Aim Assist in Video Games

Evaluating the Balance Between Controller and Mouse & Keyboard Players

PROJECT GROUP 1

Andrei Maria (s2463997)
Daniar Baialiev (s2425548)
Jelmer van Dijk (s2741784)

Table of Contents

Abstract

This study examines the impact of aim assist on the performance disparity between controller and mouse & keyboard (MK) users in first-person shooter (FPS) games, focusing on cross-platform play. We used Aimlabs to simulate common FPS tasks and measure performance metrics such as score, accuracy, and on-target rate (OTR) across different input methods. The study involved 24 participants and employed statistical tests including ANOVA and Welch's ANOVA to analyse the data collected under conditions with and without aim assist. Results indicate that MK users consistently outperform those using controllers, regardless of the aim assist status. Aim assist did not significantly alter the performance of controller users relative to MK users, suggesting it does not completely bridge the performance gap in a competitive setting. The findings raise important considerations for game design and the standardisation of input methods in esports to ensure fairness. The study's implications are discussed in the context of enhancing competitive equity in cross-platform gaming.

Keywords

- Aim Assist
- Cross-platform Gaming
- First-Person Shooter (FPS)
- Input Methods
- Esports Fairness
- Performance Analysis

Introduction

In recent years, more and more games have come with the option of cross-platform play, or cross-play, which allows users of different platforms such as PC, PlayStation or Xbox, to play together. However, PC players are generally deemed to have an advantage over console players, especially when it comes to FPS games (Ashley, 2021). To make the game more fair, console players are now often given the option to use aim assist. As the name suggests, aim assist will help the player with their aim, by assisting their crosshair towards the target, although the precise way in which this works and the strength are dependent on the game (Ashley, 2021). This has since caused many PC players to raise concerns about how aim assist is now making the game unfair to them, as they do not get this option.

Relevance to the Esports industry

Esports has long faced challenges in being recognized as a "Sport." Suits (2007) notes that sports should "involve skilful play where the outcome isn't solely determined by chance or luck." Thus, it is important to discuss the fairness and skill of using input from different peripherals in cross-platform games. According to Black (2023), some Apex Legends pro players are encouraged to use controllers instead of a mouse and keyboard since aim assist in the game is strong and can give an advantage in the gameplay. However, as mentioned further in the article, this tendency raises several concerns regarding the fairness of competitive gameplay. In extreme cases, aim assist has been so effective that it can surpass even cheating players, as noted by Hale (2014). Furthermore, Ivan Labilles (2023) highlights the example of the professional player Imperial Hal, who, after switching from mouse and keyboard to controller, achieved immediate success by winning a tournament.

These insights highlight the ongoing debate about the balance between technology and skill in esports. As the industry develops, keeping competition fair is crucial to maintaining Esports' credibility and reputation as a genuine sport.

Research questions

For this study, we will be looking at whether aim assist gives console players an unfair advantage over PC players. To answer this, three important factors need to be considered. First, which aim assistance technologies are currently being used? Second, to what extent does using aim assist impact the performance of players using a controller? Third, does aim assistance give players using a controller an unfair advantage over players using a mouse and keyboard?

Literature research

To examine existing aim assistance techniques a literature review was proposed. For a literature review databases with a large collection of publications were considered. The report uses the following digital libraries: IEEE Xplore, ACM Digital Library and Scopus.

For this study, any papers related to aim assist in FPS are relevant; thus, the primary terms for search are: “Aim assist”, and secondary terms: “Controller”, and “Mouse and Keyboard”. To include only articles related to Esports or competitive gaming the keyword “Gaming” was added to search queries.
A manual review of the resulting articles revealed that only 3 papers are suitable for the conducted research. However, two of the found articles, by Vicencio-Moreira et al. (2014) and Vicencio-Moreira, Mandryk, Gutwin, et al. (2014), truly fitted the objective of our study by researching the impact of aim assistance techniques on gaming performance. The other article, by Bateman et al. (2011), was relevant to the studies by simply providing some background information to us. Due to the small number of articles, additional searches on different databases and search engines such as Google Scholar were made, however, it yielded no new results.

Vicencio-Moreira et al. (2014) discussed the balancing of multiplayer first-person shooter games using different aiming assistance techniques. Their studies showed that certain aim assists, such as bullet magnetism and area cursors, showed effective performance across different gaming scenarios, including target-range setups and more complex 3D game environments. These findings suggest that while some aim assists can enhance novice players' abilities without significantly detracting from the game's challenge for experienced players, their effectiveness diminishes in dynamic gaming situations with multiple game elements (e.g., distractors, moving targets).

In another study, Vicencio-Moreira, Mandryk, Gutwin, et al. (2014) further examined the "effectiveness or lack thereof" of aim-assist techniques. They found that the integration of aim assist could be felt differently based on the complexity of the gaming environment. Techniques like sticky targets and target gravity, which performed well in controlled environments, often failed in realistic gaming settings due to the increased unpredictability and presence of distractors. This highlights a critical consideration for game designers: the application of aim assist must account for the intended game dynamics and player interactions.

These insights highlight a complex balance required in implementing aim assist techniques. While they can potentially equalise competition among differently skilled players, their impact varies significantly depending on game design and player dynamics. This must be considered when integrating such systems into competitive gaming platforms, ensuring that all players, regardless of input method, have equitable gaming experiences.

These findings navigated the course of our research, focusing on diverse in-game situations such as tracking moving targets, shooting multiple targets at once, tracking strafing targets and flicking to a small target. Additionally, it provided the researchers with a better understanding of aim assistance techniques.

Methods

Sample

For this study, we had a total of 24 participants, who were all recruited through convenience sampling[1], and they were all asked to fill in a consent form. All participants had at least some experience with gaming on both controllers and mouse and keyboard, which is essential to ensure the data does not get skewed too far in favour of either option. Furthermore, not all participants played FPS games regularly, which resulted in a mixture of scores. This allows for the effect of aim assist to be studied for players of different skill levels. Unfortunately, some of the collected data got corrupted, resulting in the data of 8 players becoming unusable. Thus, we ended up with 16 players whose data was usable for this study.

[1] Convenience sampling is a qualitative research sampling strategy that involves selecting participants based on their accessibility and availability to the researcher. Rather than being drawn at random from a bigger population, participants in this strategy are picked because they are easily available to the researcher.

Software

The software used for this study is Aimlabs, a program used to train various aiming skills in FPS games. These skills each have tasks the player can do to improve their abilities.
Aimlabs was chosen for this study for several reasons. Firstly, Aimlabs offers a large variety of tasks to choose from, all of which are easily accessible. Furthermore, and most importantly, Aimlabs has the option to turn on aim assist on controllers. Finally, the data gathered by Aimlabs is easily accessible. It is saved in the cloud, which allows for the use of different computers, and it is exportable to both JSON and CSV formats.

In terms of skills, Aimlabs offers the following options, with tasks often having different versions focussing on different skills:

- *Tracking*: tasks under this skill focus on improving the player's ability to track targets.
- *Flicking*: these tasks focus on improving a player's ability to rapidly aim and hit targets.
- *Speed*: tasks here focus on improving the speed of the player's aim and reaction time.
- *Precision*: tasks here focus on accuracy or percentage of targets hit by a player.
- *Perception* and *cognition*: these skills were not considered for our study and as such will not be explained here.

Furthermore, tasks can have different modes, generally labelled as one of the following:

- *Speed*: as the name implies, this mode focuses on speed. For example, so long as the player keeps hitting targets, the next target will stay up for a decreased amount of time.
- *Precision*: these focus on the accuracy of the player. For example, so long as the player keeps hitting targets, the next target will be smaller.
- *Ultimate*: these tasks combine speed and precision. In the case of the above two examples, so long as the player keeps hitting targets, the next target will be both smaller and disappear faster.

For all tasks chosen for this study, the Ultimate version was used, to ensure a focus on both speed and accuracy.

Hardware

Due to both time constraints and hardware issues, the exact hardware used was different across this study. However, all participants made use of a PC or laptop capable of running Aimlabs, a mouse and keyboard, and either an Xbox or PlayStation controller.

Measures and Variables

Players performed four different tasks three times, once using a mouse and keyboard (MK), once using a controller with aim assist turned off (CWO), and once using a controller with aim assist turned on (CW). To balance the potential learning or fatigue effects that might have influenced the results, we randomised the order of input methods used. All participants began tasks using a mouse and keyboard (MK) to set a baseline. Then even-numbered participants conducted tasks using a controller without aim assist (CWO) and then a controller with aim assist (CW). Odd-numbered participants conducted those tasks in reverse order. This counterbalancing procedure prevents our results on the performance of input methods from being confounded by sequential task exposure.

The participants all performed the following four tasks, all of which are standard tasks created by Aimlabs:

- *Microshot (Ultimate)*: a target spawns in the centre and stays up until hit. Once hit, a new target spawns, which only stays up for a limited amount of time. The more consecutive targets are hit without missing a shot, the shorter the targets stay up and the smaller they get. This task belongs to the **flicking** tasks.
- *Gridshot (Ultimate)*: three targets spawn at random points. Once a target is hit, a new one spawns, ensuring three targets are always spawned. This task belongs to the **flicking** tasks.
- *Strafetrack (Ultimate)*: the player has to keep their crosshair on a target that moves back and forth at random. Once the player has tracked a target for a certain time, that target disappears and a new one spawns. The better the player's performance, the faster the targets will move and the smaller they will get. This task belongs to the **tracking** tasks.
- *Switchtrack (Ultimate)*: many targets circle the player, changing direction at random. Once the player has tracked a target with their crosshair for a certain time, that target disappears and a new one spawns. This task belongs to the **tracking** tasks.

Aimlabs gathers various points of data for every task performed; however, these points of data do differ depending on the task. For this study, the following statistics will be focused on:

- Score: This indicates the general performance of the player. In terms of the flicking tasks, the score increases for every target the player hits, and for every missed shot, the score decreases. The tracking tasks work similarly, while the crosshair is aimed at a

target, the score increases, otherwise, it decreases. The precise way the score is calculated varies per task.
- Accuracy: this indicates the percentage of shots fired that hit a target. This means, with an accuracy of 80%, for example, that 80% of all shots fired hit a target, and 20% missed. Accuracy was only provided for the ‘Flicking’ tasks, i.e., "Microshot" and "Gridshot".
- On-Target Rate (OTR): this indicates the percentage of time the player had their crosshair aimed at any target. This means that with an OTR of 75%, the player would have had their crosshair aimed at a target for 45 seconds, and they were off target for 15 seconds. OTR was only provided for the ‘Tracking’ tasks, i.e. "Strafetrack" and "Switchtrack".

Approach for results analysis

1. **Preliminary analysis.** Involves a visual examination of the distributions for each input method—Controller with Aim Assist (CW), Controller without Aim Assist (CWO), and Mouse & Keyboard (MK)—using boxplots.

To ensure that our observations from the boxplot are statistically robust, we will proceed with a series of statistical tests:

2. **Normality Test**: We will first perform a Shapiro-Wilk test to determine if the score data for each input method conforms to a normal distribution. This step is critical, as the normality assumption underpins many statistical tests, including the ANOVA.
3. **Variance Equality Test**: Next, we will employ Levene's test to assess the equality of variances across the three groups. This test checks the assumption of homogeneity of variances, which is another prerequisite for a traditional ANOVA.
4. **Significance Testing**: Depending on the outcomes of the normality and variance tests, we will select an appropriate statistical test to determine the significance of the score differences observed among the input methods. If the data satisfies the assumptions for normality and homogeneity of variances, a one-way ANOVA will be conducted. Should the data fail to meet these assumptions, alternative tests such as Welch’s ANOVA for unequal variances or the Kruskal-Wallis H test for non-normal distributions will be used.
5. **Post-Hoc analysis:** explore further and clarify the findings from an initial statistical test that indicates overall significance. For instance, if Welch's ANOVA suggests there are differences among group means, it does not specify which specific groups differ from each other. Post-hoc tests (Games-Howell in this study) are then used to investigate these differences in detail.

These tests will validate whether the apparent advantage for MK users observed in the box plot reflects a statistically significant difference in performance scores when compared to controller users, thereby contributing to the discussion on the fairness and balance of aim assist in cross-platform play.
All of these tests will be performed in Python.

Results

Score Distribution by Input Method

1. Preliminary analysis

Figures 1 and 2 revealed that participants using Mouse & Keyboard (MK) tend to achieve higher scores compared to those using controllers, both with and without aim assist (CW and CWO). This observation suggests a significant performance advantage for MK users.

2. Assessment of Data Normality

Before assessing the statistical significance of score differences among input methods, we conducted Shapiro-Wilk tests to evaluate the normality of score distributions for each task. The test results (Table 1) indicated that the scores for the "Gridshot", "Microshot", "Strafetrack", and "Switchtrack" tasks, across all input methods (CW, CWO, MK), are consistent with a normal distribution (all p-values > 0.05).

With the normality assumption confirmed, we are positioned to employ ANOVA for a reliable comparison of mean scores between the input methods. Additionally, we will conduct Levene's test to verify the homogeneity of variances, which is another prerequisite for the ANOVA test.

3. Homogeneity of Variance

Following the assessment of data normality, we conducted Levene's test (Table 2) to verify the assumption of homogeneity of variances—a prerequisite for ANOVA. Levene's test results indicated equal variances for scores in the "Gridshot" task ($p = 0.065954$), thus meeting the assumption for conducting a traditional ANOVA.

However, for "Microshot", "Strafetrack", and "Switchtrack" tasks, the test showed unequal variances among the input methods ($p < 0.05$), which violates the assumption needed for a standard ANOVA.

Revised Analysis Plan

In light of Levene's test results, we will adapt our statistical approach. For "Gridshot", where the equal variances assumption holds, we will proceed with ANOVA for our analysis.

For "Microshot", "Strafetrack", and "Switchtrack" tasks, where variances are unequal, we will employ Welch's ANOVA—a variation of ANOVA that is robust to violations of equal variance—ensuring the validity and reliability of our findings. This approach allows us to continue with a nuanced analysis of the data, respecting the underlying statistical assumptions and maintaining the integrity of the conclusions drawn.

4a. ANOVA for "Gridshot" Task

The traditional ANOVA was applied to the "Gridshot" task (Table 3), where the assumption of equal variances was met as indicated by Levene's test. This result demonstrates that there are statistically significant differences in performance scores among the input methods for the "Gridshot" task.

4b. Welch's ANOVA for Remaining Tasks

Given the unequal variances in the "Microshot", "Strafetrack", and "Switchtrack" tasks, Welch's ANOVA was performed to determine the significance of the score differences across input methods for these tasks (Table 4). The Welch's test is more appropriate here due to its robustness in handling heteroscedasticity, which refers to the condition of unequal variances.

It confirmed significant differences among the input methods for the "Microshot", "Strafetrack", and "Switchtrack" tasks. These findings align with the preliminary visual observations from the boxplots and reinforce the notion of a performance advantage for users of Mouse & Keyboard (MK) input.

5. Post-Hoc Analysis

Following the significant findings from Welch's ANOVA, which indicated differences in scores between input methods across different tasks, it was necessary to conduct a post-hoc analysis. The Games-Howell post-hoc test was chosen because it is suitable for data with unequal variances and different sample sizes, and conditions present in our dataset. This test provides a robust means to compare all pairs of groups independently of each other.

The Games-Howell post-hoc test has revealed noteworthy differences in performance scores between various input methods. Table 5 shows a brief overview of the main results.

Alternative visualisations

1. Histograms and Density Plots

Figure 3 shows a histogram of the scores.

Purpose

To visualise the distribution of performance metrics such as accuracy across different input methods.

Benefit

These plots will help identify the shape of the distribution and compare the spread and central tendencies across groups.

2. Violin Plots

Figure 4 shows a violin plot of the scores.

Purpose

To combine the distribution view of a density plot with the summary statistics provided by boxplots.

Benefit

This will allow us to see both the median and quartile ranges along with the density distribution of the data, providing a more detailed view of the data spread.

3. Scatter Plots

Figure 5 shows a scatter plot of the scores.

Purpose

To explore potential relationships between continuous variables such as time and accuracy, and to observe how these relationships differ among the input methods. -

Benefit

Scatter plots will be useful to identify correlations and trends over time or across different sessions.

Accuracy Distribution by Input Method

1. Preliminary analysis

Figure 6 shows a boxplot of the accuracy for the three categories. Unlike the scores, there is no significant increase in performance when comparing the accuracy of CWO or CW with MK. The maximum recorded accuracy for all categories does not show a significant difference, with CW scoring marginally higher than the other options. Furthermore, there is a difference of less than 10% when comparing the lowest median, CWO, to the highest median, MK. However, there is one noticeable difference between MK and the other options, as MK appears to be more consistent. The accuracies for MK range from ~80% to ~97%, a difference of ~17%. CWO on the other hand, ranges from ~64% to ~97%, with an outlier at ~57%. Similarly, CW ranges from ~62% to ~98%. This means that both controller options show a difference 15-20% larger than MK, whilst not showing a significant difference between CW and CWO.

Figure 7 displays the same data as Figure 6, further separated by task.
For both tasks, the median increases when going from CWO to CW to MK, as it did in Figure 6. The difference between accuracies follows a similar pattern for "Gridshot", with MK being the most consistent, and CWO the least consistent. Furthermore, the maximum accuracies for "Gridshot" only display a marginal difference.
On the other hand, for "Microshot", the maximums do show a more significant difference, with CWO having a maximum of less than 90%, whilst CW and MK appear to be roughly 95% to 97%. Contrary to "Gridshot", CWO is more consistent than CW, although it does have two major outliers.

2. Assessment of Data Normality

Table 6 shows that "Gridshot" follows a normal distribution for all categories, however, "Microshot" does not. As such, whilst an ANOVA test might still give us accurate results for "Microshot" despite non-normality, we shall focus on the outcome of "Gridshot".

3. Homogeneity of Variance

Table 7 shows us that the variances for "Gridshot" are equal, whereas for "Microshot", they are not. Therefore, we will perform an ANOVA test, focussing on "Gridshot", and Welch’s ANOVA test, focussing on "Microshot", as this test does not require equal variances.

4a. ANOVA for "Gridshot" Task

Table 8 shows that according to the ANOVA test, "Gridshot" can be deemed significant.

4b. Welch's ANOVA for "Microshot" task

As with the scores, Welch’s ANOVA test was done for "Microshot", and Table 9 shows that "Microshot" can be deemed significant.

5. Post-Hoc Analysis

Finally, Table 10 shows the post hoc comparison done using the Games-Howell test, which tells us the only significant comparisons are MK compared to either controller option for "Microshot", with no significance found for "Gridshot". Furthermore, it tells us that the mean accuracy for MK is consistently higher than CW and CWO, and the mean accuracy for CW is higher than CWO.

On-target rate (OTR) distribution by input method

On-Target Rate (OTR) refers to the percentage of time that a player's crosshair is directly on a target during a tracking task. This metric is crucial in assessing a player's ability to maintain aim at a target, particularly in dynamic scenarios where the target is moving.

Here’s a basic idea of how “On Target Rate” is calculated:

OTR = (Avg Time on Target/Total Task Time)×100

1. Preliminary analysis

Figure 8 presents a graphical representation of the distribution of a dataset across different categories: CWO, CW, and MK for all the tasks combined. As follows from distribution, CWO and CW categories have similar median values, around 18-20 OTR points. On the contrary, the MK input category has a median of around 30 OTR points, meaning that users achieve higher OTR in the presented tasks while using mouse and keyboard, and not controller input (with or without aim assist). However, the results achieved from controller use lie within a smaller accuracy range. As follows from the boxplot, the outliers of CWO and CW lie closer to the middle of the data than the outliers for the MK category.

Figure 9 analyses OTR across the categories but splits collected data according to the performed tasks. For both “Strafetrack” (left subgraph) and “Switchtrack”(right subgraph) tasks, the median in CWO and CW categories is similar (~25, ~15-17 points respectively per task), suggesting that users perform equally well while using controllers both with and without aim assist. MK category suggests more OTR in both presented tasks with the largest median in both subgraphs. However, while MK has the largest OTR indicators, CWO and CW appear to provide the most consistent results in the “Strafetrack” and “Switchtrack” tasks respectively.

2. Assessment of data normality

Table 11 analyses whether the results are normally distributed to further determine what statistical tests to apply. As follows from the table, each task-category pair is normally distributed meaning that we could perform ANOVA statistical tests to interpret results.

3. Homogeneity of Variance

Table 12 displays Levene’s test results on OTR for the “Strafetrack” and “Switchtrack” tasks. Test results indicate that both tasks satisfy the constraint of equal variances, confirming the strategy to use the ANOVA statistical test.

4. ANOVA test

Table 13 suggests that the results of the ANOVA test on both tasks can be considered statistically significant.

5. Post-Hoc Analysis

The final statistical analysis, the Games-Howell test (Table 14), presents the post-hoc comparison between the categories for the task assigned. The results showed that significance is confirmed only while comparing mouse-keyboard input to controller input (both with and without aim assist) towards the “Strafetrack” task with p-value= 0.039 (0.027). Additionally, the analysis indicates that the mean OTR is generally higher for MK input. Unfortunately, comparing CW and CWO input in terms of OTR, the test didn’t show any particular relation.

Discussion

Score

Our analysis indicates a significant disparity in performance scores between Mouse & Keyboard (MK) users and controller users, both with and without aim assist (CW and CWO). This is corroborated by the findings from both the ANOVA and Welch's ANOVA tests. The MK group consistently outperformed the controller groups across all tasks, suggesting a notable advantage when using a mouse and keyboard. This advantage could be attributed to the precision and speed that mice afford, which are particularly beneficial in tasks requiring rapid and accurate target engagement. Interestingly, the addition of aim assist (CW) did not significantly alter the scores for controller users compared to when it was turned off (CWO). This suggests that while aim assist might slightly enhance the aiming capabilities of controller users, it does not equalise the playing field against MK users. This could be due to the inherent limitations of aim assist, which, although helpful, cannot fully replicate the efficiency and responsiveness of a mouse.

Key pairwise comparisons

1. Controller vs. Mouse and Keyboard (MK):

- Controllers were at a consistent disadvantage compared to the MK setup across all tasks. The MK group had much better scores, with very low p-values and large negative effect sizes. For instance:
  - In "Microshot", the comparison between controller with aim assist (CW) and MK resulted in a -24,590.00 mean score difference, a p-value of 1.599e-06, and a hedges' g of -2.487.
  - In "Gridshot", CW versus MK had a -23,388.00 mean score difference, a p-value of 1.707e-06, and a hedges' g of -2.391.
  - For "Strafetrack", the CW versus MK comparison yielded a -21,536.19 mean score difference, a p-value of 4.746e-06, and a hedges' g of -2.207.
  - Other tasks displayed a similar trend.

2. Controller With vs. Without Aim Assist:

- The comparison between controllers with and without aim assist (CW and CWO) did not show significant differences across the tasks. For example:
  - In "Switchtrack", CW versus CWO had a 2,141.88 mean score difference, a p-value of 0.626, and a hedges' g of 0.321.
  - In "Microshot", CW versus CWO showed a 434.25 mean score difference, a p-value of 0.973, and a hedges' g of 0.077.
  - In "Gridshot", CW versus CWO had a 235.81 mean score difference, a p-value of 0.994, and a hedges' g of 0.037.

Accuracy

In regards to the accuracy of the players, several things can be said. For starters, for both tasks, the accuracy of MK is the highest on average, and CWO is the lowest (Table 10). Furthermore, for "Microshot" task, this difference is three times higher than it is for "Gridshot", with a difference of 12,50% and 4,06%, respectively. This is to be expected, however, given that the distance between targets is generally much smaller for "Gridshot" than it is for "Microshot". As such, flicking from one target to another becomes somewhat easier for "Gridshot", which is likely what resulted in higher accuracy for controller users. In addition, this is likely what resulted in none of the comparisons for "Gridshot" being considered significant.
When looking at the mean accuracies for "Microshot", on average there will be a ~3% increase when going from CWO to CW, however, there is a 12,5% increase when going from CWO to MK. Given this still reasonably significant difference, using aim assist on a controller does not increase the players' aim enough to affect players using a mouse and keyboard.
On the contrary, the same cannot be said for "Gridshot". Here, the increase in accuracy when going from CWO to CW is ~2,5%, not far off the ~3% increase for "Microshot". However, when going from CWO to MK, the increase in accuracy is only ~4%, significantly lower than the 12,5% for "Microshot". This also means that whilst on average, the accuracy is still lower for CW than for MK, the difference is small enough that CW could potentially outperform MK. This is shown in Figure 6, as the highest accuracy for CW is higher than the highest for MK, albeit by only ~1-2%
A final point to make for accuracy is the fact that a higher accuracy does not inherently mean a higher score and thus higher performance. For example, someone getting an accuracy of 100% but only hitting 10 targets, will get a lower score than someone getting an accuracy of 80%, whilst hitting 80 targets. For this reason, the results of this part will not affect the conclusion as much as the score will.

OTR

Results show that Mouse & Keyboard (MK) users exhibit a consistently higher OTR compared to both controller settings - with and without aim assist (CW and CWO). Specifically, MK users achieved a median OTR of around 30 points, while controller users hovered around 18-20 points, irrespective of the aim assist status. This difference highlights the advantages of MK setups in terms of precision and responsiveness.

MK users not only had higher median OTR values but also displayed greater consistency across tasks. This suggests that the precision offered by MK input methods enhances performance in tracking tasks, where maintaining consistent aim is critical.

The analysis showed no significant differences in OTR between controllers with and without aim assist. This indicates that while aim assist may slightly enhance the controller experience, it does not significantly impact the tracking performance of a controller player. This is critical as it suggests that aim assist does not fully compensate for the agility and accuracy afforded by MK.

The post-hoc analysis indicated a significant performance gap between MK and both controller setups in the "Strafetrack" task, with MK outperforming controllers. This difference was statistically significant, proving that MK users have an advantage in tracking efficiency. There were no significant differences in OTR between CW and CWO. This suggests that the impact of aim assist on tracking efficiency is minimal, and does not significantly improve the ability of controller users to maintain aim on targets compared to using no assist at all.
To sum up, MK players still have an edge over controller players in terms of accuracy while tracking moving targets and aim assist does not compensate for this difference.

Considerations

Limitations

The study has several important limitations in understanding aim assist implications that need to be acknowledged.

- As insights from the literature suggested, it's important to research diverse gaming situations. Using AimLabs as the sole means of testing aiming mechanics is sufficient, especially with a diverse set of tasks but might fail to capture the whole variety of game mechanics that are to be found in most popular FPS games such as 'Apex Legends' or 'Call of Duty'
- Additionally, the strength of aim assist may vary across different cross-platform games and testing it within AimLabs may not provide a comprehensive understanding.
- The convenience sampling method and the relatively few participants, less than 30, are more likely not to yield findings generalizable to the larger population of gamers.
- There also were variations in gaming setups, including hardware and environmental conditions, which would have introduced unaccounted-for variability in performance.

Future research

To address these limitations, future research should consider the following recommendations:

- Expand testing to include a broader list of games that better represent the variety in popular FPS titles to understand how aim assist functions across different gaming mechanics and settings.
- Employ a more robust sampling method to increase participant diversity and number, aiming for a sample size that can provide results more representative of the general population of gamers.
- Standardise hardware and environmental conditions during testing to minimise variability and better isolate the effects of aim assist.

These steps will help ensure that future studies provide more definitive insights and practical guidelines regarding the role of aim assist in competitive gaming environments.

Conclusion

From these results, several points can be made. For starters, an increase in score, accuracy, and OTR can be found for most tasks when switching from a controller without aim assist, to a controller with aim assist, except for the OTR of "Switchtrack". However, neither this decrease for "Switchtrack", nor the increase for everything else, was deemed significant in any of the tests performed. Furthermore, the increase in these factors when switching from either controller option to mouse and keyboard was considered significant for everything except the accuracy of "Gridshot", and the OTR of "Switchtrack". Despite them not being deemed significant, the accuracy for "Gridshot" and the OTR of "Switchtrack" did still increase in all comparisons.

In conclusion, our study found no significant impact of aim assist in cross-platform gaming. The data showed an increased performance overall in mouse and keyboard players against controller players, with or without aim assist. Though aim assist is doing a decent job of helping to bridge the gap to some degree, it doesn't level performance across the board. This impacts and points to a potential area for developers within the cross-platform game design. Existing literature on this topic claims that the impact of aim assist is vastly dependent on the dynamics of the situation, thus future studies may want to test different video games and possibly include longitudinal data for determining how input methods affect players' adaption.

Acknowledgements

We would like to thank Guido Bruinsma and Olga Karageorgiou for their guidance during this project.

Conflicts of Interest

The authors confirm that there are no conflicts of interest to declare.

References

1) Ashley, J. (2021, November 15). *Aim Assist in FPS – The case for and against it in esports*. Esports.net. https://www.esports.net/news/aim-assist-in-fps-the-case-for-and-against/
2) Black, S. (2023, 29 December). ImperialHal recommends players stick to controller for "OP" Aim Assist. *Dexerto*. https://www.dexerto.com/apex-legends/imperialhal-recommends-players-stick-to-controller-for-op-aim-assist-2450251
3) Suits, B. (2007). The elements of sport. In W. J. Morgan (Ed.), Ethics in sport (pp. 9–19). *Champaign, IL: Human Kinetics.*
4) Hale J. (2024, March 13) Warzone aim assist is so strong it counters speed hackers & players can't believe it. *Dexerto*. https://www.dexerto.com/call-of-duty/warzone-aim-assist-is-so-strong-it-counters-speed-hackers-players-cant-believe-it-2541614/
5) Vicencio-Moreira, R., Mandryk, R. L., & Gutwin, C. (2014). Balancing multiplayer first-person shooter games using aiming assistance. *2014 IEEE Games Media Entertainment*. https://doi.org/10.1109/gem.2014.7048086
6) Vicencio-Moreira, R., Mandryk, R. L., Gutwin, C., & Bateman, S. (2014). The effectiveness (or lack thereof) of aim-assist techniques in first-person shooter games. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14), 937-946. Association for Computing Machinery. https://doi.org/10.1145/2556288.2557308
7) Labilles J. (2023, Feb 6) ImperialHal cements dual-input legend status with ALGS MVP win. *Dot Esports.*

https://dotesports.com/apex-legends/news/imperialhal-cements-dual-input-legend-status-with-algs-mvp-win

Figures

Figure 1: Boxplot of total score distribution across categories

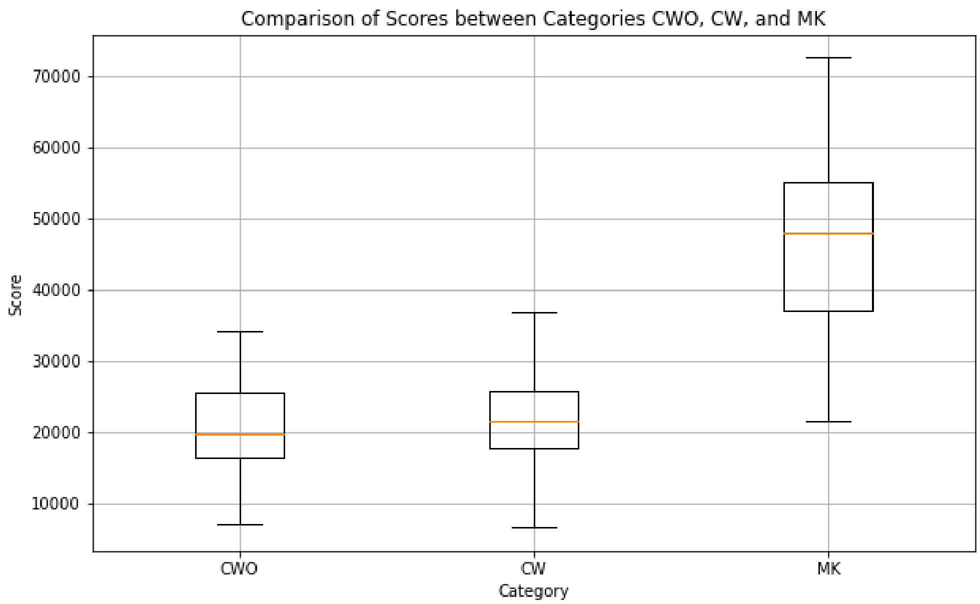


Figure 2: Boxplot of score distribution across categories per task

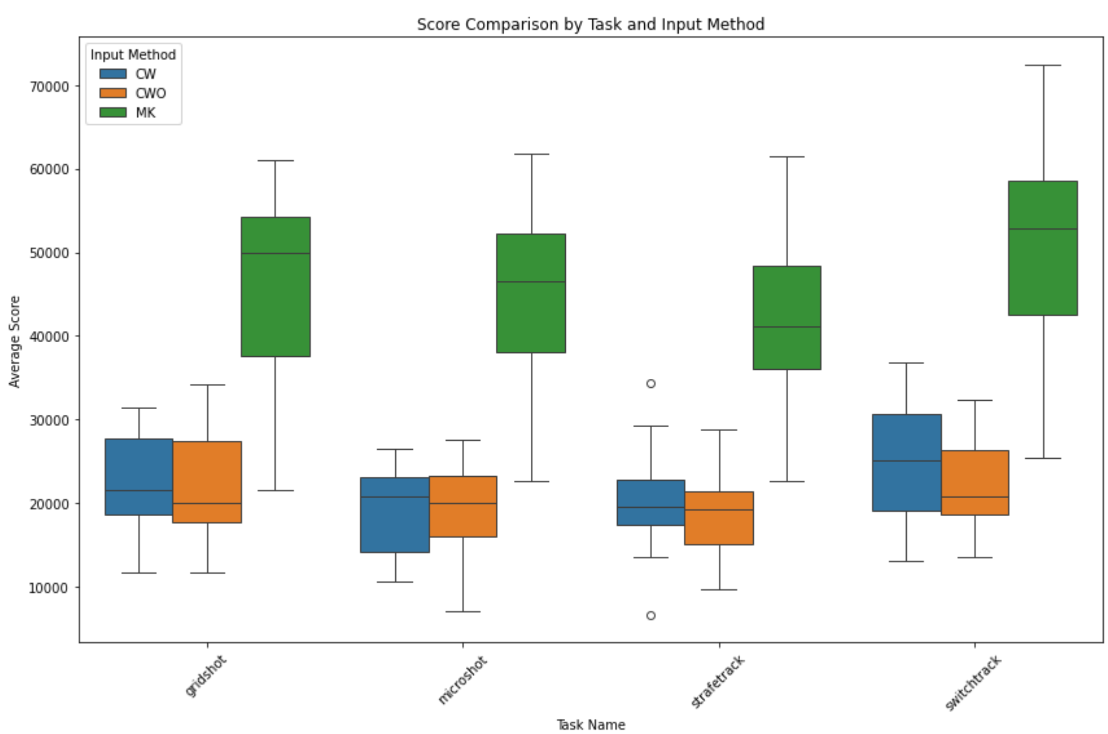


Figure 3: Histogram of the scores

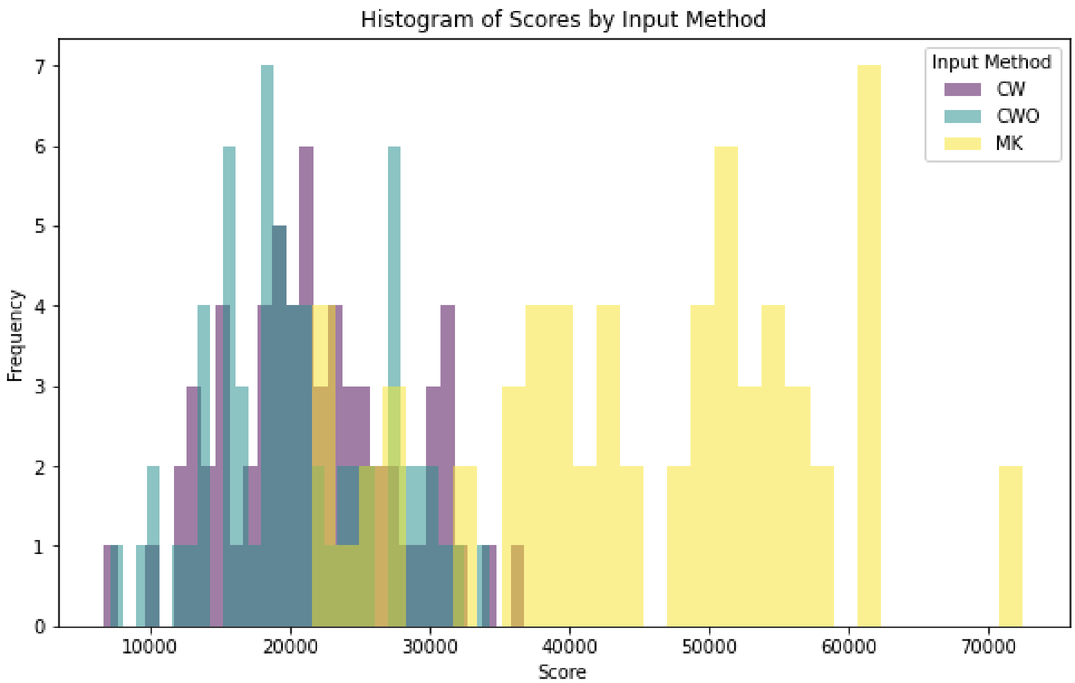


Figure 4: Violin plot of the scores

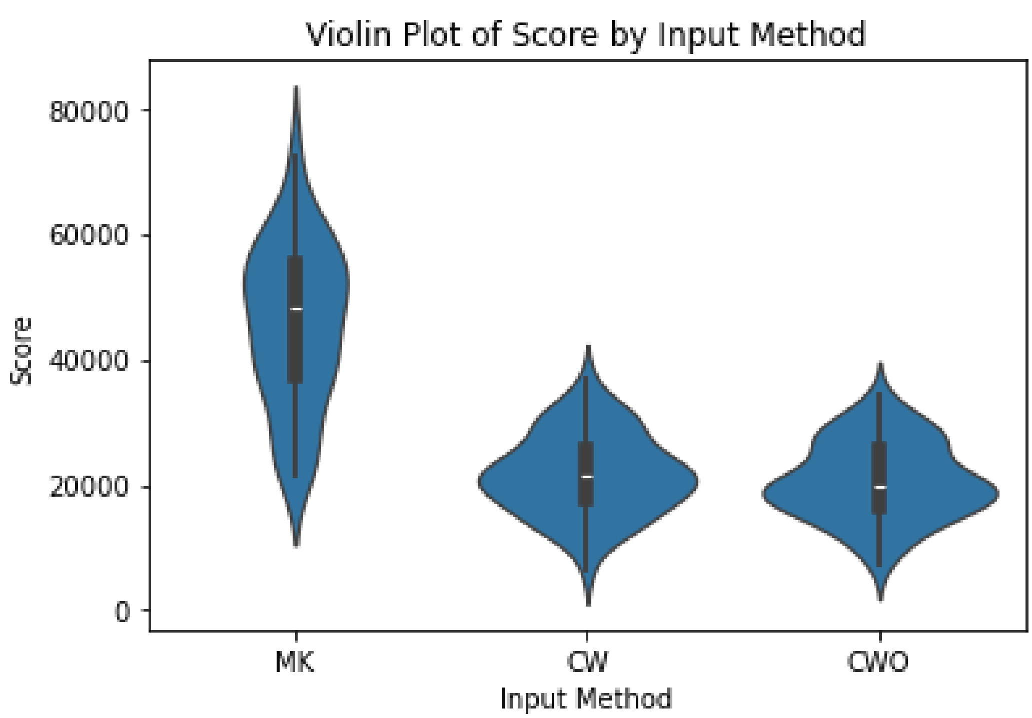


Figure 5: Scatter plot of the scores

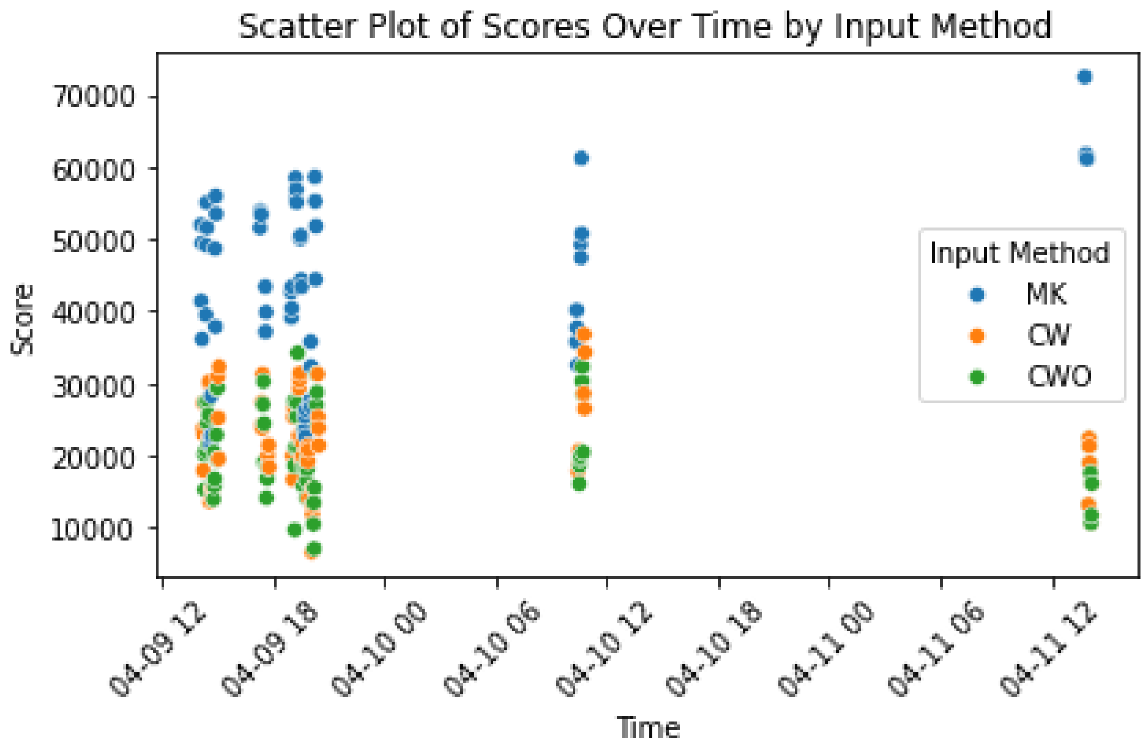


Figure 6: Boxplot of accuracy across both ‘Flicking’ tasks, separated by category

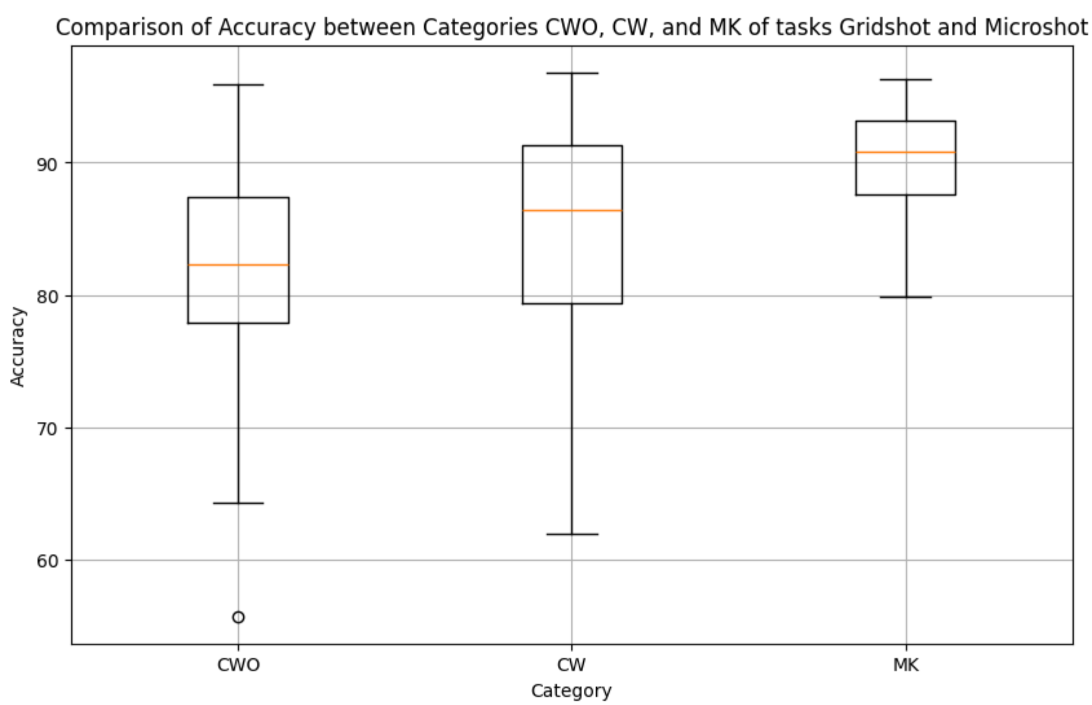


Figure 7: Boxplot of accuracy for both ‘Flicking’ tasks, separated by category

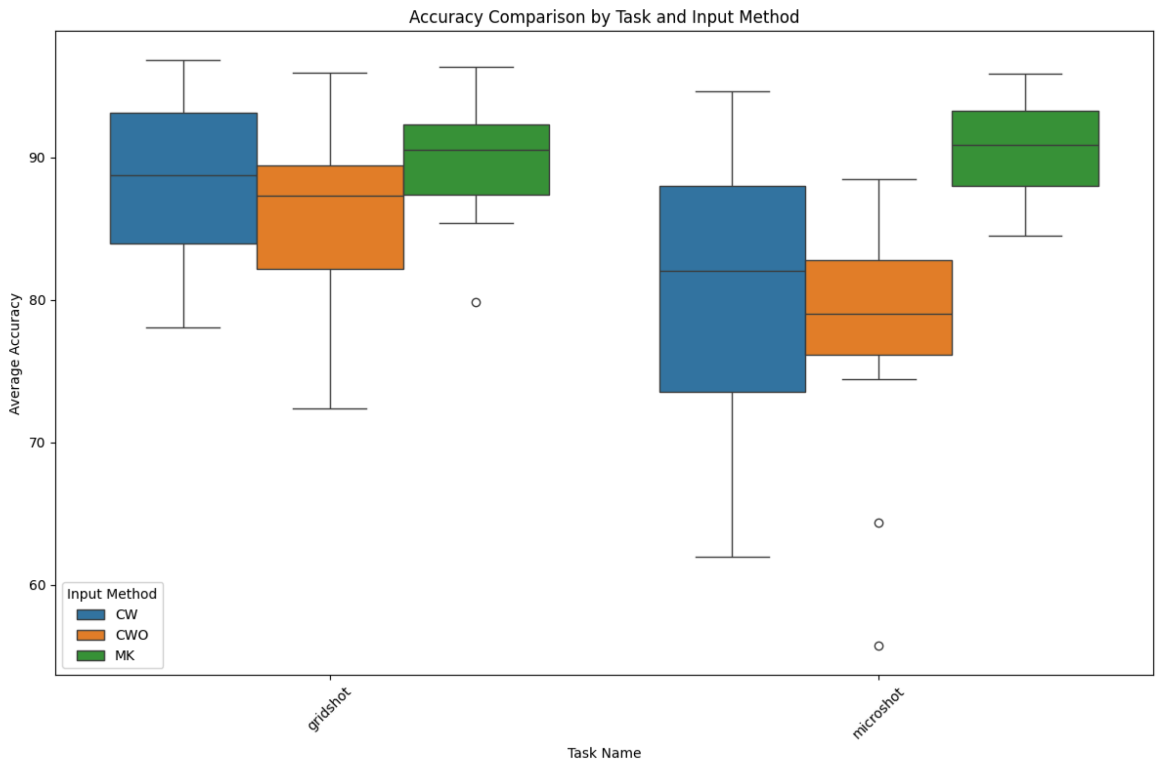


Figure 8: Boxplot of OTR across both ‘Tracking’ tasks, separated by category

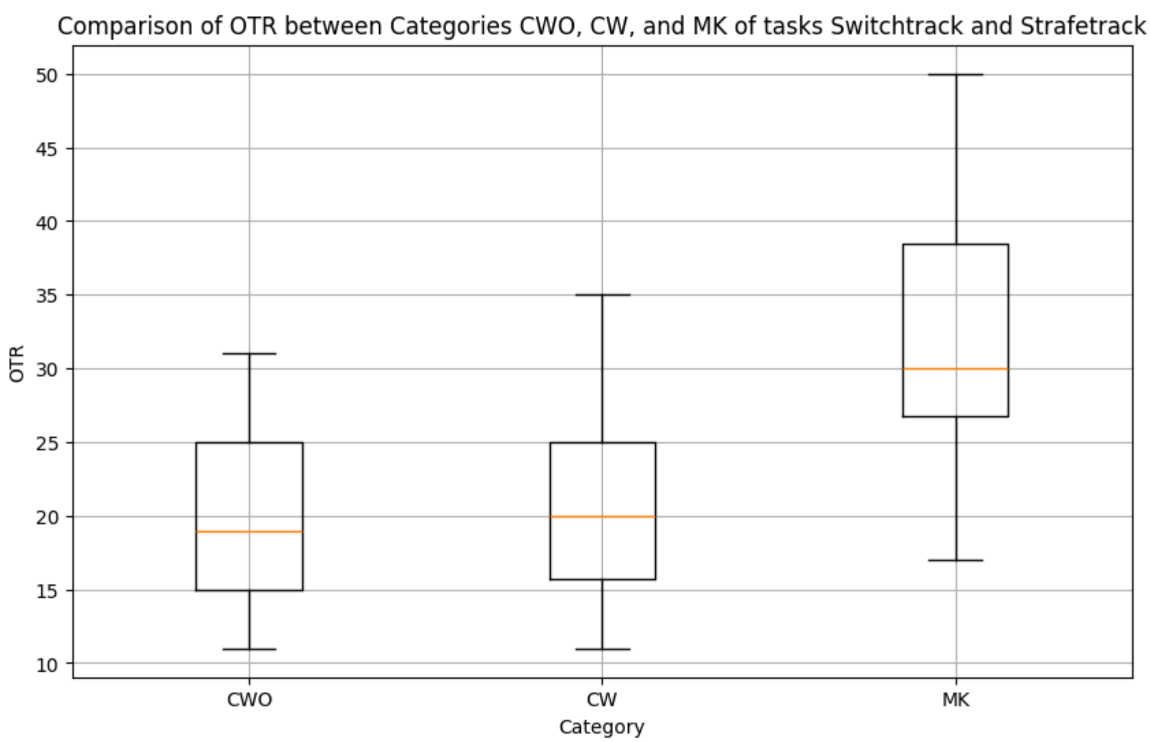
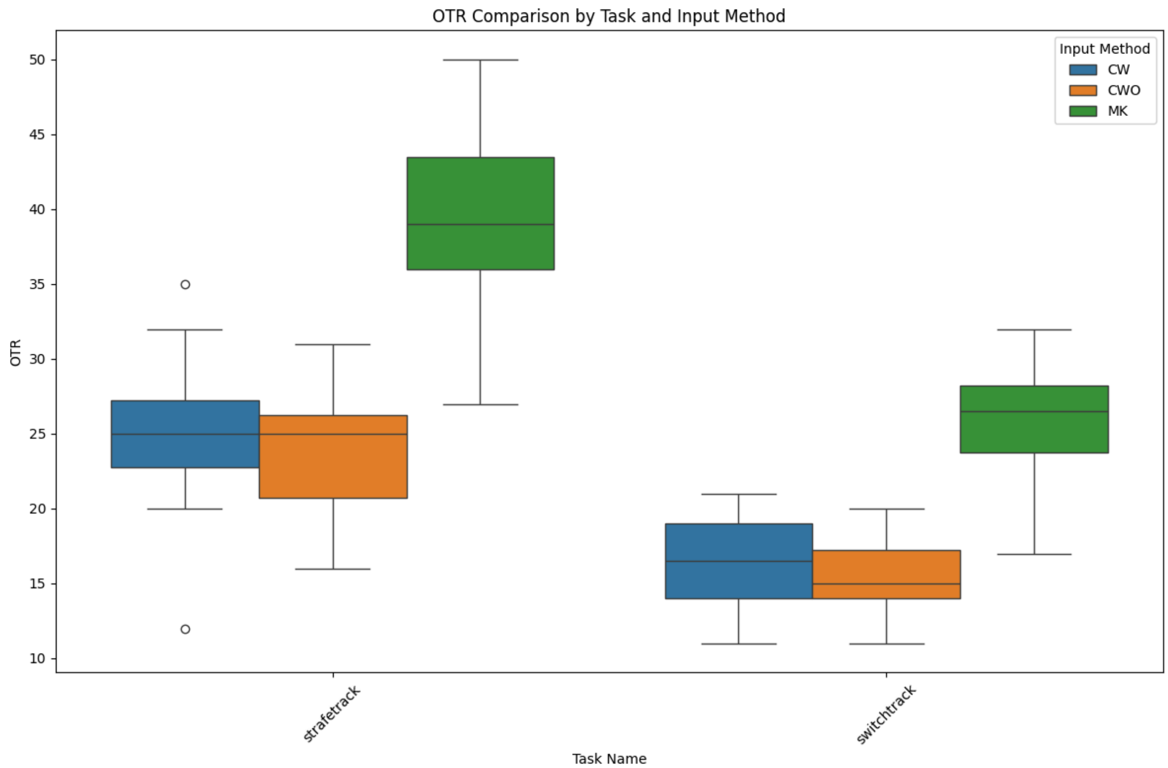


Figure 9: Boxplot of OTR for both ‘Tracking’ tasks, separated by category



Tables

Table 1: Normality assessment of the scores for each category

	Task	Category	Shapiro Stat	p-Value	Normality
0	Gridshot	CW	0.961401	0.687285	Yes
1	Gridshot	CWO	0.931652	0.258902	Yes
2	Gridshot	MK	0.914202	0.136026	Yes
3	Microshot	CW	0.914687	0.138498	Yes
4	Microshot	CWO	0.970561	0.847619	Yes
5	Microshot	MK	0.939910	0.347919	Yes
6	Strafetrack	CW	0.960391	0.668890	Yes
7	Strafetrack	CWO	0.962685	0.710684	Yes
8	Strafetrack	MK	0.956009	0.590357	Yes
9	Switchtrack	CW	0.951742	0.517785	Yes
10	Switchtrack	CWO	0.935346	0.295849	Yes
11	Switchtrack	MK	0.962571	0.708600	Yes

Table 2: Levene’s test on the scores

	Task	Levene Stat	p-Value	Equal Variances
0	Gridshot	2.889875	0.065954	Yes
1	Microshot	6.868284	0.002494	No
2	Strafetrack	4.623539	0.014922	No
3	Switchtrack	4.518224	0.016287	No

Table 3: ANOVA test on the scores of "Gridshot"

	Task	F-Statistic	p-Value	Significant
0	Gridshot	39.253802	1.361988e-10	Yes

Table 4: Welch's ANOVA test on the scores of "Microshot", "Strafetrack" and "Switchtrack"
DOF: Degrees of Freedom
BG: Between-Groups
WG: Within-Groups

	Task	**F-Statistic**	**p-Value**	**DoF (BG)**	**DoF(WG)**	**Effect Size (np2)**	**Significant**
0	Microshot	27.658019	2.350811e-07	2	27.986985	0.665782	Yes
1	Strafetrack	24.427214	6.824028e-07	2	28.332168	0.621628	Yes
2	Switchtrack	27.851378	2.231713e-07	2	27.927388	0.653582	Yes

Table 5: Post hoc comparison using Games-Howell test for the scores

Task	**Comparison**	**Mean Difference**	**p-value**	**Effect Size**	**Significant**
Gridshot	CW vs. MK	-23388.00	1.707e-06	-2.391	Yes
Gridshot	CWO vs. MK	-23623.81	1.485e-06	-2.366	Yes
Microshot	CW vs. MK	-24590.00	1.599e-06	-2.487	Yes
Microshot	CWO vs. MK	-25024.25	1.160e-06	-2.487	Yes
Strafetrack	CW vs. MK	-21536.19	4.746e-06	-2.207	Yes
Strafetrack	CWO vs. MK	-22781.12	2.009e-06	-2.385	Yes
Switchtrack	CW vs. MK	-25769.12	2.696e-06	-2.299	Yes
Switchtrack	CWO vs. MK	-27911.00	8.494e-07	-2.582	Yes
Gridshot	CW vs. CWO	235.81	9.938e-01	0.037	No
Microshot	CW vs. CWO	434.25	9.732e-01	0.077	No
Strafetrack	CW vs. CWO	1244.94	8.272e-01	0.203	No
Switchtrack	CW vs. CWO	2141.88	6.257e-01	0.321	No

Table 6: Normality assessment of the accuracies for "Gridshot" and "Microshot" for each category

	Task	Category	Shapiro Stat	p-Value	Normality
0	Gridshot	CW	0.951615	0.515693	Yes
1	Gridshot	CWO	0.962148	0.700894	Yes
2	Gridshot	MK	0.960204	0.665502	Yes
3	Microshot	CW	0.969063	0.823208	Yes
4	Microshot	CWO	0.857681	0.017710	No
5	Microshot	MK	0.940436	0.354384	Yes

Table 7: Levene's test on the accuracies for "Gridshot" and "Microshot"

	Task	Levene Stat	p-Value	Equal Variances
0	Gridshot	0.90939224	0.410041	Yes
1	Microshot	3.462486	0.039931	No

Table 8: ANOVA test on the accuracies for "Gridshot"

	Task	F-Statistic	p-Value	Significant
0	Gridshot	0.909394	0.410041	Yes

Table 9: Welch's ANOVA test on the accuracies for "Microshot"
DOF: Degrees of Freedom
BG: Between-Groups
WG: Within-Groups

	Task	F-Statistic	p-Value	DoF (BG)	DoF(WG)	Effect Size (np2)	Significant
0	Microshot	19.853273	0.000007	2	25.314508	0.355436	Yes

Table 10: Post-hoc comparison using the Games-Howell test for Accuracy

Task	Comparison	Mean Difference	p-value	Effect Size	Significant
Microshot	CW vs. MK	-9.45	3.181e-03	-1.310	Yes
Microshot	CWO vs. MK	-12.50	4.025e-05	-1.943	Yes
Microshot	CW vs. CWO	3.06	5.864e-01	0.343	No
Gridshot	CW vs. MK	-1.53	6.809e-01	-0.290	No
Gridshot	CWO vs. MK	-4.06	1.073e-01	-0.727	No
Gridshot	CW vs. CWO	2.53	4.810e-01	0.403	No

Table 11: Normality assessment of the OTR for “Strafetrack” and “Switchtrack”

	Task	Category	Shapiro Stat	p-Value	Normality
0	Strafetrack	CW	0.938388	0.329737	Yes
1	Strafetrack	CWO	0.963952	0.733688	Yes
2	Strafetrack	MK	0.952962	0.538013	Yes
3	Switchtrack	CW	0.945405	0.420556	Yes
4	Switchtrack	CWO	0.953187	0.541792	Yes
5	Switchtrack	MK	0.945970	0.428642	Yes

Table 12: Levene’s test on the OTR for “Strafetrack” and “Switchtrack”

	Task	Levene Stat	p-Value	Equal Variances
0	Strafetrack	1.567564	0.219727	Yes
1	Switchtrack	1.666667	0.200322	Yes

Table 13: ANOVA test on the OTR for “Strafetrack” and “Switchtrack”

	Task	F-Statistic	p-Value	Significant
0	Strafetrack	35.70841	5.15E-10	Yes
1	Switchtrack	44.03945	2.53994E-11	Yes

Table 14: Post-Hoc analysis using Games-Howell test for OTR

Task	**Comparison**	**Mean Difference**	**p-value**	**Effect Size**	**Significant**
Strafetrack	CWO vs. MK	-29.18	2.69E-02	-0.951	Yes
Strafetrack	CW vs. MK	-27.05	3.95E-02	-0.893	Yes
Switchtrack	CW vs. MK	-11.91	2.48E-01	-0.563	No
Switchtrack	CWO vs. MK	-10.34	4.51E-01	-0.421	No
Strafetrack	CW vs. CWO	2.14	9.65E-01	0.087	No
Switchtrack	CW vs. CWO	-1.57	9.83E-01	-0.061	No