A MULTI-AGENT SYSTEM FOR AN INTELLIGENT DRIVING INSTRUCTION APPLICATION

MODELLING ADAPTIVE SCENARIOS AND FEEDBACK TO IMPROVE THE USER EXPERIENCE

Master's Thesis

Miriam Doorn

August 29, 2023

A MULTI-AGENT SYSTEM FOR AN INTELLIGENT DRIVING INSTRUCTION APPLICATION

MODELLING ADAPTIVE SCENARIOS AND FEEDBACK TO IMPROVE THE USER EXPERIENCE

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in

COMPUTER SCIENCE

by

Miriam Doorn
born in Kampen, the Netherlands

Interactive Intelligence Group
Department of Data Science
Faculty EEMCS, Delft University of Technology
Delft, the Netherlands
http://www.ewi.tudelft.nl

Alten
Technical Software West
Fascinatio Boulevard 582
2909 VA Capelle a/d IJssel
The Netherland
http://www.alten.nl

A MULTI-AGENT SYSTEM FOR AN INTELLIGENT DRIVING INSTRUCTION APPLICATION

MODELLING ADAPTIVE SCENARIOS AND FEEDBACK TO IMPROVE THE USER EXPERIENCE

Author: Miriam Doorn
Student id: 1280333

Thesis Committee:

Chair & Supervisor: Prof. Dr. M.A. Neerincx
Committee member: Dr. C. Lofi
Assistant supervisor: M. Al Owayyed

Keywords: scenario-based training, serious games, personalization, intelligent agents, adaptive difficulty

Front: Screenshot from the driving instructor application.

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

Contents

SUMMARY

Drivers need to build up experience to learn how to deal with the varying situations that are common when driving a car. Virtual driving simulators provide a controlled, safe environment for the user to practice. State-of-the-art driving simulators are high-cost and lack personalized guidance. For this thesis project, a low-cost virtual driving instructor was developed with the goal of improving learning efficiency and user satisfaction by adapting the content to the user's skill and knowledge level.

The system design was based upon the work of Peeters, which offers a useful model for a multi-agent system where a 'scenario planner' determines the content of a scenario, the 'director' controls the different elements during the scenario, the 'monitor' measures the performance and the 'reflector' shows an evaluation of performance after finishing a scenario. In the implementation, a similar design and workflow was used.

Many driving scenarios were designed, varying in difficulty by changing the road layout, driving goals, and the behaviour of other traffic users. Each scenario trains the user in some driving skills, i.e. lane-keeping, turning corners, and/or knowledge, of traffic rules such as stopping for a stop sign. The user model is updated based on the user's performance. This then determines which scenario will be chosen next.

A research experiment was conducted to explore the effects that adaptive difficulty has on the user experience. The experiment has a within-subjects design with two equally sized groups of participants, one starting with a linear session and the other with an adaptive session. For the linear session, the difficulty rises steadily with each subsequent scenario. For the adaptive session, the difficulty of the scenarios can rise and fall depending on the user's performance.

The statistical analysis of the research data showed no significant difference in the acquisition of skills and knowledge between the two groups. There was however a medium-size significant difference found in self-efficacy. Users playing the adaptive session reported a higher self-efficacy than users playing the linear session.

Very little research has been done on the effects of personalized content on the user experience in the domain of virtual driving simulation. The prototype that was implemented for this thesis project is a good starting point to improve upon. The results of the research experiments were inconclusive in some areas. The increase in self-efficacy however validates the positive value that personalization could bring to training applications. It can be theorized that by improving the application to better adapt to the user, the user experience can be enhanced even further.

Preface

This project evolved from a concept design by Tomas Storck and Marnix Kraus, both at the time consultants at consultancy company Alten. They envisioned a low-cost driving instructor simulation built on the game Grand Theft Auto V. The project appealed to me because I love video games and have an interest in game development and educational software. Finding where the project could make a scientific contribution took some brainstorming and research, eventually settling on adaptive difficulty and its effects on the user experience.

While the game engine and environment were there as a basis to start from, the driving instructor application needed to be built from scratch. Determining the scope of the implementation was an ongoing struggle. While more time was spent on the implementation than initially planned, I believe it was necessary to be able to get any meaningful results. Working at the Alten office provided a supportive environment to conduct my research and work on the implementation.

Due to personal circumstances, I took a break from working on the thesis. While the intent was always there to finish, it took a phone call from academic counsellor Monique Draijer to get me back on track to finishing my thesis. Thanks to her and the weekly meetings of the Interactive Intelligence research group led by Willem-Paul Brinkman, I was motivated again to continue working on the final phase of the project.

I would like to thank my supervisor, Mark Neerincx, and assistant supervisor, Mohammed Al Owayyed, for their guidance and feedback. Their insight improved the quality of my work and made me think more deeply about the research. I would also like to extend thanks to Milan van Hoek, who mentored me during my time at Alten. Together we designed the many traffic scenarios that are in the application.

Finally, I would like to thank my friends, family and boyfriend for their support and kindness.

Miriam Doorn
August 29, 2023, Rotterdam

1

Introduction

Bring me my coffee or I'm gonna cut your arm off.

Trevor Philips, GTA V

In this chapter a general outline of the project is given. First, in section 1.1, an introduction is given to the domain of the project and how this project aims to improve upon the current way of working. The research questions that guide the development of the application and subsequent experiment are presented in section 1.2. In section 1.3 the scientific and societal contribution this research aims to make is discussed. This is followed by a task analysis, in section 1.4, where a user-centred approach is taken to determine the design goals of the project.

1.1. Problem Description

Drivers need to build up experience to learn how to deal with the varying situations that are common when driving a car. Traditionally this has been done by way of one-on-one driving lessons where an instructor accompanies the student-driver when operating a vehicle. Lessons are followed either in a controlled environment, safely traversing the familiar orange cones through the training area, or on the somewhat less predictable streets. During the lessons, the instructor will adjust their feedback and lesson plan according to the situation and the skills of the driver, as each student is different and learns at their own pace. This method of teaching is cost-heavy because the instructor has to be physically present during each lesson.[1]

The use of virtual driving simulators has certain advantages. In a controlled virtual environment there is freedom to design the environment, the weather conditions, and the behavior of other road users, such as cars and pedestrians. This allows the user to practice in a wide variety of situations. Even situations that would be dangerous in real life. Game-based virtual driving simulators are available for different types of vehi-

1

cles. For example, the simulators that are made by Green Dino[1]. Using these simulators, drivers can gain experience in a safe environment. While these simulators go a long way in having road users face various scripted scenarios, they lack personalized guidance.

For this thesis project, the goal is to improve learning efficiency and user satisfaction by adapting the content to the user's skill and knowledge level. To achieve this a profile has to be created modelling the user's level of knowledge and proficiency in driving. Based on this profile and the knowledge base of teaching material, a suitable training scenario will be selected automatically. On completing the scenario the user will be evaluated and the user model will be adjusted based on the user's performance in the scenario. To achieve this goal the research will focus on the fields of Procedurally Generated Scenarios, Intelligent Tutoring Systems, User Modelling, and Knowledge Representation.

1.2. RESEARCH QUESTIONS

The following research questions will guide the research, design, implementation, and evaluation of the project:

RESEARCH QUESTION 1

The goal of the research is to develop a method of personalizing the content of a virtual driving simulator to match the performance level of the user in order to improve the effectiveness and enjoyment of training. The experience should be natural and enjoyable for the user. High user satisfaction will motivate the continued use of the application.

> *What methods and techniques can be utilized to implement a virtual driving simulator that has adaptive difficulty?*

By personalizing content, the experience should be unique to each user. In looking at previous work on virtual learning we can get ideas on how to design our system.

RESEARCH QUESTION 2

In order to personalize content, know the user. What information is needed from the user and how is this data acquired? In the model, it is necessary to keep track of the knowledge and skills the user has acquired. This way scenarios can be generated where the user can learn new skills at a pace suited to them.

> *In what way can the user and learning content be modelled to enable personalized learning?*

To evaluate the driver's performance and give feedback the measures that are used for evaluation need to be defined. For example, the car's position on the road (or the deviation from the centre of the road). To procedurally generate a scenario for a driving lesson, the content must be modular. This way a scenario can be built by combining the best-fitting modules to match the user. This means a function is needed to connect the content model to the user model. The structure should be such that there is flexibility to generate a variety of scenarios. It should also be reliable in that the generated scenarios

[1] https://www.greendino.nl/simulatoren

have a logical flow and are suitable to the user's needs. The content that is generated aims to be not too hard and not too easy. By setting the difficulty somewhat above the user's skill level the user will have the right amount of stress to stay alert while not being overwhelmed. A certain amount of frustration has proven to improve learning[2]. At the end of a personalized scenario, it is expected the user will be more satisfied with how the lesson went and more confident in learning more.

RESEARCH QUESTION 3

This project requires a virtual driving simulator to be built using the GTA V game environment. As the game was not built for this purpose we could come across issues that make it difficult to implement our designs.

> *What are the challenges involved in implementing a virtual driving simulator with adaptive difficulty in the GTA V game environment?*

To what extent is the GTA V game environment a suitable one for virtual driving simulation? What difficulties can be overcome and which are impossible to mitigate?

RESEARCH QUESTION 4

To test the effectiveness of an adaptive method it needs to be compared to learning scenarios with a static lesson plan.

> *What is the impact of adaptive difficulty on the user's skill and knowledge gain and their self-efficacy when learning in a virtual driving simulation?*

A research experiment needs to be designed and conducted to find an answer to this question.

1.3. MOTIVATION

The conception of this project was possible due to ongoing improvements in technology and techniques in user interaction enabling improvements in the current way of learning how to drive. There is both a scientific motivation and a societal motivation for wanting to increase the effectiveness and efficiency of training applications for driving.

SCIENTIFIC MOTIVATION

In many fields of occupation, virtual training has proven to be an effective learning method. Flight simulators have a history of being an effective training tool for pilots [3]. A more recent study on the use of a virtual reality surgery simulator showed improved self-confidence and performance in the group that used the simulation to train [4]. Virtual driving simulation has helped novice drivers gain confidence and experience in driving while being in a safe environment [5]. Learning is at its most effective when the level of the learning material is attuned to the user's skill and knowledge level [2]. If the learning material is too easy, the student becomes bored and distracted. If the learning material is too difficult, the student becomes demotivated. Innovative methods can be employed to analyze the driving behaviour of the student and create a personalized environment that is optimized for learning.

1

Conceptual models have been conceived for a system such as this. For example, Romoser [6] proposes a simulator that utilizes an intelligent tutoring system (ITS) for driving instruction that would be capable of automatically diagnosing driver performance, providing feedback, and customizing the curriculum to fit the learning needs of the student driver. In his conclusion, Romoser explains the difficulties inherent in building such a system and the many design decisions that need to be taken, such as learning strategy, guidelines for good driving performance, information flow between the modules of the system, etc.

Considering these difficulties it is no surprise that interesting implementations that take advantage of difficulty adjustment in virtual driving instructors are hard to find. This makes it a worthwhile area of study. By implementing and evaluating such a system a valuable contribution can be made to the science of user-experience-focused training simulation not just in the domain of car driving. It can serve as a case study for other domains as well.

SOCIETAL MOTIVATION

Research has been done and is being done in the field of driving. The state of the art promises autonomous cars that are capable of traversing traffic without the input of any human driver. Modern cars offer advanced driver assistance systems. You could say a revolution in driving is on its way. It is not surprising so much effort is being put into improving the way we travel as human error while driving is the most common cause of accidents in traffic [7]. Having fully automated vehicles take care of all transportation needs is still a long way off. In the meantime, it is important that student drivers can practice efficiently and effectively. Virtual driving instructors can be a low-cost solution for users to practice driving in varying situations.

1.4. TASK ANALYSIS

A task analysis is done to understand the current situation in which the task is performed and to see where improvements can be made. In the case of driving instruction, the task of the user is to follow lessons and to learn and gain experience from these lessons. The results of the task analysis will guide the design of the application which is described in chapter 3.

1.4.1. PACT: PEOPLE, ACTIVITIES, CONTEXT, TECHNOLOGIES

The application will be designed with a human-centred approach. This means the application will be tailor-made to suit the needs of the end-users. A PACT-approach (**p**eople who undertake **a**ctivities in **c**ontexts using **t**echnologies) as defined by Benyon [8] is used to analyze the task of following virtual driving lessons. Each element: people, activities, context, and technology is discussed in a subsection.

PEOPLE

The target group of our application consists of people who are starting or have started training for a driver's license for operating a car on public roads. In the Netherlands, the minimum age to start driving lessons is 16.5 years old. It is however common for people to start lessons at a later age. There is no minimum age limit to driving in a simulation. Therefore the target age for the application is more of a loose definition. In the target group, there are a lot of differences between individuals. In this section, a list is given showing this group's general characteristics that are relevant when designing the application.

Characteristics
Here the relevant physical and psychological characteristics and skills that may apply to a large number of members of our target group are summarized:

- Age from 16 to 30
- Computer and Gaming experience varying from low to high
- Some experience as a traffic participant
- Some knowledge of traffic rules
- Learning speed varying from slow to fast
- Good eyesight and motor skills

Persona
A common design practice in user-centred design is creating so-called personas. Personas are descriptions of fictional characters: their goals and characteristics as relevant to the project. These personas represent the typical end-users of your application. Having a concrete example of the end-users gives direction in defining and designing the application. If you design your application to satisfy a persona that represents the typical user, your application will ideally satisfy a large group of users that are similar to the persona.

Here is a persona that can be seen as a typical user of our application:

> *Diana Vos is a 16-year-old girl who wants to get a head start on learning how to drive. She has never operated a car, not even in a simulation, but has experience in traffic as a pedestrian and a cyclist. She, therefore, has some knowledge of general traffic rules. She knows how to work with a computer. Apart from casual puzzle games on mobile, Diana isn't an avid gamer.*

ACTIVITIES
When following real-life driving lessons it is recommended that they are followed regularly and continuously until the student driver passes the driving exam. Distributed practice, a little and often, is more effective than cramming many lessons in a short time [1]. It is very important that the driver gets comfortable and confident in operating the vehicle.

The goal of our application is to teach people driving skills and monitor their progress. In using the application, the user can build confidence in operating a car and stays motivated to continue learning. For better retention of skills and knowledge, the application should be used several times a week. While this is not required, it is highly recommended. One learning session can take 15 to 30 minutes. The user, however, can take as much time as he needs and the session is not affected by outside interruptions. If a user makes a mistake, he will receive feedback on it and can continue using the application.

Use Scenario

This is a typical use scenario a user is envisioned to go through when using the application over multiple sessions.

- First Contact: The system has no information on the user's level. While playing the first scenario the system will evaluate the user's actions and initialize the student model with the score that was obtained.
- Next Contact: Based on the student model the system will automatically select a scenario. After playing the scenario the student model is updated.
- Continued use: The system will continue presenting scenarios with new situations optimized for the learner.
- End case: The user has acquired the skills taught by the application. The learner is welcome to return to revise and practice the knowledge and skills acquired.

CONTEXT

The application will be implemented as a modification (mod) of an existing PC game (Grand Theft Auto V). Anyone who owns the original game can theoretically install the mod to use the application. The physical location where they use the application is therefore in their homes. The video game GTA V can not be licensed by for-profit businesses. The application as implemented is considered to be a testbed to evaluate personalized learning. While, as is, the application can not be used in other driving simulators, the techniques if proven effective would be an innovative addition to the lesson plans of traditional driving simulators.

TECHNOLOGY

It is common for driving simulators used in training to have a large expensive set-up with multiple screens, large seats, and realistic controls. A long-term study using the VS500M high-fidelity driving simulator[2] showed a positive transfer of skill to on-road driving[9]. A cost-effective and accessible alternative that can run on a regular PC set-up is proposed.

To use the application a computer with the application installed is needed (further explained in section 4.1). The user controls the application either with the keyboard and the mouse or with a steering wheel and pedals. The keyboard is used to control the car. The mouse is used to look around. While not necessary to run the application, the use of a steering wheel controller with pedals is strongly recommended, as it improves immersion and helps the user get a feel for driving in a car. Ideally, the steering wheel has a large steering angle and support for force feedback. When using a steering wheel controller the view on the screen is static, always facing forward. During a driving session, the user will receive visual feedback.

Skill Transfer

The application aims to offer a realistic driving experience but is limited in that a PC application with one display will not be able to accurately capture the experience of driving

[2] https://viragesimulation.com/vs500m-car-simulator-training-and-research/

in a car. In traditional real-life driving lessons, the use of mirrors is an essential skill to learn. The application instead focuses on presenting varying traffic situations so that the user learns how to handle himself in traffic. This falls under the theory part of learning to drive, which traditionally consists of memorizing the traffic rules from a book and testing the knowledge with multiple-choice questions. The application's method aims to give the learner a more life-like experience in taking the correct action in an interactive scenario.

1.4.2. CLAIMS ANALYSIS

In the claims analysis, the effects and consequences of using the application are defined.

Claims List:

- By using the system the user will learn the skills and knowledge necessary for driving a car.
- By using the system the user will feel motivated to learn more and improve his skills.
- The system will give the user an accurate indication of his current skill/knowledge level.

REFERENCES

[1] J. Miller and M. Stacey, *The Driving Instructor's Handbook* (Kogan Page, Limited, 2005).

[2] S. D'Mello and A. Graesser, *Dynamics of affective states during complex learning,* Learning and Instruction **22**, 145 (2012).

[3] R. T. Hays, J. W. Jacobs, C. Prince, and E. Salas, *Flight simulator training effectiveness: A meta-analysis,* Military psychology **4**, 63 (1992).

[4] Y. Pulijala, M. Ma, M. Pears, D. Peebles, and A. Ayoub, *Effectiveness of immersive virtual reality in surgical training—a randomized control trial,* Journal of Oral and Maxillofacial Surgery **76**, 1065 (2018).

[5] K. Ivancic IV and B. Hesketh, *Learning from errors in a driving simulation: Effects on driving skill and self-confidence,* Ergonomics **43**, 1966 (2000).

[6] M. Romoser, *An autonomous intelligent driving simulation tutor for driver training and remediation: A concept paper,* in *Driving Assessment 2011: 6th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle DesignHonda R&D Americas, IncorporatedNissan Technical Center, North America-Toyota Collaborative Safety Research CenterFederal Motor Carrier Safety AdministrationUniversity of Iowa, Iowa City* (2011).

[7] A. Haghi, D. Ketabi, M. Ghanbari, and H. Rajabi, *Assessment of human errors in driving accidents; analysis of the causes based on aberrant behaviors,* Life Science Journal **11**, 414 (2014).

1

[8] D. Benyon, P. Turner, and S. Turner, *Designing interactive systems: People, activities, contexts, technologies* (Pearson Education, 2005).

[9] P. Hirsch and F. Bellavance, *Pilot project to validate the transfer of training of driving skills learned on a high fidelity driving simulator to on-road driving*, Vol. 16 (Cirrelt, 2016).

ALTEN

2

Domain Analysis

You forget a thousand things every day, pal. Make sure this is one of 'em.

Michael De Santa, GTA V

At the outset, the project outline, as conceptualized by an employee of the consultancy company Alten, was the development of a low-cost driving simulator using the game Grand Theft Auto V. In order to determine the research approach for the thesis a brainstorming session was conducted together with a few interested Alten consultants. For inspiration, the basic building blocks of a multi-agent system designed by Peeters[1] were used. Figure 2.1 shows the results of this brainstorm session. The white squares represent the agent roles. The keywords surrounding each agent block reveal a wide variety of interesting research subjects to further explore. Modelling the user's skills and the knowledge base will be important for the entire system.

Within the scope of the project, it is impossible to expand on each of the subjects revealed by the brainstorm. For this thesis project, we want to build an application that improves learning efficiency and user satisfaction by adapting the content to the user's skill and knowledge level. Before designing, we first consider the requirements (2.1) such an application would have in order to facilitate learning with adaptive difficulty. This is followed by research into the current state of fields related to our design goals: Driving Simulators (2.2), Procedurally Generated Scenarios (2.3), Intelligent Tutoring Systems (2.4), User Modelling and Knowledge Representation (2.5).

2.1. Requirements

From the brainstorm session, we defined the following list of requirements for the application. Requirements are defined to validate an application design. It connects the application to the user's needs and desires.

Requirement List:

2

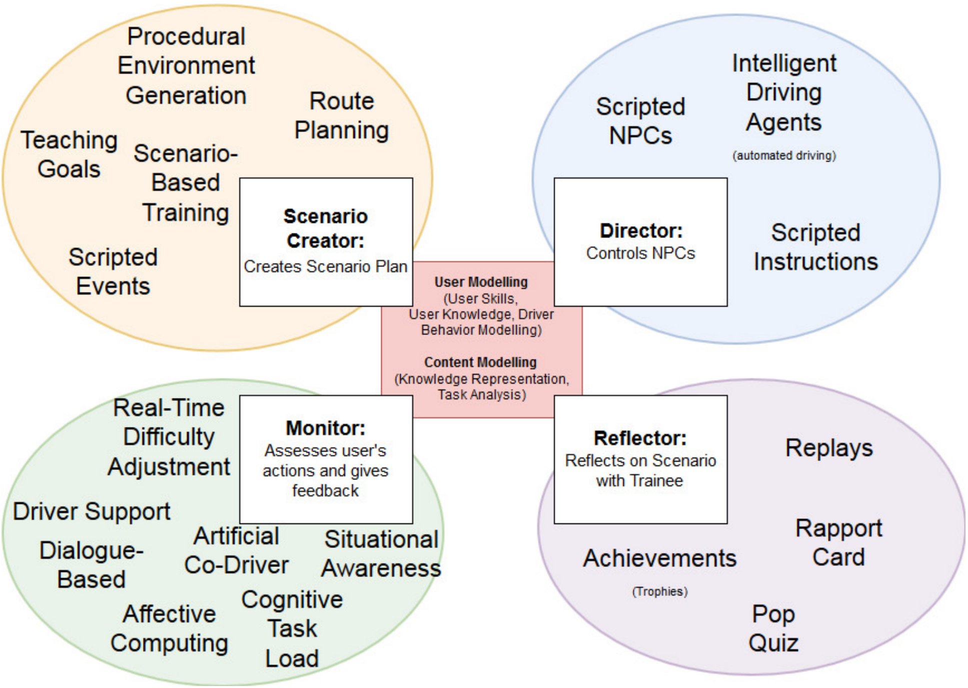


Figure 2.1: Brainstorm on possible features for a virtual driving instructor.

- The system must keep track of user data over several learning sessions.
- The system must be able to design/save traffic scenarios.
- The system must be able to pause/resume a session on interruption.
- The system must offer a semi-realistic environment (steering wheel/pedal input).

2.2. Driving Simulators

Driving simulators are used for various purposes. The reason being the ease of reproducing varying driving scenarios. Simulators have been used to study driver behaviour, such as comparing the driving performance of certain age groups [2]. In a study such as one on the effect of alcohol on driving performance a virtual environment provides a safe space to run experiments [3]. Driving simulators are also heavily used to evaluate advanced driver-assistance systems (ADAS) [4] [5]. Examples of ADAS are automated cruise control, automated braking for collision prevention, automatic parking etc.

The two biggest issues in the use of driving simulators for learning are the cost and skill transfer, whether skills learned in the simulation transfer to real-life driving.

Kappé et al. [6] suggest a good driving simulator is not necessarily one that most approaches reality in driving. The didactic value of the system is much more important. Many driving tasks can be taught using a cost-effective driving simulator. Current PCs are capable of rendering a sufficiently realistic environment for driving simulators with a high resolution, large field of view and a good frame rate.

2.3. Procedurally Generated Scenarios

Procedural generation refers to generating some form of content automatically using algorithms receiving input from a random seed or a defined variable (or number of variables). In general, procedural algorithms often have a random element, so that given a few simple parameters a large variety of content can be created. As a research field procedurally generated content is fairly new. In game development however procedural algorithms have been used since 1980 when the game *Rogue* used procedural generation to create dungeons for the player to explore.

Yannakakis and Togelius [7] introduce an approach for generating content to personalize the experience for the user, which they call Experience-Driven Procedural Content Generation (EDPCG). An important part of the process is modelling the content in a way that makes it possible to introduce variation while keeping a consistent internal logic. For example, when procedurally generating a platformer level with varying difficulty, you can have easy jumps (small gaps) or hard jumps (wide gaps), but you don't want gaps so wide that they are impossible to jump.

In 2014, Bhatti [8] wrote his thesis on generating scenarios for driving instruction and experiments within a driving simulator. His model is for creating user-generated scenarios. He provides several schemas on how to model different modules of a training scenario. By applying procedural generation algorithms on these scenario modules it is possible to generate different scenarios automatically.

In [1], Peeters presents a multi-agent architecture where four agents (the scenario planner, the director, the monitor and the reflector) work together to adapt the scenarios used in scenario-based training. This adapted scenario is tailor-made to fit the user model which is updated at the end of each scenario.

2.4. INTELLIGENT TUTORING SYSTEMS

An intelligent tutoring system (ITS) is a computer system that aims to provide immediate and customized instruction or feedback to learners, usually without requiring intervention from a human teacher. Different models exist for ITSs. The general characteristics as defined by Padayachee [9] are as follows. The Domain Model contains the knowledge to make inferences or solve problems. The Tutoring Model contains system teaching goals and plans. The Student Model maintains information about the student's knowledge, skills, learning preferences and past learning experiences. System Control provides helpful feedback on student input. The User Interface provides the learning environment, promotes ease of use, incorporates natural interaction dialogues and so on. It is clear there is some overlap with the previous section on Procedurally Generated Scenarios in that the Domain Model (Content) and the Student Model (User) determine the Tutoring Model (Scenario).

2.5. USER MODELLING AND KNOWLEDGE REPRESENTATION

To make a system that adapts itself to individual users you need to gather information about the user. What information do we gather? How do we model the data?

2.5.1. HOW TO MODEL DRIVING TASKS

The task of driving a car can be divided into three levels of skills and control, strategic (planning), tactical (manoeuvring), and operational (control) [10]. The strategic level determines the general planning of the drive, such as route planning and mode of transportation. The tactical level consists of tasks such as obstacle avoidance, overtaking etc. These tasks can then be subdivided at the operational level, where the driver determines the amount of steering and acceleration. Information can flow between the levels with the upper levels guiding the actions of the lower levels as shown in figure 2.2.

Each of the skill levels is required to get a driver's license. At the strategic level, the trainee is required to be able to read signs to follow the route to the chosen destination. At the tactical level, the trainee is required to be able to perform the tasks and make correct decisions on when to perform them. The operational level looks at low-level skills such as lane-keeping.

Michon [10] categorizes different driver behaviour models as follows. A distinction is made between Input-Output (Behavioral) and Internal State (Psychological). Another distinction is between taxonomic and functional models. Taxonomic models do not have any interaction between the components of the model, while functional models do. A taxonomic model could be seen as a collection of facts. Fleishman's taxonomy of human performance [11] is a trait model for perceptual, cognitive and motor skills. Of particular interest are the perceptual-motor abilities, such as control precision, reaction time and rate control (timing). A combination of these abilities is necessary for perform-

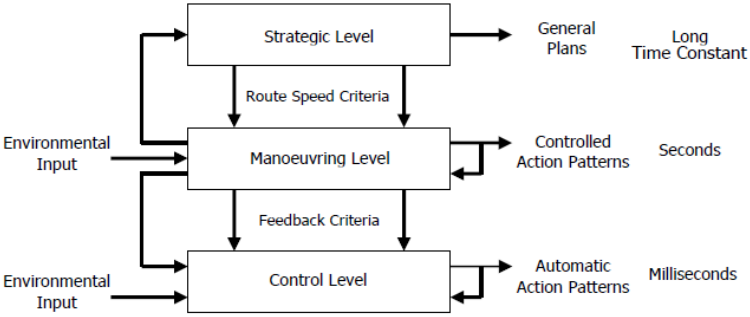


Figure 2.2: Task hierarchy for driving.

	Taxonomic	Functional
Input-Output (Behavioral)	Task Analyses	Mechanistic Models Adaptive Control Models - Servo-Control - Information Flow Control
Internal State (Psychological)	Trait Models	Motivational Models Cognitive (Process) Models

Figure 2.3: Summary of driver behaviour model types.

ing many driving tasks.

2.5.2. Common errors and violations in driving

In an effort to analyze the human factors contributing to accidents in traffic, the Driver Behaviour Questionaire (DBQ) was developed [12]. The DBQ is a self-report containing 50 items related to the frequency of risky behaviour in driving. The DBQ or variations of it have been used widely in research into the differences in driver behaviour between gender, age, driving experience and culture. The original DBQ makes a distinction between errors and violations as they are related to different cognitive processes and motivations. Violations are seen as deliberate, conscious deviations from safe practice, for example, speeding and drunk driving. Errors are divided into lapses and mistakes. Both are unintentional with the difference being that lapses are caused by inattention or slips of memory and mistakes are unintended consequences of well-intended actions, such as braking too hard on a slippery road. The occurrence of lapses while driving is closely related to the information-processing capabilities of the individual. Attempts have been

52. Misjudge the distance between oncoming vehicles when turning left and narrowly miss a collision (M)
50. Ignore a yield sign and almost collide with traffic having the right-of-way (OV)
56. Miss a stop sign and narrowly avoid colliding with traffic having the right-of-way[a]
79. Fail to notice pedestrians crossing when turning into a side-street from a main road (S/L)

Figure 2.4: A sample of the errors and violations measured by the DBQ used by [13].

made to modernize the DBQ to better account for the ageing population, aggressive driving and inattention caused by electronic devices such as the smartphone or navigational devices [13]. When adding a focus on aggressive driving, violations are divided into ordinary violations (i.e. running a red traffic light) and aggressive violations, which have an emotional component. Examples of aggressive violations are honking at other drivers when irritated and driving at unsafe following distances (tailgating) to maintain high speeds.

2.6. CONCLUSION

Both Romoser [14] and Peeters [1] provide interesting conceptual models and designs for tutoring systems that can adapt their content to the user's input. There is however a lack of research on building such a system and validating the outcome. This project aims to fill that gap by implementing a multi-agent system with a feedback loop that influences the material the user interacts with. The research experiment will explore the effect this has on the user experience.

REFERENCES

[1] M. M. Peeters, K. v. d. Bosch, J.-J. C. Meyer, and M. A. Neerincx, *Agent-based personalisation and user modeling for personalised educational games,* in *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*, UMAP '16 (ACM, New York, NY, USA, 2016) pp. 303–304.

[2] H. C. Lee, V. Drake, and D. Cameron, *Identification of appropriate assessment criteria to measure older adults' driving performance in simulated driving,* Australian Occupational Therapy Journal **49**, 138 (2002).

[3] M. A. Mets, E. Kuipers, L. M. de Senerpont Domis, M. Leenders, B. Olivier, and J. C. Verster, *Effects of alcohol on highway driving in the stisim driving simulator,* Human Psychopharmacology: Clinical and Experimental **26**, 434 (2011).

[4] S. Moten, F. Celiberti, M. Grottoli, A. van der Heide, and Y. Lemmens, *X-in-the-loop advanced driving simulation platform for the design, development, testing and validation of adas,* in *2018 IEEE Intelligent Vehicles Symposium (IV)* (IEEE, 2018) pp. 1–6.

[5] R. Rossi, M. Gastaldi, F. Biondi, F. Orsini, G. De Cet, and C. Mulatti, *A driving simulator study exploring the effect of different mental models on adas system effectiveness,*

in *International Conference on Augmented Reality, Virtual Reality and Computer Graphics* (Springer, 2020) pp. 102–113.

[6] B. Kappé, W. Van Winsum, and P. van Wolffelaar, *A cost-effective driving simulator,* in *Driving Simulator Conference, Paris* (2002).

[7] G. N. Yannakakis and J. Togelius, *Experience-driven procedural content generation,* IEEE Transactions on Affective Computing **2**, 147 (2011).

[8] G. Bhatti, *User-centered multi-layer programming approach to model scenarios on driving simulators*, Ph.D. thesis, Université de Toulouse, Université Toulouse III-Paul Sabatier (2014).

[9] I. Padayachee, *Intelligent tutoring systems: Architecture and characteristics,* in *Proceedings of the 32nd Annual SACLA Conference* (2002) pp. 1–8.

[10] J. A. Michon, *A critical view of driver behavior models: What do we know, what should we do,* Human behavior and traffic safety , 485 (1985).

[11] E. A. Fleishman, *Toward a taxonomy of human performance.* American Psychologist **30**, 1127 (1975).

[12] J. Reason, *Human error* (Cambridge university press, 1990).

[13] S. T. Cordazzo, C. T. Scialfa, and R. J. Ross, *Modernization of the driver behaviour questionnaire,* Accident Analysis & Prevention **87**, 83 (2016).

[14] M. Romoser, *An autonomous intelligent driving simulation tutor for driver training and remediation: A concept paper,* in *Driving Assessment 2011: 6th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle DesignHonda R&D Americas, IncorporatedNissan Technical Center, North America-Toyota Collaborative Safety Research CenterFederal Motor Carrier Safety AdministrationUniversity of Iowa, Iowa City* (2011).

2

3
DESIGN

You tell me exactly what you want, and I will very carefully explain to you why it cannot be.

Simeon Yetarian, GTA V

In this chapter, we lay out how we model the user and the scenarios. We start by giving an overview in section 3.1 of how the user's performance influences the selection of the scenario. Then in section 3.2 we go into the structure of a driving scenario. Section 3.3 describes all the driving skills and knowledge that will be learned and tested in the application. This is followed by section 3.4 about the ways in which we can give the user feedback on their driving during a scenario.

3.1. ADAPTIVE DIFFICULTY

Figure 3.1 shows can see the application's process flow from one scenario to the other. The design of the flow diagram is based on the flow diagram by Peeters[1], simplified to fit the scope of this project. The user model and knowledge base are used to calculate the expected difficulty a scenario has for the user. A scenario of suitable difficulty is then selected. The idea being that if the difficulty level is appropriate, the user is engaged and motivated while not being overwhelmed. The "script" of the scenario is used to apply changes to the environment and instruct the NPCs (the other traffic users). During the playing of the scenario, the user's performance is tracked and feedback is given when necessary. Additional feedback and instructions can be given at the end of the scenario. The user model is updated based on his performance during the scenario. This updated model is then used to select the following scenario. The scenarios are short so there is constant updating of the user model to determine what lesson we want to present next. The processes shown in figure 3.1 are performed by several agents. These agents will be discussed in-depth in chapter 4. In table 3.1 you can see the different parameters that have an influence on the difficulty of a scenario. Each option for the parameter is given a difficulty score. The Driving Instructor's Handbook [2] was used to determine the difficulty level of elements of a scenario.

3

Parameter	Options	Difficulty
Visibility	No Obstacle	⋆
	Obstacle	⋆ ⋆ ⋆
Maximum speed	30 km/h	⋆
	60 km/h	⋆ ⋆ ⋆
	Crossing road has different speed.	⋆ ⋆ ⋆ ⋆ ⋆
Road layout	Straight Road	⋆
	T-junction	⋆ ⋆ ⋆
	Crossroads	⋆ ⋆ ⋆ ⋆
Weather	Clear	⋆
	Overcast	⋆ ⋆
	Light Rain	⋆ ⋆
	Heavy Rain	⋆ ⋆ ⋆
	Thunderstorm	⋆ ⋆ ⋆ ⋆ ⋆
Light	Dawn	⋆
	Noon	⋆
	Dusk	⋆ ⋆ ⋆
	Night	⋆ ⋆ ⋆ ⋆ ⋆
NPCs:	Pedestrians	Differs case by case
	Cyclists	Differs case by case
	Cars	Differs case by case

Table 3.1: Influence of parameters on difficulty.

3.2. SCENARIO DESIGN

In this section, we deconstruct the elements of a driving lesson scenario. The work of Richard et al.[3] gives many examples of task analyses of driving scenarios. We used similar techniques to break down into tasks and subtasks a number of driving scenarios, which were defined in a brainstorm session with a consultant from the company Alten.

3.2.1. SCENARIO FLOW

At the start of the scenario, the car is standing still so that the user can get a feel for the controls. When the user reaches a certain speed or is a certain distance from the start position the evaluation starts. This means the user always goes through the Use Case 00 (table 3.2), but will not be graded on it. That use case is then followed by a scenario such as described in Use Case 01 (table 3.3) and Use Case 02 (table 3.4). The tasks/steps that will not be assessed are indicated in cursive font and in brackets in the use case.

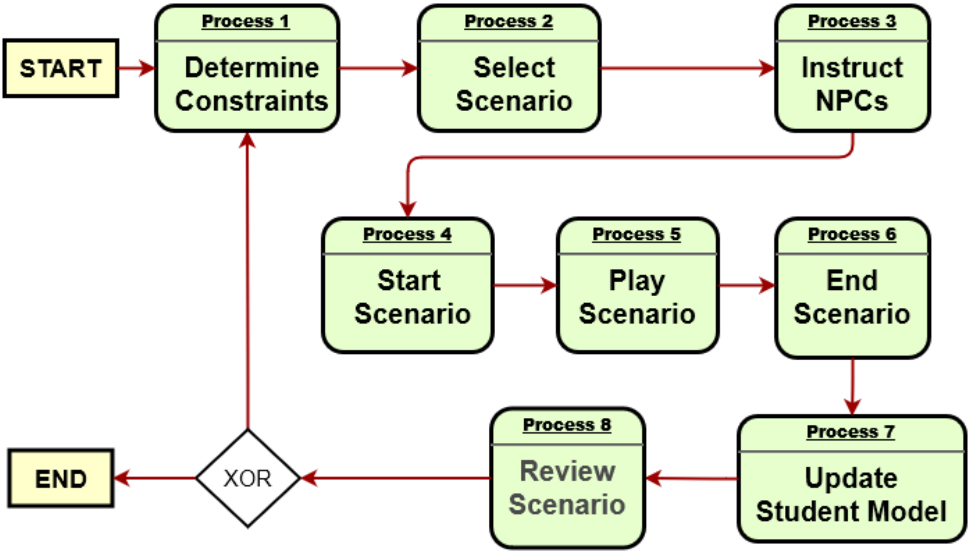


Figure 3.1: Flow diagram to automatically select scenarios.

The user has to assess the situation and respond in the correct manner. This requires knowledge of the traffic rules and good control of the car. In table 3.5 you can find a list of skills and knowledge that are trained and assessed when playing the scenario and the associated codes for them that we use in the use case. A further explanation of these skills and knowledge can be found in the sections 3.3.1 and 3.3.2.

Use Case Name:	UC00
Description:	Approach Intersection.
Actors:	Student Driver
Triggers:	Lesson starts.
Preconditions:	Vehicle is at start point.
Basic Course of Events:	0.1 Step on Gas. 0.2 Increase speed. 0.3 Follow road at proper speed.
Postconditions:	Vehicle is proper speed or a certain distance from start point.
Skills and Knowledge:	S01, S04B, S04C.

Table 3.2: Use Case of scenario 00.

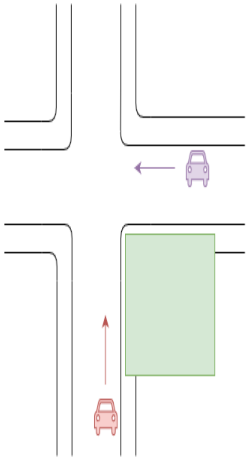
Diagram	Use Case
	Name: UC01 Descr: Turn Right at Crossroads with Obstructed Sight. Act: Student Driver Trigg: Vehicle is proper speed or a certain distance from start point. Driver sees obstruction blocking view to the right. Pre: Vehicle is moving in the direction of the intersection. Events: 1.1 [*Look in Rear View Mirror*]. 1.2 Signal to the right. 1.3 Step on the brake slowly decreasing speed. 2.1 [*Check for traffic on the opposing and left and right road*]. 2.2 Give way to the car coming from the right road. 3.1 [*Check for traffic on the opposing and left and right road*]. 3.2 Step on the gas slowly increasing speed. 3.3 Steer right following the curve of the corner you turn. 3.4 Increase speed till you reach an appropriate speed for the road you are on. Post: Vehicle is at scenario end point. S & K: S01, S02, S03B, S04A, S04B, S04C, S04D, K02A. Difficulty7 out of 10

Table 3.3: Use Case for Scenario 01

3.2.2. LEARNING GOALS

The goal of a typical driving lesson is to learn new skills, consolidate partially learned skills and assess skills already learned or partially learned. The overlapping learning goal for the prototype will be learning when and how to 'give way' to other road users. This goal can be divided into subgoals, i.e. how to act when a stop sign is present. In figure 3.3 and tables 3.4, A.1, A.2, A.3 and A.4 you can see use cases for several scenarios of varying difficulty containing a stop sign. The goals were selected with consideration for the available environments found within the game.

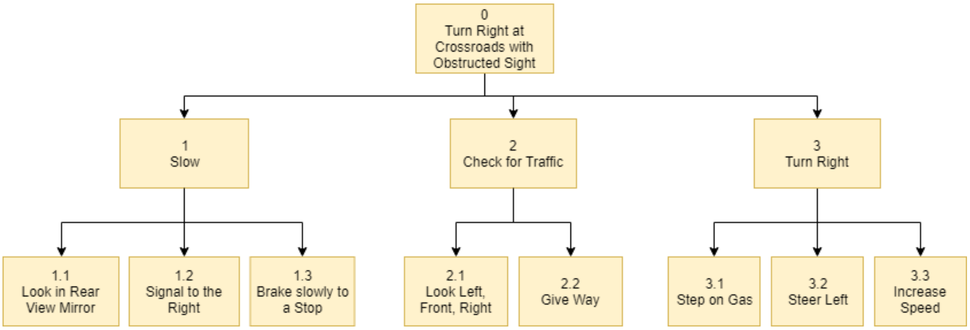


Figure 3.2: Task Hierarchy of scenario 01.

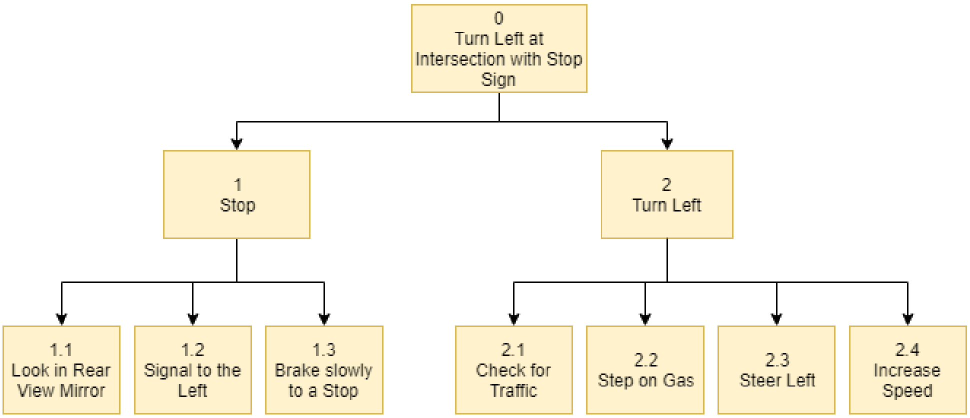


Figure 3.3: Task Hierarchy of scenario 02.

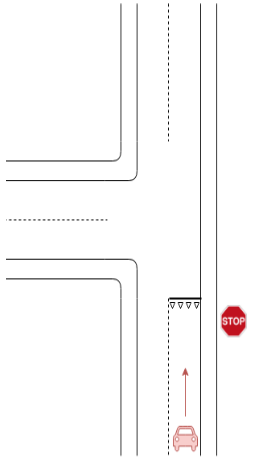
Diagram	Use Case	
	Name:	UC02
	Descr:	Turn left on an intersection with a stop sign.
	Act:	Student Driver
	Trigg:	Vehicle is proper speed or a certain distance from start point.
	Pre:	Vehicle is moving in the direction of the intersection.
	Events:	1.1 [*Look in Rear View Mirror*]. 1.2 Signal to the left. 1.3 Step on the brake slowly decreasing speed to a stop. 2.1 [*Check for traffic on the opposing and left road*]. 2.2 Step on the gas slowly increasing speed. 2.3 Steer left following the curve of the corner you turn. 2.4 Increase speed till you reach an appropriate speed for the road you are on.
	Post:	Vehicle is at scenario end point.
	S & K:	S01, S02, S03A, S04A, S04B, S04C, S04D, K03A.
	Difficulty	5 out of 10

Table 3.4: Use Case for Scenario 02

3.3. SKILLS AND KNOWLEDGE

This section contains descriptions of the skills and knowledge that the application aims to teach. The skills and knowledge were selected from a list of driving skills [4]. They were considered to be the most essential driving skills whose evaluation could be implemented within the scope of the project.

3.3.1. SKILLS: HANDLING THE CAR

LANE KEEPING (S01)

Lane-keeping refers to the user's ability to keep the car positioned correctly in a road lane when driving on straight or slightly curved road segments. It requires the user to have a feel for the environment and the car's position in it. By making micro-adjustments to the steering wheel the user has to constantly keep the car in position. This is a skill that will start out requiring deliberate action from the user but will come automatically once the user has gotten a feel for it.

How to measure it? To obtain the path we want the user to follow we have an AI agent drive from the starting point of the scenario to the end point without being obstructed by anything. We record the position points that the AI agent travels and use those to determine how much the user deviates from these points dotted along the center of a driving lane.

SIGNALLING (S02)

Signalling is the act of activating your car's signal lights to indicate in which direction you intend to turn an upcoming corner. This is one of the first things you learn when starting driving lessons as it is something that is used very often. Similar to lane keeping, this is a skill that will start out requiring deliberate action from the user but will come automatically once the user has gotten a feel for it.

How to measure it? The scoring will be based on a few points. One is, whether you signal at all (Score of 0% when you fail to signal). It's easy to forget for beginners. Another is, whether you press the correct direction (Score of 50% when you signal the wrong direction). Also, the timing of the signalling is important. We can measure the distance to the intersection and determine whether the signalling button was pressed in a timely fashion (Score from 75% to 100% depending on your timing).

TURNING CORNERS (S03)

Turning corners is the act of using the steering wheel to make the car turn at a corner. This is also one of the first things you learn when starting driving lessons. Corners can be taken sharply, where the car traces the sidewalk, and they can be taken widely.

How to measure it? Turning corners is measured similarly to lane keeping. A certain path is expected from the user and the deviation from the path is measured.

SPEED KEEPING (S04)

Speed keeping measures how well the user controls the gas and the brakes. This means accelerating, decelerating, or keeping a stable speed in a proper way, depending on the situation.

How to measure it? For each step of the scenario, we know the acceptable speed. We can check whether the user's speed is within this range. Also, we can calculate the stability in maintaining a constant speed for straight sections of road.

DISTANCE KEEPING (S05)

Distance Keeping refers to whether the user keeps a proper distance from the car in front of him. This distance is dependent on the speed of the car in front and on whether the weather conditions reduce the view. When there is a reduced view due to fog or rain, the user has to keep a larger distance from the car in front. Also if the car in front shows irregular behaviour such as swerving, the user should keep a safe distance.

How to measure it? Given the speed of the car in front and the distance to the car, we can determine whether the user adjusts his speed accordingly and keeps a safe distance to the car in front. Extra adjustments are calculated based on the quality of view and the behaviour of the car in front.

ALERTNESS (S06)

User's ability to notice and respond swiftly to unexpected events. For example, a pedestrian crossing the road where he's not supposed to or a car breaking the traffic rules.

How to measure it? The user is presented with a scenario that tests this skill. A score will be determined by how well and timely the user responds. The expected response is determined by the drive tasks of the scenario.

3

JUDGEMENT (S07)

In traffic, situations can arise that require a call of judgment to be made by the user. For example, if the user approaches a traffic light that turns orange, does the user step on the brakes or increase his speed? Other examples are finding a good opening to insert the car when entering a crowded road or determining whether a pedestrian by the side of the road is intending to cross.

How to measure it? The user is presented with a specific scenario. The decision the user makes is evaluated and feedback is given. In the case of judgement calls sometimes a decision isn't necessarily right or wrong, but you can often comment on the safety of the decision taken.

3.3.2. KNOWLEDGE: LEARNING THE TRAFFIC RULES

Aside from a list of skills and the state of proficiency the user model also keeps track of a set of knowledge. Initially, we assume the user does not have the knowledge. This assessment is not entirely accurate as it is likely that the user has experience as a traffic participant even if not in a car. The user also likely has experience riding in a car as a passenger. Each knowledge rule is scored by way of a boolean, which is true when we believe the user to have the knowledge and false when we believe the user does not have the knowledge. In table 3.5 you can see a summary of the knowledge that has been implemented in the prototype.

In the Netherlands, the rules of conduct when participating in traffic are defined by the Ministry of Infrastructure and Water Management (I&W)[1].

3.3.3. SKILL TREE: GROWING IN EXPERIENCE

To avoid overwhelming the user with too many learning goals at once we make use of a skill tree that unlocks new skills and knowledge to learn as the user plays the scenarios. The use of a skill tree is a common technique used in games to pace the progress of a user's growth. The skill tree for the skills and knowledge as defined in table 3.5 is shown in figure 3.4. As new skills and knowledge are added to the implementation the skill tree can expand as needed. For skills we consider them mastered when the skill in the user model is above a certain threshold value. For knowledge, it is considered known when a scenario related to the specific knowledge is completed successfully. Success in this case meaning the completion of the driving tasks related to the scenario. New nodes in the tree are unlocked when the skills and knowledge in all the parent nodes are deemed to be mastered. The scenarios presented to the user are selected from a pool of scenarios of which all skills and knowledge required have been unlocked. This way the user is gradually introduced to new concepts within the driving domain.

3.3.4. USER PERFORMANCE SCORES

At the end of each scenario, the user is shown the score he/she obtained. Tips are given on the skills and/or tasks that were not performed to satisfaction. The aggregated score

[1] https://www.government.nl/ministries/ministry-of-infrastructure-and-water-management

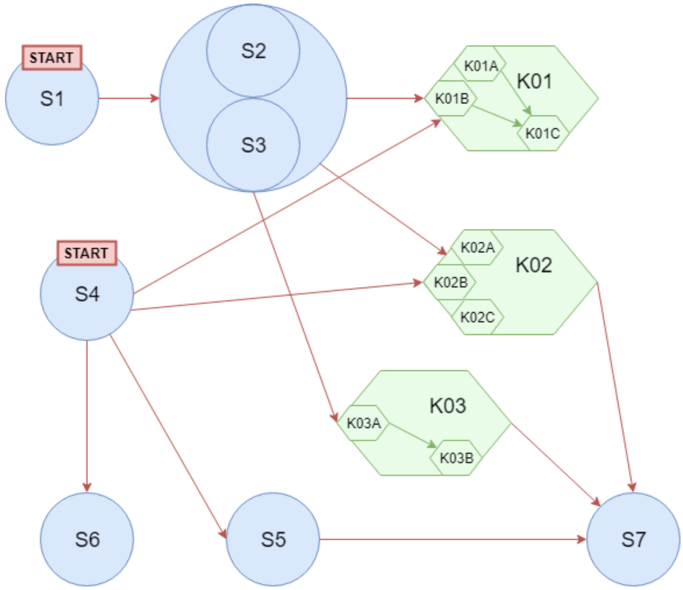


Figure 3.4: Skill tree showing the order in which the user unlocks skills.

as calculated over several played scenarios will be shown to the user at the end of the gameplay session. For skills such as S03 and S04 that are subdivided into subskills a calculated average of the subskills will be shown.

3.4. FEEDBACK

In this section, we discuss the different ways we can give feedback on the actions of the user while driving the car. The purpose of the feedback is to be informative, letting the user know what to do or what should have been done.

AUDIO FEEDBACK

Ideally for the application we would like the user to receive spoken instructions while driving in the car. For the prototype, which will be implemented, this will not be possible. The textual feedback during driving will be shown as subtitles. If these prove to be too distracting, the textual instructions will be restricted to be shown only before the start of a scenario and at the end of a scenario, when evaluating the user's performance.

VISUAL FEEDBACK

In the User Interface we can use visuals to both give instructions to the user as well as give feedback on the user's actions. For example, we can show the direction the user should turn at the next intersection by drawing an arrow on the screen. Visual feedback should not be distracting or obstruct the view of the environment.

EXAMPLE OF FEEDBACK

Given the scenario as described in Use Case 01 (table 3.3). What is the feedback that can be given for each step of the event sequence?

REFERENCES

[1] M. M. Peeters, K. v. d. Bosch, J.-J. C. Meyer, and M. A. Neerincx, *Agent-based personalisation and user modeling for personalised educational games,* in *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*, UMAP '16 (ACM, New York, NY, USA, 2016) pp. 303–304.

[2] J. Miller and M. Stacey, *The Driving Instructor's Handbook* (Kogan Page, Limited, 2005).

[3] C. M. Richard, J. L. Campbell, J. L. Brown, *et al.*, *Task analysis of intersection driving scenarios: Information processing bottlenecks*, Tech. Rep. (Turner-Fairbank Highway Research Center, 2006).

[4] A. J. McKnight and B. B. Adams, *Driver education task analysis. volume ii: Task analysis methods. final report.* (1970).

Skill Code	Skill Description
S01	Lane Keeping.
S02	Signalling.
S03	Turning Corners.
S03A	Turn Left.
S03B	Turn Right.
S04	Speed Awareness.
S04A	Stay below max speed limit.
S04B	Accelerating & Decelerating (Should not be too abrupt unless its an emergency brake).
S04C	Don't drive too slow (lower than 20 km/h below the max speed).
S04D	Speed Keeping (stability).
S05	Distance Keeping (2-second rule).
S06	Alertness.
S07	Judgment.
S07A	Merging. Make sure there is sufficient space for your vehicle.
Knowledge Code	**Knowledge Description**
K01	The rules for pedestrian crossings.
K01A	Pedestrians on a pedestrian crossing always have right of way. Always stop for pedestrians when they wish to cross the roads.
K01B	You must wait until the pedestrian has completely cleared the crossing before proceeding.
K01C	Slow down when approaching a pedestrian crossing when there are pedestrians approaching or near the crossing.
K02	The rules for crossings without priority roads.
K02A	Traffic coming from the right has priority. Slow down and stop to give way to traffic coming from the right.
K02B	Traffic users going straight ahead on the same road have right of way.
K02C	Turning a right corner has priority over vehicles on the same road turning a left corner.
K03	The rules for stop signs.
K03A	Always stop at a stop sign even when there is no traffic.
K03B	When stopping at a stop sign, give way to any vehicle coming and going to the left or right.

Table 3.5: Skill Codes and Knowledge Codes.

Feedback UC01 Step 1.1 Look in Rear View Mirror.	
Feedback Type	Description
Audio Feedback:	Instructions.
Visual Feedback:	An arrow points to the rearview mirror.
Haptic Feedback:	None
Feedback UC01 Step 1.2 Step on the brake slowly decreasing speed to a stop.	
Feedback type	Description
Audio Feedback:	Instructions.
Visual Feedback:	An indicator on the screen shows the distance to the stopping point showing when to stop completely.
Haptic Feedback:	None
Feedback UC01 Step 2.1 Check for traffic on the opposing and left road.	
Feedback Type	Description
Audio Feedback:	Instructions.
Visual Feedback:	Depending on the level of instruction you can show an icon (i.e. an eyeball) to indicate the user should check for traffic.
Haptic Feedback:	None
Feedback UC01 Step 2.2 Step on the gas slowly increasing speed.	
Common Mistake	Accelerating too fast.
Feedback Type	Description
Audio Feedback:	Instructions.
Visual Feedback:	During the move an icon on screen indicates whether you are accelerating too fast or too slow.
Haptic Feedback:	None
Feedback UC01 Step 2.3 Steer left following the curve of the corner you turn.	
Common Mistake	Corner is taken too sharp or too wide.
Feedback Type	Description
Audio Feedback:	After the move you are notified if it was too sharp, too wide or just right.
Visual Feedback:	During the move an icon on screen indicates whether you are turning too fast or too slow.
Haptic Feedback:	Depending on the control scheme, the user feels the wheel turn to adjust.
Feedback UC01 Step 2.4 Increase Speed till you near maximum.	
Common Mistake	Speeding or driving too slow.
Feedback Type	Description
Audio Feedback:	If you cross the speed limit you hear a notification sound.
Visual Feedback:	The speed of your car is shown on the screen. If you cross the speed limit you see a visual notification.
Haptic Feedback:	None.

Table 3.6: Feedback of scenario 01.

4

IMPLEMENTATION

You think we can't do that? We can. It's in our guidelines.

Michelle, GTA V

In this chapter, we'll go into the implementation of the virtual driving simulation application built using GTA V as a base environment. We start with a section (4.1) on the tools used in implementing the application. This includes the changes that were necessary to make development in GTA V possible. In section 4.2 we discuss the interface. Section 4.3 gives an overview of how the application was implemented to consist of several modules working together. Finally in section 6.2 we discuss all the limitations we came across while implementing the application.

4.1. TOOLS

The application is being developed for PC as a mod for GTA V. Github is used for storing code. Trello is used for project management.

GRAND THEFT AUTO V

The basis of the implementation will be the game Grand Theft Auto V (hereafter GTA V). GTA V is a game developed by Rockstar. It has been reported that GTA V is the most profitable entertainment product of all time as the main game continues to sell well, as well as the success of purchasable online content[1]. It was developed using their proprietary RAGE game engine. While GTA V is not open source, modding in the offline singleplayer game is tolerated by the developer as long as there is no profit being made from the mods. All modding functionalities have been authored by third-party developers. GTA V offers a realistic simulated environment. With scripting, you can call functions in

[1] https://www.gamesindustry.biz/articles/2018-04-09-gta-v-is-the-most-profitable-entertainment-product-of-all-time

the game engine, which gives a lot of freedom to make changes to the environment, the interface, the controls, character behaviour etc.

MODS
At its core GTA V is an action/RPG, not a driving simulator. As such many adaptations are needed to create a suitable environment for uninterrupted learning. The changes and solutions are summarized as follows. Support for steering wheel controllers: Install mod [2]. Realistic physics and car handling: Install mod [3]. Prevent game events: Load the game from a 100% completed save game. Prevent violence from player and non-player characters: Use Open IV [4] to change settings. Allow custom C# scripts: Use Community Script Hook V .Net [5] to access functions in the game engine.

SET-UP
For development, a high-end PC with a powerful graphics card (NVIDIA GeForce GTX 1070) will be used. For the controls, a steering wheel with pedals will be used. The wheel has buttons that can be used for signalling and navigating menus. There will not be a manual gearbox. There will not be any hardware to monitor the user's state.

4.2. INTERFACE

MUSIC AND SOUNDS
Sounds are the standard sounds used in GTA V.

CONTROLS
The controller used for the evaluation will be a Thrustmaster T150 Force Feedback. Any controller with DirectInput compatibility will be usable with the application. In figure 4.1 you can see the controls that are available to the user. For the prototype learning how to use gears is not within scope, therefore those controls will not be used.

CAMERA
One channel (screen) is used for the display of the application. The application starts and stays in a first-person view. The player character will not be visible. The user is unable to change the view to look around.

The first-person camera as implemented in the game does not give a wide enough view of the surroundings for proper driving, therefore a custom camera was created and attached to the vehicle in approximately the position of the eye-height driver of the vehicle as can be seen in figure 4.2. The car model and the position of the camera were carefully selected to give a realistic and open view of the environment.

Given the restricted visual information that is available in the surrounding environment, a wider-than-usual field of view is used for the camera. This is necessary for the user to accurately assess the traffic situation.

[2] https://www.gta5-mods.com/scripts/manual-transmission-ikt
[3] https://nl.gta5-mods.com/vehicles/realistic-driving-v
[4] http://openiv.com/
[5] https://github.com/crosire/scripthookvdotnet

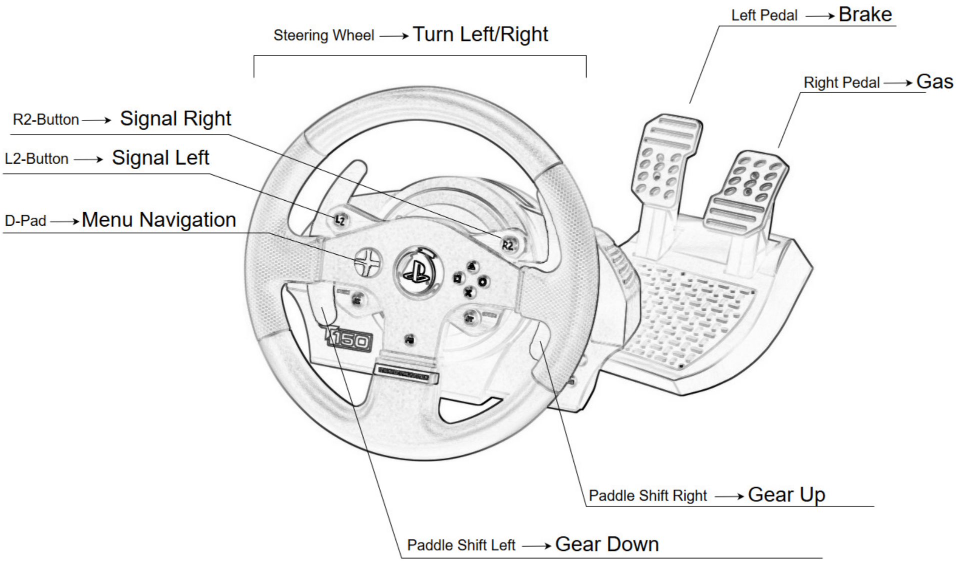


Figure 4.1: Steering wheel controls.

Figure 4.2: View when driving in the application.

Figure 4.3: Editor view for creating location and scenario data.

UI

The application as well as the documentation and implementation (code and comments) will be in English.

For finding locations and recording their data as well as for creating scenarios a separate top-down camera view is used as can be seen in figure 4.3. In this view, the camera is controlled with the Numpad. For getting position data from the map the center of the screen is used. This is due to the engine lacking a generic 'screen_to_world'-function. The locations are selected and recorded in such a way that each location can be used for several scenarios.

In this scenario-editor view, for each location, we record all possible starting positions, end positions and the locations of elements (such as pedestrian crossings, junctions etc.). Non-player characters, such as other vehicles, can be placed with a trigger added to define when their movement starts. In the interface are also options to link scenarios to specific Skills and Knowledge.

4.3. THE MODULE STRUCTURE

In figure 4.4 you can see the module structure diagram of the application. It shows an overview of all the modules in the application, the coloured blocks, and the interfaces that they have among themselves, the arrows. Each module has its own function to perform. The game module which contains the game engine, the core of the whole system, renders game scenes and executes game logic. The user agent 4.3.1 contains the data we have on the user (i.e. his performance). The knowledge base contains all the skills and knowledge we wish to teach the user. The scenario creator agent 4.3.2 takes the user model and knowledge base and selects a new scenario for the user to play through. The director agent 4.3.3 takes the "script" of the scenario selected by the scenario creator and

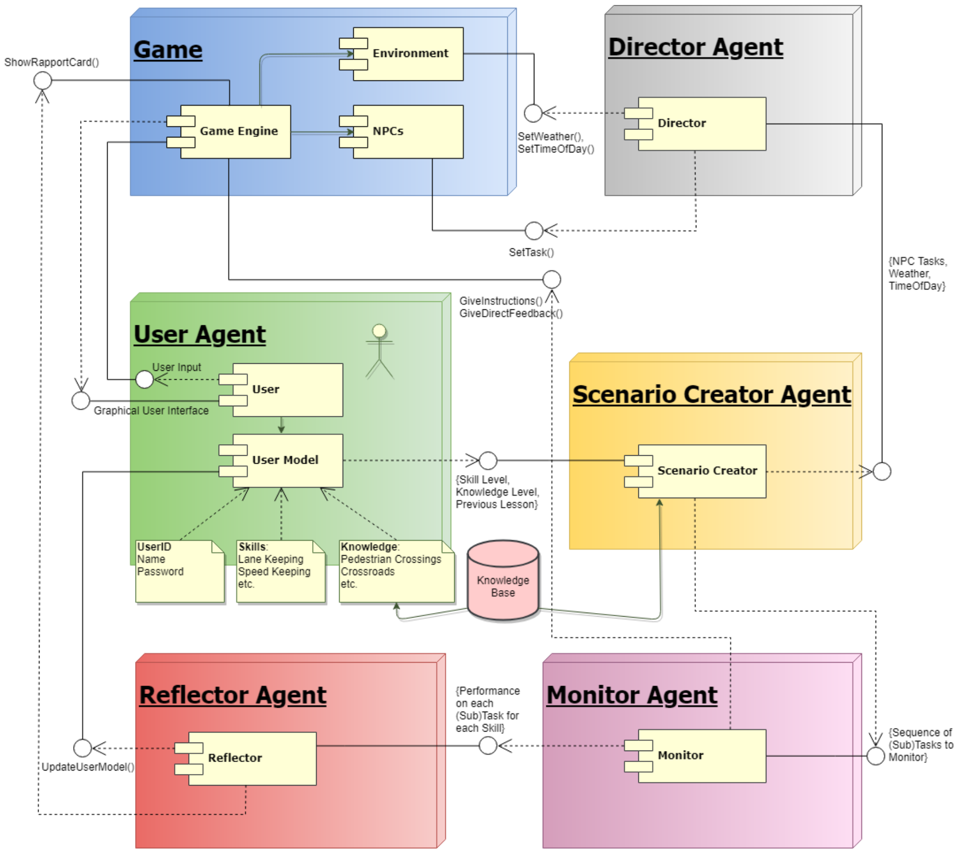


Figure 4.4: Module Structure Diagram.

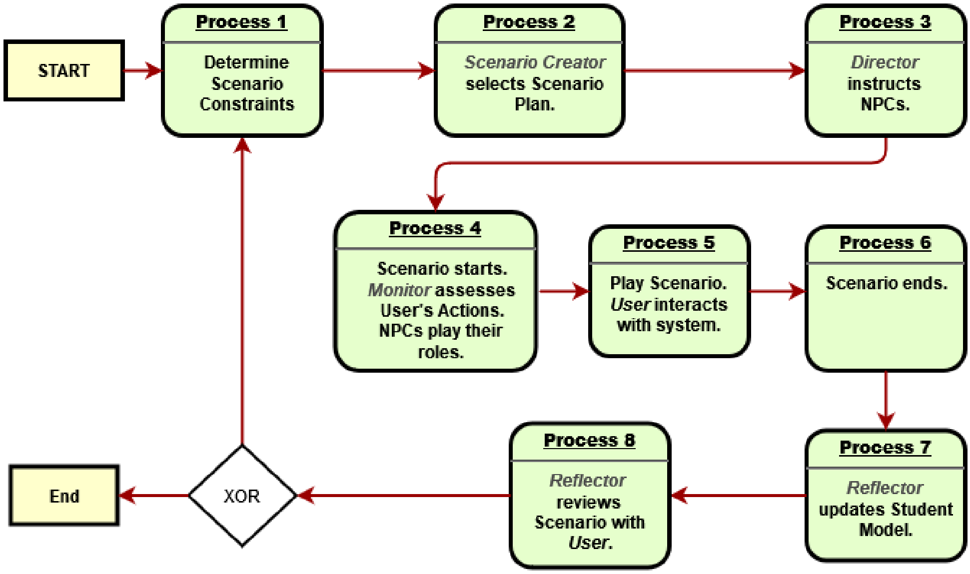


Figure 4.5: Flow diagram showing the agent's tasks.

applies the changes to the environment and instructs the NPCs, the other traffic users. During the playing of the scenario the monitor agent 4.3.4 keeps track of the user's performance and gives feedback when necessary. The reflector agent 4.3.5 gives the user a performance summary at the end of a scenario and updates the user model based on this. Each agent will be described in-depth in this chapter. Figure 4.5 shows an overview of the application flow and the tasks of each agent.

4.3.1. The User Agent

The user is modelled as an intelligent agent. He is defined by the input he gives to the system. For our application, we want to be able to evaluate the user's skill and knowledge based on his input. In the previous chapter, section 3.3 we defined the skills and knowledge that will be tested in the application.

4.3.2. The Scenario Creator Agent

The scenario creator makes a selection for a new scenario for the user to play through. There are several factors that are taken into account in determining which scenario to select. Firstly, from the user model, we can learn which skills and knowledge have been unlocked by that user. This determines the pool of scenarios that are eligible for selection. Scenarios are eligible if and only if the user has unlocked the skills and knowledge that are connected to the scenario.

For each scenario that is eligible we calculate a difficulty score that depends on the user's score for the skills, *sc(s)*, and the user's knowledge values, *sck*, for the skills and knowledge that are respectively in the scenario's skill list, *ssl*, and the knowledge list, *ssk*. The formula for this is:

$$calculated\_difficulty = \frac{\sum sc(s) \in ssl + \sum sc(k) \in ssk}{ssl.count + ssk.count}$$

This score is then combined by a weighted average with the base difficulty score that the specific scenario has. This is calculated as follows:

$$difficulty\_score = 0.9 * calculated\_difficulty + 0.1 * base\_difficulty.$$

This base difficulty score is manually determined beforehand mainly by looking at the amount of skills/knowledge needed for the scenario and the amount of NPCs and drive tasks.

The higher the difficulty score the more difficult a scenario is expected to be for the user. The difficulty most suited to the user should be one that is not too difficult and not too easy. The target difficulty for the new scenario is around 50%.

Another factor to be taken into account is the similarity to the previous scenario. Scenarios are considered more similar the more there is an overlap in skills and knowledge connected to the scenario. The similarity between two scenarios is calculated as follows, with SK referring to the skills and knowledge needed for the scenario and d referring to the difficulty level of the scenario:

$$similaritySK = \frac{SK_A \cap SK_B}{SK_A + SK_B - SK_A \cap SK_B}$$

$$similarityD = 1 - \frac{abs(D_A - D_B)}{9}$$

$$notSameID = 1 - (ID_A == ID_B)$$

$$similarity = similaritySK * (0.5) + similarityD * (0.4) + notSameID * (0.1)$$

If the user has performed well in the previous scenario, the next scenario should offer a new challenge. On the other hand, if the user has performed poorly in the previous scenario, a similar scenario can be selected for the user to practice those skills some more. A scenario is considered to be performed well if all skills have a score above 70% and all knowledge is considered known.

To avoid repeating the same scenario to users, a list is kept of the scenarios that have been played previously. Preference is given to scenarios that haven't been played yet or haven't been played recently.

Weather conditions and time of day can be changed to increase the difficulty of a certain scenario. These functions are however not used at the moment.

USE CASE TESTS

To test the formula described above a number of different use cases were generated to evaluate the results of the implementation. In figure 4.6 a representation is given of the difficulty curve that is expected for a user completing a session of 10 predefined scenarios increasing somewhat linearly in difficulty. The scenarios that are included in the selection pool for the adaptive difficulty session are labelled to match the number of the

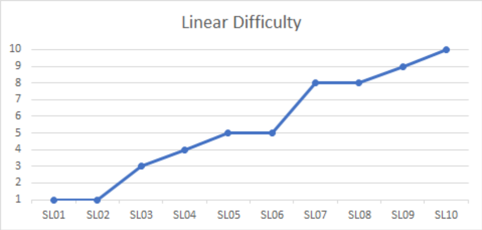


Figure 4.6: Difficulty scores for scenario sequence for different user profiles (adaptive).

scenario that is similar in the linear session. For example, variations on the scenario from the linear session SL05 would be labelled: SA05V01, SA05V02. The difficulty scores represented by the blue line are predefined difficulty values determined by the instructor. Scenarios that contain a lot of other traffic users and/or a variety of driving tasks are considered to be of high difficulty and scenarios without other traffic users and that do not require much knowledge of traffic rules are considered to be of low difficulty.

Six different use cases were created to see how the algorithm recommends different scenarios depending on the player's performance. The results can be seen in figure 4.7. As in the previous figure, the blue line represents the predefined difficulty values for the scenarios. The orange line represents the difficulty score that is calculated by the algorithm, meaning it is dependent on the user's skills and knowledge at that time.

The 'Bad' player profile consistently has a failing score for the skills and knowledge required for the scenarios. This results in the algorithm continuing to recommend the same 3 scenarios. Because no skill is mastered the player never proceeds in the skill tree and no new scenarios are unlocked, which forces the player to continue with the simple scenarios until the basic skill of lane keeping is performed proficiently.

The 'Good' player profile consistently has a good, passing score (7/10) for each scenario. This results in much more variety in the scenarios that are recommended.

4.3.3. THE DIRECTOR AGENT

The director distributes tasks to the NPCs according to the scenario received by the scenario creator. The NPCs in the prototype can be of type pedestrian or car. Each NPC has a start position and rotation where he is spawned at the start of the scenario.

The possible tasks for NPCs are walk, run, drive and wait. Each task has an end position, indicating the location an NPC should be at after completion of the task. Each task also has a trigger to determine at what timing the task should be performed. While playing the scenario the director keeps track of these triggers and checks whether a new task should be performed by the NPC.

An NPC has a variable called 'repeating'. This is a boolean used for specific scenarios where there is a continuing stream of NPCs performing the same task. Only one NPC needs to be defined and copies are made as the scenario plays out.

The director agent keeps a list of the entities generated for a specific scenario. When

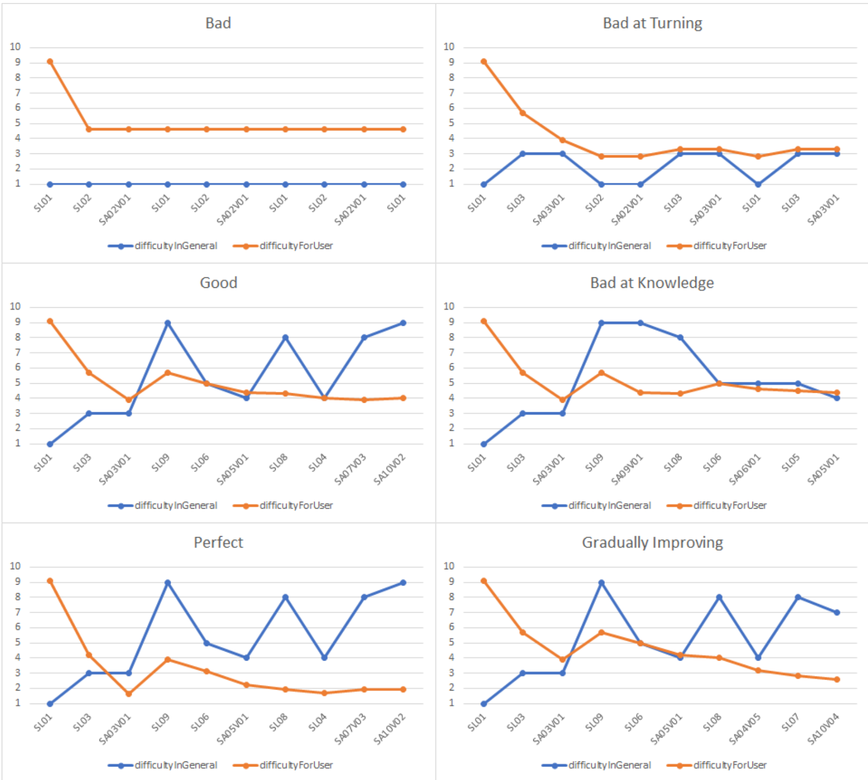


Figure 4.7: Difficulty scores for scenario sequence for different user profiles (adaptive).

the scenario ends, these entities can be removed from the game space.

The director agent also contains methods for changing the weather and time of day.

4.3.4. The Monitor Agent

In section 3.3 the skills and methods for measuring them are described. The monitor agent actually contains the functions to make these measurements. Also, it keeps track of when it is relevant to measure a certain skill. For example, you shouldn't drive too slowly (S04C), however, if there is a car driving in front of you who is driving slowly, tracking the distance to the car in front is the appropriate measure to make.

The monitor has full knowledge of the environment and the scenario being played. At each moment during the user's playing of the scenario, the monitor is aware of the intended behaviour and the driving behaviour of the user. The intended behaviour is modelled as a list of driving tasks. These driving tasks are bound to certain sections of the route the user is driving on.

Driving with the correct speed	Driving over the max limit
	

Table 4.1: In-Game Questionnaire

4.3.5. The Reflector Agent

At the end of the scenario, the reflector collects and aggregates the data on the performance of the user from the monitor. The user is shown a scorecard showing his scores on the skills. The user is also shown the mistakes he makes and the rules he failed to follow if any. The score for the scenario is then used to update the user model. The values of the scores for the skills are calculated as a weighted moving average.

After calculating the scores for the skills. These scores and the boolean values for the knowledge are used to check whether the user has unlocked new skills/knowledge or not.

Figure 4.8: Score card showing feedback and the user's score for the skills relevant to the scenario.

4

4.4. CONCLUSION

Ideally, the envisioned scenario flow would be an ongoing lesson where the scenarios are generated and executed dynamically as the user drives through the environment. Given the constraints we have concerning implementation time, we have decided to generate small single-issue scenarios. While this can disrupt the flow of driving you would get in a longer scenario, it gives the opportunity to display feedback at the end of a scenario and the user time to reflect on the situation and his actions.

5

Evaluation

[being hit by another driver] I don't like this car!

Michael De Santa, GTA V

In this chapter we will first discuss the hypotheses, section 5.1, that guide the design of the research experiment. Followed by section 5.2 on the participants and section 5.3 on the set-up and materials necessary to the experiment. In section 5.4 on design, we thoroughly explain the relevant variables and the procedure of the experiment. The statistical results are then summarized in section 5.5. The chapter finishes with section 5.6, where we discuss the results and the limitations of the research experiment.

5.1. Hypotheses

For this project, an application was developed which applies the teaching strategy of adapting lessons based on the user's skill level. In the application, the user plays short driving scenarios (lessons) and is graded on skills and knowledge.

Research Question

The research experiment was conducted to find an answer to the following research question:

> *What is the impact of adaptive difficulty on the user's skill and knowledge gain and their self-efficacy when learning in a virtual driving simulation?*

To answer the research question a comparison needs to be made between the user experience of the method of adaptive difficulty and the user experience of a more traditional method of strictly defined linear difficulty that is commonly used in digital learning applications.

HYPOTHESIS 1

Previous research has shown that significantly higher learning outcomes were achieved by using adaptive difficulty in an application for studying Spanish [1]. We believe this will be the case as well for training driving tasks in a virtual simulation. An example closer to our domain is the work of Georgiou et al[2] on procedurally generated race tracks. This research however focuses on the validity of the difficulty level generated by the system not on its effect on skill improvement.

> H1: *The participants playing the adaptive session will acquire better skills and more knowledge compared to those playing the linear session.*

The participants playing the adaptive session are expected to learn faster than those playing the linear session and perform better on the skill evaluations as they will be presented with scenarios that better match their educational needs.

HYPOTHESIS 2

The content is attuned to present the user with the difficulty best suited to them. This is expected to improve the user experience. We are particularly interested in self-efficacy. Research by Bhatti et al. [3] found a positive correlation between self-efficacy and system effectiveness for virtual training systems. Self-efficacy is a predictor of motivation and performance regardless of domain [4].

> H2: *Training with scenarios of adaptive difficulty gives a higher self-efficacy for the user compared to training with scenarios of linear difficulty.*

In the adaptive mode, the presented scenarios should be within the participant's capabilities. With the linear scenarios, the difficulty level may rise faster than the skills of the participant can handle, which would lower the self-efficacy of the participant.

5.2. PARTICIPANTS

The experiment was conducted within the office building of the technical consulting company Alten[1]. The participants were the employees who were working in the office. The number of participants was to be as many as could be recruited to participate in the time frame of two weeks. In the end, 30 people were able to participate in the experiment. No compensation or incentives were offered for participation. Around 45 minutes were spent with each participant.

The employees participating in the experiment have a software engineering background and therefore can be assumed to have a high level of proficiency in using computers. Driving experience is high, with 28 out of 30 having a driving license, half of whom have more than 5 years of driving experience. The age range of the participants is between 20 and 40, 25 male, and 5 female. Out of the 30 participants, 9 have played a virtual driving simulator in the past.

[1] https://www.alten.nl/

Figure 5.1: Hardware setup for research experiment.

5.3. SET-UP AND MATERIALS

To run the application, a desktop PC is required with the following installed on it: *Grand Theft Auto V*, the game which provides the environment the application runs on, *Open IV*, which enables the altering of the game's data files and *Script Hook V*, which enables calls to functions in the game engine. The following game mods are required: Realistic Driving V and Steering Wheel Support. Furthermore, the application that has been developed needs to be in the game folder in order to run.

Figure 5.1 shows the hardware set-up used for the research experiment.

For the questionnaires, digital surveys have been created that can be loaded into any internet browser. The survey responses are recorded in separate files on the PC.

5.4. STUDY DESIGN

In this section, a description is given of what was measured in the research experiment (5.4.1) and the procedure that was used to obtain these measurements (5.4.3).

Before the research experiment started, a pilot study was conducted with two participants to validate the procedure and to check if the data collection worked as desired. In a free-form interview, the test subjects were asked for feedback on the test session.

In this research, the conditions concern the mode of difficulty scaling, either linear or adaptive (MO). For the linear scenarios, the difficulty rises steadily with each subsequent scenario. For the adaptive scenarios, the difficulty of the scenarios can rise and fall depending on the user's performance. This experiment has a within-subjects design with two equally sized groups of participants, one starting with a linear session and the other with an adaptive session.

5.4.1. MEASURES

The measures are the variables that are expected to be affected by changing the conditions.

SKILLS & KNOWLEDGE

The skills and knowledge that are measured in the application are described in table 3.5. To summarize, the skills are lane keeping, signalling, turning corners, speed keeping, distance keeping, alertness, and judgement, and knowledge refers to several traffic rules. After finishing each scenario, the system gives the user a score (S) of 0-100% for each skill that was relevant to the scenario that was played. Knowledge (K) is represented as a boolean, 'true' for when a certain rule is assumed known and 'false' otherwise. A knowledge rule is assumed known if the user completes the drive tasks. Both the skill scores and the knowledge values are recorded after each scenario as data to track the growth of the user.

The scores for the four exam scenarios that are played after completing the first session were used to measure the participant's acquired skills and knowledge.

SELF-EFFICACY

How does the method of learning affect the user's confidence and motivation to continue playing? To measure the motivation the user has when going through the scenarios, the user is asked to self-evaluate their confidence in learning at the end of playing each scenario. The question being asked is: *How would you rate your confidence after playing this scenario?* The user gives a score of 1 to 10. A higher score indicates that the user has more confidence in their ability to perform driving tasks.

In the post-session questionnaire, there are statements related to self-efficacy (SE). The participant rated their agreement on a scale of 1 to 5, where 1 is 'very much disagree' and 5 is 'very much agree'.

5.4.2. EXPLORATORY MEASURES

The exploratory measures are variables measured during the research experiment that do not directly relate to the research question. By examining the variables, we intend to gain a deeper insight into the user experience of the application.

MENTAL EFFORT

The aim of having the difficulty of the scenarios be adaptive is to lead to a mental effort of the user that lies in a sweet spot between boredom and being overwhelmed. One very important variable for this is the difficulty of the scenario. An algorithm in the application determines the difficulty the scenario is expected to have for the user. This is explained in section 4.3.2. The calculated difficulty (CD) will be recorded for each scenario the user plays. The perceived difficulty (PD) the scenario has for a user is also recorded by way of having the user self-evaluate the difficulty level at the end of playing each scenario. The question being asked is: *How would you rate the difficulty of this scenario?*

The user gives each scenario played a score of 1 to 10 (with 10 being very difficult). By comparing the calculated difficulty, the perceived difficulty, and the user's performance of the scenario, it can be shown how well the adaptive method presents the user with suitable lesson material.

In the questionnaire the user filled in after the gameplay session there are some statements related to mental effort.

UNDERSTANDABILITY
Understandability is a variable chosen to evaluate how well the information presented is interpreted by the user. This includes the terms used for evaluation, the interface, the instructions, and the feedback. There is no difference between the linear and the adaptive group in the presentation of this information to the user. Therefore it is defined as a separate value from the somewhat related variable of Mental Effort. The results from measuring the understandability can indicate areas where the application can be improved. The post-session questionnaire has statements related to understandability. The participant has rated their agreement on a scale of 1 to 5, where 1 is 'very much disagree' and 5 is 'very much agree'.

REALISM
While the setup of the application isn't quite the same as driving a car, a certain level of realism is favorable. The realism has to be sufficient enough for the user to believe that the lessons learned in the simulation are transferable to real live traffic situations. Regarding this variable, there is no difference between the linear and adaptive session. While not directly related to the main research question it is a useful variable to help evaluate the application to give hints on future improvements. The post-session question contains several statements on realism. The participant rated their agreement on a scale of 1 to 5, where 1 is 'very much disagree' and 5 is 'very much agree'.

FUN
Using a game environment for learning should be appealing to users. Many users are familiar with the setting and have enjoyable memories of the game that's being used as a platform. The fun (FN) variable measures how much the user enjoys following lessons in the application. Together with the variable self-efficacy, it measures the user's satisfaction. The post-session questionnaire has statements related to enjoyment. The participant rated their agreement on a scale of 1 to 5, where 1 is 'very much disagree' and 5 is 'very much agree'.

5.4.3. PROCEDURE
To make sure the test sessions proceeded smoothly a moderator checklist was used. The full checklist can be found in the appendix (B.5). In the appendix, you can also find an overview of the steps of the procedure (B.6). The experiment was run one participant at a time.

INSTRUCTIONS
A general introduction to the application was given without divulging the purpose of the research. The test subject was instructed on the steps of the procedure. The in-game

controls for the vehicle were explained.

Pre-Session Questionnaire

The pre-session questionnaire contained general questions on the user's experience with video games and driving. Former experience may have an effect on the user's performance in following the lessons. The questionnaire questions can be found in appendix B.2

Gameplay Session

The group of people participating in the research have been randomly divided into two groups of equal size. Each participant went through two sessions of the application. One session playing scenarios with a linear difficulty curve and the other playing scenarios with an adaptive difficulty curve. Which of the two was played first was determined randomly with an equal amount of test subjects starting with one as with the other. The participants did not know which session they started with or even what the difference between the sessions was. The examiner did know which group the participant was in.

The linear session consisted of ten scenarios all played in order. As such, each participant played the exact same scenarios. For the adaptive session, the content was different for each participant as scenarios were selected to match the perceived skill and knowledge of the user. Only the first scenario was the same for each participant. The adaptive session also consisted of 10 scenarios. This set-up of having each test subject play two different sessions was chosen to maximize the amount of data collected per test subject. Also, it allowed the test subjects to compare the different teaching methods and show their preference.

The participant took a seat holding the steering wheel and placing their feet on the pedals. The participant was then presented with a number of driving scenarios. Instructions were shown to the participant indicating which direction to take when driving the scenario. The specific sequence of tasks that the participant performed was dependent on the scenario that was presented to the participant. Scenarios differed from each other in elements such as other road users and street signs. Chapter 3 contains a breakdown of several scenarios and the drive tasks to perform in the scenario. Aside from performing the defined drive tasks, the participant was required to adhere to certain traffic rules, such as keeping the car in the centre of the lane and staying below the maximum speed limit. These driving skills were measured throughout the scenario where appropriate.

At the completion of each driving scenario, the participant was presented with an evaluation of their performance. This consisted of a scorecard showing scores for all the skills relevant to the scenario and a text box showing in text the mistakes that were made. A simple pass or fail message, calculated by combining the performance on the skills and knowledge, gave the participant a clear indication of overall performance. The participant received a motivational message depending on their performance: i.e. "Good Job!", "You can do it!". In an in-game questionnaire, the participant was asked to rate the difficulty of the scenarios and their confidence in their performance.

In-game after each scenario, the participant was asked to rate the difficulty of the scenario and their confidence in their performance (self-efficacy). Using a single question after the completion of a task is a method commonly used in measuring usability. It is simple to understand for the user and quick to execute, giving results that are easy to process by researchers and that closely correlate to more extensive usability queries [5].

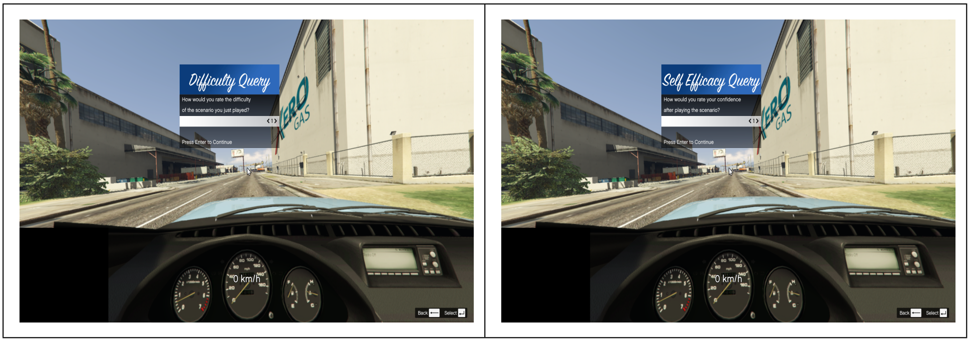


Table 5.1: In-Game Questionnaire (left for difficulty, right for self-efficacy)

Exam Scenarios

After the first session, the user was presented with 4 exam scenarios. These exam scenarios were the same for both the participants starting with a linear session and those starting with an adaptive session. Comparing the exam results between the two groups should give a fair assessment of which group shows the greatest improvement in skills on average.

Post-Session Questionnaire

The post-session questionnaire contains statements pertaining to the user's experience in playing the session. The participants answered on a 5-point Likert scale ranging from 'Strongly Disagree' to 'Strongly Agree'. The statements alternate between positive and negative to balance the response. Conflicting responses may indicate that the participant did not properly read/understand the statements. When evaluating the results, the measurements need to be inverted for the negative statements. The questionnaire statements can be found in appendix B.4. Some statements were repeated between the first and second sessions. Other statements, for example, those comparing the two sessions, were only asked after both sessions were completed.

5.4.4. Data Preparation and Statistical Analysis

The exam scores, consisting of multiple skills and knowledge scores for 4 different scenarios, were consolidated into one exam score per participant (EXS). Each relevant skill, S01, S02 and S03, was averaged over the scenarios to obtain the averaged skill scores (AS01, AS02 and AS03). The score for knowledge (K) is the percentage of knowledge scored as True. These 4 values, AS01, AS02, AS03 and K, were averaged to come to the final exam score (EXS). To compare the exam scores of the linear group and the adaptive group (each has 15 participants) an independent two-sample t-test was done on the calculated exam score means (EXS).

The self-efficacy score given by the participant for each scenario was combined into one averaged score for the linear scenarios and one averaged score for the adaptive scenarios. This means that for each of the 30 participants, we have a score for self-efficacy for linear scenarios and a score for self-efficacy for adaptive scenarios (ASE). To compare the self-efficacy for the linear scenarios and the adaptive scenarios we did an independent two-sample t-test on the averaged self-efficacy means (ASE), a total of 30 samples. The null hypothesis being that there is no significant difference between the two groups.

For both the exam scores and the self-efficacy scores, a check for normalcy was done by plotting the data in order to determine whether the data was eligible for analysis using a t-test. For the self-efficacy scores, an ANOVA (Analysis of Variance) test was done to analyse whether the interaction of learning mode and session order has a significant effect on the outcome.

The dataset containing the output of the research experiment and the code used to obtain the statistical results can be found at 4TU.ResearchData[2]. The statistical analysis was done using R (Version 4.2.2).

5.5. RESULTS

The purpose of this research experiment was to explore the difference in user experience between linear progression through driving tasks and adaptive progression. In this section, we describe the results of the statistical analysis of the data resulting from the experiment. First, we show the results of the statistical tests. This is followed by a discussion of the results and the limitations of the research experiment.

5.5.1. EXAM SCORES

To start the analysis, an F-test is first done to determine whether the variances between the groups are equal or not. Based on the obtained results ($F = 1.14$, $df = 14$, $p = 0.81$), it can be assumed that there is no statistically significant difference in the variances. After that, a two-sample t-test is conducted to compare the exam scores of participants following the linear session ($n = 15$) with those of participants following the adaptive session ($n = 15$). No significant difference was found in exam scores between linear ($M = 0.75$, $SD = 0.076$) and adaptive ($M = 0.71$, $SD = 0.081$), $t(28) = -1.29$, $p = .21$, 95% $CI = [-0.0954, 0.0216]$. The boxplot in Figure 5.2 shows an overview of the data.

5.5.2. SELF-EFFICACY

Similarly to the analysis of the exam scores, an F-test is performed first to check the equality of the variances between the groups. The obtained results indicate there is no statistically significant difference in variance ($F = 1.12$, $df = 29$, $p = 0.76$). A two-sample t-test was then performed to compare the self-efficacy of participants following the linear session ($n = 30$) with those of participants following the adaptive session ($n = 30$). The analysis found a significant difference in self-efficacy between linear ($M = 4.67$, $SD = 0.94$) and adaptive ($M = 5.27$, $SD = 1$). The t-test conducted indicated significant results ($t(58) = 2.4$, $p = .02$, 95% $CI = [0.099, 1.1]$), and the Cohen's d indicating effect size was calculated as 0.62, which is considered to be a medium-sized effect [6]. The boxplot in

[2] https://data.4tu.nl/private_datasets/VdJgsoYCnqV2bZRzr4tJLHhZ40J7VFX0jL6lscuH1DI

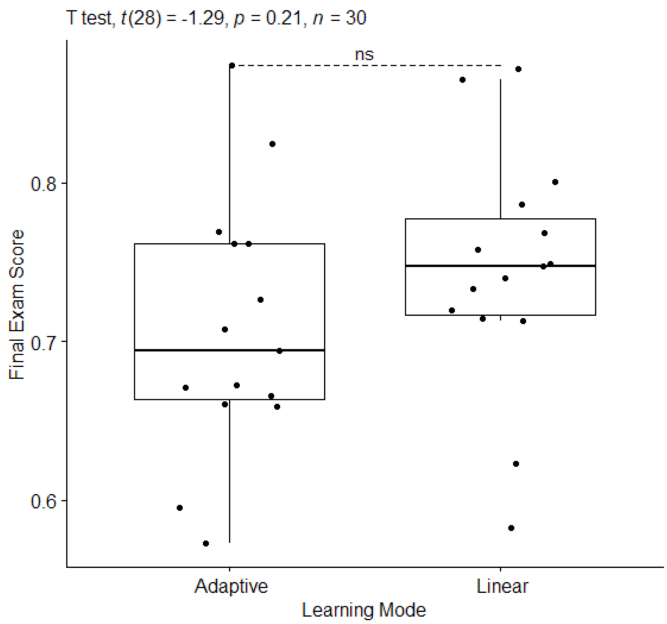


Figure 5.2: Summary of final exam score data.

Figure 5.3 shows an overview of the data. The barchart in Figure 5.4 shows a summary of the data separating the results of the first and second session. A two-way ANOVA was performed to analyze the effect of learning mode and session order on self-efficacy. The analysis showed no statistically significant interaction between the effects of the learning mode and session order ($F(1, 1) = 0.712$, $p = 0.4025$).

5.5.3. EXPLORATORY MEASURES

The boxplots in Figure 5.5 show an overview of responses to understandability-related statements. Each statement refers to a different element of the application: task understanding (Q8), instructions at the start of a scenario (Q9), score card at the end of a scenario (Q10), evaluation at the end of the session (Q11) and skills being evaluated (Q12). The results show a wide variety of responses with a mediocre average understanding on all areas.

The boxplots in Figure 5.6 show an overview of responses to realism-related statements. Each statement refers to a different element of the application: controls/car handling (Q15), environment (Q16) and car/pedestrian behaviour (Q17). The results show a wide variety of responses with a mediocre average realism on all areas.

In comparing the questionnaire response, shown in table 5.3, no significant difference was found for fun, mental effort, and self-efficacy.

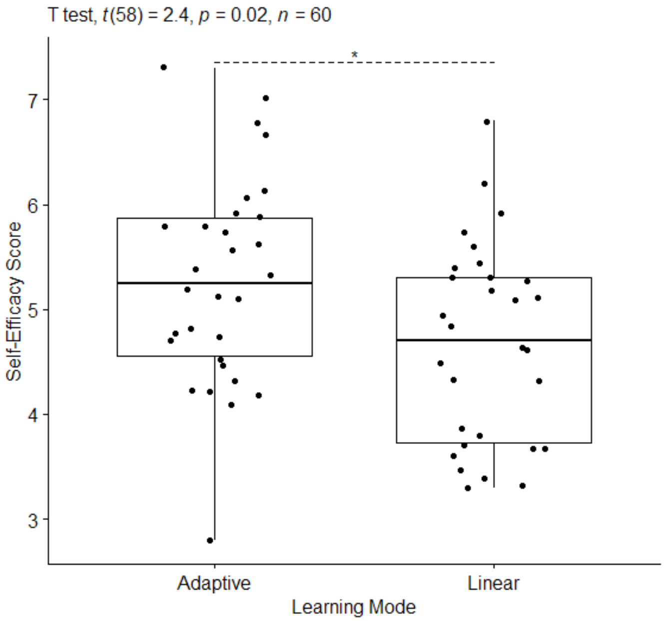


Figure 5.3: Summary of self-efficacy data.

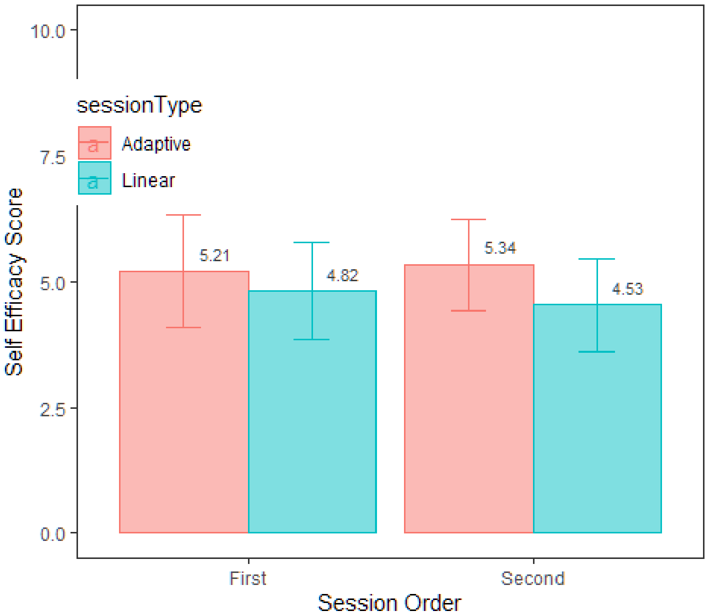


Figure 5.4: Barchart summary of self-efficacy data.

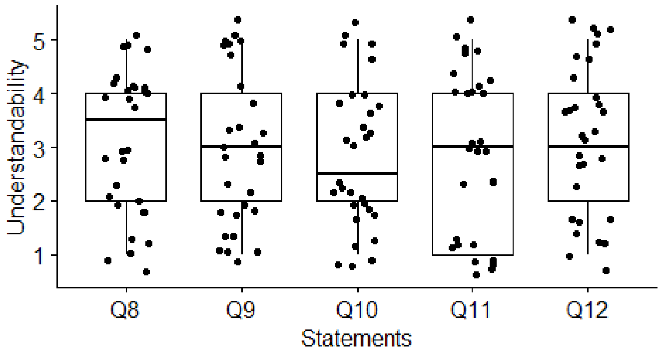


Figure 5.5: Summary of understandability data.

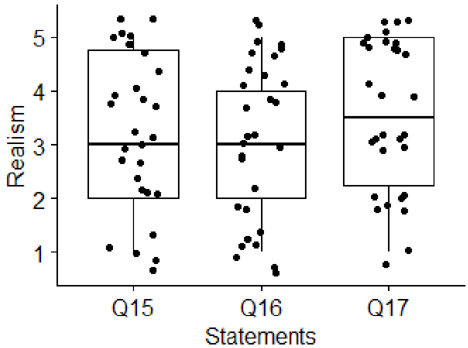


Figure 5.6: Summary of realism data.

Variable	Mode	N	M	SD
Fun	Linear	30	2.83	1.49
	Adaptive	30	2.97	1.45
Mental Effort	Linear	30	2.48	0.669
	Adaptive	30	2.35	0.803
Self-efficacy	Linear	30	3.02	1.17
	Adaptive	30	2.75	1.05

Table 5.2: Group Statistics for Questionnaire Response

Variable	t	df	p
Fun	0.35	58	0.73
Mental Effort	-0.70	58	0.49
Self-efficacy	-0.93	58	0.36

Table 5.3: Results of t-test for Questionnaire Response

5.6. DISCUSSION

In doing this research experiment, we were looking to investigate whether there was a difference in skill and knowledge acquisition between the two groups, linear and adaptive. The hypothesis was that an adaptive progression of learning scenarios would lead to greater skill and knowledge for the user as the learning material was tailored to the user's progression in skills and knowledge. To measure this we looked at the scores of the exam scenarios all participants performed after the first session. The results of the experiment however showed no significant difference in the exam scores. The difference that is observed can be caused by chance. Therefore, the hypothesis (H1) that the adaptive group shows the greatest improvement in skills and knowledge has not been proven.

An explanation for the lack of significant difference could lie in the way the research experiment was conducted. Playing ten short scenarios for about ten minutes may not be enough learning time to compare skills and knowledge acquisition. Section 5.6.1 talks more about the limitations of the research experiment that may have led to the lack of a significant difference.

It could be the case that the reasoning behind the hypothesis is flawed to start with. We theorized that participants playing the adaptive session would acquire skills and knowledge faster as the algorithm would always present scenarios that challenged them. However, the flip side of that is that if a participant struggles with a scenario, they will be presented with different, but similar scenarios in order for them to acquire the skills needed to progress. It could be that the participants that are fast and those that are slow will even out the results, which means that, on average, the skill and knowledge acquired will be the same for an adaptive session and a linear session.

The data analysis comparing the self-efficacy scores from the in-game questionnaire showed a significant difference in means with a value of $p < 0.05$. Participants playing the adaptive session indicated having a self-efficacy that was on average higher than those playing the linear session with a medium effect size. Given these results, we can confidently say that the observed difference in means is not by chance. This is in line with our hypothesis (H2). The participants who were slower in acquiring skills and knowledge would get overwhelmed by the later, more difficult scenarios when doing the linear session. In the adaptive session, this is much less likely to happen due to the use of the skill tree described in section 3.3.3.

The responses on the questionnaires show that for the application in general the perception of realism was average on a scale of non-realistic to realistic, as seen in Figure 5.6. To what extent this would influence skill transfer to real-world application is something that needs more research to be determined. The statements about understandability similarly show results to be average on a scale of difficult to understand to easy to understand, as seen in Figure 5.5. This refers to both the understanding of the tasks that were given and the results of the evaluation. This indicates that improvements need to be made to the interface design.

5.6.1. LIMITATIONS

In the time that was available a variety of scenarios was developed. The pool of scenarios from which the adaptive session could choose was limited. The result being that a session of adaptive scenarios might end up quite similar to a linear session. In fact, the response to the questionnaire shows that about half of the participants didn't notice a difference in the difficulty curve between the two sessions. To overcome this problem more scenarios would need to be available. Also, a longer session with more scenarios would more likely result in a more pronounced difference between an adaptive session and a linear session.

Due to unforeseen circumstances, the group of participants who participated in the experiment deviated somewhat from the target group of the application. The main important difference is that all participants had experience driving, many quite extensive. For a more accurate test of skills and knowledge acquired in a learning system, it would be better to have participants who are less familiar with the subject matter.

There was no budget for the research experiment, which limits what you are able to ask of people. The application was luckily quite attractive so even though there was no compensation, there were still people willing to spend about 45 minutes playing the scenarios and answering the questionnaires. There was only one set-up of the system with the application installed and a steering wheel and pedals for control. So only one participant could play at a time. This, combined with the location the experiment was conducted, led to a lower amount of participants than desired for an experiment such as this.

REFERENCES

[1] S. Sampayo-Vargas, C. J. Cope, Z. He, and G. J. Byrne, *The effectiveness of adaptive difficulty adjustments on students' motivation and learning in an educational computer game,* Computers & Education **69**, 452 (2013).

[2] T. Georgiou and Y. Demiris, *Adaptive user modelling in car racing games using behavioural and physiological data,* User Modeling and User-Adapted Interaction **27**, 267 (2017).

[3] D. Jia, A. Bhatti, and S. Nahavandi, *The impact of self-efficacy and perceived system efficacy on effectiveness of virtual training systems,* Behaviour & Information Technology **33**, 16 (2014).

[4] D. H. Schunk, *Self-efficacy, motivation, and performance,* Journal of applied sport psychology **7**, 112 (1995).

[5] J. Sauro and J. S. Dumas, *Comparison of three one-question, post-task usability questionnaires,* in *Proceedings of the SIGCHI conference on human factors in computing systems* (ACM, 2009) pp. 1599–1608.

[6] J. Cohen, *Statistical power analysis for the behavioral sciences* (Academic press, 2013).

6
Conclusion

You sure accomplished a lot today, Trevor.

Ron Jakowski, GTA V

In this chapter we first in section 6.1 reflect on and try to answer the research questions posed at the start of this project. Then we look at possible future work to improve the prototype in section 6.3. Final remarks can be found in section 6.4.

6.1. Discussion

This project was developed at the technology consultancy company Alten from an initial concept of building a virtual driving simulator in the game environment of GTA V. The idea of the project was that different students could work on different areas of study and expand upon the implementation. After brainstorming possible research angles we decided on exploring the effects of adaptive difficulty in scenario-based training. As the first project to build on this concept the application needed to be built from the start.

Our first research question in this process was:

> *What methods and techniques can be utilized to implement a virtual driving simulator that has adaptive difficulty?*

In our research we came upon the work of Peeters[1], which offers a useful model for a multi-agent system where a 'scenario planner' determines the content of a scenario, the 'director' controls the different elements during the scenario, the 'monitor' measures the performance and the 'reflector' shows an evaluation of performance after finishing a scenario. In our implementation, we used a similar design and workflow. The research also resulted in more insight into how drivers and drivers' tasks can be modelled.

We used this insight to answer the next research question:

> *In what way can the user and learning content be modelled to enable personalized learning?*

For this project, we focused on the user's acquired skills and knowledge as the main elements of the user model. Scenarios of varying difficulties were designed to teach these skills and knowledge. A skill tree is used to give the user a logical progression through the learning material. The design is made such that it can be expanded to add more content to the user model as well as the knowledge model.

> *What are the challenges involved in implementing a virtual driving simulator with adaptive difficulty in the GTA V game environment?*

When implementing the prototype we encountered several difficulties. It took quite a while to find a suitable car and camera angle that gave a good view of the centre console of the car and the environment. Building the scenarios required the creation of a scenario editor tool where scenarios could be edited and saved. Even though the scope was toned down to focus on a limited set of skills and knowledge, the implementation took a lot of time due to the number of elements that needed to be created and tested to make a functional application. There is still a lot of room for improvement in the implementation.

> *What is the impact of adaptive difficulty on the user's skill and knowledge gain and their self-efficacy when learning in a virtual driving simulation?*

In our research experiment, we compared user experiences playing adaptive and linear. For many of the measurements, the statistical results from the analysis were inconclusive. Part of this is due to the limited nature of the research experiment, limited time and participants. More than that though it seems the application needs more work in order to differentiate more between adaptive and linear. We did however find significant, positive results on the effect of adaptive learning on self-efficacy. Participants playing the adaptive session indicated having a self-efficacy that was on average higher than those playing the linear session with a medium effect size.

6.2. LIMITATIONS IN USING GTA V

The following limitations can only be mitigated by using a different environment to implement the application.

Not all skills necessary for driving can be trained and assessed by a system built in GTA V. One such skill is good viewing behaviour, i.e. looking in the mirrors at the right moment. The rear-view mirrors in the game do not function, probably due to performance considerations.

There are restrictions to the driving scenarios that can be created in the GTA V environment. Certain road layouts, such as roundabouts, are not available. Traffic lights are present, but can't be controlled or queried on status (green, orange or red).

One of the most egregious downsides of using GTA V for development is that you need an online connection to Rockstar's servers to start the game. This means that you are reliant on the stability of their servers and it may happen that the connection is bad. Also when the game has an update, the modding tools don't work anymore, so you will have to wait until the tools are updated or revert your game install to a previous version. This can be done by keeping a backup of these files: update.rpf, gta5.exe, gtavlauncher.exe, PlayGTAV.exe.

6.3. FUTURE WORK

The prototype built for this project can be improved in many ways. The participants in the research experiment indicated that it was not always clear what was expected of them and how they were being evaluated. Improving the interface and feedback is a straightforward way to make the application more appealing and effective in its goal of teaching correct driving behaviour. Adding an in-game tutorial could help a lot with understandability. An in-game tutorial is a common technique used in games to explicitly explain the elements of the interface and the skills and knowledge that are required to do well.

Initially, the idea was to procedurally generate the scenarios to achieve adaptive difficulty. Mostly due to time constraints, the company Alten allotted a specific amount of time to work on the project, this idea was abandoned in favour of hand-crafting the scenarios and using a selection algorithm to present the user with a suitable scenario. The basis needed for procedural generation has been implemented. The scenarios are made up of elements that are defined as such that they could be combined, i.e. foggy weather with a busy road, to achieve a certain difficulty. In this way, the scenarios presented to the user can be more unique and personalized to their learning needs.

The game has a large city map with a large variety of road layouts. In the prototype specific locations with suitable layouts are used for the scenarios. The environment is in an American style, which is different from that in the Netherlands. Ideally, we would want to create our own maps so that we can create scenarios as we wish in order to match learning goals. While this is possible with a map editor within the GTA V environment that is available, it is limited in options and difficult to use, making it time-consuming to create environments. Creating a better map editor or improving on the existing map editor would give more options in environment layouts to generate a larger number of possible scenarios.

In the prototype the user model is fairly simple, focusing on the skills and knowledge that are acquired while playing the virtual driving simulator. It would be interesting to further explore what user traits influence the acquirement and retainment of these skills and knowledge, and how keeping track of these traits can be used to further personalize the user experience. For example, you could track the user's learning speed and take that into account when presenting the next scenario. You could keep track of how often a user trains with the application and when the last time they played was. If it was a long time ago that the user trained, you might want to present a scenario that repeats knowledge already learned for the sake of revision.

6.4. FINAL REMARKS

Very little research has been done on the effects of personalized content on the user experience in the domain of virtual driving simulation. The prototype that was implemented for this thesis project is a good starting point to improve upon. The results of the research experiments were inconclusive in some areas but did show a medium-size significant increase in self-efficacy for the session using adaptive difficulty. This validates the positive value that personalization could bring to training applications. It can be theorized that by improving the application to better adapt to the user, the user experience can be enhanced even further.

REFERENCES

[1] M. M. Peeters, K. v. d. Bosch, J.-J. C. Meyer, and M. A. Neerincx, *Agent-based personalisation and user modeling for personalised educational games,* in *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*, UMAP '16 (ACM, New York, NY, USA, 2016) pp. 303–304.

6

ALTEN

A

DESIGN USE CASES

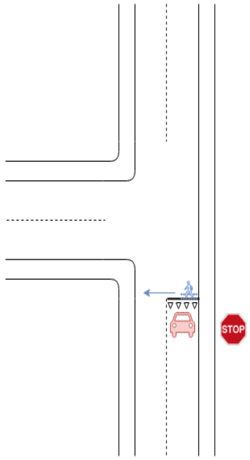
<table>
<tr><th>Diagram</th><th colspan="2">Use Case</th></tr>
<tr><td rowspan="10">
</td><td>Name:</td><td>UC03</td></tr>
<tr><td>Descr:</td><td>Go straight on an intersection with a stop sign. Pedestrian crosses road in front when the car is stopped at the stop sign.</td></tr>
<tr><td>Act:</td><td>Student Driver</td></tr>
<tr><td>Trigg:</td><td>Vehicle is proper speed or a certain distance from start point.</td></tr>
<tr><td>Pre:</td><td>Vehicle is moving in the direction of the intersection.</td></tr>
<tr><td>Events:</td><td>1.1 [Look in Rear View Mirror].
1.2 Step on the brake slowly decreasing speed to a stop.
2.1 [Check for traffic on the opposing and left road].
2.2 Give way to pedestrian crossing the road.
3.1 [Check for traffic on the opposing and left road].
3.2 Step on the gas slowly increasing speed.
3.3 Increase speed till you reach an appropriate speed for the road you are on.</td></tr>
<tr><td>Post:</td><td>Vehicle is at scenario end point.</td></tr>
<tr><td>S & K:</td><td>S01, S04A, S04B, S04C, S04D, S06, K03A.</td></tr>
<tr><td>Difficulty</td><td>8 out of 10</td></tr>
</table>

Table A.1: Use Case for Scenario 03

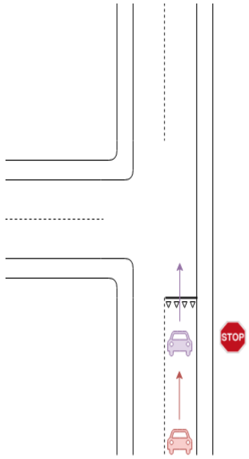
Diagram		Use Case
	Name:	UC04
	Descr:	Turn left on an intersection with a stop sign. Other car is driving in front of you.
	Act:	Student Driver
	Trigg:	Vehicle is proper speed or a certain distance from start point.
	Pre:	Vehicle is moving in the direction of the intersection.
	Events:	1.1 [*Look in Rear View Mirror*]. 1.2 Signal to the left. 1.3 Step on the brake slowly decreasing speed to a stop. 2.1 Wait for the car to drive. 2.2 Step on the gas slowly and stop at stop line. 3.1 [*Check for traffic on the opposing and left road*]. 3.2 Step on the gas slowly increasing speed. 3.3 Steer left following the curve of the corner you turn. 3.4 Increase speed till you reach an appropriate speed for the road you are on.
	Post:	Vehicle is at scenario end point.
	S & K:	S01, S02, S03A, S04A, S04B, S04C, S04D, S05, K03A.
	Difficulty:	7 out of 10

Table A.2: Use Case for Scenario 04

A

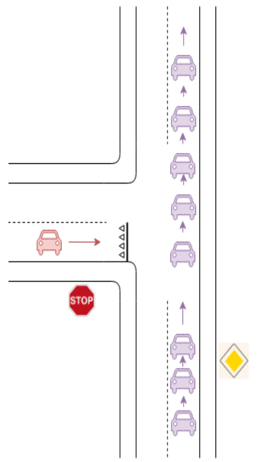
Diagram		Use Case
	Name:	UC05
	Descr:	Turn left at intersection with a stop sign. Opposite road is busy with cars in the direction the driver will turn into.
	Act:	Student Driver
	Trigg:	Vehicle is proper speed or a certain distance from start point.
	Pre:	Vehicle is moving in the direction of the intersection.
	Events:	[1.1 *Look in Rear View Mirror*]. 1.2 Signal to the left. 1.3 Step on the brake slowly decreasing speed to a stop. 2.1 [*Check for traffic on the left and right road*]. 2.2 Give way to cars crossing the road. 3.1 [*Check for traffic on the left and right road*]. 3.2 Find opportunity to join the crowded lane. 3.3 Step on the gas slowly increasing speed. 3.4 Steer left following the curve of the corner you turn. 3.5 Increase speed till you reach an appropriate speed for the road you are on.
	Post:	Vehicle is at scenario end point.
	S & K:	S01, S02, S03A, S04A, S04B, S04C, S04D, S05, S07, K03A.
	Difficulty:	10 out of 10

Table A.3: Use Case for Scenario 05

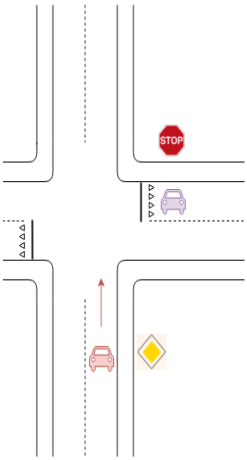
Diagram	Use Case
	Name: UC06 Descr: Go straight on an intersection an a priority road. Act: Student Driver Trigg: Vehicle is proper speed or a certain distance from start point. Pre: Vehicle is moving in the direction of the intersection. Events: 1.1 [*See the sign for priority roads*]. 1.2 [*Check for traffic on the opposing and left road*]. 1.3 Confirm the car on the right is stopping. 2.1 Maintain speed. Post: Vehicle is at scenario end point. S & K: S01, S04A,S04C,S04D, K03A. Difficulty3 out of 10

Table A.4: Use Case for Scenario 06

B

QUESTIONNAIRES

B

B.1. SUMMARY OF MEASURES

Independent Variable	
(MO) Learning Mode	Linear or Adaptive. The mode determines the method of scenario selection. For linear, with a gradually increasing difficulty, for adaptive, to match to the user's performance.
Dependent Variables	
(S01-S07) Skills	How are the skills improving over playing scenarios?
(K01-K03) Knowledge	How is the knowledge acquired playing scenarios?
(CD) Calculated Difficulty	What is the expected difficulty as calculated by the application?
(PD) Perceived Difficulty	What is the difficulty perceived and experienced?
(ME) Mental Effort	How much cognitive work is required to perform the task?
(UN) Understandability	Is the information presented clear and understandable?
(RE) Realism	Is the simulation realistic enough to be useful in learning?
(SE) Self-Efficacy	Is the user confident in the improvement of their skills?
(FN) Fun	Is the user enjoying themselves?

Table B.1: Summary of the Measures used for Evaluation.

B

B.2. PRE-SESSION QUESTIONNAIRE

1. What is your gender?

   ◯ Female ◯ Male ◯ Other

2. What is your age?

   ____

3. What is your nationality?

   ____

4. How often do you play video games?

   ◯ Daily ◯ Weekly ◯ Rarely ◯ Never

5. On which platform do you play games?

   ☐ Smartphone ☐ Console ☐ PC ☐ None of the above

6. What genre of games do you play?

   ☐ Action ☐ Racing ☐ Casual ☐ RPG ☐ None of the above

7. Have you ever played a racing game with a steering wheel?

   ◯ Yes ◯ No

8. Have you ever played a virtual driving simulator?

   ◯ Yes ◯ No

9. Have you played Grand Theft Auto V before?

   ◯ Often ◯ Sometimes ◯ Once or twice ◯ Never

10. Do you have a driver's license?

    ◯ Yes ◯ No

11. If yes, how long have you had your driver's license?

    ◯ 0-2 years ◯ 2-5 years ◯ longer than 5 years

B.3. IN-GAME QUESTIONNAIRE

1. How would you rate the difficulty of the scenario you just played?

   ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10

2. How would you rate your confidence after playing the scenario?

   ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10

B.4. POST-SESSION QUESTIONNAIRE

NR	Code	Statement
After the first Session		
Q1	ME	The lessons were easy. (Repeating)
Q2	ME	The lessons were frustrating. (Repeating)
Q3	ME	The pace of the lessons was in line with my learning speed. (Repeating)
Q4	ME	I was focused when following the lessons. (Repeating)
Q5	FN	I enjoyed following the virtual lessons in the application. (Repeating)
Q6	SE	The lessons were useful. (Repeating)
Q7	SE	I would like to continue following virtual lessons in the application. (Repeating)
Q8	UN	It was clear what was expected from me during the lessons.
Q9	UN	The instructions at the start of the lessons were clear.
Q10	UN	The score card at the end of each lesson was clear.
Q11	UN	The evaluation report at the end of the game session was clear.
Q12	UN	It is clear what skills are being tested and evaluated.
Q13	ME	It was easy to remember the controls.
Q14	ME	I had difficulties handling the car.
Q15	RE	The controls/handling of the car felt realistic.
Q16	RE	The virtual environment looks realistic.
Q17	RE	The behaviour of the other traffic users in the simulation were realistic.
Q18	SE	I felt the grades accurately reflected my performance.
Q19	SE	I have learned about traffic rules.
Q20	SE	I already knew all the traffic rules shown in the application.
Q21	SE	I have learned about car handling.
Q22	SE	I got a sense about what it's like to drive a car.
Q23	SE	I believe the virtual lessons in the application to be a good entry into driving a real car.
Q24	SE	I believe the virtual lessons in the application can prepare me for a theory driving exam.
Q25	SE	After following the virtual lessons I feel more confident to follow real driving lessons.
Q26	SE	I believe I can improve my driving skills and knowledge using this application.
After the second Session		
Q1	ME	The lessons were easy. (Repeating)
Q2	ME	The lessons were frustrating. (Repeating)
Q3	ME	The pace of the lessons was in line with my learning speed. (Repeating)
Q4	ME	I was focused when following the lessons. (Repeating)
Q5	FN	I enjoyed following the virtual lessons in the application. (Repeating)
Q6	SE	The lessons were useful. (Repeating)
Q7	SE	I would like to continue following virtual lessons in the application. (Repeating)
Q27		I could tell there was a difference in the difficulty curve between sessions.
Q28		I could tell that in one session each consecutive lesson was adapting the difficulty to match my skill level.
After explaining the difference in sessions		
Q29	FN	I enjoyed the adaptive lessons more than the linear lessons.
Q30	ME	The adaptive lessons were more challenging than the linear lessons.
Q31	ME	The adaptive lessons were less frustrating than the linear lessons.
Q32		I prefer that each consecutive lesson adapts the difficulty to match my skill level.
Q33		I would like to choose the difficulty of the lessons.

Table B.2: Post-Session Questionnaire

B

B.5. MODERATOR CHECKLIST

Before the participant arrives	
☐	Check if the machine is ready
☐	Check the position of the steering wheel, pedals, chairs and webcam
☐	Have a pen and questionnaire ready
Welcome	
☐	Introduce yourself and thank the participant for coming
☐	Offer the participant a cup of water
☐	Explain the set-up, purpose and procedure of the session
Pre-Session Questionnaire	
☐	Ask the participant if they mind being recorded
☐	Ask if the participant has any question/concerns
☐	Give the pre-questionnaire
Session	
☐	Explain the controls of the applications
☐	Brief the participant on the possibility of a crash and the procedure when a crash occurs (restart application and resume session)
☐	Emphasize that the application is being tested not the participant
☐	Explain the in-game post-scenario questionnaire
Post-Session Questionnaire	
☐	Give the Post-Session Questionnaire
☐	Thank the participant for participating

Table B.3: Moderator Checklist

B

B.6. OVERVIEW PROCEDURE RESEARCH EXPERIMENT

Step	Activity	Time (min)
1	Instructions	3
2	Pre-Session Questionnaire	2
3	Gameplay Session #1	10
4	Exam Scenarios	4
5	Post-Session Questionnaire #1	2
6	Gameplay Session #2	10
7	Post-Session Questionnaire #2	2
8	Explanation of the difference between the sessions	1
9	Post-Session Questionnaire #3	3

Table B.4: Experiment Procedure