interaction | interfaces | invention

Quantifying In-Game Task Difficulty Using Real-Time Cognitive Load

Andrew J.A. Seyderhelm
Karen L. Blackmore

Working Paper Series, Number 4, Sep 2021

ABOUT THIS SERIES

The i3 Lab is a cross-disciplinary, industry-focused research network currently based in the School of Electrical Engineering and Computing at the University of Newcastle, Australia. The principal aim of the i3 Lab Working Paper Series is to disseminate the research and/or technical outcomes of projects in an easily accessible format. The content of Working Papers generally falls into one of the following categories:

- *Preliminary findings or results, the release of which is intended to stimulate debate and/or discussion to assist in the further development of the research*
- *Technical reports associated with applied research that may be written in a less academic style than usually published in academic journals*
- *Extended versions of published works, often containing additional implementation/ application detail, figures and tables.*

The opinions or conclusions expressed in the Working Paper Series are those of the authors and do not necessarily reflect the views of the i3 Lab as a whole.

Follow @i3_lab

Email i3Lab@newcastle.edu.au
Web http://i3Lab.newcastle.edu.au

i3 Lab

Quantifying In-Game Task Difficulty Using Real-Time Cognitive Load

Andrew J.A. Seyderhelm and Karen L. Blackmore
School of Information and Physical Sciences
University of Newcastle
Callaghan, 2308, NSW, Australia

andrew.seyderhelm@uon.edu.au; karen.blackmore@newcastle.edu.eu

ABSTRACT

Serious games and simulation training is an ever growing field with adoption across a wide range of applications and industries. These games and simulations seek to create educational experiences that are engaging, challenging and enhance learning or training outcomes. Typically, these types of games have had a one-fits-all approach that may not be suitable for all players; the game may be too difficult for some or too easy for others. To tackle this problem, dynamic difficulty adjustment (DDA) is a method of adapting various game mechanics and learning content to better match the needs of the player that is common in entertainment games. However, the impact of these adjustments to the level of difficulty, from changes to the environment, mechanics or even the game aesthetic, is not straightforward, particularly in respect of cognitive load. Cognitive load refers to the mental capacity of the player and if they're either over-, or under-, burdened outcomes may be boredom, frustration, disengagement and potential failure. There are few studies exploring the effect over time of different tasks, environments and challenges on both player performance and cognitive load in complex simulated environments. In order to understand the impact of different design choices, both during game development and as applied to any form of adaptive system, it is important to understand the impact these choices are likely to have on cognitive load and player performance. This paper describes the development process and rationale for a 3D serious driving game, called the Cognitive Effect Driving Game (CEDG). The CEDG is a first person 3D game that is designed specifically to explore how different game tasks, environments, game aesthetics and challenges affect cognitive load and performance in real-time. Additionally, the CEDG will be used to validate a virtual in-game version of the detection response task (DRT), termed the Virtual DRT (VirDRT). The VirDRT is being validated as an effective measure of cognitive load, in real-time, using only the default game controller. The results of experiments such as the CEDG will enable metrics on how a wide range of weather, lighting, tasks, difficulty levels and virtual environments affect cognitive load and performance, and provide a validated tool box of methods that can be confidently used to affect difficulty in 3D games in specific ways. This document outlines the design process, decisions, technology, rationale and aesthetics for this experiment/game design and forms the foundation for future research into cognitive-adaptive serious games and simulations.

1 Introduction

Cognitive Load Theory (CLT) (Sweller, 2011) is a well-established concept for describing mental workload and is a learning theory used to guide development of effective learning content. This theory was established in the 1980's and refers to three states of the cognitive process: intrinsic, extraneous and germane cognitive load (Seyderhelm, Blackmore, & Nesbitt, 2019). There are a number of ways cognitive load can be measured including subjective post-activity measures (Hart, 2006), direct brain measures (Hairston et al., 2014), and secondary task methods (Kristina Stojmenova & Jaka Sodnik, 2018). Real-time measures have the greatest potential for serious games and simulations as they provide an immediate and accurate view as to how a participant is coping with learning and game-play content (Aldekhyl, Cavalcanti, & Naismith, 2018; Dideriksen et al., 2018; Haji, Rojas, Childs, de Ribaupierre, & Dubrowski, 2015). However, the majority of the research exploring cognitive load in simulation or serious games tends to focus on relatively constrained virtual environments or tasks (Afergan et al., 2014; Brunken, Plass, & Leutner, 2003; Fraser et al., 2012; Haji et al., 2015). Research on how cognitive load reacts in complex virtual environments to different tasks, environment conditions and aesthetics is lacking in the literature, with few examples that explore a wide range of variables and factors simultaneously. The majority of research with similarity to the goals of the research presented in this paper tend to focus on driver distraction and road safety (Brumby, Salvucci, & Howes, 2009; Horberry, Anderson, Regan, Triggs, & Brown, 2006; Just, Keller, & Cynkar, 2008; Rossi, Gastaldi, Biondi, & Mulatti, 2012). The work presented here is approaching complex simulation for the purpose of measuring cognitive load and performance towards a cognitive adaptive serious game framework (Seyderhelm et al., 2019). This is relevant to serious games and simulation research as many of these digital learning systems can be complex, take a significant amount of time to develop, and represent a broad spectrum of learning tasks. Therefore, understanding how cognitive load responds to extended complex virtual learning environments is critical to help structure the learning tasks and also for understanding what may be manipulated to engender a desired response, for example, through dynamic difficulty adjustment (DDA). This is important to understand as over complication can negatively impact player performance or engagement due to cognitive limitations (Brumby et al., 2009); conversely too little challenge may lead to a lack of engagement or boredom (Hunicke, 2005). Thus, the ultimate goal is to achieve a balance between challenge and skill (Csikszentmihalyi, Abuhamdeh, & Nakamura, 2014; Hamari et al., 2016).

There are many methods of measuring cognitive load, however few are simple to implement and are not cost effective in real-time and in real-world non-clinical settings. The more commonly used methods tend to be post-task subjective measures, for example the NASA-TLX (Horberry et al., 2006; Naismith & Cavalcanti, 2015) or real-time physiological measures (e.g. EEG) (Antonenko, Paas, Grabner, & Van Gog, 2010; Knoll et al., 2011). These methods are effective at measuring cognitive load, however subjective measures do not produce real-time results and EEG systems can be complex to understand, install, or too expensive for many use-cases (e.g. in schools or when wearing a virtual reality headset). One method that has been shown to be sensitive to cognitive load in real-time is termed a secondary task. Examples of effective systems where cognitive load is measured via a secondary task can be found in driving and vehicle manufacturing (Conti, Dlugosch, Vilimek, Keinath, & Bengler, 2012a; Harbluk, Burns, Tam, & Glazduri, 2013; Kristina Stojmenova & Jaka Sodnik, 2018). This is principally undertaken utilizing ISO 17488:2016 the Detection Response Task (DRT) (Standardization, 2016). Various secondary task methods, and also adaptions of the DRT, have been utilized in simulation, multimedia, and training research with success (Brunken et al., 2003; Haji et al., 2015; Park & Brünken, 2015). However, the DRT requires special equipment and systems to record cognitive load data

To address these limitations, a novel approach to measuring cognitive load based on the DRT is proposed in this driving serious game to make an accessible and cheap alternative that can be more widely utilized, termed the Virtual DRT (VirDRT).

The broader aim of this research is to develop a cognitive adaptive serious game framework that can be applied using any form of cognitive load measure, combined with task measures of performance, to illicit the best possible training or educational outcomes (Seyderhelm et al., 2019). The virtual environment designed for experimental research in this research aims to address several preliminary research questions that will later be explored in greater detail. This is being undertaken through a 3D driving game developed by the Author in the Unity game engine (Unity, 2021). Within this driving game, there are several tasks for the player to undertake, in addition to driving the vehicle. The tasks chosen for this research are more thoroughly detailed below (Sections 4 & 5). The game is set across three main levels: a forest, a city and a desert and also includes a tutorial level. A driving game was chosen as it more closely aligns with the typical use of the DRT and thus is more relevant for assessing the efficacy of the VirDRT. This preliminary experiment seeks to address the following research questions:

1. What happens to player cognitive load across the duration of the driving game and across the three levels with a range of secondary tasks?
2. How does cognitive load correlate and compare with player performance across the same time range?
3. Does the VirDRT accurately record cognitive load and what, if any, impact does the VirDRT have on player performance?
4. How does cognitive load compare with engagement, frustration, interest, etc.?
5. Do the aesthetics and environment conditions impact on either cognitive load or performance (e.g. rain, darkness, forest, city or desert)?
6. Can we use these comparisons to determine an optimal level of cognitive load and performance as a potential 'learning sweet spot'?

Figure 1: A screen shot from section 4 –Challenging Drive from the Forest level.

2 Measuring cognitive load and performance

For background, the following sections describe the range of tools and techniques used to measure cognitive load and other brain functions. Three main methods are identified that are used to measure cognitive load both in real-time and post experiment. All are validated methods and include: a version of the Detection Response Task (DRT), a low-cost electroencephalogram (EEG), and the NASA Task Load Index (NASA-TLX) post task assessment. In addition, the EEG is used to assess a wide range of additional brain metrics including, including but not limited to, engagement, frustration and stress (Section 2.2).

Secondary tasks, particularly the DRT, have been used to measure cognitive load in a range of settings. The DRT has been refined to work specifically well in the testing of in-vehicle interaction and control systems (Conti et al., 2012a). Cognitive load and working memory have been measured by a range of secondary tasks, from medical training (Haji et al., 2015) to multimedia learning (Park & Brünken, 2015), and may differ in their specific methodology but they all apply similar principles. Secondary task methods work by measuring the reaction time relating to a stimulus while a person performs a primary task or function; a decrease in the response time for the secondary task stimulus indicates an increased cognitive burden from the primary task (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). In simulated driving tasks it has been demonstrated that peripheral or less important tasks are shed to enable better performance of primary and other more important tasks (Beede & Kass, 2006). Thus the pre-game briefing in this experiment emphasizes that responding to the VirDRT is the least important task (Section 5.1). The VirDRT method used in the CEDG conforms to the principles and timings specified in ISO 17488:2016 (Standardization, 2016) which is a well-established and validated cognitive load measurement method (Section 2.1).

Cognitive load, along with a range of other brain-metrics, is additionally measured via a low-cost EEG. The Emotive EPOC X was selected as this and earlier models were shown to be effective devices for this purpose (Hairston et al., 2014; Knoll et al., 2011; Wang, Gwizdka, & Chaovalitwongse, 2015). This is done for the purpose of being able to validate the VirDRT to provide an accurate real-time measure of cognitive load and to provide a range other relevant metrics in order to explore other impacts different tasks and aesthetics may have on the player, for example their engagement, frustration and stress levels (Section 2.2). The EPOC X is the latest offering by the EMOTIV company and has 14-Channels, saline-based electrodes, wireless connectivity and 9-axis motion sensors (EMOTIV, 2021a).

In addition to the VirDRT and EEG measurement of cognitive load, the NASA-TLX is used post-CEDG. This instrument gathers a subjective measure of cognitive load to provide a broad sense of the overall challenge offered by the CEDG. The NASA-TLX has been used for a wide range of research, including those focusing on driver performance in simulations (Horberry et al., 2006). It's use with experiments based on the CEDG aims to validate real-time cognitive load measures.

2.1 Virtual Detection Response Task (VirDRT)

The VirDRT used in the experiment enabled by the CEDG is very similar to the remote DRT described by Harbluk et al. (2013). One of the criticisms of the remote DRT is that it can be affected by visual load, in particular, if a participant is not looking in the correct direction when the DRT is activated. The VirDRT proposed here minimises this issue by positioning the VirDRT stimulus, a red coloured dot, in a relative screen space rather than an absolute one. In other words, in previous implementations, the DRT is positioned in a fixed location in the experimental set up; for example "[…] visual stimuli can be placed in the driver's peripheral field of view, usually on the left bottom side of the windshield or on the top right side of the dashboard" (K. Stojmenova & J. Sodnik, 2018). However, with the increasing adoption of VR and 3D synthetic environments, this method may not be optimal. Instead, the VirDRT will adopt the principles of a remote DRT setup, though apply them on the user interface (UI) level of the serious game (Figure 2). In this way, the VirDRT will always be in the same screen space relative to the participants gaze irrespective of where they may be looking and will function similar to the head mounted DRT (K. Stojmenova & J. Sodnik, 2018), but without the need for any additional equipment for either PC or VR applications.

The Unity game engine enables user interface (UI) elements to be positioned relative to what are termed anchors found within the "Rect Transform" component (Unity, 2019). These anchors are linked to the UI screen space bounds and they move or scale dependent on the screen resolution. Thus, the VirDRT red dot remains in the same space on screen as defined by its position relative to the defined anchors, in this case the VirDRT adjusts to the lower left anchor maintaining its location in the lower left quadrant of the screen (Figure 3).

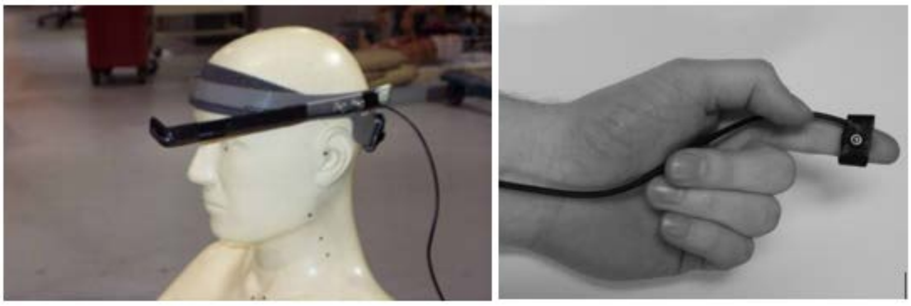

Figure 2: An example of a head mounted DRT rig (left) (Ranney et al., 2011), and a finger response button (right) (Thorpe, Nesbitt, & Eidels, 2019)

Figure 3: The VirDRT appearing bottom left of screen in the Desert level.

The VirDRT also differs in the way participants respond to the DRT stimulus. In the various iterations of the DRT described by Harbluk et al. (2013), a finger switch is used to record the DRT response. The player chooses whether the DRT switch is on the left or right hand and typically uses an index finger to operate the switch (Figure 2). This implementation is challenging when using some input devices for computer based tasks such as games. A common control device for video games are the controller types used by Play Station® and Xbox®, both of these types of controllers have shoulder buttons, that are operated by the left and right index fingers. The action of pressing these shoulder buttons, termed L1 and R1 on a Play Station® controller, is mechanically like pressing the DRT response trigger switch. As such, the VirDRT is designed and implemented to operate by pressing either shoulder button when the VirDRT stimulus is active. This also makes the implementation independent of any specific DRT equipment and more practically implementable in a broader range of applications. A generic controller, similar to the Play Station 4® device, was used in the CEDG. To support participant use, an instructional graphic (Figure 4), and a verbal description in game and at the start of the experiment session, was provided.

Figure 4: the instructional graphic provided to the player at the start of the game explaining the control schema, note the DRT response buttons.

An aim of this VirDRT implementation is to assess whether a simple and cost effective method of the DRT will be effective, making this type of cognitive load real-time measure available for a broader audience outside of a laboratory or research environments, for example in serious games, entertainment games, and simulations. The main concern with using a single controller such as this is that it requires a fair degree of hand and finger dexterity to operate. However, these controls are widely used in many console games and are unlikely to be a major hindrance as long as a tutorial section is provided to engender sufficient skill. The CEDG incorporates this VirDRT functionality into the design, however experiments should include some additional measure of cognitive load while the efficacy of the VirDRT implementation is established.

In addition to measuring cognitive load, it is also important to consider it in the context of player performance. As such, lane deviation, crashes (objects and other computer controlled vehicles), secondary task accuracy and time are recorded as metrics to correlate with the EEG and VirDRT results (Section 6), with these measures incorporated into the CEDG design and implementation. This assists in identifying where the players' cognitive load is being challenged by recording performance degradation of the primary task; that is, driving accurately and swiftly. Lane deviation has been used extensively in driver studies as a measure to indicate performance (Beede & Kass, 2006; Irwin, Monement, & Desbrow, 2015; Shinar, Tractinsky, & Compton, 2005) and a simple method of lane deviation has been adopted in this experiment. Some research focuses on drivers attempting to maintain an ideal lane position (Irwin et al., 2015; Shinar et al., 2005) based on a central lane position. However, this does not accommodate driving preference, corner apexes, or other variables. In the CEDG, the measure of lane deviation is captured when the player crosses the center line onto the wrong side of the road, however minutely. Somewhat obviously, crashing into objects or oncoming cars is an indicator of poor performance. On occasions the Unity AI waypoint system (Section 5.2) can cause the computer controlled cars to deviate into the player lane when driving around corners. The player was warned of this phenomenon and tasked to remain vigilant and to avoid other driver error, and this practice should be adopted in all use of the CEDG.

In summary, by completing the same track three times, randomizing the presentation order of the VirDRT loop, measuring cognitive load through two different techniques, and collecting performance metrics, a 'tool box' of techniques to affect cognitive load has been established. The CEDG delivers this functionality as a set of methods can then employed in additional adaptive experiments as a validated toolbox of elements that can be changed to affect difficulty.

Figure 5: A participant playing the CEDG also showing the experimental set up. Note the EEG and controller being used (included with permission).

2.2 Additional data collected

Full use of the CEDG is enhanced through pre and post surveys. This additional data can be collected to help contextualize and provide further relevant information. It is recommended that the participants/players complete a pre-experiment Demographic and Game Preferences survey (Attachment A) and a post-game Engagement Survey (Attachment B) (Wiebe, Lamb, Hardy, & Sharek, 2014). The Engagement Survey (Attachment B) assesses the level of engagement players experience while undertaking the serious game. This aids in identifying whether the serious game being tested is enjoyable and approaching the aim of achieving the outcomes desired in the three-part framework described by Hookham and Nesbitt (2019). When combined with the pre-game surveys (Attachment A), this informs how different individuals respond and engage with the serious game. This can then inform adaption strategies into the future to improve training outcomes by better adapting a serious game to the player. The post-game Engagement Survey (Attachment B), has been shown to be an effective measure of engagement and represents the most broadly accepted post-game survey currently available (Wiebe et al., 2014).

The collection and analysis of the EEG data is supported through vendor software. The EMOTIV software used, EmotivPRO 2.6.0.268, provides a range of interpretations of the full collected data that can shed further light on brain function during the CEDG, termed performance metrics in the software. These performance metrics include: engagement (En), excitement (Ex), focus (Fo), interest (In), relaxation (Re), and stress (St) (Figure 9).

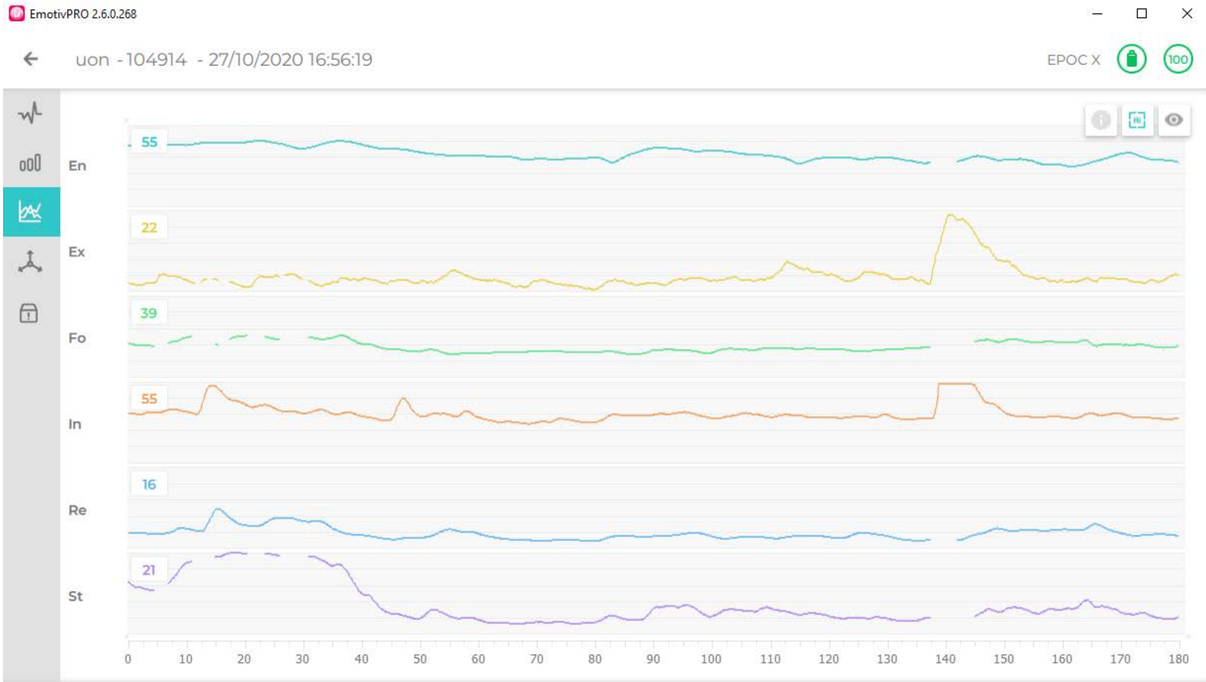

Figure 6: Example performance measure capture in the EmotivPRO software.

The EmotivPro manual describes the above metrics as follows (EMOTIV, 2021b):

> "**Stress (FRU)** is a measure of comfort with the current challenge. High stress can result from an inability to complete a difficult task, feeling overwhelmed and fearing negative consequences for failing to satisfy the task requirements. Generally, a low to moderate level of stress can improve productivity, whereas a higher level tends to be destructive and can have long term consequences for health and well-being.
>
> **Engagement (ENG)** is experienced as alertness and the conscious direction of attention towards task-relevant stimuli. It measures the level of immersion in the moment and is a mixture of attention and concentration and contrasts with boredom. Engagement is characterized by increased physiological arousal and beta waves along with attenuated alpha waves. The greater the attention, focus and workload, the greater the output score reported by the detection.
>
> **Interest (VAL)** is the degree of attraction or aversion to the current stimuli, environment or activity and is commonly referred to as Valence. Low interest scores indicate a strong aversion to the task, high interest indicates a strong affinity with the task while mid-range scores indicate you neither like nor dislike the activity.
>
> **Excitement (EXC)** is an awareness or feeling of physiological arousal with a positive value. It is characterized by activation in the sympathetic nervous system which results in a range of physiological responses including pupil dilation, eye widening, sweat gland stimulation, heart rate and muscle tension increases, blood diversion, and digestive inhibition. In general, the

greater the increase in physiological arousal the greater the output score for the detection. The Excitement detection is tuned to provide output scores that reflect short-term changes in excitement over time periods as short as several seconds.

Focus (FOC) is a measure of fixed attention to one specific task. Focus measures the depth of attention as well as the frequency that attention switches between tasks. A high level of task switching is an indication of poor focus and distraction.

Relaxation (MED) is a measure of an ability to switch off and recover from intense concentration. Trained meditators can score extremely high relaxation scores."

These measures will allow the assessment of a broad range of additional factors, including the notion of how the different tasks, challenges and environment factors impact the player.

3 Art style and aesthetic considerations

Research question four, presented in the Introduction section, seeks to understand how, or if, different aesthetics affect cognitive load and performance in complex virtual environments. The experiment detailed in this working paper seeks to explore how performance on the same track layout with the same challenges, within the Cognitive Effective Driving Game (CEDG), are impacted by entirely different virtual environment aesthetics.

Attention Restoration Theory (ART) predicts that natural environments have a restorative effect on human cognition and emotion (Ohly et al., 2016). Recent research by Jiang, He, Chen, Larsen, and Wang (2020) has demonstrated that in simulated driving tasks, natural environments, and more specifically green space, can mitigate negative mental states (p 25). The environment developed here builds on the experiment by Jiang et al., (Jiang et al., 2020) to consider if there is a significant impact of natural environments in a serious game on a players cognitive load. Jiang et al. modelled a single motorway with the same road layout and city background, to which they added increasingly green aspects, from grass, to shrubs to trees. They showed the greener the level, the more positive mental responses were recorded. Therefore, it is necessary to produce sufficient visual fidelity to render an appreciable approximation of each environment for it to be clearly distinguishable and have a recognizably different ambience. Research has shown that different environments can cause different emotional and cognitive responses (Valtchanov & Ellard, 2015) and this affect is explored further in this experiment with a particular foci on if this impacts cognitive load. Whereas Jian et al., (2020) focused on green spaces in a limited city environments, this experiment extends on their work by comparing a large environment with many different locations and structures across a city, a forest, and a desert environment. To do this, the tasks and track layout are the same in each level; figures 2, 3 and 4 provide an example of the same road section in these three different levels.

Previous research has also demonstrated that the tone of lighting impacts on player performance and affect (Knez & Niedenthal, 2008). However, that experiment utilized extreme differences in the light colour. The three levels developed for the CEDG maintain the same lighting throughout, but the background environments do have tonal differences including red (desert), blue (city) and green (forest). These differences are unlikely by themselves to be impactful due to the lower impact tonal variation of the background environment rather than significant lighting changes outlined by Knez and Niedenthal (2008), however it is worth recognising these differences as part of the broader aesthetic conditions. Understanding the impacts of these differences in level aesthetics may be informed through real-time measures of cognitive load.

Figure 7: Desert Level at the start point

Figure 8: City level at the start point

Image 9: Forest level at the start point

4 Initial Design Concepts

Prior to commencing the development of the CEDG, a design meeting with the research team was undertaken whereupon computer science and psychology methodologies were discussed in order to apply appropriate rigor and consideration to all aspects of the CEDG. From this, the following key points were collated to enable the CEDG to be as complex as possible while also providing meaningful, and manageable data:

1. There are three game levels (City, Forest and Desert), and they will each be undertaken in a random order.
2. Within each level, the road/track will be identical and take approximately 5 minutes per circuit.
3. The player will undertake two full circuits of each level, one with the VirDRT and one without, again in random order (Figure 10).
4. Throughout the levels, the player will undertake the same, or very similar, tasks in addition to driving. There will also be easier sections with no tasks, and finally sections with no secondary task but different levels of complexity of the primary task (Section 5.4). These variations are intended to manipulate the cognitive load of participants throughout the experiment.
5. In addition to biome changes in each level, light and weather conditions will also change to assess the impact of these elements on cognitive load and performance.
6. There will be a tutorial level, which takes approximately five minutes to complete, to enable the player to become familiar with driving and VirDRT controls.

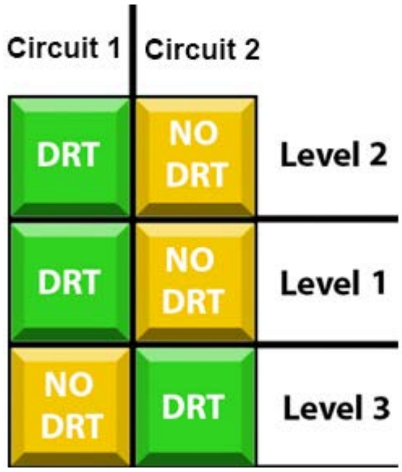

Figure 10: Example VirDRT structure - DRT is randomly selected in each level and for each player, one circuit with (green) one circuit without (yellow); the level order is also randomized for each player.

Once the above design points were finalised, an iterative track design process was undertaken with the aim of each circuit being approximately five (5) minutes in duration. Each level requires the completion of two circuits, leading to an estimated total time of ten (10) minutes per level. Figure 16 provides map overview of the final road layout and the sections and tasks. These were the same for each level and are described in more detail below (Section 5).

5 CEDG Design

The CEDG has been designed to assess the impact of different aesthetic conditions on player cognitive load and performance across a wide range of tasks in a complex virtual environment. The CEDG also tests the layering of primary and secondary tasks to allow a robust measurement of cognitive load using the VirDRT approach. To this end, the player will be given additional tasks to perform requiring different cognitive processes. For example, this may be to count or respond to certain assets, such as instructing the player to count how many blue pickup trucks are in a zone while driving and avoiding hazards. In this example, there will be other pickup trucks of different colours in the environment, requiring visual and cognitive discrimination and a likely increase in overarching cognitive load (Conti, Dlugosch, Vilimek, Keinath, & Bengler, 2012b).

As previously mentioned, to measure the impact of different aesthetics in a complex environment, three levels were designed, each with a different biome: a city, a forest and a desert. Each level has the same road layout, all that varies is the background. Within each level are sections that have the same types of tasks, with slight variations to ensure the player is required to concentrate in each level. The player is required to drive two circuits of each level, the order of each level was randomized for each player. In one circuit the VirDRT was not active and the other circuit it was, again the order of this was randomized for each level and for each player. In this way the VirDRT efficacy was measured and assessed to see if it had an impact on player performance.

5.1 Participants and Pre-commencement instruction

The aim of this experiment is to have a minimum of 30 participants play the game to be able to draw some statistically significant conclusions. The project has been approved by the University of Newcastle's Human Research Ethics Committee, Approval No. HREC #H-2020-0069. Students are recruited to participate through the University School of Psychology experimental management system, SONA. Applications to participate are extended to students in the School of Psychology and within the Discipline of Computing and IT, with student participants receiving course credit for their participation. Staff and post-graduate research students are also invited to participate, however they receive no benefits or credits for doing so. Each participant is provided with an information statement that outlines the processes and purposes of the experiment and signs a consent form. Consent can be withdrawn at any stage and participants may remove themselves and any of their data from the study at any time up until publication of results.

Prior to commencing the session, each participant was given the same overview, which detailed what should be their priorities and the goals of the game. Similar to other research using DRTs, the primary task was emphasized as most important (Bruyas & Dumont, 2013). The following list provides details of specific guidance and instructions to participants to inform future studies:

- They were to perform the primary task (driving) as their main focus as accurately and well as possible, ie. staying in their lane and not crashing, while being as fast as possible. Accuracy and quality driving was expressed as being more important than speed.
- They would have to undertake various tasks, examples were provided, however driving was the primary task.
- The levels and the VirDRT activation randomness were explained.
- Driving was more important than responding to the VirDRT.
- They would receive penalties for crashes and lane deviations, and they would be timed, and a leader board was in effect where penalties were multiplied with speed (therefore driving accuracy was reinforced as being more important as it served as a significant multiplier).
- AI cars would not try and avoid the player and would occasionally deviate into the player lane; the player must remain vigilant and avoid crashing during these incidents.
- The counting of cars section (Zone 1 – Start & Count Task) will be tested later with a penalty or bonus applied if the player got it correct or not, impacting the score and leaderboard.
- Exploring the controls and getting comfortable with them was encouraged during the tutorial section.
- If the EEG lost connection the game would be paused, connection fixed, and then the game resumed.
- Players were free to take a break or continue directly between levels.
- They could withdraw from the experiment at any time.

5.2 Game elements, controls, and NPCs

Each zone throughout the level has different challenges and considerations, however some universal concepts are applied to provide feedback, apply pressure, give information and so forth. They are described below.

Directions: whenever the player approaches a turn (excluding Zone 3) they are provided an animated pink chevron that clearly indicates the direction to turn (Figure 11). The animation is in the direction the player is supposed to go. The chevron appears in advance of the corner with sufficient time, based on robust gameplay testing, for the player to respond.

Figure 11: Pink chevron indicating direction for player to turn.

UI information: in the top right corner is a user interface (UI) element that displays current time taken in this level as well as the penalty score (Figure 12). In addition to displaying penalties in this UI element, every time the player deviates, a negative tone is played. Crashes are furnished with a crash sound combined with the negative tone.

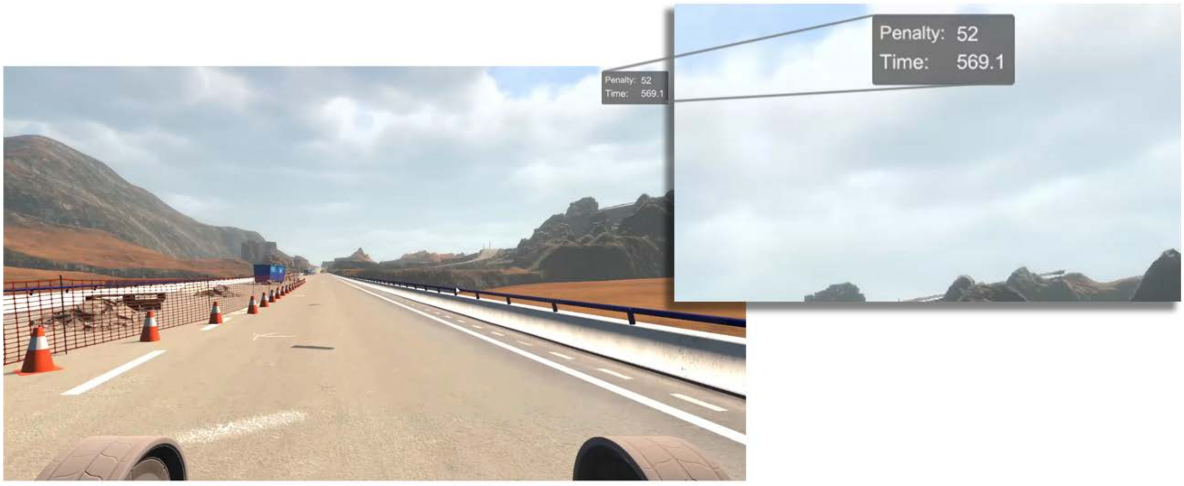

Figure 12: UI elements.

Vehicle Controls: the vehicle driven by the player in the CEDG is from the Unity Standard Assets car prefab asset (Unity, 2020). This is a relatively simple and complete control schema for physics-based driving games that was modified to be suitable for the CEDG. The primary modification is the change from a higher 3rd person camera perspective to a 1st person view. Several car cabin interior 3D models and representations were explored to make it appear like the player was sitting inside a vehicle. After considerable play testing and gathering of feedback, it was decided to remove the interior and just leave the front wheels visible. The car interior restricted the field of view too greatly and made it too difficult to drive; leaving the wheels to help the player steer, stay within the lanes and better avoid crashing. Most of the default settings in the Unity car prefab were left unchanged, however the vehicle top speed setting was reduced from 240 KPH, the original setting,, to 80KPH, the desired top speed that is a common speed prevalent on roads in Australia. A range of scripts were added to manage performance measures, audio, and UI elements. The Unity settings for the vehicle asset are displayed in Figure 13.

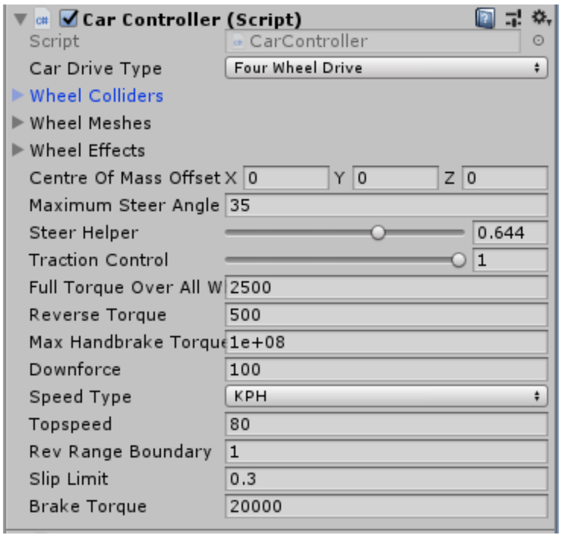

Figure 13: Unity standard assets car prefab settings used in CEDG.

NPCs: the Unity Standard Assets also contained the same vehicle control system with the addition of an AI controller, allowing the cars to drive themselves. All the vehicles in the CEDG utilized the waypoint-based car controller system. The default vehicle model was replaced by a range of different 3D representations, and the vehicle speeds, "cautious" and "Lateral Wander" settings were adjusted. These were adjusted so the vehicles were not too fast, with the top speed reduced from 225 KPH, the default setting, to a more realistic 65 KPH which is closer to standard Australian road speed limits that typically range from 60 to 70 KPH (see Figure 14). The vehicles follow a waypoint system and were set up to ignore the player and other cars; they simply drive in an endless loop on their respective waypoint-based tracks. Cars are tagged in Unity so that the player can crash into them, but they cannot crash into each other. This was done to avoid random crashes between AI cars that occasionally occurred that could block the track. At different points in the game, new cars were triggered to create player challenges (Zones 5 and 10).

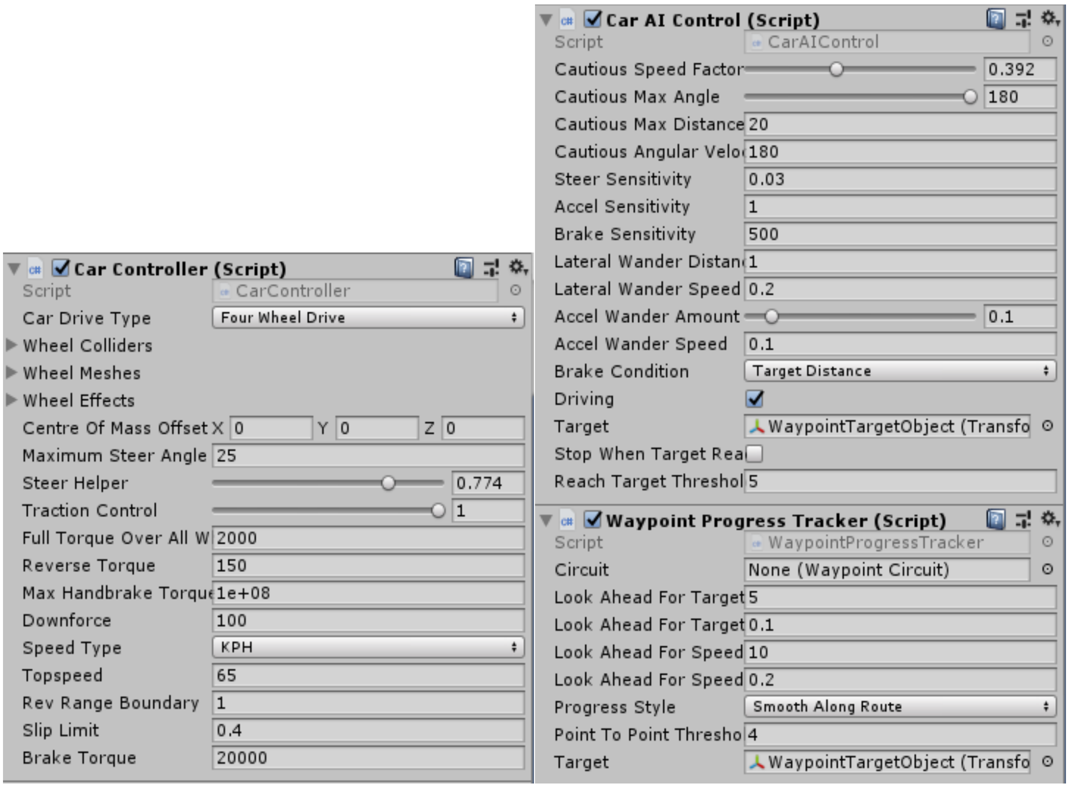

Figure 14: Waypoint car controller details, from Unity Standard Assets.

Dynamic Weather: There has been very little research on the impact of dynamic weather in games, especially from a training or education standpoint. There has been some limited study that indicates dynamic weather can positively impact player engagement (Rehnberg, 2021) and is worth exploring further (Barton, 2008). Therefore, the inclusion of dynamic weather, and the rain zones (Zones 6 and 7) in particular, may provide some interesting information on their value from a player engagement, cognitive load and/or education and training perspective.

Dynamic weather in the CEDG is not only limited to the two rain zones; throughout the game the clouds move and change, trees blow in the wind, there is lightning and thunder (in appropriate zones), and water ripples in the distance. The amount of tree movement is directly affected by the type of weather, for example the raining zones have thunder, lightning and the tree branches sway and move more vigorously. To achieve this, a Unity asset store (Unity, 2020) called Enviro Sky and Weather was used that enabled different weather zones and effects.

Dynamic Lighting: The game uses dynamic real-time lighting to affect the environment and player conditions. This is done with real-time shadows from objects that move if they animate, for example, when the weather changes, the lighting dims and turns duller, and in the dark tunnel, the only light is from the player's vehicle headlights. Lighting can impact player engagement and emotional responses in games (El-Nasr, 2006; Knez & Niedenthal, 2008), however no research was discovered on how dynamic lighting may impact the efficacy of serious games or simulations.

Sound Effects: There are various audio effects throughout the game, including vehicle engine and driving noises, rain and thunder, and a negative sound played when the player crashes or deviates from the driving lane. Crashes have additional sound effects coupled with the negative tone, these include metal crashing and glass breaking noises. Sound effects are included to add to the sense of immersion (Grimshaw, 2012) and negative tones in order to both inform the player they have done something wrong and potentially trigger an emotional response or cause the player to be more cautious.

5.3 Tutorial Level

A tutorial level was created with the purpose of having the player gain familiarity with the controls, UI, and other aspects of the game. The tutorial art aesthetic matches the general standard of the rest of the game although it is built in an environment with minimal visual interest (Figure 15). The tutorial includes instructions on game controls as well as the game goals – reiterating the briefing instructions provided in-person.

At the start of the tutorial level the player is not required to respond to the VirDRT and is encouraged to explore the vehicle controls, as well as to hear the auditory feedback tone upon crashing or deviating from a lane. The game commences with a long straight section and leads to gradual turns and ultimately a few steeper turns and hills; the aim here is to provide increasing challenge. The tutorial introduces the pink directional chevrons, and finally the player is introduced to the VirDRT and its functionality. The in-game instructions regarding the VirDRT reiterate that driving is more important than responding: "*Very soon a red dot will start to appear on your screen, when you see it respond by hitting the trigger buttons. This is the DRT. Remember though, that driving is your number one priority.*"

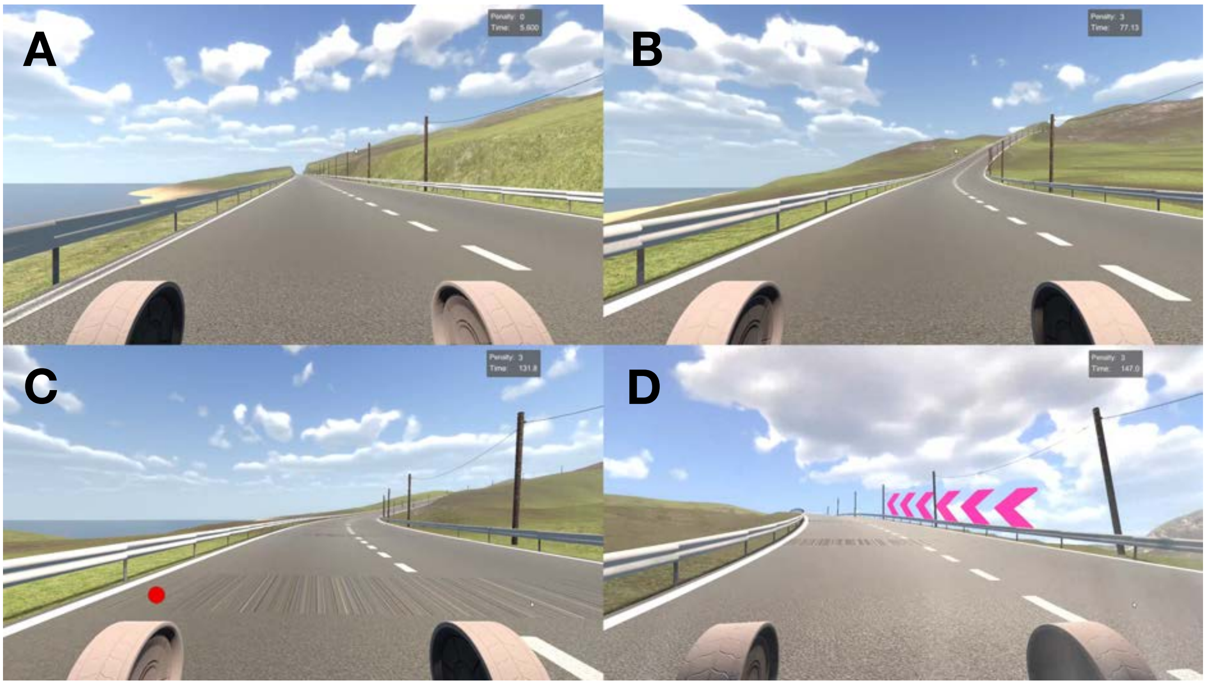

Figure 15: Tutorial level. The long straight at the start (A), the start of gentle curves in the road (B), first opportunity to respond to the VirDRT (C), and introducing pink directional chevrons (D).

The primary purpose of the tutorial level is for the player to become familiar with the controls of the vehicle so that measures of cognitive load in later levels more accurately record the cognitive load of the tasks rather than complexities involving the control schema. Learning the controls of a game can be challenging if a player has never used them before; it is expected that by the second and third level the player should be competent with the controls. In this way the various cognitive load measures will accurately record the impact of tasks and environment effects and not be affected by game control complexities.

5.4 Level zone descriptions

The following sections describe each zone, the various tasks, and the rationale behind them. The map (Figure 16) colour codes each section and the player drives the map in a clockwise direction.

Figure 16: Map of the track, coloured per zone.

5.5 Zone 1 – Start and Counting Task

The starting zone of the CEDG commences with a drive up a main street; traffic comes in the opposite direction and shortly after commencing, the player is asked to count different colour utility vehicles (pickup trucks) parked on the side of the road. The number and colour of the utility vehicles varies per level and in each case there are other similar vehicles parked nearby requiring simple visual discrimination. There are between 7 and 10 utility vehicles parked beside the road; that number range was chosen based on research indicating this is a potentially optimal range to challenge recall (Miller, 1956). During the counting task, the player must continue to drive accurately and navigate a right hand turn, while avoiding traffic (Figure 17). The player is asked to recall the number of utility vehicles counted in Zone 1 later in the level (Zone 5). It is important to make the secondary task, in this case counting the cars, challenging, to assess the impact this type of task has on cognitive load; if the task were too easy, or too hard, the impact may be so negligible as to prove unmeasurable or insignificant. The counting task is a simple task, made more complex by the requirement to maintain a primary task to sufficient accuracy, that is, by not deviating out of the driving lane or crashing. Importantly, the player is aware that later in the level they will be tested to recall the number of cars they counted, increasing pressure on the player for an accurate count (Cassady, 2004).

Level	Circuit	Vehicle Colour	Number to count
Forest	1	Purple	10
Forest	2	Yellow	8
City	1	Green	7
City	2	Blue	9
Desert	1	Red	8
Desert	2	Blue	9

Table 1: Number and colour of utility vehicles per level and circuit.

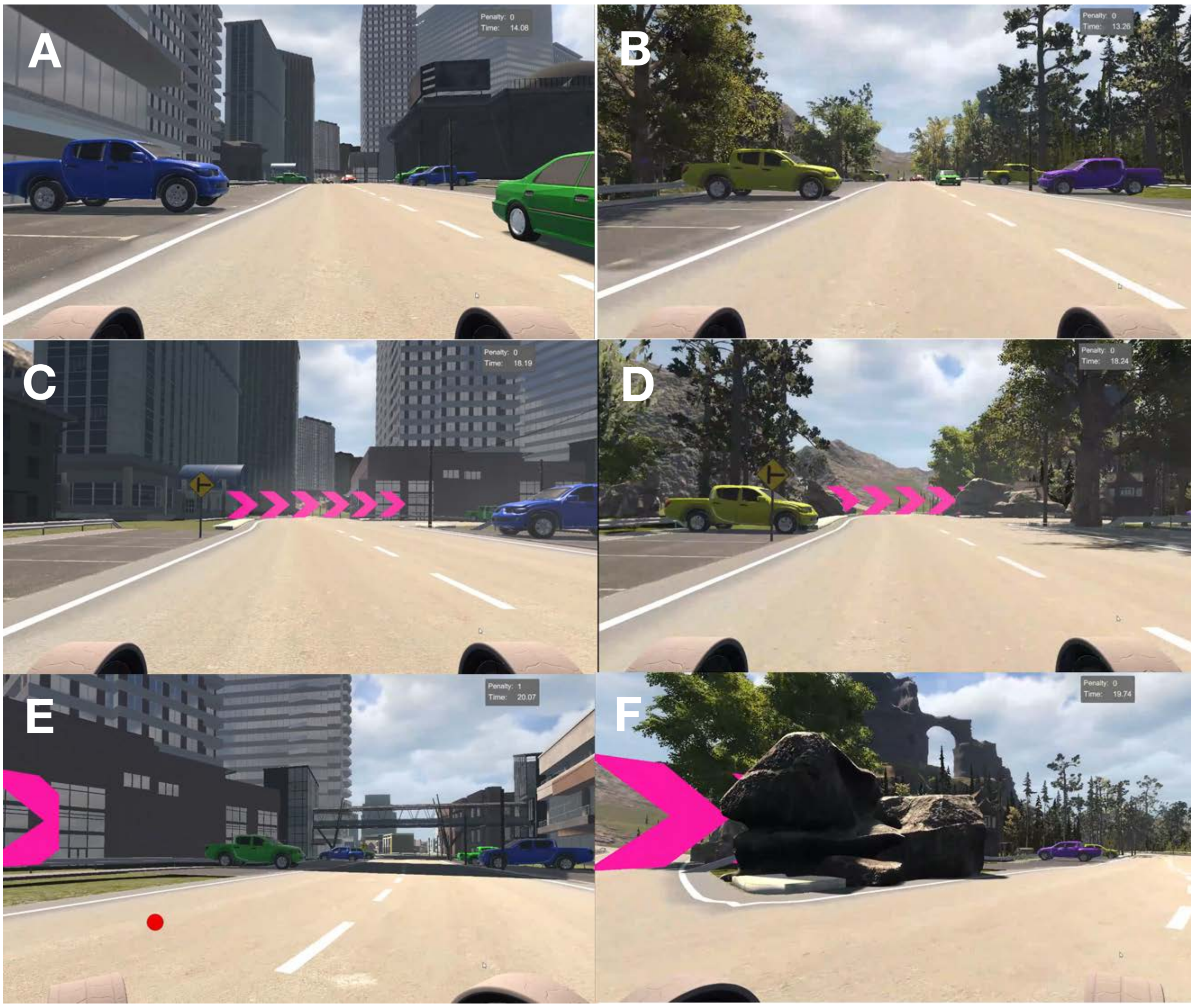

Figure 17: Zone 1 from the City (A) and Forest (B) showing similar positions from when the count task commences to the first right turn (C-F). Note the 'utility vehicles' parked beside the road (A-F), and pink chevrons indicating a turn (C-F).

5.6 Zone 2 – Audio Directions

In zone 2, *audio directions*, the player is presented with three or four recorded voice directions to follow, inspired by prior research relating to the giving and receiving of spatial directions (Vanetti & Allen, 1988). Listening to spoken words has been shown to have a significant impact on driver performance (Rossi et al., 2012). By integrating audio directions into the driving task, we aim to explore the impact on cognitive load when receiving complex directions in a driving simulation. This may be applicable in other simulation or gaming tasks, that is, the impact of listening to directions or instructions while performing a relatively complex task. These directions are only stated once and require concentration to follow. It has been shown that listening to someone speak has a significant impact on a simulated driving task performance (Just et al., 2008) and the CEDG explores how the listening and subsequent application of these directions impacts both performance and cognitive load. Instructions are provided in-game prior to receiving the directions as follows: "*Coming up you are going to be given a series of directions - follow them very closely*." Following is an example set of four audio directions indicative of all the provided directions from circuit 1 of the City level: "*Take the next right, take the first left, then the next right and at the end turn left*."

Table 2 lists the number of directions by level and circuit.

Level	Circuit	Number of Directions
Forest	1	3
Forest	2	4
City	1	4
City	2	4
Desert	1	3
Desert	2	3

Table 2: The number of directions per level and circuit.

Figure 18: Zone 2 city and desert examples of the Audio Direction zone, each pair of images from top to bottom is from a similar location.

5.7 Zone 3 – Rest

Zone 3 and Zone 9 are designated with the title Rest. Both these zones have no additional tasks and have relatively easy driving difficulty. The main challenge in these zones comes from oncoming traffic during turns. The aim of which is to observe if cognitive load reduces during short recovery period.

This will provide information about whether a short reduction in complexity may help lower cognitive load as a form of baseline or reset.

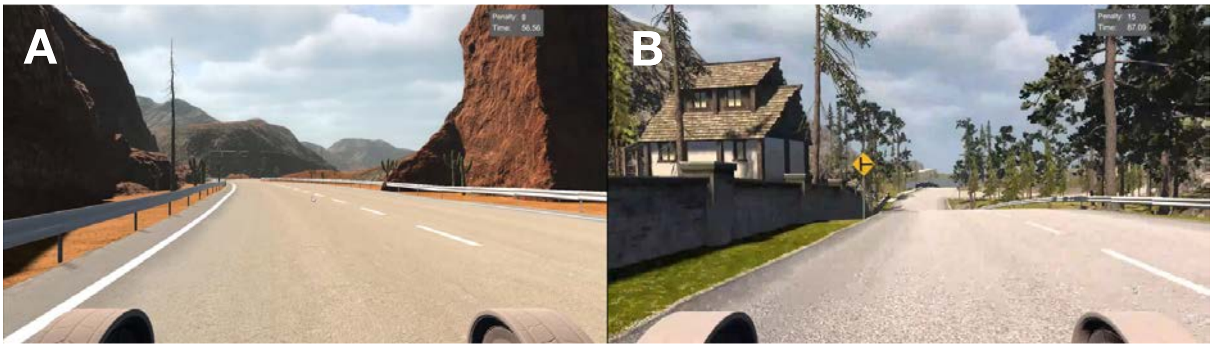

Figure 19: Zone 3, desert (A) and forest (B) near the start and end of the zone respectively.

5.8 Zone 4 – Challenging Drive

Zone 4 is a challenging driving section with no additional tasks (other than the VirDRT). However, the drive is very difficult factoring crests, steep sections, blind corners, and hair pin turns. At the start of the zone an audio instruction is played as follows: "*The following section is a very difficult road to drive on, so be careful*." This plays as the player goes past a road sign indicating a twisting road ahead (Figure 20).

The challenge of the task in Zone 4 is significantly increased and thus aims to increase the level of stress; this aligns with the concept of absolute difficulty (Adams, 2014). Adams (2014) describes absolute difficulty in terms of "...skill required and the stressfulness put together" (p 322), with stress defined by Adams as being a reflection of time pressure (p 322). Time pressure in the CEDG is applied via the initial briefing (Section 5.1), the leader board, as well as the timer display in the UI.

Within Zone 4, all the steep corners or crests are preceded by a typical road sign providing advice on what to expect, however, if they are ignored or missed then driving this section is made more difficult. The car is challenging to control and requires good use of both the brakes and acceleration. It is also one of the longest sections and combined with the difficulty level it is likely to impact both performance and cognitive load. The following image sequence presents a range of locations that exemplify the challenges encountered, with some images showing the street signs indicating the upcoming road pattern (Figure 20).

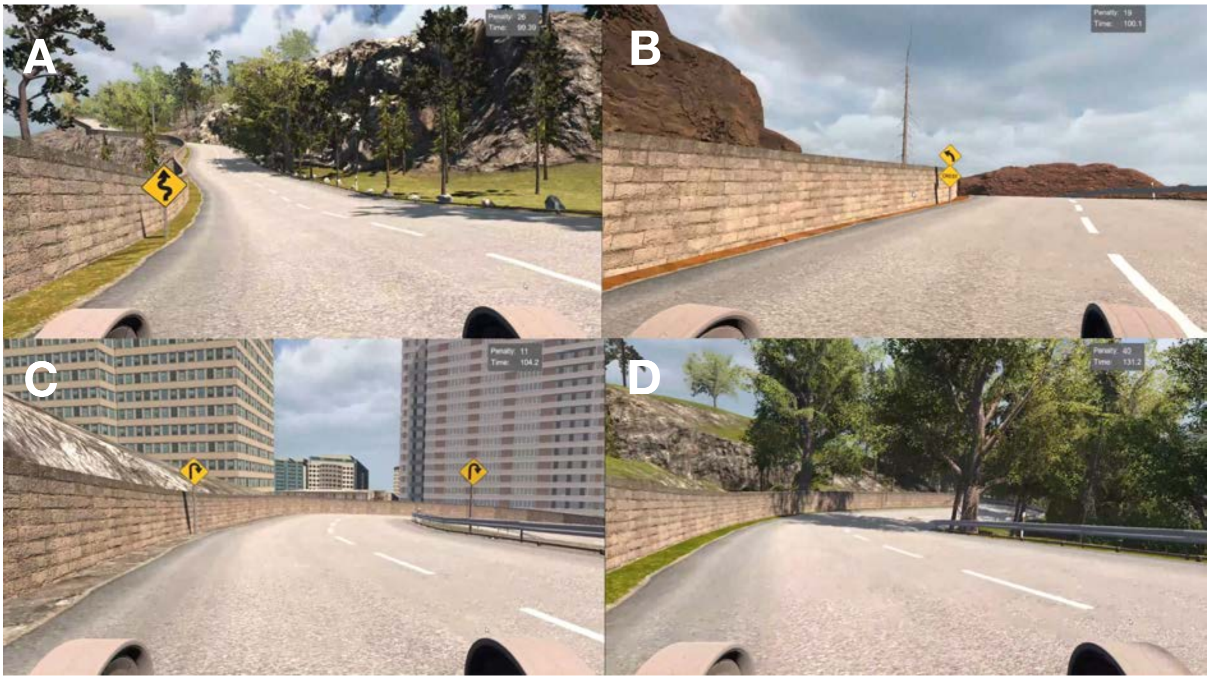

Figure 20: Zone 4, (A) is the start of the zone where audio instructions are heard; note the twisting road sign. The other three images (B – D) demonstrate some of the challenges and associated road signs (A – C) indicating what is approaching.

5.9 Zone 5 – Follow and count test

There are two challenges in this zone; the first is to follow a specified car while maintaining a distance of between 10m and 18m, while the second is to recall the number of utility vehicles counted in Zone 1. Following cars has been used as a measure of distraction and performance in a number of driving simulator research tasks (Dorn & Barker, 2005; Drews, Yazdani, Godfrey, Cooper, & Strayer, 2009) and as such, a version of this task is applied in the CEDG. During each circuit, the player is given audio and visual instructions to follow a car while maintaining a set distance range of between 10m to 18m (Figure 21). A UI element is positioned above the follow car showing the distance between the player and the follow car (Figure 21). This UI element turns red when the player deviates out of the desired distance; the player distance is sampled every frame. For each second out of range, the player receives a penalty point and a negative audio tone is played. In each circuit the AI vehicles drive at slightly different speeds and occasionally meander in their lanes due to the “cautiousness” rating of the AI system. This leads to slower driving than the player vehicle would naturally drive, as well as some randomness. These variations require the player to respond and be watchful, in other words, the two AI systems are slightly erratic. The AI vehicle to follow in circuit one is also slower than the one in circuit two. The AI follow cars are the same type in each level, with circuit one being a green sedan and circuit two being a dark blue sedan.

The second task for this zone is to recall the number of utility vehicles previously counted (Zone 1). As the player reaches a trigger point in the zone, this causes the game to pause and the player is presented with a question asking them to recall the number of cars counted in Zone 1 of that circuit. They are provided a text entry box to provide an answer; a correct answer receives a positive audio tone and green text proclaiming “Correct”, along with a 10-point bonus to the score. An incorrect answer leads to a negative sound, and red text saying “Incorrect” plus a 10-point penalty to the score

(Figure 22). These visual and audio cues along with the adjustment to the score were applied to help enhance motivation for the player to strive for correct answers.

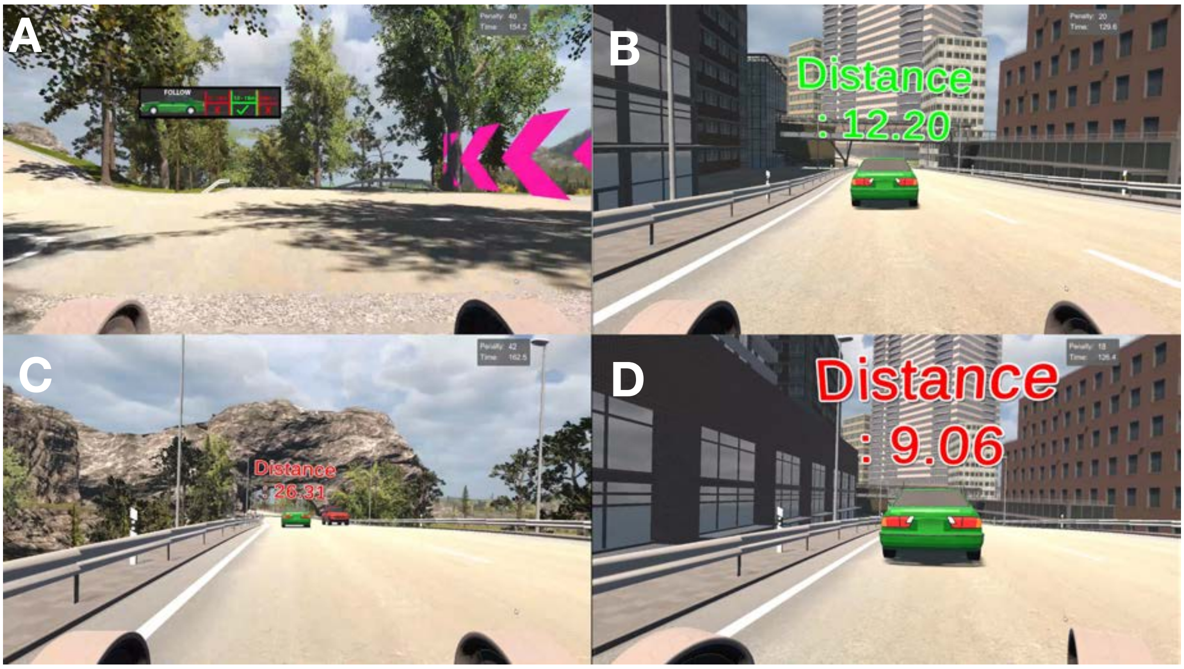

Figure 21: The visual and verbal instructions (A), maintaining the correct distance (B), falling too far behind (C) and being too close (D).

Figure 22: The correct response (A), and an incorrect response (B)

5.10 Zone 6 – Follow and Rain

Zone 6 continues the same follow task from Zone 5, however in this zone it starts to rain, there are flashes of lightening, and the sound of thunder (Figure 23). There is no effect on vehicle handling characteristics; it is purely an aesthetic and audio game characteristics to observe if there is an impact on cognitive load or performance.

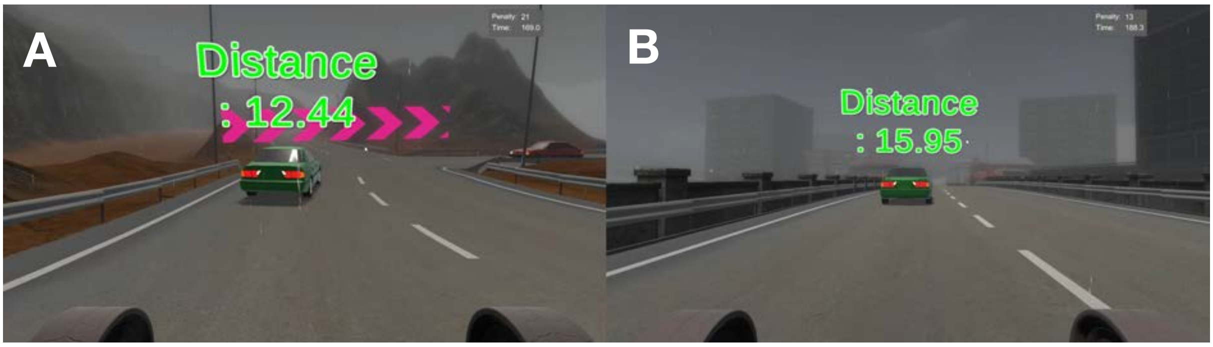

Figure 23: Zone 6, the start of the Zone in the Desert level (A), and near the end in the City level (B). Note the grey sky, rain drops and increased fog.

5.11 Zone 7 – Narrow Entry and Rain

Zone 7 commences with a very narrow entry that is only wide enough for a single vehicle (Figure 24). The rain continues throughout this zone but no other tasks are required. In this way, the impact of rain in a standard difficulty primary task is measured. The player has no secondary tasks to perform other than the VirDRT, when active. This zone may also form a cognitive load reduction or rest section as the absolute difficulty is comparatively low and will therefore provide a good measure to assess if dynamic weather impacts cognitive load, or any other metrics (engagement, surprise etc.).

Figure 24: Entry to Zone 7, showing each level.

5.12 Zone 8 – Tunnel and Dark

Similar to Zone 7, which explores the effect of weather on various player metrics, Zone 8 explores how darkness and a perception of restricted space may also impact the player. As the player drives into the tunnel (Figure 25) the environment goes dark and the vehicle headlights illuminate, albeit weakly (Figure 26). The sudden decline in light also leads to lower view distances and thus a reduction in reaction time for the player, as bends in the road are not visible until much closer, which is assumed to increase task difficulty.

In addition to this, the tunnel environment introduces a concrete median barrier and an enclosed environment. The barrier is positioned over the centre road lines and has very little impact on lane width. In combination with reduced view distance and light quality, it creates the impression of less space (Figure 25) and may impact cognitive load as the player needs to concentrate more.

Figure 25: Entry to Zone 10 – Tunnel and Dark

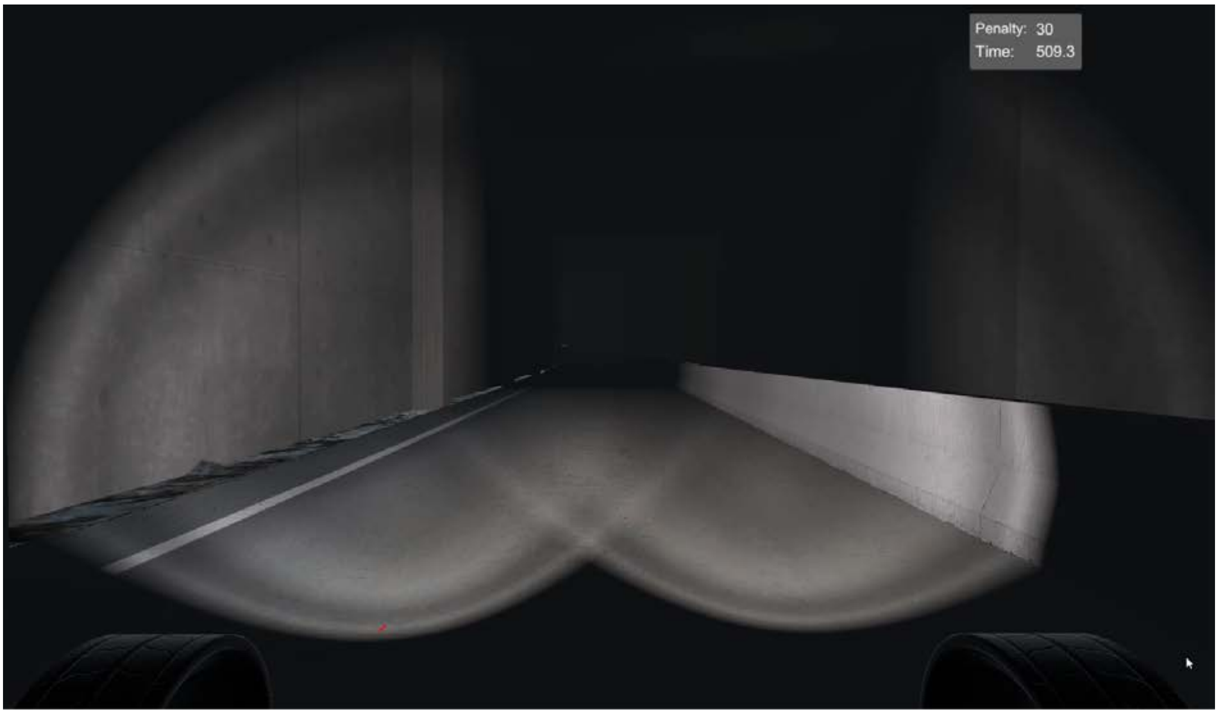

Figure 26: Inside the dark tunnel, lights on and the centre-road crash barrier.

5.13 Zone 9 – Rest

Zone 9 is a short section with low driving difficulty similar to Zone 3 (Section 5.7).

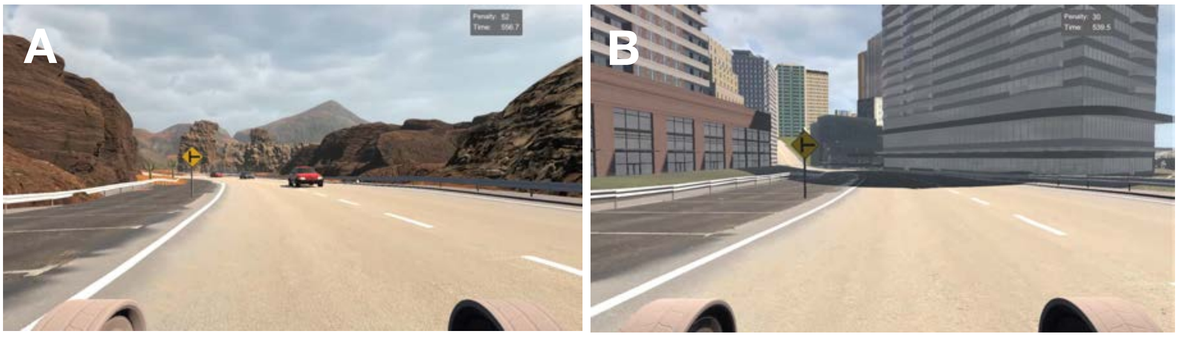

Figure 27: Zone 9 in the Desert (A) and City (B)

5.14 Zone 10 – Bridge

The final zone of the circuit (Zone 10) presents the player with a long straight bridge that has roadworks and oncoming traffic. The roadwork blocks the left-hand lane, requiring the player to cross into oncoming traffic when a gap appears and move forward to a waiting bay further along the bridge; no lane deviations are recorded in this zone. Various driving spatial judgment tasks, for example judging gaps in traffic, have been explored in virtual driving environments that indicate that this type of task is relatively complex (Alexander, Barham, & Black, 2002; Feldstein, 2019). Overtaking maneuvers are frequently misjudged and are complex tasks, requiring the ability to rapidly judge the player vehicle acceleration capability, distance and oncoming vehicle speed (Gray & Regan, 2005). However judgements, or risk behaviors, are frequently different in virtual versus real-world tasks (Feldstein, 2019) and this zone explores the cognitive effects of judging distance, speed and time and may also shed light on player risk characteristics and whether there are demographic differences in these risk attitudes.

As the player approaches the bridge, they are given verbal instructions on what will occur. To encourage more realistic risk acceptance behavior, additional consequences for crashing were introduced for this zone only. The instructions are: "*You must drive across the bridge without crashing into other cars. Be very careful, if you do crash you will be sent back to the start to try again*." As the player approaches the bridge several cars are spawned with various distances between them, in addition, any AI cars currently driving across the bridge on their waypoint path act as additional obstacles. By having AI waypoint cars included makes the challenge more variable and less predictable, meaning across the six (6) circuits, the players are less likely to identify a fixed strategy based on vehicle spacing. The bridge and the obstacles remain unchanged, only the AI waypoint cars cause variability.

The following Figures (28 and 29) show Zone 10 in more detail.

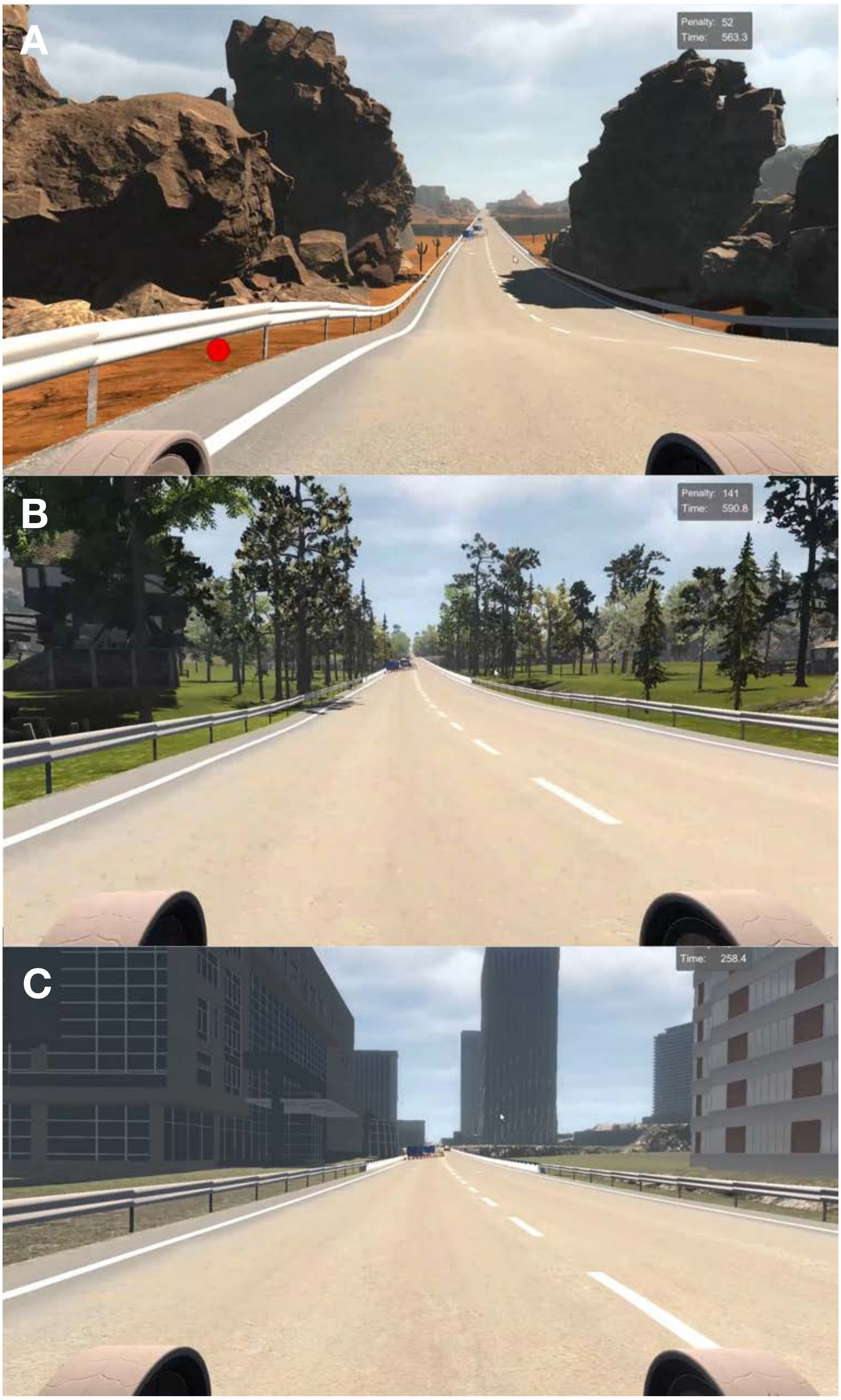

Figure 28: Approach to the bridge in the three levels; desert (A), Forest (B), City (C).

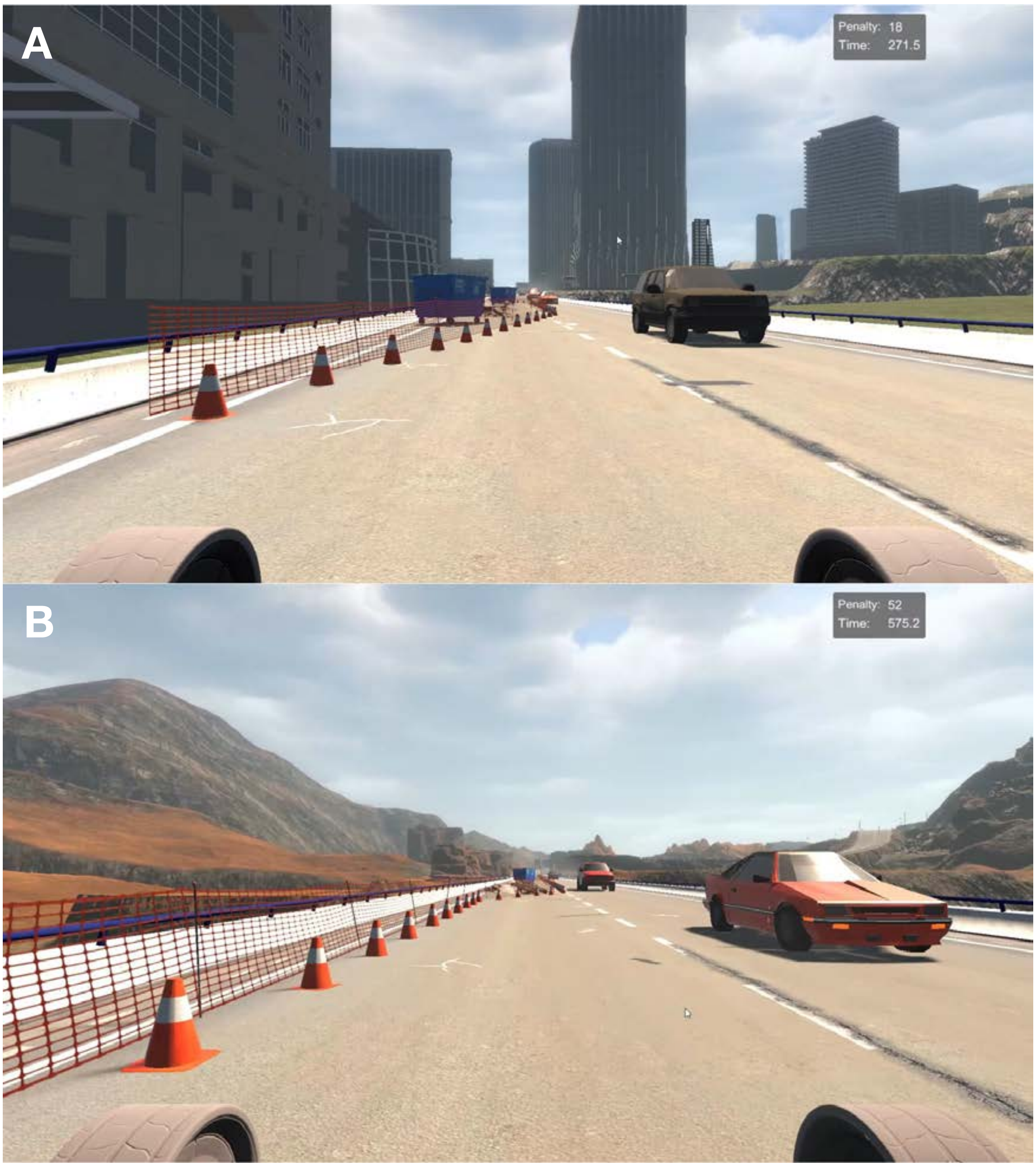

Figure 29: Driving across the bridge, the start of the road works (A) and the middle bay (B).

6 Data recording and output

Throughout the game, data is recorded and stored for later use and assessment (Table 3). Constant game output (ie. player position in the environment and timestep) is recorded every 0.2 seconds, and event based output occurs whenever a player deviates out of a lane, crashes into an object, or crashes into another vehicle. In addition, game video and audio capture tools are used to create a full recording of each participant's game-play. No recording of the player or audio is taken using room-based recording equipment; the recording is solely from in-game screen recording.

Constant Game Output	**Event based output**	**VirDRT**
Participant ID	Participant ID	Participant ID
Level	Level	Level
Player position (X,Y,Z coordinates)	Player position (X,Y,Z coordinates)	Player position (X,Y,Z coordinates)
System time	System time	System time
	Game duration so far	Game duration so far
	Event: lane deviation, object crash, or vehicle crash	VirDRT time ON
		VirDRT time OFF
		VirDRT HIT or MISS
		VirDRT response time

Table 3: In-game data recorded

In addition to the video and in-game data recorded, the EEG records 14 channels via the Emotive X and EmotivPRO software (Figure 29). The channels recorded are: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4.

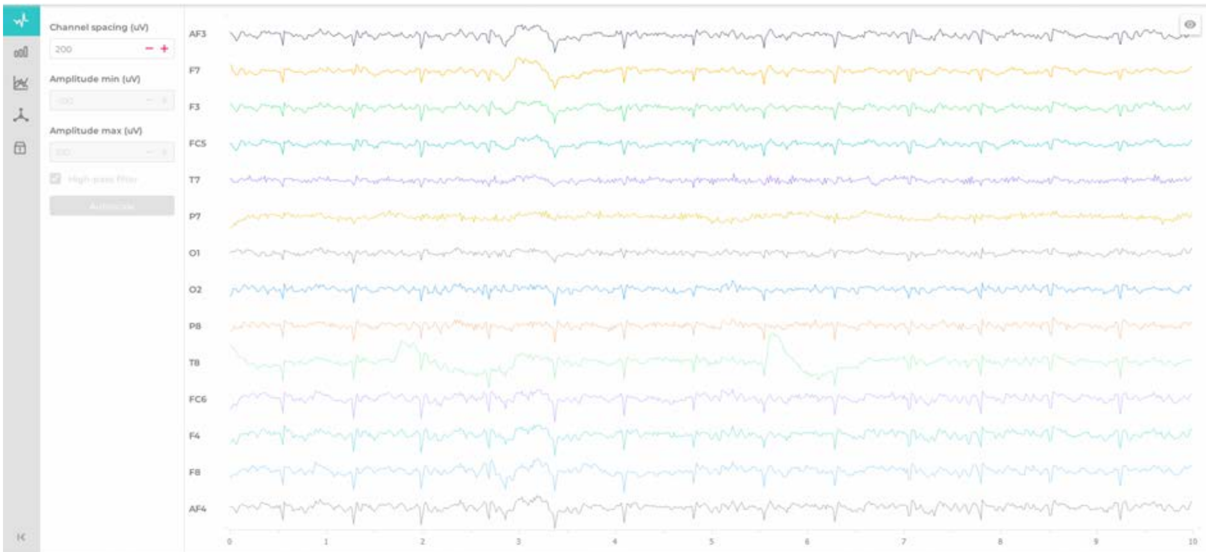

Figure 30: Sample of real-time EEG capture of 14 channels.

7 Conclusions

This game environment, the CEDG, is designed to address several questions that deal with the effect of complex virtual driving environments and tasks on cognitive load and player performance. The purpose of this is to be able to assess how these aspects interact in a game more akin to what a player, or learner, may encounter in the real-world contexts rather than less feature-rich research settings. The ultimate purpose is to garner a better understanding of the impacts that extraneous effects, task difficulties, and task stacking may have in virtual immersive simulation training applications. From this, a toolbox of difficulty adjustment parameters can be developed to better understand what, and how, in-game elements and tasks can be manipulated to affect cognitive load and performance, and thus dynamically adjust game or task difficulty to meet the needs of the player and learning goals. This is an important first step for the development of effective dynamically responsive virtual immersive simulation scenarios.

A secondary aim of the experiment environment is to assess if an accessible version of the DRT can be effective without the need for specialised equipment. This version is termed the VirDRT, and in this CEDG, it is designed to operate from a standard games console controller. When validated, and if successful, this may enable a cost effective and easy to implement method for measuring real-time cognitive load in virtual environments, and can later be transferred to a range of different game controller types (e.g. VR, steering wheels etc.). This further enables the development of dynamic virtual immersive simulation scenarios that respond to changes in player cognitive load.

Thus, the data collected from the CEDG forms the foundation for understanding how to apply techniques and methods to future cognitive-adaptive serious games and simulation training. The finding here will inform what and how to apply changes in-game and which types of changes are likely to have the greatest impact on cognitive load, challenge, and player performance.

The CEDG enables the recording of a large amount of data that has the potential to provide information beyond the scope of the research aims informing its design, and thus may be useful in other contexts. Some examples of how this data can be explored based on different ages, genders, and games experience data include (but are not limited to) how they:

- Have differing risk profiles
- Cope with virtual spatial and time reasoning
- How memory is affected in virtual environments
- Respond to different video game aesthetics
- Have differing player types and preferences

8 Acknowledgement

This work was supported by the Defence Innovation Network (DIN) NSW PhD Grant in collaboration with Bohemia Interactive Simulations, and the University of Newcastle Postgraduate Research Scholarship.

9 References

Adams, E. (2014). *Fundamentals of game design* (Third Edition ed.). United States of America: Pearson Education.

Afergan, D., Peck, E. M., Solovey, E. T., Jenkins, A., Hincks, S. W., Brown, E. T., . . . Jacob, R. J. K. (2014, 26 April 2014 - 1 May 2014). *Dynamic difficulty using brain metrics of workload.* Paper presented at the Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14, Toronto, ON, CA.

Aldekhyl, S., Cavalcanti, R. B., & Naismith, L. M. (2018). Cognitive load predicts point-of-care ultrasound simulator performance. *Perspectives on medical education, 7*(1), 23-32.

Alexander, J., Barham, P., & Black, I. (2002). Factors influencing the probability of an incident at a junction: results from an interactive driving simulator. *Accident Analysis & Prevention, 34*(6), 779-792.

Antonenko, P., Paas, F., Grabner, R., & Van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational psychology review, 22*(4), 425-438.

Barton, M. (2008). How's the weather: Simulating weather in virtual environments. *Game Studies, 8*(1).

Beede, K. E., & Kass, S. J. (2006). Engrossed in conversation: The impact of cell phones on simulated driving performance. *Accident Analysis & Prevention, 38*(2), 415-421.

Brumby, D. P., Salvucci, D. D., & Howes, A. (2009, April 4 - 9, 2009). *Focus on driving: How cognitive constraints shape the adaptation of strategy when dialing while driving.* Paper presented at the Proceedings of the SIGCHI conference on human factors in computing systems, Boston, MA, USA.

Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational psychologist, 38*(1), 53-61.

Bruyas, M.-P., & Dumont, L. (2013, 18 June 2013). *Sensitivity of Detection Response Task (DRT) to the driving demand and task difficulty.* Paper presented at the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, New York.

Cassady, J. C. (2004). The influence of cognitive test anxiety across the learning–testing cycle. *Learning and instruction, 14*(6), 569-592.

Conti, A., Dlugosch, C., Vilimek, R., Keinath, A., & Bengler, K. (2012a, 21-25 July, 2012). *An assessment of cognitive workload using detection response tasks.* Paper presented at the Proceedings of the Fourth International Conference on Applied Human Factors and Ergonomics, San Francisco, California, USA.

Conti, A., Dlugosch, C., Vilimek, R., Keinath, A., & Bengler, K. (2012b). An assessment of cognitive workload using detection response tasks. *Advances in human factors and ergonomics series. Advances in human aspects of road and rail transportation*, 735-743.

Csikszentmihalyi, M., Abuhamdeh, S., & Nakamura, J. (2014). Flow. In *Flow and the foundations of positive psychology* (pp. 227-238): Springer.

Dideriksen, A., Reuter, C., Patry, T., Schnell, T., Hoke, J., & Faubert, J. (2018). Define Expert - Characterizing Proficiency for Physiological Measures of Cognitive Workload. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2018*. Retrieved from http://www.iitsecdocs.com/

Dorn, L., & Barker, D. (2005). The effects of driver training on simulated driving performance. *Accident Analysis & Prevention, 37*(1), 63-69.

Drews, F. A., Yazdani, H., Godfrey, C. N., Cooper, J. M., & Strayer, D. L. (2009). Text messaging during simulated driving. *Human Factors, 51*(5), 762-770.

El-Nasr, M. S. (2006, 14-16 June, 2006). *Projecting tension in virtual environments through lighting.* Paper presented at the Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology, Hollywood, California, USA.

EMOTIV. (2021a, 2021). Epoc X. Retrieved from https://www.emotiv.com/epoc-x/

EMOTIV. (2021b). Performance Metrics. Retrieved from https://emotiv.gitbook.io/emotivpro-v2-0/data-streams/performance-metrics-needs-image

Feldstein, I. T. (2019). Impending collision judgment from an egocentric perspective in real and virtual environments: A review. *Perception, 48*(9), 769-795.

Fraser, K., Ma, I., Teteris, E., Baxter, H., Wright, B., & McLaughlin, K. (2012). Emotion, cognitive load and learning outcomes during simulation training. *Medical education, 46*(11), 1055-1062.

Gray, R., & Regan, D. M. (2005). Perceptual processes used by drivers during overtaking in a driving simulator. *Human Factors, 47*(2), 394-417.

Grimshaw, M. (2012). Sound and player immersion in digital games. In *The Oxford handbook of sound studies*.

Hairston, W. D., Whitaker, K. W., Ries, A. J., Vettel, J. M., Bradford, J. C., Kerick, S. E., & McDowell, K. (2014). Usability of four commercially-oriented EEG systems. *Journal of neural engineering, 11*(4), 046018.

Haji, F. A., Rojas, D., Childs, R., de Ribaupierre, S., & Dubrowski, A. (2015). Measuring cognitive load: performance, mental effort and simulation task complexity. *Medical education, 49*(8), 815-827.

Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior, 54*, 170-179. doi:10.1016/j.chb.2015.07.045

Harbluk, J. L., Burns, P. C., Tam, J., & Glazduri, V. (2013, 17-20 June, 2013). *Detection response tasks: Using remote, headmounted and Tactile signals to assess cognitive demand while driving.* Paper presented at the PROCEEDINGS of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Bolton Landing, New York, USA.

Hart, S. G. (2006, 1 October 2006). *NASA-task load index (NASA-TLX); 20 years later.* Paper presented at the Proceedings of the human factors and ergonomics society annual meeting.

Hookham, G., & Nesbitt, K. (2019, 29-31 January, 2019). *A Systematic Review of the Definition and Measurement of Engagement in Serious Games.* Paper presented at the Proceedings of the Australasian Computer Science Week Multiconference - ACSW 2019, Sydney, NSW, Australia.

Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., & Brown, J. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis & Prevention, 38*(1), 185-191.

Hunicke, R. (2005, 15 - 17 June, 2005). *The case for dynamic difficulty adjustment in games.* Paper presented at the Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology, Valencia, Spain.

Irwin, C., Monement, S., & Desbrow, B. (2015). The influence of drinking, texting, and eating on simulated driving performance. *Traffic injury prevention, 16*(2), 116-123.

Jiang, B., He, J., Chen, J., Larsen, L., & Wang, H. (2020). Perceived Green at Speed: A Simulated Driving Experiment Raises New Questions for Attention Restoration Theory and Stress Reduction Theory. *Environment and Behavior*, 0013916520947111.

Just, M. A., Keller, T. A., & Cynkar, J. (2008). A decrease in brain activation associated with driving when listening to someone speak. *Brain research, 1205*, 70-80.

Knez, I., & Niedenthal, S. (2008). Lighting in digital game worlds: Effects on affect and play performance. *CyberPsychology & Behavior, 11*(2), 129-137.

Knoll, A., Wang, Y., Chen, F., Xu, J., Ruiz, N., Epps, J., & Zarjam, P. (2011, 5-9 September, 2011). *Measuring cognitive workload with low-cost electroencephalograph.* Paper presented at the Ifip conference on human-computer interaction, Lisbon, Portugal.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review, 63*(2), 81.

Naismith, L. M., & Cavalcanti, R. B. (2015). Validity of cognitive load measures in simulation-based training: a systematic review. *Academic Medicine, 90*(11), S24-S35.

Ohly, H., White, M. P., Wheeler, B. W., Bethel, A., Ukoumunne, O. C., Nikolaou, V., & Garside, R. (2016). Attention Restoration Theory: A systematic review of the attention restoration potential of exposure to natural environments. *Journal of Toxicology and Environmental Health, Part B, 19*(7), 305-343.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist, 38*(1), 63-71.

Park, B., & Brünken, R. (2015). The rhythm method: A new method for measuring cognitive load—An experimental dual-task study. *Applied Cognitive Psychology, 29*(2), 232-243.

Ranney, T. A., Baldwin, G., Parmer, E., Domeyer, J., Martin, J., & Mazzae, E. N. (2011). *Developing a test to measure distraction potential of in-vehicle information system tasks in production vehicles*. ((No. HS-811 463)). The National Academies of Sciences, Engineering, and Medicine Retrieved from https://trid.trb.org/view/1136519

Rehnberg, O. (2021). *Static vs dynamic weather systems in video games.* (Bachelor Dissertation). Malmo University, Malmo University Publications. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:mau:diva-40220

Rossi, R., Gastaldi, M., Biondi, F., & Mulatti, C. (2012). Evaluating the impact of processing spoken words on driving: Experiments with driving simulator. *Transportation research record, 2321*(1), 66-72.

Seyderhelm, A. J., Blackmore, K. L., & Nesbitt, K. (2019, 11-15 November, 2019). *Towards Cognitive Adaptive Serious Games: A Conceptual Framework.* Paper presented at the Joint International Conference on Entertainment Computing and Serious Games, Arequipa, Peru.

Shinar, D., Tractinsky, N., & Compton, R. (2005). Effects of practice, age, and task demands, on interference from a phone task while driving. *Accident Analysis & Prevention, 37*(2), 315-326.

Standardization, I. O. f. (2016). Road vehicles — Transport information and control systems — Detection-response task (DRT) for assessing attentional effects of cognitive load in driving. In.

Stojmenova, K., & Sodnik, J. (2018). Detection-Response Task-Uses and Limitations. *Sensors (Basel), 18*(2). doi:10.3390/s18020594

Stojmenova, K., & Sodnik, J. (2018). Detection-Response Task—Uses and Limitations. *Sensors, 18*(2), 594.

Sweller, J. (2011). Cognitive load theory. In *Psychology of learning and motivation* (Vol. 55, pp. 37-76): Elsevier.

Thorpe, A., Nesbitt, K., & Eidels, A. (2019, 29-31 January, 2019). *Assessing game interface workload and usability: A cognitive science perspective.* Paper presented at the Proceedings of the australasian computer science week multiconference, Sydney, NSW, Australia.

Unity. (2019, 19 August 2019). Unity User Manual. Retrieved from https://docs.unity.cn/2019.4/Documentation/Manual/UnityManual.html

Unity. (2020, 8 April 2020). Unity Asset Store. *Standard Assets (for Unity 2018.4).* Retrieved from https://assetstore.unity.com/packages/essentials/asset-packs/standard-assets-for-unity-2018-4-32351#content

Unity. (2021). Unity Game Engine. Retrieved from https://unity.com/

Valtchanov, D., & Ellard, C. G. (2015). Cognitive and affective responses to natural scenes: effects of low level visual properties on preference, cognitive load and eye-movements. *Journal of Environmental Psychology, 43*, 184-195.

Vanetti, E. J., & Allen, G. L. (1988). Communicating environmental knowledge: The impact of verbal and spatial abilities on the production and comprehension of route directions. *Environment and Behavior, 20*(6), 667-682.

Wang, S., Gwizdka, J., & Chaovalitwongse, W. A. (2015). Using wireless EEG signals to assess memory workload in the $ n $-back task. *IEEE Transactions on Human-Machine Systems, 46*(3), 424-435.

Wiebe, E. N., Lamb, A., Hardy, M., & Sharek, D. (2014). Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. *Computers in Human Behavior, 32*, 123-132.

Attachment A – Demographic and Game Preference Surveys

The following demographic and game preferences surveys were made available via a Survey Monkey online form as well as a hard copy, depending on player preference.

DEMOGRAPHIC SURVEY:

Project Title: Real Time Cognitive Measures for Enhanced Human Performance.

Date: ______________________

Participant ID: _______________

1. What is your gender?
   ☐ Female ☐ Male ☐ Other

2. What is your age in years? ______________

3. Do you have a driver's license (including Provisional)
   ☐ Yes ☐ No

4. Do you feel well rested today, i.e. did you have enough sleep last night?
   ☐ Yes ☐ No ☐ Other__________________

5. Do you feel well enough to participate in this experiment?
   ☐ Yes ☐ No

6. Do you currently feel affected by alcohol or drugs?
   ☐ Yes ☐ No ☐ Other_______________________

7. Which is your dominant hand?
   ☐ Left ☐ Right ☐ Ambidextrous

8. What program/degree are you currently studying? ________________________________

9. What is the highest educational qualification you have completed?

☐ High School Certificate

☐ TAFE

☐ Bachelor's Degree

☐ Postgraduate

☐ Other (please specify): ______________

GAME PREFERENCES

(For questions 10 – 16 please consider gameplay across *all* electronic platforms, e.g. smart phone, PC, game consoles, etc)

10. How many games do you play weekly? Enter number of games:

Please answer question 11 if your answer to Question 10 was more than 0.

11. For each game, on average, how long do you play in a gameplay session:

☐ Less than 15 mins
☐ 15 – 60mins
☐ 1 – 2 hours
☐ 2 – 4 hours
☐ 4 – 6 hours
☐ 6+ hours

Comments: ________________________

12. How regularly do you have gameplay sessions?

☐ Frequently (I play nearly every day)

☐ Often (I play more days in a week than not)

☐ Casually (I play 2-3 days out of the week)

☐ Irregularly (Varies significantly from week to week)

☐ Sporadically (You may go several weeks without playing)

☐ Don't Play

Comments: ________________________

13. Based on how often you play, would you identify as: (check one that applies)

☐ Casual Gamer

☐ Core Gamer

☐ Hardcore Gamer

☐ I do not identify as a gamer

Comments: ________________________

14. Which platform/s have you played games on before? (check all that apply)

☐ PC

☐ Sony Playstation

☐ Microsoft Xbox

☐ Nintendo Wii/Switch

☐ Smart phone or tablet

☐ Other ____________

15. Please tick all genres of games you play across all platforms:

☐ First Person Shooter (Call of Duty, Half Life 2 etc.)
☐ Third Person Shooter (e.g. Fortnight, Uncharted,The Division etc.)
☐ Role Playing Games including MMO (World of Warcraft, Divinity 2 etc)
☐ Car Racing Games or Simulators (e.g. Gran Turismo, Dirt, Forza etc.)
☐ Casual (Angry Birds, Clash of Clans, Candy Crush etc.)
☐ Strategy Games (e.g. Civilization, Total War etc.)
☐ Sports Games (e.g. FIFA, Madden NFL etc.)
☐ Fighting Games (e.g. Street Fighter, Tekken, UFC etc.)
☐ Open World Games (e.g. Zelda, Minecraft, Grand Theft Auto etc.)
☐ Other ____________

Please answer question 16 if you have played car racing games in the past.

16. How often do you use a steering wheel or other driving related peripheral device for racing games, including arcade systems (e.g. Daytona)?
☐ Frequently (I use one on at least a weekly basis)
☐ Casually (I use one at least on a monthly basis)
☐ Irregularly (I use one at least on a yearly basis)
☐ Hardly ever (I have only ever used one a couple of times)
☐ Haven't used one
Comments: ______________________

17. Did you answer the questions in this study truthfully?
☐ Yes ☐ No

Thank you for completing this survey.

Attachment B – Post-Game Engagement Survey

The post-game Engagement Survey is provided in a Google Survey online format as well as in paper form depending on which the participant prefers.

Engagement Survey

* Required

1. **When I was playing the game, I lost track of the world around me.** *

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

2. **I blocked out things around me when I was playing the game.** *

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

3. **The time I spent playing the game just slipped away.** *

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

4. **I was absorbed in my gaming task.** *

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

5. **I was so involved in my gaming task that I lost track of time.** *

Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

6. **During this gaming experience I let myself go.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

7. **I lost myself in this gaming experience.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

8. **I was really drawn into my gaming task.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

9. **I felt discouraged while in the game.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

10. **I felt annoyed while using the game.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

11. **Using the game was mentally taxing.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

12. **I found the game confusing to use.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

13. **I felt frustrated while using the game.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

14. **I could not do some of the things I needed to do in the gaming environment.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

15. **The gaming experience was demanding.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

16. **This gaming experience did not work out the way I had planned.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

17. **I liked the graphics and images used in the game environment.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

18. **The game appealed to my visual senses.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

19. **The game was aesthetically appealing.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

20. **The screen layout of the game environment was visually pleasing.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

21. **The game was attractive.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

22. **The content of the gaming environment incited my curiosity.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

23. **I would continue to go play this game out of curiosity.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Strongly Agree

i3 Lab

24. **I would recommend playing the game to my friends and family.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

25. **Playing the game was worthwhile.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

26. **I felt interested in my gaming task.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

27. **My gaming experience was rewarding.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree

28. **This gaming experience was fun.** *
Mark only one oval.

	1	2	3	4	5	6	7	
Strongly Disagree	◯	◯	◯	◯	◯	◯	◯	Strongly Agree