**The effectiveness of an online driving simulator:**

**Transfer effects of driving skills examined with the Tweak-Finder Model**

Maran Voskes (s2005549)

Master Thesis Human Factors & Engineering Psychology

Faculty of Behavioural, Management, and Social Science, University of Twente

1st Supervisor: Martin Schmettow, Prof. Dr.

2nd Supervisor: Jorrit Kuipers, Ir., founder Green Dino

**Abstract**

Within the first months after licencing, new drivers experience their highest crash risk, which is mainly due to inexperience with the driving task and lack of feedback after licensing. Driving simulators demonstrate to be a potential learning resource to gain a larger variety of driving experience within the training period. Within a risk-free environment, trainees can experience a high variety of traffic situations and can explore risky situations that cannot be trained on the road. This study examines the effectiveness of an online simulator developed by Green Dino, a pioneer in designing driving simulators. The online simulator allows students to train their driving skills at their own personal computer at any time they want. This allows students to gain experience in an easy, safe and cost-effective way. By exploring transfer effects between on-road driving and simulator driving, the effectiveness of the online simulator will be determined within a two-part study.

The first part of the study focuses on three datasets: (1) an unconstrained dataset gathered by the online simulator in the original way, (2) a highly controlled dataset gathered in a physical driving simulator, and (3) a medium-controlled dataset gathered in the online simulator within an experiment. These datasets were explored for learning curves and used the Tweak-Finder Model (TFM) to estimate learning curves. The unconstrained dataset demonstrated no learning curve indications, but the other two did, implying that more controlled datasets are better suitable for learning curve analyses. Also indications for the symmetry of transfer assumption between the physical simulator and on-road driving were found in the highly controlled dataset.

Based on the results of the first part, an experiment was designed for the second part of the study. The online simulator was updated and 33 students were asked to perform a driving training of approximately 5 hours, including 3 tasks (taking turns, roundabouts, and intersections) each performed 20 times. Indications of learning curves were explored and when positive, learning curves were run. Differences in amplitude parameter were analysed to demonstrate possible transfer effects of driving skills.

Results showed that the simulator data was not able to produce learning curves, which is possibly caused by a mismatch between the online simulator data metrics and the TFM or still a too high level of freedom. Outcome variable workload was able to demonstrate learning effects, indicating that trainees learn to decrease their mental workload within the simulator. If learning curve analyses are desired, it is recommended that the online simulator reconstructs its metrics in a way that matches the requirements of the TFM model.

*Keywords:* driving training, (online) driving simulator-based training, Tweak-Finder model, symmetry of transfer assumption

**Table of Contents**

# 1. Introduction

Within the first six to twelve months after licencing, drivers experience their highest crash risk (Mayhew, Simpson, & Pak, 2003). This is mainly caused by inexperience of the driver (Beanland et al., 2013) and the loss of external feedback from the driving instructor (Kuipers & Wieringa, 2014). Analyses of police reports by McKnight & McKnight (2003) conclude that this inexperience is largely due to failures in visual scanning, attention maintenance, and speed-management, which are responsible for around 87.1% of the crashes among young drivers. Especially hazard perception is considered to be a critically important safety-related driving skill. Inexperienced drivers tend to scan less broadly and move their fixations less than experienced drivers and therefore are more inclined to fail at detecting risks on time.

Driving training intends to train unlicenced drivers the required skills for driving safely, which includes vehicle handling skills, but also cognitive skills like hazard perception. However, reviews on the effectiveness of standard driving training are uniform in concluding that there is no reduction observed in the crash rates among newly licenced drivers (Pollatsek et al., 2011). Beanland et al. (2013) state that some evidence demonstrates that most of the unsafe driving behaviour is a result of factors like overconfidence, ignorance, and poor hazard perception, which could potentially be addressed during driving training. Furthermore, Weiss et al. (2013) state that traditional driver training is not able to provide the means required to gain more experience in the driving task. It is only to a small extent possible to control the novice driver's exposure to different traffic situations within on-road driving training (Pollatsek et al., 2006), preventing them from gaining experience in various traffic scenarios.

This indicates that driving training could be improved by enabling trainees to gain more experience within the training process and so cope better with the first months of higher crash risk. Training in driving simulators could be a solution. Within a driving simulator, trainees can train their skills within a risk-free environment in which they can gain experience in a high amount of different traffic situations (Kappé & Van Emmerik, 2005; Käppler, 2008; SWOV, 2019).

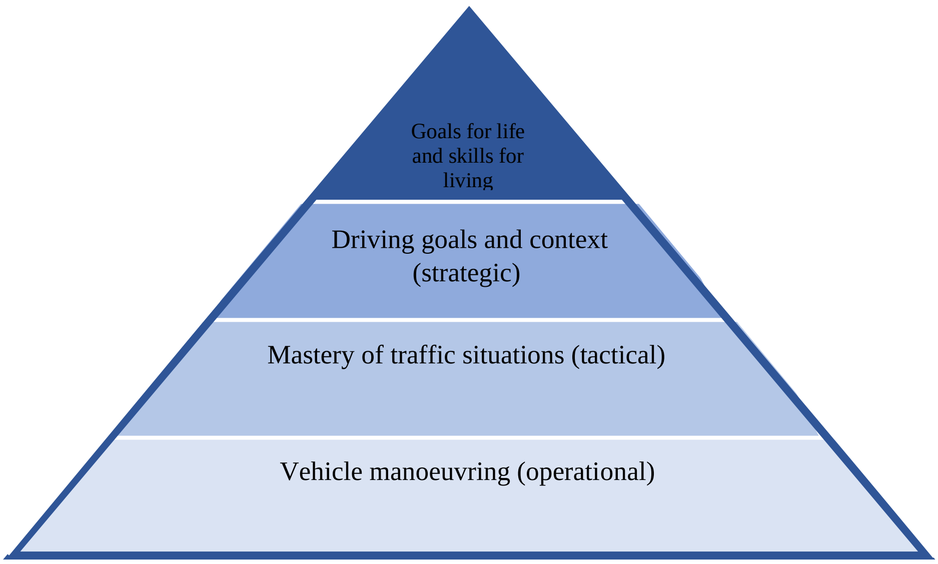
This study will focus on acquiring driving skills and gaining experience within a computer-based online driving simulator. To allow driving trainees to practice their driving easily at home, Green Dino (a driving simulator developer, https://www.greendino.nl/) designed a computer-based driving simulator in which students could partake driving lessons at home on their own computers. It will be examined whether this simulator is effective in aiding driving trainees in their acquisition of driving skills.

### 1.1 Current driving training

The primary goal of current driving training is to learn new drivers to drive safely. By providing the basic knowledge, attitudes and skills of driving, drivers are intended to learn how to deal with various (risky) traffic situations (Beanland et al., 2013, NHTSA, 1994). Therefore, drivers should understand that safe driving goes beyond being able to control the vehicle and also includes cognitive skills such as hazard perception skills and risk awareness. The Goals for Driver Education (GDE) matrix identifies a hierarchy with four levels that need to be addressed in driver education to be considered effective in forming safe drivers: vehicle manoeuvring (operational), mastery of traffic situations (tactical), driving goals and context (strategic), and goals for life/skills for living (Hatakka et al., 2002) (Figure 1).

**Figure 1**

The Goals for Driver Education (GDE) framework adapted from Hatakka et al. (2002). The operational level contains basic skills for controlling the vehicle such as braking, steering, and switching gears. The tactical level includes interaction with objects and other traffic users (e.g., manoeuvring around obstacles and merging into traffic). The strategical level comprises the route and time of driving that are chosen by the driver. Goals for life and skills for living contain the control over how lifegoals and personal tendencies, such as peer pressure and sensation seeking, affect driving behaviour. The framework operates as a hierarchy ranging from basic operational vehicle driving skills to higher-order skills which means that operational skills should be developed sufficiently to support executing higher-order skills (Voskes, 2021).



However, within most driving education programmes, the main focus is on the lowest level of the matrix (vehicle manoeuvring) and traffic rule knowledge, not properly covering the higher-level skills in practical training (Beanland et al., 2013; De Winter et al., 2009; Dols et al., 2001; Pollatsek et al., 2011). Additionally, it is unknown how much practice these higher-level skills require to develop in safe driving behaviour (Simons-Morton & Ehsani, 2016). This lack of training the higher levels of the hierarchy is primarily caused by the difficulty of systematically handling them within practical training (Dols et al., 2001). The driving environment (i.e., roads, traffic, weather, etc) is highly uncertain and risky situations are scarce and usually avoided, which offers less opportunities to develop higher level skills like hazard perception skills. Furthermore, overload of cognitive capacities by vehicle control of novice drivers might impede development of these cognitive skills.

Simons-Morton & Ehsani (2016) suggest that for driving training to be more effective, novices should be exposed to more complex driving conditions and focus more on higher level skill development. Simulators are a potentially useful learning resource that can facilitate this process. They can expose the trainees to various traffic scenarios and so allow trainees to develop driving skills within a safe, risk-free environment.

### 1.2 Driving simulators

In the Netherlands, over 100 simulators are currently used within driving training to facilitate the development of driving skills (Kappé & van Emmerik, 2005). Simulator trainings are usually implemented in one of two ways: replacing the first on-road driving lessons with simulator lessons or integration within the driving training in which particular tasks, such as driving at a roundabout, are first trained in the simulator and later on the real road. These simulator lessons can potentially provide experience over a significant portion of the driving task, including sensory-perceptual, psychomotor, and cognitive skills (Allen, Cook, & Rosenthal, 2001).

Simulators have several advantages that are impossible to incorporate in on-road training and have the potential to elevate the current driving training curriculum. Firstly, simulators can offer a broader and more various exposition of traffic situations (Jamson, 2011; Käppler, 2008; SWOV, 2019). Many different traffic scenarios with many educational moments can be provided to the student in a brief period. Additionally, situations that happen only scarcely, but which are essential for safe driving, can be trained within the simulator without being dependent on encountering the right situation on the road. Secondly, certain traffic situations and driving skills can be repeated unlimitedly when a student seems to have difficulty learning them (SWOV, 2019; Van Emmerik, 2004), improving the individual adaptability of driving training (Kappé & Van Emmerik, 2005; Käppler, 2008). Task demands can be decreased so that it is possible for the student to solely focus on a specific driving skill (ST Software, 2010). Thirdly, a simulator has the possibilities to demonstrate how to act in a certain situation, either by showing how they are expected to act by a general example (SWOV, 2019), by replaying the driving activities of the drivers themselves (Käppler, 2008) or by projecting visual cues while driving (Vlakveld, 2005). Fourthly, simulators offer a safe practice environment in which errors are not coming with health-risks (Jamson, 2011; Kappé & Van Emmerik, 2005; Käppler, 2008; SWOV, 2019). Dangerous and difficult traffic situations or environmental circumstances, such as dense fog or a near-crash experience, can be trained without posing severe risks to the trainee and other road users. Furthermore, error-training can be provided without safety risks (Ivancic & Hesketh, 2000). Within error-training, trainees are forced to make errors to learn which strategy works best and how to recover from their errors. Lastly, driving performance can be assessed objectively and accurately in a standardised fashion (De Winter et al., 2009; Käppler, 2008; SWOV, 2019). The simulator can present trainees with the same scenarios, including the same traffic density and environmental circumstances, allowing for better performance comparisons.

Nevertheless, there is an ongoing discussion identifiable between researchers about the utility of driving simulators as training devices (Pollatsek, 2011). These discussions mostly concern topics such as fidelity, transfer, and content. These topics will be discussed in the respective order, followed by a description of the online driving simulator of Green Dino.

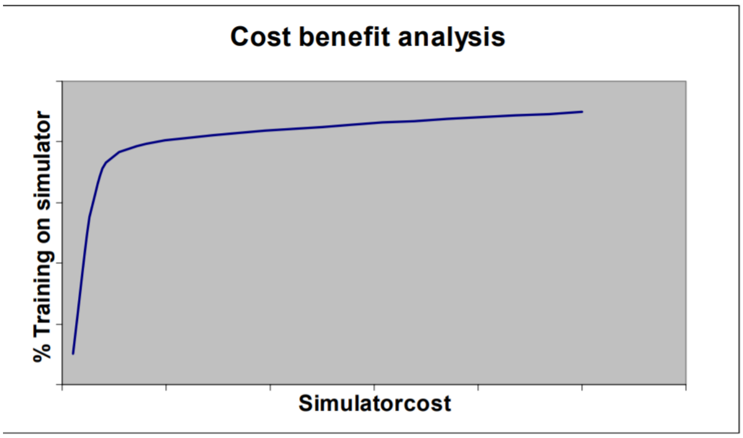
### *1.2.1 Fidelity*

Fidelity is defined as the realism or representativeness of the simulation (Allen, Park, & Cook, 2010). Based on the hardware used, three categories within fidelity can be identified: low, medium, and high-fidelity simulators (Caird & Horrey, 2011). Low-fidelity simulators are mostly equipped with simple components such as a single computer screen and basic controls. They have either a steering wheel and pedals or game controls and lack a motion system. Medium-fidelity simulators also lack a motion system but offer a wider field of view by projecting the simulation around the driver. High-fidelity simulators posses a motion system, presenting feelings of motion while driving. Additionally, full vehicle cabs are set up around the base. High-fidelity simulators are assumed to represent the highest form of realism possible and is therefore presumed to be the most effective type of simulator (SWOV, 2010). However, software seems to play a more essential role for the perceived realism of the simulation (Kappé & Van Emmerik, 2005; Kappé, Van Winsum, & Wolffelaar, 2002). Furthermore, the impact of fidelity on training appears complex and dependent on the training goal and didactics used.

Vlakveld (2005) states that simulations only provide a representation of the reality, not approximation itself, mostly caused by lack of motion. Some research on the topic show conclusions confirming this statement. For example, Jamson (2011) states that high-fidelity simulators are required for achieving high quality simulator training of specific driving tasks and that low- and medium-fidelity simulators can only be used for relatively unchallenging tasks. However, the majority of research supports also the use of medium- and low-fidelity simulators but note that there are some limitations. The ELSTAR project (Kappé et al., 2002) concluded that a large proportion of driving tasks, such as vehicle handling and traffic participation, can be learned with a low-fidelity simulator, but that it might be impossible to do the entire driving training in a simulator (Figure 2). Tasks likely not possible to learn in a simulator (e.g., negotiating fast curves and full-stop braking) are more cost-effective to train on the road. They also indicate that copying on-road driving lessons for simulator training does not take the full potential out of the unique capabilities of the simulator, such as the possibility to replay, react and preview driving skills. Allen et al. (2011) demonstrated that low-fidelity simulators might facilitate training of hazard perception skills by showing that novice drivers receiving training were able to avoid crashes better than non-trained novices. Also, Kappé and Van Emmerik (2005) state that basic skills and procedures on vehicle operation and traffic participation, but also cognitive skills like hazard perception and situation awareness can be trained sufficiently in low-fidelity simulators. They also stress that a very realistic representations can even complicate learning for novice drivers since the focus is removed from the core of the skill (Vlakveld, 2005). Lastly, non-significant differences between simulator measures of accident and graduation rates indicate that fidelity is not an essential element in simulator-based training programmes (Park et al., 2005). Participants in the low-fidelity simulator performed as well in terms of avoiding collisions and meeting standard performance criteria for graduating the training programme as participants in the high-fidelity simulator.

**Figure 2**

Curve showing that most driving tasks do not require a high-fidelity driving simulator (Kappé et al., 2002)




### *1.2.2 Transfer*

Transfer of skills concerns the question whether driving skills learned in the restricted conditions of the simulator will also be evident in on-road driving where a higher unpredictability of circumstances is present (Groeger & Banks, 2007). A positive transfer from prior learning allows the student to use learned skills in situations not experienced before. However, Groeger and Banks (2007) state that it is highly unlikely that positive transfer happens in the driving domain. They suggest that prior knowledge is situation-specific, that people do not learn from abstract principles, and that psychological models are inadequate to describe performance, preventing transfer between situations. Especially the degree of overlap between learning and transfer contexts are stressed as important factors. Groeger and Banks (2007) conclude by stating that “transfer to more novel circumstances, which would be sufficient to enable appropriate more or less instantaneous reactions, as might be required in hazardous situations, does not take place” (p. 1261). However, this paper focused on transfer within on-road driving situations and disregards the possibilities of a simulator to provide the student with a high number of various traffic situations.

Other research papers have gathered more promising results regarding transfer between on-road driving and simulator driving. Lintern (1991) explained that the practice task does not have to be identical to the transfer task. Critical perceptual similarities are more important than physical similarity for evoking prior knowledge regarding behaviour in the situation. De Winter et al. (2009) demonstrated this by showing a positive effect of transfer between simulator training and graduating the driving exam. Students that followed simulator training had a 4-5% higher chance of passing the driving test than students that only followed on-road training. Furthermore, Vlakveld (2005) showed that although immediate transfer was poor between simulator and on-road driving, novices managed to learn very rapidly to manoeuvre the real car the first time. Therefore, practising different traffic scenarios in a simulator might broaden the mental models of novice drivers, subsequently providing them with more experience before driving independently on the road.

#### *1.2.3 Content*

As mentioned above, simulators are nowadays mainly implemented in the beginning of driving training to get used to driving, or in between on-road lessons. However, there are many differences between the content of the lessons. Some simulator lessons solely contain training on operational skills whereas in other lessons the focus is merely on cognitive driving skill acquisition.

Research also disagrees in the content used for simulator-based driving training. Groeger and Banks (2007) argue that it is unlikely to develop cognitive skills within a simulator since the transfer from simulator to real driving is too large. Contradictorily, Wheeler and Triggs (1996) state that basic psychomotor skills are unlikely to develop due to their dependence on feedback in the dynamic environment. Nevertheless, other research shows that it is possible to train both these operational and cognitive skills within a simulator. For example, Pollatsek et al (2011) mentions that the recognition of a potential hazard is often simple and can be trained with repeated practice, which is easily done with simulators. Furthermore, research suggests that low-cost PC-based simulators have the potential of offering training in cognitive skills, such as situation awareness and risk perception, required for safe driving (Allen et al., 2001; Allen et al., 2007; Divekar et al., 2016; Pradhan et al., 2009). Especially interesting is the finding of De Winter et al. (2007), who state that lower fidelity simulators can be used to train and develop cognitive skills that newly licenced drivers usually develop over time while driving independently. This indicates that it is possible to gain the experience and develop cognitive skills within a simulator that drivers are lacking in the first six to twelve risk-prone months after licencing.

### *1.2.4 Online driving simulator*

Green Dino (www.greendino.nl), a pioneer in designing driving simulators, developed an online driving simulator in which students can train their driving skills at home on their own computer (Figure 3, https://online-rijlessen.virtual-reality-lms.com/). This online driving simulator is based on the demonstrated to be effective physical driving simulator training produced by Green Dino (De Winter et al., 2009). More than forty driving lessons are available, ranging from taking turns to entering a highway, divided into 3 categories: beginner, advanced, and specialist. Additionally, several driving exams are presented. Different from the physical simulator in which students are provided with a specific lesson, students can choose the lessons they desire to do and can stop at any moment they want within the online simulator. The car is controlled by the computer mouse and keyboard keys and therefore focusses on developing procedural and cognitive driving skills. After each completed lesson, students receive a performance score between 1 and 10 which informs them about how well they had performed the task. An adaptive feedback system provides the students with spoken and written feedback during the lessons. Trainees start with a high amount of extensive feedback which guides them through the driving lesson (level 1). When the trainee shows improvement, the feedback system provides fewer and less extensive comments, and the students are only guided where needed (level 2). The moment the trainees are able to perform the task almost independently and show that they have the knowledge about the procedure, the feedback system stops with providing feedback regarding this specific task (level 3). However, when the student shows deteriorating performance, more feedback is provided again (See Figure 4 for an overview of the feedback levels).

**Figure 3**

Interface of the online driving simulator of Green Dino. The left side mirror view field has been opened by the respective trainee, but can also be closed. Additionally, the right side mirror view field can be opened and closed.

**Figure 4**

Overview of the feedback levels used in the online driving simulator. Generally, when trainees develop skills, they enter a higher feedback level and therefore get less feedback. If driving skills deteriorate, trainees go back to a lower feedback level and thus get more feedback again.


The online simulator saves several measures that assess the driving style (safety scores) and the driving performance (task scores) of the trainees within lessons. These driving performance measures are divided into scores that represent the overall score for the entire lesson (OverallTaskScore) and in scores that demonstrate the performance on a specific task (TaskScore). This allows students and driving instructors to see how proficient the student performs the variety of lessons and what tasks are most relevant to focus on within training.

Safety scores represent how safe a student has performed a driving task but are independent of task scores of the driving task. This means that low safety scores are not an indication of low task performance, because the trainee might be capable of performing the task well, but because of unsafe behaviour (e.g., speeding, not being attentive, etc) is inhibited to show this capability. These safety scores are used to determine the driving style of the student and do not measure learning. The safety scores for all tasks start at the maximum score of 10 and decrease when the trainee shows unsafe driving behaviour. This means that a safety score of 10 can demonstrate two possible outcomes: the driver has driven perfectly safe, or the driver did not encounter a task related to this safety element. Safety scores are person specific scores and do not show a comparison to a reference group.

Task scores represent how well a specific task within the lesson was performed and are represented as percentile scores. Individual task scores are based on the comparison with results of the average student. Data of more than 10.000 students that performed the entire driving training on a physical simulator was gathered and formed a reference score for each specific task. So, the final task score represents how much better the trainee performed compared to this reference group and is therefore not determined by the number of errors made. To illustrate, a task score of 5.5 means that 55% of the baseline group made more mistakes than the respective trainee, which is considered to be a mediocre score. A task score of 9.2 means that 92% of the baseline group made more mistakes, and the trainee is considered to perform very well on this task. The OverallTaskScore represents the general performance within a driving lesson: all individual task scores are summed and the average is retrieved.

Within the knowledge of the researcher, there is no research involving this type of simulator: an online driving simulator without actual driving controls. There is research done on computer-based e-learning that includes videos of traffic situations in which students had to learn cognitive skills from interacting with the videos (Fisher, Pollatsek, & Pradhan, 2006; Pollatsek et al., 2006; Pradhan, Fisher, & Pollatsek, 2006; Regan, Triggs, & Godley, 2000; Wang, Zhang, & Salvendy, 2010). These studies all showed positive results regarding cognitive skill (e.g., hazard perception and situation awareness) development. However, the students were not actually operating the car in any of these computer-based e-learnings as they do in the online driving simulator of Green Dino. Therefore, this simulator is considered to be a combination of the computer-based e-learning and a low-fidelity simulator, benefiting from the advantages of both if applied effectively.

### 1.3 Models of learning: acquiring driving skills

Driving is defined as a complex activity containing a highly dynamic environment. Perceptual motor skills, procedural skills, and cognitive skills have to be time-shared to constantly attend on managing the vehicle while identifying and mitigating potential safety threats (Groeger & Banks, 2007; Simons-Morton & Ehsani, 2016; Van Emmerik, 2004). Therefore, both procedural skills (e.g., vehicle manoeuvring and manipulation of the vehicle controls) and cognitive skills (e.g., situation monitoring, hazard perception, and response planning and execution) have to be acquired to drive safely (Beanland et al., 2013).

Fitts and Posner (1967, as cited in Groeger & Banks, 2007) distinguish three phases within general skill acquisition: the cognitive, associative, and autonomous phase. In the cognitive phase, performance is slow and error-prone. Many cognitive capacities are required to fulfil the task. With practice, learners gradually enter the associative phase in which performance is more accurate, gross errors are eliminated, and patterns or sequences of performance elements are beginning to emerge. However, when cognitive demands are high, performance can be largely deteriorated. In the final autonomous phase, performance is reliable, fluent and efficient, and cognitive load is low. Within this last phase, drivers are thought to develop mental models allowing routine, non-executive functions to control established (over-learned) skills. Hall and West (1996) suggests that the basic vehicle controlling skills and the basic traffic rules can already reach the autonomous phase after 15 hours of driving. However, this does not imply safe driving capabilities after 15 hours, since cognitive driving skills need longer to develop and require a lot of practice and experience (Dols et al., 2001).

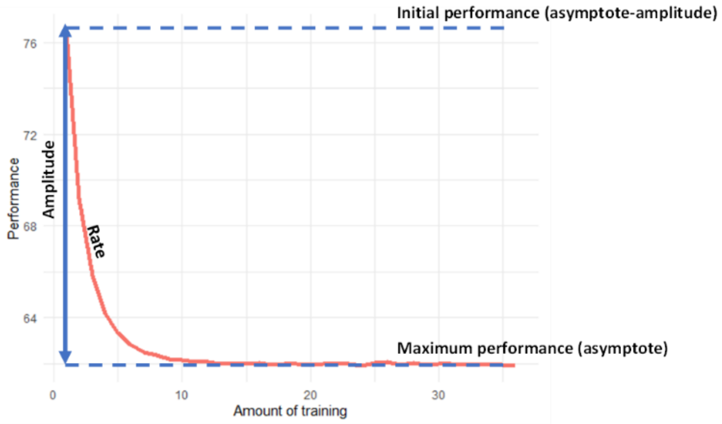
Anderson (1982) states that many more hours of learning and practise are required to acquire any significant cognitive skill. He developed a theory on the acquisition of cognitive skills which can be applied to learning driving skills (Vlakveld, 2011). He divided cognitive skill acquisition in three stages: the declarative, knowledge compilation, and procedural stage. In the declarative stage, performance is relatively unstable since possible strategies are tested and rejected, and focus is consciously on isolated components of the driving task. Once the knowledge compilation stage is reached, these isolated components get chunked together. Associations between action patterns in familiar conditions have become stronger and can be applied to situations recognised as being similar. After months to years of practicing, the procedural stage is attained and separate chunks including components of the task performance have formed procedures captured in mental models. These procedures can be executed in a seemingly effortless manner without much awareness of the separate elements of the skill. However, at some moments (e.g., when the environment is highly adaptive), attentional monitoring is required to check whether the procedure is applied well. Therefore, it is stated that even though a large amount of situational driving tasks (i.e., situation monitoring and hazard perception) can reach the procedural stage, they are never fully automated due to the constantly changing traffic environment (Kappé & Van Emmerik, 2005).

### 1.4 Learning curves

Learning curves can be used to examine, visualise, and predict skill acquisition. Learning curves enable quantitative exploration of an individual's learning process of skills over time. According to Heathcote, Brown, and Mewhort (2000), learning is represented by an exponential law of practice and consists of three individual-specific parameters: amplitude, rate, and asymptote. The amplitude demonstrates the individual's amount of improvement. The amplitude is generally higher for individuals who have a small amount of previous experience because they have more room to improve their performance. The rate represents how fast an individual learns. Maximum performance is represented by the asymptote: performance is stabilised and more learning is unlikely within this specific training. Figure 5 displays a traditional learning curve pattern.

**Figure 5**

Representation of a traditional three-parameter exponential curve learning curve (red line). The x-axis represents the amount of training the trainee has had and the y-axis the performance level. The rate displays the speed of improvement, the amplitude the amount of improvement, and the asymptote the maximum performance level within that specific training.




Several master theses of students of the University of Twente already explored learning by applying these learning curves within the surgical domain, such as bronchoscopy (Küpper, 2018; Westerhof, 2018) and laparoscopy (Arendt, 2017; Kaschub, 2016; Weimer, 2019). They investigated surgical skill acquisition of inexperienced participants and compared individual learning curves to support development of surgical training. With their results, they could visualise and predict the performance of individuals and potentially talented surgeons could be discovered. Furthermore, differences in performance were used to develop adaptive training methods, providing individual-specific training to optimise learning.

### *1.4.1 The Tweak-Finder Model*

Schmettow (n.d.) developed a new learning curve model: the tweak-finder model of building skills (TFM). This model describes acquiring skills as exploring a pool of undiscovered tweaks within task execution to improve one's performance and is based on three assumptions: (1) the pool of tweaks is finite, (2) finding a tweak is irreversible, (3) every tweak has a fixed probability to be found. Because there is no infinite number of possible tweaks, the pool of tweaks diminishes over time. Subsequently, tweaks will be harder to find within the respective training, leading to a decrease of learning per exercise. This results in a learning process equal to the learning curve presented in Figure 5. Students start their training at their initial performance level which depends on their previous experience related to the task. For less experienced students, the amplitude is high and the learning rate is large in the beginning of the training since the trainee has a higher likelihood to find tweaks in the initially large pool of undiscovered tweaks. For more experienced students, this amplitude and learning rate are lower. They have already discovered more tweaks and have more difficulties finding the last remaining tweaks for the task. After a certain amount of training, performance stabilises and the maximum performance level is reached when (nearly) all possible tweaks are found and applied.

This new theory corresponds with the previously mentioned theories about skill acquisition by Fitts and Posner (1967) and Anderson (1982). The first two phases of Fitts and Posner (cognitive and associative) are both represented within this model. In the beginning of the learning process when many tweaks are still undiscovered, more cognitive capacities are required to perform the task since the tweaks cannot be used yet to ease the task. Once more tweaks are discovered and can be applied while performing the task, performance improves and gross errors are eliminated. Performance starts to stabilise and trainees do not demonstrate learning effects anymore. However, reaching the maximum performance level by discovering most of the tweaks does not imply that the autonomous stage of the Fitts and Posner model is entered. The maximum performance is an individual specific measure that demonstrates the best possible performance for that specific task within that context. However, this does not directly mean that the task can be performed autonomously in the same way as Fitts and Posner refer to in what they define as the autonomous phase. Moreover, Anderson's (1982) stages can be recognised in the TFM. Within the declarative stage of Anderson's theory, many different strategies are tested and rejected, which might explain the fast discovery of many tweaks within the beginning of the learning process. Once more tweaks are discovered, associations between them can be made that form parts of a useful strategy as is done in the knowledge compilation stage. Trying completely new strategies is not necessary anymore since the discovered tweaks form an effective strategy. As a consequence, performance gets more reliable and forms a stabilised performance level like in the TFM. The task can be performed in a more effortless manner and less awareness of the separate elements of the procedure is required.

However, also a large difference can be detected between the TFM and the theories of Fitts and Posner (1967) and Anderson (1982). As already hinted by the comparisons above, it is difficult to clearly define when which phase of the older theories are entered when trying to explain skill acquisition in terms of the TFM. Whereas these prior theories see skill as an amorphous mass that follows an exact path when being acquired, the TFM composes the skill in identifiable elements (tweaks) that can be discovered at any moment within the training, depending on the number of undiscovered tweaks that are left.

To calculate the individual learning curves and so display the learning process, the TFM uses the LACY (log-scale amplitude, catch rate, asymptote) formula: $P_t = exp\ Asym + exp\ Ampl + logit^{-1} (1\text{-}Ctch)^t$. Continuous variable *T* represents the trial number within the respective training. The asymptote and amplitude parameters share nearly the same definition as in the model of Heathcote et al. (2000). The function of the catch rate parameter (Ctch) is similar to the rate parameter of Heathcore et al.'s (2000) model but differs in its definition. Whereas Heathcore et al. define the rate parameter just as the speed of learning, the TFM defines the Ctch parameter as the chance to catch a tweak within the training. These three parameters have all a unique story and functionality.

Firstly, the asymptote parameter demonstrates the proficiency of the trainees' performance once all possible tweaks have been found. This provides the possibility to detect the stabilised performance of the best performing trainees and avoids the risk to base conclusions on one best result achieved within the training. Moreover, it aids in detecting the students that need more training or to determine when the training is not effective anymore. When trainees are not able to stabilise their performance within the training task, more training or a different type of training might be required to reach this maximum performance level. Moreover, when the maximum performance is reached, the training is not as effective anymore as in the beginning of the learning process. The specific training task can be stopped and the trainee can switch to another training to use valuable training time to its maximum. Lastly, retention effects can be explored with the asymptote parameter. Once the asymptote is reached, it is expected that the trainee will remain performing on that level. However, it might be possible that performance deteriorates after some time in which the task is not performed. As a consequence, the trainee performs worse than the maximum performance that was reached previously and the asymptote has to be reached over again. By comparing the learning curves of training sessions at different moments in time within individuals, retention effects of the acquired driving skills can be determined.

Secondly, the amplitude parameter shows the magnitude of the tweaks found and can aid in discovering trainees that already have gathered some experience on the respective task. A low amplitude indicates that no learning is happening which can be due to personal factors such as previous experience with the task or natural talent. This means that individuals with a low amplitude in their learning curve have more experience with the respective task than individuals with a high amplitude. Therefore, the amplitude parameter can be used to determine transfer effects from for example other training tasks or other life experiences in different situations. However, a low amplitude might also be caused by external factors like a too simplistic training task.

Lastly, the catch rate parameter predicts how fast trainees are able to catch all the tweaks in the pool. It can be predicted how much training is required for the respective trainee: Some students are able to catch more tweaks in a shorter training duration whereas others might need more training to discover all the tweaks. To illustrate, an individual starting with a larger set of undiscovered tweaks (higher amplitude) can reach approximately the same maximum performance (asymptote) at the same moment as someone starting with a smaller set of undiscovered tweaks when he or she has a higher chance to catch a tweak (rate).

**1.4.1.1 The symmetry of transfer assumption.** The TFM states that skill acquisition is composed to identifiable elements which are called tweaks. This is in contradiction to prior theories (e.g., Fitts and Posner and Anderson) that see skill as an amorphous mass. Moreover, the TFM assumes that finding tweaks is irreversible and that particular tweaks show overlap in tasks. Therefore, it is expected that once a tweak is found in a certain task, this tweak could be used as well in other tasks where application of the tweak is efficient. This notion makes transfer a symmetrical effect:

transfer of skills from domain A to domain B signifies transfer of skills from domain B to domain A. As a result, a way to determine the transfer from simulator to on-road driving performance is by establishing the transfer from on-road to simulator driving performance.

Think for example of somebody that has been riding the bike since he or she was a child. During these years, this person has already gained a lot of knowledge about various traffic situations and traffic rules by exploring traffic related tweaks. Once he or she decides to start driving, these previously discovered tweaks during these cycling years that are also useful in the car are likely used during driving to facilitate performance. During the driving, new tweaks will be discovered that ease the driving task and can be combined with the tweaks discovered while riding the bike. However, new tweaks also useful for cycling will be discovered as well within driving the car. Once the person is riding his or her bike again, the newly discovered tweaks of the driving can reversely be applied to the cycling performance again. So, when a learning element is part of another task, transfer of certain elements happens between these tasks. This way, discovered tweaks transfer back and forth between these two tasks and facilitate the learning process within them both.

Experienced drivers are shown to have more detailed elaborated mental models and a more refined operational skill set compared to inexperienced drivers (Shallice, 1998; Simons-Morton & Ehsani, 2016; Vlakveld, 2011), resulting in better and safer driving performance for experienced drivers (Beanland et al, 2013; McKnight & McKnight, 2003). Since transfer is anticipated to be symmetrical, it is expected that when this difference in experience transfers from on-road to simulator driving, this transfer is also apparent from simulator to on-road driving performance. This expectation is defined as the *symmetry of transfer assumption* and is within the knowledge of the researchers not explored before. This assumption will be examined by using the factorial amplitude model of the TFM which focusses on differences between amplitudes among individuals. As mentioned above, the amplitude shows the amount of learning that has happened. It is expected that the amplitude of experienced on-road drivers will be smaller than the amplitude of inexperienced on-road drivers since they already discovered more tweaks that facilitate their driving performance and therefore have less to learn left.

### 1.5 The present study: an overview

The present study aims to examine the effectiveness of a computer-based online simulator by means of (1) establishing the requirements for estimating learning curves on driving skill acquisition within a simulator, (2) exploring transfer effects of acquired driving skills, and (3) testing the LACY formula of the TFM. This will be approached by a two-part study. Firstly, existing datasets will be explored for learning curves to determine the required level of control and suitable performance measurements for learning curve analyses (Chapter 3). Phase 1 describes the results of an existing dataset gathered in the original manner from the actual online simulator retrieved from Green Dino and will be explored for learning curves. Since this dataset is relatively unconstrained – the data contains many variables with highly adaptive factors intervening – two additional previously gathered datasets will be explored for differences: one highly controlled dataset (Phase 2) and one somewhere in the middle of this continuum (Phase 3). The highly controlled dataset was gathered by Voskes (2021) within a physical driving simulator located at the University of Twente, aiming to compare driving performance between two conditions (accuracy-based training and speed-based training). The other dataset was gathered by Van Wijk (2020) within the online driving simulator of Green Dino. However, this dataset was obtained in an experimentally controlled manner and is therefore more controlled than the unconstrained dataset of the online simulator. Results of the analyses will be used as advice about the existing online driving simulator and the data it is gathering and will be used as basis for the second part of the study.

Secondly, an experiment will be performed within the online driving simulator (Chapter 4). The symmetry of transfer assumption will be explored by comparing the amplitude parameter of a sample of drivers that form a continuum of driving experience. Results will contribute to the establishment of the effectiveness of the online driving simulator, but will also demonstrate the first indications of the effectiveness of the TFM model for representing driving skill acquisition.

## 2. Data analysis

The initial data analysis plan was the same for all datasets in Chapter 3 and 4 and was performed in statistical programme Rstudio and the programming language R (version R 3.4.4.) using packages 'brms' (version 2.15.0) and 'bayr' (version 0.9.4). However, after exploring the datasets, this initial plan was not always feasible. Any deviations from this data analysis plan are mentioned in the respective chapters.

First, variables were added and mutated when required, and subsequently, the raw data was explored for signs of learning curves. When learning curve patterns were shown, statistical models were estimated, and the suitability of the fit was analysed.

This study uses the tweak-finder model of building skills from Schmettow (n.d.) to compute multi-level learning curves and analyse the gathered data. The aforementioned parameters (i.e., amplitude, rate, and asymptote) are included in this model, forming the following ACY formula: $P_t = Asym + Ampl + (1\text{-}Ctch)^t$ (see Chapter *1.4.1 The Tweak-Finder Model* for more detailed information regarding the parameters). Since random effects of multi-level models are usually modelled Gaussian, an unbound space is required and a highly advanced MCMC (Markov Chain Monte Carlo) algorithm – Hamiltonian MC sampling – is used to require all parameters to run without boundaries. Therefore, two types of conversion need to be performed: Ampl and Aysm need conversion from non-negative to unbound and Ctch needs double-bound to unbound conversion. Usual pairs of transformations (log/exp and logit/inv_logit) can be used for this conversion and all non-linear parameters are transformed to an unbound space, forming the LACY formula: $P_t = exp\ Asym + exp\ Ampl + logit^{-1}\ (1\text{-}$

*Ctch)*$^t$. With this formula, the Asym and Ampl parameter is always positive and Ctch represents an odds-ratio.

To examine the symmetry of transfer assumption, learning curves of experienced and inexperience drivers were compared. Especially the amplitude parameter is interesting for these purposes: the difference in amplitudes demonstrates the difference in the amount of learning that took place within the training. These results can be used to indicate whether previously acknowledged experience within a specific task, in this case driving on the road, is transferred to the trained task, in this case driving in an online simulator. When this difference in amplitude is apparent, it can be indicated that driving skills are transferred from on-road driving to simulator driving. Subsequently, according to the symmetry of transfer assumption, it can be indicated as well that acquired skills transfer from simulator driving to on-road driving

# 3. Exploration of existing datasets

## 3.1 Introduction

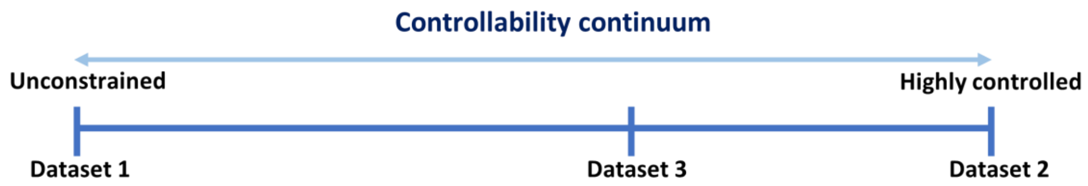
Current on-road driving training makes use of a predetermined curriculum which includes four phases that are trained in the respective order: (1) vehicle control, (2) acting in low complexity traffic situations, (3) acting in high complexity traffic situations, and (4) safe and responsible driving (CBR, n.d.). This division of phases ensures that driving trainees can learn how to drive by gradually increasing the complexity level that they should be able to handle according to their previously acquired skills. However, within on-road driving training, other traffic and environmental circumstances can only be predicted and not controlled to follow the perfect order of complexity. Therefore, it is not possible to train very specific traffic situations. Simulators have the unique possibility to do this, and specific lessons can be developed that gradually increase in complexity.

The online driving simulator of Green Dino possesses these gradually increasing complexity lessons and orders them accordingly. However, students are given the freedom to select a lesson they desire to perform and do not have to follow the predetermined order. Additionally, students do not have to repeat certain lessons until sufficient performance is achieved. Therefore, it is questioned how controlled the provided training should be to enable learning and to measure learning effects.

This analysis makes use of three different datasets that are located at different points on the controllability continuum (Figure 6). The controllability continuum represents a continuum between uncontrolled and highly controlled circumstances in which data can be collected. The dataset on one end of the continuum (study 1) is obtained in the original online simulator of Green Dino and was gathered in a relatively unconstrained manner: students had a lot of freedom in what to train, in what order, for how long, and when. The dataset on the highly controlled end of the continuum (study 2) was obtained in an experimental setting using a physical driving simulator. The participants were provided with a specific training for a specific number of trails, were constantly in the presence of researchers, and had to come to the University of Twente to perform the training. The dataset in the middle (study 3) was gathered in a manner that falls somewhere between these two extremes: it was also gathered in an experimental manner with a specific training for a specific number of trials, but students were provided more freedom because they were able to perform the training within their own environment (i.e. at their own computers in their own homes) and no instructor/researcher was present to watch their behaviour during the training.

**Figure 6**

The controllability continuum.




All datasets were observed for signs of learning curves and learning curve analyses were performed when these indications were positive. Results will demonstrate the effectiveness of the manner in which these datasets are obtained and will be used to determine the point on the controllability continuum which appears to be best suited for learning curve analyses within simulator-based driving training. Furthermore, this paper is the first to analyse driving simulator data with the LACY formula of the TFM. The measurements used within these datasets will be examined to see which variables are best suited for learning curve analyses with this model.

### 3.2 Phase 1: Exploring unconstrained data

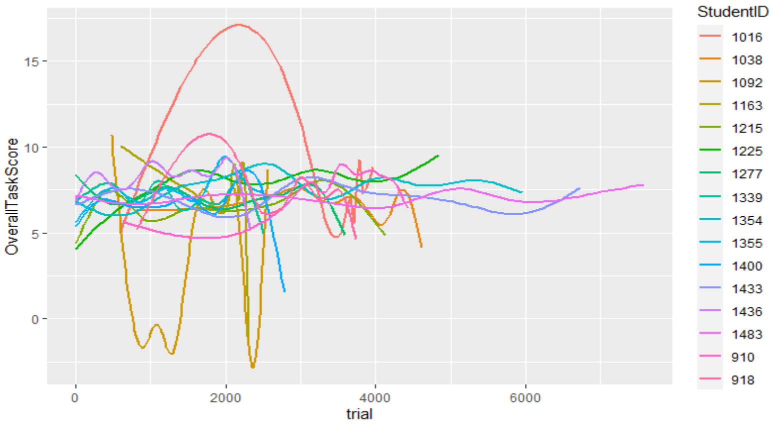
Within the first phase, the data of the original online driving simulator of Green Dino was explored for suitability for learning curve analyses. The data was gathered by the online driving simulator which automatically stores the data of all students. The sample included students that were in the middle of their driving training, being the perfect representatives since they are at the beginning of their driving skill acquisition process. The dataset included different variables with information about the performance of the trainees on specific lessons and tasks. The performance of each lesson results in different measurements falling in two categories: safety scores and task scores (see Chapter *1.2.4 Online driving simulator* for more information regarding these scores).

The most important variables for this analysis were considered to be: identifying variables Student ID, Lesson ID (specifying the main focus of the lesson, e.g., taking turns) and Task ID (specifying the specific task students performed within a lesson, e.g., road position or starting the car), and outcome variables OverallTaskScore (task performance measure for the entire lesson, between 0-10) and TaskScore (score for a specific task, such as lane position, between 0-10). To be able to display learning effects, a continuous variable *trial* was added which cumulatively counted the task performance within the different lessons for each individual. Important to note is that if learning curves are present, they will be reversed to the traditional form described previously. Opposite to outcome measures number of errors or Time on Task (ToT) where a decrease reflects a better performance, an increase in score reflects a better performance for these outcome measurements.

When trying to visualise learning curves, it was noticed that the trials all represented different parts of a lesson (e.g., looking behaviour, turning right, keeping distance, etc) which were ordered alphabetically. Moreover, outcome variable OverallTaskScore was mentioned for each specific task within one lesson, resulting in a multiplication of the occurrence of this variable. This resulted in wobbly curves not providing specific information regarding the learning effects (Figure 7). After careful exploration of the data, it was discovered that the trial variable did not demonstrate a process of learning but plainly showed the scores for each task within a lesson due to its alphabetical order. Therefore, it was concluded that the trial variable was not representative of the number of performed lessons. It cumulatively counted the performance measures that were gathered for all tasks within a lesson and did not represent the number of lessons performed.

**Figure 7**

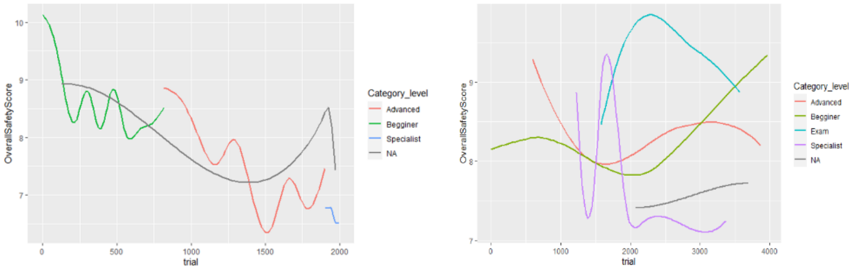
Visualisations of wobbly curves due to repetition of OverallTaskScore within the trial numbers.




After further exploration, it was remarked that there were many differences between the training processes of the trainees. Firstly, students differ in the number of lessons completed in the online simulator. Some students only performed the introduction lesson, others just tried to do an exam lesson, and others focused on completing as much lessons as possible. Per lesson, students got one outcome score per specific task included. Therefore, for the students that only performed a small number of lessons, a difficulty for examining learning curves originated, since more performance measures are required to visualise the process. Secondly, the order of lessons highly differed between the students. The lessons within the programme increase in difficulty, so the first lessons are the easiest ones, and the students gradually continue to the more difficult ones. Some students followed this exact order as presented in the simulator, however others just seemed to do the lessons they think are necessary for them (Figure 8). Thirdly, the students differed in their performance between lessons, indicating that some lessons were harder or were intended to evoke errors. This might be caused by the adaptive feedback system the simulator uses. The adaptive feedback system offers feedback for specific tasks within the lesson that are not sufficient yet, but does not give feedback to those already properly performed, which might make these tasks harder to perform.

**Figure 8**

Visualisations representing the performed order of lesson difficulty: a student performing the predetermined order with increasing difficulty (left) and a student performing different difficulty lessons mixed (right)




To be able to visualise development of skills better, a sample of students that had 2000 or more outcome measurements was created, which accounted for around 45 lessons. Furthermore, the new variables trial_lesson and trial_task were made. These variables cumulatively counted the specific lessons and tasks, allowing to analyse the development between specific lesson or of a specific task within different lessons. Again, due to the alphabetical order of the tasks within a lesson, it was not possible to visualise the development within a specific lesson, but the trial_task variable seemed suitable to visualise the development of specific tasks over lessons.

However, these visualisations also did not show indications of learning curves. In all the learning processes visualised with the overall performance score of a lesson, deviant patterns were found such as dips in the learning process or a decrease in performance (Figure 9). These effects seem to appear for all students in this created sample, indicating that a moderator might be intervening with the relation. This effect could be caused by the adjustment of task difficulty between the lessons and the adaptive feedback system. Moreover, the visualisations of task specific learning processes in this sample did not show potential learning curves (Figure 10). This can have different causations like fatigue of the student, difficulty of certain lessons, the feedback provided in the lessons, or the order in which the lessons are performed, but this are only thoughts since no hard conclusions can be drawn from this data.

**Figure 9**

Raw OverallTaskScore data plotted on trial_task displaying the development over overall task performance within the entire learning process.

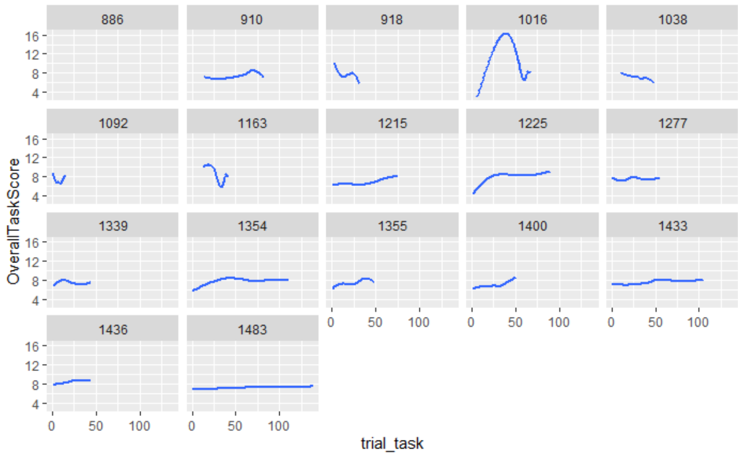
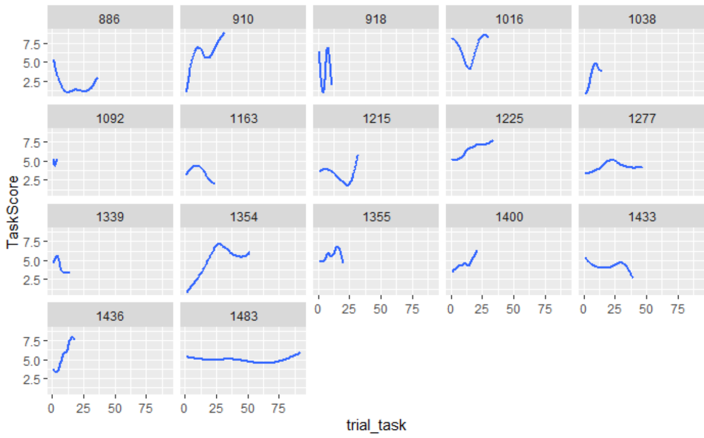



**Figure 10**

Raw Taskscore data plotted with trial_task for task taking sharp turns.




### *3.2.1 Conclusion*

It is not possible to conduct learning curve analyses with the data gathered from the online driving simulator. The data is too unconstrained due to the great amount of freedom students get in deciding which lesson to perform and in which order. Furthermore, the system itself is highly adaptive. It provides the students with different levels of feedback within different lessons, making it difficult to compare learning effects within and between students. However, this does not imply that the online simulator is not useful in acquiring driving skills and we should not jump to the conclusion that the simulator is invalid. The unsuitability could be in the way the scores are computed and the training itself can be effective.

A point for improvement is the freedom given to students. Students can completely determine by themselves what lessons they will perform, in which order, and how many times. This freedom brings two issues. Firstly, data is difficult to compare within and between students. Within-subject analyses are difficult to perform since some driving trainees might start with lessons for specialists, skipping the acquisition of essential basic skills in the lessons with lower difficulty. This is one of the causes for the difficulty of representing learning effects. Additionally, between-subject analyses are hard to execute. All students have different learning paths and number of lessons performed, making it difficult to observe a specific and general learning effect for all students. This makes it for example difficult to evaluate the lesson itself: someone performing well might do so because he or she already did a lot of lessons, but somebody else performing well might have done this as first lesson. Secondly, students are allowed to continue with other lessons when they did not score sufficiently. In this way, students will not train their skills till a sufficient level and will stop their learning process without having reached their maximum performance level.

Therefore, it is advised to add a specific order to the lessons which students are obligated to follow, so that learning effects may be shown which can be analysed easier. Moreover, training particular skills more by means of repeating certain lessons might be helpful in developing the specific skill set.

### 3.3 Phase 2: Exploration of highly controlled data

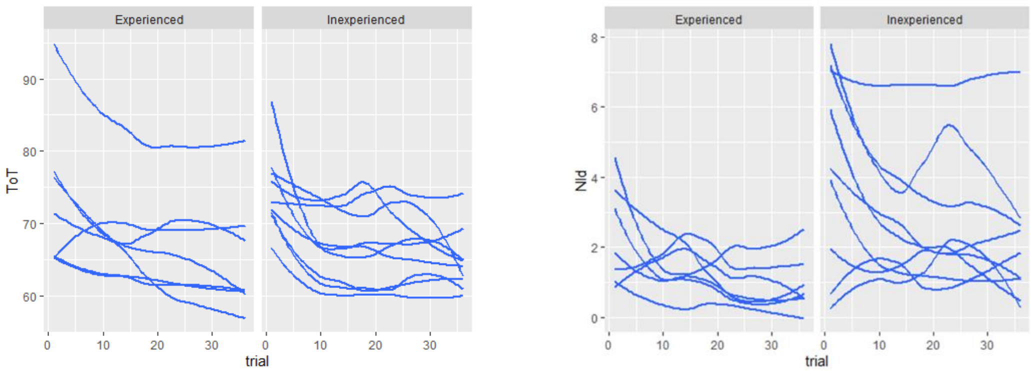
Results of Phase 1 show that learning curve patterns are difficult to analyse when the data is collected in an unconstrained manner. To see whether learning curves analyses could be possible when the data is gathered in a more controlled manner, data gained in a controlled experiment using a physical driving simulator will be analysed. This dataset was obtained by Voskes (2021) and focused on the effectiveness of including speed-episodes in driving training. Differences in driving skill acquisition were examined between speed-focused training and accuracy-focused training. It is expected that this dataset is suitable for learning curve analysis because it uses the same parameters (ToT and number of errors) as the surgery simulator master thesis studies (Arendt, 2017; Kaschub, 2016; Küpper, 2018; Weimer, 2019; Westerhof, 2018) and is obtained similarly controlled. When learning curves can be found in this dataset, it can be suggested that these types of analyses are also suitable for acquiring skills within the driving domain. Moreover, it could be an indication that the online simulator data could also produce data suitable for learning curve analyses when conducted in a more controlled way.

Firstly, the raw data was explored for learning curve patterns for outcome variables ToT and number of lane departures with trials as continuous variable. Since the participants performing the speed training were instructed to drive faster and disregard accuracy between trial 13 and 24, data of this group was removed from the dataset and the analysis continued with the participants in the accuracy training group.

Next, this new sample was divided in experienced and inexperienced drivers. Their learning patterns were plotted for both outcome variable ToT and number of lane departures (Nld). Within these plots, patterns of traditional learning curves were displayed for both outcome variables: a fast decrease in ToT and Nld in the first couple of trials and gradually a stabilisation of the performance (Figure 11). Additionally, the graphs seem to display differences in amplitude between experienced and inexperienced drivers. Whereas experienced drivers seem to reach their maximum performance relatively fast, inexperienced drivers seem to display more learning before reaching their maximum performance. Learning curve model estimates will be calculated to determine whether this observed difference is actually there.

**Figure 11**

Plotted raw data on ToT (left) and number of lane departures (right). A division was made between experienced and inexperienced drivers for both outcome variables.




Subsequently, the tweak-finder model of building skills was used to calculate learning curves (Schmettow, n.d.). Two models were formed: (1) outcome variable ToT with predictor driving experience and (2) outcome variable Nld with predictor driving experience. The three parameters were attached with priors, an outcome variable and predictor dependent multi-level model indicators. The family used for outcome variable ToT was Gamma and for outcome variable Nld Poisson (Schmettow, n.d.). Since the observed difference in amplitude between experienced and inexperienced drivers is of main interest, the amplitude parameter includes the driving experience predictor in its formula (See Appendix A for the Rscript).

The population-level fixed effect estimates of model 1 are represented in Table 1. When applied to the formula, they give the following results for the population-level based estimates: $ToT = 63.10 + 9.68 + (3.64*Inexperienced) + (1- 0.39)^t$. This means that, on population-level, inexperienced drivers were able to improve their performance with 3.64 (*95% CL* [1.32, 38.06]) seconds more than experienced drivers, which meets the expectations that inexperienced drivers have more room to learn new skills. A caterpillar plot was created representing the exact participant-level estimates (Figure 12).

**Table 1**

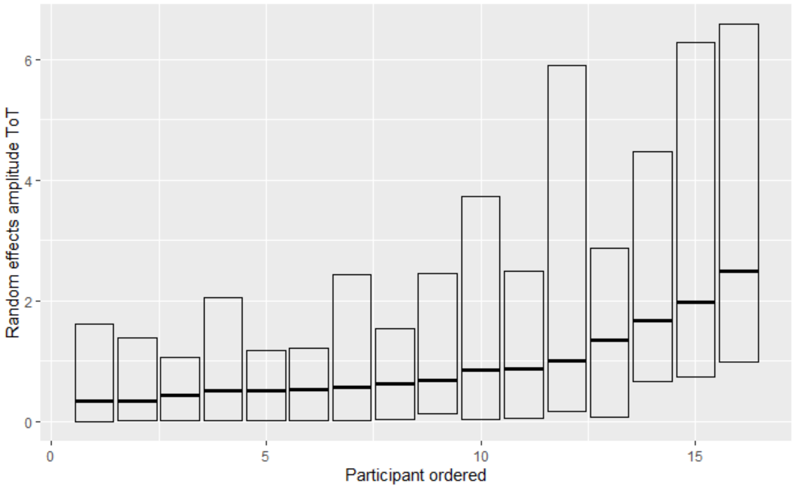
Coefficient estimates with 95% credibility limits of the fixed-effects for outcome variable ToT

| Parameter | Centre | Lower | Upper |
|---|---|---|---|
| Amplitude | 9.68 | 3.84 | 24.58 |
| Amplitude Experience* | 3.64 | 1.32 | 38.06 |
| Catch rate | .39 | .06 | 2.01 |
| Asymptote | 63.10 | 58.53 | 66.99 |

*Added for inexperienced drivers

**Figure 12**

Predicted random effects for the amplitude parameter on outcome variable ToT. Participants are ordered by increasing amplitude.

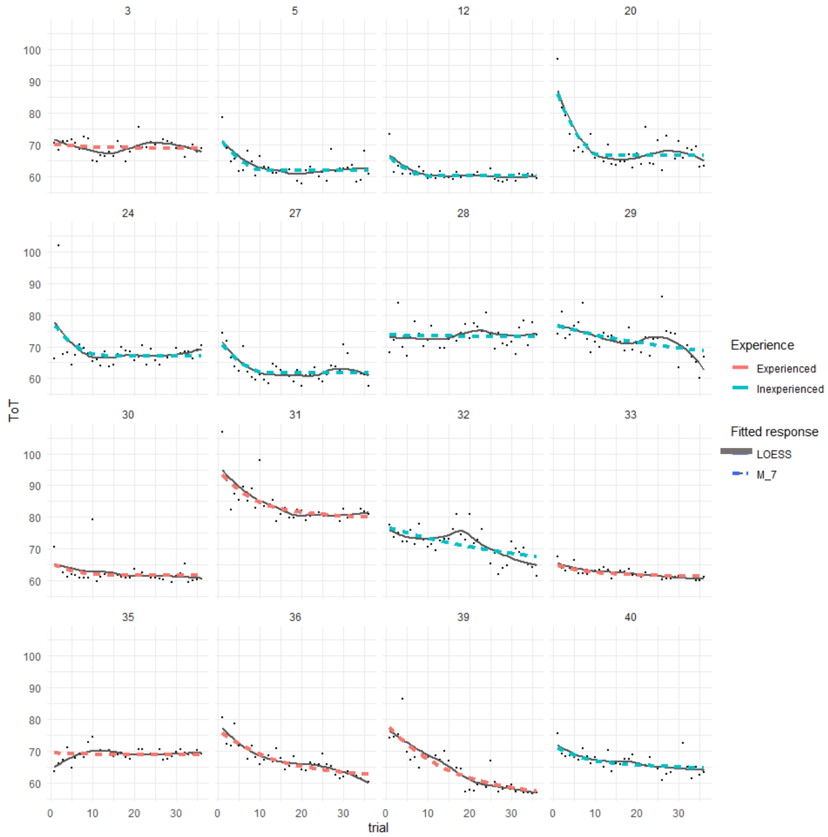



To visualise the participant-level estimates, the predicted estimates were visualised as learning curves and a model fit was performed (Figure 13). Overall, the model seems to fit the gathered raw data: the majority of datapoints are on the predicted learning curve and follow the same trend as the predicted learning curve.

As indicated by the fixed effects, predicted learning curves from most inexperienced drivers, specifically participant 5, 12, 20, and 27, demonstrate a higher amplitude than most experienced drivers. An interesting observation is that three out of the seven experienced drivers (participant 33, 36, and 39) do not seem to stabilise their performance yet. This means that the training was too short for them to reach their maximum performance. The same holds for two out of the nine inexperienced drivers (participant 29 and 32). Participant 3, 28, and 35 showed a relatively stabilised performance from their initial trial on, suggesting that they did not find many new tweaks within the simulator.

**Figure 13**

Predicted learning curves and model fitting for outcome variable ToT (s). The red dotted graphs are from experienced drivers and the blue dotted graphs from inexperienced drivers. The gray line represents the visualised curve on the observed data.




The population-level fixed effect estimates of model 2 are represented in Table 2. After these numbers were applied to the LACY formula, the following equation was formed: *Nld = 1.35 + 2.57 + (1.96*Inexperienced) + (1- 1.02)$^{t}$*. This shows that inexperienced drivers started with 1.96 (*95% CL* [0.91, 7.70]) more errors than experienced drivers. So, as expected, inexperienced drivers have a higher amplitude on population-level than experienced drivers and therefore demonstrate more learning. However, estimated learning curves of the participant-level estimates (Figure 14) do not show this difference in amplitude as clearly as the population-level estimates (see Figure 15 for overview of the participant-level estimates). Nevertheless, the initial performance of inexperienced is in general worse than that of experienced drivers, suggesting the appearance of the reversed-transfer effect.

**Table 2**

Coefficient estimates with 95% credibility limits of the fixed-effects for outcome variable number of lane departures

| Parameter | Centre | Lower | Upper |
|---|---|---|---|
| Amplitude | 2.57 | .97 | 5.36 |
| Amplitude Experience* | 1.96 | .91 | 7.70 |
| Catch rate | 1.02 | .15 | 5.69 |
| Asymptote | 1.35 | .87 | 2.36 |

*Added for inexperienced drivers

**Figure 14**

Predicted learning curves and model fitting for outcome variable number of lane departures (Nld). The red dotted graphs are from experienced drivers and the blue dotted graphs from inexperienced drivers. The grey line represents the visualised curve on the observed data.

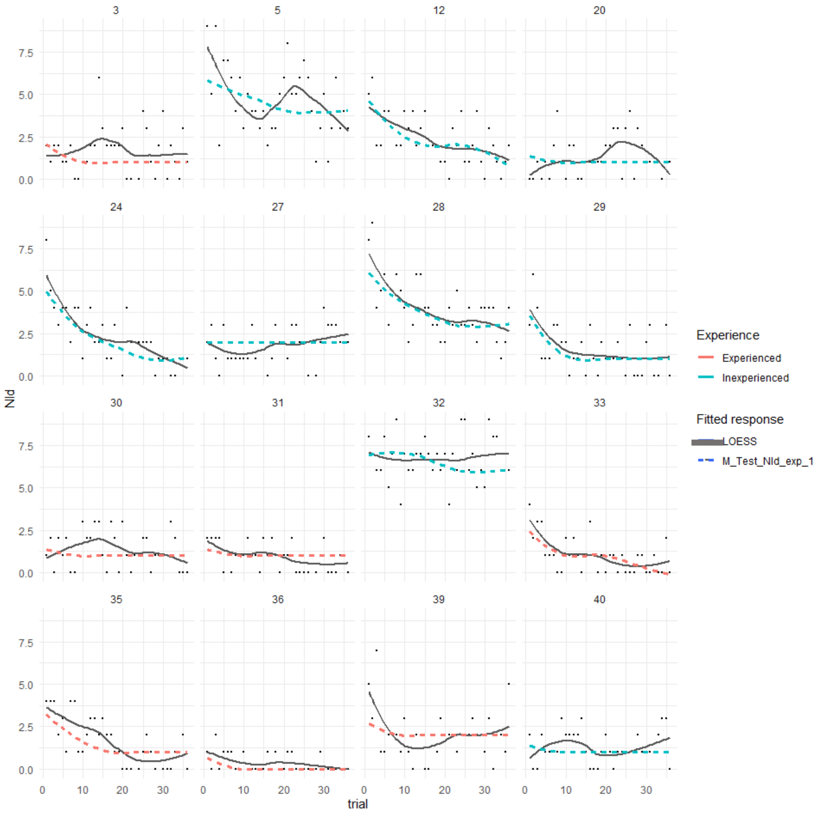
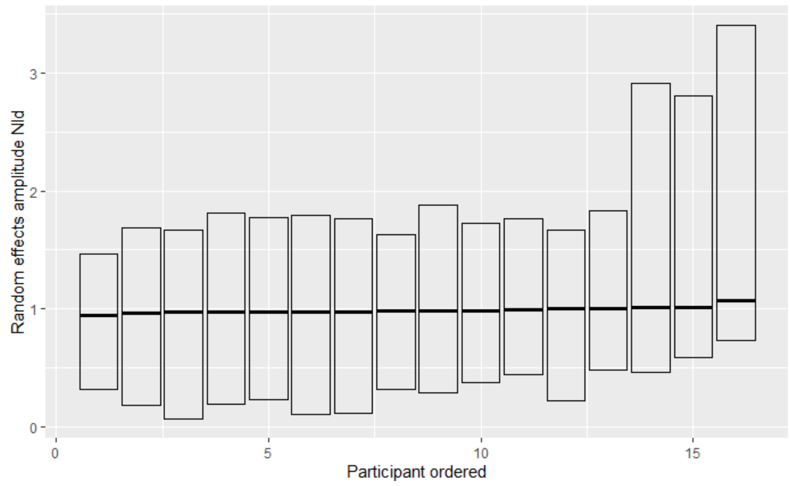


**Figure 15**

Predicted random effects for the amplitude parameter on outcome variable ToT. Participants are ordered by increasing amplitude.




Inexperienced drivers 12, 24, 28 and 29 indeed show a relatively high amplitude compared to the other drivers, however their learning curves do not all have a typical pattern. To illustrate, participants 24 and 28 do not seem to reach their asymptote yet, indicating that the training was too short for them to catch all the possible tweaks. Moreover, participant 12 starts to stabilise his/her performance, but then improves him/herself even more. This pattern is also observed for participant 32 and 33: first these participants stabilise their performance, but later seem to find new tweaks that are useful for improve task performance. Furthermore, participants 20, 27, 30, 39 and 40 show a relatively low amplitude, suggesting that they did not find many new tweaks in the training that facilitated their driving performance.

A model fit analysis shows that the model fits the observed data reasonably (Figure 14). Most of the participants show the same trends within their estimated and observed curves. Exceptions are participants 3, 5, 20 and 32. Their observed curves seem to display some serious deviations from the estimated curves, indicating a less reasonable fit for these participants.

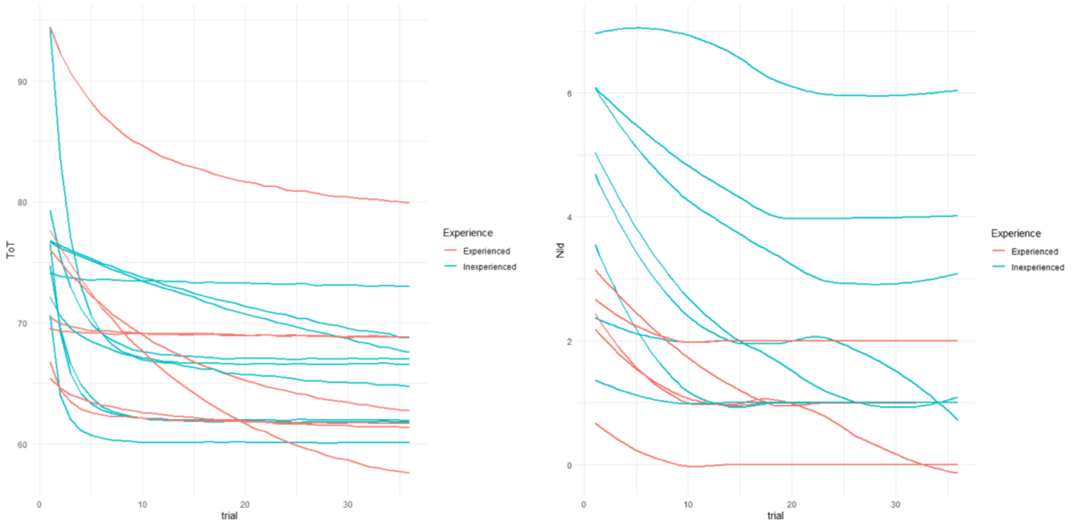
***3.3.1 Conclusion***

Both outcome variables show to be suitable for driving skill-based learning curve analysis within a highly controlled dataset gathered in a physical simulator. Most participants showed traditional patterns of learning curves, especially for outcome variable ToT. Traditional learning curve patterns were less obvious for outcome variable Nld. Nevertheless, all fitted responses showed an large or small improvement of performance anyhow. This suggests that learning happened within the simulator.

An interesting finding in this dataset are the indications of the symmetry of transfer assumption for both outcome variables (Figure 16). The initial performance of inexperienced drivers turned out to be worse than that of the experienced drivers, demonstrating a lower amount of previous experience. This suggests that inexperienced drivers have a larger set of undiscovered tweaks to explore. Once these tweaks are found, drivers can apply them and performance improves. Within this experiment, inexperienced drivers indeed demonstrated a higher amplitude, displaying a higher number of caught tweaks and thus more learning. From these findings, it can be concluded that driving expertise regarding both outcome variables ToT and Nld is transferred from on-road driving to simulator driving. Accordingly, based on the symmetry of transfer assumption, it can be expected that driving skills also transfer from the simulator to the road.

**Figure 16**

Spaghetti plot of the fitted response models for outcome variable ToT (left) and Nld (right) divided by driving experience.




Furthermore, ToT and number of errors seem to be a good outcome measurement for learning curve analyses within the driving domain. It was already proven that ToT and number of errors are suitable outcome variables for learning curves for surgery skill acquisition (Arendt, 2017; Kaschub, 2016; Küpper, 2018; Weimer, 2019; Westerhof, 2018), but this analysis demonstrates that the measurements can also work for demonstrating learning effects for driving skills.

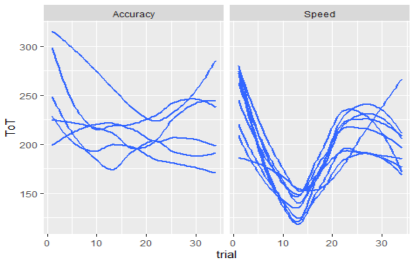
### 3.4 Phase 3: Exploration of data from a medium-controlled setting

Phase 2 shows that learning curve analysis is possible within a highly controlled dataset gathered in a physical driving simulator. It even already gives an indication that transfer effects are apparent between simulator- and on-road driving. To check whether similar results can be found in a medium controlled dataset gathered in the online driving simulator, a third dataset was explored on fitness for learning curve analyses. This dataset was obtained simulator by Van Wijk (2020) using the original Green Dino online and focused also on the effectiveness of speed-episodes in acquiring driving skills. This dataset is considered to be medium controlled because it was gathered in an experimental way, but participants were given more freedom in performing the training than in the dataset of Voskes (2021). The experiment was performed in the personal environment of the participant: participants were instructed on how to perform the training and when to do which task, however, they were able to do the training at their own home and computer, did not have to come to another location, and were not watched by a physically presented instructor. This gives them somewhat more freedom compared to the dataset of phase 2 in which participants had to come to a laboratory environment to execute the training and where a researcher was present during the entire training.

First, the data was visualised and signs of learning curves were explored. ToT (seconds) was used as outcome variable and trial (n) as continuous variable. Half of the plotted visualisations showed a dip at the middle of the training, which was due to the performed speed-episode in these trials (Figure 17). Therefore, to reduce noise, it was decided to remove these participants from the dataset and focus on the participants that performed accuracy training.

**Figure 17**

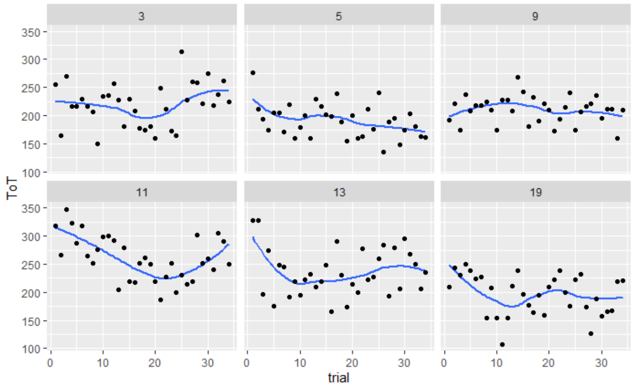
Division of plotted data between accuracy and speed training of number of trails on ToT performance.




Second, data was plotted individually (Figure 18). These graphs do not display as clear patterns of learning as the highly controlled dataset did. However, participants 5, 11, 13, and 19 do show the rapid learning phase in the beginning and seem to reach their maximum performance level. Therefore, it was decided to run a learning curve model. To eliminate possible retention effects, the last five trials that were measured a week later were removed from the dataset.

**Figure 18**

Individual ToT (s) measurements plotted for trials to search for learning curve patterns.




Next, the tweak-finder model of building skills was used to calculate learning curves (Schmettow, n.d.). The three parameters were attached with priors and a multi-level model indicator (See Appendix B for the Rscript). Population-level fixed effect estimates are represented in Table 3. Their credibility limits display a large range in which the centre estimate can fall, making the numbers less reliable. When applied to the formula, they give the following results for the population-level based estimates: $ToT = 201.29 + 95.09 + (1- 0.12)^t$. Participant-level learning curve visualisations were created by adding their belonging participant-level estimates (see Figure 19 for an overview of the estimates) obtained from the model to the population-level fixed effects (Figure 20). As expected from the pre-analysis, participants 5, 13, and 19 show traditional learning curve patterns: A fast decrease in ToT in the first few trials, followed by gradually reaching an stabilisation of performance. Participant 11 shows this fast decrease in the beginning too, however, this individual is not yet able to stabilise his/her performance. Participants 3 and 9 do not display learning effects. Their performance is relatively stabilised from the beginning on.

**Table 3**

Coefficient estimates with 95% credibility limits of the fixed-effects for outcome variable ToT (s).

| Parameter | Centre | Lower | Upper |
|---|---|---|---|
| Amplitude | 95.09 | 7.45 | 6.59e+02 |
| Catch rate | .12 | .00 | 5.35 |
| Asymptote | 201.29 | 0.82 | 2.31e+02 |

**Figure 19**

Predicted random effects for outcome variable ToT (s).

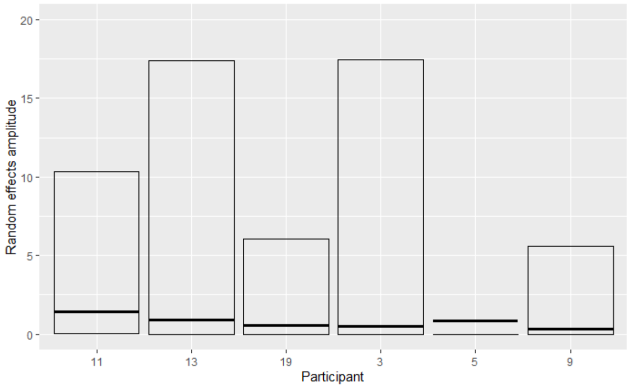
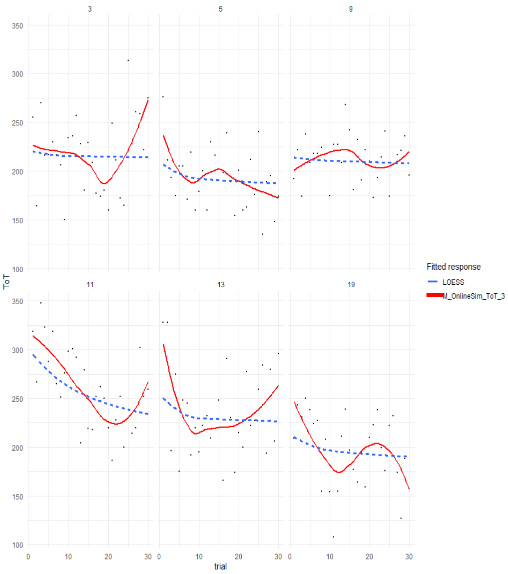



**Figure 20**

Predicted estimate learning curves and model fitting of outcome variable ToT. The red line represents the visualised curve on the observed data. The blue dotted line represents the fitted response estimates from the model.




Regarding model fit, the raw data and the predicted estimates do not seem well aligned (Figure 20). A remarkable observation for the majority of the participants is the unfit of raw data between approximately trial 10 and 30. Since the same observations are not apparent in the high controlled dataset, a possible explanation for these unfitting datapoints is an extreme amount of noise in the dataset. This noise might be caused by the higher level of freedom students had in this training compared to the highly controlled training due to the absence of an instructor and the personal environment in which the training was performed.

### *3.4.1 Conclusion*

These results suggest that medium controlled datasets and data gathered in an online driving simulator are suitable for learning curve analyses, but that the estimations fall within a large range. Four out of six individually estimated learning curves show clear learning patterns and display traditional learning curves. However, the credibility limits of the estimated coefficients are relatively large for both the fixed-effects and the participant-level effects. This indicates that the estimates are less reliable compared to smaller credibility ranges. Furthermore, the model does not seem to fit the observed data well.

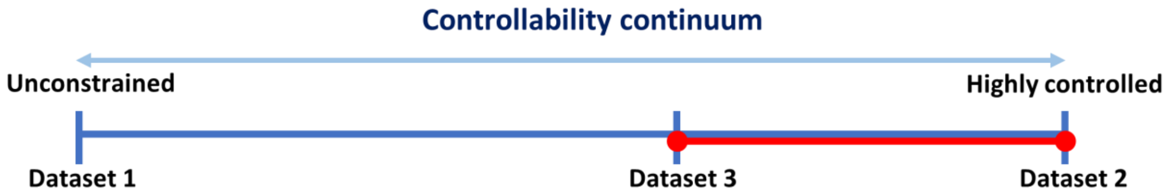
## 3.5. General conclusion

The results of these three phases provide information about the level of controllability driving training should have to perform learning curve analyses and which outcome variables are suitable. Phase 1 demonstrates that unconstrained data is not suitable for these analyses. No learning effects could be observed from the data, indicating that training should be provided with less freedom in order to observe learning effects. Furthermore, the lack of learning effects suggests that the current measures of the online simulator are not suitable for learning curve analyses. The task scores are designed to provide feedback to the trainee and not for learning curve purposes, which might make them not suited to represent learning processes. Highly controlled data including the performance measures ToT and number of errors turned out to be better suitable for learning curve analyses. Results in Phase 2 show that when the data is gathered in a controlled environment in which trainees perform the exact same training in a laboratory setting, learning effects are observable and learning curve analyses can be performed. However, it is not desirable to obtain the online simulator data in a highly controlled manner similar as was done in Phase 2. A great benefit of the online simulator is that it provides an easy way to practice driving at home, so it would not be desirable if students have to go to a laboratory setting to perform these online simulator lessons. Therefore, a third dataset falling between these extremes was examined in Phase 3. This dataset also demonstrated learning effects and showed to be suitable for learning curve analyses This indicates that training does not have to be performed in a highly controlled setting, but can be obtained less controlled.

However, what should be considered is that the learning curve estimates of the medium-controlled dataset were less reliable than the estimates of the highly controlled dataset. Therefore, based on these results, it is concluded that learning curve analyses are possible for data gathered in an online driving simulator, but that the location of optimal online simulator driving training should be somewhere between the location of these two datasets (Figure 21).

**Figure 21**

Controllability continuum. The red line represents the range in which optimal driving training is likely located.




Another conclusion we can draw from these results is that ToT and number of errors are suitable measurement for learning curve analyses on driving skill acquisition. In both the highly controlled and medium controlled dataset, ToT and number of errors represented learning effects and could be visualised as individual learning curves.

## 4. Experimental study

### 4.1 Introduction

Learning curves display the learning process of a trainee for the performed training. So, when a learning curve can be observed within the data of a trainee, it can be inferred that the training has been effective. The unconstrained data gathered by the online simulator of Green Dino in Phase 1 demonstrated that the current way data is gathered in the online driving simulator is not suited for learning curve analyses. This makes it difficult to demonstrate the effectiveness of the simulator. The highly controlled and medium controlled dataset of Phases 2 and 3 show that learning curve analyses are possible with a higher controllability of data obtainment and different outcome measurements. Additionally, as seen in the highly controlled dataset, transfer effects of previously acquired driving skills can be examined.

Based on the results of the previous chapter, a study was designed which uses the Green Dino online driving simulator (Phase 1) and its original outcome measurements. However, data is collected under more controlled conditions similar to Phase 2 and 3. This way, the suitability of the original outcome measures of the online simulator for learning curve analyses under controlled conditions can be examined. When this manner proves suitable, learning curves can be estimated and the effectiveness of the simulator can be examined.

It was shown that training located at the controlled part of the controllability continuum resulted in a better analysis of learning effects. Therefore, the training used for this experiment was designed to not go lower in controllability than the medium-controlled dataset used in the previous chapter. Additionally, based on the results of the first data exploration of the online simulator, the proposed specific order and repetition of lessons was implemented. However, to adhere to one of the benefits of the online simulator - the freedom for individually practicing specific skills -, the experiment was designed to not be as controlled as in Phase 2. By proving freedom of training moment and moments of breaks, the controllability of data gathering was lowered.

This study has 2 aims: (1) examine the effectiveness of the training provided in the online driving simulator of Green Dino by determining transfer effects of previously acquired skills, and (2) testing the TFM. By exploring the symmetry of transfer assumption – transfer from on-road driving to online simulator driving is symmetrical to transfer from online simulator driving to on-road driving -, the effectiveness of the provided training can be examined. The research question being central is formulated as: *To what extent do driving skills acquired in on-road driving transfer to online simulator driving?*. Learning curves will be estimated and the amplitude parameter will be compared between drivers on a continuum of driving experience. It is expected that driving skills acquired with on-road driving will transfer into the simulator since transfer is assumed to be symmetrical (Schmettow, n.d.) and that drivers with low levels of experience will therefore demonstrate a higher amplitude due to their larger pool of undiscovered tweaks. For the second aim, the usability for learning curve analyses with the TFM of the online simulator scores (OverallTaskScore and TaskScore) and outcome variable workload will be explored.

**4.2 Pilot study**

A pilot study was performed to determine the effectiveness of the proposed experiment. A small sample (*n* = 9) of representative individuals was recruited that performed an online driving training. The training contained three lessons (taking turns, roundabouts, crossings) with each 15 trials and were performed in a fixed order (1-2-3-1-2-3 etc.). One trial had a duration of 75 seconds, making a total duration of approximately one hour. The data was explored for suitability for learning curve analyses and used *OverallTaskScore* and workload as outcome variables. Outcome variables *NrFailed* and *TaskScore* were not used since the range of their value options was too low due to the short duration of the lessons.

At first glance, OverallTaskScore seemed to be a useful outcome variable. However, results showed that the OverallTaskScore measurement was not useful for learning curve analyses due to the short duration of the lessons. As indicated in Chapter *1.2.4 Online driving simulator*, OverallTaskScore represents the average score of all tasks encountered in the lesson and since the participants were not able to perform a large number of tasks within the 75 seconds trials, this score was not as representative as desired. Therefore, it was decided to extend the duration of the trials for the definite experiment to five minutes per trial. By extending the lesson duration, participants are able to perform more tasks and outcome variables OverallTaskScore, NrFailed, and TaskScore are better representative of the performance of the participants. Especially NrFailed is seen as a potentially suitable outcome measurement because it represents the individual's performance level without comparing it to scores of others. Additionally, this outcome measurement is similar to number of errors which is proven to be suitable for learning curve analyses in the medical domain (Arendt, 2017;

Kaschub, 2016; Küpper, 2018; Weimer, 2019; Westerhof, 2018), but also in the previously analysed driving simulator datasets.

Outcome variable workload showed promising results for learning curve analyses and will therefore be used in the definite experiment. The data represented traditional learning curve patterns and a successful model estimation was run. Moreover, transfer effects of previously acquired driving skills towards the simulator were observed.

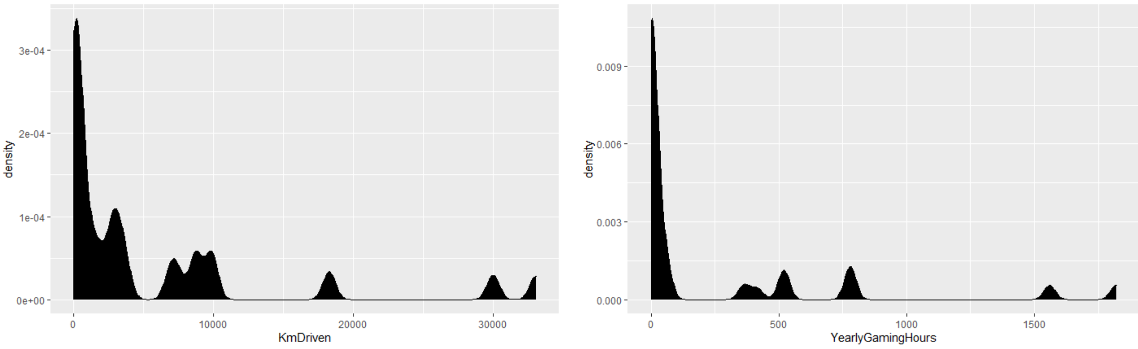
## 4.3 Methods

### *4.3.1 Participants*

33 individuals participated in this study. The sample consisted of psychology students from the University of Twente who were participating in the study as part of their course. Participants had different levels of driving experience, creating a continuum of driving experience based on an approximate of individual kilometres driven (Figure 22A). The continuum ranged from an estimate of 0 to 33072 kilometres driven in one's licenced driving days (*M* = 5219.63, *SD* = 8264.25). Regarding gaming, the yearly gaming hours of the participants formed another continuum displaying gaming experience (Figure 22B). This continuum ranged from an average of 0 to 1820 hours of gaming per year (*M* = 229.01, *SD* = 450.61).

**Figure 22**

Density plots displaying the distrubtion of kilometres driven (A) and yearly gaming hours (B) among the sample.




**A** **B**

### *4.3.2 Task*

Participants performed three different lessons: (1) turns, (2) roundabouts, and (3) intersections. Within all these lessons, normal traffic rules were operative (e.g., driving within the right lane and giving way to traffic from the right). These lessons had different difficulty levels and included different required driving skills. The first lesson was mainly operational and procedural based. Participants were driving on a road without other traffic and solely had to focus on taking the turns safely by steering and adjusting speed correctly. The second lesson focused on driving on roundabouts and included cognitive skills (e.g., situation awareness, interaction with other traffic) in addition to the required operational and procedural skills. The participants had to drive towards a roundabout and were instructed which exit they had to take. Medium traffic density was present, so they had to interact with them to be able to safely cross the roundabout. The third lesson was almost similar to the second lesson regarding task type and interactivity with other traffic. However, instead of crossing a roundabout, they had to cross an intersection and high traffic density was presented. Therefore, for these last two lessons, situation awareness was more essential than in the first lesson.

Each lesson performance counted as one trial. One trial consisted of 5 minutes, which allowed drivers to drive towards specific traffic situations and subsequently perform the main task (taking turns, crossing a roundabout, crossing an intersection) several times within one lesson. Each lesson was performed 20 times, forming 20 trials per lesson and 60 trials in total. Participants were instructed about the speed limits with traffic signs. A virtual instructor provided personal adaptive feedback to the driver based on their performance within all tasks.

#### *4.3.3 Design*

A mixed design was used for this study (See Table 4 for an overview of the within- and between factors). Over 20 trials per task, individual driving performance was measured and used to create learning curves. These learning curves represent the individual development of driving skills of a participant. The outcome variables were the driving performance of an individual, the amount of learning displayed (specific measures are described in Section *4.3.4 Measures*) and experienced workload. The continuous variable required to form learning curves was represented by the trial numbers and shows the amount of training. Predictors used were driving experience, gaming experience, and skill level.

**Table 4**

Overview of within- and between factors used in the experiment

| Within Factors | Between Factors |
|---|---|
| Individual driving performance over trials | Driving experience |
| | Gaming experience |
| | Skill level (different lessons) |

#### *4.3.4 Measures*

Based on the insights from the pilot study (see Chapter 4.2 *Pilot Study*), four outcome variables that represent driving performance were selected: *OverallTaskScore, the number of errors (NrFailed), TaskScore*, and *workload*. Firstly, the pilot study showed that OverallTaskScore was not suitable for learning curves analysis, however, effects will be measured again because of the larger sample size and the longer duration of trials that is used in the definite experiment. OverallTaskScore represents the average score of all individual tasks performed in one lesson and therefore demonstrates how well the lesson in general was performed. Scores are between 0 and 10 and represent the percentage of students in the baseline group that make more mistakes (e.g., a OverallTaskScore of 5.5 means that 55% of the students in the baseline group made more mistakes than the respective trainee, see Chapter *1.2.4 Online driving simulator* for more detailed information). The performance on OverallTaskScore and its respective learning curve is measured individually for each lesson separately.

Secondly, number of errors was selected as outcome variable. This variable was not used in the pilot study since the duration of lessons was too short for representative values but is thought to be suitable for the longer duration lessons. This variable is person and task specific which means that the outcome value is no comparison to other drivers or tasks.

Thirdly, outcome variable TaskScore seemed to be a suitable variable. TaskScore is calculated in the same manner as OverallTaskScore, so it demonstrates how well the trainee has performed the task in comparison to the drivers in the reference group. However, this variable represents a single outcome for a single task within one lesson instead of the average performance of the lesson.

Fourthly, outcome variable workload showed to be suitable for learning curve analyses in the pilot study. In self-reports, participants indicated the amount of mental effort put into the task after execution of every trial with a score between 0 (very low mental demand) and 21 (very high mental demand).

Other included measurements were the continuous variable *trial*, predictor variables *kilometres driven* and *yearly gaming hours*, and identifying variables *StudentID*, *LessonID* and *TaskName*. The continuous variable trial represented how many times the participant had completed the task, resulting in a representation of the amount of training. Trials were lesson and task specifically calculated, resulting in 20 trials for every lesson and the tasks represented within these lessons. Predictor measurement kilometres driven represents the driving experience of the participants and was calculated by multiplying the frequency of driving with the years of being licenced (obtained in the pre-questionnaire) and the national average of kilometres driven per day (Eurostat, 2021). Information regarding gaming experience was represented as yearly gaming hours and was retrieved from self-report results. Both predictor measurements formed a continuum (See 3.3.1 *Participants*). Identifying variable StudentID represented the individual that produced the respective scores and LessonID indicated which lesson was performed (taking turns, roundabouts, or crossings). Identifying variable TaskName displayed the specific task the performance outcome was linked to within the lesson. For each lesson, the one or two best representative tasks for the lesson performance were chosen to analyse the previously mentioned outcome measures. For the lesson taking turns TaskNames *Taking a turn* and *Right speed while turning* where analysed. For the lesson about roundabouts, TaskName *Taking a* *roundabout* and *Crossing a roundabout* were used and the lesson regarding crossings used TaskName *Crossing an intersection*.

***4.3.5 Materials***

**4.3.5.1 Online driving simulator.** The online computer-based driving simulator used is developed by Green Dino. Participants could log in on their own computer and download the software. The car was controlled with computer controls such as the mouse and the keyboard. By moving the mouse up, the car accelerated, and by moving down, the car decelerated. Therefore, an additive mouse was required, and a laptop mousepad was not sufficient. Furthermore, clicking the left or right mouse button activated the respective indicator and the left and right arrow keys (or the z and c keys) opened a viewport displaying the mirrors and a view to the left and right of the car. The virtual environment could contain 21 visual models and contained a logic 3D Roadnet. Virtual traffic could be added.

**4.3.5.2 Manual.** Since the experiment was performed online and the researchers were not present, a manual was designed in which all information regarding the experiment was described step-by-step (Appendix C). The manual included three sections: (1) account creation and software download, (2) ethical consent/pre-questionnaire, and (3) driving experience and workload assessment. Section 1 described in detail how an account should be created at the online driving website (https://rijlessen-online.nl/) and how the software should be downloaded on the participant's computer. The second section informed the participant about the pre-questionnaire and how to give ethical consent. Section 3 explained how the participants were expected to perform the experiment, how the vehicle could be operated, and where they could note their experienced workload level during driving.

**4.3.5.3 Questionnaires and informed consent.** This study used two questionnaires that were provided via the Qualtrics platform. The first questionnaire included assessments for driving experience and gaming experience and was completed before the driving part started (Appendix D). Additionally, it included the informed consent which included information about the nature of the study, possible minor risks, and the rights of the participant. Participants signed this form by answering yes to the question "I give consent". The second questionnaire assessed the experienced workload of participants and was filled out after each trial.

Driving experience was determined with two self-selected questions from the Driver Behaviour Questionnaire (DBQ) measuring the number of years of being licenced and frequency of driving. Individuals that were within their driving training process were asked to indicate how many hours of driving training they had had so far.

Gaming experience was assessed using proprietary items forming a questionnaire. It consisted of questions measuring hours spend on gaming and frequency of playing (i.e., How many hours do you play games per week?).

Workload was measured with self-reports on the NASA task load index (Hart & Staveland, 1988), which ranged from 0 (not mentally demanding) to 21 (highly mentally demanding) after each trial (Appendix E). This means that participants indicated their experienced workload 60 times in total. Only mental demand was measured, the other measurements were disregarded.

### *4.3.6 Procedure*

This experiment had a duration of approximately five hours and could be performed at the participant's own computer. Participants were instructed that they could do the experiment in different time slots. They received an email containing information about the study such as the deadline for completion, how to start and contact details from the researcher (see Appendix F) and a manual with more information about the experiment itself (see Section *4.3.5.2 Manual*). This manual guided the participants through the process of creating an account, downloading the required software, giving informed consent, and filling out the pre-questionnaire in the respective order.

After completion of these steps, participants could log in with their personal account at https://online-rijlessen.virtual-reality-lms.com/login/index.php and start the simulation. Three lessons were represented: (1) Taking bends, (2) Mini roundabouts with traffic, and (3) Unmarked junctions with dense traffic. Participants performed each lesson 20 times in a fixed order (1-2-3-1-2-3 etc., see Figure 23), making a total of 60 trials. After each trial, participants indicated their experienced workload in the questionnaire. After 60 trials and answering the workload question 60 times, participants forwarded their results and were shown a message explaining that this was the end of the experiment and that they were thanked for their participation.

**Figure 23**

Fixed order of the experiment. Participants started the sequence with taking bends, continued with mini-roundabouts, and ended with unmarked junctions (1-2-3) and repeated this sequence 20 times.


### *4.3.7 Data analysis*

This analysis uses the initial data analysis plan as described in Chapter *2. Data analysis*. Learning curves over the driving experience continuum will be compared to investigate whether a transfer effect is present from on-road driving to driving performance in the online simulator. For these purposes, amplitudes are the best representable parameters and will be compared. If data demonstrates that participants that have driven more kilometres in their driving career display a lower amplitude within their learning process than participants that have driven less kilometres, one can indicate that the symmetry of transfer assumption is presented.

### 4.4 Results

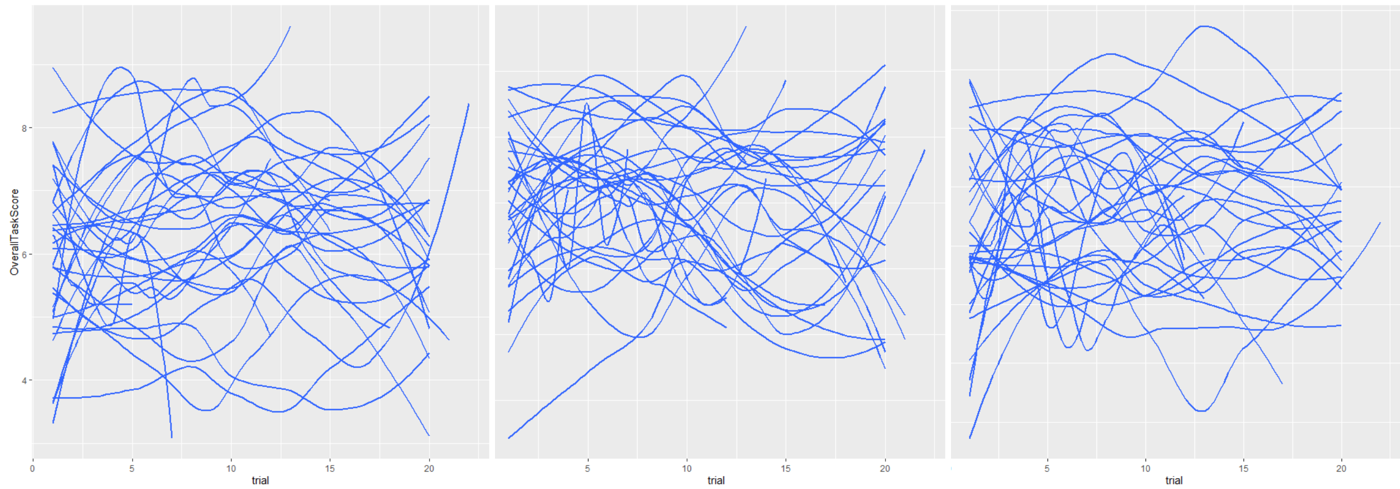
Once the experiment was finished, the simulator data was provided by Green Dino. The students that did at least 5 trials per lesson were selected, resulting in a sample population of 33 participants. Personal information, such as email addresses and names, and redundant variables were removed. Continuous variable *Trial* and identifying variables *KmDriven* and *YearlyGamingHours* were added to the dataset. Moreover, workload data was retrieved from the Qualtrics platform and rescaled to values between 0 and 1. The same continuous and identifying variables as for the rest of the data were added to this dataset.

#### *4.4.1 Simulator data*

Firstly, the observed data from the simulator was explored for individual learning curve patterns within the three lessons on outcome variable *OverallTaskScore* (Figure 24). No traditional learning curve patterns were found for each of the lessons: plotted curves show different trends between and within participants. Additionally, no other clear common pattern that might explain this wobbliness was observed. Therefore, no model estimation was run for this outcome variable.

**Figure 24**

Visualisations of the observed data from outcome variable OverallTaskScore for the three lessons separately.




**Taking turns** **Roundabouts** **Crossings**

To gain more insight in the individual tasks within the lessons and see whether these tasks are suitable for learning curve analysis, outcome variable *NrFailed* was explored. Due to problems with the data saving process of the simulator, the TaskScore outcome variable was not measured for all individual tasks, and therefore could not be used for the analysis. Within each lesson, the best representative task(s) were selected, and graphs were plotted. Firstly, individual task specific plots were calculated for the lesson *Taking turns* (Figure 25). Tasks *taking a turn* and *right speed while turning* both do not demonstrate traditional learning curve patterns and do not display another common pattern among individuals. Subsequently, individual task specific plots were created for the lessons *Roundabouts* (Figure 26) and *Crossings* (Figure 27), but these tasks also did not demonstrate traditional learning curves or other common patterns. Therefore, it was not possible to estimate learning curve models for any of these outcome variables. Appendix G contains the complete data exploration results in the form of an Rscript.

**Figure 25**

Task specific graphs for outcome variable Nrfailed on the lesson Taking turns. Each participant represents one graph. The left graphs represent the task *taking a turn*, and the right graph displays developments for the task *right speed while turning*.

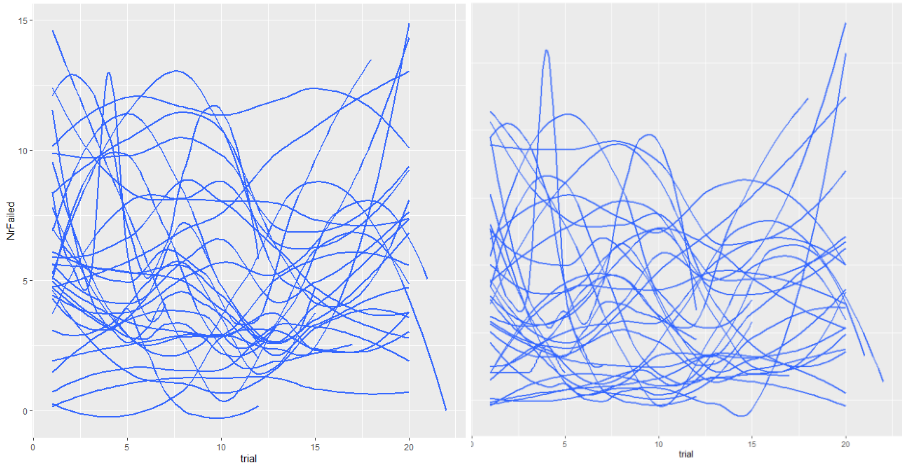



**Figure 26**

Task specific graphs for outcome variable Nrfailed on the lesson Roundabouts. Each participant represents one graph. The left graphs represent the task *taking a roundabout*, and the right graphs display developments for the task *crossing a roundabout*.

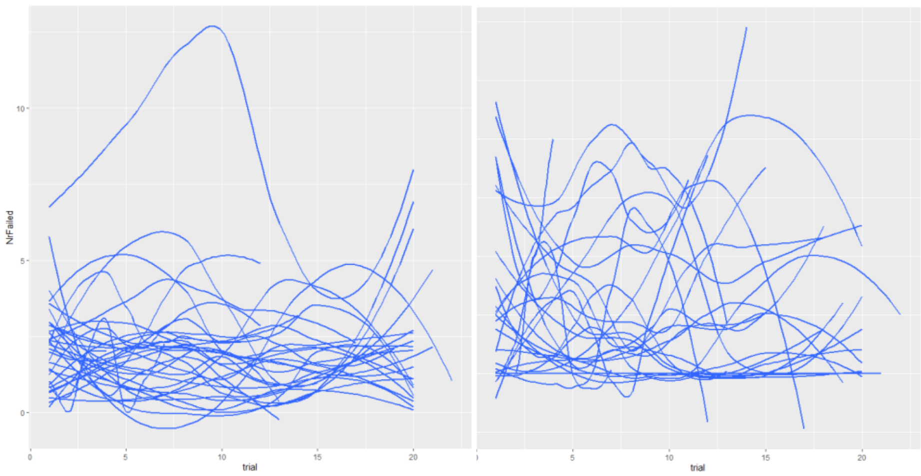
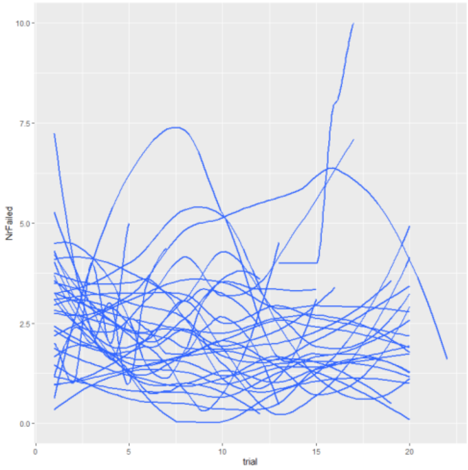


**Figure 27**

Task specific graphs for outcome variable Nrfailed on the lesson Crossings. Each participant represents one graph. The graph represents the task *crossing an intersection*.

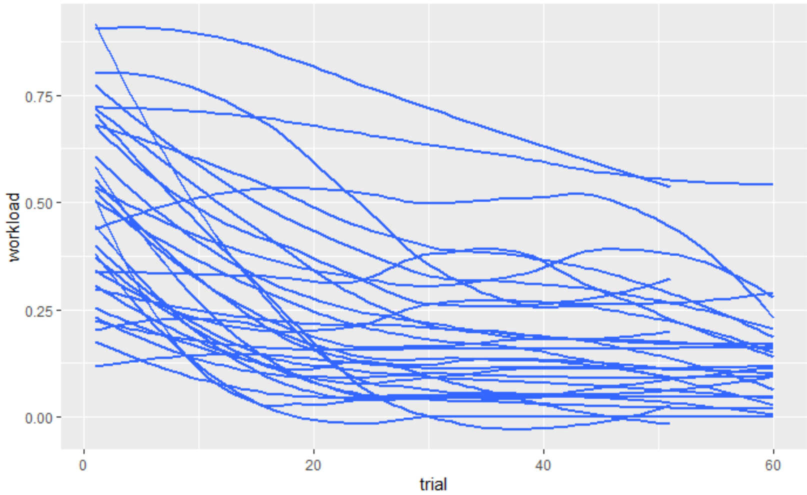



### *4.4.2 Workload data*

Workload data was retrieved from self-reports and originally contained measures between 0 and 21. To prepare the data for learning curve model estimations, the values were rescaled to values between 0 and 1, not including these boundaries. Plots of raw data showed promising results (Figure 28): the majority of the participants seems to display a traditional learning curve. Therefore, a learning curve model estimation was performed.

**Figure 28**

Observed workload data plotted with the number of trials. A clear common trend is apparent, following a traditional learning curve pattern.




Outcome variable workload was implemented in the LACY formula of the TFM. The three parameters were attached with priors and predictor dependent multi-level model indicators. To estimate the effects of driving experience and gaming experience on the amplitude parameter, their measurements were included in the formula for this parameter. The family used for this analysis was Beta which requires values between 0 and 1 (Schmettow, n.d.). See Appendix G for the Rscript.

The population-level fixed effect estimates are represented in Table 5. When applied to the formula, they give the following results for the population-level based estimates: $Workload = 0.08 + (0.58(1.00001*km\_driven)(0.9999*hrs\_gaming)) + (1\text{-}0.33)^t$. For predictor km_driven, this means that for every new trial, the amplitude increases by factor 1.00001 (*95% CL* [0.99998, 1.00004]) per driven kilometre, which is a very small effect. Furthermore, the 95% credibility range includes values lower and higher than 1, indicating both an increase and decrease for predictor km_driven. This makes the model estimation for this predictor questionable. For predictor hrs_gaming, a decrease was demonstrated. For every new trail, the number of deviations decreases by factor 0.9999 (*95% CL* [0.9994, 1.0004]), meaning that the amplitude decreases per gaming hour within a year. However, this is again a very small effect. A caterpillar plot was created representing the exact participant-level amplitude estimates that have to be added to the fixed-effects to estimate the individual learning curves (Figure 29).

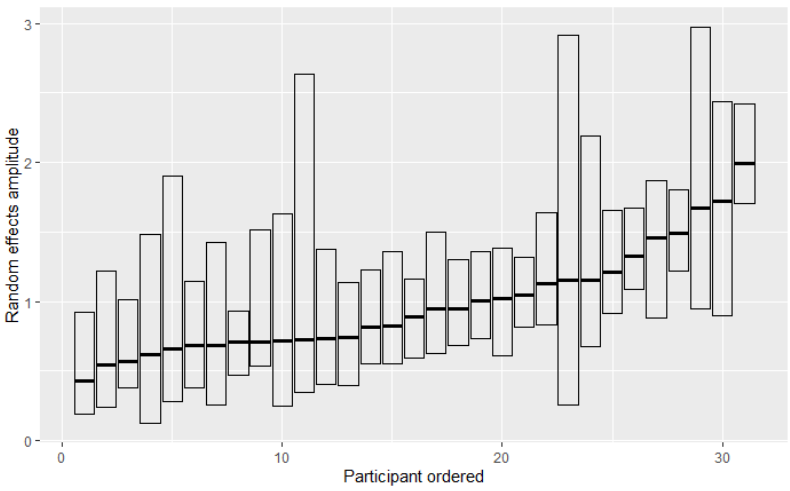
**Table 5**

Coefficient estimates with 95% credibility limits of the fixed-effects for outcome variable workload

| Parameter | Centre | Lower | Upper |
|---|---|---|---|
| Amplitude | 0.58 | 0.41 | 1.10 |
| Amplitude km_driven | 1.00001 | 0.99998 | 1.00004 |
| Amplitude hrs_gaming | 0.9999 | 0.9994 | 1.0004 |
| Catch rate | 0.33 | 0.24 | 0.77 |
| Asymptote | 0.08 | 0.07 | 0.09 |

**Figure 29**

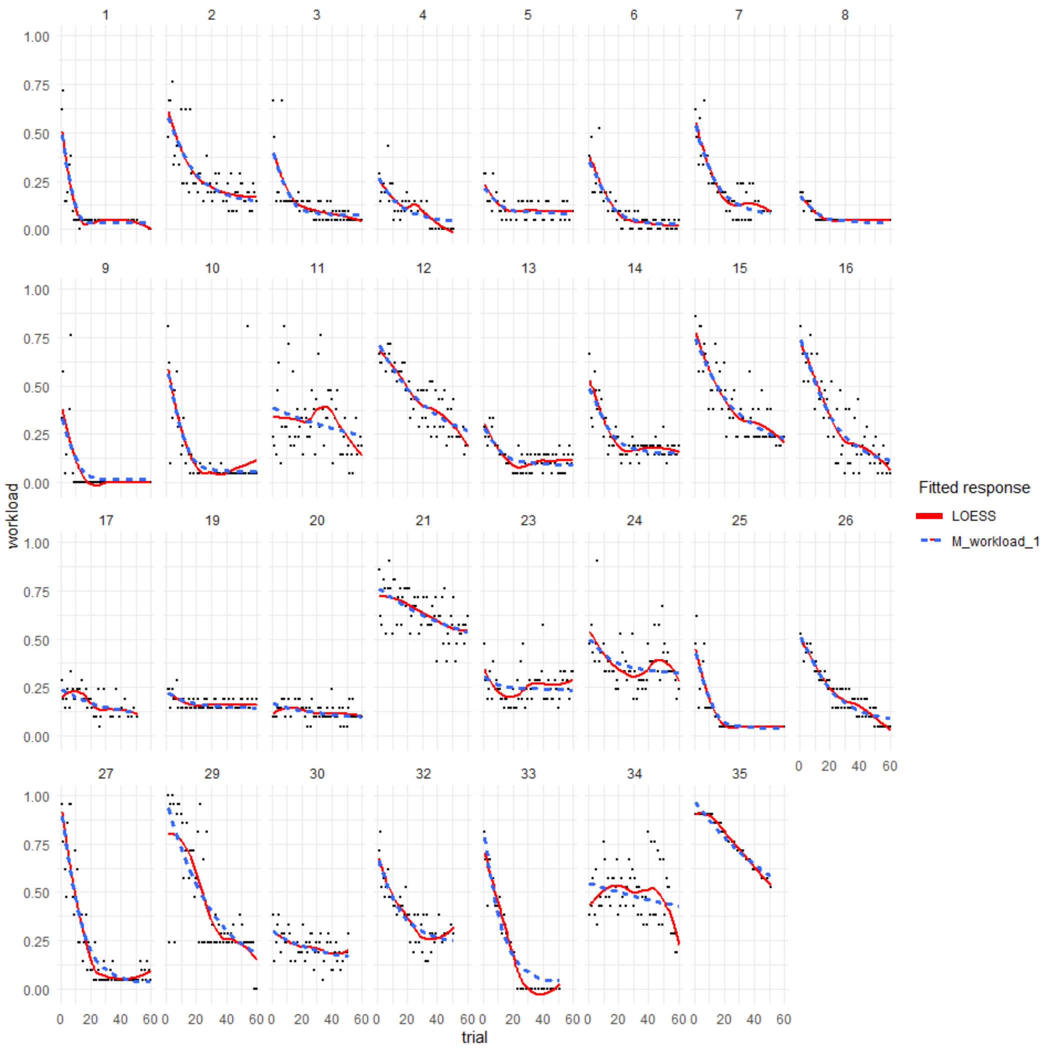
Predicted random effects for the amplitude parameter on outcome variable workload. Participants are ordered by increasing amplitude.



The predicted estimates were used to compute learning curves and a model fit was attached to measure the reliability of the predicted learning curves (Figure 30). It can be concluded that most of the participants quickly and easily find tweaks to decrease their cognitive load in the beginning, and that their pool of tweaks is empty at the moment their mental effort is stabilised. Some participants tend to show a small amplitude, indicating that their mental workload did not decrease much. This could either be explained by a very low initial mental workload, or a relatively high stabilisation of experienced mental effort. Generally, the predicted model shows a good fit with the observed data: for most of the sample, the predicted learning curves follow the same trend as the observed data. Two exceptions are participants 11 and 34, who seem to display a large deviation from the predicted model at the beginning and end of the graph. This could be caused by moments of lack of concentration or fatigue, either increasing or decreasing experienced mental effort at those moments.

**Figure 30**

Model fit of the estimated learning curves on outcome variable workload. The red line represents the observed data and the blue dotted line represents the estimated learning curve data.



However, as already indicated by the low effects of predictors kilometres driven and yearly gaming hours, individual curves demonstrate the same conclusion: there are no common differences on amplitude between participants on the kilometres driven and gaming hours (Figures 31 and 32). The amplitude is not higher or lower for a specific group of participants and seems to deviate strongly within both driving- and gaming experience.

**Figure 31**

Predicted learning curves for outcome variable workload. The colour of the graph indicates the amount of km driven, with red curves as the most km driven and blue curves as the least km driven.

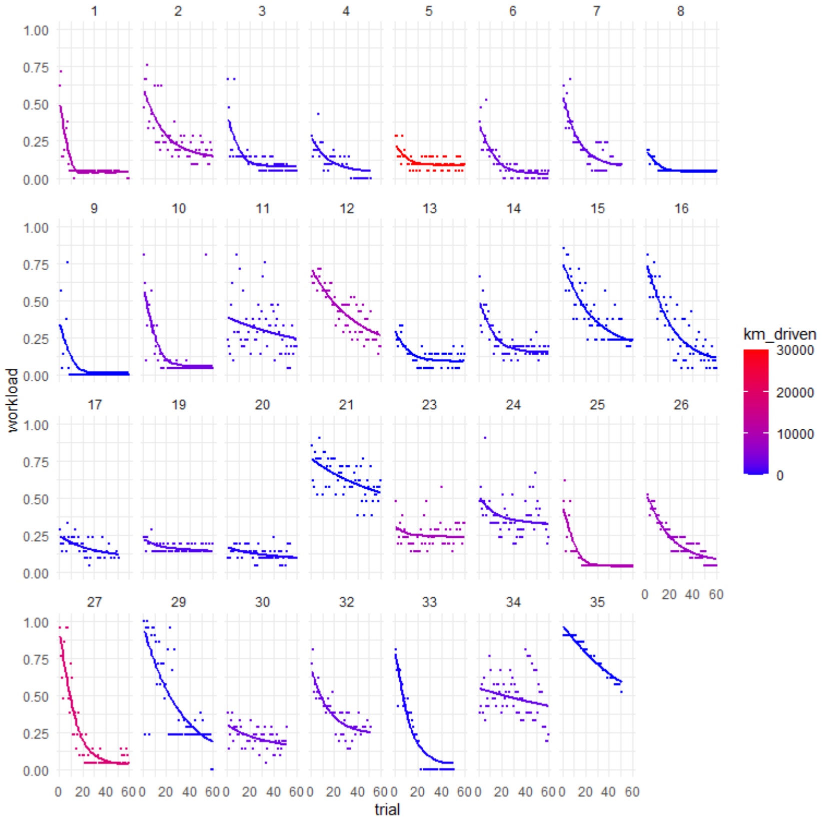
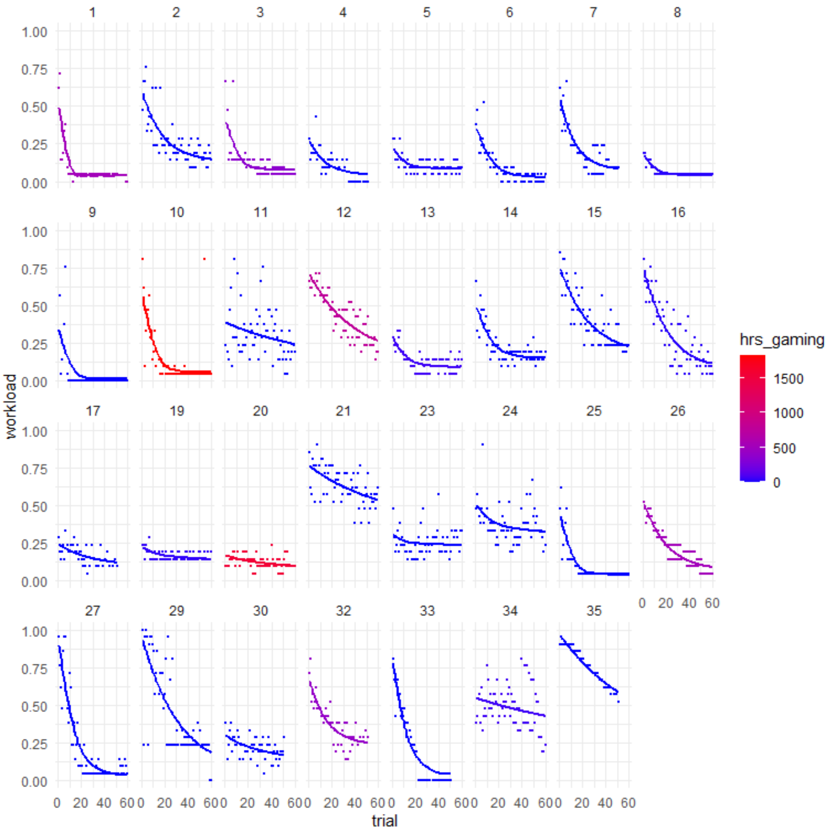


**Figure 32**

Predicted learning curves for outcome variable workload. The colour of the graph indicates the amount of gaming hours, with red curves as the most gaming hours and blue curves as the least gaming hours.




## 4.5 Discussion

This study aimed to (1) examine the effectiveness of online driving simulator training by means of exploring transfer effects of acquired driving skills, and (2) testing the LACY formula of the TFM. A performed pilot study showed that the lesson duration should be a minimum of five minutes to allow the simulator to produce representative scores and performance outcomes. The gathered workload data appeared successful and even showed some indications for transfer effects. Results of the definite study did not show learning curve analysis possibilities for outcome variables OverallTaskScore, NrFailed, and TaskScore. Again, workload showed to be useful for learning curve analyses, however, did not show the transfer effects that were found in the pilot study.

The most interesting finding of this study is the misalignment between the simulator data and the workload data. Whereas the workload data shows learning curves that indicate that the participants have developed skills that facilitated task performance, simulator data does not represent learning curves. Because of these workload based learning curves, the possibility that students did not show progress on the driving task is removed and it can carefully be assumed that the online simulator training has been effective. This leaves us with the conclusion that the TFM and the simulator scores are a mismatch and that therefore no learning curves could have been observed on these scores. The online simulator scores are developed for feedback purposes: showing the trainee how well he or she performed the task compared to a reference group. However, the TFM requires other measurements (e.g., ToT or workload as was concluded before) that measure performance directly. The following paragraphs present some possible explanations for this mismatch.

#### *4.5.1 Simulator data*

Based on the results gathered in the previous analyses within this paper, it was expected that outcome variable number of errors would be useful for learning curve analyses on driving skill acquisition. Furthermore, (Overall)TaskScore seemed to be an interesting variable to perform learning curve analyses with. It includes a general score for the entire lesson but also scores for separate tasks within the lesson. However, both types of scores did not show any indications of learning curves within this experiment.

The failure to use the data gathered by the simulator for learning curve analyses might have different causes. Firstly, (Overall)TaskScore is a score computed based on a comparison with the performance of a reference group of more than 10.000 students obtained within a physical simulator (Green Dino BV, 2021). The online simulator is developed based on the effectiveness of the physical simulator and is more or less a copy of the physical simulator. The scores are computed the same and lessons performed are nearly the same, except for the operational part (see chapter *1.2.4 Online driving simulator*). Therefore, it can be possible that individual performance is not precisely measured as for performance measures ToT, workload, and number of errors, but that the comparison aspect converts the measurement already too much. However, perhaps more importantly, the data of the reference group used for this comparison score was gathered in a physical simulator and not within the online simulator. Although these simulators have nearly the same software, they cannot be considered exactly the same. The hardware fidelity-levels of the respective simulators are very different due to their diverse vehicle controls (Caird & Horrey, 2011). The physical simulator developed by Green Dino can be categorised as a medium-fidelity simulator that requires the student to operate the car with a steering wheel, pedals and clutch. Meanwhile, the online driving simulator is considered to be part of the lowest fidelity category due to its computer-based operational system. Therefore, it cannot be concluded that scores based on this reference group data gathered within the physical simulator represent the same driving skill level in the online driving simulator. Moreover, it cannot be said with confidence that mental effort required to perform the lessons is the same in both simulators. Therefore, this reference group might not have been the best comparison and obtaining data for a reference group within the same online simulator might be more suitable.

However, it might even be better to base the scores solely on individual performance and eliminate the comparison aspect. Although a big sample of drivers is used for the reference group, it cannot accurately be concluded what exact driving skill level this group has. It might be possible that this reference group mainly included good drivers who drive better than the average driver. If so, a low task score does not necessarily mean that the task was performed poorly, but that it is also possible that the task was performed at a sufficient level but worse than the very good driving reference group. However, it is also possible that the group mainly contained bad drivers. This way, a high task score does not directly indicate that the trainee is becoming a good driver, but only that they drive better than these bad drivers. Therefore, the task scores do not provide direct information about the driving skill level of the respective individual The trainee only knows how he or she performs compared to the reference group, but is not informed what this means for their driving performance. Therefore, it would be recommended to create scores that directly measure individual performance such as the previously proven effective performance measures ToT and number of errors, and not indirectly by means of a comparison.

Secondly, the data output of the online simulator misses an essential element that the physical simulator includes: a menu displaying which mistake was made the most within a task. Within the data output of the physical simulator, task specific measurements (e.g., taking a turn or position inside the lane) are divided again into different sub-tasks, and the sub-tasks that was performed the worst is highlighted. This way, it is possible to obtain knowledge about the most specific part of the task that was performed insufficient. However, within the online simulator, this extra layer of task specific measurements is missing. This makes it difficult to conclude why the task within the lesson is not performed sufficiently and whether the same sub-task is not performed well over and over or whether a variety of insufficient sub-tasks causes the trainee to perform bad.

Thirdly, OverallTaskScore is calculated as the average lesson score based on the sum of all individual TaskScores. However, not all TaskScores were presented within the retrieved dataset, lacking some important measurements. To illustrate, for the lesson Taking Turns, solely scores for individual tasks *taking a gentle turn, taking a right-angle turn, position inside the lane*, and *driving away* were provided, and scores for some important sub-tasks within the lesson taking a turn (e.g., *taking a turn* and *adhering to the right speed within a turn*) were not calculated. This might have caused a distorted average score for OverallTaskScore since not all important measurements are included and thus not taken into consideration for the overall score.

Fourthly, outcome variable NrFailed also has possible causations for its failure to produce learning curves. As seen in other studies within the medical field (Arendt, 2017; Kaschub, 2016; Küpper, 2018; Weimer, 2019; Westerhof, 2018) and in the driving data gathered by Voskes (2021), number of errors should be a suitable outcome variable for learning curve analyses with the TFM. Additionally, this measurement was not converted to a comparison score and therefore represents a purely individual performance measure. However, no indications for learning curves were found in this dataset. A possible explanation can be that the tasks within the lessons are too general and that the tasks need to be sub-divided into more specific tasks to be able to use the number of errors as outcome measurement. To illustrate, Voskes (2021) uses the number of lane departures as error measurement. This is a more specific task than are used in the online driving simulator. An example is the task taking a turn, which includes more subtasks such as lowering speed when approaching the bend or keeping a right position inside the lane. Another causation can be the duration or the content of the training. The total training had a duration of approximately five hours which students could perform at different time slots. Furthermore, the training included only three types of almost identical lessons that had to be repeated for several times. This might have caused retention effects due to the stops the participants had in between the training, or effects caused by lack of concentration due to the monotonous and long training.

To conclude, it is most likely that the way in which the scores are calculated is not suitable for learning curve analyses with the LACY formula of the TFM, with a mismatch between the data and the model as a result. Outcome variables suitable for learning curve analyses appear to require person specific and sub-task specific measurements. This does not imply that no learning is happening within the online simulator since the workload data shows learning effects. However, additional learning effects can simply not be demonstrated with the TFM based on these simulator scores.

#### *4.5.2 Workload data*

It was expected that the experienced workload of the trainees would follow the same pattern as a learning curve: students quickly find a lot of tweaks to decrease their experienced cognitive load and facilitate their task performance, and gradually find less tweaks till mental workload is stabilised since the pool of undiscovered tweaks shrinks. Results of this study confirmed this expectation and showed that learning curve analyses were possible with workload as outcome measurement. Furthermore, they show that trainees are learning in the simulator since the decrease in workload indicates that the tasks get easier to perform due to acquired skills.

These results accord with the theory of Fitts and Posner (1967) on general skill acquisition. They state that many cognitive capacities are required to fulfil the task in the first phase of acquiring skills (the cognitive phase). The moment trainees proceed through the acquisition process, cognitive load decreases. The TFM of Schmettow (n.d.) agrees with Fitts and Posner (1967), but also adds some interesting points. He mentions that within the first phase, cognitive load is high because trainees are trying to comprehend the instructions of the task. Once the instructions are understood, a more or less discrete learning function drops from 0 to 1. This is seen in the fast decrease of mental effort within the first few trials of the learning curves of the majority of the sample and implies that most of the tweaks to decrease mental workload have been found through comprehension of the task. Students do not have to continuously and consciously think about how to perform the task because they know what is expected from them due to instruction comprehension. Then the learning process continues with finding the remaining tweaks, which is a continuous process of refining the action plan. However, since most tweaks to decrease mental workload have been discovered, mental effort is lower as the process continues.

An important note to make is that even though self-reported workload data tends to be relatively reliable compared to other self-reported measurements (Noordzij, Dorrestijn, & van den

Berg, 2016), the measurements are still not perfectly reliable due to the self defining nature. When providing self-reported answers on rating scales, participants try to convert their gut feeling into a number (Schmettow, n.d.). They establish an idea of their experienced workload on the two endpoints of the rating scale, defining their personal absolute range between these anchors. When indicating their experienced workload, they intuitively assess the intensity of their feelings and compare these to the anchors, giving the values of the rating scale a conception. Since this defining process is individual specific and therefore out of control of the researcher, response patterns can vary between participants. A result is that some participants might think that a workload score of 15 is already very high, whereas others think that only the endpoint of the rating scale (21) represents a high workload. However, clear patterns seem to be formed over time for all participants, compensating for this difference of perception (Noordzij et al., 2016).

#### *4.5.3 Transfer effects*

Based on the promising results on the symmetry of transfer assumption within the highly controlled dataset used in chapter 3.3, it was expected that these effects would also arise within this dataset. However, learning curve model estimation was not possible for the simulator data and an exploration on this effect was not possible. Therefore, no conclusions can be drawn within this study about transfer of driving skills between on-road driving and online simulator driving.

Moreover, although learning curve estimations were possible for the workload data, no indications for the transfer effect were found. The estimated effect for kilometres driven and gaming hours was very small, and the height of the amplitude parameter varied largely between driving- and gaming experience. This inability to demonstrate transfer effects is likely due to the mismatch between the simulator scores and the TFM. Perhaps with better matching outcome variables, learning curves on driving performance in the simulator could have been estimated using the TFM. Therefore, it is recommended that the symmetry of transfer assumption is tested again for the online simulator with the proven to be suitable outcome measures ToT and a more specific number of errors variable than the online simulator is currently using.

#### *4.5.4 Limitations*

This study includes some limitations. Firstly, the high amount of repetition of the same task is good for research purposes. However, it is not ensured that this many repetition is also the best way to train drivers. It might have become boring for the trainees to do the same task repeatedly, which could have reduced their concentration levels and subsequently their performance. Secondly, due to this high number of repetition and the long duration of the lessons to collect representative data, the training might have been too long. This also might have caused reduction in concentration, especially since the students are aware that less concentration has no severe consequences for them as would be in a real car. Solving these potential problems is not as easy as just reducing the number of trials or lesson duration. As observed in the pilot study, measurements were not as reliable as desired when the lessons were less long. Moreover, decreasing the number of trials causes less accurate measurements of learning developments. So, the best solution might be to divide the training in fixed blocks that have to be performed within a specific time frame. For example, students perform a set of 12 trials at once, and continue with the next set of trials after a break of some hours. Thirdly, different levels of feedback were provided. Due to the adaptive feedback system the online simulator uses, amount and extensiveness of feedback varied both between and within tasks. Therefore, it is proposed to fix the amount of feedback per trials within future research. Students could for example get extensive and a lot of feedback in the first 5-10 trials which decreases in a fixed way in extensiveness and amount over the training process. Lastly, participants were pre-occupied by two classes on the topic of learning curves. They learned about the traditional pattern of learning curves and therefore potentially knew which answers they had to give to the experienced workload self-reports to be able to form a learning curve. This makes the self-reported workload scores less reliable.

# 5. General discussion

This study had the central aim to examine the effectiveness of the training provided in the online driving simulator of Green Dino. Three sub-aims were identified throughout the process: (1) establishing the requirements to estimate learning curves on driving skill acquisition within a simulator, (2) exploring transfer effects of acquired driving skills, and (3) testing the LACY formula of the TFM. By analysing various datasets that gathered data in different ways and interpreting their results, several conclusions and recommendations can be provided.

## 5.1 Aim 1: Establishing the requirements to estimate learning curves on driving skill acquisition within a simulator

A variety of four datasets with a different level of controllability were analysed. The original dataset provided by Green Dino was gathered in an unconstrained way: students were offered a lot of freedom within their training (Chapter *3.2 Phase 1*). The level of controllability and the used outcome measures showed to be not suitable for learning curve analyses. Therefore, two other datasets on the controllability continuum were explored. The highly controlled dataset gathered in the physical simulator showed that learning curves were demonstrated and that driving skill acquisition within a simulator was thus possible (Chapter *3.3 Phase 2*). The medium-controlled dataset gathered in the online simulator also showed its potential for learning curve analyses and thus learning within the simulator, however, these effects were very small and therefore not very reliable (Chapter *3.4 Phase 3*). A fourth dataset was retrieved from an experiment in the online simulator based on these results (Chapter *4 Experimental study*). The level of controllability was considered to be not lower in controllability than the medium-controlled dataset, but not as high as in the highly controlled dataset. This dataset also did not show learning curves, so learning effects could not be measured using this type of analysis. Outcome variable workload showed more promising results regarding learning curves and indicated that student's cognitive load decreased as the training continued. This suggests that less mental effort was needed to perform the task due to the useful tweaks that were discovered during the training.

This disagreement of simulator data and workload data produces odd conclusions. Whereas the workload data shows clear patterns of learning curves indicating that learning is happening, simulator data does not show the effects of the discovered tweaks to decrease mental load within the driving performance development. Since learning curves could not be estimated within two separate studies on the online simulator data using different levels of controllability (Chapter 3.2 and 4), a mismatch between this simulator data and the TFM is currently the best explanation for the failure to find learning curves. The TFM needs other measurements to be able to estimate learning curves than the online simulator currently offers, whereas the simulator scores have unique qualities that the TFM is not able to translate into learning curves. Another possible declaration for this dissemination is that students are not able to apply the discovered tweaks that lowered their mental load to their task performance because their lack other essential driving skills to complete the driving task sufficiently.

***5.1.1 Recommendations***

If Green Dino wishes to use the TFM to establish learning effects within their online driving simulator, they should consider changing its data metrics. Research in the medical field (Arendt, 2017; Kaschub, 2016; Küpper, 2018; Weimer, 2019; Westerhof, 2018) and Chapter 3.3, 3.4 and 4 of this paper demonstrate that ToT, number of errors and workload are useful measures to estimate learning curves with the TFM. Therefore, it is suggested to add these respective measurements to the online simulator and perform learning curve analyses to explore its effectiveness. Another recommendation would be to explore other metrical options such as the absolute proportion of task failures within the occurrence of the specific task. It is highly recommended to use these metrices not in the form of a comparison to a reference group as is done now for the OverallTaskScore variable, but as an absolute individual performance score. If Green Dino does not want to use the TFM to analyse learning effects, a different way of determining learning effects could be explored, but that goes beyond the scope of this paper.

Moreover, it is recommended to use a relatively highly controlled training protocol. This paper shows that the driving training provided in a more controlled manner is better suitable for learning curve analyses than data gathered within uncontrolled training. This way, data is better structured and follows a gradually increasing driving task difficulty instead of a variety of different lessons that vary a lot in difficulty.

Adapting the metrics to fit learning curve analyses poses some additional analytical benefits that go beyond the analysis used for this paper. In addition to establishing transfer effects by exploring differences in amplitudes as was done in this paper, other possibilities arise. Together with the other two learning curve parameters (asymptote and rate), different forms of research can be performed. To illustrate, effectiveness of specific lessons can be examined (e.g., difficulty level or duration of the lesson), students that need more training can be detected, and a prediction of required training can be estimated (see Chapter *1.4.1 The Tweak-Finder Model* and Chapter *5.4 Future research* for a more detailed explanation of how the parameters fit in these research ideas). Furthermore, because the TFM decomposes tasks in subsets of tweaks, an underlying task analysis could provide useful information for the design of training. For example, particular sets of tweaks used within different lessons can be detected and a specific training with different lessons to avoid exhaustion could be designed.

**5.2 Aim 2: Exploring transfer effects of acquired driving skills**

The TFM assumes that intersecting sets of tweaks exist between tasks. Since intersection is a symmetric operator (from A to B = from B to A), transfer can be considered to be a potential symmetrical effect as well (Schmettow, n.d.). Therefore, it was expected that the on-road driving experience individuals have gathered within several traffic situations would be transferred to simulator driving and that they would therefore demonstrate less learning. The analysis on the physical simulator data indeed displayed this advantage, however, within the online simulator this expected advantage caused by experience was not observable. A possible causation for these differences in results could be the way in which the simulator is operated. As already mentioned in Chapter *4.5.3 Transfer effects*, the lack of transfer effects can potentially be ascribed to the difference in vehicle operation. The results of the physical simulator regarding the transfer effects provide additional evidence for this notion: when operational equipment is the same, transfer is more evident than when vehicle operation differs from the original task.

However, it should be noted that the discovered transfer effects within the physical simulator were based on driving performance and that the online simulator results were based on the experienced workload. It might be possible that the workload decrease was mainly based on the process of getting used to the new manner in which the vehicle was operated and not on actual driving skill acquisition that could be applied to the task. This potential explanation is strengthened by the lack of differences for predictor gaming experience. Participants frequently playing games did not show a smaller amplitude than participants that never game, indicating that they were not able to operate the vehicle more easily because of their gaming experience.

A more plausible explanation for the lack of indications of the symmetry of transfer assumption could be the personal reference points used for rating experienced workload (see Chapter *4.5.2 Workload data*). It might have been possible that the individual reference points posed difficulties for between-subject analyses. Differences in amplitude might be caused by these differences in individual reference points rather than by the learning process (Schmettow, n.d.). Therefore, it is expected that transfer effects could still be observed in the online simulator when the performance-based metrics allow learning curve analyses.

Based on these results, it can be concluded that transfer effects are observable from on-road driving to simulator driving, however, these effects are more evident for physical simulator driving than for online simulator driving. Future research could consider providing a preparing lesson in which students are trained to operate the virtual car. Trainees could for example play a game where the operational aspect is equal to that in the online simulator but where the context is different. To give an example, students could move around with an animal within a barn and already train the required operational skills. By training the operational skills in another context, students are not provided with extra training that already supports in driving skill acquisition but are only able to catch the tweaks concerning the simulator controls. This way, the process of getting acquainted to operating the online simulator is eliminated and obtained results are not moderated by this process. This potentially results in a better observation of differences in actual driving skills between experienced and inexperienced drivers and not simulator controlling skills.

### 5.3 Aim 3: Testing the LACY formula of the TFM

The TFM is a newly developed learning curve model that compares learning to finding useful tweaks that facilitate performance. Based on three assumptions (the pool of tweaks is finite, finding a tweak is irreversible, and every tweak has a fixed probability to be found), the learning process of an individual is explained. This study was de first study that explored the effectiveness of the TFM on learning effects within the driving domain. Results indicate that the TFM is suitable for learning curve analyses on acquiring driving skills when specific outcome measurements are used. Outcome variables ToT, number of errors and workload prove to be useful and learning curve models could be calculated with the LACY formula (see Chapters 3.3, 3.4, and 4.4.2). Also, the explanation of the learning process caused by the discovery of tweaks that can be used to improve performance is clear and suitable for acquiring driving skills. However, the TFM could not be used to demonstrate learning effects with the current online simulator metrics. This suggests that not all measurements are suitable for the model. For now, it is concluded that the LACY formula of the TFM is limited to outcome measures ToT, workload, and to an extent the number of errors. Future research could explore the suitability of other outcome measures for the TFM learning curve estimation model. For example, the effectiveness of absolute proportion scores based on passes/failures within the number of occurrences could be examined.

### 5.4 Future research

This study introduces some ideas for future research. Firstly, as already indicated above, a metrics that matches the TFM could be constructed and examined. The proven to be useful outcome measures (ToT and more specific number of errors) could be added to the metrics of the online simulator and a study similar to the one in Chapter 4 could be performed. The fixed duration of the lesson should be removed, and trainees could be instructed to perform a single driving task for which the ToT and number of errors (e.g., lane departures) are measured. Differences in amplitude potentially caused by experience could again be examined to find proof for the symmetry of transfer assumption from on-road driving to online simulator driving.

Secondly, outcome variables ToT, workload and number of errors prove to be suitable for learning curve analyses within situations where no other traffic is present (see Chapter 3 *Exploration of existing datasets*). However, it is not yet determined which measurements fit learning curve analyses on driving tasks where interaction with other traffic is essential. The experiment in Chapter 4 included interactional tasks, but could not demonstrate learning effects because of the mismatch between the simulator scores and the TFM. Therefore, based on this study, no conclusions can be drawn regarding suitable performance measurements for interactional tasks. Outcome variable number of errors seems to be a logical candidate. For example the number of wrong interactions, or the number of errors on sub-tasks within this interaction could be counted and used to plot performance. However, something that should be considered is that ToT might not be the best outcome variable for tasks where other traffic is present. When other traffic is presented, drivers have to interact with them and adhere to the traffic rules about certain procedures (e.g., giving the right of way to the person from the right or waiting for a pedestrian that wants to cross at a marked crossing). This might impede their ToT performance and make it therefore less reliable. Additionally, measuring ToT performance within a situation with other traffic present might make the students focus on the riskier part of the speed-accuracy trade-off (Gas et al., 2018) while performing this task – going as fast as possible, instead of focussing on driving accurately and safely. Nevertheless, this should be explored in another study together with other potential outcome variables candidates.

Thirdly, future research could focus on other parameters of the TFM. This study mainly focused on the amplitude of the learning curve to predict the transfer of the already acquired driving skills. However, there are many more possibilities for learning curve research in driver training. Once the valid performance measures are present, research for different purposes could be performed. One example would be to use the asymptote parameter to examine within which amount of training the student is able to reach his or her maximum performance. This can be used to predict the number of lessons a student needs or to predict how much driving lessons need to be addressed to a particular skill (e.g., taking turns). Based on these predictions, adaptive training including tasks from which the respective trainee benefits the most could be provided to the students. Additionally, asymptote estimations could be used to examine retention effects of the acquired skills. Some time after the initial training moment, the same training could be reperformed and it could be tested how much training is required to reach the initial asymptote again. Furthermore, research on the chance to catch a tweak parameter can present valuable information for driving skill acquisition. To illustrate, information about how fast inexperienced drivers can reach the same performance level as experienced drivers can be provided, and the time frame required to show these performance levels can be predicted. Additionally, with this information, more insights can be gathered about the possibility to obtain the lacking driving experience during the first months after licencing that causes the high crash risk of inexperienced drivers (Mayhew et al, 2003). Moreover, different types of training (e.g., speed episodes or accuracy episodes) could be tested for effectiveness. Within these specific episodes, trainees are instructed to focus specifically on one of the conditions of the speed-accuracy trade-off (Gas et al., 2018). It could be examined whether these episodes facilitate driving skill acquisition by comparing the estimated parameters of the TFM between different types of training.

Lastly, a study could be designed that investigates the overlap of tweaks used between different driving tasks. The participants within this study performed three tasks in which they could find tweaks to improve their performance. It is expected that discovered tweaks within one task will transfer to other tasks in which these tweaks might be helpful too (Schmettow, n.d.). However, the focus of this paper was on transfer between simulator- and on-road driving and not on transfer between different tasks. Future research could look at the similarities and differences within driving tasks and see whether discovered tweaks within one training task transfer to another training task. The online simulator measures task specific performance scores of tasks that are represented among a larger set of lessons. Therefore, the development of a particular sub-skill can be observed within a larger set of lessons and the amount of tweak overlap between this set of lessons can be predicted. To illustrate, the task *taking a gentle bend* can be measured within a big variety of lessons such as taking turns, roundabouts, and crossings. By varying the amount of expected tweak overlap in the experiment, differences in the amplitude parameter could be observed. If the lessons that are expected to have a large amount of tweak overlap demonstrate a reduced amplitude when performed after each other, these lessons could be used to train the same driving skill. As a result, training for a particular skill can be made more divers. Furthermore, a certain amount of training on a particular task that already prepares students properly for a more advanced task can be determined. Subsequently, results could be applied to the training process by providing information about which specific driving tasks influence other driving tasks and provide information about the order of the training procedure.

## 6. Conclusion

This study aimed to examine the effectiveness of the Green Dino online driving simulator and to explore the symmetry of transfer assumption of the TFM. Additionally, different outcome variables were explored for suitability for learning curve analyses with the LACY formula of the TFM. Data on experienced workload was suitable for learning curve analyses and showed that participants were able to show signs of learning by decreasing and stabilising their experienced mental effort as the training progressed. Additionally, data gathered in a controlled environment proved suitable. However, results showed that learning curve analyses were not possible on both unconstrained and medium-controlled data with the current metrics of the online simulator. The most compelling explanation is that the scores of the simulator are a mismatch with the TFM and that therefore learning curve analyses could not be performed. Reconstructing the data metrics is required to be able to perform learning curve analyses.

**References**

Allen, R. W., Cook, M. L., & Rosenthal, T. J. (2001). Low-Cost PC-Simulation Technology Applied to Novice Driver Training. In: Proceedings of the First International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, 14-17 August 2001, Aspen, Colorado. Iowa City, IA: Public Policy Center, of Iowa, 2001: 37-41. https://doi.org/10.17077/drivingassessment.1007

Allen, R., Park, G. D., & Cook, M. L. (2010). Simulator fidelity and validity in a transfer-of-training context. *Transportation Research Record*, *2185*(2185), 40–47.

Allen, R. W., Park, G. D., Cook, L. K., & Fiorentino, D. (2007, September 12-14). The Effect of Driving Simulator Fidelity on Training Effectiveness. Paper presented at the Driving Simulator Conference (DSC) North America, Iowa City, US

Allen, R. W., Park, G., Terrace, S., & Grant, J. (2011, June 27-30). Detecting Transfer of Training Through Simulator Scenario Design: A Novice Driver Training Study. Paper presented at the Driving Assessment Conference, Lake Tahoe, CA, US.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369–406. https://doi.org/10.1037//0033-295X.89.4.369

Arendt, A. (2017). Towards reliable and valid prediction of MIS-performance with basic laparoscopic tasks in the LapSim and low-fi dexterity tasks(Unpublished master's thesis). University of Twente, Enschede.

Beanland, V., Goode, N., Salmon, P. M., & Lenné Michael G. (2013). Is there a case for driver training? a review of the efficacy of pre- and post-licence driver training. *Safety Science*, *51*(1), 127–137. https://doi.org/10.1016/j.ssci.2012.06.021

Caird, J. K., & Horrey, W. J. (2011). Twelve practical and useful questions about driving simulation. In D. L., Fisher, M. Rizzo, J. Caird, & J. D. Lee. (Eds.). Handbook of driving simulation for engineering, medicine, and psychology (pp. 5.1-5.17). Boca Raton, Florida; CRC Press

Centraal Bureau Rijvaardigheidsbewijzen [CBR]. (n.d.). *Wat is de rijopleiding in stappen (RIS)?*. Retrieved December 12, 2021, from https://www.cbr.nl/nl/veelgestelde-vragen/wat-is-de-rijopleiding-in-stappen-ris.htm

De Winter, J., De Groot, S., Mulder, M., Wieringa, P., Dankelman, J., & Mulder, J. (2009). Relationships between driving simulator performance and driving test results. *Ergonomics -London-*, *52*(2), 137-153.

De Winter, JCF., Wieringa, PA., Dankelman, J., Mulder, M., van Paassen, MM., & de Groot, S. (2007). Driving simulator fidelity and training effectiveness. In SN (Ed.), *Proceedings of the European Annual Conference on Human Decision Making and Manual Control* (pp. 1-9). Technical University of Denmark.

Divekar, G., Samuel, S., Pollatsek, A., Thomas, F. D., Korbelak, K., Blomberg, R. D., & Fisher, D. L. (2016). Effects of a pc-based attention maintenance training program on driver behavior can last up to four months: simulator study. *Transportation Research Record: Journal of the Transportation Research Board*, *2602*(1), 121–128. https://doi.org/10.3141/2602-15

Dols, J. F., Pardo, J., Falkmer, T., Uneken, E., & Verwey, W. (2001). The Trainer Project: A new simulator-based driver training curriculum. Proceedings of the first international driving symposium on human factors in driver assessment, training and vehicle design, 119-124. https://doi.org/10.17077/drivingassessment.1021

Eurostat (2021). *Passenger mobility statistics.* Retrieved December 15, 2021, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Passenger_mobility_statistics#Distance_covered

Fisher, D. L., Pollatsek, A. P., & Pradhan, A. (2006). Can novice drivers be trained to scan for information that will reduce their likelihood of a crash? *Injury Prevention : Journal of the International Society for Child and Adolescent Injury Prevention*, *12*, 25–9.

Gas, B., Buckarma, E., Cook, D., Farley, D., & Pusic, M. (2018). Is speed a desirable difficulty for learning procedures? an initial exploration of the effects of chronometric pressure. Academic Medicine : Journal of the Association of American Medical Colleges, 93(6), 920-928. doi:10.1097/ACM.0000000000002167

Green Dino BV. (2021). *Operator Handleiding Rijsimulator V24*. https://drivemasterportal.com/handleidingen/nl/OperatorManual.pdf

Groeger, J., & Banks, A. (2007). Anticipating the content and circumstances of skill transfer: Unrealistic expectations of driver training and graduated licensing? *Ergonomics, 50*(8), 1250-63.

Hall, J. & West, R. (1996). Role of formal instruction and informal practice in learning to drive, Ergonomics 1996, 39(4) 693-706. https://doi.org/ 10.1080/00140139608964490

Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Advances in psychology, 52, 139-183.

Hatakka, M., Keskinen, E., Gregersen, N. P., Glad, A., & Hernetkoski, K. (2002). From control of the vehicle to personal self-control: Broadening the perspectives to driver education. *Transportation Research Part F: Traffic Psychology and Behaviour, 5*(3), 201-215. https://doi.org/10.1016/S1369-8478(02)00018-9

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: the case for an exponential law of practice. Psychonomic Bulletin & Review, 7(2), 185–207. https://doi.org/10.3758/BF03212979

Ivancic, K., & Hesketh, B. (2000). Learning from errors in a driving simulation: Effects on driving skill and self-confidence. Ergonomics, 43(12), 1966-1984. https://doi.org/10.1080/00140130050201427

Jamson, H. (2011) Cross-platform validation issues. In D. L., Fisher, M. Rizzo, J. Caird, & J. D. Lee. (Eds.). *Handbook of driving simulation for engineering, medicine, and psychology* (pp. 12.1-12.13). Boca Raton, Florida; CRC Press

Kappé, B., Van Emmerik, M. L. (2005, December 1). *Mogelijkheden van rijsimulatoren in de rijopleiding en het rijexamen*. TNO Defensie en Veiligheid. https://www.politieacademie.nl/kennisenonderzoek/kennis/mediatheek/PDF/45001.pdf

Kappé, B., Van Winsum, W., Van Wolffelaar, P. (2002). A cost-effective driving simulator. DSC 2002, Paris, France, Sept 2002.

Käppler, W. D. (2008). Smart driver training simulation : Save money, prevent (2008 springer e-books). Berlin: Springer. https://doi.org/10.1007/978-3-540-77070-1.

Kaschub, V., L. (2016). Learning complex motor procedures: can the ability to learn dexterity games predict a person’s ability to learn a complex task? (Unpublished master's thesis). University of Twente, Enschede.

Kuipers, J., Wieringa, P.A. (2018). Automated feedback on viewing skills lowers accident involvement. Proceedings of the 6th Humanist Conference. The Hague: Humanist.

Küpper, A. (2018). Individual learning curves in bronchoscopy: exploring skill acquisition with a low-fidelity prototype (Unpublished master's thesis). University of Twente, Enschede.

Lintern, G. (1991). An informational perspective on skill transfer in human-machine systems. Human Factors, 33, 251–266

Mayhew, D. R., Simpson, H. M., & Pak, A. (2003). Changes in collision rates among novice drivers during the first months of driving. Accident Analysis and Prevention, 35(5), 683–691. https://doi.org/10.1016/S0001-4575(02)00047-7

McKnight, A. J., & McKnight, A. S. (2003). Young novice drivers: careless or clueless? *Accident Analysis and Prevention*, *35*(6), 921–925. https://doi.org/10.1016/S0001-4575(02)00100-8

National Highway Traffic Safety Administration [NHTSA]. (1994). Research agenda for an improved novice driver education program. Report to Congress, May 31, 1994 (DOT-HS-808-161). Washington, DC: US Department of Transportation.

Noordzij, M. L., Dorrestijn, S. M., & van den Berg, I. A. (2017). An idiographic study into the physiology and self-reported mental workload of learning to drive a car. Proceedings of the Human Factors and Ergonomics Society Europe.

Park, G. D., Allen, R. W., Rosenthal, T. J., & Fiorentino, D. (2005). Training effectiveness: how does driving simulator fidelity influence driver performance? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *49*(25), 2201–2205. https://doi.org/10.1177/154193120504902518

Pollatsek, A., Narayanaan, V., Pradhan, A., & Fisher, D. L. (2006). Using eye movements to evaluate a pc-based risk awareness and perception training program on a driving simulator. *Human Factors*, *48*(3), 447–64.

Pollatsek, A., Vlakveld, W., Kappé, B., Pradhan, A. K., & Fisher, D. L. (2011) Driving simulators as training and evaluation tools: Novice drivers. In D. L., Fisher, M. Rizzo, J. Caird, & J. D. Lee. (Eds.). *Handbook of driving simulation for engineering, medicine, and psychology* (pp. 30.1-19-9). Boca Raton, Florida; CRC Press

Pradhan, A. K., Fisher, D. L., & Pollatsek, A. (2006). Risk perception training for novice drivers: evaluating duration of effects of training on a driving simulator. *Transportation Research Record: Journal of the Transportation Research Board*, *1969*(1), 58–64. https://doi.org/10.1177/0361198106196900108

Pradhan, A. K., Pollatsek, A., Knodler, M., & Fisher, D. L. (2009). Can younger drivers be trained to scan for information that will reduce their risk in roadway traffic scenarios that are hard to identify as hazardous? *Ergonomics*, *52*(6), 657–73. https://doi.org/10.1080/00140130802550232

Regan, M. A., Triggs, T. J., & Godley, S. T. (2000). Evaluation of a novice driver cd-rom based training program: a simulator study. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *44*(11), 334–337. https://doi.org/10.1177/154193120004401101

Schmettow, M. (n.d.). *New statistics for design researchers*. Github Pages. https://schmettow.github.io/New_Stats/LCM.html

Shallice, T. (1988). From neuropsychology to mental structure. Cambridge, UK: Cambridge University Press

Simons-Morton, B., & Ehsani, J. (2016). Learning to drive safely: reasonable expectations and future directions for the learner period. *Safety*, *2*(4), 20–20. https://doi.org/10.3390/safety2040020

ST SOFTWARE Simulator Systems BV (2010). Manual and curriculum driver training The Load Reduction method. Retrieved on February 26, 2020 from https://www.stsoftware.nl/data/doc/ManualCurriculum.pdf

SWOV (2010). Simulators in driver training. SWOV Fact sheet, December 2010 , Leidschendam. Retrieved from https://www.swov.nl/en/publication/simulators-driver-training

SWOV (2019). Driver training and driving tests. SWOV Fact sheet, March 2019. SWOV, The Hague.

Van Emmerik, M. (2004). Beyond the simulator : Instructions for high-performance Tasks. University of Twente, Enschede.

Van Wijk, L. (2020). Testing the effect of a training method in an online driving simulator: replication of the speed episode effect and retention. (Master Thesis). University of Twente, Enschede

Vlakveld, W. P. (2005). The use of simulators in basic driver training. SWOV. Retrieved from https://www.researchgate.net/publication/228996846_The_use_of_simulators_in_basic_driver_training

Vlakveld, W. P. (2011). Hazard anticipation of young novice drivers : assessing and enhancing the capabilities of young novice drivers to anticipate latent hazards in road and traffic situations (dissertation). Stichting Wetenschappelijk Onderzoek Verkeersveiligheid. Retrieved from https://www.swov.nl/publicatie/hazardanticipation-young-novice-drivers-assessing-and-enhancing-capabilities-young

Voskes, M. (2021). *Simulator-based driving training : the effect of speed-episodes on acquiring driving skills.* (Bachelor Thesis). University of Twente, Enschede

Wang, Y., Zhang, W., & Salvendy, G. (2010). Effects of a simulation-based training intervention on novice drivers' hazard handling performance. *Traffic Injury Prevention*, *11*(1), 16–24. https://doi.org/10.1080/15389580903390631

Weimer, C. (2019). Towards an effective MIS simulator-based training with basic laparoscopic tasks: The impact of time pressure on the learning process thesis (Unpublished master's thesis). University of Twente, Enschede.

Weiss, T., Petzoldt, T., Bannert, M., & Krems, J. (2013). Calibration as side effect? computer-based learning in driver education and the adequacy of driving-task-related self-assessments. *Transportation Research Part F: Psychology and Behaviour*, *17*, 63–74. https://doi.org/10.1016/j.trf.2012.10.001

Westerhof, M. (2018). Towards predicting bronchoscopic skill acquisition using basic bronchoscopic simulator tasks (Unpublished master's thesis). University of Twente, Enschede.

Wheeler, W. A., & Triggs, T. J. (1996). A task analytical view of simulator based training for drivers. Proceedings of the road safety research and enforcement conference “e□ective partnerships” (pp. 217–221), Coogee Beach, New South Wales, Australia.

# Appendices

## Appendix A: Rscript Phase 2: exploration of highly controlled data

```
D_SimPar <- read_delim("~/Master/Thesis/Data/Bachelor data/Data_Bachelor_Ma
ster.csv",
        ";", escape_double = FALSE, trim_ws = TRUE, col_types = cols(ToT =
col_number(), X8 = col_skip()))

## Warning: Missing column names filled in: 'X8' [8]

D_SimPar %>%
  sample_n(10)

D_SimPar %>%
    ggplot(aes(x = trial, y = ToT)) +
    geom_smooth(se = F, scale = "free_y") +
  geom_point() +
    facet_wrap(~Participant)

## Warning: Ignoring unknown parameters: scale

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

D_SimPar %>%
    ggplot(aes(x = trial, y = ToT, group = Participant)) +
    geom_smooth(se = F)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

D_SimPar %>%
    ggplot(aes(x = trial, y = ToT, group = Participant, colour = Training))
+
    geom_smooth(se = F)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

D_SimPar %>%
    ggplot(aes(x = trial, y = Nld, group = Participant)) +
    geom_smooth(se = F) +
    facet_wrap(~Training)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 204 rows containing non-finite values (stat_smooth).

D_SimPar %>%
    ggplot(aes(x = trial, y = Nld, group = Participant, colour = Training))
+
    geom_smooth(se = F)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 204 rows containing non-finite values (stat_smooth).

D_SimPar %>%
    ggplot(aes(x = trial, y = Nc, group = Participant, colour = Training))
```

```
+
    geom_smooth(se = F)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 204 rows containing non-finite values (stat_smooth).

D_SimPar %>%
    ggplot(aes(x = trial, y = Nc, group = Participant)) +
    geom_smooth(se = F) +
    facet_wrap(~Training)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 204 rows containing non-finite values (stat_smooth).

D_SimPar %>%
    ggplot(aes(x = trial, y = Nc, group = Participant)) +
    geom_smooth(se = F) +
    facet_wrap(~Experience)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 204 rows containing non-finite values (stat_smooth).

D_SimPar %>%
    ggplot(aes(x = trial, y = Nld, group = Participant)) +
    geom_smooth(se = F) +
    facet_wrap(~Experience)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning: Removed 204 rows containing non-finite values (stat_smooth).

D_SimPar %>%
    filter(Training == "Accuracy") %>%
    ggplot(aes(x = trial, y = ToT, group = Participant)) +
    geom_smooth(se = F) +
    facet_wrap(~Experience)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

D_SimPar %>%
    filter(Training == "Accuracy") %>%
    ggplot(aes(x = trial, y = Nld, group = Participant)) +
    geom_smooth(se = F) +
    facet_wrap(~Experience)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

D_SimParAcc <-
  D_SimPar %>%
  filter(Training == "Accuracy")

D_SimParAcc %>%
  sample_n(10)
```

#MODEL ESTIMATION

```
F_lacy_prior <- c(set_prior("normal(5.25, 0.576)", nlpar = "ampl"),
                  set_prior("normal(-2.76, 2.07)", nlpar = "ctch"),
                  set_prior("normal(1.84, 0.576)", nlpar = "asym"))

F_lacy_prior_1 <- c(set_prior("normal(5.25, 0.875)", nlpar = "ampl"),
                  set_prior("normal(-2.76, 2.07)", nlpar = "ctch"),
                  set_prior("normal(1.84, 0.576)", nlpar = "asym"))

F_lacy_prior_2 <- c(set_prior("normal(5.25, 1.05)", nlpar = "ampl"),
                  set_prior("normal(-2.76, 2.07)", nlpar = "ctch"),
                  set_prior("normal(1.84, 0.576)", nlpar = "asym"))

F_lacy <- formula(ToT ~ exp(asym) + exp(ampl) * inv_logit((1-ctch))^trial)

F_acy_ef_4 <- list(formula(ampl ~ 1 + Experience + (1|Participant)),
                   formula(ctch ~ 1 + (1|Participant)),
                   formula(asym ~ 1 + (1|Participant)))

M_7 <-
  D_SimParAcc %>%
  brm(bf(F_lacy,
         flist = F_acy_ef_4,
         nl = T),
      prior = F_lacy_prior_1,
      family = Gamma(link = identity),
      iter = 4000,
      data = .)

## Compiling Stan program...

## Start sampling

## Warning: There were 73 divergent transitions after warmup. See
## http://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup
## to find out why this is a problem and how to eliminate them.

## Warning: Examine the pairs() plot to diagnose sampling problems

coef(M_7, mean.func = exp)
```

*Coefficient estimates with 95% credibility limits*

| parameter | type | nonlin | fixef | re_factor | re_entity | center | lower | upper |
|---|---|---|---|---|---|---|---|---|
| b_ampl_Intercept | fixef | ampl | Intercept | NA | NA | 9.5936844 | 3.8164777 | 22.7077998 |
| b_ampl_ExperienceInexperienced | fixef | ampl | ExperienceInexperienced | NA | NA | 3.7101714 | 1.2822731 | 35.1462193 |
| b_ctch_Intercept | fixef | ctch | Intercept | NA | NA | 0.4304830 | 0.0660745 | 2.2480975 |
| b_asym_Intercept | fixef | asym | Intercept | NA | NA | 63.1273758 | 58.2648549 | 66.8221400 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| r_Participant__ampl[3,Intercept] | ranef | ampl | Intercept | Participant | 3 | 0.5719200 | 0.0224164 | 2.2452980 |
| r_Participant__ampl[5,Intercept] | ranef | ampl | Intercept | Participant | 5 | 0.8578548 | 0.0647212 | 2.5702140 |
| r_Participant__ampl[12,Intercept] | ranef | ampl | Intercept | Participant | 12 | 0.8467500 | 0.0567920 | 3.7854208 |
| r_Participant__ampl[20,Intercept] | ranef | ampl | Intercept | Participant | 20 | 1.3540433 | 0.1002308 | 2.8992582 |
| r_Participant__ampl[24,Intercept] | ranef | ampl | Intercept | Participant | 24 | 0.5195112 | 0.0353933 | 1.2075160 |
| r_Participant__ampl[27,Intercept] | ranef | ampl | Intercept | Participant | 27 | 0.6279157 | 0.0413148 | 1.5138297 |
| r_Participant__ampl[28,Intercept] | ranef | ampl | Intercept | Participant | 28 | 0.3379070 | 0.0021905 | 1.5900211 |
| r_Participant__ampl[29,Intercept] | ranef | ampl | Intercept | Participant | 29 | 0.4345902 | 0.0249093 | 1.0784712 |
| r_Participant__ampl[30,Intercept] | ranef | ampl | Intercept | Participant | 30 | 1.0211906 | 0.1727637 | 6.2246444 |
| r_Participant__ampl[31,Intercept] | ranef | ampl | Intercept | Participant | 31 | 2.0132844 | 0.7781757 | 6.0904874 |
| r_Participant__ampl[32,Intercept] | ranef | ampl | Intercept | Participant | 32 | 0.5082666 | 0.0291096 | 1.1728936 |
| r_Participant__ampl[33,Intercept] | ranef | ampl | Intercept | Participant | 33 | 0.6886364 | 0.1293093 | 2.3410304 |
| r_Participant__ampl[35,Intercept] | ranef | ampl | Intercept | Participant | 35 | 0.5022472 | 0.0089201 | 2.1787732 |
| r_Participant__ampl[36,Intercept] | ranef | ampl | Intercept | Participant | 36 | 1.6780066 | 0.6955674 | 4.4067702 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| r_Participant__ampl[39,Intercept] | ranef | ampl | Intercept | Participant | 39 | 2.5293614 | 1.0690883 | 6.4419275 |
| r_Participant__ampl[40,Intercept] | ranef | ampl | Intercept | Participant | 40 | 0.3429966 | 0.0197034 | 1.3418345 |
| r_Participant__ctch[3,Intercept] | ranef | ctch | Intercept | Participant | 3 | 1.2063928 | 0.0024876 | 364.7019670 |
| r_Participant__ctch[5,Intercept] | ranef | ctch | Intercept | Participant | 5 | 6.3171438 | 0.8770068 | 61.7415517 |
| r_Participant__ctch[12,Intercept] | ranef | ctch | Intercept | Participant | 12 | 10.5151801 | 1.0563509 | 166.8539428 |
| r_Participant__ctch[20,Intercept] | ranef | ctch | Intercept | Participant | 20 | 4.0455579 | 0.7213391 | 27.7810890 |
| r_Participant__ctch[24,Intercept] | ranef | ctch | Intercept | Participant | 24 | 2.6559216 | 0.4586747 | 19.7783900 |
| r_Participant__ctch[27,Intercept] | ranef | ctch | Intercept | Participant | 27 | 3.9802075 | 0.6726269 | 32.9970384 |
| r_Participant__ctch[28,Intercept] | ranef | ctch | Intercept | Participant | 28 | 0.1160097 | 0.0008306 | 1473.0994494 |
| r_Participant__ctch[29,Intercept] | ranef | ctch | Intercept | Participant | 29 | 0.1693461 | 0.0219629 | 1.2551516 |
| r_Participant__ctch[30,Intercept] | ranef | ctch | Intercept | Participant | 30 | 4.2080839 | 0.1006847 | 95.2197999 |
| r_Participant__ctch[31,Intercept] | ranef | ctch | Intercept | Participant | 31 | 0.7344288 | 0.0689020 | 8.1023671 |
| r_Participant__ctch[32,Intercept] | ranef | ctch | Intercept | Participant | 32 | 0.1508321 | 0.0229511 | 0.9741203 |
| r_Participant__ctch[33,Intercept] | ranef | ctch | Intercept | Participant | 33 | 0.8401029 | 0.0390905 | 45.7246070 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| r_Participant__ctch[35,Intercept] | ranef | ctch | Intercept | Participant | 35 | 0.5114837 | 0.0009512 | 938.6552413 |
| r_Participant__ctch[36,Intercept] | ranef | ctch | Intercept | Participant | 36 | 0.4693697 | 0.0698949 | 3.3307233 |
| r_Participant__ctch[39,Intercept] | ranef | ctch | Intercept | Participant | 39 | 0.4292605 | 0.0739247 | 2.8021512 |
| r_Participant__ctch[40,Intercept] | ranef | ctch | Intercept | Participant | 40 | 0.7959421 | 0.0563292 | 28.6892246 |
| r_Participant__asym[3,Intercept] | ranef | asym | Intercept | Participant | 3 | 1.0693271 | 0.9040377 | 1.1582604 |
| r_Participant__asym[5,Intercept] | ranef | asym | Intercept | Participant | 5 | 0.9823121 | 0.9250029 | 1.0631700 |
| r_Participant__asym[12,Intercept] | ranef | asym | Intercept | Participant | 12 | 0.9528531 | 0.8974264 | 1.0314282 |
| r_Participant__asym[20,Intercept] | ranef | asym | Intercept | Participant | 20 | 1.0559974 | 0.9945104 | 1.1443880 |
| r_Participant__asym[24,Intercept] | ranef | asym | Intercept | Participant | 24 | 1.0625233 | 1.0010995 | 1.1528203 |
| r_Participant__asym[27,Intercept] | ranef | asym | Intercept | Participant | 27 | 0.9800887 | 0.9219529 | 1.0627195 |
| r_Participant__asym[28,Intercept] | ranef | asym | Intercept | Participant | 28 | 1.0961570 | 0.8561201 | 1.2138822 |
| r_Participant__asym[29,Intercept] | ranef | asym | Intercept | Participant | 29 | 0.9889381 | 0.8121807 | 1.1136354 |
| r_Participant__asym[30,Intercept] | ranef | asym | Intercept | Participant | 30 | 0.9740775 | 0.8982001 | 1.0546682 |
| r_Participant__asym[31,Intercept] | ranef | asym | Intercept | Participant | 31 | 1.2505541 | 0.9811333 | 1.3750148 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| r_Participant__asym[32,Intercept] | ranef | asym | Intercept | Participant | 32 | 0.9497017 | 0.7884067 | 1.0675774 |
| r_Participant__asym[33,Intercept] | ranef | asym | Intercept | Participant | 33 | 0.9606982 | 0.8445289 | 1.0349494 |
| r_Participant__asym[35,Intercept] | ranef | asym | Intercept | Participant | 35 | 1.0629509 | 0.9003802 | 1.1504440 |
| r_Participant__asym[36,Intercept] | ranef | asym | Intercept | Participant | 36 | 0.9706301 | 0.8580963 | 1.0564173 |
| r_Participant__asym[39,Intercept] | ranef | asym | Intercept | Participant | 39 | 0.8725559 | 0.7786111 | 0.9518561 |
| r_Participant__asym[40,Intercept] | ranef | asym | Intercept | Participant | 40 | 1.0103154 | 0.8447700 | 1.0963212 |

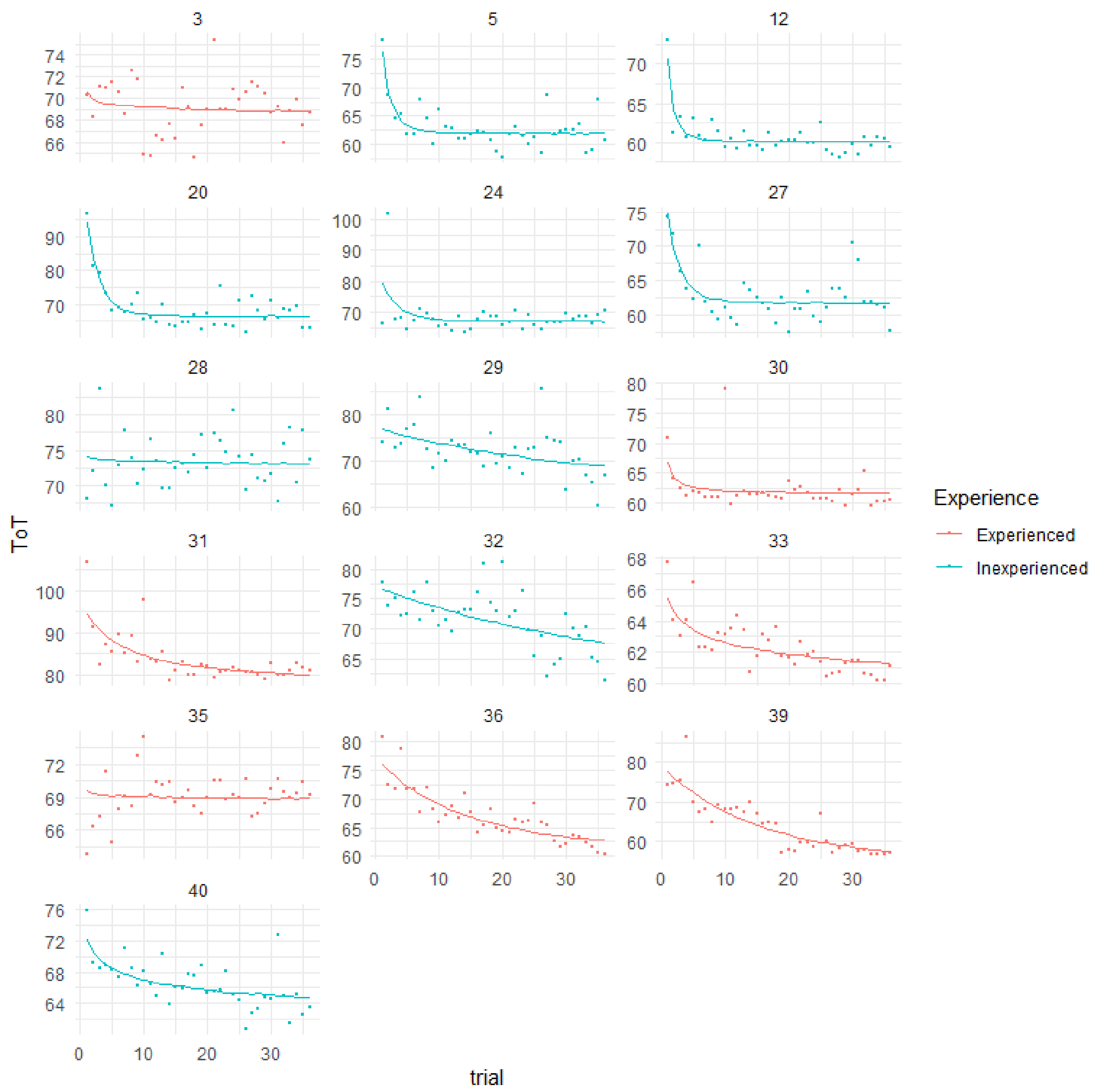
```
P_M_7 <- posterior(M_7)
PP_M_7 <- post_pred(M_7)

T_pred_M_7 <- PP_M_7 %>%
  group_by(Obs) %>%
  summarize(center = median(value))

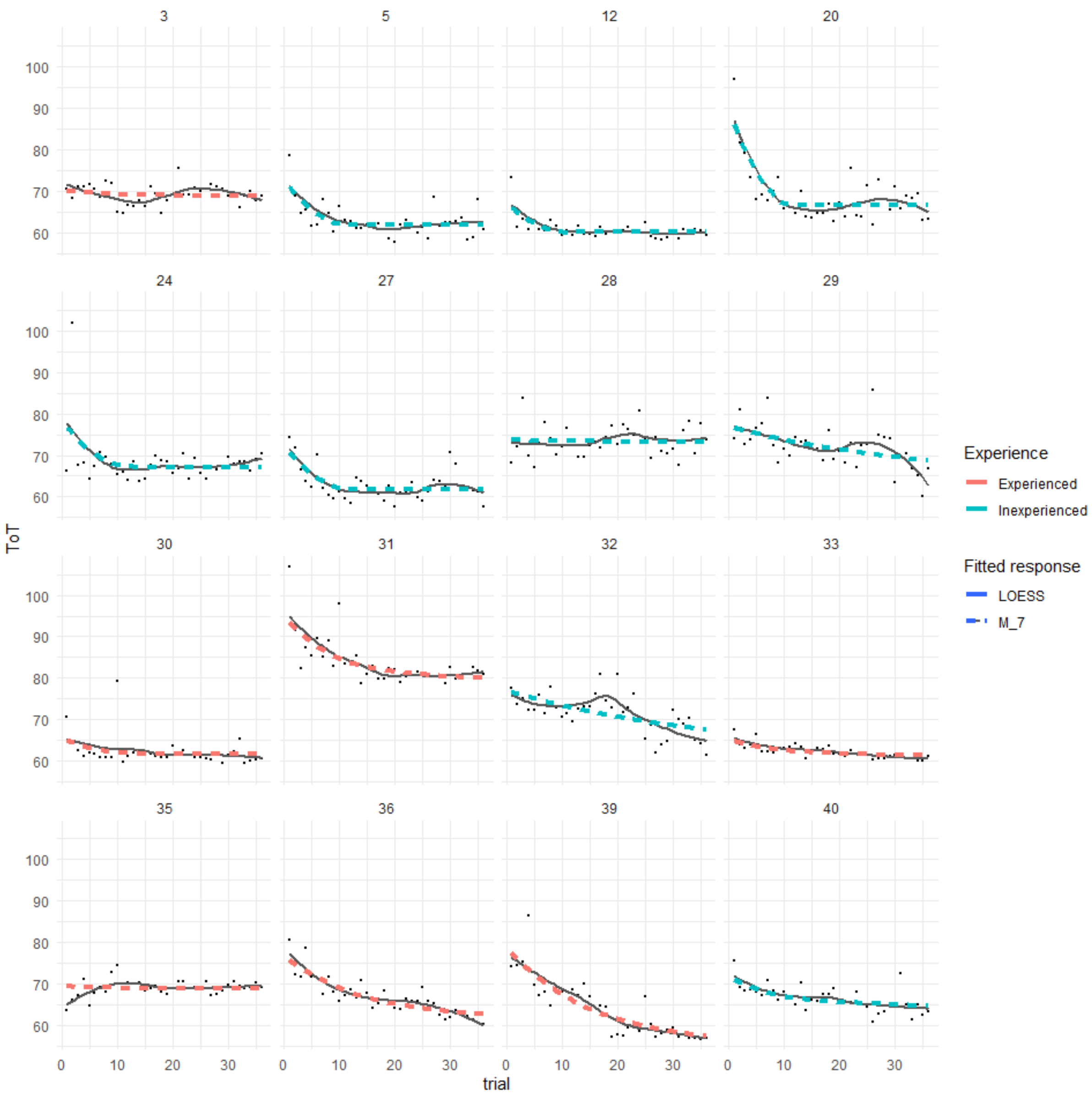
D_SimParAcc$M_7 <- T_pred_M_7$center
D_SimParAcc$M_7_resid <- D_SimParAcc$ToT - D_SimParAcc$M_7

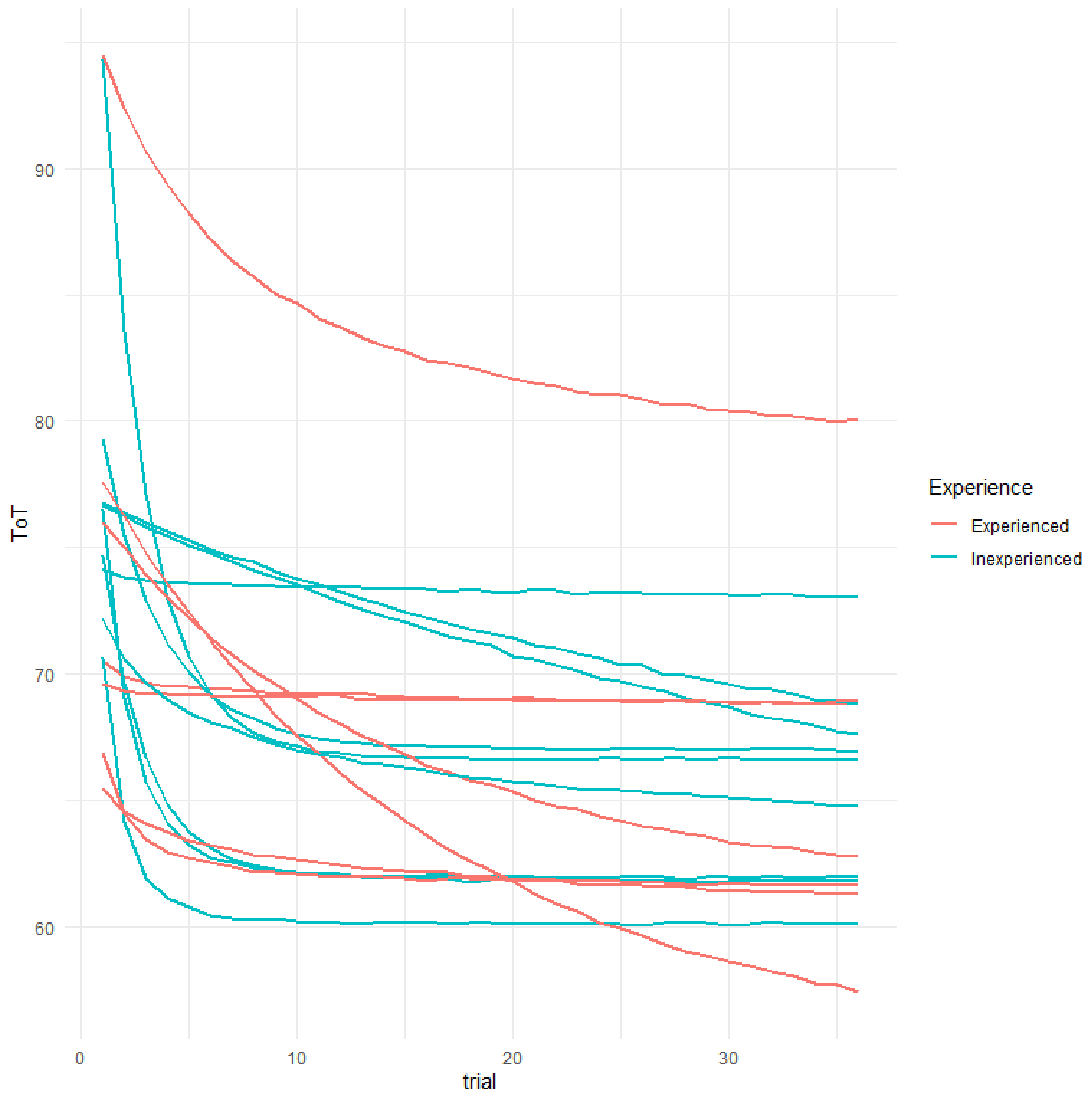
D_M_7 <-
  as_tibble(M_7$data) %>%
  mutate(M_7 = T_pred_M_7$center)

D_SimParAcc %>%
  ggplot(aes(x = trial, y = ToT, col = Experience)) +
  facet_wrap(~ Participant, nrow = 7, scales = "free_y") +
  geom_point(size = .2) +
  geom_line(aes(y = M_7)) +
  theme_minimal()
```

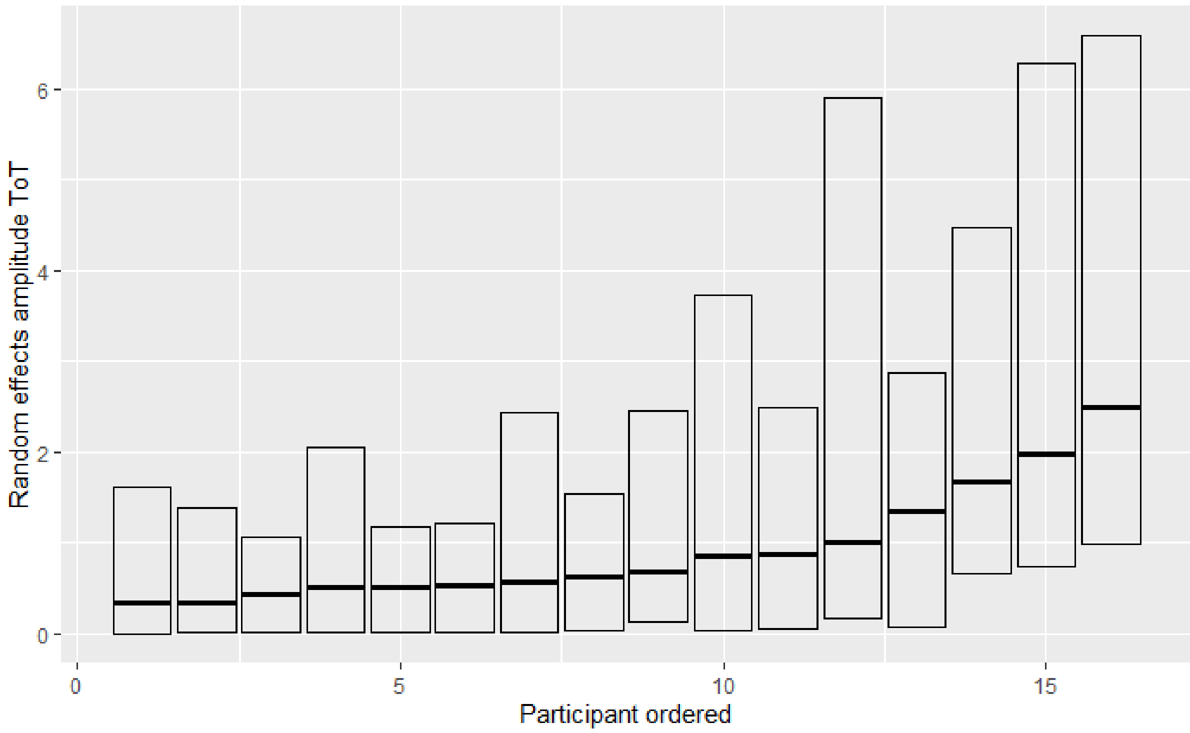



```
D_SimParAcc <- D_SimParAcc %>%
  mutate(M_7 = T_pred_M_7$center)

```{r, fig.height= 10, fig.width= 10}

D_SimParAcc %>%

  ggplot(aes(x = trial, y = ToT)) +

  geom_point(size = 0.5) +

  geom_smooth(aes(linetype = "LOESS"), se = F, color = "gray34", size = 0.7
5) +

  geom_smooth(aes(y = M_7, color = Experience, linetype = "M_7"), se = F, s
ize = 1.5) +

  labs(linetype = "Fitted response", M_7 = "Predicted", LOESS = "Observed")
+

  facet_wrap(~Participant) +

  theme_minimal()
```
```

```
```## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
D_SimParAcc %>%
  ggplot(aes(x = trial, y = ToT, group = Participant, col = Experience)) +
  geom_line(aes(y = M_7), size = 1) +
  theme_minimal()
```
```




```{r}
coef(M_7, mean.func = exp) %>%
  filter(type == "ranef", nonlin == "ampl") %>%
  mutate(re_entity = rank(center)) %>%
  ggplot(aes(x = re_entity, ymin = lower, y = center, ymax = upper)) +
  geom_crossbar() +
  labs(y = "Random effects amplitude ToT", x = "Participant ordered" )
```




NUMBER OF LANE DEPARTURES

```
F_lacy_Nld <- formula(Nld ~ exp(asym) + exp(ampl) * inv_logit((1-ctch))^tri
al)

M_Test_Nld_exp_1 <-
  D_SimParAcc %>%
  brm(bf(F_lacy_Nld,
         flist = F_acy_ef_4,
         nl = T),
      prior = F_lacy_prior_2,
      family = poisson(link = identity), iter = 4000,
      data = .)

## Compiling Stan program...

## Start sampling

## Warning: There were 31 divergent transitions after warmup. See
## http://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup
## to find out why this is a problem and how to eliminate them.

## Warning: Examine the pairs() plot to diagnose sampling problems

## Warning: Bulk Effective Samples Size (ESS) is too low, indicating poster
ior means and medians may be unreliable.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#bulk-ess

coef(M_Test_Nld_exp_1, mean.func = exp)
```

*Coefficient estimates with 95% credibility limits*

| parameter | type | nonlin | fixef | re_factor | re_entity | center | lower | upper |
|---|---|---|---|---|---|---|---|---|
| b_ampl_Intercept | fixef | ampl | Intercept | NA | NA | 2.5669188 | 1.0046929 | 5.2330310 |
| b_ampl_ExperienceInexperienced | fixef | ampl | ExperienceInexperienced | NA | NA | 1.9772203 | 0.9254745 | 7.0010166 |
| b_ctch_Intercept | fixef | ctch | Intercept | NA | NA | 1.0193128 | 0.1465544 | 5.4653561 |
| b_asym_Intercept | fixef | asym | Intercept | NA | NA | 1.3383014 | 0.8726383 | 2.2795326 |
| r_Participant__ampl[3,Intercept] | ranef | ampl | Intercept | Participant | 3 | 0.9649559 | 0.1202078 | 1.7413016 |
| r_Participant__ampl[5,Intercept] | ranef | ampl | Intercept | Participant | 5 | 0.9818646 | 0.2651434 | 1.8225039 |
| r_Participant__ampl[12,Intercept] | ranef | ampl | Intercept | Participant | 12 | 0.9340415 | 0.2539352 | 1.5368235 |
| r_Participant__ampl[20,Intercept] | ranef | ampl | Intercept | Participant | 20 | 0.9649391 | 0.0422965 | 1.7229734 |
| r_Participant__ampl[24,Intercept] | ranef | ampl | Intercept | Participant | 24 | 1.0014977 | 0.3845757 | 1.8981063 |
| r_Participant__ampl[27,Intercept] | ranef | ampl | Intercept | Participant | 27 | 0.9663013 | 0.0716409 | 1.7732598 |
| r_Participant__ampl[28,Intercept] | ranef | ampl | Intercept | Participant | 28 | 0.9946865 | 0.3604393 | 1.8983840 |
| r_Participant__ampl[29,Intercept] | ranef | ampl | Intercept | Participant | 29 | 0.9742173 | 0.2527391 | 1.6851309 |
| r_Participant__ampl[30,Intercept] | ranef | ampl | Intercept | Participant | 30 | 0.9589915 | 0.1295602 | 1.7673588 |
| r_Participant__ampl[31,Intercept] | ranef | ampl | Intercept | Participant | 31 | 0.9831695 | 0.2866223 | 2.0507512 |

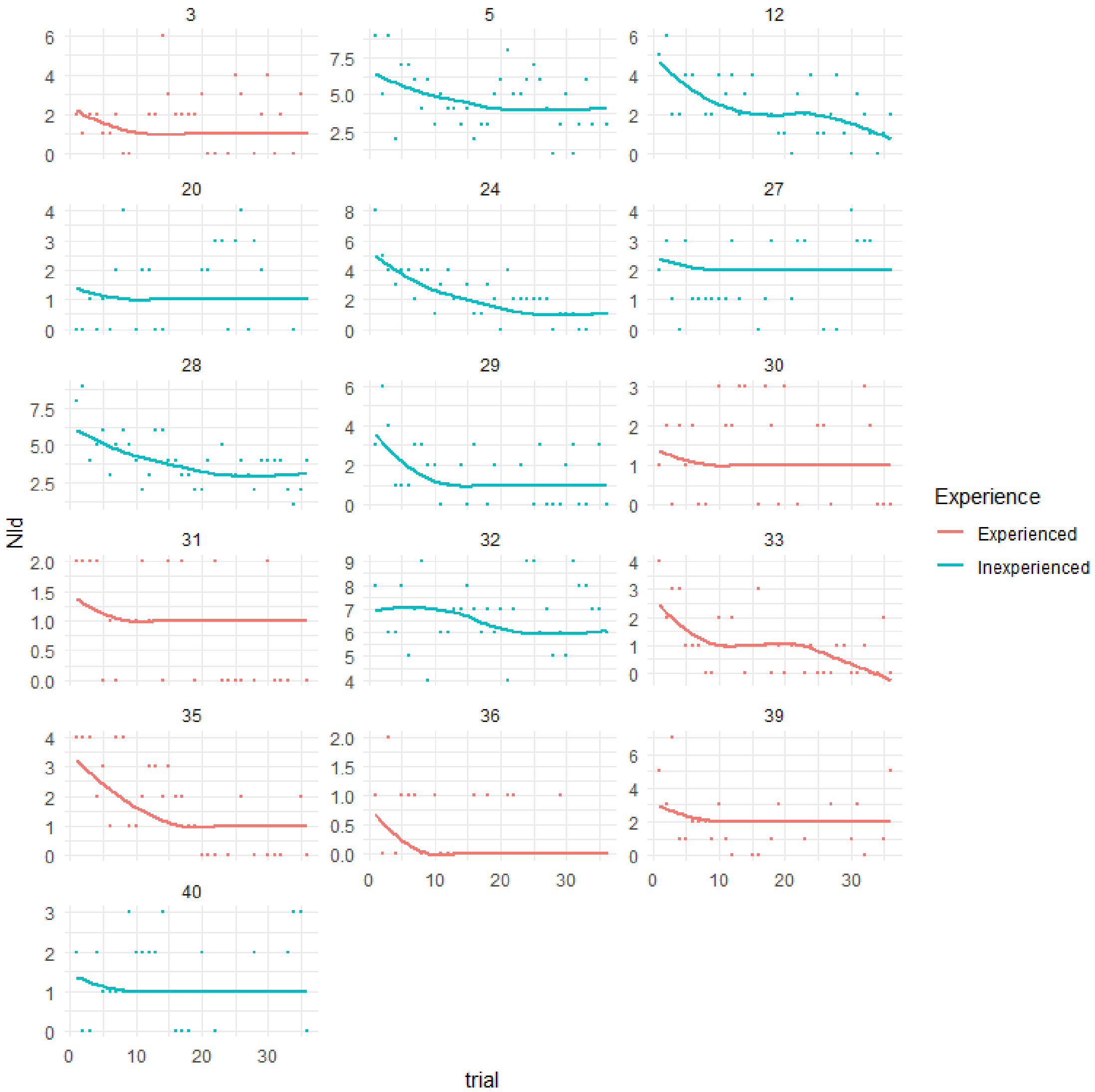
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| r_Participant__ampl[32,Intercept] | ranef | ampl | Intercept | Participant | 32 | 0.9994587 | 0.1381888 | 1.7529003 |
| r_Participant__ampl[33,Intercept] | ranef | ampl | Intercept | Participant | 33 | 1.0262154 | 0.5652447 | 2.9203119 |
| r_Participant__ampl[35,Intercept] | ranef | ampl | Intercept | Participant | 35 | 1.0968491 | 0.7270852 | 3.6915294 |
| r_Participant__ampl[36,Intercept] | ranef | ampl | Intercept | Participant | 36 | 0.9709845 | 0.1937871 | 1.8394362 |
| r_Participant__ampl[39,Intercept] | ranef | ampl | Intercept | Participant | 39 | 1.0186883 | 0.4334252 | 3.0345950 |
| r_Participant__ampl[40,Intercept] | ranef | ampl | Intercept | Participant | 40 | 0.9674657 | 0.0538010 | 1.8654701 |
| r_Participant__ctch[3,Intercept] | ranef | ctch | Intercept | Participant | 3 | 4.5492891 | 0.0217775 | 681.2526431 |
| r_Participant__ctch[5,Intercept] | ranef | ctch | Intercept | Participant | 5 | 0.1739470 | 0.0073578 | 7.7426016 |
| r_Participant__ctch[12,Intercept] | ranef | ctch | Intercept | Participant | 12 | 0.3736587 | 0.0479965 | 5.0217484 |
| r_Participant__ctch[20,Intercept] | ranef | ctch | Intercept | Participant | 20 | 24.9356100 | 0.3487958 | 2732.1230230 |
| r_Participant__ctch[24,Intercept] | ranef | ctch | Intercept | Participant | 24 | 0.2965541 | 0.0477761 | 2.2510189 |
| r_Participant__ctch[27,Intercept] | ranef | ctch | Intercept | Participant | 27 | 9.2885098 | 0.3154155 | 1148.1995494 |
| r_Participant__ctch[28,Intercept] | ranef | ctch | Intercept | Participant | 28 | 0.2354033 | 0.0229213 | 2.9858299 |
| r_Participant__ctch[29,Intercept] | ranef | ctch | Intercept | Participant | 29 | 0.9276379 | 0.1349490 | 8.6050623 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| r_Participant__ctch[30,Intercept] | ranef | ctch | Intercept | Participant | 30 | 5.2971891 | 0.0592066 | 1235.5383274 |
| r_Participant__ctch[31,Intercept] | ranef | ctch | Intercept | Participant | 31 | 1.4331329 | 0.1283244 | 126.9867980 |
| r_Participant__ctch[32,Intercept] | ranef | ctch | Intercept | Participant | 32 | 0.0307839 | 0.0001749 | 45.9660065 |
| r_Participant__ctch[33,Intercept] | ranef | ctch | Intercept | Participant | 33 | 0.7233554 | 0.1032212 | 6.8502098 |
| r_Participant__ctch[35,Intercept] | ranef | ctch | Intercept | Participant | 35 | 0.2852903 | 0.0470773 | 2.2054622 |
| r_Participant__ctch[36,Intercept] | ranef | ctch | Intercept | Participant | 36 | 3.0532409 | 0.2931557 | 616.2280553 |
| r_Participant__ctch[39,Intercept] | ranef | ctch | Intercept | Participant | 39 | 1.0858989 | 0.0615592 | 19.2650559 |
| r_Participant__ctch[40,Intercept] | ranef | ctch | Intercept | Participant | 40 | 13.4474381 | 0.3646939 | 2025.3778861 |
| r_Participant__asym[3,Intercept] | ranef | asym | Intercept | Participant | 3 | 1.1174112 | 0.3725567 | 1.8290425 |
| r_Participant__asym[5,Intercept] | ranef | asym | Intercept | Participant | 5 | 1.7452925 | 0.3131903 | 4.0080156 |
| r_Participant__asym[12,Intercept] | ranef | asym | Intercept | Participant | 12 | 1.0075651 | 0.2565346 | 1.9410036 |
| r_Participant__asym[20,Intercept] | ranef | asym | Intercept | Participant | 20 | 0.9272568 | 0.4792665 | 1.5112540 |
| r_Participant__asym[24,Intercept] | ranef | asym | Intercept | Participant | 24 | 0.7071507 | 0.1773432 | 1.4661948 |
| r_Participant__asym[27,Intercept] | ranef | asym | Intercept | Participant | 27 | 1.2782487 | 0.6631250 | 2.0520808 |

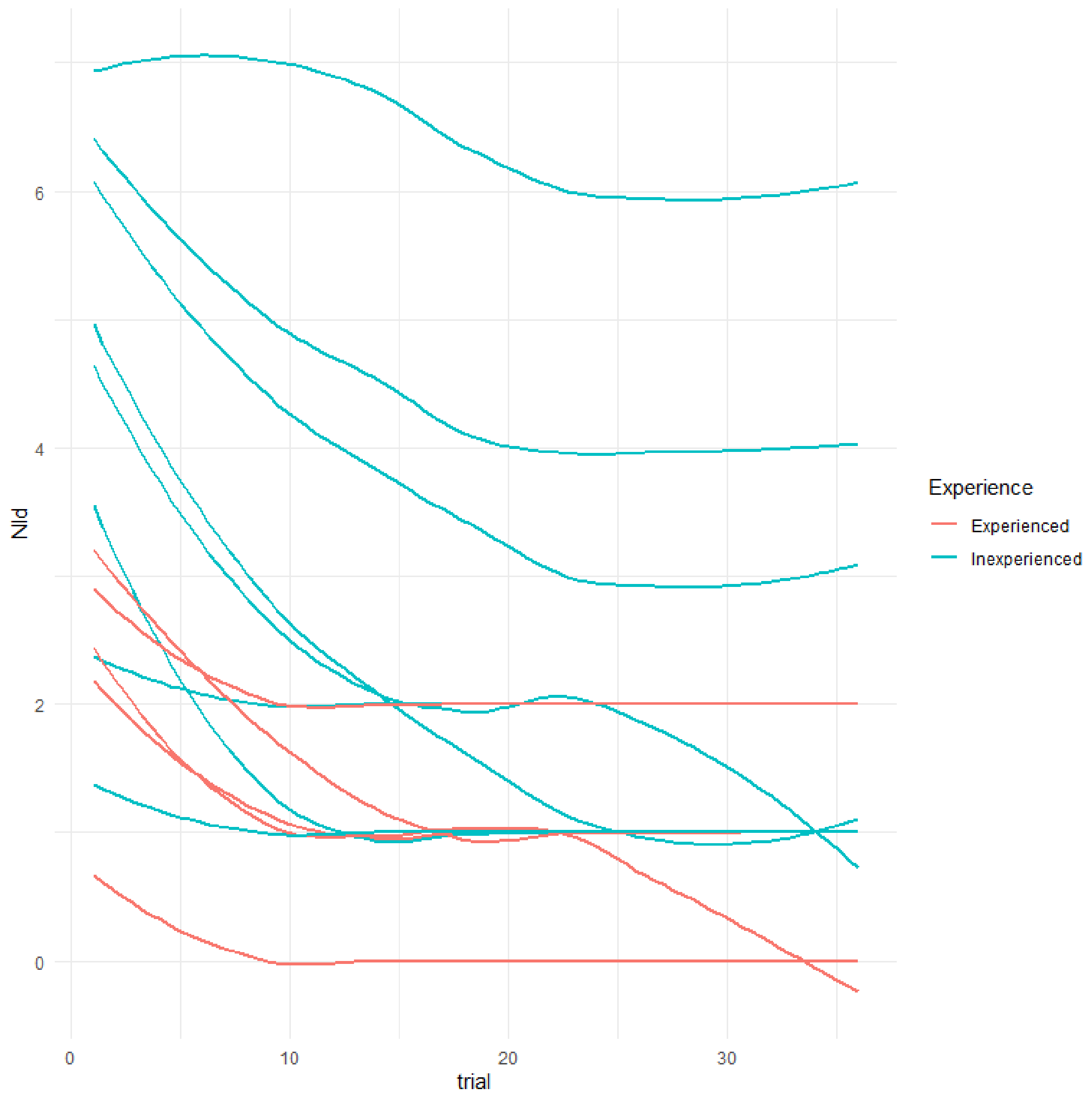
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| r_Participant__asym[28,Intercept] | ranef | asym | Intercept | Participant | 28 | 1.5777789 | 0.3537748 | 3.2077841 |
| r_Participant__asym[29,Intercept] | ranef | asym | Intercept | Participant | 29 | 0.8047921 | 0.3757183 | 1.3912537 |
| r_Participant__asym[30,Intercept] | ranef | asym | Intercept | Participant | 30 | 0.8961852 | 0.3044669 | 1.4799771 |
| r_Participant__asym[31,Intercept] | ranef | asym | Intercept | Participant | 31 | 0.5581470 | 0.2013306 | 0.9881416 |
| r_Participant__asym[32,Intercept] | ranef | asym | Intercept | Participant | 32 | 2.0097228 | 0.3908040 | 5.9721175 |
| r_Participant__asym[33,Intercept] | ranef | asym | Intercept | Participant | 33 | 0.4880841 | 0.1675197 | 0.9130086 |
| r_Participant__asym[35,Intercept] | ranef | asym | Intercept | Participant | 35 | 0.5282689 | 0.1336718 | 1.1082363 |
| r_Participant__asym[36,Intercept] | ranef | asym | Intercept | Participant | 36 | 0.2554529 | 0.0863497 | 0.5238289 |
| r_Participant__asym[39,Intercept] | ranef | asym | Intercept | Participant | 39 | 1.2502402 | 0.5818614 | 2.0852853 |
| r_Participant__asym[40,Intercept] | ranef | asym | Intercept | Participant | 40 | 0.9156072 | 0.4731199 | 1.4935520 |

```
P_M_Test_Nld_exp_1 <- posterior(M_Test_Nld_exp_1)
PP_M_Test_Nld_exp_1 <- post_pred(M_Test_Nld_exp_1)

T_pred_M_Test_Nld_exp_1 <- PP_M_Test_Nld_exp_1 %>%
  group_by(Obs) %>%
  summarize(center = median(value))

D_SimParAcc$M_Test_Nld_exp_1 <- T_pred_M_Test_Nld_exp_1$center
D_SimParAcc$M_Test_Nld_exp_1_resid <- D_SimParAcc$Nld - D_SimParAcc$M_Test_
Nld_exp_1

D_M_Test_Nld_exp_1 <-
  as_tibble(M_Test_Nld_exp_1$data) %>%
  mutate(M_Test_Nld_exp_1 = T_pred_M_Test_Nld_exp_1$center)

D_SimParAcc %>%
  ggplot(aes(x = trial, y = Nld, col = Experience)) +
  facet_wrap(~ Participant, nrow = 7, scales = "free_y") +
```

```
  geom_point(size = .2) +
  geom_smooth(aes(y = M_Test_Nld_exp_1), se = F) +
  theme_minimal()

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

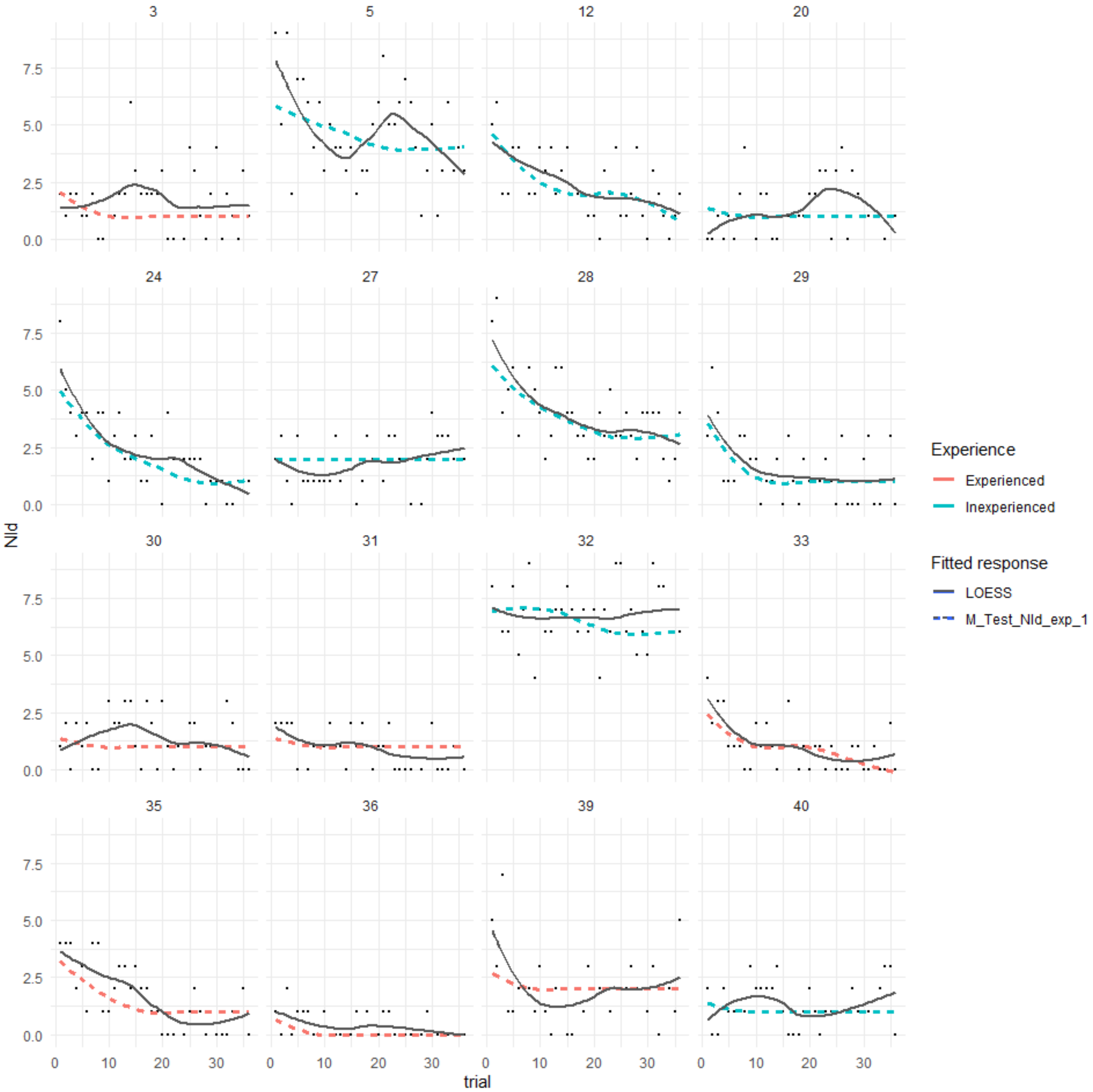



```
D_SimParAcc %>%
  ggplot(aes(x = trial, y = Nld, col = Experience, group = Participant)) +
  geom_smooth(aes(y = M_Test_Nld_exp_1), se = F) +
  theme_minimal()

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

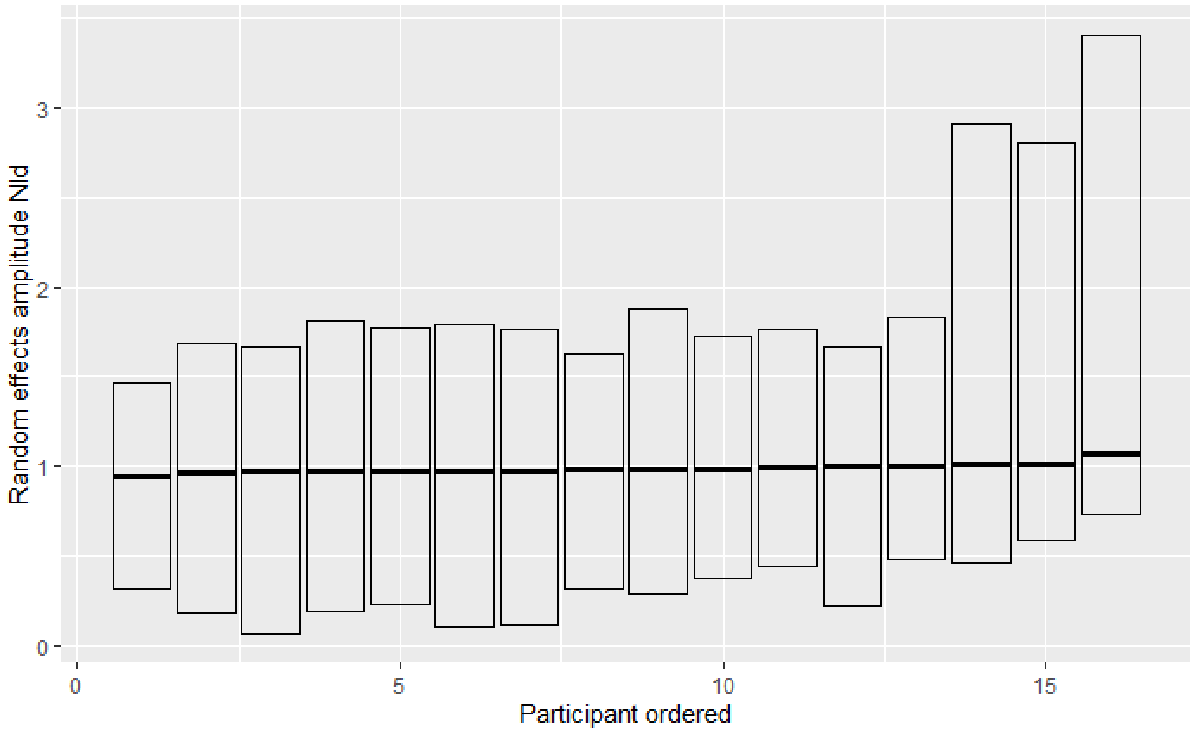



```
D_SimParAcc <- D_SimParAcc %>%
  mutate(M_Test_Nld_exp_1 = T_pred_M_Test_Nld_exp_1$center)

D_SimParAcc %>%

  ggplot(aes(x = trial, y = Nld)) +

  geom_point(size = 0.5) +

  geom_smooth(aes(linetype = "LOESS"), se = F, color = "gray34", size = 0.7
5) +

  geom_smooth(aes(y = M_Test_Nld_exp_1, linetype = "M_Test_Nld_exp_1", colo
r = Experience), se = F, size = 1.2) +

  labs(linetype = "Fitted response", M_Test_Nld_exp_1 = "Predicted", LOESS
= "Observed") +

  facet_wrap(~Participant) +

  theme_minimal()
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
coef(M_Test_Nld_exp_1, mean.func = exp) %>%
  filter(type == "ranef", nonlin == "ampl") %>%
  mutate(re_entity = rank(center)) %>%
  ggplot(aes(x = re_entity, ymin = lower, y = center, ymax = upper)) +
  geom_crossbar() +
  labs(y = "Random effects amplitude Nld", x = "Participant ordered" )
```



Random effects amplitude NId
0
1
2
3
0
5
10
15
Participant ordered

**Appendix B: Rscript Phase 3: exploration of medium controlled data**

```
D_OST <- read_csv("~/Master/Thesis/Data/AH_1SECONDS_1.csv")

##
## -- Column specification ------------------------------------------------
--------
## cols(
##   Part = col_double(),
##   Training = col_character(),
##   Block = col_double(),
##   block = col_double(),
##   Blk_type = col_character(),
##   trial = col_double(),
##   crashes = col_double(),
##   speed = col_double(),
##   steer = col_double(),
##   ToT = col_double()
## )

D_OST %>%
  sample_n(10)
```

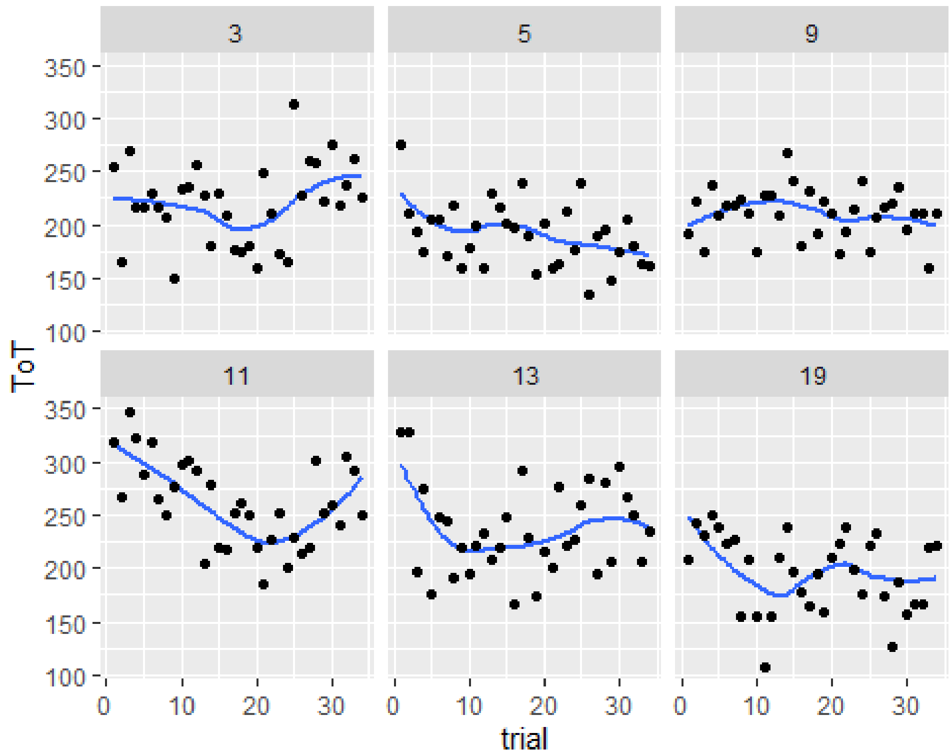
| Part | Training | Block | block | Blk_type | trial | crashes | speed | steer | ToT |
|---|---|---|---|---|---|---|---|---|---|
| 5 | Accuracy | 3 | 3 | Accuracy | 24 | 0 | 11.615480 | -0.012670 | 176 |
| 9 | Accuracy | 4 | 4 | Accuracy | 27 | 0 | 9.225628 | -0.022370 | 217 |
| 13 | Accuracy | 4 | 4 | Accuracy | 32 | 0 | 7.981857 | 0.014188 | 250 |
| 3 | Accuracy | 2 | 2 | Accuracy | 9 | 0 | 13.488240 | 0.001329 | 150 |
| 2 | Speed | 3 | 3 | Accuracy | 22 | 0 | 11.490860 | -0.027510 | 176 |
| 19 | Accuracy | 1 | 1 | Accuracy | 8 | 0 | 12.823300 | 0.003369 | 155 |
| 11 | Accuracy | 2 | 2 | Accuracy | 14 | 0 | 7.228878 | 0.018741 | 279 |
| 17 | Speed | 3 | 3 | Accuracy | 24 | 0 | 11.214700 | 0.004707 | 178 |
| 10 | Speed | 1 | 1 | Accuracy | 4 | 1 | 13.248980 | 0.001389 | 151 |
| 2 | Speed | 3 | 3 | Accuracy | 20 | 0 | 10.686310 | 0.018840 | 188 |

```
D_OST_ACC <- D_OST %>%
  filter(Training == "Accuracy")

D_OST_ACC %>%
  sample_n(10)
```

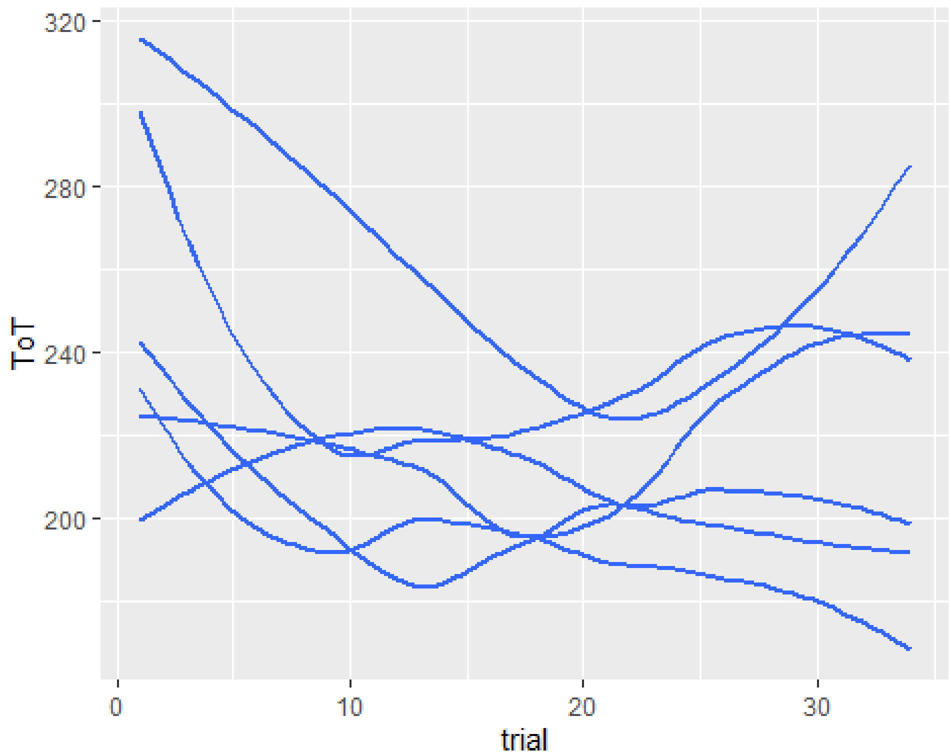
| Part | Training | Block | block | Blk_type | trial | crashes | speed | steer | ToT |
|---|---|---|---|---|---|---|---|---|---|
| 9 | Accuracy | 2 | 2 | Accuracy | 9 | 0 | 9.557164 | 0.015569 | 210 |
| 19 | Accuracy | 4 | 4 | Accuracy | 31 | 0 | 12.129550 | 0.013979 | 166 |
| 11 | Accuracy | 1 | 1 | Accuracy | 2 | 0 | 7.542732 | 0.011001 | 267 |
| 19 | Accuracy | 4 | 4 | Accuracy | 33 | 0 | 9.115585 | 0.023584 | 220 |
| 11 | Accuracy | 2 | 2 | Accuracy | 16 | 0 | 9.132105 | 0.005223 | 218 |
| 19 | Accuracy | 2 | 2 | Accuracy | 16 | 1 | 11.502620 | -0.028050 | 177 |
| 13 | Accuracy | 3 | 3 | Accuracy | 20 | 0 | 9.308915 | 0.002763 | 215 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 9 | Accuracy | 4 | 4 | Accuracy | 27 | 0 | 9.225628 | -0.022370 | 217 |
| 9 | Accuracy | 3 | 3 | Accuracy | 23 | 0 | 10.114680 | -0.021840 | 214 |
| 19 | Accuracy | 3 | 3 | Accuracy | 23 | 0 | 9.939046 | -0.004640 | 199 |

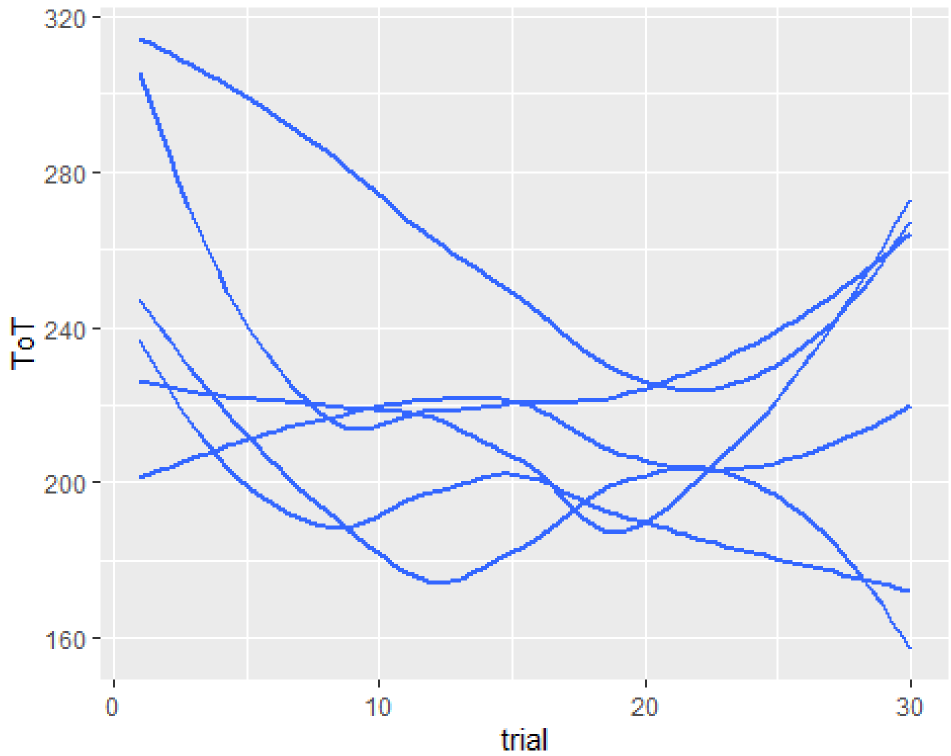
```
D_OST_ACC %>%
    ggplot(aes(x = trial, y = ToT)) +
    geom_smooth(se = F, scale = "free_y") +
  geom_point() +
    facet_wrap(~Part)

## Warning: Ignoring unknown parameters: scale

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
D_OST_ACC %>%
  filter(ToT> 140) %>%
    ggplot(aes(x = trial, y = ToT, group = Part)) +
    geom_smooth(se = F)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
D_OST_ACC <-
  D_OST %>%
  filter(Training == "Accuracy", trial <=30)

D_OST_ACC %>%
  ggplot(aes(x = trial, y =, ToT, group = Part)) +
  geom_smooth(se = F)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
D_OST_ACC %>%
  sample_n(10)
```

| Part | Training | Block | block | Blk_type | trial | crashes | speed | steer | ToT |
|---|---|---|---|---|---|---|---|---|---|
| 19 | Accuracy | 1 | 1 | Accuracy | 5 | 0 | 8.434799 | 0.014712 | 238 |
| 3 | Accuracy | 2 | 2 | Accuracy | 13 | 0 | 8.874301 | -0.000950 | 228 |
| 5 | Accuracy | 3 | 3 | Accuracy | 17 | 0 | 8.538765 | 0.000217 | 239 |
| 13 | Accuracy | 1 | 1 | Accuracy | 4 | 0 | 9.361368 | 0.007796 | 275 |
| 13 | Accuracy | 3 | 3 | Accuracy | 19 | 0 | 11.473720 | -0.000220 | 174 |
| 3 | Accuracy | 1 | 1 | Accuracy | 7 | 0 | 9.272910 | -0.003050 | 217 |
| 9 | Accuracy | 4 | 4 | Accuracy | 30 | 0 | 10.155710 | -0.011860 | 196 |
| 3 | Accuracy | 4 | 4 | Accuracy | 28 | 0 | 7.796170 | -0.006340 | 259 |
| 13 | Accuracy | 3 | 3 | Accuracy | 18 | 0 | 8.809698 | 0.025327 | 230 |
| 3 | Accuracy | 3 | 3 | Accuracy | 24 | 0 | 12.373340 | -0.009010 | 165 |

```
F_lacy_prior <- c(set_prior("normal(5.25, 0.576)", nlpar = "ampl"),
set_prior("normal(-2.76, 2.07)", nlpar = "ctch"),
set_prior("normal(1.84, 0.576)", nlpar = "asym"))

F_lacy <- formula(ToT ~ exp(asym) + exp(ampl) * inv_logit((1-ctch))^trial)

F_lacy_ef_ToT <- list(formula(ampl ~ 1|Part),
              formula(ctch ~ 1|Part),
              formula(asym ~ 1|Part))

F_lacy_prior_1 <- c(set_prior("normal(5.25, 0.875)", nlpar = "ampl"),
set_prior("normal(-2.76, 2.07)", nlpar = "ctch"),
set_prior("normal(1.84, 0.576)", nlpar = "asym"))
```

```
F_lacy_prior_3 <- c(set_prior("normal(5.25, 1.05)", nlpar = "ampl"),
set_prior("normal(-2.76, 2.07)", nlpar = "ctch"),
set_prior("normal(1.84, 0.875)", nlpar = "asym"))

F_lacy_prior_4 <- c(set_prior("normal(5.25, 1.43)", nlpar = "ampl"),
set_prior("normal(-2.76, 2.07)", nlpar = "ctch"),
set_prior("normal(1.84, 1.05)", nlpar = "asym"))

F_lacy_prior_5 <- c(set_prior("normal(5.25, 1.76)", nlpar = "ampl"),
set_prior("normal(-2.76, 2.07)", nlpar = "ctch"),
set_prior("normal(1.84, 1.43)", nlpar = "asym"))

M_OnlineSim_ToT_3 <-
D_OST_ACC %>%
brm(bf(F_lacy,
flist = F_lacy_ef_ToT,
nl = T),
prior = F_lacy_prior_5,
family = Gamma(link = identity), iter = 4000,
data = .)

## Compiling Stan program...

## Start sampling

## Warning: There were 291 divergent transitions after warmup. See
## http://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup
## to find out why this is a problem and how to eliminate them.

## Warning: There were 781 transitions after warmup that exceeded the maxim
um treedepth. Increase max_treedepth above 10. See
## http://mc-stan.org/misc/warnings.html#maximum-treedepth-exceeded

## Warning: Examine the pairs() plot to diagnose sampling problems

## Warning: The largest R-hat is 1.21, indicating chains have not mixed.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#r-hat

## Warning: Bulk Effective Samples Size (ESS) is too low, indicating poster
ior means and medians may be unreliable.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#bulk-ess

## Warning: Tail Effective Samples Size (ESS) is too low, indicating poster
ior variances and tail quantiles may be unreliable.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#tail-ess

coef(M_OnlineSim_ToT_3, mean.func = exp)

## Warning: `funs()` was deprecated in dplyr 0.8.0.
## Please use a list of either functions or lambdas:
##
##   # Simple named list:
##   list(mean = mean, median = median)
##
```

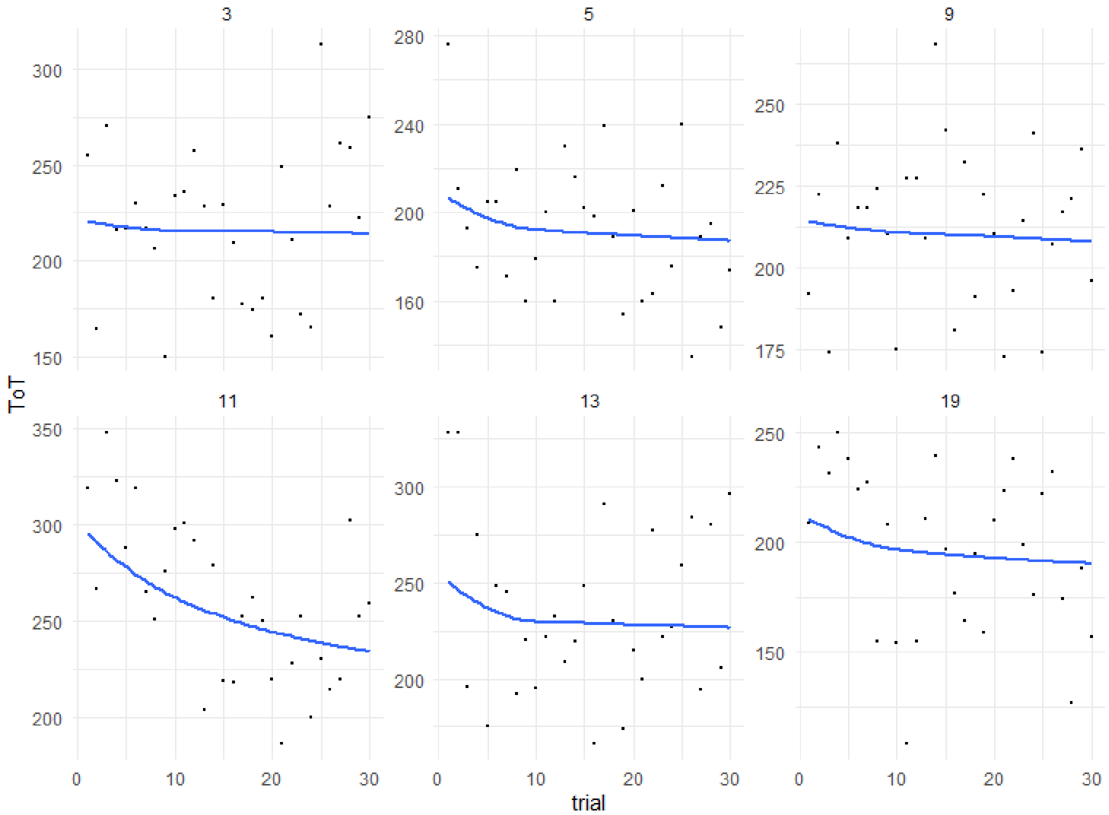
```
##   # Auto named with `tibble::lst()`:
##   tibble::lst(mean, median)
##
##   # Using lambdas
##   list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
```

*Coefficient estimates with 95% credibility limits*

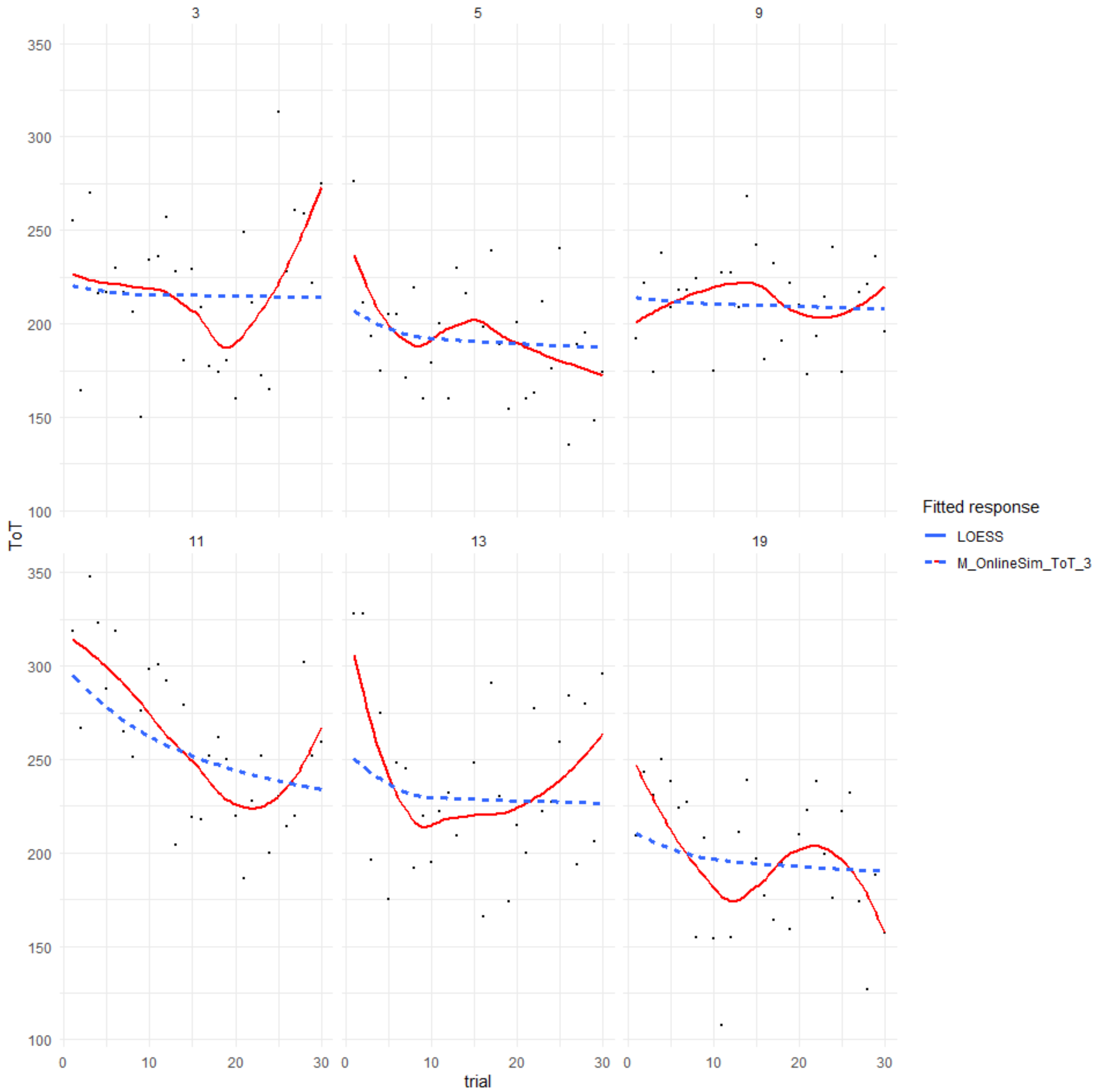
| parameter | type | nonlin | re_factor | re_entity | center | lower | upper |
|---|---|---|---|---|---:|---:|---:|
| b_ampl_Intercept | fixef | ampl | NA | NA | 95.0929248 | 7.4503574 | 6.594453e+02 |
| b_ctch_Intercept | fixef | ctch | NA | NA | 0.1194718 | 0.0017256 | 5.350162e+00 |
| b_asym_Intercept | fixef | asym | NA | NA | 201.2938220 | 0.8221717 | 2.308816e+02 |
| r_Part__ampl[3,Intercept] | ranef | ampl | Part | 3 | 0.8873212 | 0.0000000 | 2.417145e+01 |
| r_Part__ampl[5,Intercept] | ranef | ampl | Part | 5 | 0.9580594 | 0.0000000 | 7.357631e+02 |
| r_Part__ampl[9,Intercept] | ranef | ampl | Part | 9 | 0.8231365 | 0.0000000 | 1.047952e+01 |
| r_Part__ampl[11,Intercept] | ranef | ampl | Part | 11 | 1.1791998 | 0.0004975 | 1.755630e+01 |
| r_Part__ampl[13,Intercept] | ranef | ampl | Part | 13 | 1.0534266 | 0.0000001 | 6.755397e+01 |
| r_Part__ampl[19,Intercept] | ranef | ampl | Part | 19 | 0.9028609 | 0.0000000 | 1.746443e+01 |
| r_Part__ctch[3,Intercept] | ranef | ctch | Part | 3 | 1.0440996 | 0.0000000 | 2.814140e+21 |
| r_Part__ctch[5,Intercept] | ranef | ctch | Part | 5 | 2.6265113 | 0.0000000 | 7.753867e+19 |
| r_Part__ctch[9,Intercept] | ranef | ctch | Part | 9 | 1.0557070 | 0.0000000 | 2.288086e+18 |
| r_Part__ctch[11,Intercept] | ranef | ctch | Part | 11 | 0.9665635 | 0.0000000 | 1.457287e+11 |
| r_Part__ctch[13,Intercept] | ranef | ctch | Part | 13 | 1.2290528 | 0.0000000 | 2.379252e+16 |
| r_Part__ctch[19,Intercept] | ranef | ctch | Part | 19 | 1.5076059 | 0.0000000 | 3.022580e+18 |
| r_Part__asym[3,Intercept] | ranef | asym | Part | 3 | 1.0206984 | 0.0000085 | 3.292763e+00 |
| r_Part__asym[5,Intercept] | ranef | asym | Part | 5 | 0.9164346 | 0.0000083 | 1.993443e+00 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| r_Part__asym[9,Intercept] | ranef | asym | Part | 9 | 0.9971745 | 0.0000476 | 2.984567e+00 |
| r_Part__asym[11,Intercept] | ranef | asym | Part | 11 | 1.0881639 | 0.0111607 | 3.546197e+01 |
| r_Part__asym[13,Intercept] | ranef | asym | Part | 13 | 1.0843770 | 0.0000385 | 7.324366e+00 |
| r_Part__asym[19,Intercept] | ranef | asym | Part | 19 | 0.9257099 | 0.0000027 | 2.155788e+00 |

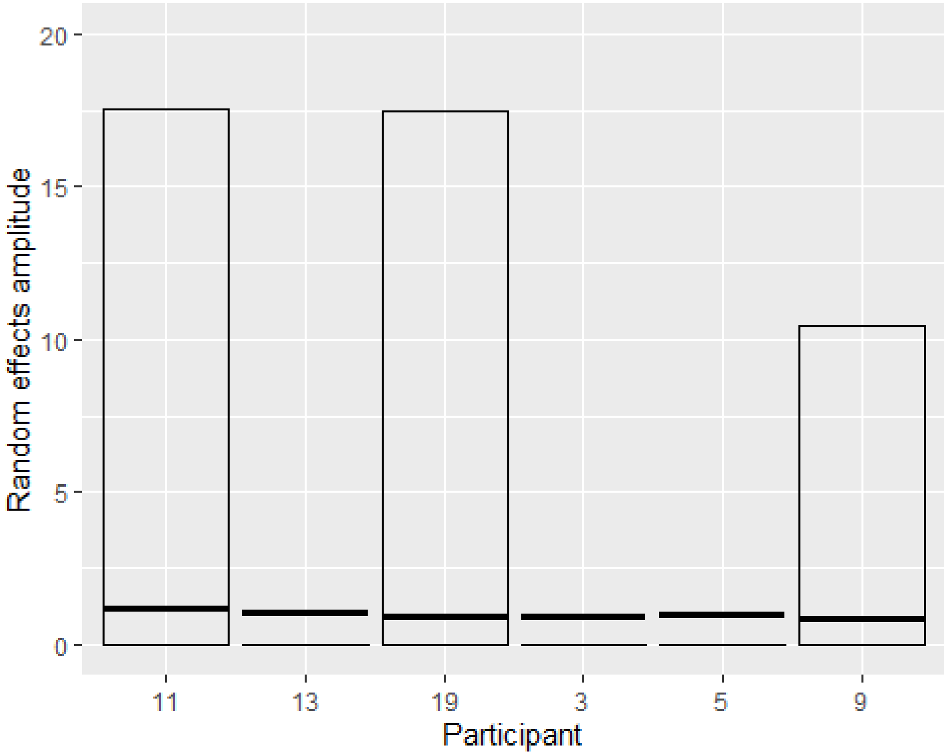
```
P_M_OnlineSim_ToT_3 <- posterior(M_OnlineSim_ToT_3)
PP_M_OnlineSim_ToT_3 <- post_pred(M_OnlineSim_ToT_3)

T_pred_M_OnlineSim_ToT_3 <- PP_M_OnlineSim_ToT_3 %>%
  group_by(Obs) %>%
  summarize(center = median(value))

D_OST_ACC$M_OnlineSim_ToT_3 <- T_pred_M_OnlineSim_ToT_3$center
D_OST_ACC$M_OnlineSim_ToT_3_resid <- D_OST_ACC$ToT - D_OST_ACC$M_OnlineSim_
ToT_3

D_M_OnlineSim_ToT_3 <-
  as_tibble(M_OnlineSim_ToT_3$data) %>%
  mutate(M_OnlineSim_ToT_3 = T_pred_M_OnlineSim_ToT_3$center)

D_OST_ACC %>%
  ggplot(aes(x = trial, y = ToT)) +
  facet_wrap(~ Part, scales = "free_y") +
  geom_point(size = .2) +
  geom_smooth(aes(y = M_OnlineSim_ToT_3), se = F) +
  theme_minimal()

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
D_OST_ACC %>%
  ggplot(aes(x = trial, y = ToT)) +
  geom_point(size = 0.5) +
  geom_smooth(aes(linetype = "LOESS"), se = F, color = "red", size = 0.75)
+
  geom_smooth(aes(y = M_OnlineSim_ToT_3, linetype = "M_OnlineSim_ToT_3"), s
e = F, size = 1.2) +
  labs(linetype = "Fitted response", M_OnlineSim_ToT_3 = "Predicted", LOESS
= "Observed") +
  facet_wrap(~Part) +
  theme_minimal()

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
coef(M_OnlineSim_ToT_3, mean.func = exp) %>%
  filter(type == "ranef", nonlin == "ampl") %>%
  ggplot(aes(x = re_entity, ymin = lower, y = center, ymax = upper)) +
  geom_crossbar() +
  labs(y = "Random effects amplitude", x = "Participant" ) +
  ylim(0,20)
```


20
15
10
5
0
Random effects amplitude
11
13
19
3
5
9
Participant

**Appendix C: Manual experiment**

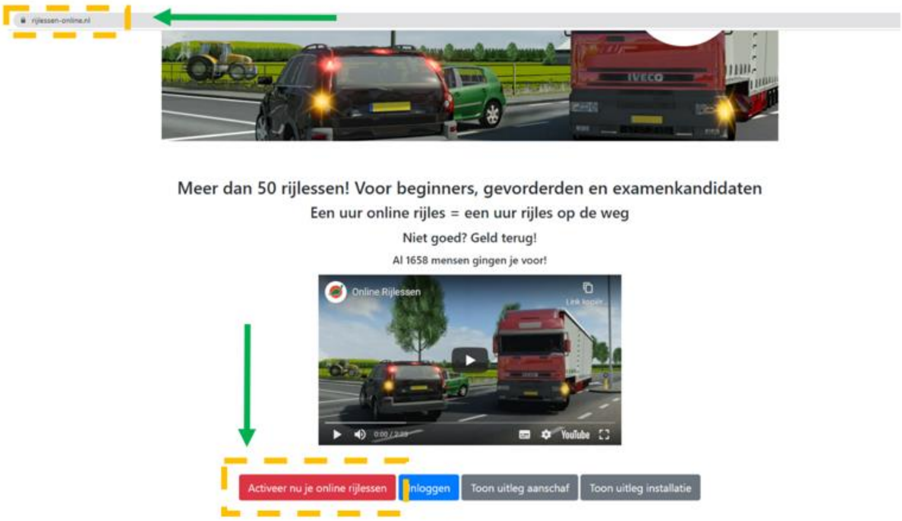
# Manual

Thank you for being willing to participate in this study! This manual will show you step by step what needs to be done in order to successfully complete the driving experience. Please read it thoroughly. You will see the words "THIS IS THE END OF THE MANUAL" when you have reached the end. After that, you can start with driving, so not before. The following steps will be performed.

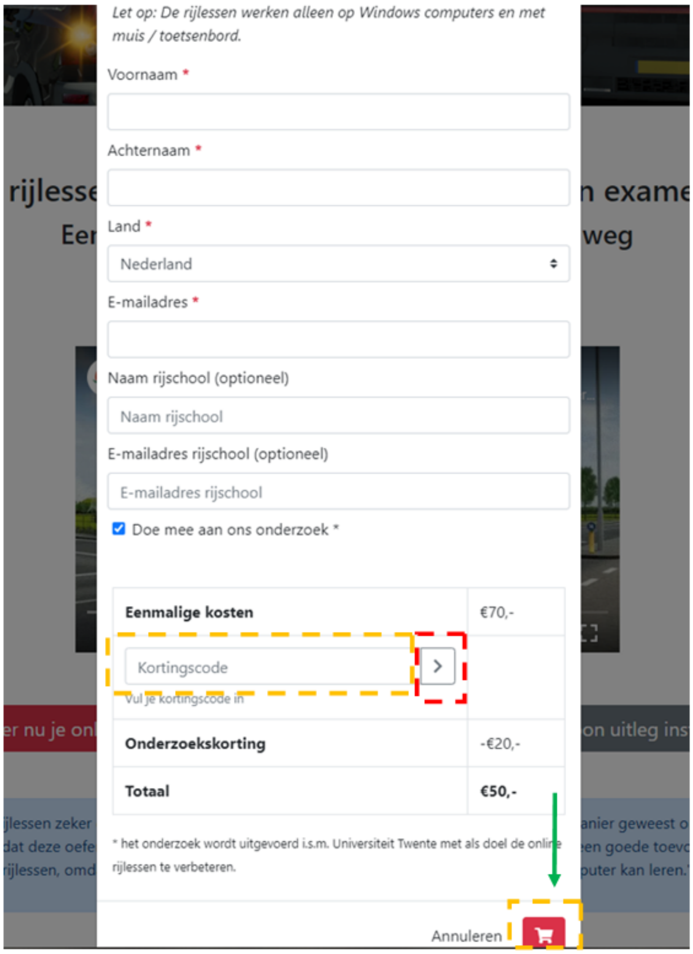
1. Account creation and software download
2. Ethical consent/ Pre-questionnaire
3. Driving experience and workload assessment

## Section 1: Account creation and software download

In order to create an account please go to the link [https://rijlessen-online.nl/](https://rijlessen-online.nl/) , scroll down a bit and select the red button "Activeer nu je online rijlessen" or "activate your online driving lessons" in case your browser is in English.

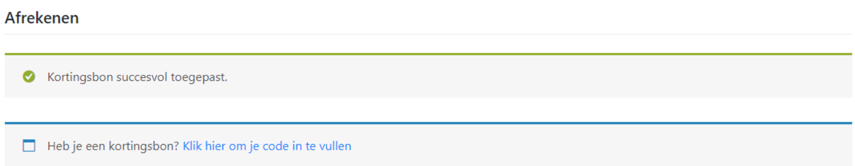


A pop-up will appear which looks like this:

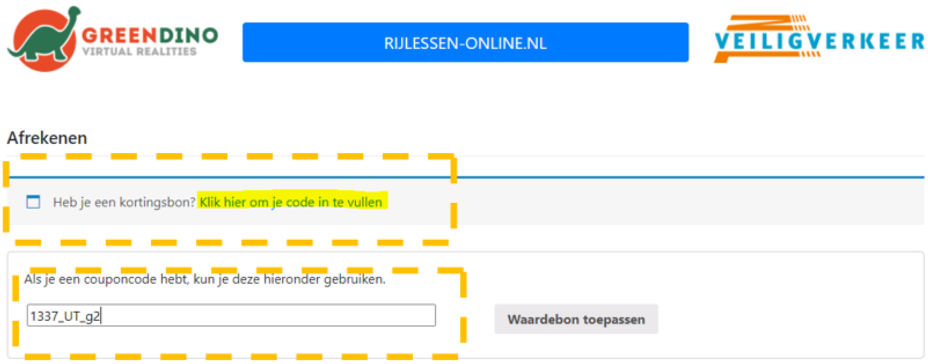



Please fill in your personal information and <u>make sure that the box for “Doe mee aan ons onderzoek” is ticked.</u> The information about the driving school can be disregarded. Fill the discount code or “kortingscode” **1337_UT_g2** and <u>click at the arrow next to it</u> to activate it. This removes the to be paid fee and enables you to drive for free. To continue, click on the red shopping car.

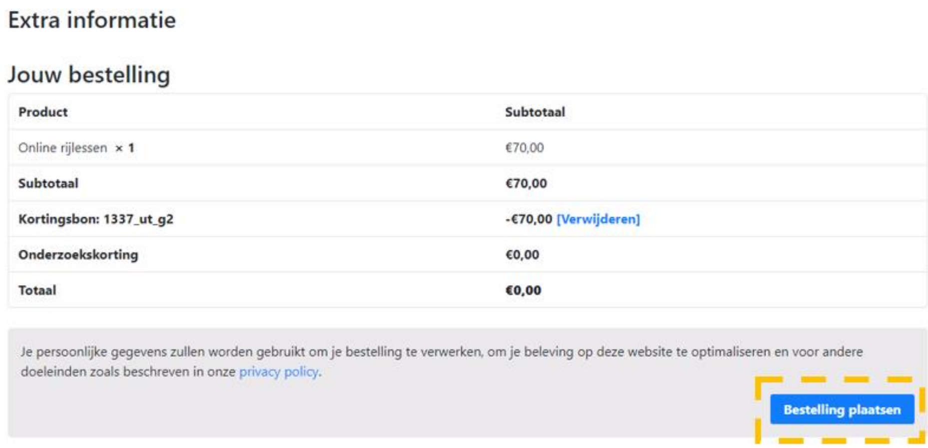
If the discount code worked, you will see this at the top of the screen




If the discount code did not work, you will see this. Fill out the discount code again and click on “waardebon toepassen”.




Now you can scroll down and click on "bestelling plaatsen"




Now you have created an account. More information will follow in your mailbox. A first email will tell you your order has been processed. The second email provides you with your personal log in information. Receiving the second email can take a few minutes (normally +- 15 minutes, but it might take longer). The second email will look like this:

Er is voor jou een nieuw account gemaakt op 'Online Rijlessen' en je hebt daarvoor een nieuw, tijdelijk wachtwoord gekregen.

De informatie die je nodig hebt om op de site in te loggen is:
Gebruikersnaam: mtest
Wachtwoord: m5+deyl!eU
(je zult je wachtwoord moeten wijzigen wanneer je je de eerste keer aanmeldt op de site)

Om te beginnen met je 'Online Rijlessen', klik op
https://online-rijlessen.virtual-reality-lms.com/login/

In de meeste e-mailprogramma's zal de bovenstaande regel een blauw, onderlijnde link zijn, waarop je gewoon kunt klikken. Als dat niet werkt, knip en plak dan dit adres in de adresbalk bovenaan in je internetbrowser.

Als je hulp nodig hebt, neem dan op werkdagen contact op met onze helpdesk door een e-mail te sturen naar online-rijlessen@greendino.nl

You can see your username (gebruikersnaam) and password (wachtwoord) here which you can use to log in and start the driving experience at https://online-rijlessen.virtual-reality-lms.com/login/

GREENDINO VIRTUAL REALITIES ONLINE RIJLESSEN VEILIGVERKEER


**Heb je nog geen account? Ga dan naar https://rijlessen-online.nl/ om je Online Rijlessen aan te schaffen.**

Je sessie bleef te lang zonder activiteit. Je moet opnieuw inloggen.

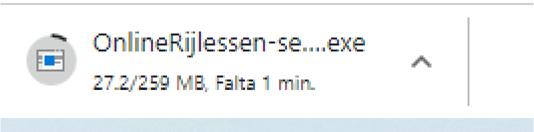
Gebruikersnaam
Wachtwoord
Login

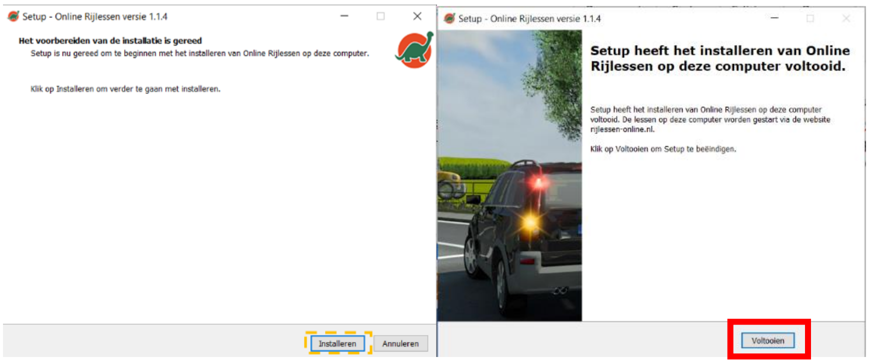
Gebruikersnaam onthouden

Ben je je gebruikersnaam of wachtwoord vergeten?

After you have logged in, the programme asks you to adjust your password. Thereafter, you will be led to the simulator. Expand the explanation or uitleg section by clicking on the arrow next to it. Here you can find the button to download the software. Follow the instructions given by your computer to complete the downloading process.


You will then see the program downloading. After the download, click on the file and allow your computer to install the programme. Don't forget to click on finish/voltooien to complete the download.



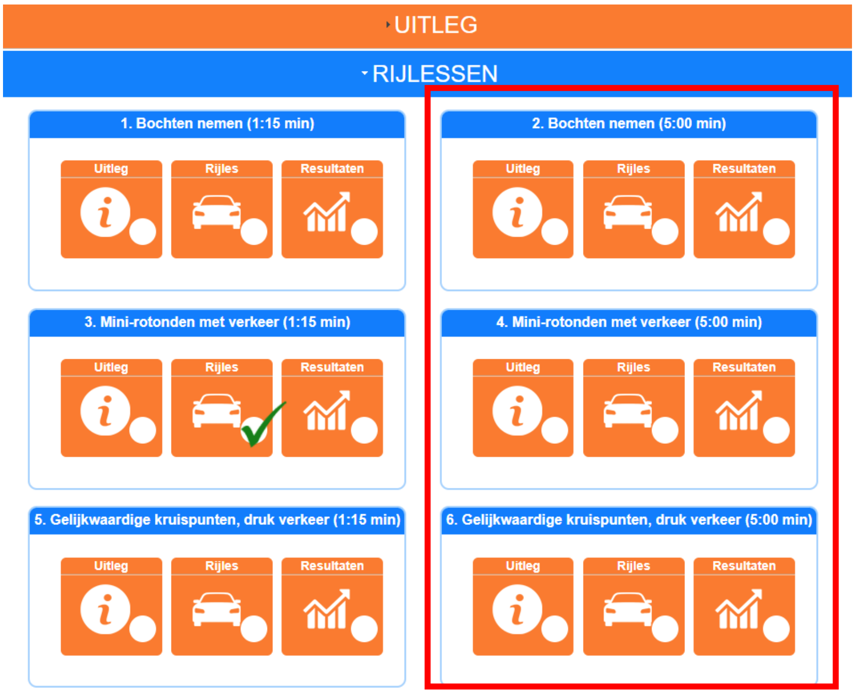

Now you have created an account and downloaded the programme. Please continue to section 2

## Section 2: Ethical consent and prequestionnaire

Before you start driving, we would like to ask you to go to this link. (https://utwentebs.eu.qualtrics.com/jfe/form/SV_25eUbgqoK9i1Egu ). Please give your ethical consent and fill out the questionnaire. After this, come back to this manual and continue with section 3.

## Section 3: Driving experience and workload assessment

Now the driving experience can start! Make sure that you have read section 3 entirely before you start driving. Go back to the webpage of the online simulator. Click on the arrow next to Rijlessen or driving lessons to expand the section. 6 lessons will appear. Please do not click yet on any of these but read further in this manual.




You will be performing the **5 minute lessons**, which are displayed at the right side of the screen (number 2, 4 and 6, see the previous screenshot). Be aware that there is a fixed order in which you should complete the lessons: (1) Taking bends or bochten nemen , (2 ) mini-roundabouts with traffic or mini-rotonden met verkeer and (3) unmarked junctions with dense traffic or gelijkwaardige kruispunten, druk verkeer (1-2-3). **You follow this order every time till you have completed each lesson 20 times.** To start a lesson, you click on the "driving lesson" button displaying a car.


After every lesson, you have to indicate your experienced workload during the lesson. To do this, go to this questionnaire (https://utwentebs.eu.qualtrics.com/jfe/form/SV_cuqQFicVH4jaARg ). The order of the questionnaire corresponds with the order of the lessons. So, if you are lost on the lesson you are supposed to do and how many times you already performed it, you can look back at this questionnaire. It is important that you leave the questionnaire open until the moment you have completed all trials, otherwise your data of the previous trials will not be saved. After you have answered the last question in this questionnaire, please forward the results by clicking on the yellow arrow in the right bottom of the page.

Once the workload questionnaire is opened, it is time to start driving! Here are the basic mechanics on how to operate the car. Keep in mind that you will need an external mouse, so a laptop mouse is not sufficient.

- Start the car by pressing spacebar (please do this directly when the car is presented on the screen so that you can start driving immediately)

- Moving the mouse forward results in acceleration

- Moving the mouse down results in deceleration,

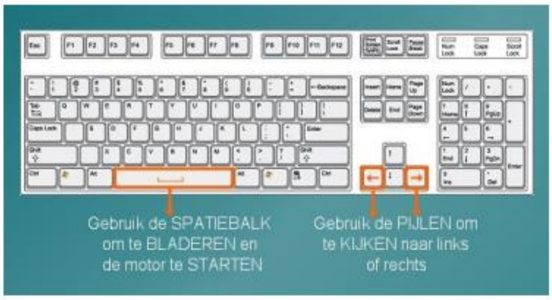
- Left and right controls the steering wheel direction.

- Clicking the mouse buttons controls the indicators.

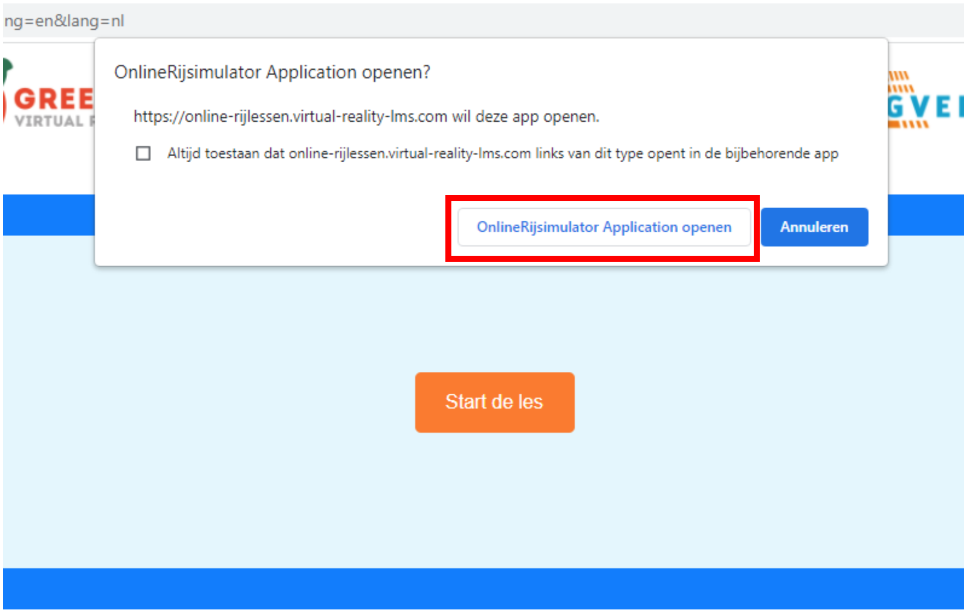
- The left and right arrow, or the z and c keys, are used to open a viewport which displayed the mirrors and a view to the left and right of the car.







Go back to the webpage of the online driving simulator and click on the first lesson and then click on start de les or start the lesson. A pop-up will appear which asks you whether you want to open the OnlineRijsimulator Application.




Please click on OnlineRijsimulator Application openen to be able to start the lesson. A screen will pop-up which will disappear after some time.

After this screen, you will see a car and will be able to drive! Please start directly by pressing the spacebar. You can perform the experiment at different moments if you like. <u>But leave the workload questionnaire opened</u> if you do, because otherwise your data will be gone.

THIS IS THE END OF THE MANUAL

**Appendix D: Pre-questionnaire experiment**

# Online driving simulator study pre-questionnaire

---

Start of Block: Consent

Q22 The following experiment is about the effectiveness of an online driving simulator. You have received a manual with all relevant information. This experiment contains no risks for you, however, there is a minor possibility that you might start to feel uncomfortable while driving. If so, please take a break and try to continue later. Be aware that if you do not give your consent, you will not be able to participate in the study.

Hereby, I give my consent to participate in the Online Driving Simulator Study that investigates individual learning behaviour regarding driving and is run by the department of Psychology at the University of Twente in Enschede. I declare that I have been informed about the nature, method, purpose and risks and burden of the research in a manner that is clear to me. I know that the data and results of the survey will only be disclosed to third parties anonymously and confidentially. My questions are answered satisfactorily. I voluntarily agree to participate in this research. I reserve the right to terminate my participation in this study at any time without giving reasons. If I would like any further information about the study, now or in the future, I can contact the researcher.

Contact information researcher: m.voskes@student.utwente.nl

I give consent

○ Yes (1)

○ No (2)

*Skip To: End of Survey If The following experiment is about the effectiveness of an online driving simulator. You have rece... = No*

End of Block: Consent

---

Start of Block: Introduction

Q18
Welcome to the online driving simulator experience. We would like to thank you for participating and

we wish you a fun learning experience. At this point you have created an account and downloaded the app as explained in the manual. If you did not, please go back to the manual and perform the steps required. Before you actually perform the driving tasks, we would like you to answer some questions regarding your driving experience and gaming experience. This questionnaire will take no more than 5 minutes.

Please mention the email address you used to create an account for the Green Dino driving simulator.

________________________________________________________________

**End of Block: Introduction**

---

**Start of Block: Driving experience**

Q20 Now we are going to assess your driving experience. Please try to answer the questions as accurate as possible. Are you ready to continue?

○ Yes (1)

---

Page Break

Q5 What is your current driving license status?

◯ Not licenced (1)

◯ Currently licensing (taking driving training on road) (2)

◯ Licenced in the Netherlands (3)

◯ Licenced in another country, namely (4) ____________________

*Skip To: Q8 If What is your current driving license status? = Currently licensing (taking driving training on road)*
*Skip To: End of Block If What is your current driving license status? = Not licenced*
*Skip To: Q6 If What is your current driving license status? = Licenced in the Netherlands*
*Skip To: Q6 If What is your current driving license status? = Licenced in another country, namely*

Q8 How many hours of driving training did you get so far?

| | 0 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Training hours () | | | | | | | | | | | |

*Skip To: End of Block If How many hours of driving training did you get so far? [ Training hours ] >*

Q6 How many years have you been driving?

Longer than 12 years

| | 0 | 1 | 2 | 4 | 5 | 6 | 7 | 8 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of years () | | | | | | | | | | | |
| Number of months () | | | | | | | | | | | |

Q7 How frequent do you drive?

- ○ Almost every day (1)
- ○ A few days a week (2)
- ○ A few days a month (3)
- ○ A few times a year (4)
- ○ Never (5)

End of Block: Driving experience

Start of Block: Gaming experience

Q19 Now we are going to assess your gaming experience. Please answer the following questions as accurate as possible. Are you ready to continue?

- ○ Yes (1)

Page Break

Q14 How many hours do you normally spend on playing computer games during a week? Select 0 if you do not play computer games (Note: no cellphone, tablet or console games)

| | 0 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of hours () | | | | | | | | | | | |

*Skip To: End of Block If How many hours do you normally spend on playing computer games during a week? Select 0 if you do... [ Number of hours ] =*

Q17 How often do you play?

- Daily (1)
- Weekly (2)
- Monthly (3)
- Semesterly (4)
- Yearly (5)
- Never (6)

**End of Block: Gaming experience**

**Start of Block: Block 4**

Q18 Now we are going to assess your biking experience. Please try to answer the questions as accurate as possible. Are you ready to continue?

- Yes (1)

Page Break

Q19 How many years have you been riding a bike?

| | More than 20 years |
| --- | --- |
| | 0 1 2 3 4 5 6 7 8 9 1011121314151617181920 |
| Number of years () | |

Q20 How frequent do your ride a bike?

- ◯ Almost every day (1)
- ◯ A few days a week (2)
- ◯ A few days a month (3)
- ◯ A few times a year (4)
- ◯ Never (5)

End of Block: Block 4

**Appendix E: Workload questionnaire experiment**

# Online driving simulator study workload questionnaire 20 trials

Start of Block: Instruction

Q8
We are interested in how you experienced your workload during the driving lessons. Please answer the question after each lesson (60 times) you performed in the simulator.

Please mention the email address you used to create an account for the Green Dino driving simulator.

________________________________________________________________

End of Block: Instruction

Start of Block: Workload

Q2 Lesson 1, trial 1: How mentally demanding was the task?

| | Very Low Very High |
|---|---|
| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 |
| Mental demand () | |

Q12 Lesson 2, trial 1: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q24 Lesson 3, trial 1: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q23 Lesson 1, trial 2: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q22 Lesson 2, trial 2: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q21 Lesson 3, trial 2: How mentally demanding was the task?

| | Very Low | Very High |
|---|---|---|
| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 | |
| Mental demand () | | |

Q20 Lesson 1, trial 3: How mentally demanding was the task?

| | Very Low | Very High |
|---|---|---|
| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 | |
| Mental demand () | | |

Q25 Lesson 2, trial 3: How mentally demanding was the task?

| | Very Low | Very High |
|---|---|---|
| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 | |
| Mental demand () | | |

Q26 Lesson 3, trial 3: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q27 Lesson 1, trial 4: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q28 Lesson 2, trial 4: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q29 Lesson 3, trial 4: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q30 Lesson 1, trial 5: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q31 Lesson 2, trial 5: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q32 Lesson 3, trial 5: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q33 Lesson 1, trial 6: How mentally demanding was the task?

| | Very Low | Very High |
|---|---|---|
| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 | |
| Mental demand () | | |

Q34 Lesson 2, trial 6: How mentally demanding was the task?

| | Very Low | Very High |
|---|---|---|
| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 | |
| Mental demand () | | |

Q35 Lesson 3, trial 6: How mentally demanding was the task?

| | Very Low | Very High |
|---|---|---|
| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 | |
| Mental demand () | | |

Q36 Lesson 1, trial 7: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q37 Lesson 2, trial 7: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q38 Lesson 3, trial 7: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q39 Lesson 1, trial 8: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q40 Lesson 2, trial 8: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

Mental demand ()

Q41 Lesson 3, trial 8: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

Mental demand ()

Q42 Lesson 1, trial 9: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

Mental demand ()

Q43 Lesson 2, trial 9: How mentally demanding was the task?

Very Low Very High

| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 |
|---|---|
| Mental demand () | |

Q44 Lesson 3, trial 9: How mentally demanding was the task?

| | Very Low | Very High |
|---|---|---|

| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 |
|---|---|
| Mental demand () | |

Q45 Lesson 1, trial 10: How mentally demanding was the task?

| | Very Low | Very High |
|---|---|---|

| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 |
|---|---|
| Mental demand () | |

Q46 Lesson 2, trial 10: How mentally demanding was the task?

| | Very Low | Very High |
|---|---|---|

| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 |
|---|---|
| Mental demand () | |

Q47 Lesson 3, trial 10: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

| | |
|---|---|
| Mental demand () | |

Q48 Lesson 1, trial 11: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

| | |
|---|---|
| Mental demand () | |

Q49 Lesson 2, trial 11: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

| | |
|---|---|
| Mental demand () | |

Q50 Lesson 3, trial 11: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

Mental demand ()

Q51 Lesson 1, trial 12: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

Mental demand ()

Q52 Lesson 2, trial 12: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

Mental demand ()

Q53 Lesson 3, trial 12: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q38 Lesson 1, trial 13: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q39 Lesson 2, trial 13: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q40 Lesson 3, trial 13: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q41 Lesson 1, trial 14: How mentally demanding was the task?

| | Very Low | Very High |
|---|---|---|
| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 | |
| Mental demand () | | |

Q42 Lesson 2, trial 14: How mentally demanding was the task?

| | Very Low | Very High |
|---|---|---|
| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 | |
| Mental demand () | | |

Q43 Lesson 3, trial 14: How mentally demanding was the task?

| | Very Low | Very High |
|---|---|---|
| | 0 1 2 3 4 5 6 7 8 9 101112131415161718192021 | |
| Mental demand () | | |

Q44 Lesson 1, trial 15: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q45 Lesson 2, trial 15: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q46 Lesson 3, trial 15: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q47 Lesson 1, trial 16: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q49 Lesson 2, trial 16: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q48 Lesson 3, trial 16: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q50 Lesson 1, trial 17: How mentally demanding was the task?

Very Low Very High

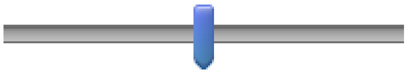
0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q51 Lesson 2, trial 17: How mentally demanding was the task?

Very Low Very High

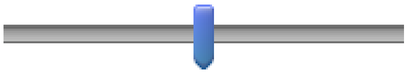
0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () |  |

Q52 Lesson 3, trial 17: How mentally demanding was the task?

Very Low Very High

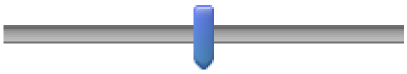
0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () |  |

Q53 Lesson 1, trial 18: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () |  |

Q54 Lesson 2, trial 18: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

Mental demand ()

Q55 Lesson 3, trial 18: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

Mental demand ()

Q56 Lesson 1, trial 19: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

Mental demand ()

Q57 Lesson 2, trial 19: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

Mental demand ()

Q58 Lesson 3, trial 19: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q59 Lesson 1, trial 20: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q60 Lesson 2, trial 20: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

Q61 Lesson 3, trial 20: How mentally demanding was the task?

Very Low Very High

0 1 2 3 4 5 6 7 8 9 101112131415161718192021

| | |
|---|---|
| Mental demand () | |

End of Block: Workload

## Appendix F: Information email experiment

Hey!

Thanks for showing interest in participating in the online driving simulator study! In addition to supporting research, you will be training your driving skills for free! The manual attached will guide you through the driving experience and explains step by step how it works and what is expected from you. This manual is very important for correct execution of the experiment, **so please read this manual properly**. If you face any difficulties during preparing the experiment or while you are performing the driving, you can always contact me via email (m.voskes@student.utwente.nl). I will try to reply as fast as possible to help you out!

The driving will take around 5 hours. Therefore, you are allowed to divide the experiment in multiple time moments. The deadline for completion will be on Sunday the 12th at 24:00. Be aware that the simulator does not work on a Mac device. So if you have a Mac, please try to find an alternative device on which you can do the experiment.

Have fun with the driving and thanks in advance for participating!

Kind regards,

Maran

**Appendix G: Rscript exploration simulator and workload data experimental study**

```
D_experiment <- read_delim("~/Master/Thesis/Experiment/experiment_data_13de
c (1).csv",
    ";", escape_double = FALSE, trim_ws = TRUE)

##
## -- Column specification ------------------------------------------------
--------
## cols(
##   LessonDate = col_character(),
##   StudentID = col_double(),
##   LessonID = col_double(),
##   OverallSafetyScore = col_double(),
##   OverallTaskScore = col_double(),
##   TaskName = col_character(),
##   TaskId = col_double(),
##   NrFailed = col_double(),
##   NrSucces = col_double(),
##   TaskScore = col_double(),
##   SafetyName = col_character(),
##   SafetyScore = col_double(),
##   DrivingExperience = col_character(),
##   GamingExperience = col_character(),
##   BikingExperience = col_character()
## )

D_experiment <- D_experiment %>%
group_by(StudentID, SafetyName, TaskId, LessonID) %>%
  arrange(LessonDate) %>%
  mutate(trial = row_number(LessonDate)) %>%
  ungroup()

write_xlsx(D_experiment, "~\\Master\\Thesis\\Experiment\\experiment_data.xl
sx")

D_exp <- read_delim("~/Master/Thesis/Experiment/experiment_data_1.csv",
    ";", escape_double = FALSE, trim_ws = TRUE)

##
## -- Column specification ------------------------------------------------
--------
## cols(
##   LessonDate = col_character(),
##   StudentID = col_double(),
##   LessonID = col_double(),
##   OverallSafetyScore = col_double(),
##   OverallTaskScore = col_double(),
##   TaskName = col_character(),
##   TaskId = col_double(),
##   NrFailed = col_double(),
##   NrSucces = col_double(),
##   TaskScore = col_double(),
##   SafetyName = col_character(),
##   SafetyScore = col_double(),
##   DrivingExperience = col_character(),
```

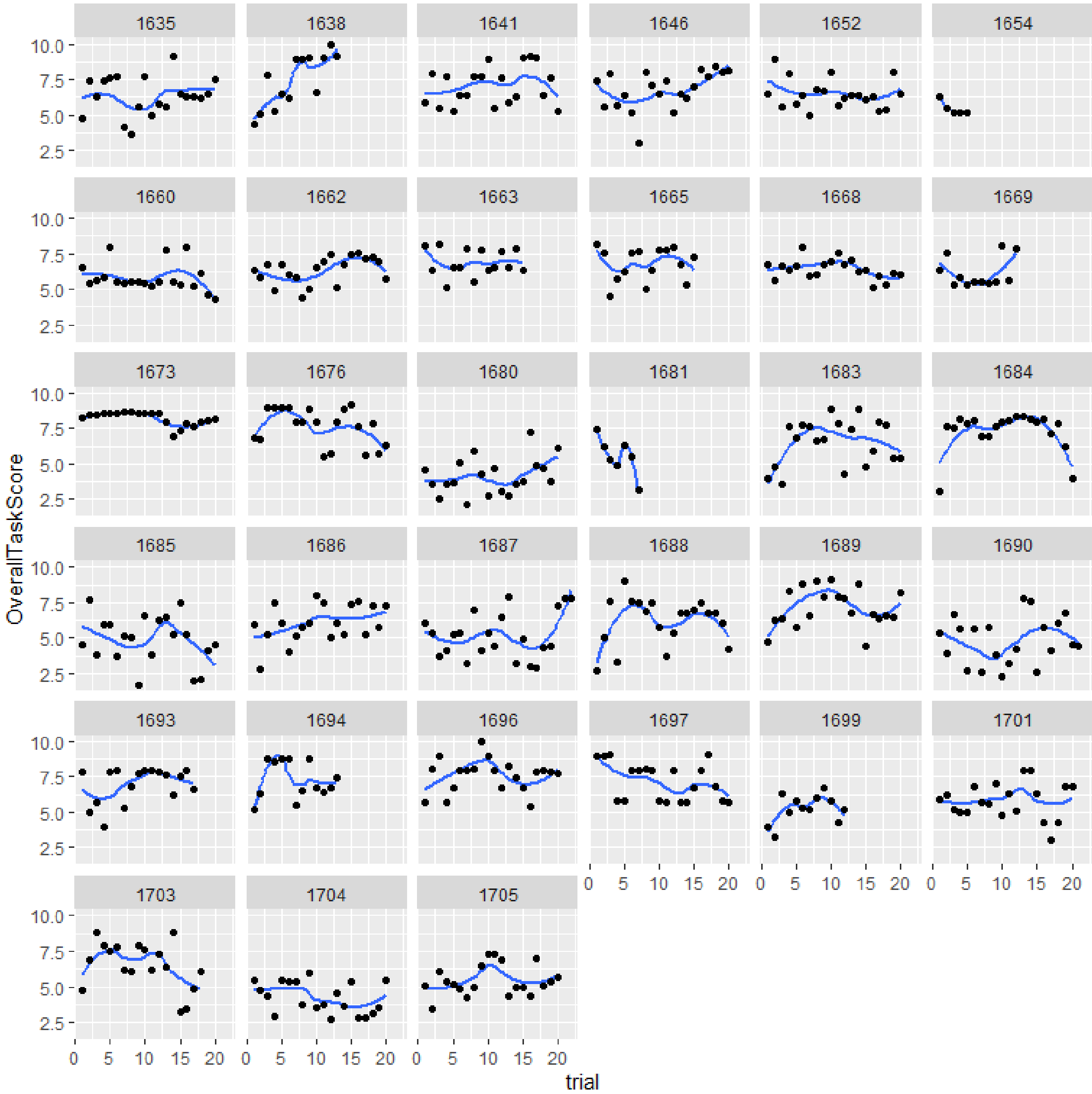
```
##   GamingExperience = col_character(),
##   BikingExperience = col_character(),
##   trial = col_double(),
##   KmDriven = col_double(),
##   YearlyGamingHours = col_double()
## )
```

```
D_exp %>%
  sample_n(10)

D_exp %>%
   summarise(mean(KmDriven),
             sd(KmDriven),
             mean(YearlyGamingHours),
             sd(YearlyGamingHours),
             max(KmDriven),
             min(KmDriven),
             max(YearlyGamingHours),
             min(YearlyGamingHours))
```

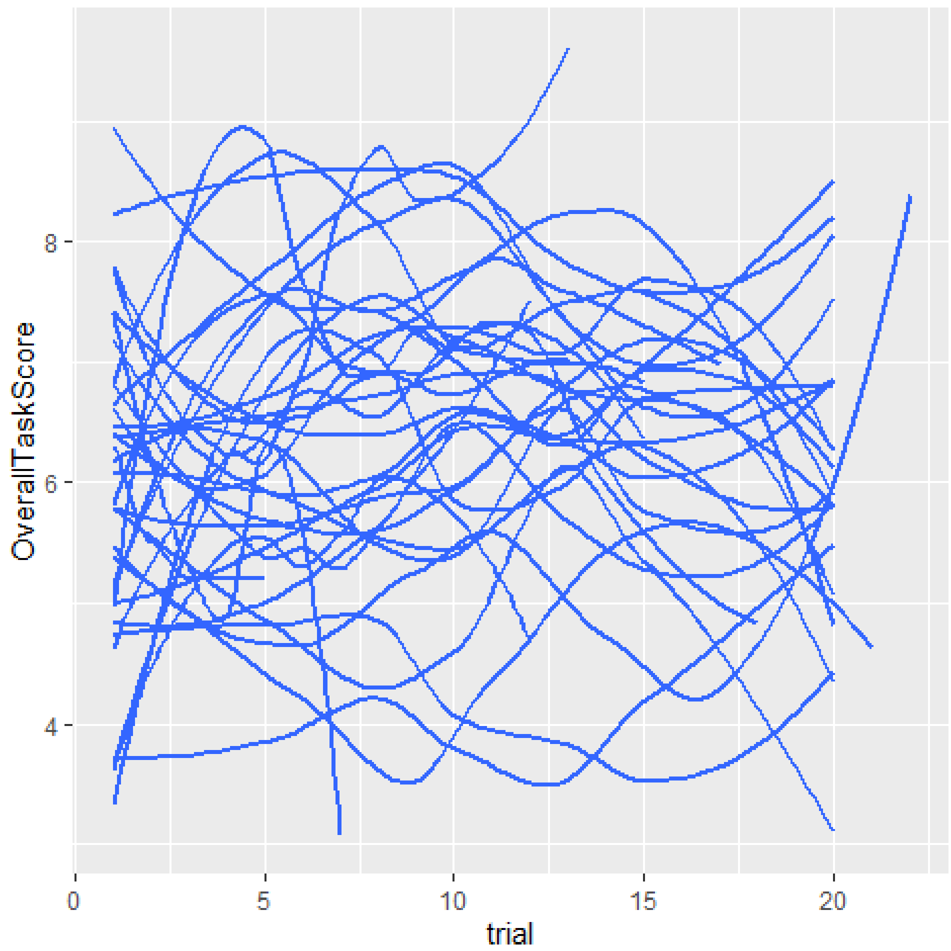
| mean(KmDriven) | sd(KmDriven) | mean(YearlyGamingHours) | sd(YearlyGamingHours) | max(KmDriven) | min(KmDriven) | max(YearlyGamingHours) | min(YearlyGamingHours) |
|---:|---:|---:|---:|---:|---:|---:|---:|
| 5219.632 | 8264.246 | 229.0064 | 450.6127 | 33072 | 0 | 1820 | 0 |

#OverallTaskScore

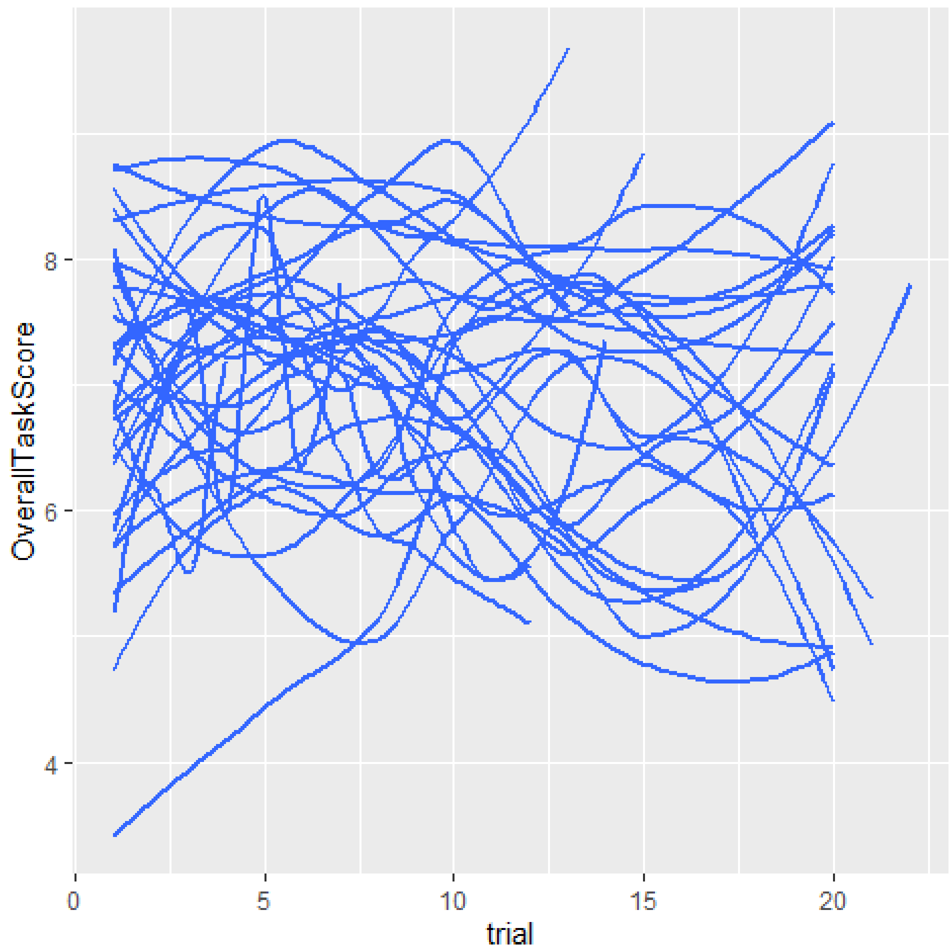
```
D_exp %>%
  filter(SafetyName == "Houden aan veilige snelheid", LessonID == "235") %>
%
  ggplot(aes(x = trial, y = OverallTaskScore)) +
  geom_smooth(se = F, method = "loess") +
  geom_point() +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```



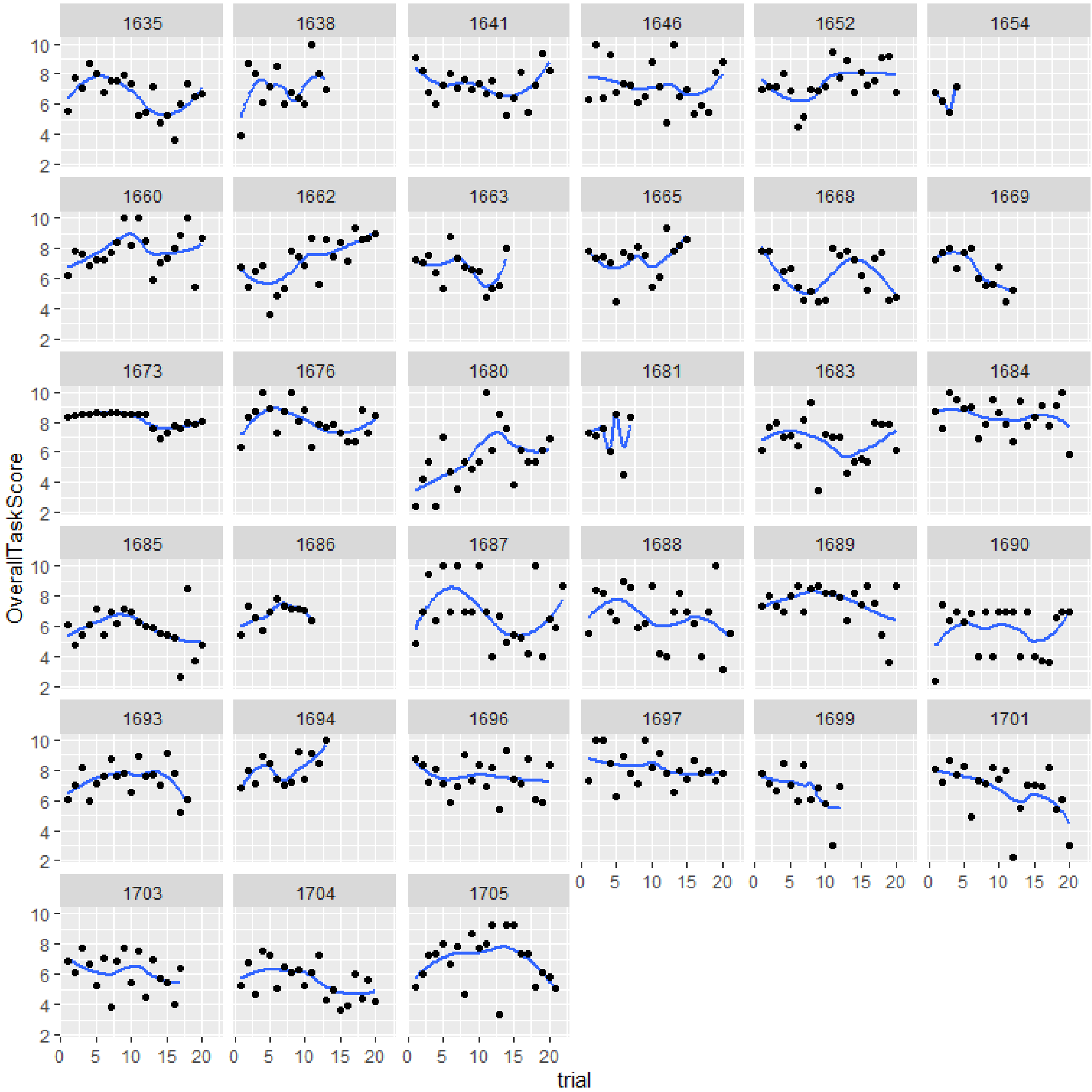

```
D_exp %>%
  filter(SafetyName == "Voertuigbeheersing", LessonID == "235") %>%
  ggplot(aes(x = trial, y = OverallTaskScore, group = StudentID)) +
  geom_smooth(se = F)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```



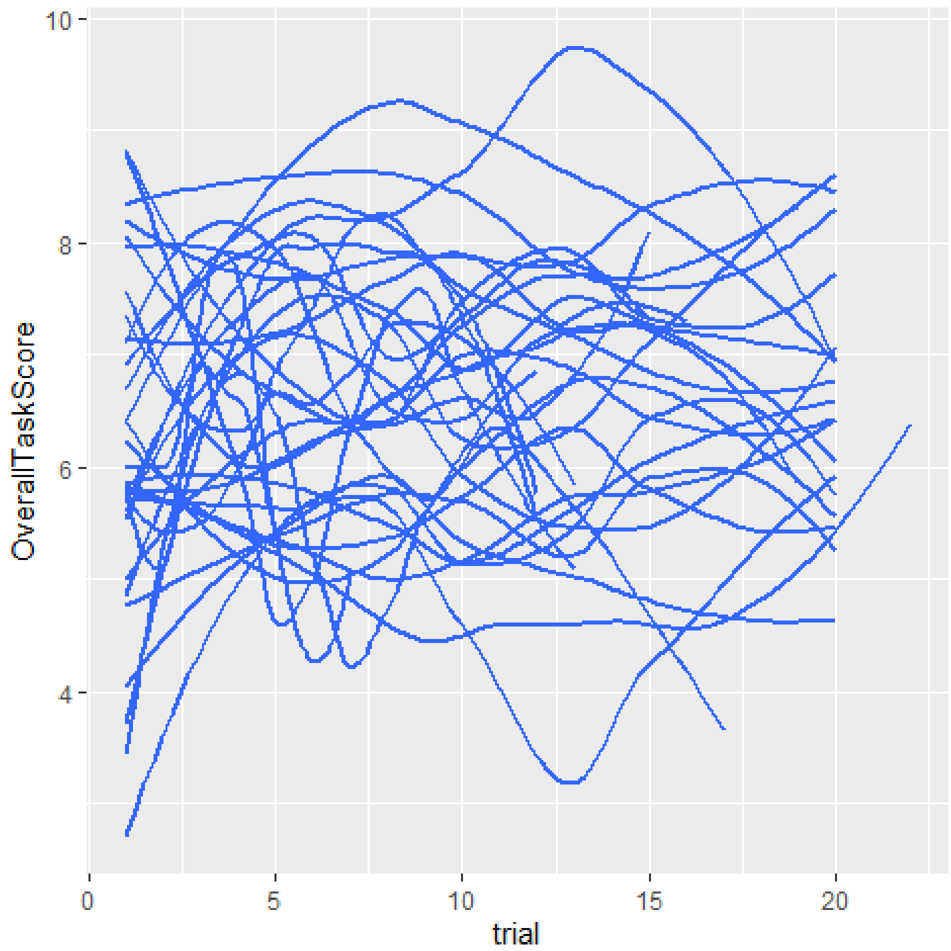

```
D_exp %>%
  filter(SafetyName == "Houden aan veilige snelheid", LessonID == "237") %>
%
  ggplot(aes(x = trial, y = OverallTaskScore, group = StudentID)) +
  geom_smooth(se = F, method = "loess")

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.985

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.015

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0602
```



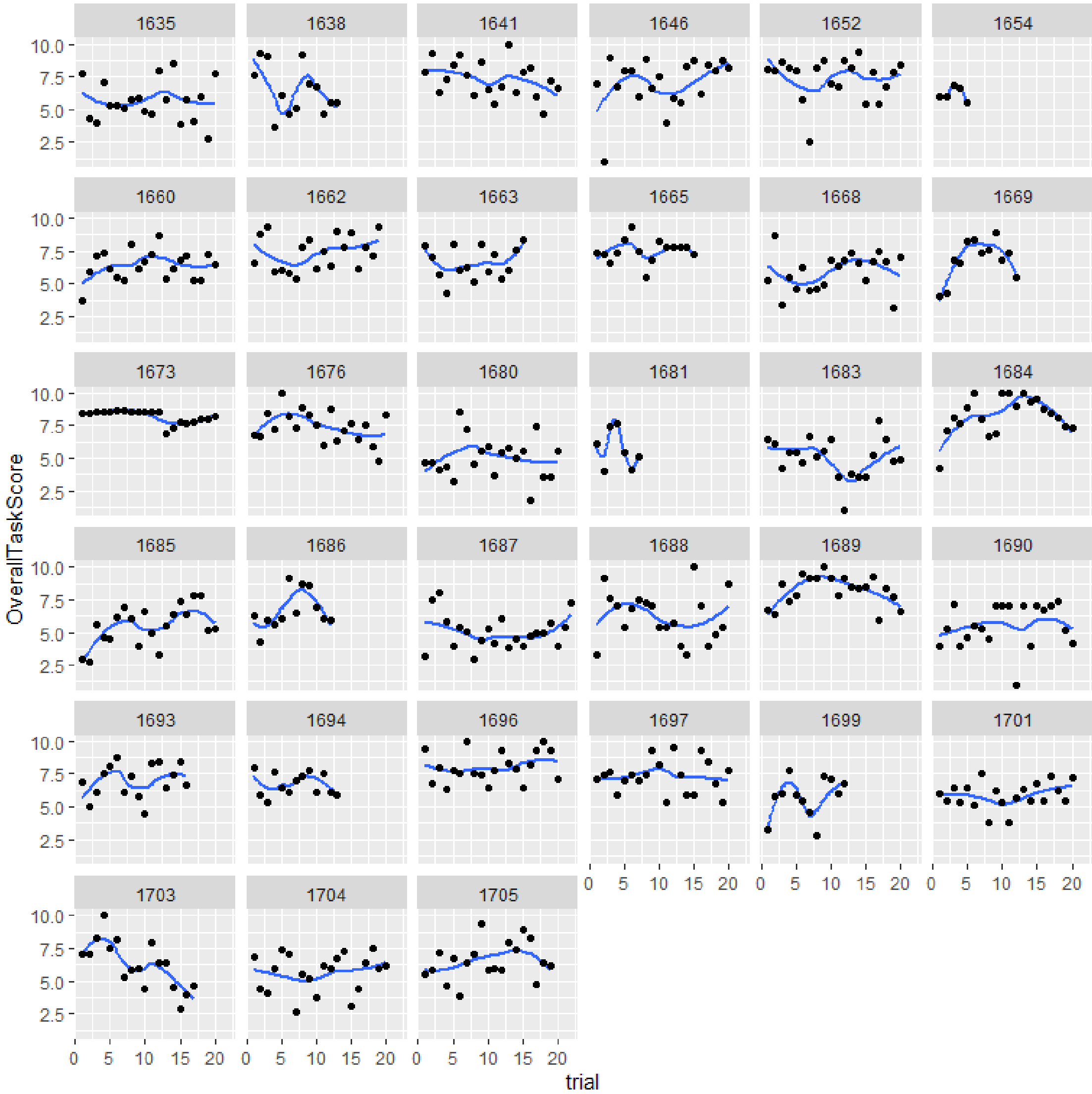

```
D_exp %>%
  filter(SafetyName == "Houden aan veilige snelheid", LessonID == "237") %>
%
  ggplot(aes(x = trial, y = OverallTaskScore)) +
  geom_smooth(se = F, method = "loess") +
  geom_point() +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.985

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.015

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0602
```




```
D_exp %>%
  filter(SafetyName == "Houden aan veilige snelheid", LessonID == "239") %>
%
  ggplot(aes(x = trial, y = OverallTaskScore, group = StudentID)) +
  geom_smooth(se = F, method = "loess")

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```



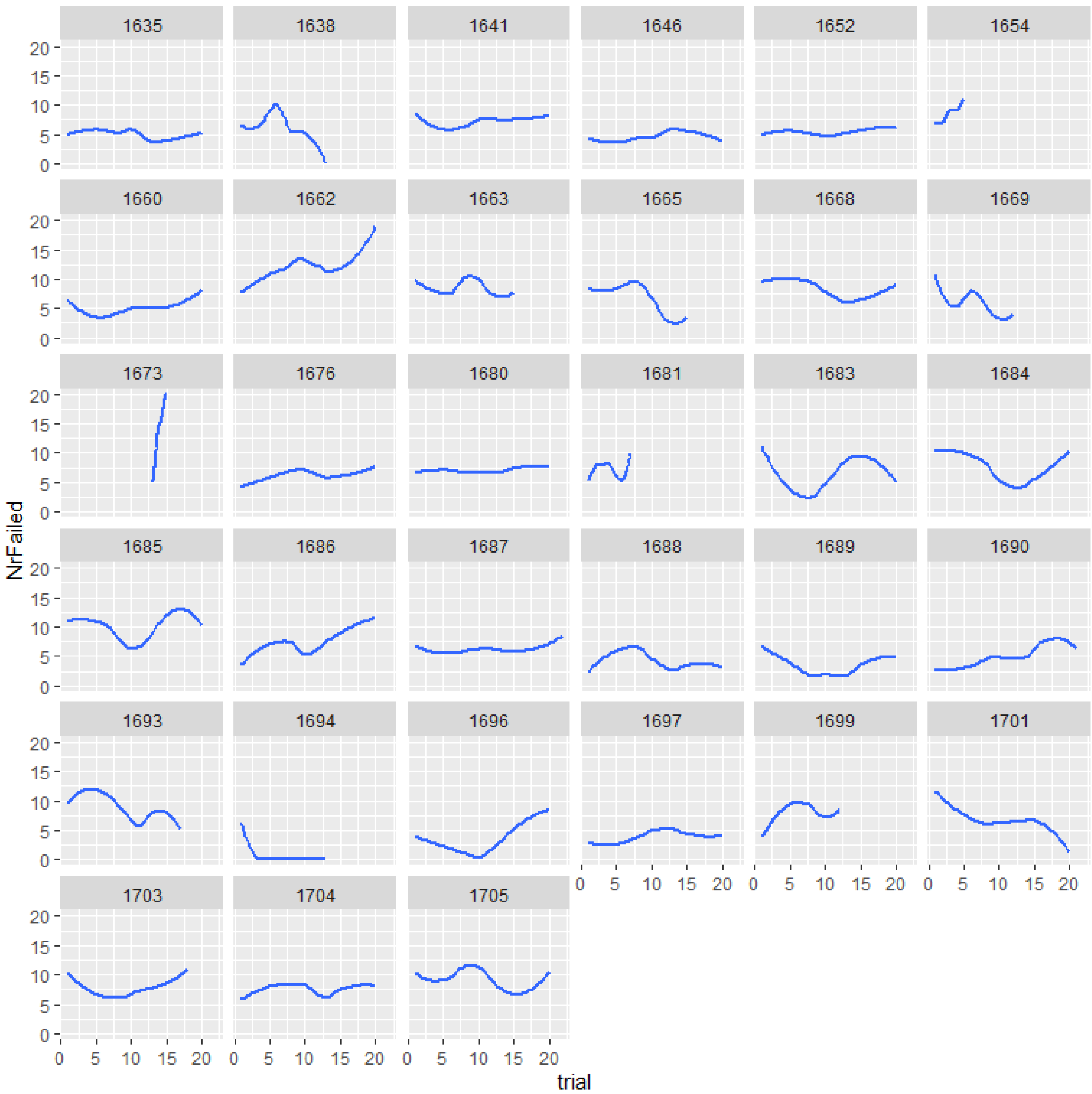

```
D_exp %>%
  filter(SafetyName == "Houden aan veilige snelheid", LessonID == "239") %>
%
  ggplot(aes(x = trial, y = OverallTaskScore)) +
  geom_smooth(se = F, method = "loess") +
  geom_point() +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```



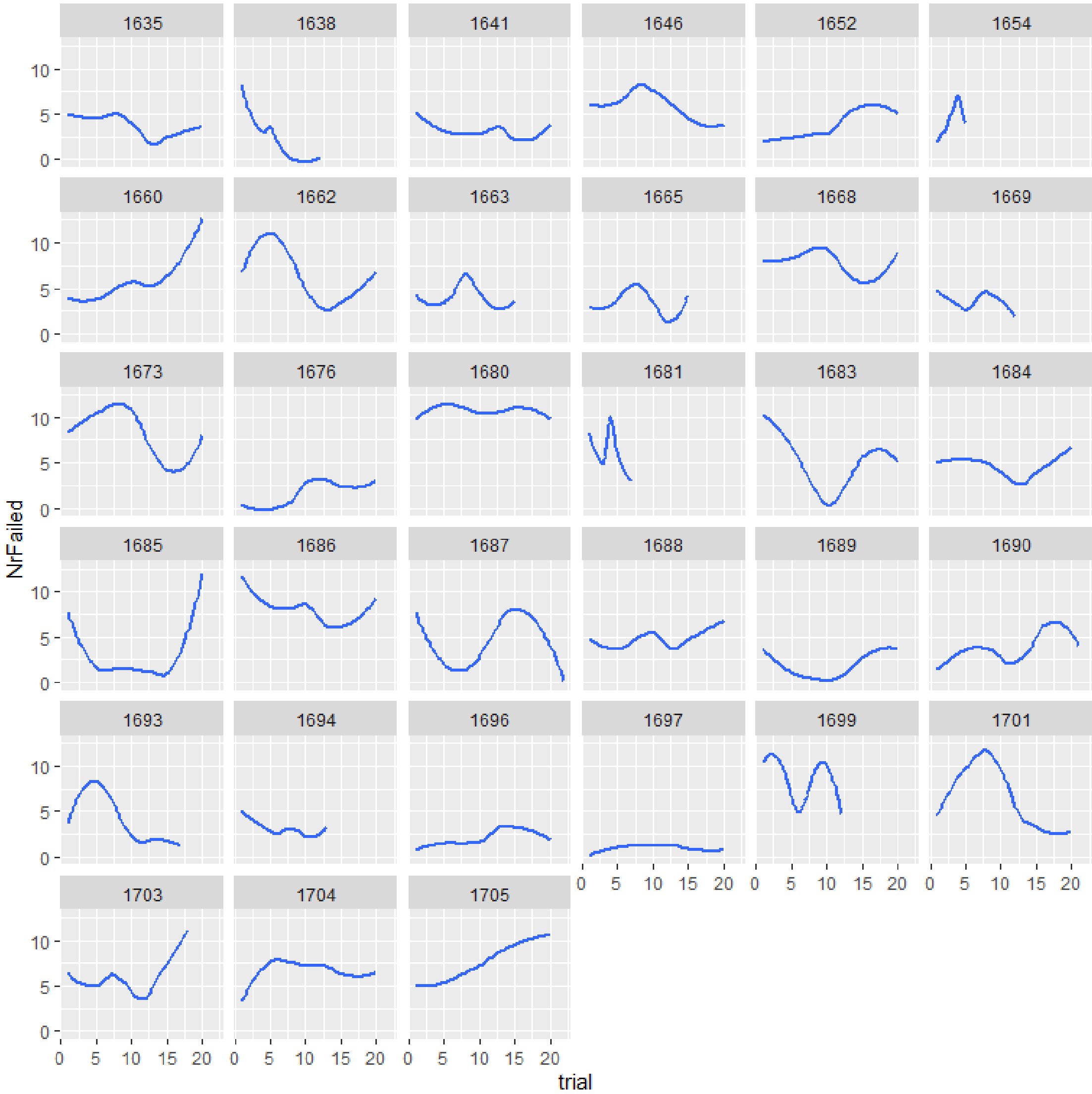

#TASK SPECIFIC ##TAKING TURNS

```
D_exp %>%
  filter(TaskName == "Positie binnen de rijbaan", LessonID == "235") %>%
  ggplot(aes(x = trial, y = NrFailed)) +
  geom_smooth(se = F, method = "loess") +
  facet_wrap(~StudentID) +
  ylim (0,20)

## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 23 rows containing non-finite values (stat_smooth).

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 12.99

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 1.01

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 1.0201

## Warning: Removed 11 rows containing missing values (geom_smooth).
```

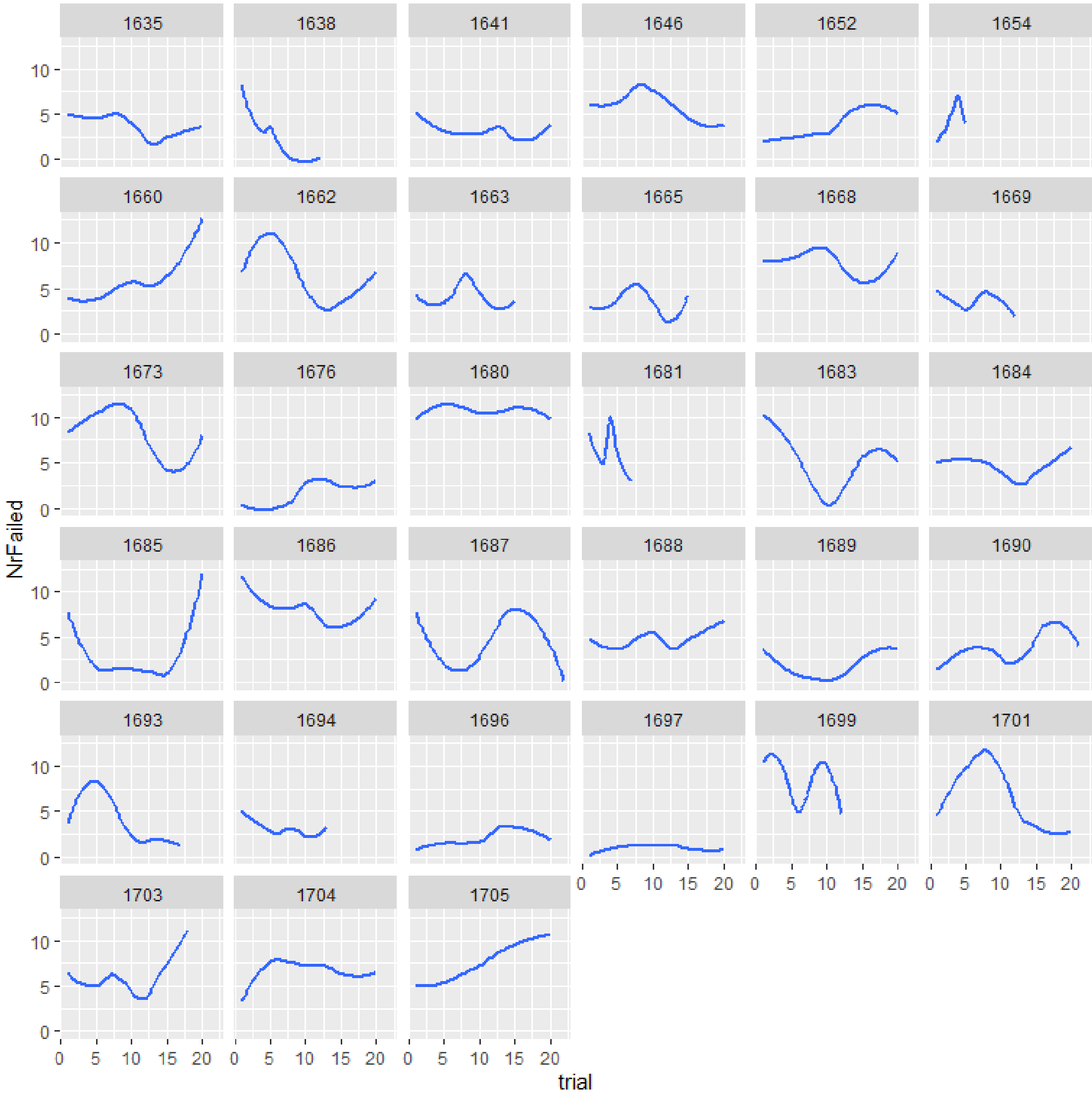


```
D_exp %>%
  filter(TaskName == "Het benaderen en nemen van een scherpe bocht", Lesson
ID == "235") %>%
  ggplot(aes(x = trial, y = NrFailed)) +
  geom_smooth(se = F, method = "loess") +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```

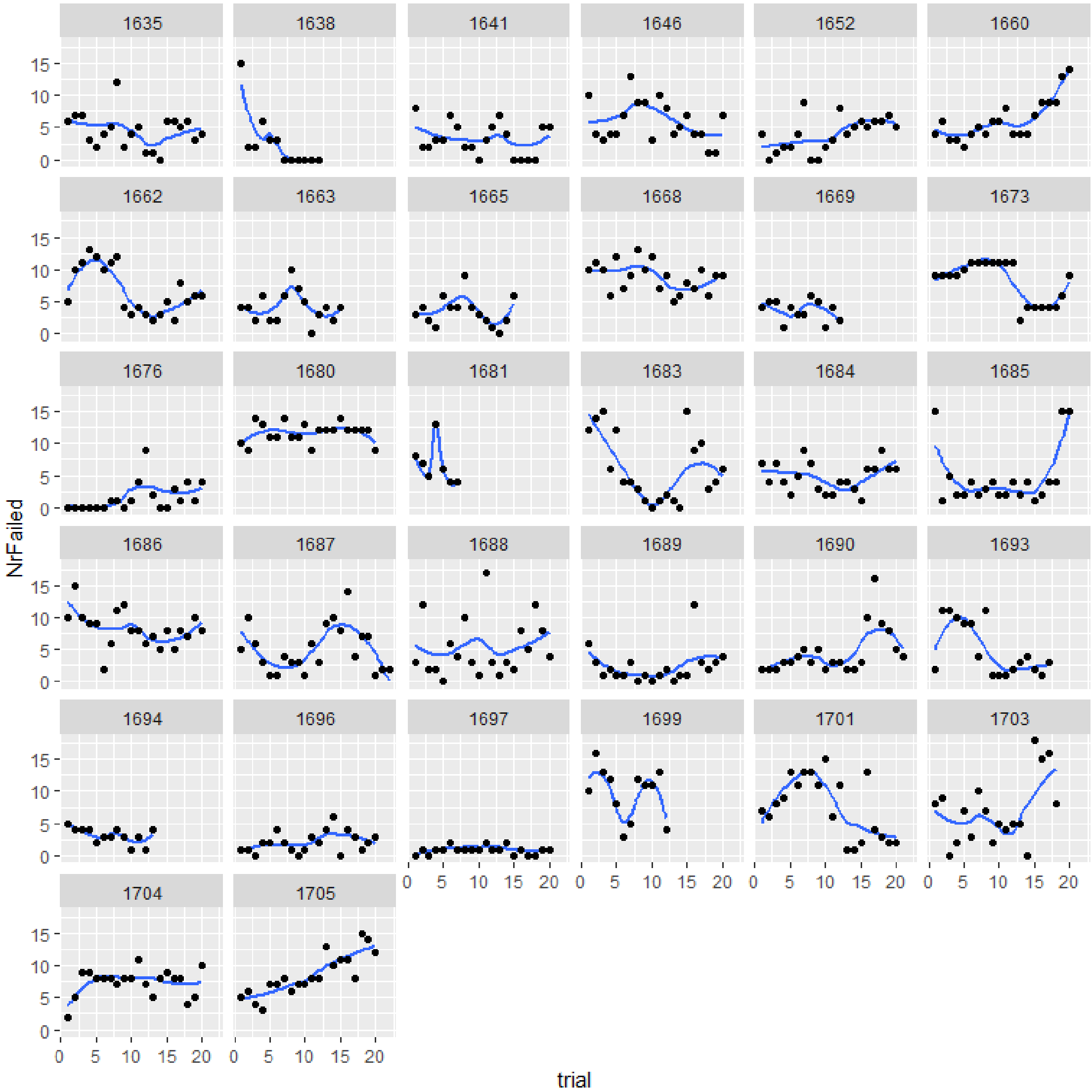



```
D_exp %>%
  filter(TaskName == "Het benaderen en nemen van een scherpe bocht", Lesson
ID == "235") %>%
  ggplot(aes(x = trial, y = NrFailed, group = StudentID)) +
  geom_smooth(se = F, method = "loess") +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0
```

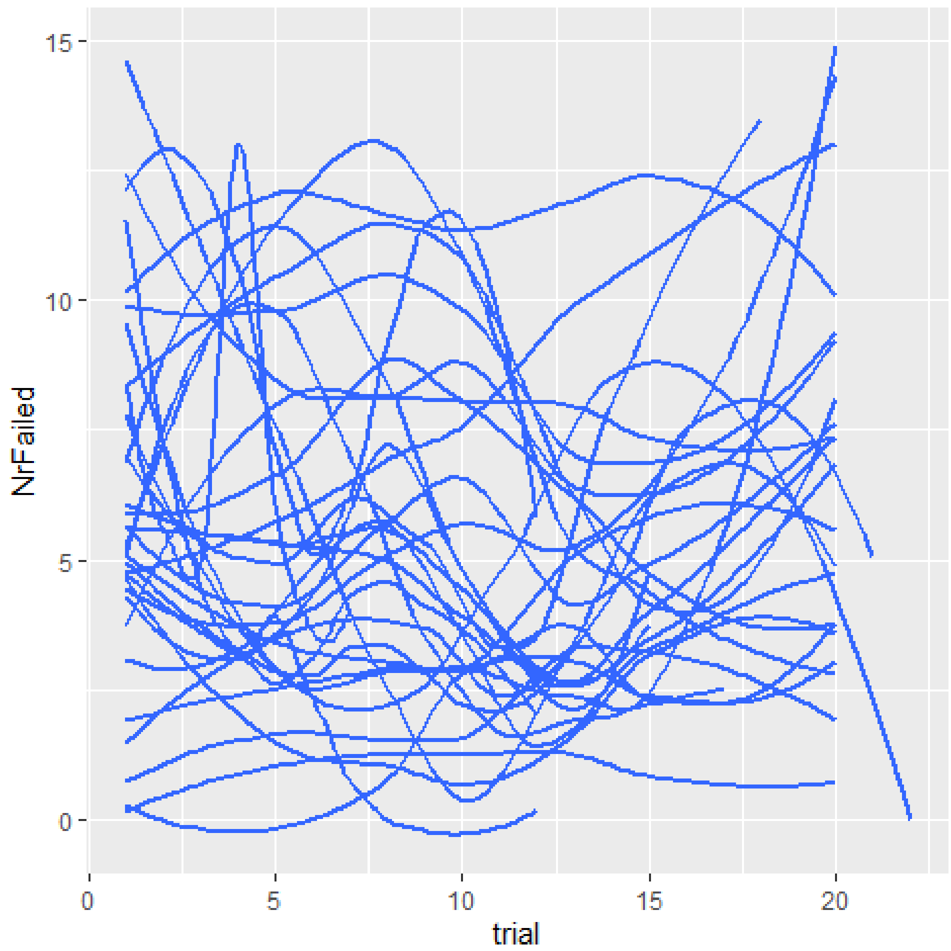
```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```



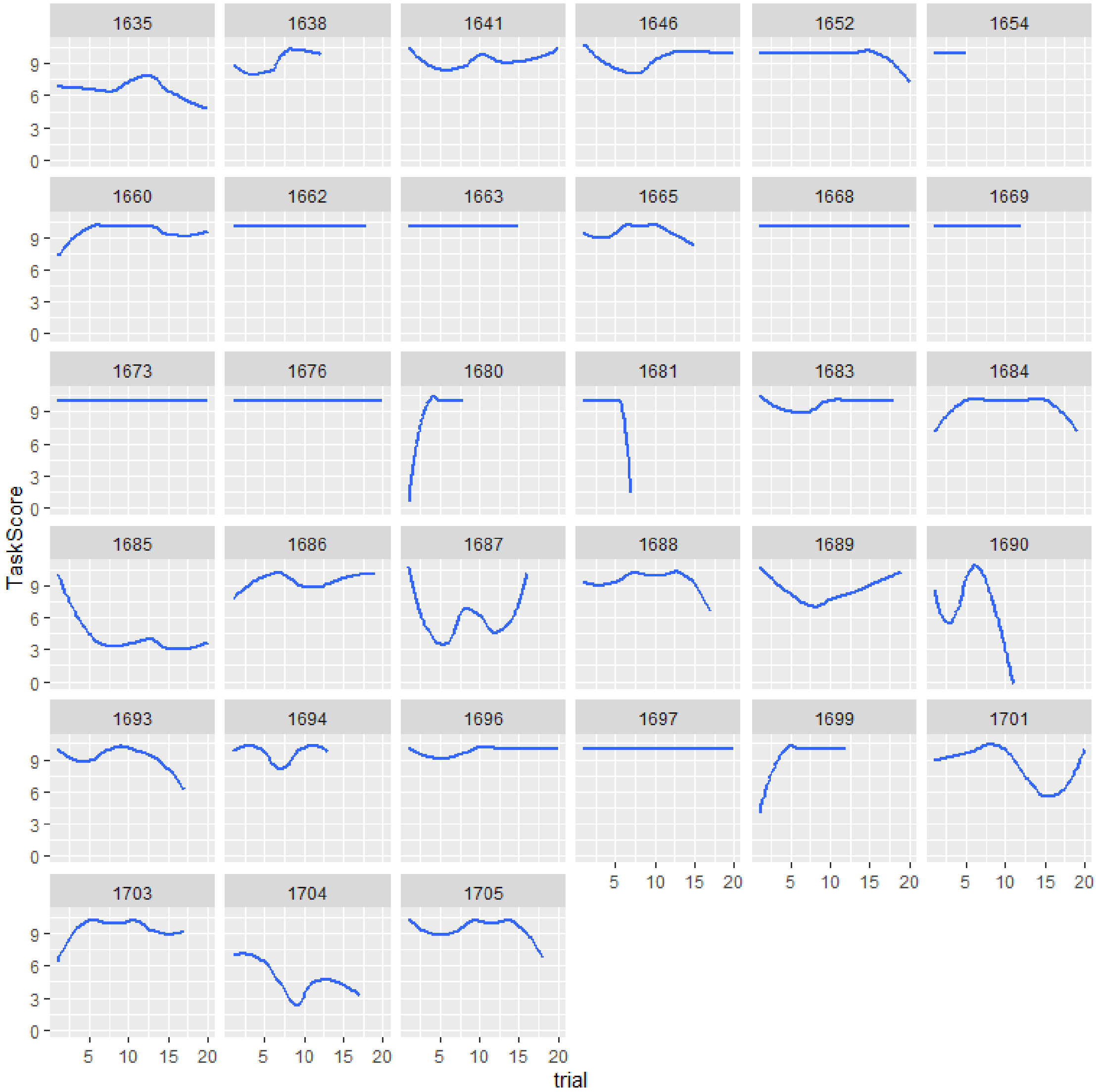

```
D_exp %>%
  filter(TaskName == "Het nemen van een bocht", LessonID == "235") %>%
  ggplot(aes(x = trial, y = NrFailed, group = StudentID)) +
  geom_smooth(se = F, method = "loess") +
  geom_point() +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'
```



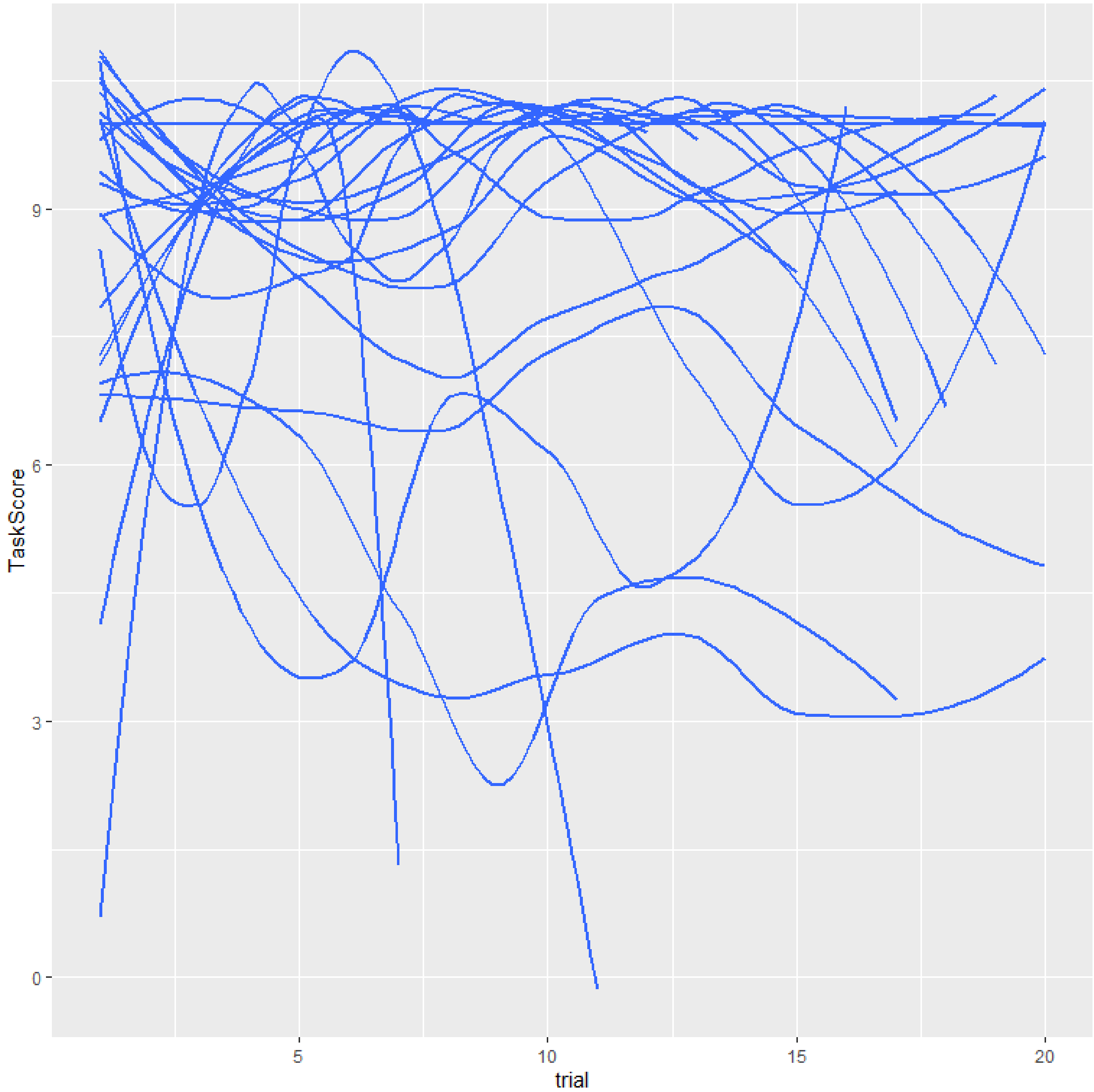

```r
D_exp %>%
  filter(TaskName == "Het nemen van een bocht", LessonID == "235") %>%
  ggplot(aes(x = trial, y = NrFailed, group = StudentID)) +
  geom_smooth(se = F, method = "loess")

## `geom_smooth()` using formula 'y ~ x'
```




```
D_exp %>%
  filter(TaskName == "Het nemen van een flauwe bocht", LessonID == "235") %
>%
  ggplot(aes(x = trial, y = TaskScore)) +
  geom_smooth(se = F, method = "loess") +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```

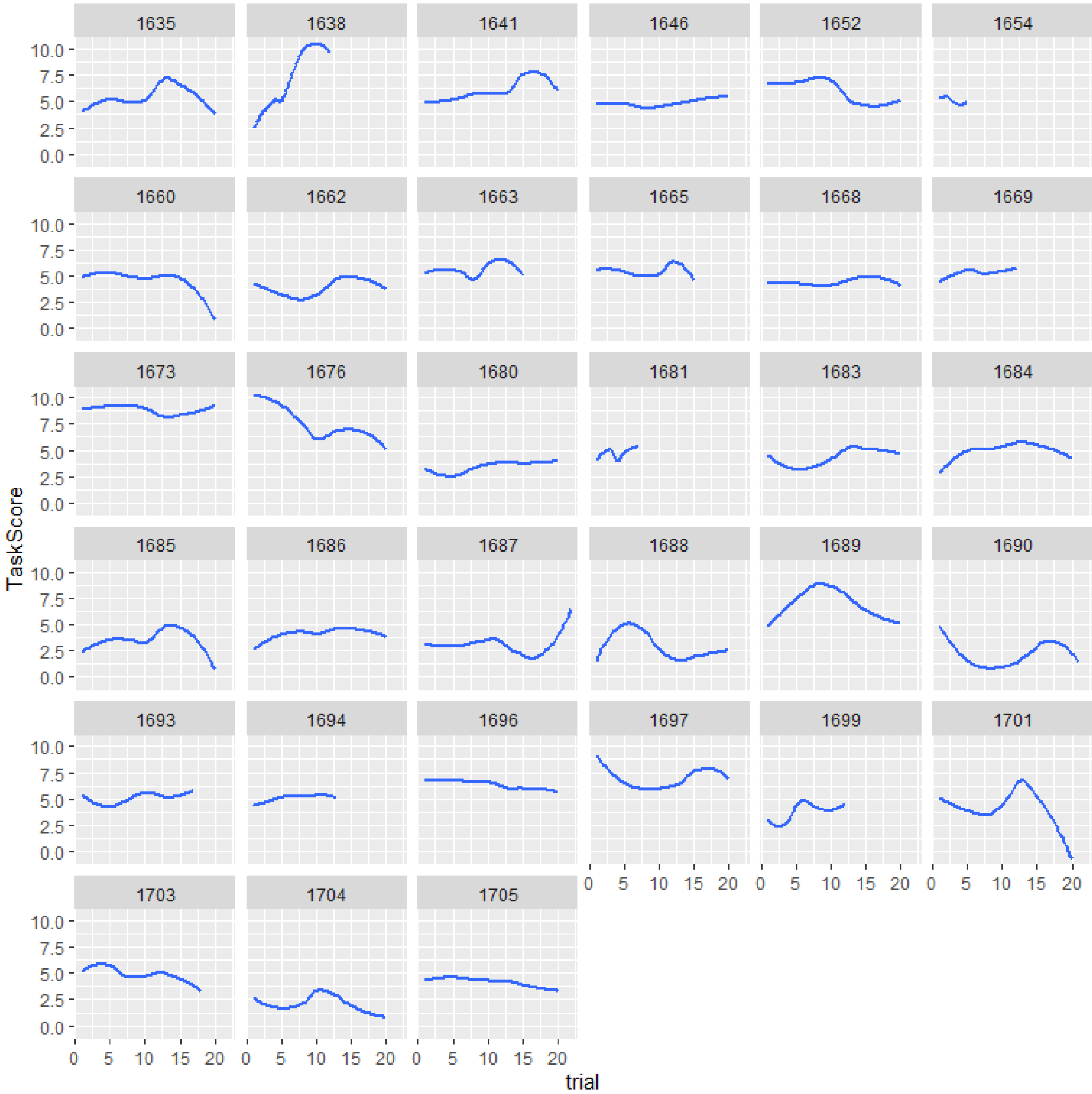



```
D_exp %>%
  filter(TaskName == "Het nemen van een flauwe bocht", LessonID == "235") %
>%
  ggplot(aes(x = trial, y = TaskScore, group = StudentID)) +
  geom_smooth(se = F, method = "loess")

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```



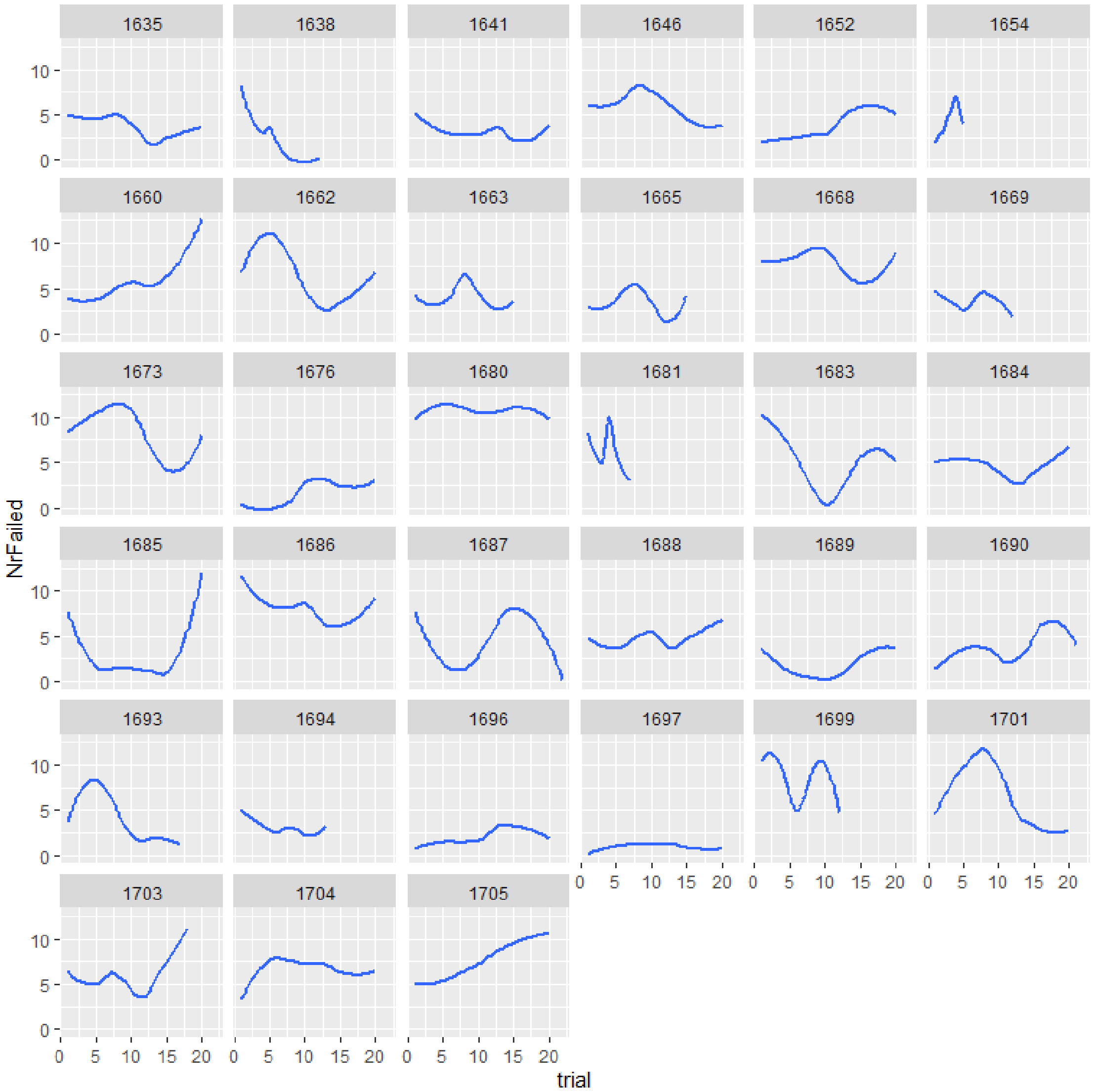

```
D_exp %>%
  filter(TaskName == "Het nemen van een haakse bocht", LessonID == "235") %
>%
  ggplot(aes(x = trial, y = TaskScore)) +
  geom_smooth(se = F, method = "loess") +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```



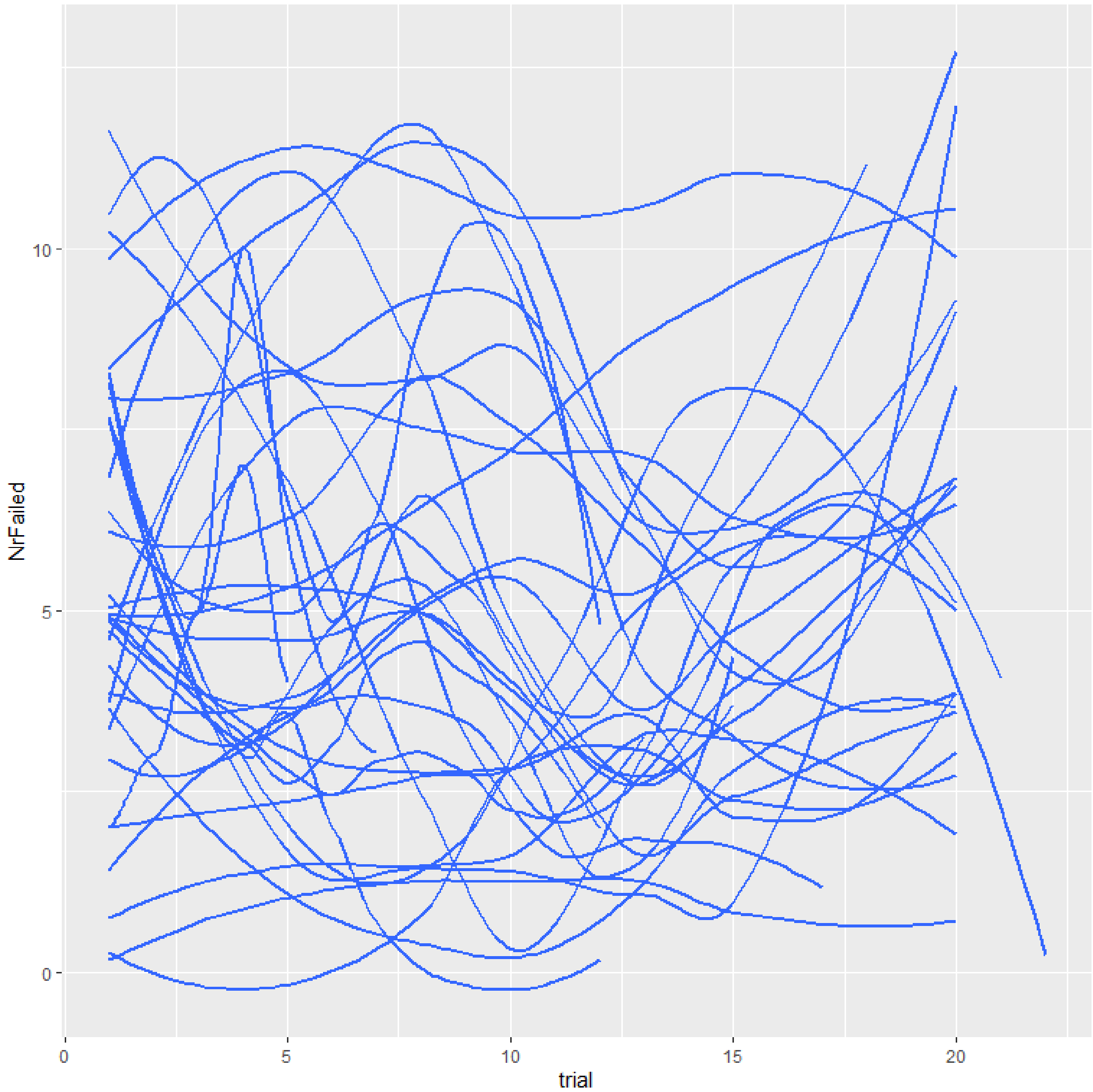

```
D_exp %>%
  filter(TaskName == "Het nemen van een haakse bocht", LessonID == "235") %
>%
  ggplot(aes(x = trial, y = NrFailed)) +
  geom_smooth(se = F, method = "loess") +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```



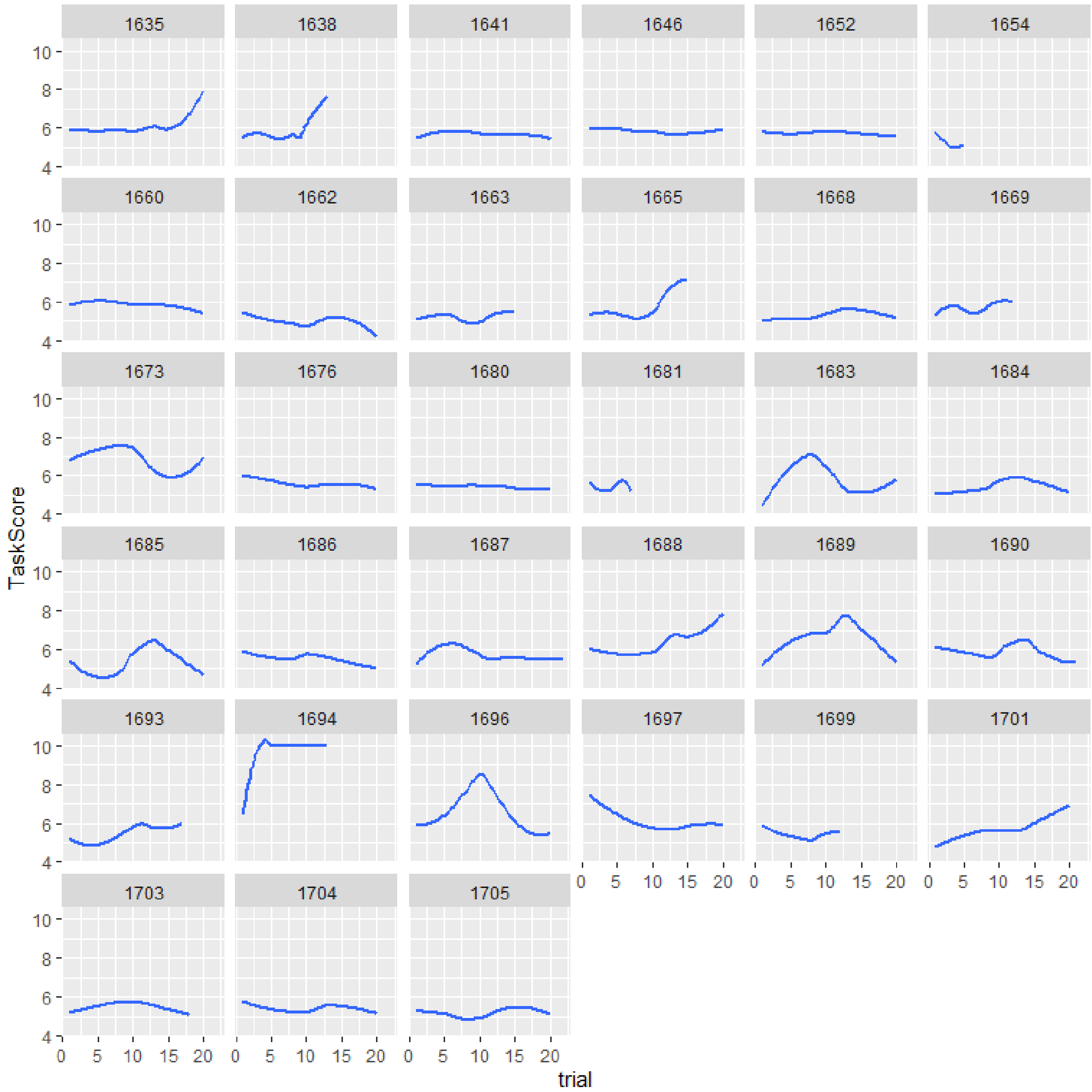

```
D_exp %>%
  filter(TaskName == "Het nemen van een haakse bocht", LessonID == "235") %
>%
  ggplot(aes(x = trial, y = NrFailed, group = StudentID)) +
  geom_smooth(se = F, method = "loess")

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```



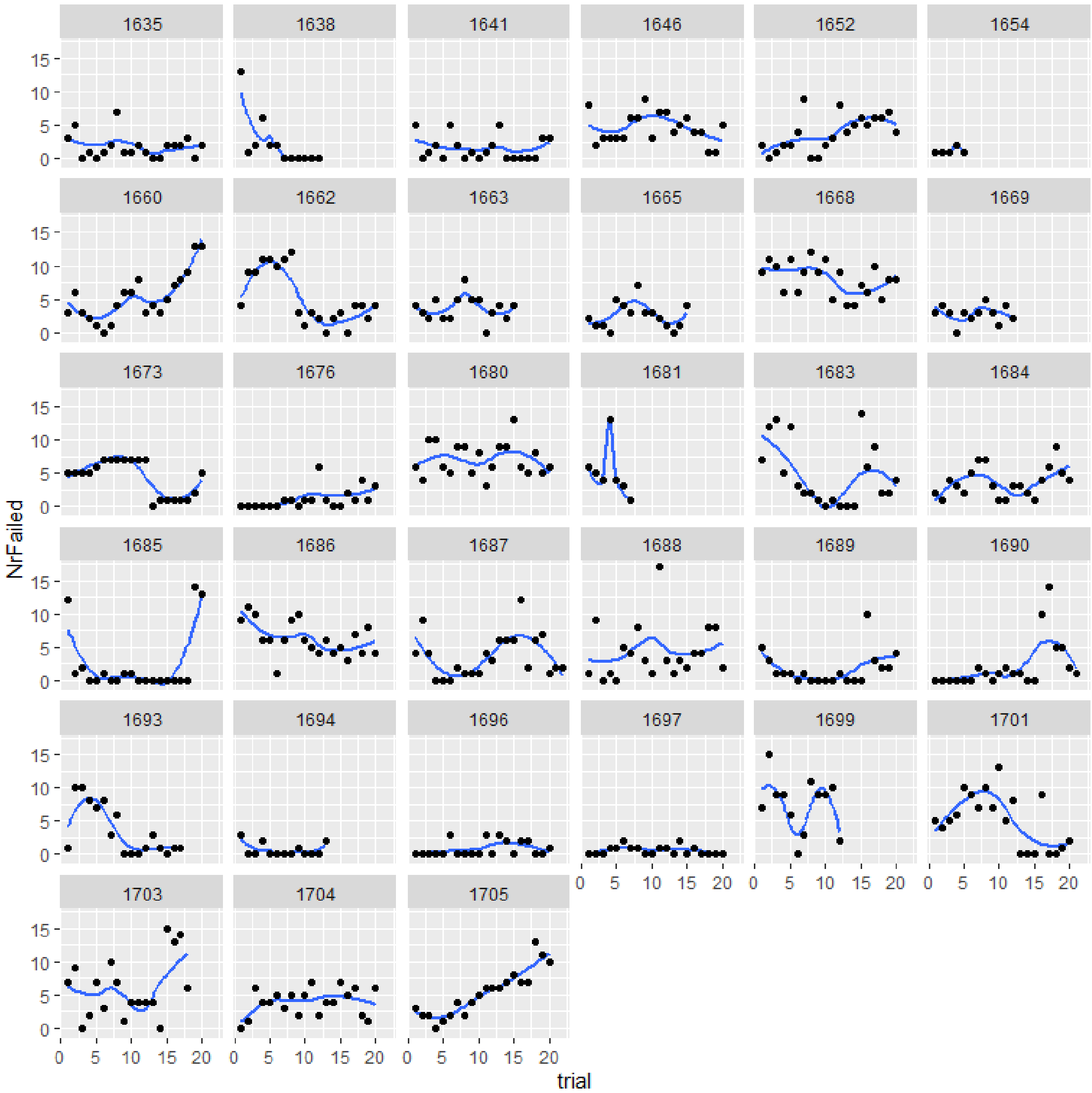

```
D_exp %>%
  filter(TaskName == "Positie binnen de rijbaan", LessonID == "235") %>%
  ggplot(aes(x = trial, y = TaskScore)) +
  geom_smooth(se = F, method = "loess") +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```



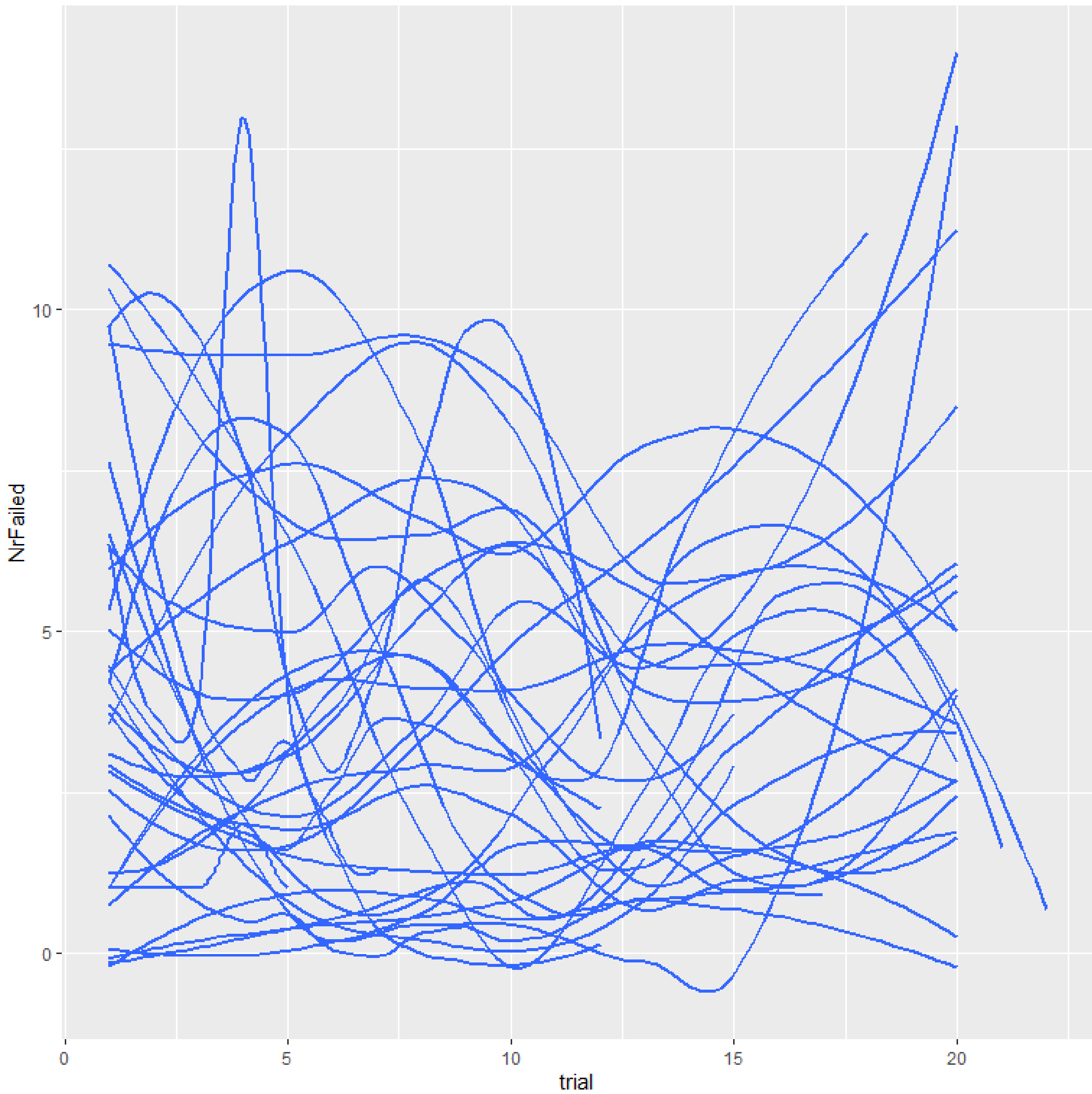

```
D_exp %>%
  filter(TaskName == "Het rijden van de juiste snelheid in de bocht", Lesso
nID == "235") %>%
  ggplot(aes(x = trial, y = NrFailed)) +
  geom_smooth(se = F, method = "loess") +
  geom_point() +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```




```
D_exp %>%
  filter(TaskName == "Het rijden van de juiste snelheid in de bocht", Lesso
nID == "235") %>%
  ggplot(aes(x = trial, y = NrFailed, group = StudentID)) +
  geom_smooth(se = F, method = "loess")

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98
```

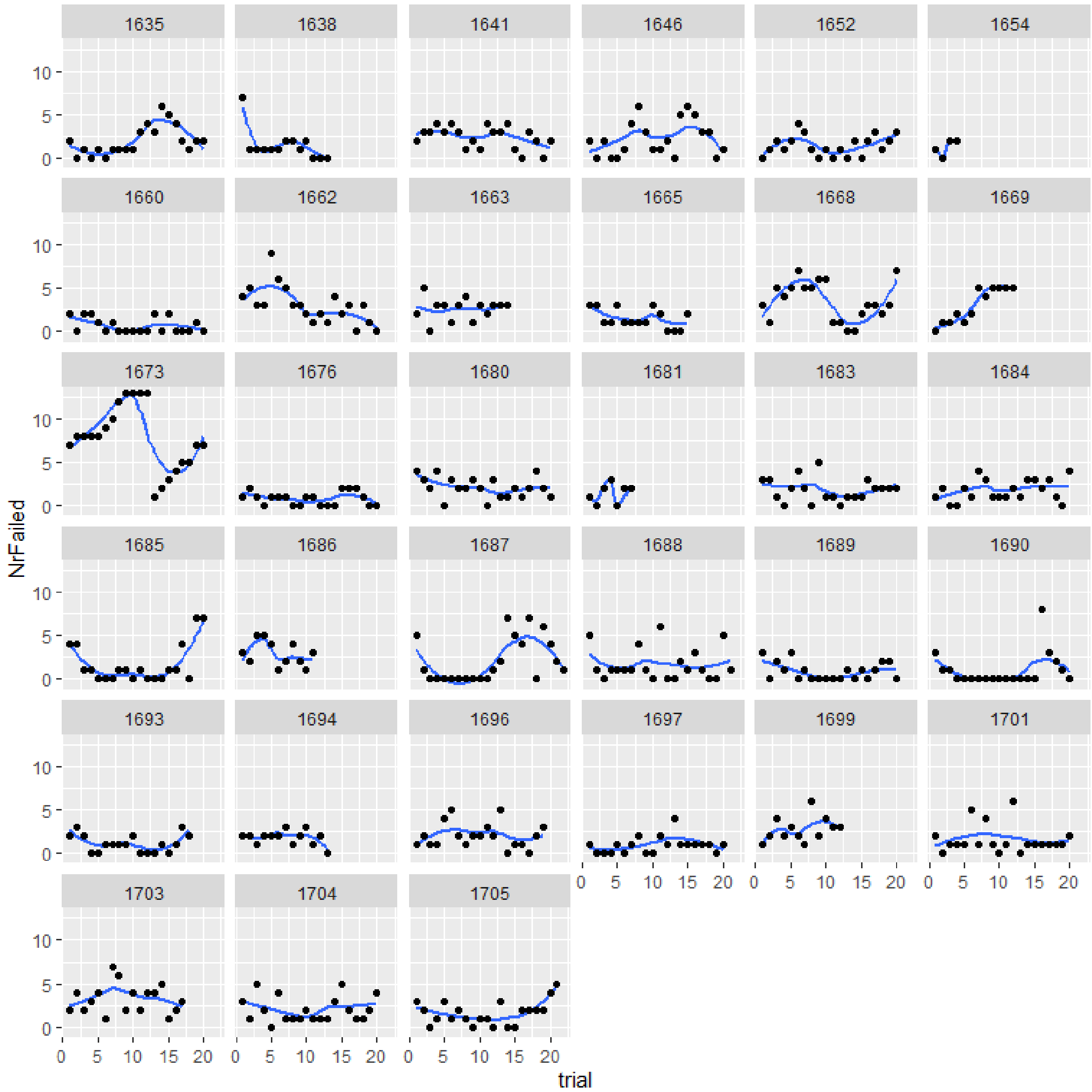
```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```




##ROUNDABOUTS

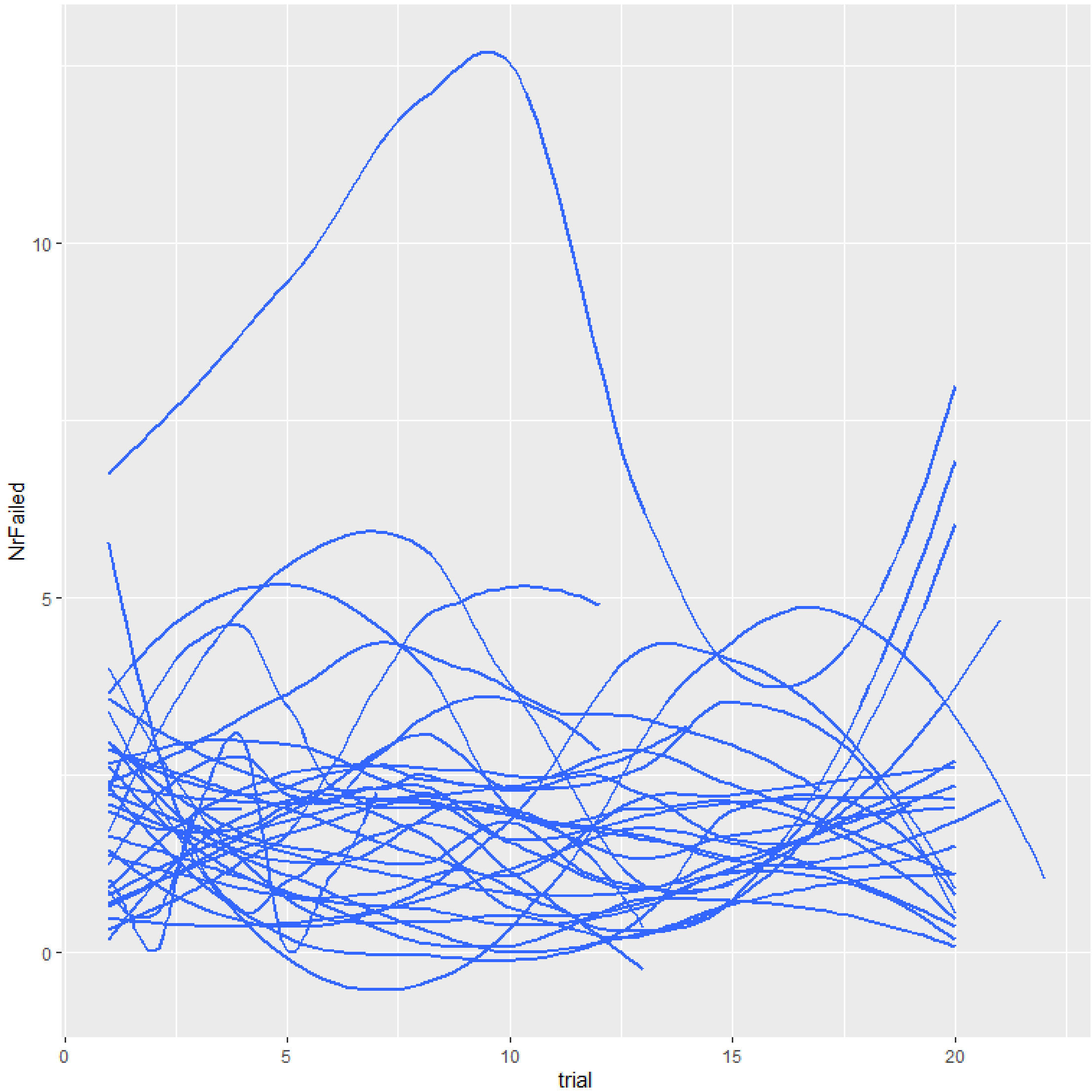
```
D_exp %>%
  filter(TaskName == "Een rotonde nemen", LessonID == "237") %>%
  ggplot(aes(x = trial, y = NrFailed)) +
  geom_smooth(se = F, method = "loess") +
  geom_point() +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.985

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.015

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0602
```



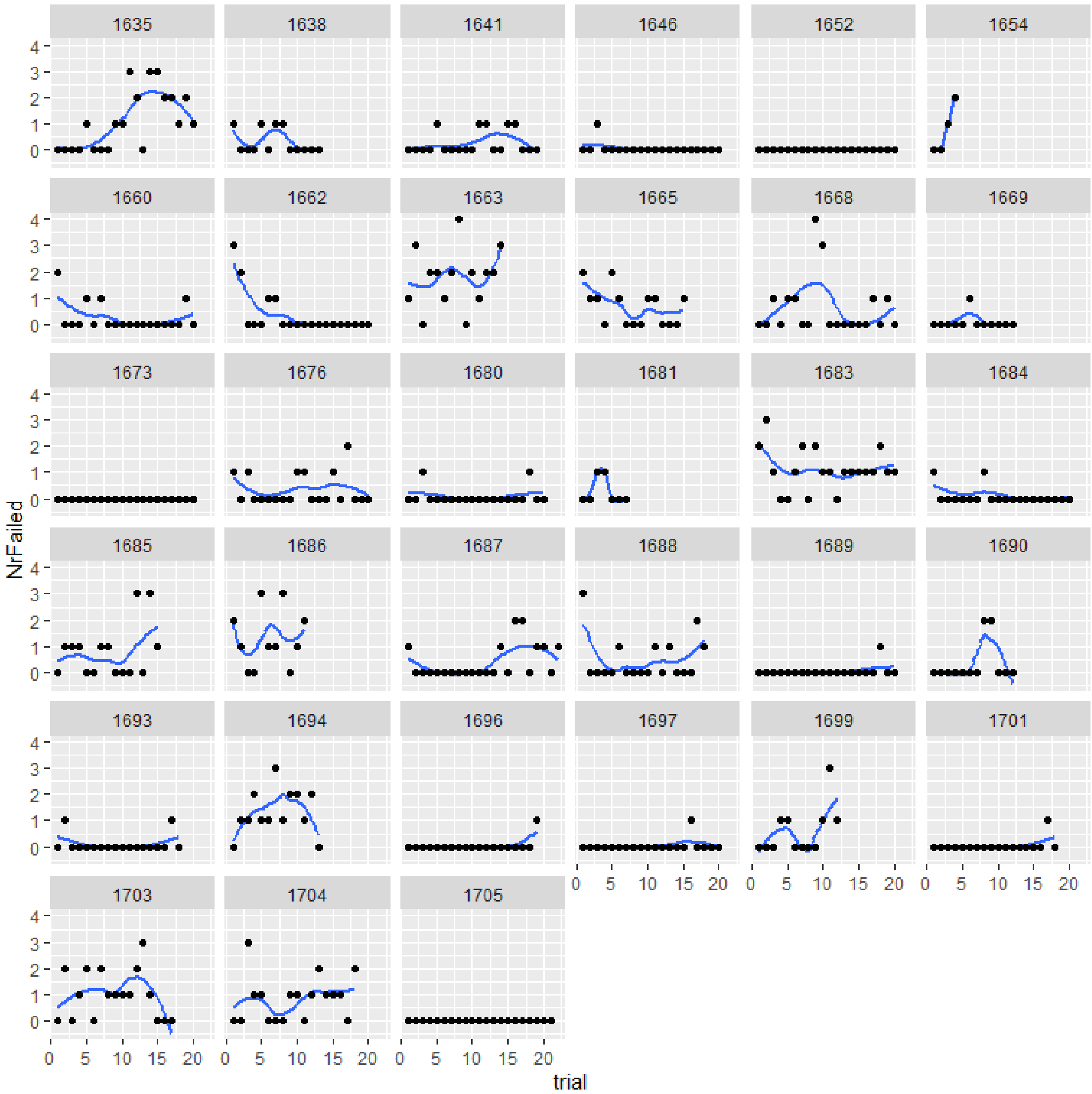

```
D_exp %>%
  filter(TaskName == "Een rotonde nemen", LessonID == "237") %>%
  ggplot(aes(x = trial, y = NrFailed, group = StudentID)) +
  geom_smooth(se = F, method = "loess")

## `geom_smooth()` using formula 'y ~ x'

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.985

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.015

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0602
```



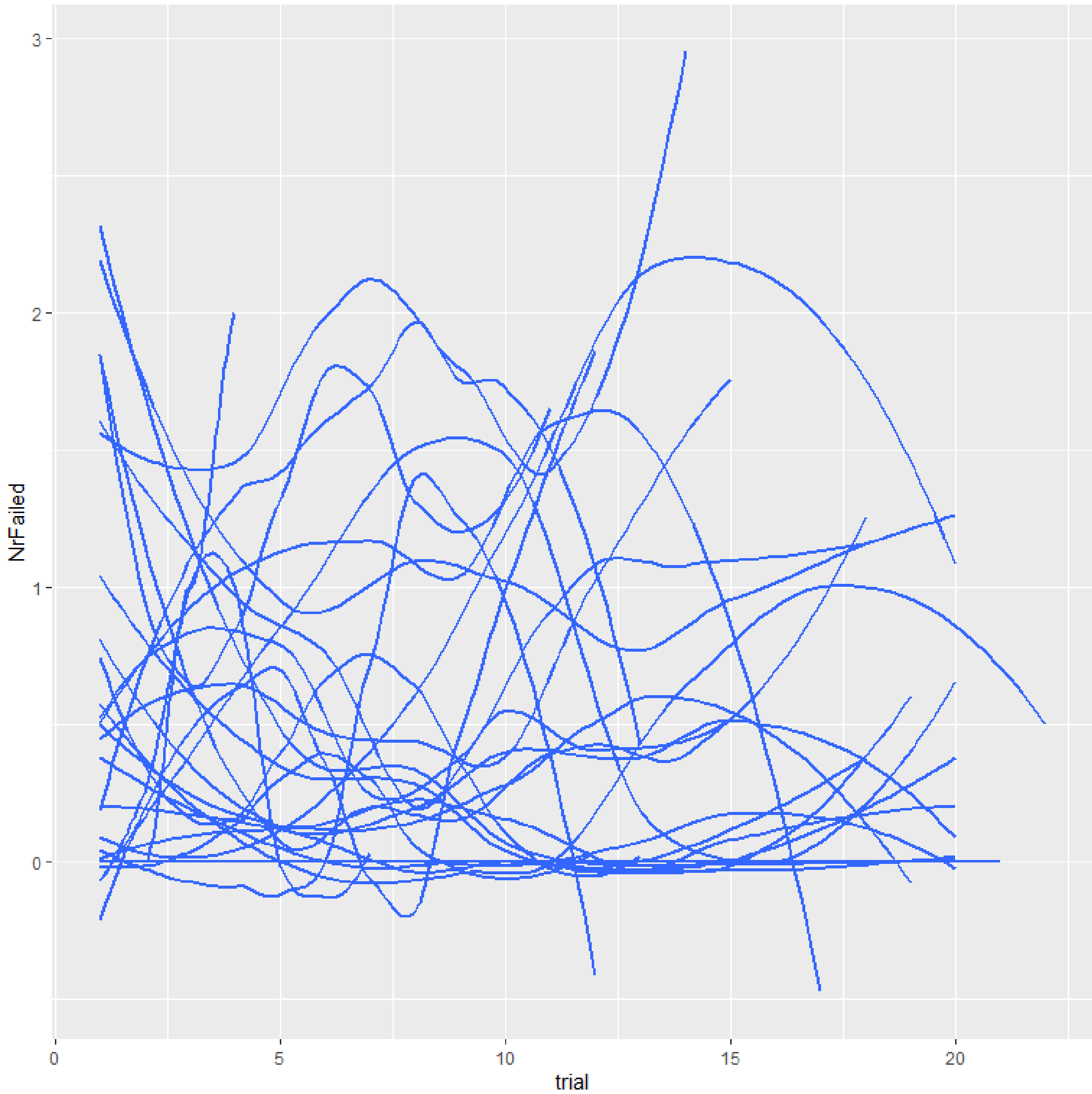

```
D_exp %>%
  filter(TaskName == "Het oversteken van de rotonde", LessonID == "237") %>
%
  ggplot(aes(x = trial, y = NrFailed)) +
  geom_smooth(se = F, method = "loess") +
  geom_point() +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.985

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.015

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0602
```




```
D_exp %>%
  filter(TaskName == "Het oversteken van de rotonde", LessonID == "237") %>
%
  ggplot(aes(x = trial, y = NrFailed, group = StudentID)) +
  geom_smooth(se = F, method = "loess")

## `geom_smooth()` using formula 'y ~ x'
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.985

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.015

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0602
```

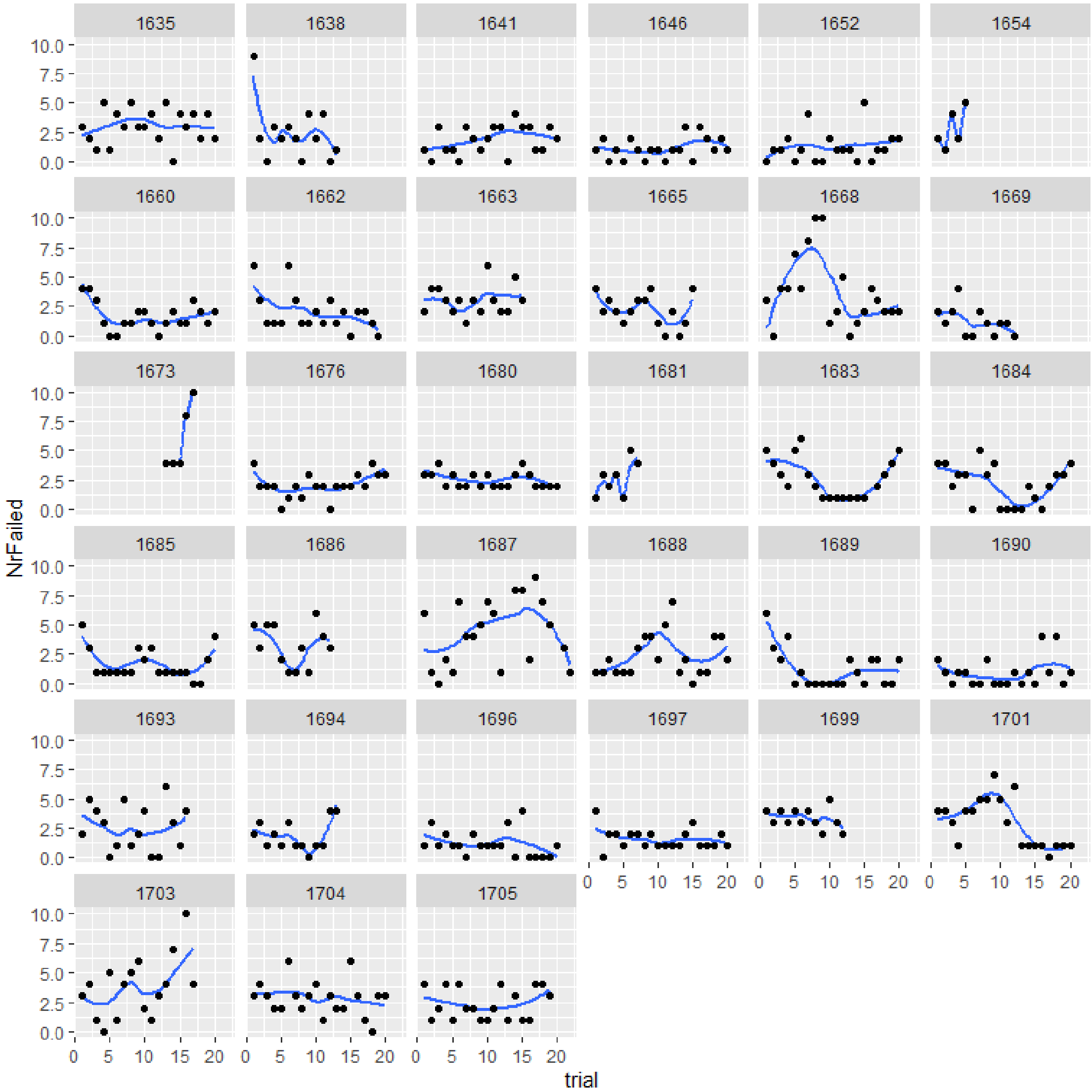



##CROSSINGS

```
D_exp %>%
  filter(TaskName == "Een kruispunt nemen", LessonID == "239") %>%
  ggplot(aes(x = trial, y = NrFailed)) +
  geom_smooth(se = F, method = "loess") +
  geom_point() +
```

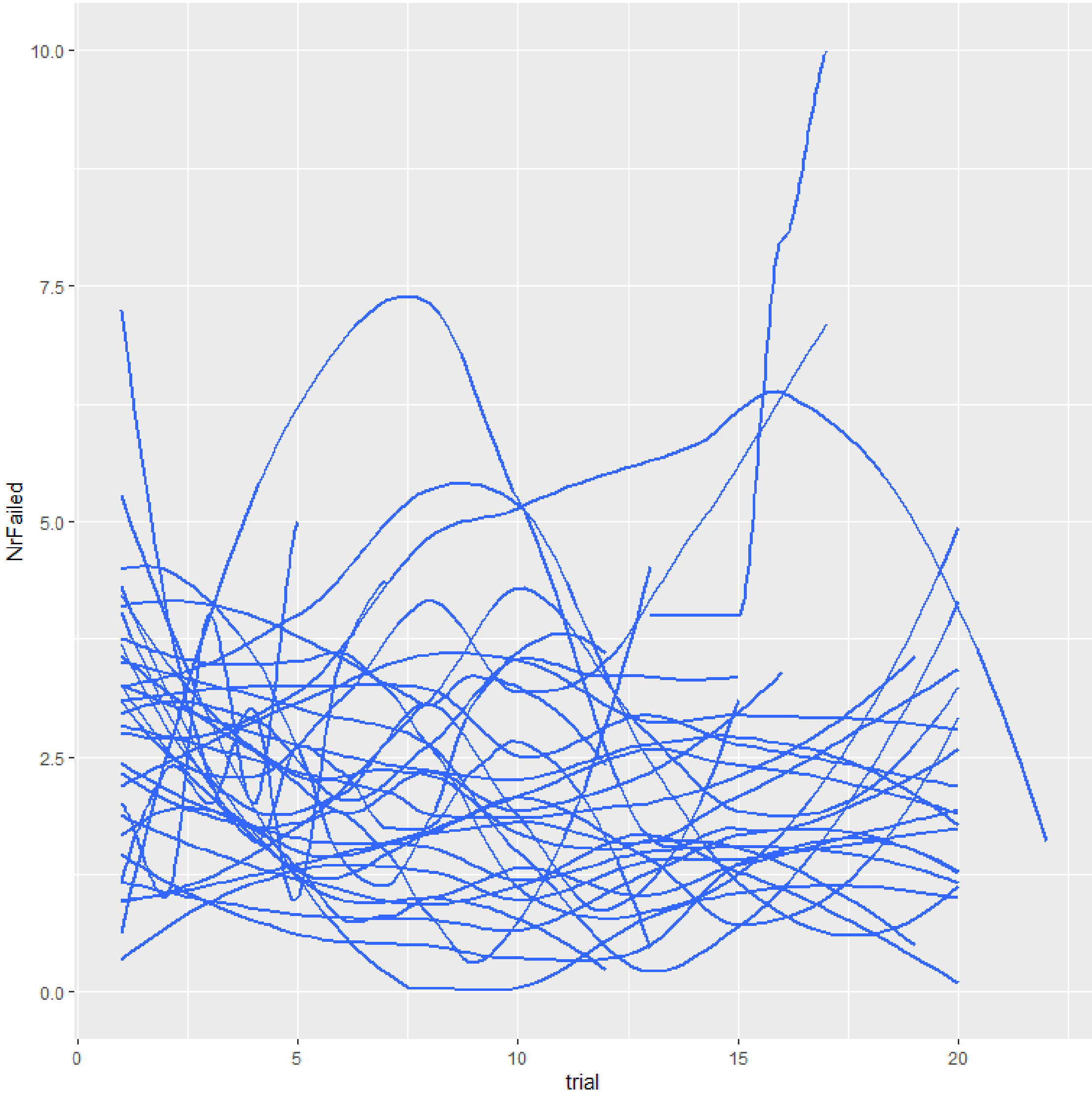
```
  ylim(0, 10) +
  facet_wrap(~StudentID)
```

```
## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 19 rows containing non-finite values (stat_smooth).

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 12.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804

## Warning: Removed 8 rows containing missing values (geom_smooth).

## Warning: Removed 19 rows containing missing values (geom_point).
```



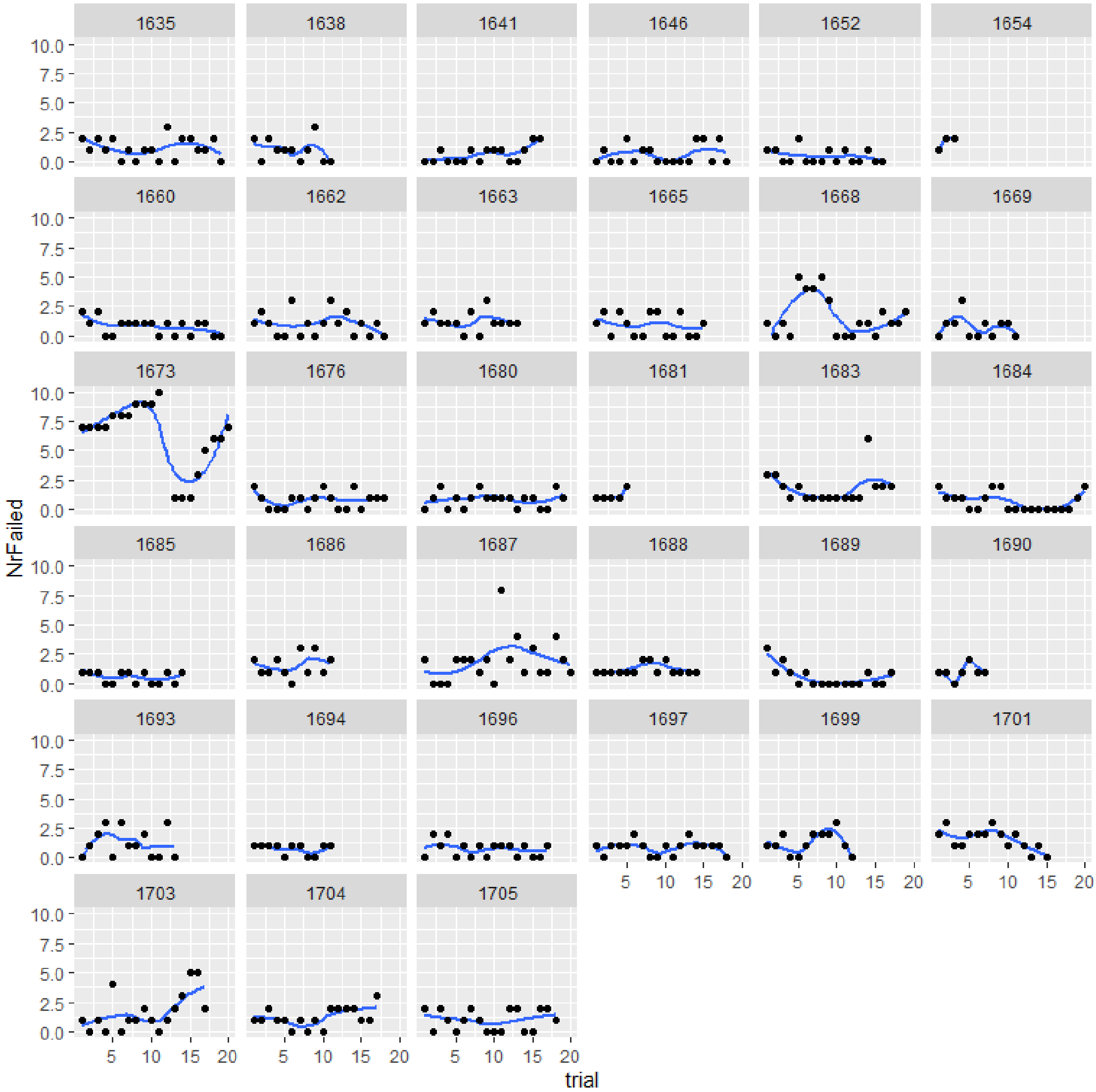

```
D_exp %>%
  filter(TaskName == "Een kruispunt nemen", LessonID == "239") %>%
  ggplot(aes(x = trial, y = NrFailed, group = StudentID)) +
  geom_smooth(se = F, method = "loess") +
  ylim(0, 10)

## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 19 rows containing non-finite values (stat_smooth).

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 12.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804

## Warning: Removed 8 rows containing missing values (geom_smooth).
```

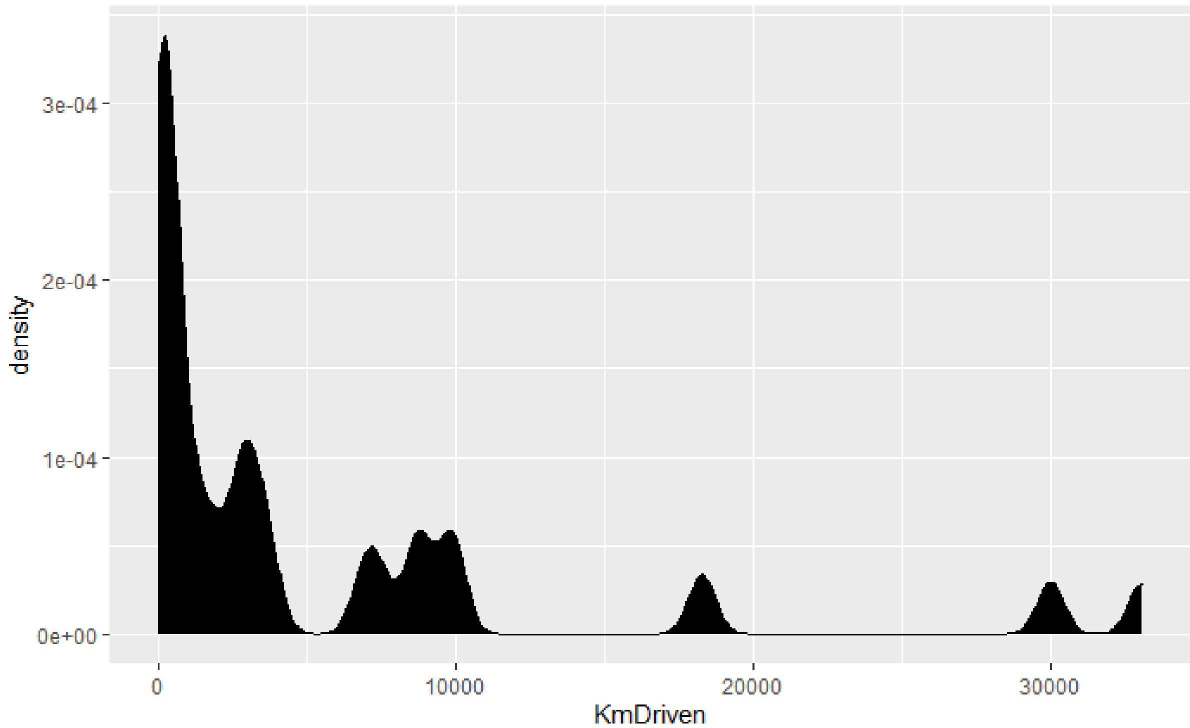



```
D_exp %>%
  filter(TaskName == "Linksaf slaan op gelijkwaardig kruispunt", LessonID =
```

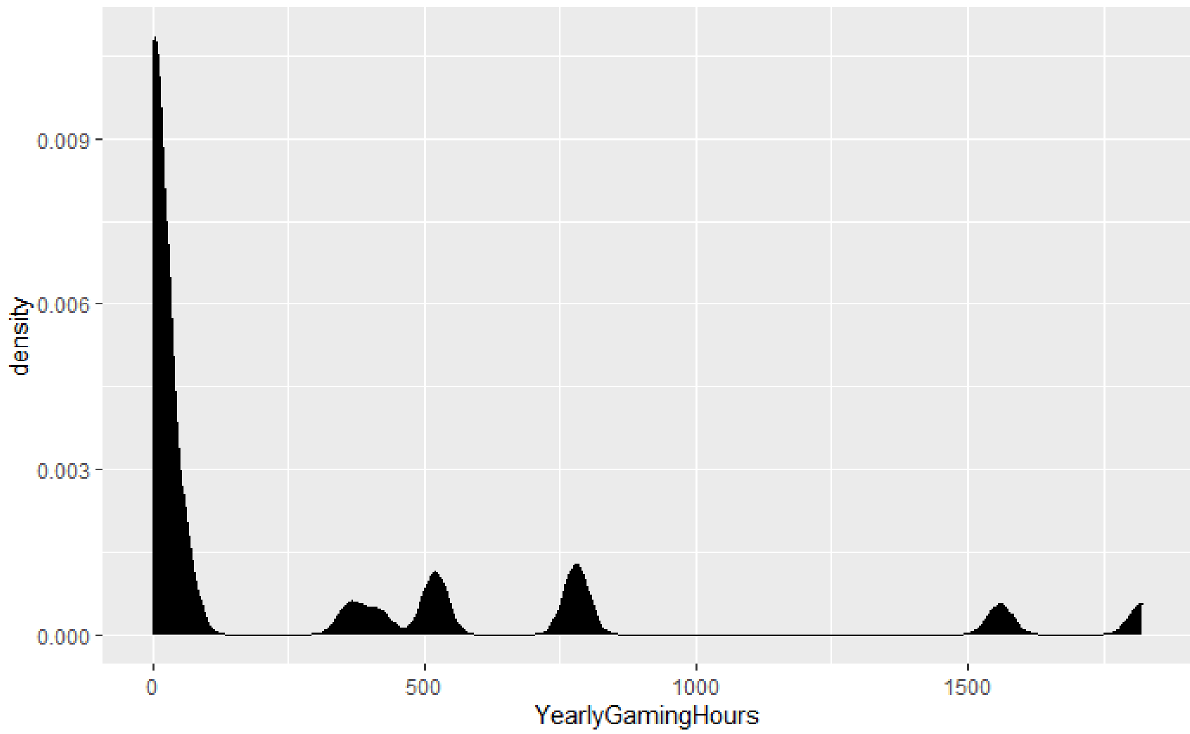
```
= "239") %>%
  ggplot(aes(x = trial, y = NrFailed, group = StudentID)) +
  geom_smooth(se = F, method = "loess") +
  geom_point() +
  ylim(0, 10) +
  facet_wrap(~StudentID)

## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.99

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 1.01

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 1.0201

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 0.98

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804

## Warning: Removed 31 rows containing missing values (geom_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).
```




```
D_exp %>%
  ggplot(aes(x = KmDriven)) +
  geom_density(fill = 1)
```




```
D_exp %>%
  ggplot(aes(x = YearlyGamingHours)) +
  geom_density(fill = 1)
```




#Workload

```
D_workload <- read_delim("~/Master/Thesis/Results/data_workload.csv",
    ";", escape_double = FALSE, trim_ws = TRUE)
```

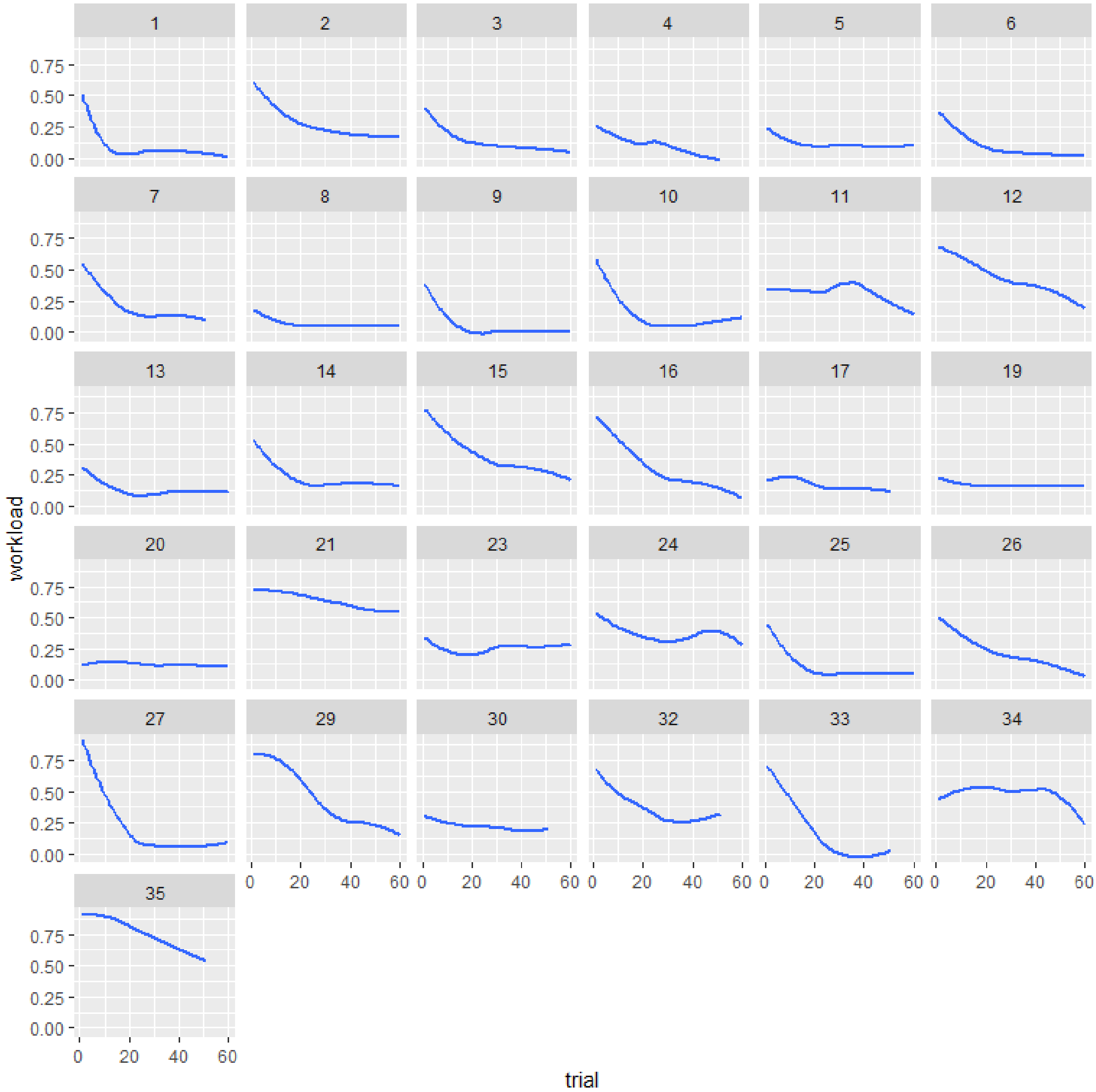
```
##
## -- Column specification --------------------------------------------------
--------
## cols(
##   workload = col_double(),
##   trial = col_double(),
##   Part = col_double(),
##   km_driven = col_double(),
##   hrs_gaming = col_double(),
##   exp_biking = col_character()
## )
```

```
D_workload %>%
  sample_n(10)
```

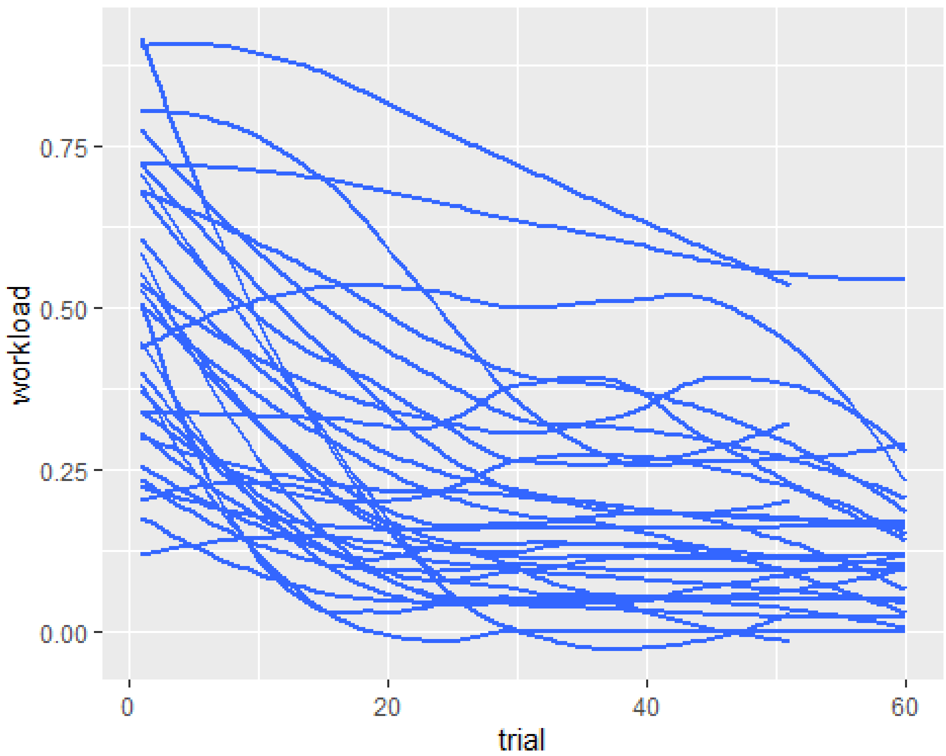
| workload | trial | Part | km_driven | hrs_gaming | exp_biking |
|---|---|---|---|---|---|
| 0.4759795 | 44 | 11 | 848 | 0 | Frequent biker |
| 0.2856938 | 56 | 34 | 2652 | 60 | Frequent biker |
| 0.0954081 | 28 | 14 | 339 | 0 | Frequent biker |
| 0.4284081 | 12 | 32 | 3111 | 416 | Frequent biker |
| 0.0002652 | 26 | 9 | 0 | 2 | Frequent biker |
| 0.0478367 | 30 | 8 | 0 | 24 | Frequent biker |
| 0.0478367 | 34 | 30 | 3339 | 0 | Frequent biker |
| 0.2381224 | 29 | 26 | 7398 | 520 | Frequent biker |
| 0.1429795 | 35 | 30 | 3339 | 0 | Frequent biker |
| 0.2856938 | 35 | 32 | 3111 | 416 | Frequent biker |

```
D_workload %>%
    ggplot(aes(x = trial, y = workload)) +
    geom_smooth(se = F) +
    facet_wrap(~Part)
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
D_workload %>%
    ggplot(aes(x = trial, y = workload, group = Part)) +
    geom_smooth(se = F)

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




##Model estimation

```
F_lacy_workload <- formula(workload ~ exp(asym) + exp(ampl) * inv_logit((1
- ctch))^trial)

F_lacy_prior_workload <- c(set_prior("normal(-0.7, 2)", nlpar = "ampl"),
                          set_prior("normal(-3, 2)", nlpar = "ctch"),
                          set_prior("normal(-1, 2)", nlpar = "asym"))

F_lacy_prior_workload_1 <- c(set_prior("normal(-0.7, 2.7)", nlpar = "ampl")
,
                          set_prior("normal(-3, 2)", nlpar = "ctch"),
                          set_prior("normal(-1, 2)", nlpar = "asym"))

F_lacy_ef_workload <- list(formula(ampl ~ 1 + km_driven + hrs_gaming + exp_
biking + (1|Part)),
                     formula(ctch ~ 1 + (1|Part)),
                     formula(asym ~ 1 + (1|Part)))

M_workload_1 <-
  D_workload %>%
  brm(bf(F_lacy_workload,
         flist = F_lacy_ef_workload,
         nl = T),
      prior = F_lacy_prior_workload,
      family = beta(link = identity), iter = 4000,
      data = .)

## Compiling Stan program...

## Start sampling
```

```
## Warning: There were 8000 transitions after warmup that exceeded the maxi
mum treedepth. Increase max_treedepth above 10. See
## http://mc-stan.org/misc/warnings.html#maximum-treedepth-exceeded

## Warning: Examine the pairs() plot to diagnose sampling problems

## Warning: The largest R-hat is 1.82, indicating chains have not mixed.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#r-hat

## Warning: Bulk Effective Samples Size (ESS) is too low, indicating poster
ior means and medians may be unreliable.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#bulk-ess

## Warning: Tail Effective Samples Size (ESS) is too low, indicating poster
ior variances and tail quantiles may be unreliable.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#tail-ess
```

```
fixef(M_workload_1, mean.func = exp)
```

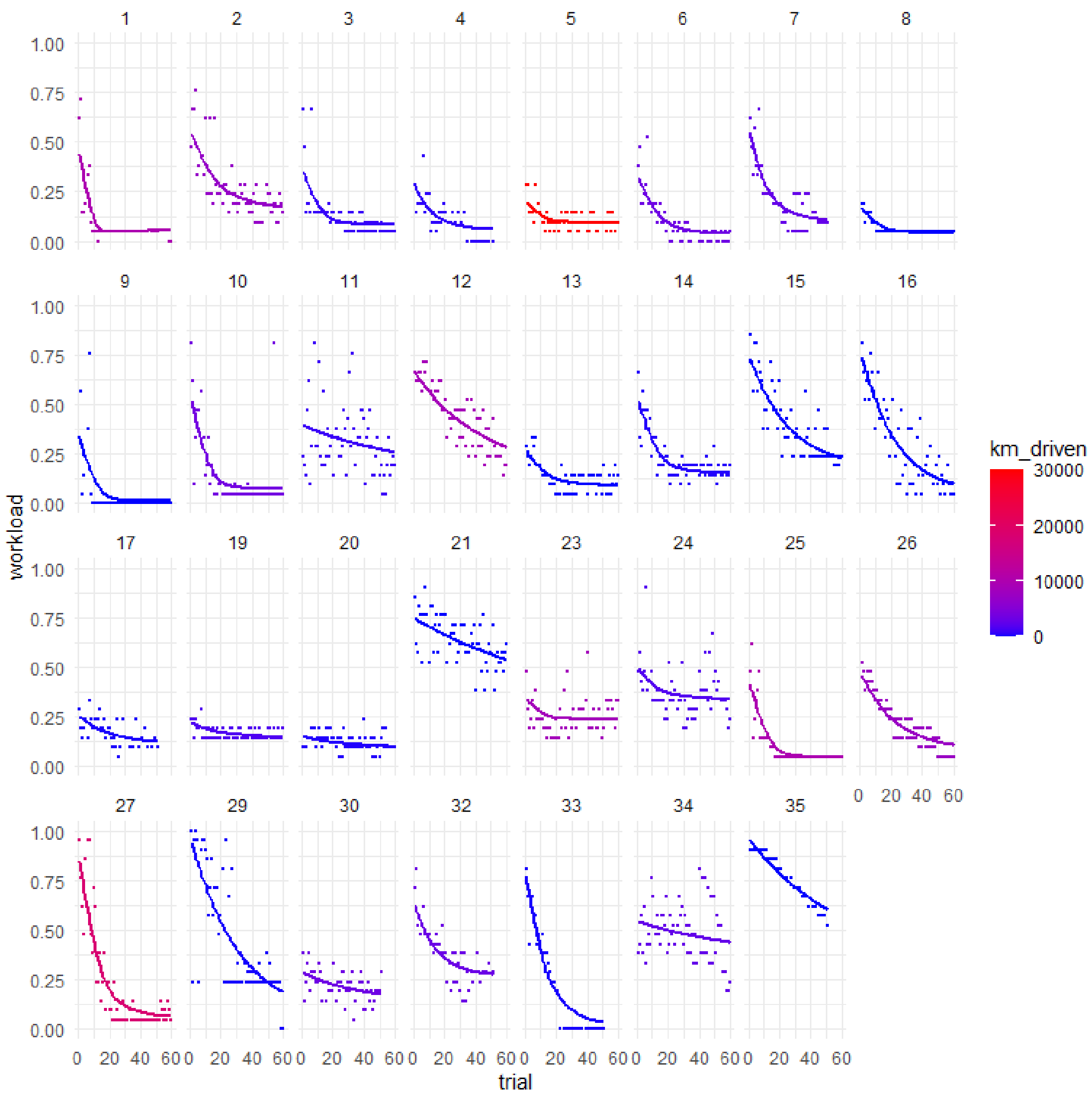
```
## Warning: `funs()` was deprecated in dplyr 0.8.0.
## Please use a list of either functions or lambdas:
##
##   # Simple named list:
##   list(mean = mean, median = median)
##
##   # Auto named with `tibble::lst()`:
##   tibble::lst(mean, median)
##
##   # Using lambdas
##   list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
```

*Coefficient estimates with 95% credibility limits*

| nonlin | fixef | center | lower | upper |
|---|---|---|---|---|
| ampl | Intercept | 0.5767261 | 0.4133821 | 1.0981831 |
| ampl | km_driven | 1.0000104 | 0.9999820 | 1.0000403 |
| ampl | hrs_gaming | 0.9998850 | 0.9994984 | 1.0003911 |
| ampl | exp_bikingFrequentbiker | 0.8083654 | 0.4110529 | 1.2089220 |
| ctch | Intercept | 0.3298381 | 0.2420029 | 0.7733222 |
| asym | Intercept | 0.0782651 | 0.0678640 | 0.0947970 |

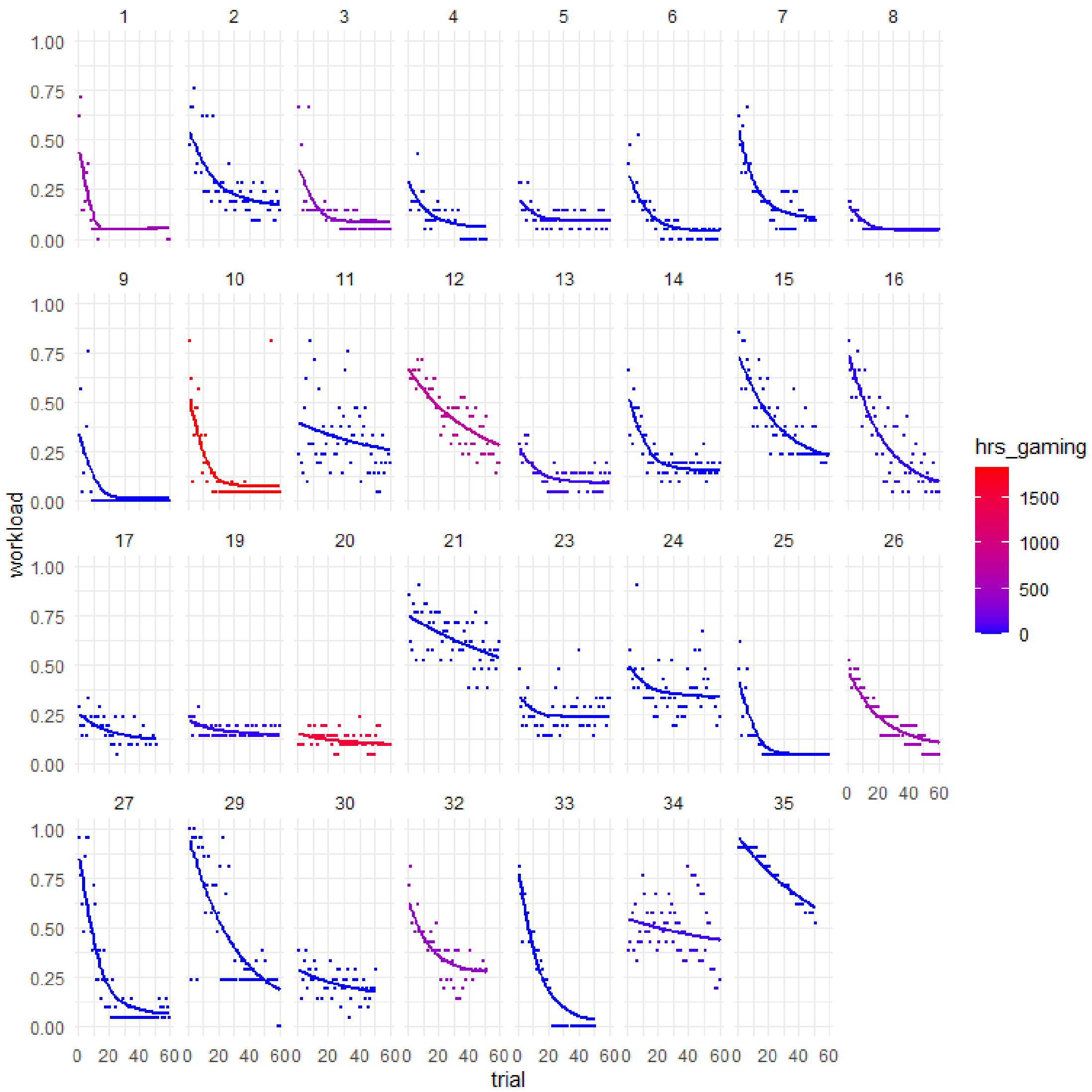
```
P_M_workload_1 <- posterior(M_workload_1)
PP_M_workload_1 <- post_pred(M_workload_1)

T_pred_M_workload_1 <- PP_M_workload_1 %>%
  group_by(Obs) %>%
  summarize(center = median(value))

D_workload$M_workload_1 <- T_pred_M_workload_1$center
D_workload$M_workload_1_resid <- D_workload$workload - D_workload$M_workloa
d_1
```

```
D_M_workload_1 <-
  as_tibble(M_workload_1$data) %>%
  mutate(M_workload_1 = T_pred_M_workload_1$center)

D_workload %>%
  ggplot(aes(x = trial, y = workload, color = km_driven)) +
  facet_wrap(~Part, nrow = 4) +
  geom_point(size = .2) +
  geom_smooth(aes(y = M_workload_1), se = F) +
  scale_color_gradient(low = "blue", high = "red") +
  theme_minimal()

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



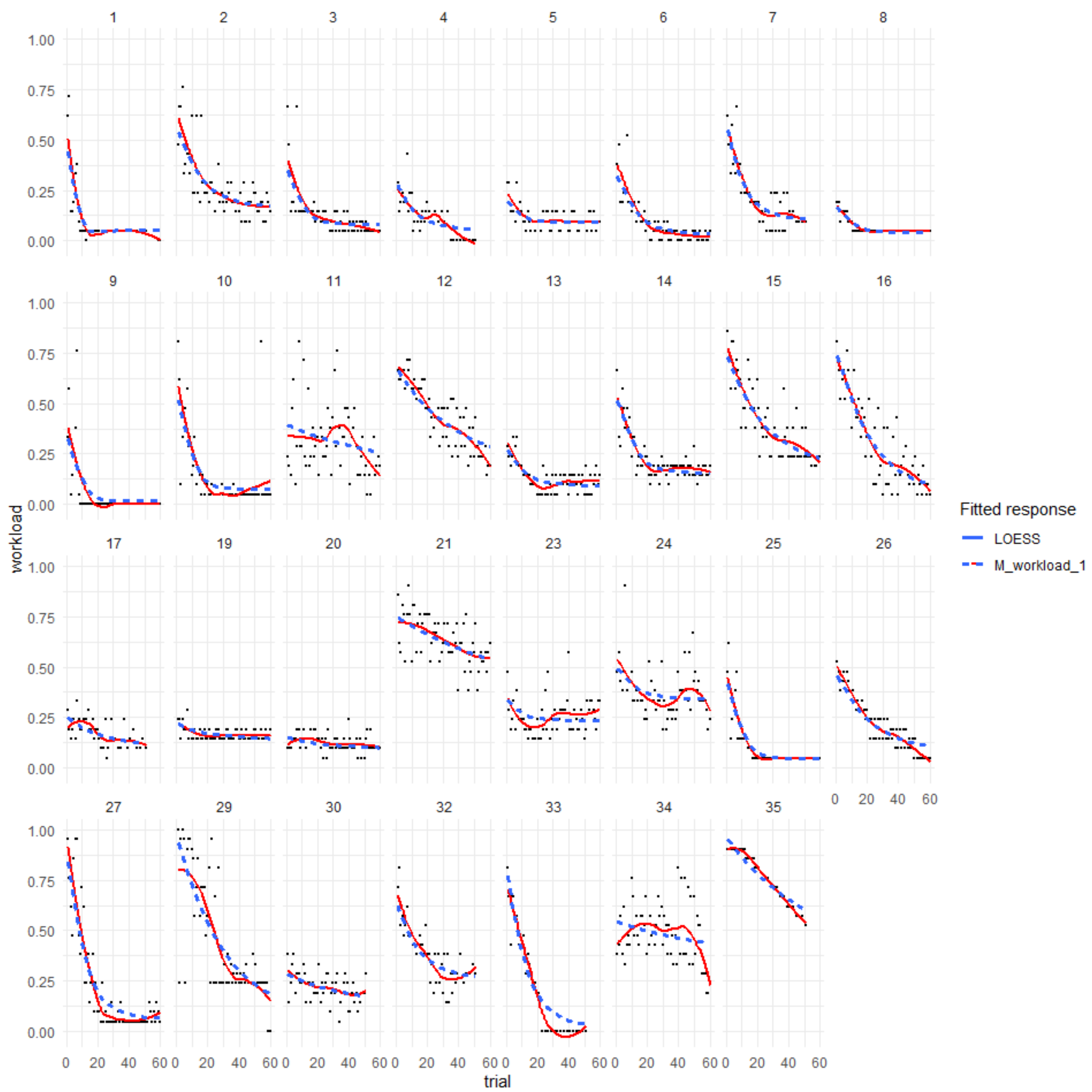

```
D_workload %>%
  ggplot(aes(x = trial, y = workload, color = hrs_gaming)) +
  facet_wrap(~Part, nrow = 4) +
  geom_point(size = .2) +
  geom_smooth(aes(y = M_workload_1), se = F) +
  scale_color_gradient(low = "blue", high = "red") +
  theme_minimal()
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

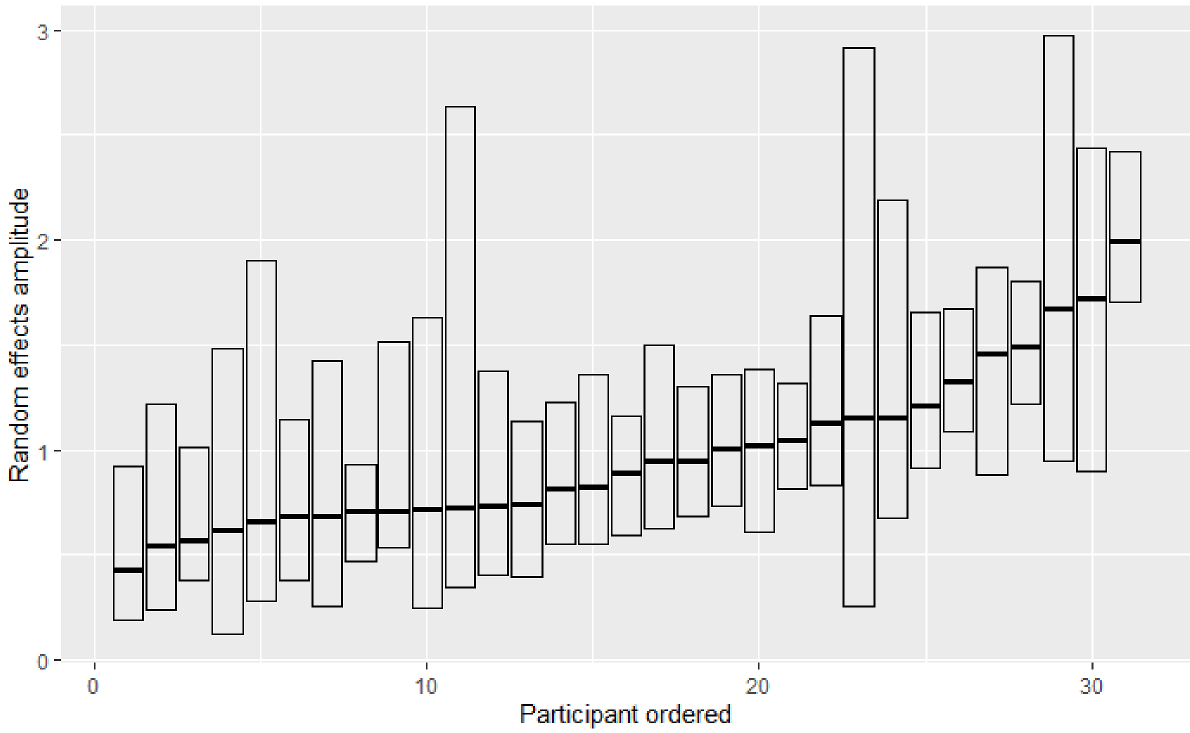



```
D_workload %>%
  ggplot(aes(x = trial, y = workload)) +
  geom_point(size = 0.5) +
  geom_smooth(aes(linetype = "LOESS"), se = F, color = "red", size = 0.75)
+
  geom_smooth(aes(y = M_workload_1, linetype = "M_workload_1"), se = F, siz
e = 1.2) +
  labs(linetype = "Fitted response", M_workload_1 = "Predicted", LOESS = "O
bserved") +
  facet_wrap(~Part, nrow = 4) +
  theme_minimal()
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
coef(M_workload_2, mean.func = exp) %>%
  filter(type == "ranef", nonlin == "ampl") %>%
  mutate(re_entity = rank(center)) %>%
  ggplot(aes(x = re_entity, ymin = lower, y = center, ymax = upper)) +
  geom_crossbar() +
  labs(y = "Random effects amplitude", x = "Participant ordered" )
```


3
2
1
0
0
10
20
30
Random effects amplitude
Participant ordered