze does not have to be on Pac-Man. This represents a representational shift from an egocentric representation to an allocentric representation. This theory is further corroborated by Figure 3 and **Appendix A:**. They each show the proportion of fixations on Pac-Man, ghosts, or neither and the correlation between Pac-Man's location and gaze location. The proportion of fixations on Pac-Man falls incredibly fast in the first 120-150 seconds, as the stimuli are still novel to the participant, and there is not yet a task-specific schema for playing Pac-Man. Around the 150-200 second mark the proportion of fixation on Pac-Man stabilizes around 0.3 (30 %). We also see that the proportion of fixation on ghosts remains low throughout the whole trial, hovering around 0.1 (10 %). This is further confirmed in Figure 2, as the distance from the gaze position to the ghost position is relatively high throughout the whole trial at 60-80 coordinate points. This is quite high when the gameplay screen only is 160 x 160 coordinate points. Therefore, the attention does not seem to go over to the ghost, but somewhere else. This can also be seen in plot 2 as the plotted lines for 'Pac-Man and 'neither' seem to be negatively correlated, where one might assume it would rather be the lines for 'Pac-Man and 'ghost'. This indicates that there is something else in the visual field which occupies most of the attention of the participant.

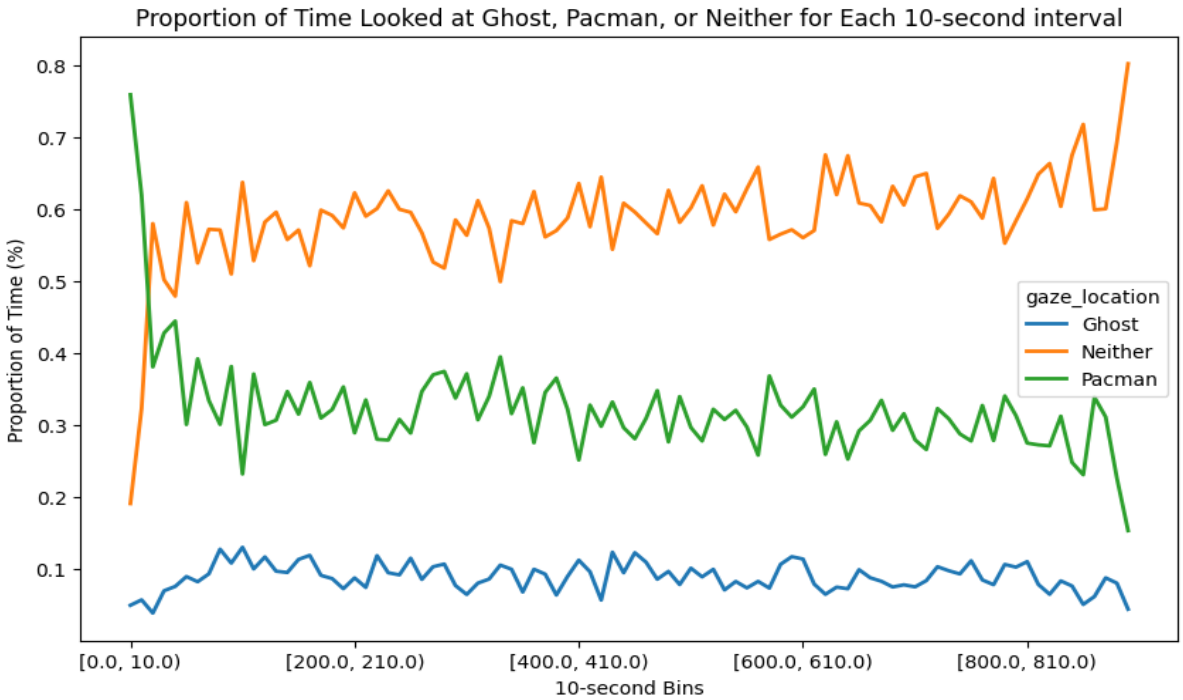

Figure 3. This figure is a plot of the proportion of fixations on Pac-Man, Ghost, or neither in 10-second intervals

4.2. Correlation tests

We then calculated the correlation between Pac-Man's location and the gaze location. Spearman's rank correlation was computed to assess the relationship between the X coordinate of gaze location and the X coordinate of Pac-Man location. There was a positive correlation between the two variables, $r(1134451) = 0.73, p < 0.001$. This finding suggests a moderately strong positive association between the horizontal gaze position and the corresponding horizontal position of Pac-Man. Spearman's rank correlation was computed to assess the relationship between the Y coordinate of gaze location and the Y coordinate of Pac-Man location. There was a positive correlation between the two variables, $r(1134451) = 0.77, p < 0.001$. This finding indicates a positive relationship between the vertical gaze location and the corresponding vertical position of Pac-Man. The strong correlation in both dimensions suggests a strong spatial association between the participant's gaze and Pac-Man's location during gameplay.

4.3. Regression modelling

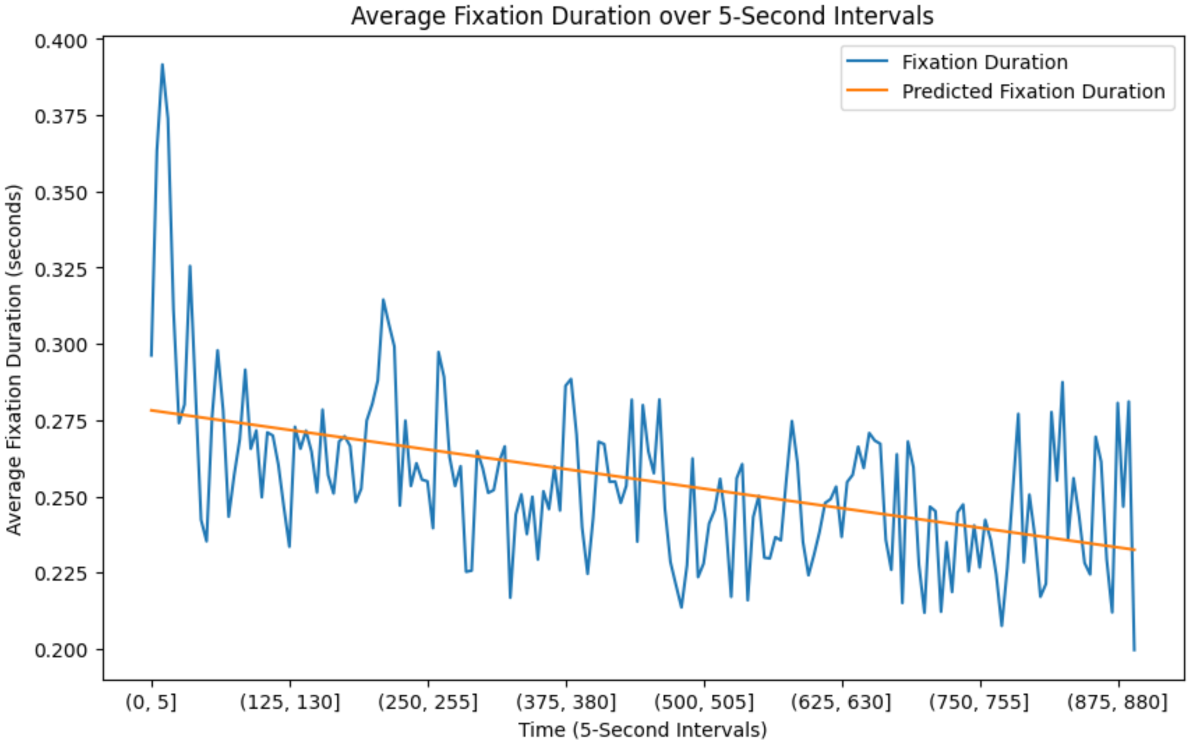

Figure 4. This figure shows the average fixation duration for 5-second intervals throughout the trial. It also shows the predicted values from the Ordinary-least-squares linear regression model

Next, the predictive relationship between fixation duration and the predictors time since the start of the trial (time_minutes) and distance to Pac-Man from gaze location(distance_to_pacman) was examined using a simple OLS linear regression model. Multiple linear regression was used to test if time_minutes (time since the start of the trial in minutes), and distance_to_pacman (distance from gaze location to Pac-Man location) significantly predicted segment_duration (duration of fixation). The fitted regression model was: $segment_{duration} = 0.3170 - 0.0019 \cdot time_{minutes} - 0.0015 \cdot distance_{to_{pacman}}$. The overall regression was statistically significant ($R_2 = 0.056, F(2, 1134450) = 3.384 \cdot 10^4, p < 0.001$). It was found that all predictors were significant predictors of the duration of fixation ($p < 0.001$), however since the R-squared value is so minuscule, we cannot say anything concrete from this regression analysis. This is the best model that was possible from the data points, which shows that the data is not suited for modelling of fixation duration, and cognitive load as a result. One can also see from inspection of Figure 4, that the predicted values are not remotely close to the actual values, although the trendline is in the right direction. Figure 4 shows that the average fixation duration is declining over time. This further strengthens the theory of a learning curve, as the participants are only casual players of the Pac-Man game. Therefore, the first few minutes are harder, as they must come to grips with the controls, rules, strategy, etc. One could say that the trendline seems to suggest declining cognitive load over time, as the participants progress in their learning curve and begin creating a task-specific schema for gaze dynamics. However, since the model has such a low R-squared value, this is not definite, and more data or different data points are needed to verify or decline this notion.

4.4. Saccade Analysis

We also analysed saccades to try and interpret cognitive load from saccades since previous literature has hinted at a relation. Previous literature claims that saccade duration and saccade amplitude are related to the cognitive load. Literature also mentions mean peak velocity as a highly discriminatory factor of cognitive load, but this dataset does not include the data necessary to calculate the mean peak velocity. This means that we unfortunately cannot interpret cognitive load, as the saccade data is spotty and messy, resulting in no correlation or pattern visible in the data. Nevertheless, through the exploration of relationships between variables, we have uncovered a relationship between the saccade duration and the saccade amplitude.

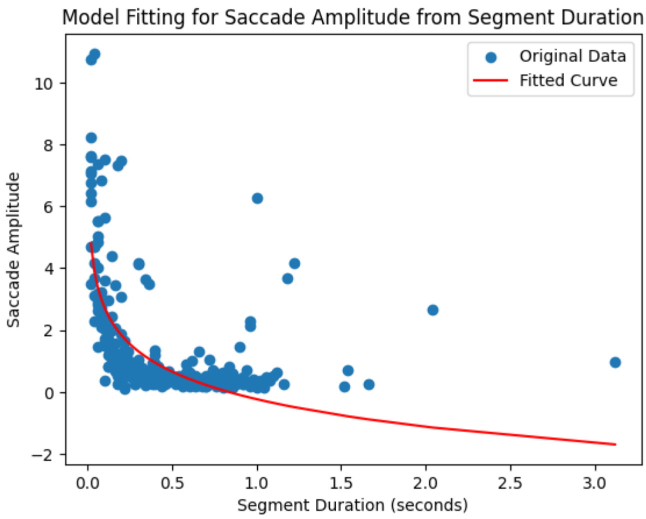

Figure 5. This figure shows the relationship between saccade amplitude and saccade duration. We have also plotted a negative logarithmic model, that fits the relationship quite well

We fit a negative logarithmic function to model the relationship between segment duration and saccade amplitude. The fitted model explained $R_2 = 0.47\ (47\%)$ proportion of the variance in saccade amplitude. The fitted negative logarithmic function is $saccade\ amplitude = a - b \cdot \log(saccade\ duration)$. The estimated parameters are $a = -0.23$ and $b = 1.29$ The parameter a represents the convergent value of saccade amplitude, and b represents the rate at which saccade amplitude decreases with increasing segment duration. The scatter plot displays the original data points for segment duration and saccade amplitude, while the red curve represents the fitted negative logarithmic function. It's important to note that the negative logarithmic function assumes a specific form of relationship between segment duration and saccade amplitude. This relationship is not documented in any known previous literature, but we theorize some reasons for this apparent optimization. The first reason could be that longer saccades are more optimal and therefore more natural for the eye muscles. The second reason could be that longer eye movements use other eye muscles, which are quicker, or lastly, it could be due to the current cognitive task, as the participant might have developed a task-specific schema, entailing quick, long saccades for optimal gameplay in Pac-Man.

5. Discussion and future directions

The findings of our study revealed people's behaviour during Pac-Man gameplay. Visual inspections revealed a learning curve in the trial through the participant's transition from an egocentric perspective to an allocentric perspective. Initially, players fixated more on Pac-Man, but this decreased to around 30% over time, while attention to ghosts remained consistently low. Correlation tests further confirmed a strong spatial association between gaze location and Pac-Man's horizontal and vertical positions, supporting the idea of an initial egocentric perspective. Regression modelling showed a statistically significant relationship between fixation duration and time since the start of the trial, as well as distance to Pac-Man. The model however reported a low R-squared value suggesting caution in drawing definitive conclusions about this relationship. Saccade analysis, despite limitations, revealed an unexpected negative logarithmic function relationship between saccade amplitude and duration, despite a lack of validation from existing literature. These findings collectively highlight participants' evolving gaze dynamics, the establishment of a task-specific schema, and the presence of a learning curve during Pac-Man gameplay, providing some insight into the cognitive processes involved.

This study still possesses some limitations. The lack of pupil dilation data, a key indicator of cognitive load, represents a notable gap in the exploration of cognitive processes during Pac-Man gameplay. Incorporating pupil diameter data could tell us more about phenomena such as fatigue and inhibition. Secondly, the data's messy and incomplete structure is challenging, especially considering the typically high-quality data provided by the equipment used. Additionally, the experimenters' main focus on imitation learning (IL) and reinforcement learning (RL) optimization, rather than cognitive load exploration, resulted in relatively short trials, for cognitive load exploration. The game Ms. Pac-Man is deliberately “too” easy in the first few rounds, so the player can get acclimated to the controls, rules, and gameplay. The difficult and exhausting levels seem to first appear later on with multiple ghosts and worse power-ups, according to online sentiment (*Ms. Pacman Scores*, 2002). Future research with better experimental structures and high-quality data sources may better address these limitations, providing a clearer understanding of cognitive processes in gaming environments.

6. Conclusion

In conclusion, this study has unveiled insights into the gaze behaviour of participants during Pac-Man gameplay, shedding light on their dynamics and the emergence of a learning curve. The visual inspections showed a transition from an egocentric to an allocentric perspective, further corroborated by the increasing distance from the gaze position to Pac-Man throughout the trial. Fixation analysis revealed a decline in the proportion of fixations on Pac-Man, stabilising at around 30%, while fixation on ghosts remained low. Correlation tests provided empirical support for a strong spatial association ($r_s = 0.73, 0.77$) between gaze location and Pac-Man's position, affirming the participants' egocentric perspective.

Despite insights into Pac-Man gameplay, limitations include the absence of pupil dilation data for cognitive load exploration, messy data structure, and the focus on short trials optimized for imitation and reinforcement learning. Future applications may involve monitoring the attentive state of drivers, where understanding gaze dynamics could contribute to enhancing road safety. Additionally, this analysis could prove valuable in contexts related to truth-telling and deception, where insights into gaze patterns may provide cues about cognitive processes.

7. References

Bellemare, M. G., Naddaf, Y., Veness, J., & Bowling, M. (2013). The Arcade Learning Environment: An Evaluation Platform for General Agents. *Journal of Artificial Intelligence Research*, *47*, 253–279. https://doi.org/10.1613/jair.3912

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

Chen, S., Epps, J., Ruiz, N., & Chen, F. (2011). Eye activity as a measure of human mental effort in HCI. *Proceedings of the 16th International Conference on Intelligent User Interfaces*, 315–318. https://doi.org/10.1145/1943403.1943454

Gütlin, D. (2021). *CatEyes* (0.0.5) [Jupyter Notebook]. https://github.com/DiGyt/cateyes (Original work published 2021)

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Kramer, A. (1990). *Physiological metrics of mental workload: A review of recent progress*.

Ms. Pacman scores. (2002, July 8). AtariAge Forums. https://forums.atariage.com/topic/10879-ms-pacman-scores/

Pekkanen, J., & Lappi, O. (2017). A new and general approach to signal denoising and eye movement classification based on segmented linear regression. *Scientific Reports*, *7*(1), 17726. https://doi.org/10.1038/s41598-017-17983-x

Pomplun, M., & Sunkara, S. (n.d.). *Pupil Dilation as an Indicator of Cognitive Workload in Human-Computer Interaction*.

Porta, M., Ricotti, S., & Perez, C. J. (2012). Emotional e-learning through eye tracking. *Proceedings of the 2012 IEEE Global Engineering Education Conference (EDUCON)*, 1–6. https://doi.org/10.1109/EDUCON.2012.6201145

Rizzo, A., Ermini, S., Zanca, D., Bernabini, D., & Rossi, A. (2022). A Machine Learning Approach for Detecting Cognitive Interference Based on Eye-Tracking Data. *Frontiers in Human Neuroscience*, *16*, 806330. https://doi.org/10.3389/fnhum.2022.806330

Rudmann, D. S., McConkie, G. W., & Zheng, X. S. (2003). Eyetracking in Cognitive State Detection for HCI. *Proceedings of the 5th International Conference on Multimodal Interfaces*, 159–163. https://doi.org/10.1145/958432.958464

Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. *9th Python in Science Conference*.

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Zhang, R., Walshe, C., Liu, Z., Guan, L., Muller, K. S., Whritner, J. A., Zhang, L., Hayhoe, M. M., & Ballard, D. H. (2019). *Atari-HEAD: Atari Human Eye-Tracking and Demonstration Dataset* (arXiv:1903.06754). arXiv. http://arxiv.org/abs/1903.06754

Appendix A:

Temporal analysis of the correlation between gaze location and Pac-Man location: