Mindlessly Driving or Singing Along: Measuring Driving Performance in ACT-R

Bachelor's Project Thesis

Andreea Minculescu, s3932222, a.minculescu@student.rug.nl
Supervisors: Dr J.P. Borst and Moritz Held

Abstract: Driving is a complex and highly demanding task. Since multitasking typically leads to decreased performance of the main task, it is expected that any concurrent task would interfere with driving performance. Instead, a human study showed that paying full attention to the road while in a monotonous environment results in marginally worse driving performance than if one were driving and listening to the radio at the same time. One possible explanation is that a monotonous driving environment stimulates mind-wandering, which may be more demanding, from a cognitive standpoint, than a simple-enough secondary task. The present study tries to verify this hypothesis by augmenting an ACT-R model of human driving behaviour with i) mind-wandering behaviour and ii) a secondary task of listening to the radio. Results show that the listening model performed significantly better than the mind-wandering model, thus validating the findings of the human study. Overall, this study demonstrated how computational models can be used to provide insights into common misconceptions regarding in-vehicle device design and, more generally, driver safety.

1 Introduction

Driving has become a ubiquitous activity in modern society - for example, American drivers spend on average 87 minutes per day behind the wheel [Langer, 2005]. Driving is a highly demanding task, which requires dynamic execution of multiple concurrent sub-tasks. These sub-tasks range in difficulty from low-level steering and basic maneuvers such as lane changing or lane merging, to high-level decision-making and planning.

It is of no surprise, therefore, that a secondary task, such as a phone conversation [e.g. Strayer and Johnston, 2001] or listening to music [e.g. Brodsky, 2001], can cause distraction and negatively affect driving performance. However, other studies show evidence that multi-tasking can be beneficial under certain circumstances [e.g. Atchley and Chan, 2011] and this might be caused by a reduction of mind-wandering behaviour [e.g. Nijboer et al., 2016]. The current study investigates whether mind-wandering can explain this discrepancy in the literature, by building and evaluating a computational model of human driving behaviour.

1.1 Driving in a monotonous environment

Around 90%-95% of road accidents can be attributed to human error, and inattention to the road is one of the leading contributing factors [Treat et al., 1979]. This implies that road safety interventions should be focused more towards discouraging driver inattention. Instead, current countermeasures are more oriented towards road infrastructure improvement or laws and regulations [Eoh et al., 2005]. As a consequence, driving has rather been reduced to a simple lane-keeping task on highways and this led to a new type of contemporary car crashes, caused by lapse of vigilance due to monotony [Cerezuela et al., 2004].

It is possible that this lapse of vigilance manifests as a result of an increase of mind-wandering. It has been shown that people tend to mind-wander when they have enough cognitive resources to do so [Mooneyham and Schooler, 2013], and a boring driving environment, where there is little to no external engagement, is the perfect breeding ground for this phenomenon.

1.2 Detriments of mind-wandering

Mind-wandering is a mental activity characterized by decoupling of attention from the external environment and redirecting said attention to internal unrelated events [for a more detailed account, see Smallwood and Schooler, 2015]. Due to this decoupling from the external world, mind-wandering is usually accompanied by a decrease in performance of the task in progress. This is due to the fact that most activities occur in direct relation to the external environment and, by definition, mind-wandering obstructs that [Smallwood and Schooler, 2006].

Yanko and Spalek [2014] measured how mind-wandering affects driving performance by probing participants at random times during a simulated driving task, to indicate whether or not they were focused on the task at that moment. The authors noted that reports of mind-wandering were consistent with worse driving performance on multiple levels: participants exhibited longer reaction times to unexpected events, drove at a higher speed and maintained a shorter headway distance to the car in front.

Overall, mind-wandering constitutes a severe risk with regards to driver safety: since it is associated with reduced awareness of the external environment, reacting in time to obstacles and unexpected situations becomes more difficult.

1.3 Benefits of a secondary task

Two concurrent tasks requiring the same cognitive resource will contend for said resource and this, consequently, will lead to decreased performance for both tasks [Wickens, 2002]. Therefore, undertaking a secondary task while driving can have a negative impact on performance, especially when the task interference occurs at a perceptual level [e.g. Gherri and Eimer, 2011], motor level [e.g. Janssen et al., 2012] or working memory level [e.g. Strayer and Johnston, 2001].

On the other hand, it has been shown that a secondary task can lead to improved performance of the primary task. For example, Nijboer et al. [2016] conducted a study in which they tested the effects of four different secondary tasks on performance during a monotonous driving scenario. They found that participants' driving performance was overall better when the driving task was combined with a non-demanding secondary task, such as listening to the radio, as compared to the no-secondary-task baseline.

The authors hypothesize that such a secondary task reduces the chance of intrusion of another more demanding, involuntary phenomenon, such as mind-wandering. Ultimately, this does not mean that multi-tasking improves performance, per se, but rather that a simple enough secondary task is the smaller of two evils in terms of cognitive resource contention.

1.4 Computational driving model

The computational model proposed by Salvucci [2006] implements two basic vehicle control tasks: i) lateral control to maintain a central position with respect to the center of the current lane (i.e. steering) and ii) longitudinal control to maintain constant speed and appropriate distance from a lead vehicle (i.e. acceleration/breaking). The mathematics behind the corresponding control laws will be expanded upon in section 2.2.

The Salvucci driving environment is simple: a straight road with three lanes and a lead vehicle (see Fig 1.1). The lead vehicle maintains a constant speed and does not interfere with the driving task - one should rather think of it as a landmark in the distance. This driving scenario is consistent with the monotonous environment required as part of the current study.

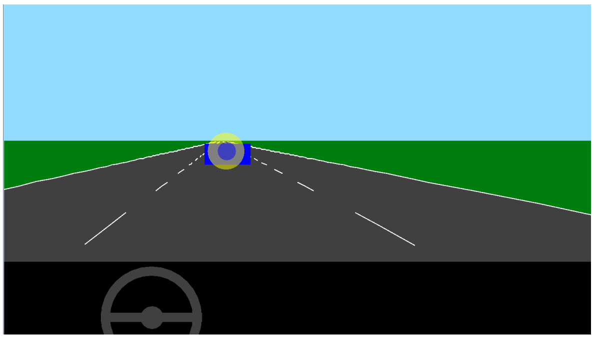

Figure 1.1: Driving environment of the Salvucci [2006] simulation. The blue rectangle represents the lead vehicle and the yellow circle represents the point of focus. The road is colored with gray and the three lanes are delimited by white striped lines. The perspective is that of the simulated car's driver.

1.5 Current study

1.5.1 Description

The present study aims to investigate the results found in the Nijboer et al. [2016] study, by building two cognitive computational models within the ACT-R architecture. The model proposed by Salvucci [2006] will be used as a baseline of human driving behaviour and augmented with i) mind-wandering behaviour and ii) a secondary task of listening to the radio. The models will be compared against each other and against the baseline driving model. The aim is to better understand, from a behavioural standpoint, if and how multi-tasking decreases the chance of mind-wandering and, consequently, how that leads to improved overall driving performance.

1.5.2 Motivation

The phenomenon found by Nijboer et al. [2016] is puzzling. Driving is a task with such high cognitive workload that one would expect any concurrent task to affect driving performance, since any task, large or small, involves some procedural steps [Salvucci and Taatgen, 2010]. However, it seems that, under certain circumstances, a secondary task can indirectly lead to improved driving performance, perhaps by obstructing the most intrusive aspects of mind-wandering.

The current study will contribute to a better understanding of driver behaviour and will serve to invalidate common misconceptions regarding in-vehicle device design and, more generally, driver safety.

2 Model

2.1 ACT-R architecture

The cognitive models of the present study were developed using the Adaptive Control of Thought (ACT-R) architecture [Anderson, 2007]. ACT-R is a general psychological theory and provides a simulation environment that can be used to develop cognitive models that adhere to this theory.

The architecture contains modules for cognitive processing and external world interaction. The procedural module is used to coordinate actions across different modules through *production rules*, based on an if-then approach. Particularly, *if* the current state of the modules matches the configuration specified on the left-hand side of the rule, *then* the respective production rule fires and the current state of the modules is modified as specified by the right-hand side configuration. A production rule is executed in 50 ms and multiple rules are executed serially.

Other modules of interest here are the declarative module (to store factual knowledge), the visual module (for visual perception), the manual module (for manual manipulation) and the aural module (for aural processing). Each module has a buffer associated with it. A buffer can be understood as the interface between the module and the procedural memory system. Each buffer can contain at most one chunk at a time, which are building structures that store the declarative knowledge of the system.

2.1.1 Threaded cognition

Threaded cognition Salvucci and Taatgen [2008] is an extension of the ACT-R architecture, which implements a theory of concurrent multitasking (i.e. performing multiple tasks at the same time). The underlying assumption of threaded cognition is that tasks can be represented as independent processing threads and coordinated by the procedural system. The theory allows for concurrent execution, resource retrieval and conflict resolution across all threads.

The theory specifies a mechanism to interleave the execution of production rules associated with different tasks. This mechanism is said to be "greedy" and "polite". It is *greedy* in the sense that threads request resources as soon as possible, when needed. If multiple rules are eligible at the same time, the one corresponding to "the most urgent task" is chosen first. The title of "most urgent" corresponds to the task which has had a rule chosen for execution least recently. The mechanism is *polite* in the sense that a thread releases resources for other threads as soon as possible after execution had been completed.

Overall, threaded cognition allows for concurrent task execution by having the procedural system send requests to various modules and supervising completion of requests. This implies that, as long as there is no resource interference, the threads complete their respective tasks independently. Otherwise, the greedy and polite resource usage assumption mediates conflicts. The theory of threaded cognition is useful in implementing the listening model, since task interplay between mind-wandering and listening to the radio is controlled by the theory.

2.2 Baseline model: driving

The basic driving model [Salvucci, 2006] implements the tasks of lateral and longitudinal control.

Lateral control involves a continuous process of surveying the road and adjusting the steering wheel in response, to maintain the car in a central position relative to the lane. Several research studies define driver vision of the road in terms of two regions (see Fig 2.1) a near point close to the front of the car, which provides information regarding the car's position in relation to the road (e.g. how close the car is to the center of the lane) and ii) a far point set at an arbitrarily distant point, which provides information regarding the curvature of the road [e.g. Donges, 1978]. The far point is characterized as: i) the vanishing point on a straight road, up to a maximum distance equivalent to 2 seconds of headway, ii) the tangent point to a curved road or iii) a lead vehicle [Salvucci, 2006]. In the current driving scenario, the simulation uses only the third option.

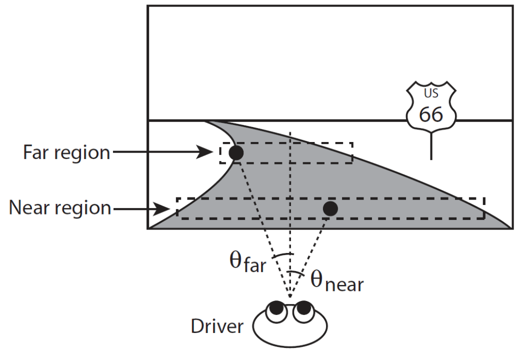

Figure 2.1: Division of driver visual field into a near and a far region. The black dots in the near and far regions represent the near point and the far point, respectively. Diagram taken from Salvucci and Taatgen [2010].

Using this definition of the visual field of a driver, Salvucci and Gray [2004] defined a mathematical steering-control law. Let θ_{near} be the horizontal angle to the near point with respect to the driver's direction of travel and $\Delta\theta_{near}$ the change in θ_{near} compared to the last control check. Assume θ_{far} and $\Delta\theta_{far}$ to be the equivalent for the far point. According to the control law, after an arbitrarily short period of time Δt, the steering angle φ is modified by $\Delta\varphi$, as specified by the equation:

$$\Delta\varphi = k_{far}\Delta\theta_{far} + k_{near}\Delta\theta_{near} + k_l\theta_{near}\Delta t \quad (2.1)$$

where k_{far}, k_{near} and k_l are scaling constants. The control-law imposes three constraints, such that the far point is stable (i.e. $\Delta\theta_{far} \to 0$), the near point is stable (i.e. $\Delta\theta_{near} \to 0$) and the near point is at the center of the lane (i.e. $\theta_{near} \to 0$).

Longitudinal control involves a process of adjusting the accelerator/brake depression in accordance to the time-headway to the far point (in this case, the lead vehicle). This adjustment is computed by considering the current time-headway (let us define it as thw_{car}), the change of time-headway as compared to the last control check (i.e. Δthw_{car}) and some arbitrary desired time-headway (i.e. thw_{follow}). Thus, the model computes the update value of the car's acceleration (ψ), over an arbitrarily short period of time (Δt) as follows:

$$\Delta\psi = k_{car}\Delta thw_{car} + k_{follow}(thw_{car} - thw_{follow})\Delta t \quad (2.2)$$

where k_{car} and k_{follow} are scaling constants. Note that the longitudinal control law is an extension of the lateral control law. Similarly, the longitudinal control law imposes two constraints, such that the car maintains a steady time-headway (i.e. $\Delta thw_{car} \to 0$) and the current time-headway approaches the desired time-headway value (i.e. $thw_{car} - thw_{follow} \to 0$).

Overall, the driving task is embedded in ACT-R as an iterative loop of four procedural steps:

1. Determine the horizontal visual angle to the near point (θ_{near})
2. Determine the horizontal visual angle to the far point (θ_{far})*
3. Adjust steering angle (φ) and acceleration (ψ) based on their respective control laws
4. Check vehicle stability† and go back to the first step

The first two steps require visual processing and the third step requires the employ of the motor module. Considering the fixed execution time of 50 ms of an ACT-R production rule [Anderson, 2007], one iteration of the process runs in 200 ms.

2.3 Mind-wandering model

The model described by van Vugt et al. [2015] was used as inspiration for this study's mind-wandering model. Fig 2.2 demonstrates the embedding of the mind-wandering model into the Salvucci [2006] driving model.

At any time, the model can be in one of two possible goal states: "attend" or "wander". These goal states are encoded as chunks in the declarative memory, following the structure `(chunk-type goal state)`. If the model is in the attending state, then the process described in section 2.2 ensues. Otherwise, the model is said to be in the mind-wandering state.

The mind-wandering process consists of continuous retrieval of memory chunks from the declarative module. A memory chunk is of the form `(chunk-type memory awareness)`, where the "awareness" slot serves as an identifier. The model contains 30 memory chunks of type "wander" awareness and 1 chunk of type "attend" awareness with the same baseline activation level, following the approach described in Moye and van Vugt [2019]. As a result, the model has a 30/31 chance to mind-wander. If a memory chunk of type "wander" awareness is retrieved, then the model will retrieve another memory chunk, entering a continuous loop of memory retrievals, similar to the concept of a "train of thoughts". This loop, however, can be interrupted in two distinct cases.

Firstly, the loop can be interrupted by the retrieval of a memory chunk of type "attend", in which case the model switches to the attending state. This is equivalent to a person remembering that there is a main task to be done (driving in this case).

Secondly, the loop can be interrupted by surpassing a threshold value of maximum lane deviation. The model makes the assumption that any driver, experienced or not, would regain focus to the main driving task once the car deviates by a large enough

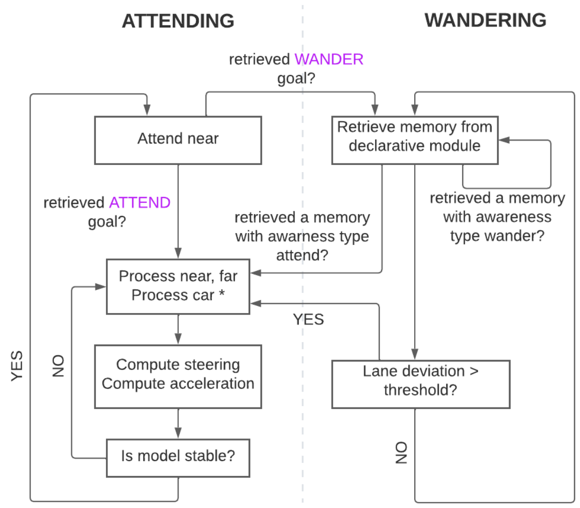

Figure 2.2: The implementation of the mind-wandering driving model. The diagram should be read starting from the upper left corner, (i.e. attend near) *This production rule fires only when the location of the lead vehicle is in the visual-location buffer.

*Note that the current time-headway (i.e. thw_{car}), necessary for the longitudinal control law computation, can be obtained if one knows the far point.

†Defined as a situation in which the velocity of the two visual angles and the near visual angle are all below some preset thresholds. The thresholds were defined experimentally.

factor from the middle of the lane. In this driving scenario, the large-enough-factor is assumed to be the event of crossing over into a different lane. The assumption is justified by the fact that highways often have rumble strips applied on the lane separators, which cause vibration and audible rumbling to be sent through the wheels to the vehicle interior. Since these have been shown to consistently prevent run-off-the-road crashes [see e.g. Khan et al., 2015], the model assumes that a rumble strip provides intrusive enough behavior that it should override mind-wandering inattention and instead return driver attention back to the task. Based on observed behavior within the simulation, the threshold value was set to 1 foot of lane deviation towards left or right, from the middle of the current lane.

Overall, note that the model supports a wide variety of driver behavior. This is due to the fact that entering the mind-wandering process has a 50% chance (i.e. the model can either retrieve the "attend" or "wander" goal chunk, with relatively equal chance) and the duration of the mind-wandering episode varies, depending on the 3.33% retrieval chance of the memory of "attend" awareness. Thus, the model is a comprehensive representation of mind-wandering effects on driving performance.

2.4 Listening model

The listening model simulates behaviour akin to listening to the radio while driving. For this, an audio event is periodically injected into the simulation. An audio event encodes characteristics of a sound, such as duration, content and location. In the current study, audio events are spaced 300ms apart (resulting in a speech rate of approximately 200 words per minute) as an approximation of human speech rates, which can range from 120-150 words per minute, during typical conversations, to 250-300 words per minute during animated conversations [Ray and Zahn, 1990]. All audio events encode words, which represent lyrics of a song that the model "hears" on the radio. Once an audio event is perceived by the model, two production rules can fire: i) attend sound or ii) access meaning. The implementation of these production rules was based on the methodology described in Borst et al. [2010].

The *attend sound* production rule is only a preparatory step. It matches a set of constraints against the audio module (in this case, it checks if there is an audio event in the audio-location buffer), requests an attention shift to the audio event and places a chunk in the aural buffer. Overall, this process is akin to realising that there is a sound present in the environment and it lasts for 50ms.

The *access meaning* production rule processes the audio event chunk in the aural buffer and retrieves the information in the content slot. Next, the model retrieves one chunk from the declarative memory of form `(chunk-type word spelling)`, whose spelling slot matches with the information from the content slot of the audio event. Note that the retrieval will always be successful, since the audio event content and the word chunks are both hardcoded (based on the assumption that humans would not encounter difficulties in understanding lyrics of songs). Overall, this process is akin to retrieving the meaning of a word from memory and it lasts for 550ms (i.e. the duration of a chunk retrieval from declarative memory).

The listening production rules were built upon the mind-wandering model (see section 2.3) and the order of firing of all production rules is regulated by threaded cognition (see section 2.1.1). Note that the listening production rules create a bottle-neck at the level of memory chunk retrieval of the mind-wandering model. This is caused by the fact that there is resource contention for declarative chunks (memory chunks in the case of the mind-wandering model and word chunks in the case of the listening model). As a reminder, according to threaded cognition conflict resolution implementation, if two production rules corresponding to two different tasks match at a point in time, then the rule corresponding to the task used the longest time ago fires first. Thus, it is expected for the wandering loop to be interrupted by a listening production rule whenever an audio event occurs in the environment.

Since I did not have a clear hypothesis regarding the extent to which participants in the human study processed sound, assumptions regarding the degree of information processing of audio events had to be made. Accordingly, the listening model was divided into four sub-models. An overview is presented in Table 2.1, where the "attending" and "wandering" labels correspond to the states in Fig. 2.2.

Model	*Attending*		*Wandering*	
	Attend sound	Access meaning	Attend sound	Access meaning
Listen 1	X		X	
Listen 2	X	X	X	X
Listen 3			X	X
Listen 4	X		X	X

Table 2.1: Degree of processing audio event information across 4 different listening sub-models. "Attend sound" is considered light processing (awareness of sound), while "access meaning" is considered heavy processing (retrieval of word meaning). Note that accessing meaning cannot occur without attending sound but the reverse can be the case.

Listen 1 involves light processing of information, as it only attends to the sound. This is akin to listening to a song in a foreign language - one is aware of the presence of sound but cannot understand the meaning of the words.

Listen 2 is on the opposite side of the spectrum from the Listen 1, since it involves heavy processing. This is akin to intently listening to a piece of news on the radio.

Listen 3 involves heavy processing only during the wandering state. This model assumes that driving is such an intense activity that the driver would completely "phase out" all sound.

Listen 4 involves heavy processing during the wandering state and light processing during the attending state. This model assumes that driving is an intense enough activity that focus would be allocated mostly to it, but the driver would still be partially aware of the presence of sound in the environment.

Note that the four listening sub-models account for fundamentally different types of dual-tasking integration behaviour and generate considerably and consistently different results.

3 Results

Model results were generated and compared against the original human study by Nijboer et al. [2016]. The data was generated by running the ACT-R simulation 24 times (consistent with the human study) for the mind-wandering and listening sub-models and recording the state of the simulated car once every 50ms. The plots were generated in python (version 3.7.3), using the matplotlib‡ library. All statistical analyses were performed in R (version 4.0.2).

3.1 General behaviour

Fig. 3.2 shows measurements of one random simulation run of the mind-wandering model. These measurements are juxtaposed against Fig. 3.1, which shows measurements of one random human participant of the Nijboer et al. [2016] study in the single task condition. The figures are comparable against each other because mind-wandering is involuntary and unavoidable, which suggests that a human would inevitably mind-wander while driving in the single task condition.

Both plots investigate the same indices of driving performance: i) lane deviation, which indicates the distance from the car to the center of the current driving lane and ii) steering angle, which records data from the steering wheel. While the human study also investigated driving speed, in the Salvucci [2006] driving model the speed control mechanism did not seem to be sensitive enough to show any visible deviation as a result of mind-wandering. For that reason, driving speed analysis was excluded from the current study.

‡ `https://matplotlib.org/`

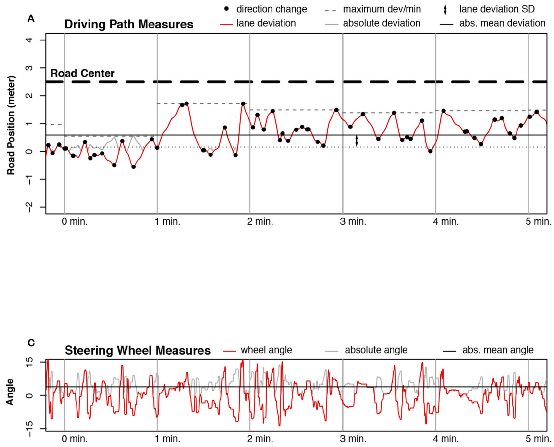

Figure 3.1: Measurements taken from one random human participant. The red line (raw data) is the main focus in the context of the present study. Source: Nijboer et al. [2016]

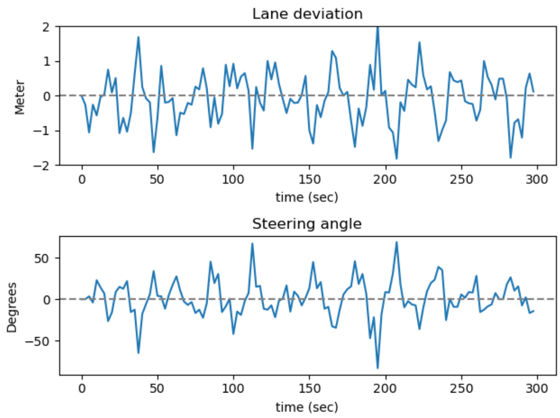

Figure 3.2: Measurements taken from one random mind-wandering model run.

Both figures show irregular behaviour and, while any strong conclusion related to the compatibility of the model to the human data is unreliable, one can claim that the general trend of the model is humanly plausible. This is further reinforced by the fact that the scales of the measurements are somewhat consistent across the two studies.

3.2 Mind-wandering behaviour

The main interest of the study was to investigate how the mind-wandering behaviour is influenced by the listening task. Fig. 3.3 shows the distribution of the count of mind-wandering occurrences, where an "occurrence" is defined as one retrieval of a memory chunk (in the context of this paper). The graph was obtained by counting the number of occurrences for each of the 24 runs and plotting the distribution of the obtained 24 data points - the same procedure was used for each model. Fig. 3.4 shows the distribution of the average duration of mind-wandering episodes, where an "episode" is defined as multiple consecutive occurrences from start until interruption (due to lane deviation threshold, remember-to-attend chunk retrieval or audio event interference). The data was obtained by counting the number of mind-wandering occurrences within each episode and taking the mean of the counts. Thus, the graph was obtained by repeating the operation for all 24 runs of each model and plotting the distribution of the obtained 24 data points. Note that in both plots, all values of the baseline (i.e. original Salvucci [2006] driving model) distributions are 0, since no mind-wandering occurs.

From both plots, it is clear that the mind-wandering model wanders most and has the most diverse behaviour. This is to be expected because the listening sub-models implement an extra condition in which the wandering episode is interrupted (through contention for the declarative module). This a first confirmation that the hypothesis proposed by Nijboer et al. [2016] could make sense.

Fig. 3.3 suggests that there is an inverse relationship between the level of audio event processing and the number of mind-wandering occurrences: the heavier the processing, the lower the number of mind-wandering episodes. This can be observed by comparing the distributions of the Listen 1 model (which only attends to sound, hence the lightest kind of processing out of all models) and Listen 2 model (which retrieves the meaning of each word, hence the heaviest kind of processing out of all models).

The aforementioned phenomenon can be explained by the fact that accessing meaning takes considerably longer than simply attending sound (550ms vs 50ms) and, overall, there is less time for mind-wandering. Moreover, Fig. 3.4 suggests that this difference in number of occurrences is due to the frequency of mind-wandering, rather than duration of episodes, since the durations are similar across different listening sub-models (except for Listen 3, see further). This is because audio events occur at the same intervals across all listening sub-models and, therefore, it stands to reason that wandering interruption patterns would be consistent across models.

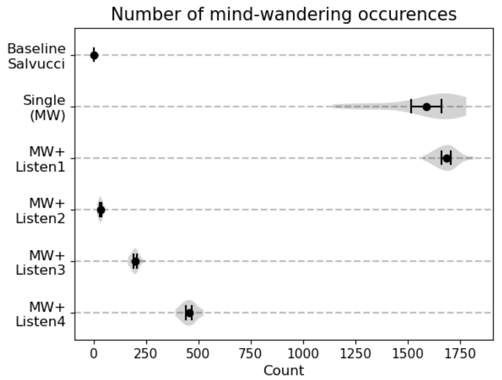

Figure 3.3: Distribution of total number of mind-wandering occurrences for each of the models. "Single" is the mind-wandering (MW) model and the order of the listening sub-models corresponds to table 2.1. Black dots represent the mean across all model runs and bars denote 95% CI.

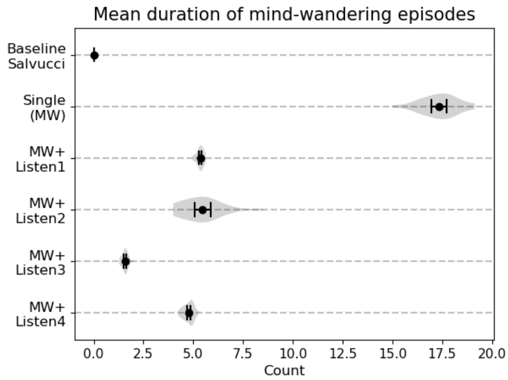

Figure 3.4: Distribution of average duration of mind-wandering episodes for each of the models. "Single" is the mind-wandering (MW) model and the order of the listening sub-models corresponds to table 2.1. Black dots represent the mean across all model runs and bars denote 95% CI.

According to the observations made in the current section, one would expect Listen 3 to have a higher occurrence count distribution than Listen 4 and the distribution of episode duration is significantly different from the other three models. However, this does not seem to be the case. One hypothesis is that this is caused by the decay time of the audio event. Sounds are available in the environment for a short amount of time, after which they decay and are not accessible anymore. One important side effect of processing audio information in the attending state is that the audio event is essentially "cleared out" of the environment before the mind-wandering episode begins. However, if no listening production rules fire during the attending state, as is the case for only Listen 3, the audio event still may linger in the environment at the beginning of the mind-wandering episode (supposing it had not decayed by that point). As a result, due to this still lingering audio event, the listening production rules of Listen 3 will fire and interrupt the episode immediately, as soon as the mind-wandering episode begins. Overall, this would result in fewer mind-wandering occurrences and shorter mind-wandering episodes, as can be seen in Fig. 3.3 and 3.4, respectively.

3.3 Human-model comparison

Driving performance cannot be measured directly, but rather through indices of performance. Two such indices are lane deviation and steering wheel angle (also discussed in section 3.1). The averaged model data is shown in Fig. 3.6 and 3.8. It is relevant to compare the trends in the data generated by the models against those generated by human drivers. Generally, values close to 0 indicate good driving performance, since it means that the car maintains a constant position close to the middle of the current driving lane. The baseline model is an indicator of almost "perfect" driving behaviour and all other models are expected to perform worse (i.e. the underlying distribution should contain higher values).

Fig. 3.5 plots the distribution of mean lane deviation for the 24 human participants in the single and listening conditions. While the distributions of the conditions overlap considerably, it is clear that the mean of the listening sample is marginally smaller than the one for the single condition.

Comparing these human results against the results of the model (Fig. 3.6), one can see that the same trend is mostly preserved: listening submodels are generally lower in value than the mind-wandering model (note that the data labeled under the "single" condition is both figures should be compared against each other). The mean lane deviation distributions differ between models: the heavier the processing of an audio event, the higher the values and the more diverse the behaviour. This can be seen by comparing the distribution of Listen 1 (light processing) against the distribution of Listen 2 (heavy processing).

This phenomenon indicates that high levels of information processing produce disrupted driving behaviour. This is to be expected both from a i) cognitive standpoint - the more cognitive resources allocated to audio processing the fewer allocated to driving - and ii) design standpoint - accessing meaning takes 550ms, during which no driving occurs. The authors of the original human study also investigated the standard deviation and maximum lane deviation of the human participants. Refer the Appendix in order to note similar trends to those discussed above.

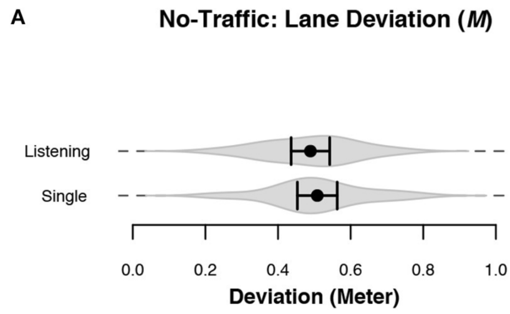

Figure 3.5: Mean lane deviation of human participants. Black dots represent the mean across subject and bars denote 95% CI. Source: Nijboer et al. [2016].

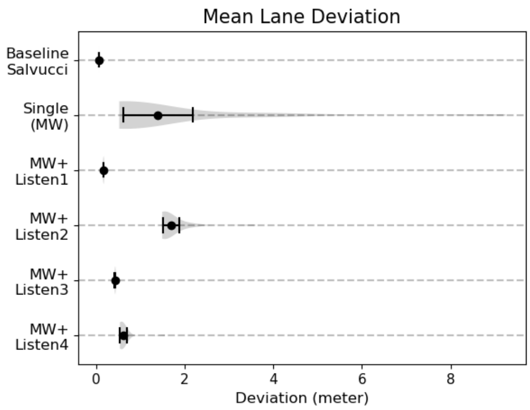

Figure 3.6: Mean lane deviation of model runs. "Baseline" is the original Salvucci [2006] driving model, "single" is the mind-wandering (MW) model and the order of the listening sub-models corresponds to table 2.1. Black dots represent the mean across subject and bars denote 95% CI.

Fig. 3.7 and 3.8, illustrating steering wheel behaviour, show similar trends: the listening condition demonstrates largely better performance than the single condition and there seems to be a direct relationship between the level of audio processing and level of disrupted driving behaviour.

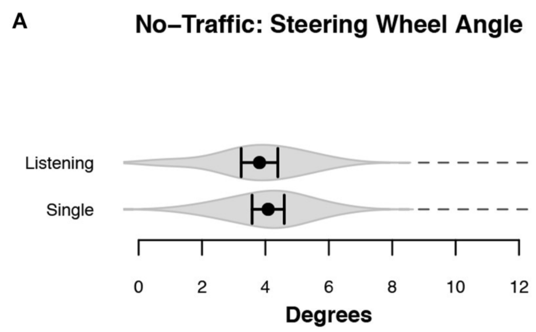

Figure 3.7: Mean steering angle of human participants. Black dots represent the mean across subject and bars denote 95% CI. Source: Nijboer et al. [2016].

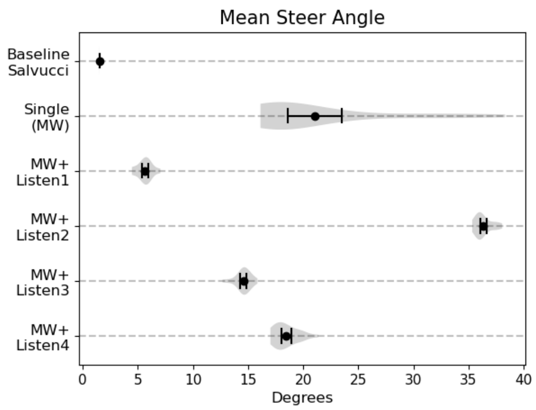

Figure 3.8: Mean steering angle of model runs. "Baseline" is the original Salvucci [2006] driving model, "single" is the mind-wandering (MW) model and the order of the listening sub-models corresponds to table 2.1. Black dots represent the mean across subject and bars denote 95% CI.

Another possible driving performance index is the number of changes in car heading. Fig. 3.9 and 3.10 show that the listening task improves consistency. However, the same trend between the level of audio processing and the level of disrupted driving behaviour does not seem to hold anymore. Instead, higher processing levels seem to be associated with more consistent behaviour. This can be explained by the fact that, during the 550ms of access meaning processing there is no driving and, overall, there is less time for swerving. Another interesting observation is the fact that the baseline model performs worst out of all models. This, again, has a design-related explanation: throughout the entire simulation time only driving ensues and, overall, there is more time for swerving.

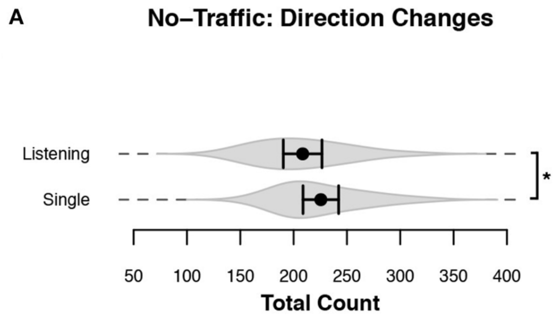

Figure 3.9: Count of direction changes of human participants. Black dots represent the mean across subject and bars denote 95% CI. Source: Nijboer et al. [2016].

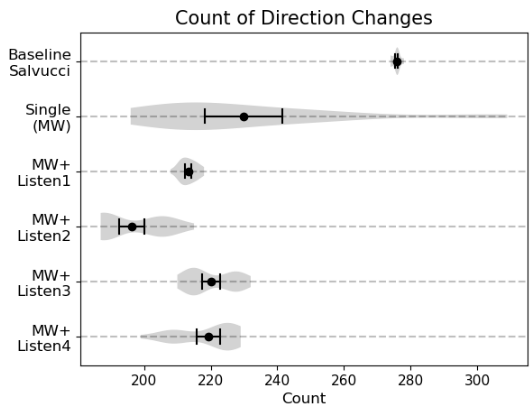

Figure 3.10: Count of direction changes of model runs. "Baseline" is the original Salvucci [2006] driving model, "single" is the mind-wandering (MW) model and the order of the listening sub-models corresponds to table 2.1. Black dots represent the mean across subject and bars denote 95% CI.

Overall, all metrics point towards one conclusion: a purely qualitative analysis suggests that driving performance is mostly improved by the listening dual task, as compared to the single task condition (i.e. purely mind-wandering). This is consistent with the findings in Nijboer et al. [2016]. Additionally, the level of audio information processing influences driving performance, since Listen 2 (the heaviest processing model) performs consistently worse than all other models.

3.4 Statistical analysis

3.4.1 Most realistic listening sub-model

The four listening sub-models are based on different assumptions regarding the extent to which drivers process audio information. Therefore, it is interesting to formally determine which of the four models is most realistic. Here, the term “realistic” is defined as *the most similar to human data* [i.e. the data form the Nijboer et al., 2016, study].

Thus, the root mean square deviation (RMSD) for all indices of performance of each of the four listening sub-models were computed and used to select the most realistic model (see Appendix for exact values). RMSD is used as a measure of similarity between the predictions of a model and the actual observations.

In order to compare the human data against the listening sub-models, the following formula was applied:

$$RMSD = \sqrt{\begin{matrix}(\mu_{modelSingle} - \mu_{humanSingle})^2 + \\ (\mu_{modelListen} - \mu_{humanListen})^2\end{matrix}} \quad (3.1)$$

where μ_X represents the mean of the X distribution. *humanSingle* corresponds to the single condition of the human study and *humanListen* corresponds to the listening condition of the human study. *modelSingle* and *modelListen* are the equivalent for the respective listening sub-model.

The above equation was applied to all indices of driving performance (the term $\mu_{modelListen}$ in the equation) and, thus, for each of the four listening sub-models five different RMSD values were obtained. Next, a scoring system was created to select the best model: for each index of performance the smallest value is identified (through the nature of the RMSD definition, smaller values indicate higher similarity to the human data) and the corresponding model received one point; the model with the most points was considered the best. Thus, the **Listen 1 model** seems to be most alike human data. Since Listen 1 is the least processing intensive, this might suggest that human drivers tend to allocate few cognitive resources to processing the audio information coming from the radio while driving.

Note that the scoring system used to select the best model attributes the same level of importance to the all indices of performance and is not sensitive to the magnitude of the RMSD values. However, other scoring systems may also be appropriate for the task: for example, one could compute the average of the five performance index RMSD values and choose the model with the smallest average value (interestingly, in the present study, using this scoring system would yield the same result). Moreover, some could argue that, since three of the performance indices compute lane deviation in some form, attributing the same level of importance to all indices may be undesirable. However, since the authors of the human study also attributed the same level of importance to all indices of performance, the present study follows the same design, for consistency.

3.4.2 Statistical comparison between mind-wandering and listening models

The figures in section 3.3 qualitatively show evidence that listening improves driving performance. However, one should wonder if this difference in performance is statistically significant.

The statistical analysis consisted of performing one-to-one comparisons of all performance indices of the mind-wandering model data against all performance indices of the best listening sub-model data (i.e. Listen 1 model). Thus, five Wilcoxon signed-rank test were performed, using the R supported function[§]. The Wilcoxon signed-rank test is the non-parametric equivalent of a paired Student’s t-test, meaning that it is commonly used to determine whether the means of two sets of data are significantly different from each other. The Wilcoxon signed-rank test was used because the data does not follow a normal distribution (as revealed by applying Shapiro-Wilk normality tests).

The Wilcoxon tests revealed that the difference in means between the mind-wandering and best listening sub-models is statistically significant ($p <$

[§] `https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/wilcox.test`

0.05 for all indices of performance; for exact values refer to the Appendix). This shows that the data associated with the listening condition is significantly lower than the data associated with the mind-wandering condition for all indices of driving performance. Since low values are indicative of better driving performance, this is strong evidence that listening to the radio improves driving performance.

4 Discussion

Driving is a highly demanding task from a cognitive standpoint, since it involves the interplay of multiple concurrent sub-tasks. In this context, one would expect that undergoing a secondary task while driving would have a negative impact on performance. Instead, Nijboer et al. [2016] found the opposite effect in their study. The authors hypothesized that a simple enough secondary task (such as listening to the radio) counteracts the intrusion of mind-wandering, a more demanding and involuntary process.

The current study investigated the phenomenon found by Nijboer et al. [2016] by building and analysing two ACT-R computational models. Both models augmented an already existing model of human driving behaviour proposed by Salvucci [2006]. The first model implemented mind-wandering behaviour as a process of continuous retrieval of memory chunks. The second model further augmented the mind-wandering model and was divided into four sub-models, which implemented different levels of audio information processing. A qualitative analysis of a random, but representative, model run demonstrated that the mind-wandering model implements realistic behaviour. Further statistical analysis of driving performance indices revealed that i) the least audio information processing intensive listening sub-model behaves most similarly to the human data and ii) the best listening sub-model performs better on the driving task than the mind-wandering model.

Overall, the present study provides strong evidence in favour of the idea that listening to the radio while driving in a monotonous environment has a positive effect on driving performance. Since the single condition model involved mind-wandering behaviour and the results of the study point towards reasonable behaviour, this suggests that mind-wandering is a plausible explanation for the phenomenon observed by Nijboer et al. [2016]. Moreover, the present study provided insight into how audio information processing influences driving performance: the more intensive the processing level, the worse the performance on the driving task.

4.1 Model evaluation

The mechanisms embedded into the models built as part of the current study are based on assumptions and simplifications of reality. As a result, the more realistic the models are, from a biological standpoint, the more useful and reliable the results of the study. The models of the current study are grounded in past research and reasonable assumptions of biological processes [see e.g. Salvucci and Taatgen, 2010], thus they can claim good biological plausibility. However, points of improvement with regards to biological plausibility can be identified.

The driving model only implements basic vehicle-control behaviour. Thus, it lacks an extra level of metacognition, defined as the ability to reason about one's own performance and adjust one's own behaviour appropriately. In reality, drivers have been shown to be aware that their own attention is impaired while performing a dual task and allow themselves a larger margin of error. This type of adaptation of driving behaviour in distracted situations is important if one wishes to develop a realistic model of human driving behaviour [Salvucci and Taatgen, 2010].

The mind-wandering model is based on a process of continuous retrieval of memory chunks. The present model considers that all memory chunks have the same activation baseline and the same chance of being retrieved at a time. However, a more biologically plausible approach would be to implement spreading activation between memories, ensuring that memories of the same valence (positive, negative or neutral) will be retrieved in a sequence [van Vugt et al., 2012].

The listening model was divided into four sub-models based on assumptions made regarding the level of audio information processing performed by a human driver. Although reasonable, these assumptions were not based on any former research but rather on understanding of phenomena and intrinsic human intuition. More specifically, the listening sub-model assumptions were gradually formulated in order to better match the human study behaviour. Thus, it is unclear if Listen 1, deemed the most realistic out of all four sub-models, is also the most realistic sub-model overall or if different assumptions would match the human data even better.

The main strengths of the models lie in generality and simplicity. Firstly, both models implement diverse human behaviour, since the exit conditions of the mind-wandering model can fire at diverse moments in time and in diverse situations. Secondly, both models implement simplified versions of the underlying biological mechanisms and no rules were introduced unless necessary from a cognitive standpoint.

4.2 Further research

The present study found statistical significance in terms of the difference in behaviour between the mind-wandering and listening models, for all indices of performance. However, this result is inconsistent with the human study, where the authors found evidence of statistical significance for only one of the indices. This does not necessarily mean that the studies generated contradictory results. It could be the case that a larger sample of human drivers would reveal statistical significance - in the end, human behaviour is more diverse than the behaviour of a computational model. Thus, further investigation would be required in order to determine the cause of this discrepancy.

The present study implements a simulation in a boring environment. However, it is unclear what a "boring" environment entails: it could be considered as *easy* (as is the case in this study), *familiar* to the driver or *non-stimulating* in terms of the landscape. For example, one might study the reaction times of drivers stuck in a traffic jam, in the same single and dual conditions. It could be that listening to the radio would not have the same effect on driving performance anymore since driving might not be perceived as the main task anymore - in the end, it would mainly involve waiting in line.

The present study does not investigate driving performance directly, but indirectly through indices of performance. Perhaps other indices of performance might be more appropriate, depending on which aspect of driving one wishes to investigate. While the present study investigates something akin to driving accuracy and stability, it could be that investigating reaction times in life-threatening situations or brain imaging techniques would reveal interesting patterns, which might contradict or further reinforce the present study. Approaching an issue from different angles gives a more thorough narrative of underlying mechanisms.

4.3 Bigger picture

There are many misconceptions regarding secondary task interference and the appropriate safety regulations. For example, although it has been shown that there is almost no difference in terms of driving performance between handheld and hands-free phone use [Horrey and Wickens, 2006], efforts are still allocated to designing in-vehicle devices for phone mounting purposes. This is further reinforced by the Listen 2 model, which suggests that paying full attention to audio information is associated with worse driving performance. From this it follows that a hands-free phone conversation, a hand-held phone conversation or an in-vehicle conversation would not make a difference in terms of performance on the driving task. In fact, this is exactly the pattern displayed by human participants in a study by Collet et al. [2009], where the authors measured electrodermal activity, heart rate and reaction times.

One hope is that cognitive models, such as those built as part of this study, can be used to better pinpoint the true source of secondary task interference while driving and to test solutions.

References

John R Anderson. *How can the human mind occur in the physical universe?* Oxford University Press, 2007.

Paul Atchley and Mark Chan. Potential benefits and costs of concurrent task engagement to maintain vigilance: A driving simulator investigation. *Human factors*, 53(1):3–12, 2011.

Jelmer P Borst, Niels A Taatgen, and Hedderik Van Rijn. The problem state: a cognitive bottleneck in multitasking. *Journal of Experimental Psychology: Learning, memory, and cognition*, 36 (2):363, 2010.

Warren Brodsky. The effects of music tempo on simulated driving performance and vehicular control. *Transportation research part F: traffic psychology and behaviour*, 4(4):219–241, 2001.

Gemma Pastor Cerezuela, Pilar Tejero, Mariano Chóliz, Mauricio Chisvert, and M José Monteagudo. Wertheim's hypothesis on 'highway hypnosis': empirical evidence from a study on motorway and conventional road driving. *Accident Analysis & Prevention*, 36(6):1045–1054, 2004.

C Collet, A Clarion, M Morel, A Chapon, and C Petit. Physiological and behavioural changes associated to the management of secondary tasks while driving. *Applied ergonomics*, 40(6):1041–1046, 2009.

Edmund Donges. A two-level model of driver steering behavior. *Human factors*, 20(6):691–707, 1978.

Hong J Eoh, Min K Chung, and Seong-Han Kim. Electroencephalographic study of drowsiness in simulated driving with sleep deprivation. *International Journal of Industrial Ergonomics*, 35(4): 307–320, 2005.

Elena Gherri and Martin Eimer. Active listening impairs visual perception and selectivity: an erp study of auditory dual-task costs on visual attention. *Journal of cognitive neuroscience*, 23(4): 832–844, 2011.

William J Horrey and Christopher D Wickens. Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human factors*, 48(1):196–205, 2006.

Christian P Janssen, Duncan P Brumby, and Rae Garnett. Natural break points: The influence of priorities and cognitive and motor cues on dual-task interleaving. *Journal of Cognitive Engineering and Decision Making*, 6(1):5–29, 2012.

Mubassira Khan, Ahmed Abdel-Rahim, and Christopher J Williams. Potential crash reduction benefits of shoulder rumble strips in two-lane rural highways. *Accident Analysis & Prevention*, 75:35–42, 2015.

Gary Langer. Abc news poll: Traffic in the united states. *ABC News*, 2005.

Benjamin W Mooneyham and Jonathan W Schooler. The costs and benefits of mind-wandering: a review. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 67(1):11, 2013.

Amir J Moye and Marieke K van Vugt. A computational model of focused attention meditation and its transfer to a sustained attention task. *IEEE Transactions on Affective Computing*, 12(2):329–339, 2019.

Menno Nijboer, Jelmer P Borst, Hedderik van Rijn, and Niels A Taatgen. Driving and multitasking: the good, the bad, and the dangerous. *Frontiers in psychology*, 7:1718, 2016.

George B Ray and Christopher J Zahn. Regional speech rates in the united states: A preliminary analysis. *Communication Research Reports*, 7(1): 34–37, 1990.

Dario D Salvucci. Modeling driver behavior in a cognitive architecture. *Human factors*, 48(2): 362–380, 2006.

Dario D Salvucci and Rob Gray. A two-point visual control model of steering. *Perception*, 33 (10):1233–1248, 2004.

Dario D Salvucci and Niels A Taatgen. Threaded cognition: an integrated theory of concurrent multitasking. *Psychological review*, 115(1):101, 2008.

Dario D Salvucci and Niels A Taatgen. *The multitasking mind.* Oxford University Press, 2010.

Jonathan Smallwood and Jonathan W Schooler. The restless mind. *Psychological bulletin*, 132(6): 946, 2006.

Jonathan Smallwood and Jonathan W Schooler. The science of mind wandering: empirically navigating the stream of consciousness. *Annual review of psychology*, 66:487–518, 2015.

David L Strayer and William A Johnston. Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological science*, 12(6):462–466, 2001.

John R Treat, Nicholas S Tumbas, Stephen T McDonald, David Shinar, Rex D Hume, RE Mayer, RL Stansifer, and N John Castellan. Tri-level study of the causes of traffic accidents: final report. executive summary. Technical report, Indiana University, Bloomington, Institute for Research in Public Safety, 1979.

Marieke K van Vugt, Niels A Taatgen, Jérôme Sackur, and Mikaël Bastian. Modeling mind-wandering: a tool to better understand distraction. In *Proceedings of the 13th international conference on cognitive modeling*, pages 252–257. University of Groningen Groningen, Netherlands, 2015.

MK van Vugt, P Hitchcock, B Shahar, and W Britton. The effects of mbct on affective memory associations in depression: measuring recall dynamics in a randomized controlled trial. *Frontiers in Human Neuroscience*, 6:257, 2012.

Christopher D Wickens. Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, 3(2):159–177, 2002.

Matthew R Yanko and Thomas M Spalek. Driving with the wandering mind: The effect that mind-wandering has on driving performance. *Human factors*, 56(2):260–269, 2014.

A Appendix

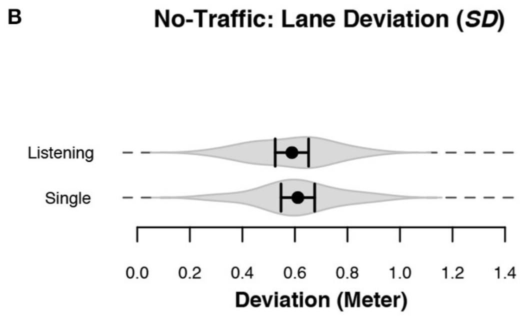

Figure A.1: Standard deviation of lane deviation of human participants. Black dots represent the mean across subject and bars denote 95% CI. Source: Nijboer et al. [2016].

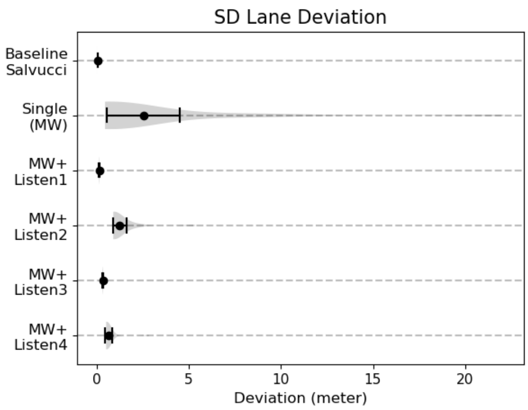

Figure A.2: Standard deviation of lane deviation of model runs. "Baseline" is the original Salvucci [2006] driving model, "single" is the mind-wandering (MW) model and the order of the listening models corresponds to table 2.1. Black dots represent the mean across subject and bars denote 95% CI.

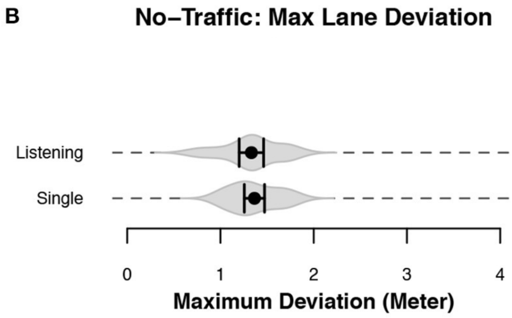

Figure A.3: Maximum of lane deviation of human participants. Black dots represent the mean across subject and bars denote 95% CI. Source: Nijboer et al. [2016].

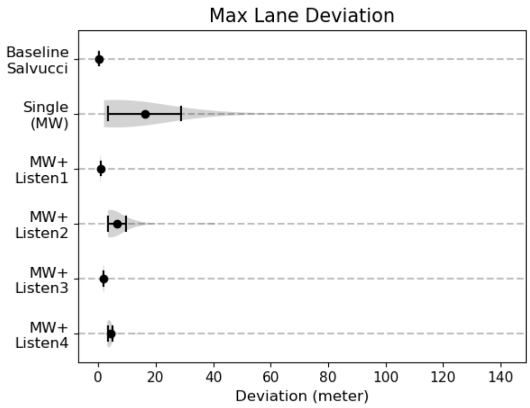

Figure A.4: Maximum of lane deviation of model runs. "Baseline" is the original Salvucci [2006] driving model, "single" is the mind-wandering (MW) model and the order of the listening models corresponds to table 2.1. Black dots represent the mean across subject and bars denote 95% CI.

B Appendix

	Listen model 1	Listen model 2	Listen model 3	Listen model 4
Mean Lane Deviation	0.945	1.498	0.892*	0.899
SD Lane Deviation	1.969	2.025	1.932	1.916*
Max Lane Deviation	14.799*	15.705	14.804	15.077
Direction Change Count	6.776*	12.634	12.793	11.899
Mean Steering Angle	17.057*	36.674	20.081	22.396
Score	**3**	0	1	1

Table B.1: The root mean square deviation values of each index of performance for each of the listening sub-models. Stars mark the smallest values per each index of performance. The score was obtained by counting the number of index minimal values per each model.

	Shapiro test	Wilcoxon-test p-value
Mean Lane Deviation	**5.099e-08**	**0.005788**
SD Lane Deviation	**4.724e-08**	**0.02495**
Max Lane Deviation	**5.139e-08**	**0.0286**
Direction Change Count	**0.006692**	**0.0109**
Mean Steering Angle	**6.957e-05**	**1.319e-11**

Table B.2: The exact p-values of the Shapiro-Wilk normality test and of the Wilcoxon signed-rank test. Bold marks significance (i.e. values smaller than the significance level $\alpha = 0.05$).