# Evaluation

The last chapter discussed the key implementation details and concepts. We also conducted a study to evaluate the implemented learning tools and how they work together. Participants were divided into two groups and asked to carry out a test run in the Virtual Rival World. Given the broad scope of functionalities and measurements it was not possible to test everything in one test session. For that reason only the players’ standpoint was taking into account.

The remainder of this section will discuss the methodology and results of this study. The outcome of this study will help to understand racing games, improve Education, Enjoyment and Performance of drivers.

# Research focus

The Virtual Rivals development focus was on player usability. The data is recorded during the runs and feed into the Analysis Tool. We initially defined two research questions (See Section~\ref{}):

We want to understand players in racing games and improve Education, Enjoyment and Performance of drivers. For this evaluation, we defined six more detailed questions:

The next section introduces the study process in more detail.

1. What is the difference between self-assessment and real driving skill?
   1. Automatically skill-adjustment is necessary to improve \textit{Engagement}, \textit{Education} and \textit{Performance}
2. How is the relationship between personality and speed?
   1. Personalised racing simulations to improve \textit{Education} and \textit{Performance}.
3. How is the relationship between personality and accidents?
   1. Personalised racing simulations to improve \textit{Education} and \textit{Performance}.
4. How does ghost car / virtual rival influence \textit{Engagement}?
5. How does ghost car / virtual rival influence \textit{Education}?
6. How does ghost car / virtual rival influence \textit{Performance}?

# Methodology

## Big Five

Chapter~\ref{} introduces the theoretical background of the Big Five personal trait. For our study, we implemented the short evaluation questionnaire illustrated in Table~\ref{} based on \citep{Rammstead2007}. The Big Five theory presents a model in which personality is organized into five factors:

* Extraversion: Manifests in an outgoing and energetic behaviour.
* Agreeableness: Perceived as kind and cooperative.
* Conscientious: Implies a desire to do a task well, being careful and efficient.
* Neuroticism: Tend to be emotionally unstable e.g. more likely to feel anger and frustration
* Openness: More likely to be creative and tolerant.

The Big Five evaluation questionnaire should help to find a relationship between personalities, driving and gaming. In the next section, we discuss our second personality metric specialised on risk taking.

## Sensation Seeking

Chapter~\ref{} discusses Sensation Seeking personal trait as indicator for risk taking. For our study, we implemented the Brief Sensation Seeking Scale illustrated in Table~\ref{} based on \citep{ Donohew}. The Brief Sensation Seeking Scale significantly predicts intention to and actual engagement in a number of health risk behaviours. We are particularly interested to find a relationship between Sensation Seeking, driving and gaming. In the next section, we introduce the integrated driving measurements.

## Driving Data

Chapter~\ref{} introduced driving metrics to measure performance. For our study, we measured various driving metrics e.g. speed, trajectories, driving errors and resets. Figure~\ref illustrates one of the recorded driving trajectories. Data points are captured every 200ms. Accidents where the player touched the barrier are marked in red. Reset points are indicated in green. Additionally we record speed for every time-step, lap times and sector times. The driving data is vital to measure and analyse performance. The next section introduces a method to capture Emotions.

## Emotions Capture Method

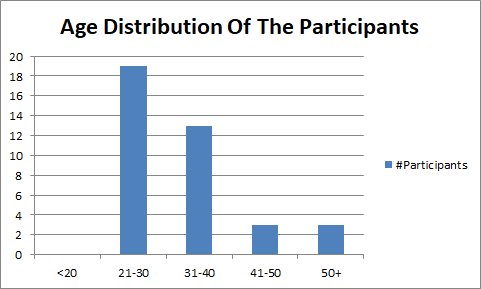
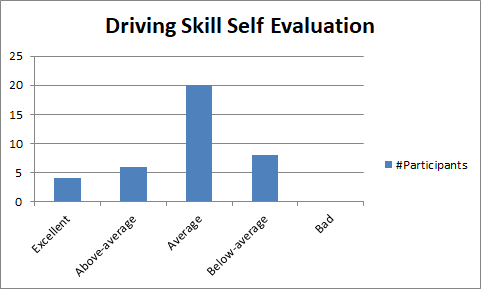
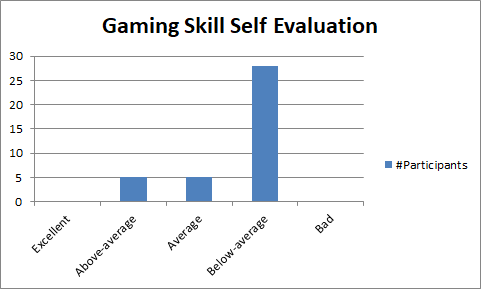
Chapter~\ref{} discussed the Wheel of Emotions as tool to measure Enjoyment and Motivation. Figure~\ref{} illustrates the integration of the Wheel of Emotions in our Virtual Rival Framework based on \citep{Scherer}. The most important emotions we need to track in race games are:

* Pride: When winning in general.
* Interest: Desire and little control over the situation.
* Challenge: The desired goal takes a lot of effort but is still reachable.
* Surprise: Unexpected situations with little effort.
* Boredom: When the mind is not challenged results in low effort and attention.
* Anger: Comes in unfair situations.
* Frustration: When success is expected, failure is often accompanied by frustration.

Racing simulations have to control the challenge and the certainty of the situation to optimize Enjoyment, Motivation and Performance. The next section introduces the background of the participants in this study.

# Participants

A total of 48 people participated in this study (14 female / 34 male). Participants came from different backgrounds and between the ages of 20 and 50 years. Figure~\ref{} illustrates the age distribution. In order to participate, they only needed a computer or laptop. The participants were recruited using Amazon Mechanical Turk. All of the participants have real world driving and gaming experience. Figure~\ref{} illustrates the results of the real world driving skill self-evaluation. Most participants rated their driving skill as average. The gaming skill self-evaluation had similar results as shown in Figure~\ref{}.



# Procedure

Two test sessions were carried out over the course of two weeks. In each session we invited a group of users to try our game. First, the participants were greeted and informed about the structure of the study. In a short tutorial level, we showed how to control the car.

Each task was given to the students only after they have completed the previous one. The participants received their tasks from Amazon Mechanical Turk. Every participant receives a unique ID. Entering the ID confirms the completion of the tasks. The procedure is described in Figure \ref{}.

Tasks

1. The first task is to drive around and explore the race track. This is used to learn to control the car. The participants were not given any specific tasks or opponents.
2. The second task is to enter basic information used to characterise the participant e.g. age, driving skill, experience with video games.
3. In the third tasks user have to complete two personality tests. We implemented the Big Five and Sensation Seeking personality test, consisting out of multiple questions. The Big Five personality labels the human personality broadly in five dimensions. The Sensation Seeking personality test is specialised on risk taking.
4. The fourth tasks are two evaluation rounds. Participants have to drive as fast as possible. During the rounds we estimate the initial skill.
5. The main task consists of three parts. Participants are randomly assigned in two groups. The first group is the reference group racing against a classic race ghost, a shadow of the last round. The second group races against a virtual rival with automatic adjusted difficulty. After each lap the participants have to self-evaluate their performance. The last part is a short evaluation of the current emotions and motivation of the participant. This task will be performed multiple times which needs the most time for completion.
6. The last task is to confirm the successful completion by entering the participant ID.

# Result

We used Pearson’s r to assess the relationship between variables. We assessed the normality of the data with the Shapiro-Wilk test. For five variables—neuroticism, sensation seeking, fatalistic present time perspective, mean speed in the 80/60 scenario, and mean speed in the 80/80 scenario—we found the p-value to be less than 0.05 and rejected the null hypothesis about their normal distribution. Because Pearson’s correlation coefficient assumes the normality of variables, we used Spearman’s correlation coefficient q to assess the relationship between pairs of variables where one of the variables would not meet the assumption of normality. We used Scipy.stats version 1.3.0 to compute the correlations. - Linkov

We used Pearson’s r to investigate the relationship between variables. Pearson’s correlation coefficient assumes the normality of variables. We assessed the normality of the data with the p-value to test for non-correlation. In cases where one of the variables would not meet the assumption of normality we used Spearman’s correlation coefficient q to find correlations. We used Scipy.stats version 1.3.0 to compute the correlations.

Because of the large variations commonly present in human behaviour and the large number of factors influencing this behaviour (personality, intelligence and learned associations) psychologists consider the following correlations to be indicative for effect sizes in a relationship between personality and the participants' game behaviour (Cohen, 1988, 1992).

* Weak Correlation: r=0.1 (1% of variance explained)
* Medium Correlation: r=0.3 (9% of variance explained)
* Strong Correlation: r=0.5 (25% of variance explained)

In this section, we evaluate the data gathered during the study. Our analysis tool helps to process the data (See Section). The focus is on answering the six core research questions of this work.

1. What is the difference between self-assessment and real driving skill?
2. How is the relationship between personality and speed?
3. How is the relationship between personality and accidents?
4. How does ghost car / virtual rival influence \textit{Engagement}?
5. How does ghost car / virtual rival influence \textit{Education}?
6. How does ghost car / virtual rival influence \textit{Performance}?

# Useful

Video games as a tool to train visual skills

Does driving experience in video games count? Hazard anticipation and visual exploration of male gamers as function of driving experience

# Future Work

Personality-based game adaptation If a game is filled with content that coincides with a player’s preference (compared to a generic game) then the preference-filled game will be rated more positively. Following this line of reasoning we could consider games that can adapt their content to a player’s preferences as desirable games. In our research in the Chapters 4 through 7 we see that personality leads players to choose some specific options in the game over others. We believe that we may state that these patterns of choice indicate what a player’s preferences are. Moreover, it may be possible to develop a game with lists of potential situations and choices for each personality configuration. The choices can be provided to the player as soon as a player’s personality is indicated by his behaviour in the game. Dynamically altering a game in this way should increase the game’s attractiveness to each player for which it can provide adequate content for his personality. Several techniques have been developed for adapting games. An example of such a technique is dynamic scripting Spronck et al. (2006). Matching a technique to the task of personality-based adaptation will require further consideration. An investigation by Schreurs (2011) indicates that adaptation based on personality is challenging. In Schreurs’ investigation, the participant’s personality was measured and correlated to the reported appreciation of three in-game situations (a fight, a puzzle, and a social challenge). The participant’s demographics were also measured. The investigation shows a few correlations between personality and situation appreciation. However, the investigation shows that demographic factors such as skill and gameplay experience correlate significantly in multiple situation-personality combinations (for the full details please consult the work by Schreurs (2011)). Therefore, demographics should be considered as control variables. However, we maintain that personality-based game adaptation shows much promise and that this line of research deserves attention in the future.

Games as personality profiling tools – Lankveld