Are you sure? Modelling the Confidence of a Driver in Left-Turn Gap Acceptance Decisions

Thesis Report

F. Bontje

Student number:	4436504
Supervisor:	A. Zgonnikov (TU Delft)
Date:	15 May 2023

Are you sure? Modelling the Confidence of a Driver in Left-Turn Gap Acceptance Decisions

Thesis Report

by

F. Bontje

to obtain the degree of Master of Science at the Delft University of Technology, to be defended publicly on Tuesday May 30, 2023 at 12:30 PM.

Student number:	4436504	
Date:	15 May 2023	
Project duration:	July 1, 2022 – May 15, 2023	
Thesis committee:	Dr. A. Zgonnikov,	TU Delft, supervisor
	Prof. dr. ir. J.C.F. de Winter,	TU Delft

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

TU Delft

Contents

1 Abstract

When a person makes a decision, it is automatically accompanied by a subjective probability judgement of the decision being correct, in other words, a (local) confidence judgement. Confidence judgements have, among other things, an effect on justifications of future decisions and behaviour. A better understanding of the metacognitive processes responsible for these confidence judgements could improve behaviour models. To date, confidence judgements are mostly studied in a fundamental manner. Little to no research has been done into confidence in more dynamic tasks. Such applied research could render insights on whether fundamental principles also hold for real-life tasks. It could also have practical relevance for several applications. Driving is amongst the areas for which an improved understanding of the decision making and accompanied confidence judgements can be useful, for instance in order to improve driving assistance systems. However, current studies on driving behaviour are merely focused on decision making and do not take confidence into account.

In this study, we made a first attempt of connecting these two fields of research by investigating the confidence of drivers in left turn gap acceptance decisions in a driver simulator experiment (N=17). The study showed that confidence can be related to the gap size with respect to the oncoming vehicle, described by the time-to-arrival and the distance gap. Confidence increases with the gap size for gap accepting decisions and decreases with the gap size for gap rejecting decisions. In addition, we concluded that confidence can be related to the driving behaviour, and that confidence is negatively related to the decision response time. Moreover, we found that confidence judgements can best be captured with the use of an extended dynamic drift diffusion decision model of which the drift rate of the evidence accumulator as well as the decision boundaries are functions of the time-to-arrival and distance gap. Furthermore, we demonstrated that allowing for post-decision evidence accumulation in the model increases its ability to describe confidence judgements in gap rejecting decisions. Overall, the study confirmed that principles known from fundamental confidence research can be used to describe confidence judgements in a dynamic and applied task.

Topics: Confidence, Decision making, Driver behaviour, Evidence accumulation, Metacognition

2

The paper

Are you sure? Modelling the Confidence of a Driver in Left-Turn Gap Acceptance Decisions

Floor Bontje (4436504) - Master Thesis

Abstract—When a person makes a decision, it is automatically accompanied by a subjective probability judgement of the decision being correct, in other words, a (local) confidence judgement. Confidence judgements have, among other things, an effect on justifications of future decisions and behaviour. A better understanding of the metacognitive processes responsible for these confidence judgements could improve behaviour models. To date, confidence judgements are mostly studied in a fundamental manner. Little to no research has been done into confidence in more dynamic tasks. Such applied research could render insights on whether fundamental principles also hold for real-life tasks. It could also have practical relevance for several applications. Driving is amongst the areas for which an improved understanding of the decision making and accompanied confidence judgements can be useful, for instance in order to improve driving assistance systems. However, current studies on driving behaviour are merely focused on decision making and do not take confidence into account.

In this study, we made a first attempt of connecting these two fields of research by investigating the confidence of drivers in left-turn gap acceptance decisions in a driver simulator experiment (N=17). The study showed that confidence can be related to the gap size with respect to the oncoming vehicle, described by the time-to-arrival and the distance gap. Confidence increases with the gap size for gap accepting decisions and decreases with the gap size for gap rejecting decisions. In addition, we concluded that confidence can be related to the driving behaviour, and that confidence is negatively related to the decision response time. Moreover, we found that confidence judgements can best be captured with the use of an extended dynamic drift diffusion decision model of which the drift rate of the evidence accumulator as well as the decision boundaries are functions of the time-to-arrival and distance gap. Furthermore, we demonstrated that allowing for post-decision evidence accumulation in the model increases its ability to describe confidence judgements in gap rejecting decisions. Overall, the study confirmed that principles known from fundamental confidence research can be used to describe confidence judgements in a dynamic and applied task.

Index Terms—Confidence, Decision making, Driver behaviour, Evidence accumulation, Metacognition

1 INTRODUCTION

Our decisions are automatically and subconsciously accompanied by a confidence judgement [1], [2]. A confidence judgement is defined as a metacognitive judgement of an individual describing the estimated subjective probability that his/her decision is correct [3], [4], [5], [6], [7], [8].

Previous research has demonstrated that the processes responsible for decision making and confidence judgements are similar and closely related to each other [1], [5], [9]. Confidence judgements affect for example the justifications of future decisions and behaviour [10], [1], [11]. Moreover, the processes responsible for decision making [12] and confidence judgements are both based on the accumulation of evidence towards or against a decision [5], [9], [10], [13]. One of the key findings in confidence research is that confidence in a decision increases when more evidence towards this decision is accumulated, and that simultaneously, the amount of evidence for – and therefore also confidence in – an alternative decision decreases [14], [5], [15].

• *Supervisor: Dr. A. (Arkady) Zgonnikov*

Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands

Earlier experimental studies and confidence models provided many insights into confidence judgements. However, these studies have been carried out in a fundamental manner [7], [16], [11], [17], [18], [10] using standard laboratory setups that consider confidence in simplistic perceptual or preferential tasks. Examples of these kinds of tasks are direction discrimination tasks [14], [19] and perceptual categorisation tasks [18], [20].

In addition, limited research has addressed confidence in dynamic tasks until now. Such research is crucial to examine the extent to which earlier findings on confidence in simple tasks also hold in more real-life tasks. A better understanding of confidence judgements for real-life tasks, such as driving, can be of practical relevance. When it comes to driving, an improved understanding of drivers' decision behaviour and accompanied (confidence) judgements is invaluable when designing more sophisticated driving assistance systems [21], [22]. However, earlier studies into the decision making of drivers have not taken into account confidence judgements [23], [24], [25], [26], [27]. The focus of these studies has mainly been on factors influencing decision outcomes and on the modelling of the decision-making process. They proved that internal factors as such personality traits, experiences, and emotional state's as well as external factors like the environment or the traffic situation affect decision outcomes.

In this study, we made a first attempt of connecting these two fields of research by investigating the confidence of human drivers in left-turn gap acceptance decisions at unprotected intersections. This allowed us to obtain new insights into confidence judgements. The main objective of this study was to investigate how to model the confidence of a driver in left-turn gap acceptance decisions. To do so, we examined the influence of the dynamics of an oncoming vehicle on the decision behaviour (Research Question 1a, RQ1a) and response times (RQ1b) of drivers. We defined the dynamics of the oncoming vehicle as the time-to-arrival (TTA) and the distance gap to the ego vehicle. Additionally, we assessed how confidence is related to the time-to-arrival and distance gap of the oncoming vehicle (RQ2) and how confidence is related to the indicated decision response time (RQ3). We also assessed whether the initial throttle operation moment – the moment a participant starts driving – is related to the decision response time (RQ 4). Moreover, we investigated how driving behaviour action dynamics relate to confidence (RQ 5). We thereby expressed driving behaviour with the use of two action dynamic measures: the velocity profile and the distance to the centre of the intersection.

Lastly, we assessed whether confidence judgements of a driver for left-turn gap acceptance decisions can be described by an extended cognitive decision behaviour model (RQ 6a). Thereby, we examined whether the evidence accumulation process can best be examined through one evidence accumulator describing both decisions (drift-diffusion model) or by two independent competing evidence accumulators (race model) (RQ 6b) [26], [28], [29]. Also, we investigated for each model whether it performs best with or without making use of post-decision evidence accumulation (RQ 6c and RQ 6d) [9], [11], [5], [10], [20], [14], [15].

2 Methods

In this study, we investigated the confidence of drivers in left-turn gap acceptance decisions through a within-subjects designed fixed-base driver simulator experiment ($N = 17$; 9 male; 8 female; age= 31+/-11 years). All participants were in possession of a valid driving license. The Human Research Ethic Committee of the TU Delft approved the study and participants signed an informed consent form prior to the experiment.

2.1 Setup

For the driving simulation, we used the simulator software Carla and Unreal Engine4 in combination with Python on a Windows-based desktop computer. The driver simulator consisted of a 65-inch screen (Samsung UE65MU7000) and a Logitech G29 steering wheel and pedals set. Participants were seated at a distance of about 1.3 meters from the screen.

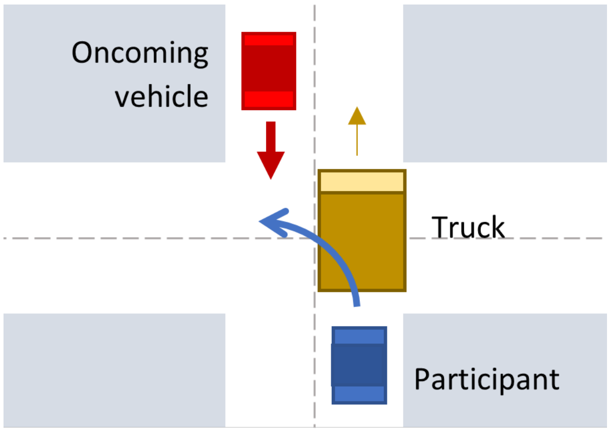

Fig. 1: Illustration of the traffic situation scenario (Top view)

2.2 Experiment protocol

During the experiment (Figure 1), the participants were instructed to drive an auditory navigated route through a virtual urban area, which consisted of a square grid of roads of 150 by 150 meters. Before each intersection, at a distance of 120 meters and 30 meters, navigation prompts were provided, instructing the participant to drive straight ahead or to make a left or right turn.

Each participant drove five different randomly generated routes, with each route comprising 20 intersections. At 80% of the intersections, the participant had to make a left-turn across the path of an oncoming vehicle. The participants were first forced to stop behind a truck at the centre of the intersection before performing the turn (Figure 2). The truck was present in order to limit the view of the driver and to make the oncoming car appear in a natural manner. Participants had to make a gap acceptance decision, deciding to perform the turn in front of the oncoming vehicle ("go" decision) or wait until it had passed ("wait" decision). They had to indicate the decision at the decision moment by pressing a button on the steering wheel. Furthermore, participants had to submit a confidence rating regarding their decision after performing the turn.

During the driving task, we altered the traffic situation by varying the distance gap and the time-to-arrival (TTA) conditions of the oncoming vehicle. The distance gap and time-to-arrival were randomly chosen between 70 and 90 meters and 5.5 and 6.5 seconds, respectively. On each route, each combination of distance and time-to-arrival was present four times. The ratio between the initial time-to-arrival and distance gap determined the constant velocity of the oncoming vehicle, which ranged between 38.77 km/h and 58.91 km/h.

To become comfortable with the task and to limit the effects of learning, the participants drove at least one and if necessary, multiple practice trials before starting the experiment. We reduced fatigue and habituation effects by planning a break halfway through the experiment and by offering the participants the opportunity to take more breaks if needed.

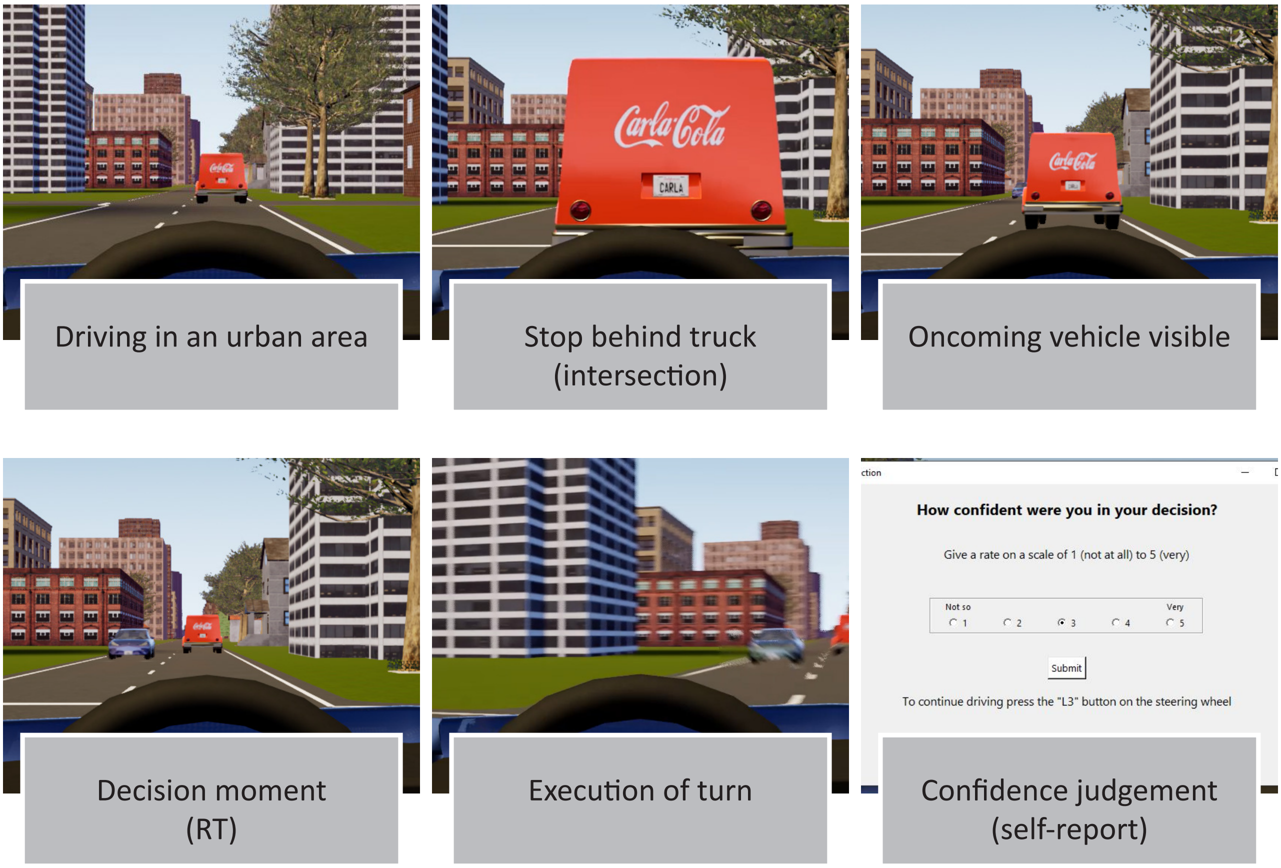

Fig. 2: Screen shots of the six phases of a trial in the driver simulator experiment – 1) Driving through the urban environment, 2) Stop at intersection, whereby vision is blocked by the truck, 3) Truck drives away and oncoming vehicle becomes visible, 4) Participant has to decide to "go" or "wait" (decision response time, RT), 5) Execution of the turn, 6) Self-report on the amount of confidence the participant had in the made decision.

During the experiment, we recorded the positions, velocities, and accelerations of all the vehicles (frequency: 100 Hz). We also recorded the gas throttle input of the participants' vehicle.

2.3 Data analysis

We restricted our data analysis to examining only the left-turn trials in which the indicated decision was actually conducted (Appendix F). We excluded all changes of mind, situations in which the participant carried out a different decision than indicated (3.4% of all left-turn decisions), as well as the trials in which the participants did not indicate their decision (by failing to press the designated button) (2.2% of all left-turn decisions).

2.4 Metrics

The independent variables in this research were the time-to-arrival and distance gap, and the dependent variables are the decision outcome ("go"/ "wait"), the response times (RT), initial throttle operation moment, and confidence judgements. Moreover, we assessed two action dynamics - the velocity profile and the distance to the centre of the intersection - as dependent variables of confidence and the decision outcome.

2.4.1 Decision behaviour, response time, the initial throttle operation moment and confidence

We instructed participants to indicate their decision at the moment of their decision through a button press on the steering wheel. We defined the time between the moment the oncoming vehicle occurred and the decision indication moment as the decision response time (RT). As such, we defined the first moment in time in which the oncoming vehicle became visible as the start of the decision-making process. We also investigated the relation between the decision response time and the initial throttle operation moment. The initial throttle operation moment is the first moment in time – counted from the start of the decision-making process – in which the participants used the gas throttle after having stopped behind the truck.

We measured the amount of confidence participants had in their decision at the decision moment by presenting them with the following question for self-reflection – a so-called self-report: “How confident were you in your decision? Give a rate on a scale of 1 to 5”. We posed the question after the turn was performed, at a distance of ten meters from the centre of the intersection, in order to minimise the interruption of the driving task as much as possible. Participants had to provide the confidence judgement before they could resume the driving task.

2.4.2 Action dynamics

In order to analyse how the decision response time and decision confidence are reflected in the subsequent driving behaviour, we employed two measures of the observed action dynamics related to the start and the execution of the turn: the velocity profile during the turn and the distance to the centre of the intersection.

The velocity profile during the turn described the absolute velocity of the participants’ vehicle over time. In all analyses of the velocity profile, we used the initial throttle operation moment as the zero-point in time in order to negate the effect of the decision response time on the action dynamics. For the analyses, the endpoint in time was the moment the confidence judgement was reported in 75% of the trials for “go” and “wait” decisions separately. The 75% cut-off served to account for the varying duration of the performed turns in the comparison of the different velocity profiles. The confidence judgment was reported when the vehicle was ten meters from the centre of the intersection.

To analyse the relation between confidence and the velocity profile, we used the following metrics: the maximum value over time, the root mean square deviation ($RMSD_{Indiv}$, eq. 1), the deviation from the individual mean (DM_{Indiv}, eq. 2) and the deviation from the group mean (DM_{Gr}, eq. 3).

The root mean square deviation (eq. 1) as well as the deviation from the mean (eq. 2 and eq. 3) indicate how the velocity profile of a given trial deviates from an average trial. We computed the deviation from the individual mean (eq. 2) and root mean square deviation by comparing the velocity profile of the trial with the personal average velocity profile of this kind of decision (“go”/ ”wait”) over all the turns performed by a participant. For the deviation from the group mean (eq. 3), we compared the velocity profile of a trial to the overall decision average for both specific decision outcomes.

The root mean square deviation is the square root of the sum of the squared deviations of the action dynamic measure (y) from the individual mean ($\mu_{decision,indiv}$), divided over the total number of time steps (N), and is described by the following equation:

$$RMSD_{indiv} = \sqrt{\frac{\sum_{i=1}^{N_t}(y-\mu_{decision,indiv})^2}{N_t}} \tag{1}$$

The deviation from the mean describes the sum of the mean deviations, or the sum of the differences between an individual input value of the action dynamic measure and the average value at each time step, divided over the total number of time steps. The deviation from the individual mean and the deviation from the group mean are respectively described by the following equations:

$$DM_{indiv} = \frac{\sum_{i=1}^{N_t}(y-\mu_{decision,indiv})}{N_t} \tag{2}$$

$$DM_{Gr} = \frac{\sum_{i=1}^{N_t}(y-\mu_{decision,gr})}{N_t} \tag{3}$$

The main advantage of using deviations from the mean is that they provide an indication of the amount and the direction of deviation. There is, however, also a risk that deviations from the mean level out the values of the deviation. The root mean square deviation prevents the positive and negative deviations from cancelling each other but does not provide an indication of the direction of the deviation.

When it comes to the analysis of the distance to the centre of the intersection, we used a similar approach as for the velocity profile. Also in this case, the zero-point in time was the initial throttle operation moment and the endpoint in time was the moment the confidence judgement was reported in 75% of the trials for “go” and “wait” decisions separately. We used the root mean square deviation ($RMSD_{Indiv}$, eq. 1), the deviation from the individual mean (DM_{Indiv}, eq. 2) and the deviation from the group mean (DM_{Gr}, eq. 3) as metrics to analyse the relation between confidence and this measure. An additional metric that we took into account is the minimum value within the time frame. That is because the distance to the centre of the intersection is a measure of the travelled trajectory, which in particular describes the corner cutting behaviour performed during the turn.

2.5 Statistical analysis

For the statistical analysis, we made use of MATLAB R2020a. To perform the linear regression analysis, we used the fitlme (fit linear mixed effects model) function. We checked for convergence by analysing the positive definiteness of the Hessian, and we set the level of significance at $\alpha = 0.05$. In the decision behaviour model, we described the probability of making a gap acceptance decision of different time-to-arrival and distance gap conditions of the oncoming vehicle. The regression models that allow for decision dependence of the fixed effects used the “go” decision as reference class, which means that all fixed effects described for “wait” decisions are relative to the ones found for “go” decisions. We checked the significance of the fixed effects (for “go” decisions) and the cumulative fixed effect (for “wait” decisions) of the lme models by using the hypothesis test coefTest, an F-test for which we used a significance of alpha 0.05 (CI 95%) resulting in a critical F-value of 3.8476.

For the data analyses, we used the aggregate data of all participants, thereby accounting for the influence of individual differences through the random intercept term (1 | ID) or (decision | ID), in which ID refers to a single participant. This random effects term allowed for a potential (correlated) random effect of the intercept as well as for the random effect of the decision (Appendix B).

We made use of linear mixed effects models to describe the relation between 1) the decision outcomes on the one hand and the time-to-arrival (TTA) and the distance gap on the other (eq. 4); 2) the decision response time (RT) on the one hand and the time-to-arrival and the distance gap on the other (eq. 5); 3) confidence judgements on the one hand and the time-to-arrival, the distance gap and the decision response time on the other (eq. 6), and; 4) confidence judgements on the one hand and the time-to-arrival, the distance gap and the initial throttle operation moment on the other (eq. 7).

$$Pr_{go} \sim distance + TTA + (1|ID) \tag{4}$$

$$\begin{aligned} RT \sim distance * decision + \\ TTA * decision + (decision|ID) \end{aligned} \tag{5}$$

$$\begin{aligned} Conf \sim RT * decision + distance * decision + \\ TTA * decision + (decision|ID) \end{aligned} \tag{6}$$

$$\begin{aligned} Conf \sim Thr_{int} * decision + distance * decision + \\ TTA * decision + (decision|ID) \end{aligned} \tag{7}$$

In addition, we used linear mixed effects models to investigate the effect of confidence on the metrics of the different action dynamic measures. In general, these models can be described by the following formula:

$$Metric \sim Conf * decision + (1|ID) \tag{8}$$

2.6 Cognitive modelling

For the cognitive modelling, the independent variables were the time-to-arrival and distance gap conditions. The dependent variables, obtained from the experiment, comprised the decision output, the response times, and the confidence judgements.

The premise of the investigated potential confidence models is that they are an extension of decision models that capture the decision behaviour and response times. Before we constructed and evaluated the confidence models, we first created and evaluated two different decision models: a race model and a dynamic drift diffusion model.

2.6.1 Decision models

The baseline model of both decision models was the dynamic drift diffusion model for left-turn gap acceptance decision presented by Zgonnikov et al. 2020 [26]. This dynamic drift diffusion model differs from the classical drift-diffusion model in two aspects. Firstly, the drift-rate is developing over time as a function of the time-varying perceptual information (time-to-arrival and distance). Secondly, the decision boundaries are decreasing over time and defined as a function of the time-to-arrival. Our proposed dynamic drift diffusion model differed from the baseline model by describing both left-turn gap acceptance ("go") and gap rejecting ("wait") decisions. Furthermore, it accounted for the effect of the distance gap on the decision boundary in addition to the effect of the time-to-arrival. As such, our models' decision boundary is a function of the time-to-arrival as well as of the distance gap. The race model describes the decision process by two competing independent decision accumulators for both "go" and "wait" decisions [29], instead of using one evidence accumulator for both decisions as in the dynamic drift diffusion model. Our race model described the drift-rates as well as the decision boundaries as a function of the time-varying perceptual information, the time-to-arrival and distance gap of the oncoming vehicle. In order to train both decision models for "go" as well as "wait" decisions, we defined a new loss function. This function was the sum of the value of the weight least squares (WLS) for the four different traffic conditions. The weight least squares were defined as the sum of the "go" and "wait" weight least squares in each traffic condition. We calculated them with the use of vincentized distributions [30].

2.6.2 Confidence models

We built the confidence models based on the two decision models by using the same evidence accumulator for confidence judgements as for the decision making. After obtaining the value of the evidence accumulator at a confidence response time (CT), we translated it into a confidence judgement. This translation was described by a linear equation with two parameters, accounting for the sensitivity to evidence and the intercept/bias. We trained the confidence models using the root mean square error (RMSE, eq. 9) as a metric of the performance. The RMSE describes the difference between the predicted mean confidence judgements and the observed mean confidence judgements:

$$RMSE = \sqrt{\frac{(Conf_{pr} - Conf_{exp})^2}{N_{conf}}} \tag{9}$$

$$Conf_i = [\mu_{conf,go,i}, \mu_{conf,wait,i}] \tag{10}$$

$\mu_{conf,\ decision,i}$ contains the mean confidence values for the four conditions in the specified decision, of the prediction or the experiment. N_{conf} is the number of measuring points, so $N_{conf} = 4 * 2 = 8$. For optimisation of the model parameters, we made use of the "fmincon" function of MATLAB R2020a.

3 Results

This section elaborates on the results of our research. To do so, it first presents the hypotheses we formulated. Thereafter, the section discusses the findings.

3.1 Hypotheses

The main objective of this research was to investigate how to model the confidence of a driver in left-turn gap acceptance decisions. Therefore, we formulated various hypotheses and sub-hypotheses that were assessed in this study. These are elaborated upon in this section. The hypotheses correspond to the corresponding research questions as presented in the introduction.

3.1.1 Decision behaviour and response times

Our first hypothesis related to the relation between the time-to-arrival and distance gap of the oncoming vehicle and the decision outcome and response times. We expected to observe a positive relation between the "go" decision probability and the time-to-arrival and distance gap conditions (Hypothesis 1a, H1a), and a positive relation between the decision response time and the time-to-arrival for both decisions (H1b). This is similar to the findings of Zgonnikov et al. 2020 [26] that the probability of making a gap acceptance ("go") decision as well as the decision response times are affected by the perceptual information capturing the time-to-arrival and distance gap. Their study showed a positive relation between "go" decision probability and the time-to-arrival and distance gap. Regarding decision response times, they found a positive relation with the time-to-arrival.

3.1.2 Confidence: time-to-arrival and distance gap

Secondly, we expected to observe a relation between confidence and the present perceptual evidence, the time-to-arrival, and the distance gap of the oncoming vehicle, for both decision outcomes (H2). Earlier research has shown that confidence judgements relate to the available perceptual evidence towards a decision [8], [14], [5].

We hypothesised that confidence positively relates with the time-to-arrival and the distance gap of the oncoming vehicle in "go" decisions, and that it negatively relates with these factors in "wait" decisions (H2.1). That is because a larger time-to-arrival and distance gap gives the driver more time and space to perform the left-turn, and a smaller time-to-arrival and distance gap decreases the chance of causing a collision. In other words, the time-to-arrival and distance gap of the oncoming vehicle positively relate to the amount of evidence in favour of the "go" decision and negatively relate to the amount of evidence in favour of the "wait" decision. However, we expected that the magnitude of the relation between confidence and the perceptual evidence depends on the decision outcome. To be specific, we hypothesised that the time-to-arrival and distance gap have a stronger effect on the confidence judgements for "go" decisions as on the confidence judgements for "wait" decisions (H2.2).

3.1.3 Confidence: Decision response time and initial throttle operation moment

Our third hypothesis was that decision response time is negatively related to confidence (H3). Earlier confidence research has shown that the longer participants need for their evidence accumulation to reach a decision boundary, the more their confidence decreases [14], [20]. We also expected this to hold for left-turn gap acceptance ("go") decisions.

Moreover, we expected that the initial throttle operation moment is related to the decision response time (H4). We hypothesised that it can be used as a behaviour measure of the decision response time for "go" decisions (H4.1). That is because the initial throttle operation moment is the moment in which the vehicle enters the intersection, which can be considered the start of the execution of the turn. For "go" decisions, that is the start of the execution of the task. Moreover, we expected a negative relation between the initial throttle operation moment and confidence (H4.2). This is a result of the combination of our expectation that the initial throttle operation moment is a measure of the decision response time and the expected negative relation between the indicated decision response time and confidence (see H3).

3.1.4 Confidence: Action dynamics

We expected to observe a relation between the driving behaviour action dynamics and the confidence judgements. Specifically, we expected that the velocity profile and the distance to the centre of the intersection are related to confidence (H5), since the level of confidence is related to task performance. Worse performance corresponds to lower post-decision confidence judgements [7]. As a result, we expected:

- to observe a negative relation between confidence on the one hand and the root mean square deviation of the velocity profile and distance to the centre of the intersection on the other (H5.1). In other words, that a worse performance is related to a decrease in confidence, which we expected to be able to observe by the enhanced deviation from a participants' average driving behaviour;
- that the level of confidence affects the driving behaviour action dynamic measures of "go" decisions to a greater extent than those of "wait" decisions (H5.2). That is because we assumed that the action dynamics are only a measure of the execution of the decision for "go" decisions;
- to observe a negative relation between confidence on the one hand and the maximum velocity and the deviation from the individual mean for "go" decisions on the other (H5.3), as we expected a higher velocity during the execution of low confidence left-turn decisions. We assumed that lower confidence judgements in "go" decisions result in a more rushed execution, as participants seek to avoid a collision;
- to observe a positive relation between confidence and the deviation from the mean of the velocity profile and that of the distance to the centre of the intersection (H5.5), since poor performance and high-risk driving behaviour (i.e., increased corner cutting driving behaviour and high velocities) are related to overestimation of performance [31], [32], [33].

3.1.5 Cognitive modelling

We hypothesised that the same cognitive evidence accumulation process can be used for modelling confidence judgements as for modelling the decision-making process (H6a). Earlier confidence studies have shown that these two cognitive processes are related to each other [2], [10], [15], [9] [34]. Furthermore, earlier research has concluded that both processes can be described by an evidence accumulation process [14], [9], [28]. For the task in this study, the sources of the perceptual evidence were equal for both decision making and confidence judgement processes.

We expected that the evidence accumulation process can better be examined through a drift-diffusion model than through a race model, as the drift-diffusion uses one evidence accumulator describing both decisions, and thereby is able to account for the interaction between both decision outcomes (H6b). On the contrary, the race model makes use of two independent accumulators describing both decision outcomes, meaning it cannot account for a potential interaction between the two.

Lastly, we hypothesised that the drift diffusion model performs better when making use of post-decision evidence accumulation (H6c). That is because in a drift-diffusion model without post-decision evidence accumulation, the confidence response time is equal to the decision response time. As such, the amount of evidence will per definition be equal to the decision boundary, which implies that confidence is only described by the decision response time [14], [28]. Allowing for post-decision evidence accumulation will lead to different time moments for the confidence judgement and the response, which in drift-diffusion models makes it possible to describe how confidence relates to the evidence accumulation process. For the race model, this problem does not occur as the value of the losing evidence accumulator can be used for the computation of confidence [14]. Therefor, we expected that the effect of allowing for post-decision evidence accumulation for the race model is indifferent (H6d). The hypothesis that post-evidence accumulation is used for confidence judgments is in line with findings of earlier research [9], [11], [5], [10], [20].

We reflect on the obtained results regarding these hypotheses in the sections below. All the presented results are based on the aggregate data across all participants.

3.2 Decision behaviour and response times

We found that the drivers' decision behaviour (Figure 3) and response times (Figure 4A) are influenced by the time-to-arrival and distance gap of the oncoming vehicle.

3.2.1 Decision behaviour

We performed a linear regression analysis between the probability of making a "go" decision and the time-to-arrival and distance gap of the oncoming vehicle. As expected (H1a), we found a positive relation between the probability of making a "go" decision and the time-to-arrival (t = 9.2, p = 1.7e-19) and the distance gap (t = 19, p = 4.5e-73) (table 1). As such, the probability of making a "go" decision increased with larger time-to-arrival and distance gap conditions. Consequently, the reverse proved true for "wait" decisions: The probability of making a "wait" decision decreased with larger time-to-arrival and distance gap conditions.

	Estimate	Std. Error	t-score	pValue
(Intercept)	-2.175	0.1562	-13.93	1.336e-41
TTA	0.1860	0.02031	9.157	1.659e-19
Distance	0.01940	0.001016	19.09	4.508e-73

TABLE 1: Logistic regression table, describing the relation between the decision probabilities and the TTA and distance gap for "go" decisions, lme: $Pr_{go} \sim dist + TTA + (1|ID)$

.

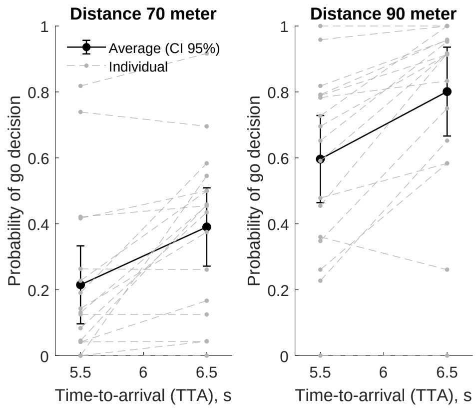

Fig. 3: Average and individual mean gap acceptance decision behaviour for different conditions of the oncoming vehicle (time-to-arrival and distance gap).

3.2.2 Response times

Regarding the decision response times (RT), the results of the linear regression analysis (Figure 4B, table 2) demonstrated that for "go" decisions, response times are positively related to the time-to-arrival (t = 4.0, p = 6.5e-05) and to the distance gap (t = 4.4, p = 9.9e-06).

We found that the intercept and the effect of the time-to-arrival are independent of the made decision. The results showed an additional positive effect of 0.011 (t = 4.6, p = 5.5e-06) of the distance gap on the decision response time for "wait" decisions compared to "go" decisions 0.0073 (F = 20, p = 9.9e-06). This resulted in a cumulative fixed effect of the distance gap in "wait" decisions of 0.018 (F = 116, p = 4.6e-26). This additional positive effect of the distance gap can potentially explain the observed difference between the mean decision response times for "go" and "wait" decisions (Figure 4A), of respectively M = 1.7 (SD = 0.43) and M = 2.3 (SDS = 0.66).

We assessed the influences of personal differences on the reaction time by allowing for random effects in the lme model. These results (Appendix B) show an additional slope for the "wait" decision (M = -0.036, SD = 0.66) for seven participants.

In short, response times positively relates to the time-to-arrival for both decision outcomes as expected (H1b). In addition, the distance gap also positively relates to the response time. The distance gap influenced the decision response time more for "wait" decisions than for "go" decisions.

The relation between decision response time and the distance gap has not been found in the earlier study into left-turn gap acceptance decisions by Zgonnikov et al. 2020[26], which only found a relation between response times and the time-to-arrival. The used time-to-arrival and distance gap conditions in the experiment potentially explain this difference, as they result in different velocities of the oncoming vehicle. In this study, the distance gap varied between 70 and 90 meters and the time-to-arrival between 5.5 and 6.5 seconds whereas the earlier study used a distance gap variation between 90 and 150 meters and a time-to-arrival variation between 4 and 6 seconds. As a result, in this study the velocity of the oncoming vehicle ranged in between 38.77 km/h and 58.91 km/h, while in the earlier study the velocity ranged between 55 km/h and 135 km/h. However, further research is required to investigate how the time-to-arrival and distance gap conditions influence the dependency of the decision response time on the distance gap condition.

	Estimate	Std. Error	t-score	pValue
(Intercept)	-2.175	0.1562	-13.93	1.336e-41
TTA	0.1860	0.02031	9.157	1.659e-19
Distance	0.01940	0.001016	19.09	4.508e-73

TABLE 2: Results of regression analysis describing the effect of the TTA and distance gap conditions on the decision response time for "go" decisions and the effect on "wait" decisions with respect to "go" decisions, $RT \sim dist * dec + dec * TTA + (dec|ID)$.

3.3 Confidence

In this section, we evaluate our findings concerning: a) the relation between confidence on the one hand and time-to-arrival and distance gap on the other, b) the relation between confidence on the one hand and the decision response time and initial throttle operation moment on the other, and c) the relation between confidence and measures of action dynamics.

3.3.1 Time-to-arrival and distance gap

The results of the linear regression analysis of the relation between confidence and respectively the time-to-arrival and the distance gap conditions confirmed that the time-to-arrival as well as the distance conditions affect confidence (table 3, Figure 4B).

For "go" decisions (table 3), the results displayed a positive relation between confidence on the one hand and the time-to-arrival (b = 0.74, t = 12, p = 1.8e-32) and distance gap (b = 0.034, t = 10, p = 1.85e-23) conditions

	Estimate	Std. Error	t-score	pValue
Intercept	-2.277	0.4975	-4.576	5.126e-06
TTA	0.7413	0.06104	12.14	1.830e-32
Distance	0.03361	0.003304	10.17	1.465e-23
RT	-0.7411	0.08663	-8.554	2.841e-17
Wait decision	10.26	0.6763	15.17	1.599e-48
Wait decision: TTA	-1.123	0.0886	-13.06	5.231e-37
Wait decision: Distance	-0.04432	0.004790	-9.253	7.133e-20
Wat decision: RT	0.1349	0.1055	1.279	0.2013

TABLE 3: Results of regression analysis of the effect of the RT, the distance gap and TTA conditions on confidence judgements for different decision outcomes The reference class in the lme are the "go" decisions, so the found fixed effects coefficients for "wait" decisions are relative to the "go" decision coefficients, lme: $Conf \sim RT * decision + TTA * decision + distance * decision + (decision|ID)$.

on the other. The hypothesis test of the fixed effects for "go" decisions confirmed these positive relations (critical F-value, $F < 3.85$) for the time-to-arrival (F = 147, p = 1.8e-32) and the distance gap (F = 103, p = 1.4e-23). For "wait" decisions, we observed, relative to "go" decisions, a negative relation between confidence and respectively the time-to-arrival (t = -13, p = 5.2e-37) and the distance gap (t = -9.3, p = 7.1e-20). This negative relative relation resulted in negative net estimate values of the fixed effects of the time-to-arrival (b = -0.38, F = 40, p = 3.9e-10) and distance gap (b = -0.011, F = 9.5, p = 0.0021) on confidence in "wait" decisions. With these findings, we proved that the relation between confidence on the one hand and the time-to-arrival and the distance gap of the oncoming vehicle on the other is positive for "go" decisions and negative for "wait" decisions (H2.1).

The net fixed effects of the time-to-arrival and the distance gap for "wait" decisions have a smaller absolute magnitude and F-value compared to the fixed effects for "go" decisions. This indicates that confidence judgements were less affected by the time-to-arrival and distance gap conditions in "wait" decisions compared to "go" decisions. This finding is in line with the hypothesis that "go" decisions are more strongly affected by the time-to-arrival and distance gap conditions (H2.2).

A potential explanation of this stronger effect of the time-to-arrival and distance gap of the oncoming vehicle on "go" confidence judgements can be found in the different implications of "incorrect" decisions for both decision outcomes. Incorrect "wait" decisions involve no risks, as an unnecessary waiting decision will not be punished. On the other hand, an incorrect "go" decision comes with the risk of causing a collision which is related to the time-to-arrival and distance gap of the oncoming vehicle.

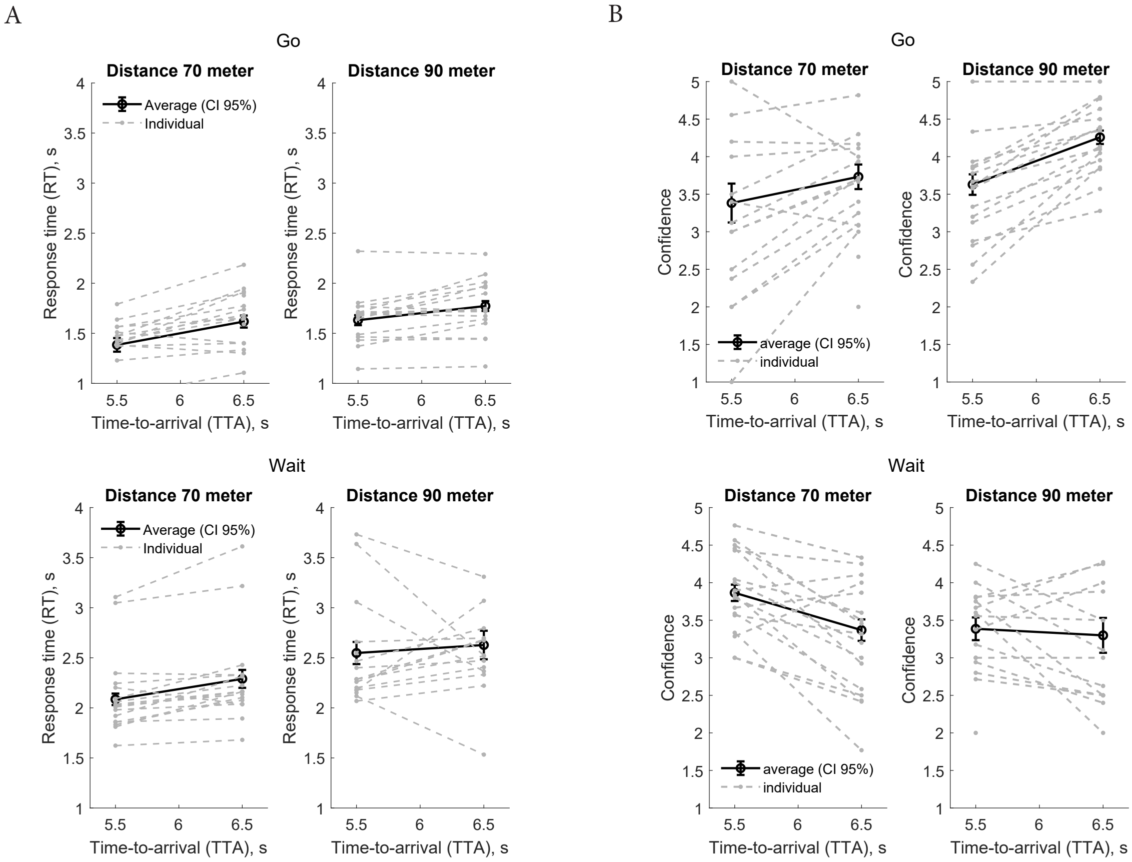

Fig. 4: Average and individual mean A) decision response times, and B) confidence judgements, for different decision outcomes (“go”/ ”wait”) and different time-to-arrival and distance gap of the oncoming vehicle.

3.3.2 Decision response time and initial throttle operation moment

We investigated the relation between the decision response time and the initial throttle operation moment on the one hand and confidence on the other for the decision behaviour in left-turn gap acceptance/rejection decisions. To do so, we firstly assessed the relation between confidence and the moment of indicating the decision, the decision response time. Second, we examined the connection between the latter and the initial throttle operation moment. Last, we assessed the relation between confidence and the initial throttle operation moment.

Confidence and decision response time

We found a negative correlation between the decision response time and confidence (r = -0.27, p = 1.3e-27) for all decisions. The linear regression model of the relation between confidence, the response time, the distance gap, and time-to-arrival conditions (eq. 6) accounts for the fixed effect of the decision response time on the confidence judgements (table 3). The results substantiated the hypothesised (H3) negative relation (b = -0.74, t = -8.6, p =2.84e-17) between confidence and response time for “go” decisions. In addition, we observed no additional significant effect of the response times in “wait” decisions, meaning the effect of the response on confidence is independent of the made decision.

The negative relation between confidence and decision response time is in compliance with earlier confidence research, which argues that less (qualitative) evidence towards a decision results in a longer duration of evidence accumulation before making a decision, which results in a reduced amount of confidence [14], [20].

Decision response time and initial throttle operation moment

The start of the execution of the turn – the moment the participant enters the intersection - is indicated by the initial throttle operation moment. We observed that participants started driving before indicating their decision in 47% of all the “go” and in 27% of all the “wait” decisions (Figure 5). As a result, in this task it was not possible to consider the initial throttle operation moment as a “clear” behaviour-based indication of the response time, which contradicts our hypothesis (H4.1). This finding was substantiated by the moderate positive correlation found between the decision response time and initial throttle operation moment for “go” decisions (r = 0.28, p = 1.6e-15) and the moderate negative correlation for “wait” decisions (r = -0.24, p = 1.46e-11).

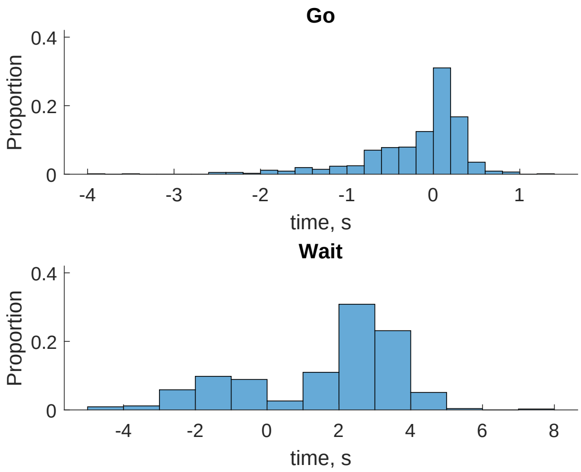

Fig. 5: Distribution of the initial throttle operation moment relative to the response time (button press), for ”go” and ”wait” decisions separately.

Confidence and initial throttle operation moment

Regarding the relation between confidence and the initial throttle operation moment, we observed a negative correlation (r = -0.22, p = 1.16e-18) over all decisions. To further investigate the relation, we performed a linear regression analysis (eq. 7), which is similar to the one used for the analysis of the relation between confidence and the response time. The results (table 4) demonstrated a negative relation between the initial throttle operation moment and confidence (b = -0.34, t = -5.1, p = 4.1e-07) for “go” decisions. For “wait” decisions, we found a relatively positive fixed effect (b = 0.33, t = 4.7, p = 3.3e-06) of the initial throttle operation moment on confidence compared to “go” decisions. This additional effect levelled out the effect of the initial throttle operation moment on confidence found for “go” decisions (F = 213, p = 3.3e-45), resulting in a net effect of -0.015 (F = 226, p = 1.2e-47) in “wait” decisions. This indicates that the effect of the initial throttle operation moment was strongly reduced for “wait” decisions.

3.3.3 Action dynamics

The following section evaluates the relation between decision response time, confidence, and action dynamic measures. We considered two measures of action dynamics: the velocity profile during the turn and the distance to the centre of the intersection.

Velocity profile

We analysed the effect of confidence on the following four metrics of the velocity profile: the root mean square deviation, the deviation from the individual mean, the deviation from the group mean and the maximum value.

The linear regression analyses of these metrics demonstrate (Appendix C) that we did not observe relations between confidence and respectively the deviation from the

	Estimate	Std. Error	t-score	pValue
Intercept	-2.406	0.5163	-4.660	3.445e-06
TTA	0.6925	0.06356	10.90	1.125e-26
Distance	0.02978	0.003414	8.722	6.993e-18
Thr_{int}	-0.3413	0.0671	-5.0871	4.086e-07
Wait decision	10.32	0.7073	14.58	3.325e-45
Wait decision: TTA	-1.147	0.09024	-12.71	2.839e-35
Wait decision: Distance	-0.05141	0.004853	-10.60	2.319e-25
wait decision: Thr_{int}	0.3260	0.06981	4.670	3.273e-06

TABLE 4: Results of regression analysis of the effect of the initial throttle operation moment (Thr_{int}), the distance gap and TTA conditions on confidence judgements for different decision outcomes. The reference class in the lme are the “go” decisions, so the found fixed effects coefficients for “wait” decisions are relative to the “go” decision coefficients, lme: $Conf \sim Thr_{int} * decision + TTA * decision + distance * decision + (decision|ID)$.

individual mean, root mean square deviation and the maximum value of the velocity profile. We only observed a positive relation between the deviation from the group mean of the velocity profile and confidence (b = 0.11, t = 2.2, p = 0.028). For “wait” decisions, confidence had a similar positive effect on the group mean as no additional effect of the deviation from the group mean was observed. The figure (Figure 6A) substantiates this positive relation between confidence and the velocity profile for “go” decisions. However, it also shows that the deviation is minor.

The data thus suggest that decisions rewarded with higher confidence judgements related to faster turns with respect to the mean velocity profile. They should, however, be interpreted with caution, given their relatively low significance.

The independence of the relation between confidence and the deviation from the group mean of the decision outcome suggested that this metric indicates the relation between confidence and someone’s general driving style, rather than the relation between confidence in a particular decision and the subsequent driving behaviour. If it had indicated the relation between the two in a particular decision, a different effect of the decision confidence on the subsequent driving behaviour should have been observed between “go” and “wait”. That is because for “go” decisions, the subsequent driving behaviour, the performance of the turn, is a direct measure of the execution of the decision which, like confidence, is likely to be affected by the dynamics of the oncoming vehicle. For “wait” decisions, the driving behaviour can be considered a measure of the general driving style, since the turn is performed after the decision (waiting) has been made and the oncoming vehicle has passed.

As a result, the relation between confidence and the deviation from the group mean of the velocity profile suggests that participants with higher confidence judgements have a faster driving style compared to

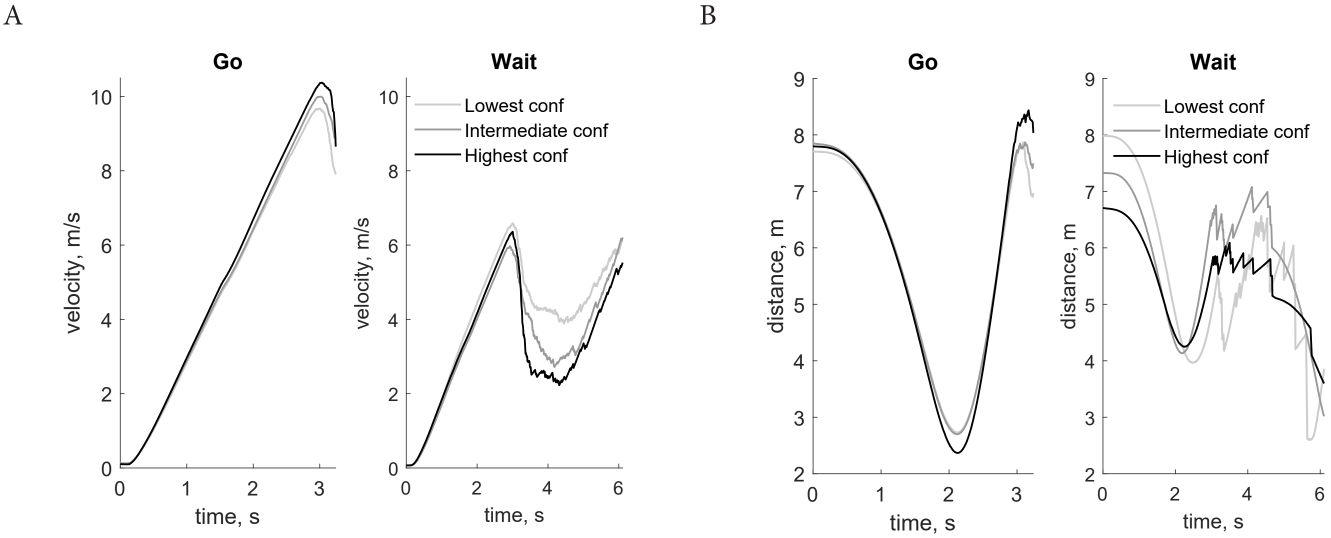

Fig. 6: Illustration of the group means values over time of the A) distance to the centre of the intersection, and the B) velocity. The tertiles contains 1/3 of all trails over all participants within the specified decision ("go"/ "wait"), the tertiles represent the lowest, intermediate and highest confidence judgements.

other individuals. This is consistent with our hypothesis regarding overconfidence in poor performance and high-risk driving behaviour (H5.5). The phenomenon described by Kruger and Dunning (1999) [31] that a stronger overestimation of performance more often occurs with lower performances could potentially explain the observed behaviour. However, further research is needed to draw any conclusions about the role of overestimation and self-perception on the observed driving behaviour.

Distance to the centre of the intersection

The distance to the centre of the intersection provides an indication of the travelled trajectory. In particular, this measure indicates the presence of corner cutting behaviour. For the distance to the centre of the intersection, we examined the following four metrics: the root mean square deviation, the deviation from the individual mean, deviation from the group mean and the minimum value.

The performed regression analyses show (Appendix C) that the deviation from the group mean and the minimum value of the distance to the centre of the intersection were affected by confidence.

Confidence showed a negative relation with the deviation from the group mean (b = -0.16, t = -2.12, p = 0.035) as well as with the minimum distance (b = -0.24, t = -2.40, p = 0.017) for "go" decisions.

For "wait" decisions, we observed a positive relative effect of confidence for both metrics (DM_{Gr}: b = 0.31, t = 2.3, p = 0.022, minimum: b = 0.58, t = 3.29, p = 0.0011). The hypothesis test showed that we cannot draw conclusions from the accumulated fixed effect of the deviation from the group mean for "wait" decisions. That is because this effect (b = 0.15, F = 1.7, p = 0.20) was found not to be significantly different from zero. On the other hand, regarding the minimum distance, a positive relation was found between the minimum distance (b = 0.35, F = 5.2, p = 0.024) and confidence for "wait" decisions.

The results hint towards a negative relation between confidence and the deviation from the group mean of the distance to the centre of the intersection, for "go" decisions. This finding contradicts our hypothesis that increased corner cutting behaviour, in general, is related to higher confidence due to the overestimation of performance (H5.5). On the contrary, the negative relation implies that higher confidence decisions resulted in a turn closer to the centre of the intersection (less corner cutting) relative to the general average turn. For "wait" decisions, no relation could be observed between this metric of the travelled trajectory and confidence. The absence of this relation is in line with our hypothesis (H5.2) that decision confidence mainly affects "go" decisions.

In addition, we observed a relation between confidence and the minimum distance to the centre of the intersection. For "go" decisions, the minimum distance related negatively to confidence, which confirms our hypothesis (H5.4). This finding hints towards the presence of corner cutting behaviour at low confidence "go" decisions. For "wait" decisions, the minimum distance positively related to confidence. The relation between a "risk-full" driving style and overestimation of performance as described in earlier studies may serve as explanation why participants who had higher confidence judgements [31], [33], [32] also tended to perform corner cutting behaviour in "wait" decisions. However, further research is needed to confirm this potential explanation.

It is important to note that the significance for all findings regarding the relation between confidence and the distance to the centre of the intersection was relatively low, meaning additional research is required for verification. Figure 6B, which illustrates the average distance to the centre of the intersection for different confidence judgments, shows that the observed relations are marginal for "go" decisions and that for "wait" decisions distortions are present.

3.4 Cognitive modelling

In this paper, we presented and evaluated four potential confidence models. The presented confidence models assumed that the same evidence accumulator is responsible for the decision-making process as for the confidence judgement. The confidence models differed in terms of the kind of decision model they were based on, as well as whether or not they allowed for post-decision evidence accumulation.

All the presented confidence models were an extension of two different cognitive evidence accumulation decision models (Figure 7). The decision models were based on a dynamic drift-diffusion left-turn gap acceptance decision model developed in an earlier study [26]. In the following section, we first discuss the different decision models. Thereafter, we evaluate the different potential confidence models.

3.4.1 Decision models

Baseline decision model

The baseline drift-diffusion left-turn gap acceptance model presented by Zgonnikov et al. 2020 [26] captures the cognitive process responsible for the decision-making, based on perceptual evidence. This section discusses the main principles of this model. It describes that a decision is made when the decision variable (x) reaches the decision boundary (b). The authors defined the moment in time when the decision boundary is reached as the decision response time. They stated that the available perceptual evidence is constructed of the time-to-arrival and the distance gap of the oncoming vehicle. The model captured the gap acceptance behaviour with one evidence accumulator. The study only accounted for "go" decisions and the associated response times, while "wait" decisions were not included.

In the baseline dynamic drift-diffusion model, the rate of change of the decision variable ($dx(t)$, eq. 11) and the decision boundaries ($b(t)$, eq. 12) were functions of the time-varying perceptual evidence, described by the time-to-arrival (TTA(t)) and the distance gap (d(t)). The decision boundaries were solely a function of the time-to-arrival, whereas the drift-rates were a function of the generalised gap ($g(t)$, eq. 13). The generalised gap was defined as a linear combination of the time-to-arrival and the distance gap, in which the contribution of the distance gap relative to the time-to-arrival is described by β.

$$dx(t) = \alpha(g(t) - \theta_{crit})dt + dW \tag{11}$$

$$b(t) = \frac{b_0}{(1 + e^{-k(TTA(t)-\tau)})} \tag{12}$$

$$g(t) = TTA(t) + \beta d(t) \tag{13}$$

The noise present in the evidence accumulation was described by a stochastic Wiener process (W). The drift-rate parameter, $\alpha \geq 0$, indicated the influence of the perceptual information on the evidence accumulation process relative to the noise. The critical value of the generalised gap for which the drift-rate changes sign is described by θ_{crit}. The boundary parameters b_0, k>0 and τ indicated respectively the boundary scale parameter, the sensitivity of the boundary and the value of the time-to-arrival at which the boundary reaches the value of $\pm\frac{1}{2}b_0$. Furthermore, the model accounted for the present perceptual and response delays by adding a non-decision time (t_{ND} eq. 14) to the predicted decision times of the model.

$$t_{ND} = N(\mu_{ND}, \sigma_{ND}) \tag{14}$$

The non-decision time was normally distributed over all samples with mean (μ_{ND}) and standard deviation (σ_{ND}).

Modified decision models

The modified decision models developed in this study hold that the drift-rates as well as the decision boundaries depend on the time-varying perceptual evidence and are thereby similar to the baseline model. The manner in which the rate of change of the decision variable ($dx(t)$, eq. 11) is described remains unchanged. However, we employed an adjusted definition for the decision boundaries (b(t), eq. 15).

We described the decision boundaries as a function of the generalised gap ($g(t)$, eq. 13). We did so in order to let the model capture our findings that besides the decision output, the response times are also related to the time-to-arrival and the distance gap conditions. This differs from the baseline model that described that the decision boundaries are solely a function of the time-to-arrival [26]. The modified decision boundaries can be described by the following equation:

$$b(t) = \frac{b_0}{(1 + e^{-k(g(t)-\theta_{crit})})} \tag{15}$$

In this research, we considered two different modelling methods – the dynamic drift-diffusion model and the race model – in order to investigate what method can most adequately describe the decision-making process responsible for the left-turn gap acceptance and rejection decisions. The main difference between the dynamic drift-diffusion model and the race model is whether they translate perceptual evidence into one or two decision variables to predict the decision outcome.

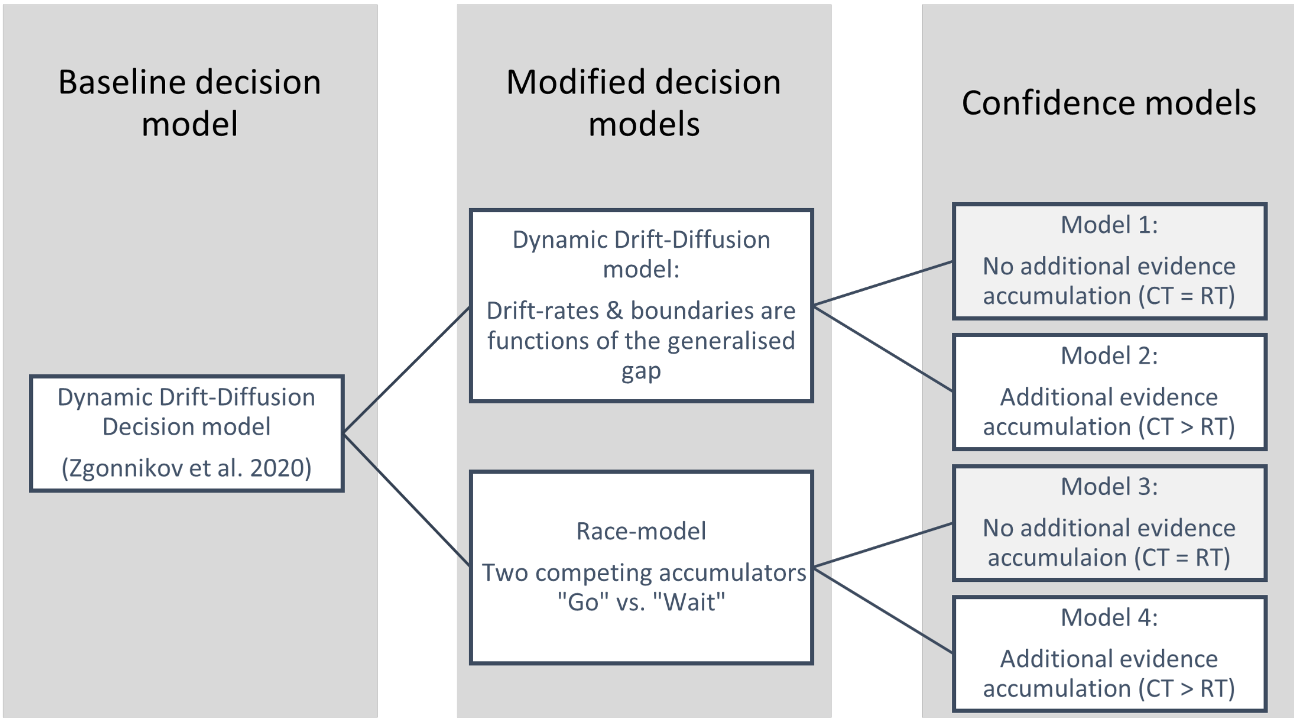

Fig. 7: Overview chart of the applied modelling method: the construction of four potential confidence models based on one baseline decision model.

Drift-diffusion model

The dynamic drift-diffusion decision model described that evidence for both potential decision outcomes can be captured by one evidence accumulation process (Figure 8A). The process was defined by the following equations which capture the change in the decision variable (dx, eq. 17) and the decision boundaries ($b(t)$, eq. 17):

$$dx(t) = \alpha(g(t) - \theta_{crit})dt + dW \tag{16}$$

$$b(t) = \pm\frac{b_0}{(1+e^{-k(g(t)-\theta_{crit})})} \tag{17}$$

A "go" decision is made when the decision variable (x) first reaches the positive decision boundary and a "wait" decision is made when it first reaches the negative decision boundary. The first moment in which the decision variable reaches the decision boundary defines the decision response time.

The models' drift-rate describes that the present perceptual information contributes to evidence accumulation in favour of the "go" decision (towards the positive decision boundary) until the moment in time in which the generalised gap is smaller than its critical value ($g(t) < \theta_{crit}$). The perceptual evidence contributes to the evidence accumulation process in favour of the "wait" decision (towards the negative decision boundary).

This model has seven free parameters ($\alpha, \beta, \theta_{crit}$, $b_0, k, \mu_{ND}, \sigma_{ND}$).

Race model

The race model (Figure 8B) included two independent competing decision variables which are responsible for the decision outcome [29]. Both decision variables are based on the same present perceptual evidence and noise. However, the manner in which the present perceptual evidence advocates in favour or against a decision outcome differs between the two decision variables.

To simplify the model, we assumed that the boundary scale parameters (b_0) and the relative contribution of the time-to-arrival and distance gap to the perceptual evidence (β) are similar for both decisions. Regarding the decision boundaries, we found that using independent boundary scale parameters only led to negligible improvements in the decision model (Appendix E). This indicates that a similar amount of evidence is needed to make "go" and "wait" decisions for similar values of the generalised gap. Moreover, we considered the ratio between the time-to-arrival and the distance gap in the observation of the generalised gap identical for both decisions. That is because the drivers' observation of the generalised gap is likely to be independent from the decision, which it precedes.

In the race model, the available perceptual evidence in the left-turn gap acceptance decision is summarised by the value of the generalise gap ($g(t)$), which constitutes the time-varying time-to-arrival and distance gap (eq. 13). The results of our regression analysis indicated that the time-to-arrival and the distance gap positively relate to the probability of making gap acceptance decisions. This can be explained by the decreasing chance of causing a collision for larger generalised gap sizes. As a result, the chance of causing a collision increases over time as the generalised

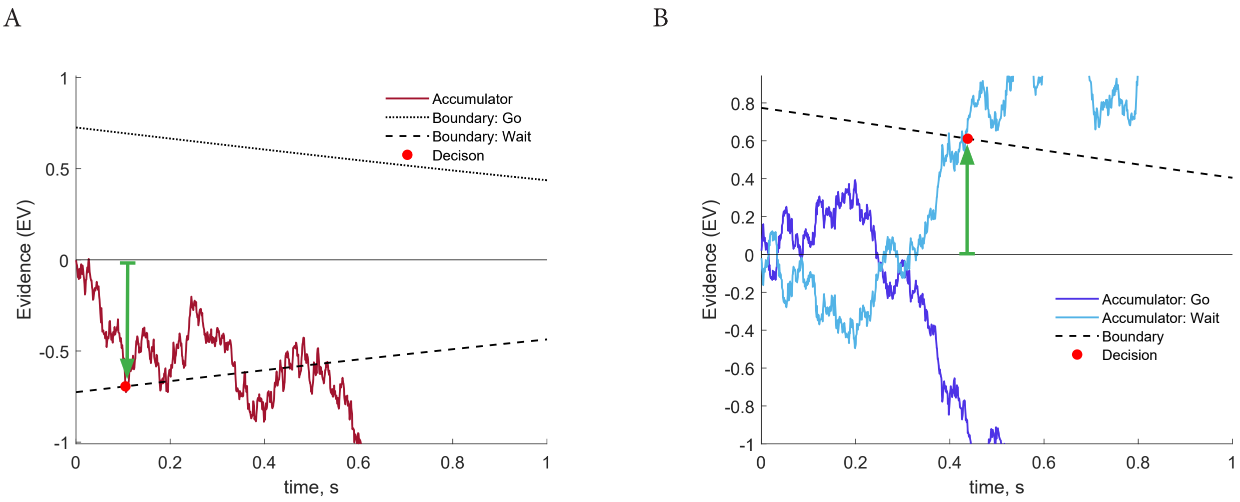

Fig. 8: Illustration of evidence accumulation process for both decision models A) drift-diffusion model, B) race model. The arrow indicates the evidence used for the confidence model V_c (the direction of the arrow is defined as positive).

gap size decreases. The increasing probability of causing a collision is reflected in the decreasing drift-rate for "go" decisions and the increasing drift-rate for "wait" decisions.

To capture this behaviour, we defined the following two equations for the rate of change in evidence ($dx_{decision}$) for the two accumulators:

$$dx_{go}(t) = \alpha_{go}(g(t) - \theta_{crit,go})dt + dW \tag{18}$$

$$dx_{wait}(t) = \alpha_{wait}(\theta_{crit,go} - g(t))dt + dW \tag{19}$$

$$b(t) = \frac{b_0}{(1 + e^{-k(g(t)-\theta_{crit})})} \tag{20}$$

As it is unknown whether the relative contribution of the perceptual information is equal in both decisions, we assumed the drift-rate parameter ($\alpha_{decision}$) as well as the critical value ($\theta_{crit,decision}$) of the generalised gap to be decision dependent. It can be reasoned that perceptual information potentially contributes less to the decision-making process in "wait" decisions than in "go" decisions ($\alpha_{go} > alpha_{wait}$). That is because the execution of the decision and the risk associated with "go" decisions are affected by the time-to-arrival and distance gap of the oncoming vehicle, which does not apply for "wait" decisions.

In addition, the critical value of the generalised gap is potentially influenced by the nature of the decision. The critical value of the generalised gap for "go" decisions describes the value of the generalised gap at which additional perceptual evidence will disfavour the "go" decision. When the generalised gap reaches this critical value, the drift-rate becomes negative. On the other hand, for "wait" decisions, the critical value of the generalised gap indicates the value of the generalised gap after which additional perceptual evidence will contribute to evidence accumulation in favour of a "wait" decision, i.e., the value of the generalised gap in which the drift-rate becomes positive. "Go" decisions are accompanied by the risk of causing a collision, which may result in a more conservative decision process. This leads to a high critical value of the generalised gap, implying that strong evidence is needed before someone makes a "go" decision.

The race model has nine free parameters (α_{go}, α_{wait}, β, $\theta_{crit,go}$, $\theta_{crit,wait}$, b_0, k, μ_{ND}, σ_{ND}).

Results

We compared the performance of the decision models with the use of the weighted least square (WLS, table 5). The results show that both models had a somewhat similar performance. However, the race model performed slightly better. The higher performance of the race model does not necessarily mean that the race model best captures the cognitive process of the decision making. Its better performance could be explained by the fact that the model offers a more general description of the evidence accumulation process due to the extra two free parameters.

The parameters of the race model show that the relative contribution of the perceptual evidence (α) in the change of the decision variable relative to noise was similar for the two different decision variables. The contribution of perceptual evidence was just slightly less for the decision variable describing the "wait" decisions. Moreover, we observed a different critical value of the generalised gap size (θ_{crit}) for the two decisions. The critical value of the gap size was larger for the "go" decisions evidence accumulator compared to the "wait" decisions evidence accumulator. This difference indicates that for "go" decisions, the sign of the drift-rate changes (from positive to negative) at a greater generalised gap compared to the gap size at which the sign of the drift-rate for "wait" decisions changes (from negative to positive). This difference cannot be captured by a drift-diffusion model, in which by definition the change in sign occurs at the same critical value as both decisions are described by one accumulator.

	WLS	α	β	b_0	k	μ_{ND}	σ_{ND}	θ_{crit}
DDM	1.54	1.12	0.109	1.41	0.396	1.51	0.140	14.0
Race	1.33	Go: 1.06 Wait: 1.05	0.108	1.21	0.454	1.54	0.128	Go: 14.0 Wait: 13.5

TABLE 5: Drift-diffusion and race decision models: a) comparison of performance of the decision models, mean WLS over 10 models, and b) parameter values found through optimisation.

3.4.2 Confidence models

In this section, we propose and evaluate different confidence models. The different confidence models were based on the two main theories of how evidence accumulation is used in confidence judgements. The first theory states that the same evidence is used for decisions as for confidence judgements [14], [15], [34]. The second theory argues that decisions and confidence judgements are based on different sets of evidence and that confidence is based on post-decision evidence accumulation [9], [11], [5], [10], [20] [13].

We applied both theories to the two different decision models, as defined in the previous section. This resulted into four potential confidence models (Figure 7). Models 1 and 2 were based on the dynamic drift-diffusion decision model and models 3 and 4 were based on the race decision model. The first theory regarding evidence accumulation in confidence judgements applied for models 1 and 3, whereas the second theory applied for models 2 and 4. These latter two models allowed for post-decision evidence accumulation.

Model principles

For the four proposed confidence models, we defined as a premise that the evidence accumulation process responsible for the decision-making process can be used to model the process responsible for confidence judgements. The manner in which the model uses the evidence accumulation process depends on the used decision model and the theory used to describe confidence. Specifically, the decision model defines how the evidence accumulator(s) are used as an input variable for the confidence model and the used confidence theory determines the moment in time when the confidence judgement is made (confidence response time, CT).

Input variable

The drift-diffusion model-based confidence models (model 1 and 2) made use of the single decision variable accumulator of the decision model. These models imply that confidence judgements are related to the present evidence in favour of the final decision outcome. We defined the input variable ($V_c(t)$) for these confidence models as the amount of evidence in favour of the made decision at time moment t (eq. 21).

$$V_c(t) = \begin{cases} \text{EV(t)} & decision = \text{"}Go\text{"} \\ \text{-EV(t)} & decision = \text{"}Wait\text{"} \end{cases} \tag{21}$$

In the race-model-based confidence models (model 3 and 4), we described that the confidence judgement depends on the value of the evidence towards the losing/alternative decision. For these models, we used as input variable the negative value of the losing/alternative evidence accumulator, described by the decision model, at time moment t (eq. 22).

$$V_c(t) = -EV_{alt}(t) \tag{22}$$

Confidence response time

The time moment t at which the input variable is evaluated, the confidence response time (CT), is dependent on whether or not post-decision evidence accumulation is allowed. When post-decision evidence accumulation is not allowed (model 1 and 3), the confidence response time is equal to the decision response time (eq. 23) and is thus solely defined by the decision model.

$$CT = RT \tag{23}$$

This implies that we use the same value(s) of the evidence accumulator(s) for both the decision-making process and for the confidence judgement. For drift-diffusion models, this implies that confidence relates to the value of the decision boundary at the response time. The decision boundary is a function of time and as a result confidence is only dependent on the response time [14]. For the race model, this is not the case as the losing accumulator is defined as the input variable ($V_c(t)$) which makes confidence depending on the decision response time as well as on the accumulated evidence. When additional evidence accumulation is allowed (model 2 and 4), the evidence accumulation process continues for an addition time interval after the decision is made, the inter-judgement time (τ).

$$CT = \tau + RT \tag{24}$$

As a result, the value(s) of the evidence accumulator(s) used for the decision making differ(s) from the value(s) used for confidence judgements. This makes that confidence depends on the accumulated evidence and confidence response time for both decision models.

To describe how confidence judgements ($Conf$) are based on the input variable ($V_c(t)$), we defined the following equation (eq. 25):

$$Conf = c_{0,dec} + c_{dec}V_c(t) \tag{25}$$

The intercept $c_{0,dec}$ represents the bias of the confidence judgement. The sensitivity parameter $c_{dec} > 0$ describes the ratio between the contribution of the input variable $V_c(t)$ and the confidence bias to the confidence judgement. Both model parameters, describing the bias and the sensitivity, depend on the decision. This relates to our earlier finding that for “go” decisions, confidence judgements are differently biased and influenced in a different manner by perceptual information than “wait” decisions.

Results

We compared the performance of the different confidence models with the use of the RMSE relative to the confidence ratings measured in the experiment. This comparison showed that additional evidence accumulation improved the model (model 2 and 4 in Figure 9 and table 6). The results showed that additional evidence accumulation in particular improved the ability of the models to predict confidence judgements in “wait” decisions. This holds for both the race as well as the dynamic-drift-diffusion based confidence model. This confirmed our hypothesis that allowing for post-decision evidence accumulation improves the performance of the drift-diffusion model (H6c). Our hypothesis that allowing for post-decision evidence accumulation in the race model would not have an effect (H6d) does not hold, as doing so lead to improved performance. This finding substantiates the confidence modelling theory that holds that confidence is based on post-decision evidence accumulation, and is in line with earlier, more fundamental research into confidence [5], [10], [10], [9].

For the two models, we used 1.2 seconds as inter-judgement time (τ, Appendix E). Inter-judgement times larger than this value did not improve or reduce the performance of the confidence models. This indicates that the evidence accumulation only continued for a relative short period of time after the decision is made. To put things in perspective, the average duration of making the turn (time between the first moment of seeing the oncoming vehicle and giving the confidence judgement) was 3.06 seconds (*SD* = 0.498) for “go” decisions and 6.26 seconds (*SD* = 1.092) for “wait” decisions.

Furthermore, we observed that the confidence judgement process is best predicted using the dynamic drift-diffusion decision model (model 2), confirming our hypothesis (H6b). This implies that confidence judgements are better described by one evidence accumulator accounting for both decision outcomes, instead of two independent evidence accumulators for both decisions as described by the race model. The higher performance of the drift-diffusion model thus suggests that the interaction between both decision processes plays a role in confidence judgements. Moreover, this result shows that confidence is related to the positive amount of evidence in favour of the made decision. The research of Zylberberg et al. (2012) [28] did already show that post-decision confidence is more strongly affected by positive evidence in favour of the made decision than by the negative evidence in favour of the alternative decision in a simple perceptual task.

Finally, the fact that the drift-diffusion model that allows for post-decision evidence accumulation functioned well proves our hypothesis that the same evidence accumulation process can be used for modelling confidence judgements as for modelling decision outcomes (H6a).

	DDM	Race model
CT = RT	Model 1: 0.260	Model 3: 0.354
CT = RT + τ	Model 2: 0.0951	Model 4: 0.144

TABLE 6: Performance of the different confidence models: RMSE relative to the confidence ratings measured in the experiment, average over 4 separate models.

4 Discussion and conclusions

The cognitive processes responsible for decisions and confidence judgements are closely related to each other [1], [18], [6], [28]. However, until now, limited applied research had been conducted into decision confidence. Insights about confidence judgements in dynamic real-life tasks can be of value for several applications, for example improving driver assistance systems that consider the confidence of drivers.

In addition, an improved understanding of human cognition can lead to better human machine interaction as argued in previous literature [21] [35]. In this study, we focused on one aspect of human cognition: the decision confidence. We did so in order to obtain a better understanding of the decision confidence of drivers for left-turn gap acceptance decisions. Such an increased understanding may contribute to an improved ability to model (associated) cognitive processes which may increase the safety and efficiency in traffic [22].

The study comprised analyses of the relation between self-reported confidence and the time-to-arrival and the distance gap, the relation between decision response time and the initial throttle operation moment, and the effect of confidence on the velocity profile and on the distance to the centre of the intersection. Moreover, we modelled the cognitive process that describes decision making behaviour, response times and the accompanied confidence judgements.

4.1 Measuring confidence through self-reports

In this study, we measured the amount of confidence participants had in their left-turn gap acceptance/rejection decisions for different time-to-arrival and distance gap conditions with the use of post-decision self-reports. The relation between the present amount of perceptual evidence and observed confidence judgements can be used to validate the use of post-decision self-reports as a measurement method of confidence in a dynamic driving task. In order to validate the measurement method, our results should show similarities with the results of earlier confidence studies on the relation between the present (perceptual) evidence and confidence. Earlier confidence studies describe that a higher quality as well as a higher quantity of evidence towards a decision should result into higher confidence

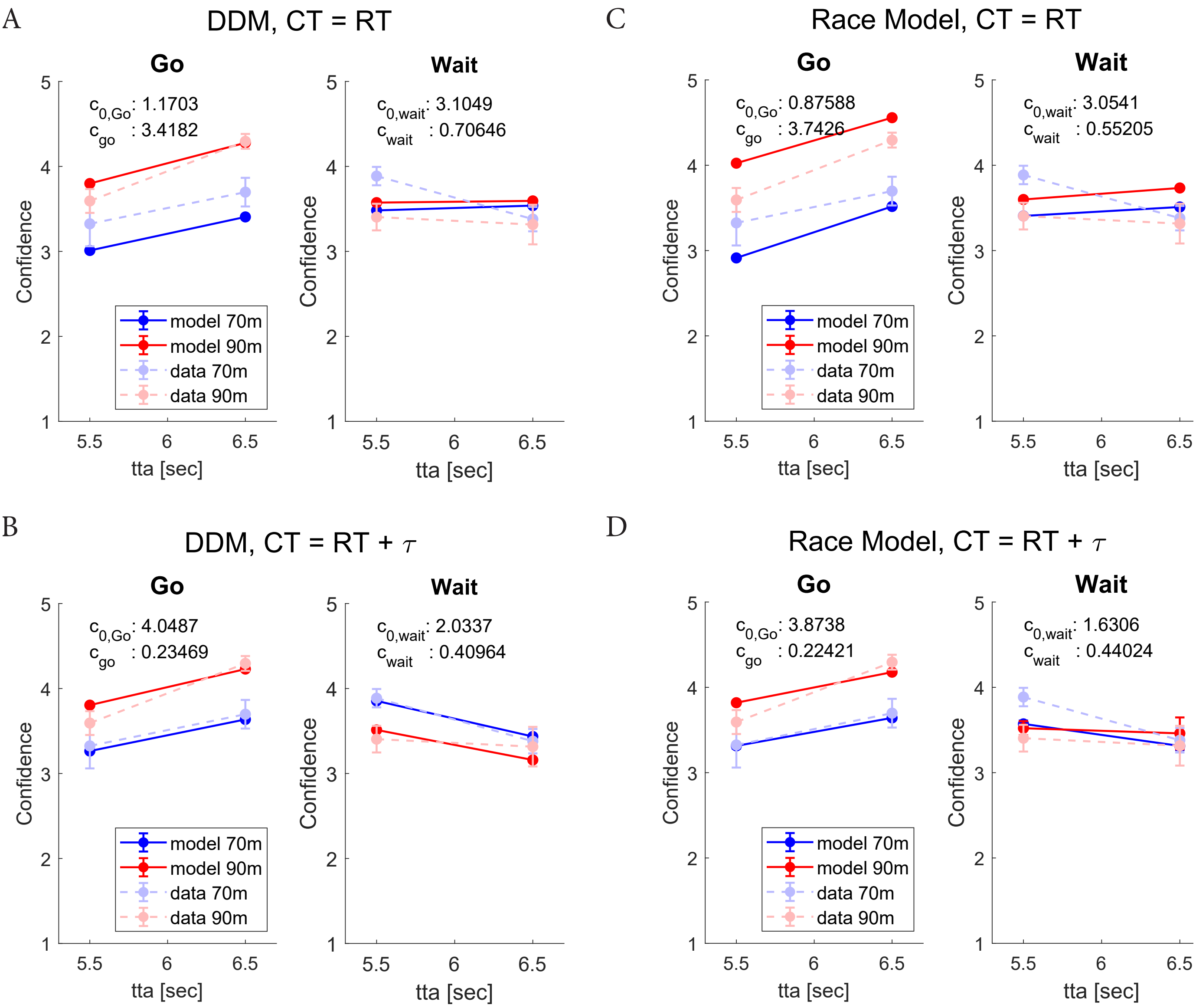

Fig. 9: Comparison of different confidence models, based on different decision models (drift-diffusion and race models) and with or without allowance for additional evidence accumulation. The additional evidence accumulation for the confidence response time (CT) is described by the presence of an additional inter-judgement time (τ). The figures show the mean and 95% CI of the observed (data) and predicted (model) confidence judgments; A) drift-diffusion based confidence model without additional evidence accumulation; B) drift-diffusion based confidence model with additional evidence accumulation of 1.2 seconds (τ = 1.2 s); C) race model without additional evidence accumulation; D) Race model with additional evidence accumulation of 1.2 seconds (τ = 1.2 s).

judgements [14], [5], [9].

For our research, this means that for example, the combination of a large time-to-arrival and a large distance gap should result in more solid confidence judgements than a combination of a large time-to-arrival and a small distance gap. That is because large time-to-arrival and distance conditions advocate for "go" decisions and against "wait" decisions, and vice versa. In practical terms, there is more space available to make a turn at larger time-to-arrival and distance conditions. The data confirmed that in large time-to-arrival and distance conditions, the participants rewarded "go" decisions with a higher confidence judgement than in smaller time-to-arrival and/or distance gap conditions. For "wait" decisions, we observed the reverse relation (Figure 4B). This observed pattern between perceptual information and confidence judgements supports the validity of our approach of measuring confidence as it corresponds to the previous findings in the literature.

4.2 Response time and initial throttle operation moment

We assessed the use of the initial throttle operation moment as a behaviour measure of the response time. The results of the regression analysis showed that the initial throttle operation moment only relates with confidence for "go" decisions (b = -0.341, t = -0.341, p = 4.09e-07), and that this relation is weaker than the relation observed between confidence and the reported response times (b = -0.700, t = -8.49, p = 4.99e-17). These two findings indicate that a) the initial throttle operation moment cannot be used as a measure of the decision response time for wait decisions, and that b) when measured by using the initial throttle operation moment, the decision response time, and thus its relation with confidence, differs from the reported decision response time.

The first finding can be linked to the difference in the execution of the two decisions. That is because for "go" decisions, the execution of the decision starts with entering the intersection by operating the gas throttle, while for "wait" decisions the execution of the decision means remaining in the same position on the road. The second finding, regarding the difference between the reported decision response time and the initial throttle operation moment reported for "go" decisions, can be explained by the tendency of some participants to drive along with the truck and thereby enter the intersection without seeing the oncoming vehicle. This occurred in 47.01% of all analysed "go" decisions.

4.3 Action dynamics: velocity profile and the distance to the centre of the intersection

The analysis of the action dynamic measures shows us that not only the execution of the decision was affected by confidence, but also that it is likely that confidence affects general driving behaviour as we observed in "wait" decisions. The results suggest that people with more confidence have in general a faster driving style and tend to cut corners more often in "wait" decisions.

For both the velocity profile as well as for the distance to the centre of the intersection, we could not observe a relation between confidence on the one hand and RMSD and the deviation from the individual mean on the other. The absence of these relations may suggest that driving performance of individual participants is not (strongly) affected by confidence. However, the absence does not necessarily disprove the effect of confidence on the driving behaviour of a participant, as it could be due to the small sample size. In our study, the number of measurements per participant per condition/choice combination was limited. A more extensive study should be conducted to acquire a solid picture of the effects of confidence on the driving behaviour of an individual participant and to be able to draw conclusions.

In addition, it is important to mention that all the results regarding the relation between confidence and the action dynamic measures in "wait" decisions should be interpreted with caution. This is because these analyses were made with respect to the initial throttle operation moment without taking into consideration the fact that some participants started driving before seeing the oncoming vehicle and thus, before they could make a decision. This results in a distortion in the mean value over time of the action dynamic measures.

4.4 Cognitive modelling

We showed that the cognitive process responsible for confidence judgements can be modelled with the same evidence accumulator that is used to model the decision-making process when accounting for additional evidence accumulation. In order to translate the present evidence expressed by the evidence accumulator into a confidence judgement, we defined two additional confidence model parameters that describe the sensitivity towards the present evidence and the confidence bias. These additional model parameters can potentially be related to the two main factors that determine someone's metacognitive ability: the metacognitive sensitivity representing the reliability or accuracy, and the metacognitive bias which can be described as the calibration [36]. This can be used to explain the observed differences between individuals, something that may be of interest for future research.

4.5 Limitations

This study has certain limitations due to the set-up we used and the defined scope of the research. These limitations concern the experiment set-up, the influence of personal differences, and the aspects of the cognitive process responsible for confidence judgements as accounted for in the confidence models.

Regarding the measuring of confidence, the study indicated that confidence in a dynamic task can be measured with post-decision confidence ratings. The main advantage of using a self-report-based measurement method is that confidence judgements are directly measured [18].However, the disadvantage of this method, which has not been taken into account in this study, is that measurement errors and strategic biases can cause distortion in the measurement [27], [18]. For example, participants may not be willing to give an honest confidence judgement because of pride or perceived expectations they want to live up to.

In our experiment, participants were asked explicitly to report their confidence at the decision moment (decision response time). However, the modelling results suggest that participants were not able to do so. They show that confidence judgements in this task are best explained when taking into account additional evidence accumulation. In other words, self-reports cannot be used to measure confidence at a specific past moment in time as participants arguably (unconsciously) continue with evidence accumulation which affects their retrospective confidence judgement. Further research is required to assess to what extent additional post evidence accumulation affects confidence. Such research could make a comparison between self-reported confidence at the decision response time and self-reported confidence at the end of the turn.

Another limitation of the use of self-reports is that they can only be used in a controlled experimental set-up, as they interrupt the driving task. This disadvantage also applies to the use of button presses in order to measure decision response times, as the button press cannot be directly applied in every driving task.

The results of this study demonstrated that the confidence of a driver in left-turn gap accepting decisions is affected by the time-to-arrival and distance gap of the oncoming vehicle, besides individual differences. We made use of the aggregate data for the regression analysis and allowed for individual differences with random intercepts. However, we did not investigate or account for the causes of these individual differences. Earlier research has shown that individual differences can be caused by differences in self-concept, the level of metacognitive skills, the level of skills and (the amount of) experience [8], [7], [17], [31], [27].

Regarding the modelling of confidence, we limited the scope of the research by only considering dynamic drift-diffusion and race models as basis models for our confidence models. Thereby, the evidence accumulation process was only described by one decision variable accounting for both decisions (dynamic drift-diffusion model) or by two independent decision variables for both decisions (race model). However, other decision models are known in the literature that describe that the evidence accumulators of the two decisions are constructed from a dependent and independent part [28], [29]. Additional research would be needed to investigate whether using these kinds of models leads to an improved confidence model.

Moreover, in order to simplify the models, we assumed for our confidence models that allow for post-decision evidence accumulation that the drift-rate can continuously be described by the same equation. However, it can be questioned to what extent this assumption is true. For instance, it can be argued that the contribution of new perceptual evidence reduces after the decision is made. Follow-up research could explore how the evidence accumulation after the decision has been made can best be described.

Lastly, the values of the model parameters of the decision models as well as the confidence models were found with the use of the "fmincon" function of MATLAB (see Appendix E). This optimisation method searches for a local minimum which makes the outcome dependent on the initial parameter values used. In this study, we chose not to use an optimisation method that is able to find the global minimum, which in hindsight would have been preferable to obtain more certainty about the parameter values.

4.6 Conclusions

The study addressed, among other things, the relation between confidence and the time-to-arrival and distance gap of the oncoming vehicle, the decision response time, and the action dynamics of the performed turn. Moreover, we investigated whether confidence judgements can be captured by a cognitive model which has as a premise that confidence and decisions are based on the same evidence accumulation process.

We can draw the following conclusions regarding these topics:

- the time-to-arrival and distance gap of the oncoming vehicle affect the confidence judgements. In particular, confidence in "go" decisions is positively related to the time-to-arrival and the distance gap. Confidence in "wait" decisions relates negatively to the time-to-arrival and distance gap. Furthermore, confidence in "go" decisions is more strongly affected by the time-to-arrival and distance gap of the oncoming vehicle compared to confidence in "wait" decisions.
- confidence relates negatively to the indicated decision response time regardless of the decision outcome. In addition, the start of the execution of the turn, indicated by the initial throttle operation moment, relates negatively to confidence for "go" decisions.
- the velocity profile during the turn and the distance to the centre of the intersection seem to relate to the confidence of the participant. Participants who were more confident in their decision drove faster through the intersection. For "go" decisions, low confidence judgements appeared to be associated with corner cutting behaviour.
- participants' reported confidence in left-turn gap acceptance/rejection decisions can best be explained by a confidence model based on a drift-diffusion decision model which accounts for post-decision evidence accumulation.

Data Availability

All raw and analysed data as well as analysis and modelling scripts are available at https://osf.io/tgexp/?view_only=4d789506318d4bda92f8e090e3fc9236

Acknowledgment

I would like to thank Dr. Arkady Zgonnikov, Assistant Professor in Human-Robot Interaction at the Department of Cognitive Robotics at the Technical University of Delft, for his supervision during the process of conducting this study. His guidance was marked by useful substantive feedback and practical advice. I thoroughly enjoyed the thought-provoking meetings and being able to build on his earlier studies on modelling driving behaviour.

I am furthermore grateful to the Cognitive Robotics Lab at the TU Delft for providing me with the necessary means to conduct this research.

Lastly, I would like to thank my family, who have consistently supported me throughout my studies.

3

References

REFERENCES

[1] R. Kiani and M. N. Shadlen, "Representation of confidence associated with a decision by neurons in the parietal cortex," *Science*, vol. 324, no. 5928, pp. 759–764, 2009.

[2] P. G. Middlebrooks and M. A. Sommer, "Neuronal correlates of metacognition in primate frontal cortex," *Neuron*, vol. 75, no. 3, pp. 517–530, 2012-08.

[3] L. Aitchison, D. Bang, B. Bahrami, and P. E. Latham, "Doubly bayesian analysis of confidence in perceptual decision-making," *Computational Biology*, vol. 11, no. 10, 2015.

[4] C. R. Fetsch, R. Kiani, and M. N. Shadlen, "Predicting the accuracy of a decision: A neural mechanism of confidence," *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 79, pp. 185–197, 2014.

[5] N. Yeung and C. Summerfield, "Metacognition in human decision-making: confidence and error monitoring," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 367, no. 1594, pp. 1310–1321, 2012.

[6] A. Pouget, J. Drugowitsch, and A. Kepecs, "Confidence and certainty: distinct probabilistic quantities for different goals," *Nature Neuroscience*, vol. 19, no. 3, pp. 366–374, 2016.

[7] M. Händel, A. B. H. de Bruin, and M. Dresel, "Individual differences in local and global metacognitive judgments," *Metacognition and Learning*, vol. 15, no. 1, pp. 51–75, 2020.

[8] Kahneman, "Variants of uncertianty," vol. 11, pp. 143–157, 1982.

[9] P. R. Murphy, I. H. Robertson, S. Harty, and R. G. O'Connell, "Neural evidence accumulation persists after choice to inform metacognitive judgments," *eLife*, vol. 4, 2015.

[10] S. M. Fleming and H. C. Lau, "Self-evaluation of decision-making: A general bayesian framework for metacognitive computation." *Frontiers in Human Neuroscience*, vol. 124, 2017.

[11] M. Rouault, P. Dayan, and S. M. Fleming, "Forming global estimates of self-performance from local confidence," *Nature Communications*, vol. 10, no. 1, p. 1141, 2019.

[12] R. Ratcliff and G. McKoon, "The diffusion decision model: Theory and data for two-choice decision tasks," *Neural Computation*, vol. 20, no. 4, pp. 873–922, 2008.

[13] K. Desender, L. Vermeylen, and T. Verguts, "Dynamic influences on static measures of metacognition," *nature communications*, 2022.

[14] R. Kiani, L. Corthell, and M. N. Shadlen, "Choice certainty is informed by both evidence and decision time," *Neuron*, vol. 84, no. 6, pp. 1329–1342, 2014.

[15] F. Meyniel, D. Schlunegger, and S. Dehaene, "The sense of confidence during probabilistic learning: A normative account," *PLOS Computational Biology*, vol. 11, no. 6, 2015.

[16] V. Liberman, "Local and global judgments of confidence." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 30, no. 3, pp. 729–732, 2004.

[17] E. F. Chua and L. A. Solinger, "Building metamemorial knowledge over time: insights from eye tracking about the bases of feeling-of-knowing and confidence judgments," *Frontiers in Psychology*, vol. 6, 2015.

[18] A. Kepecs and Z. F. Mainen, "A computational framework for the study of confidence in humans and animals," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 367, no. 1594, pp. 1322–1337, 2012.

[19] A. Zgonnikov, A. Kenny, D. Ohora, and K. Wong-Lin, "Capturing decision confidence through response trajectories and willingness to gamble," p. 4, 2017.

[20] T. J. Pleskac and J. R. Busemeyer, "Two-stage dynamic signal detection: A theory of choice, decision time, and confidence." *Psychological Review*, vol. 117, no. 3, pp. 864–901, 2010.

[21] F. Flemisch, J. Kelsch, C. Löper, A. Schieben, and J. Schindler, "Automation spectrum, inner / outer compatibility and other potentially useful human factors concepts for assistance and automation," *Human Factors for assistance and automation*, p. 16, 2008.

[22] G. Markkula, R. Romano, R. Madigan, C. W. Fox, O. T. Giles, and N. Merat, "Models of human decision-making as tools for estimating and optimizing impacts of vehicle automation," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 37, pp. 153–163, 2018-12.

[23] T. Toledo, "Driving behaviour: Models and challenges," *Transport Reviews*, vol. 27, no. 1, pp. 65–84, 2007.

[24] J. Pekkanen, O. Giles, Y. Mun, R. Madigan, T. Daimon, N. Merat, and G. Markkula, "Variable-drift diffusion models of pedestrian road-crossing decisions," *Computational Brain & Behavior*, p. 34, 2021.

[25] M. Schmidt-Daffy, "Prospect balancing theory: Bounded rationality of drivers' speed choice," *Accident Analysis & Prevention*, vol. 63, pp. 49–64, 2014.

[26] A. Zgonnikov, D. Abbink, and G. Markkula, "Should i stay or should i go? cognitive modeling of left-turn gap acceptance decisions in human drivers," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 2022.

[27] E. Paschalidis, C. F. Choudhury, and S. Hess, "Modelling the effects of stress on gap-acceptance decisions combining data from driving simulator and physiological sensors," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 59, pp. 418–435, 2018-11.

[28] A. Zylberberg, P. Barttfeld, and M. Sigman, "The construction of confidence in a perceptual decision," *Frontiers in Integrative Neuroscience*, vol. 6, 2012.

[29] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen, "The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks." *Psychological Review*, vol. 113, no. 4, pp. 700–765, 2006.

[30] R. Ratcliff and F. Tuerlinckx, "Estimating parameters of the diffusion model:approaches to dealing with contaminant reaction times and parameter variability," *Psychonomic Bulletin Review*, vol. 9, no. 3, pp. 438 – 481, 2002.

[31] J. Kruger and D. Dunning, "Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments," p. 14, 1999.

[32] D. A. Schlueter, K. L. Austerschmidt, P. Schulz, T. Beblo, M. Driessen, S. Kreisel, and M. Toepper, "Overestimation of on-road driving performance is associated with reduced driving safety in older drivers," *Accident Analysis and Prevention*, 2023.

[33] ""what were they thinking?": Metacognition and impulsivity play a role in young driver risk-taking," *Journal of Psychiatry and Behavioral Sciences*, 2021.

[34] Z. Boundy-Singer, C. M. Ziemba, and R. L. Goris, "Confidence reflects a noisy decision reliability estimate," *nature human behaviour*, 2022.

[35] T. Schürmann and P. Beckerle, "Personalizing human-agent interaction through cognitive models," *Frontiers in Psychology*, vol. 11, 2020.

[36] S. M. Fleming and H. C. Lau, "How to measure metacognition," *Frontiers in Human Neuroscience*, vol. 8, 2014.

4

Appendices

In this chapter the following appendices are presented:

A. Figures, lme Models

B. Random Effects

C. Action Dynamics, lme

D. Correlations

E. Optimisation of Model Parameters

F. Excluded Left-Turns

Appendix A - Figures, LME models

Decision behaviour

$$Pr_{go} \sim distance + TTA + (1|ID)$$

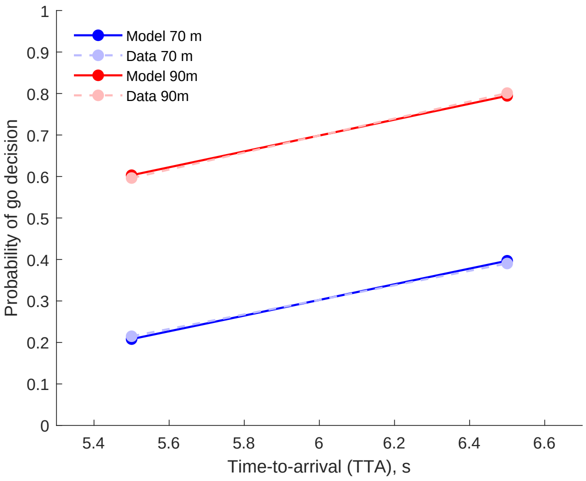

Fig. 10: Linear mixed effects model of the probability of "go" decisions

Decision response time

$$RT \sim distance * decision + TTA * decision + (decision|ID)$$

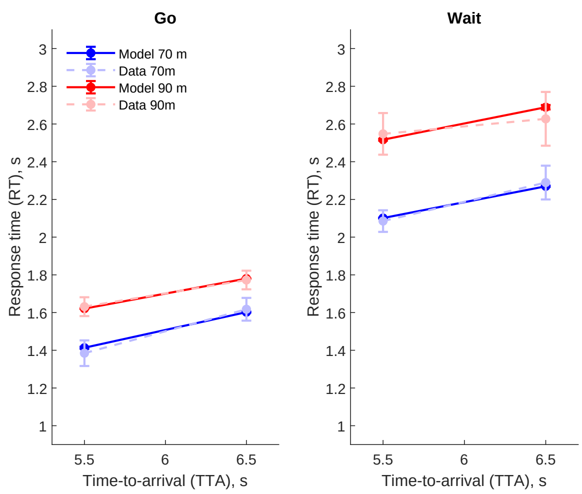

Fig. 11: Linear mixed effects model of the decision response time

Confidence

$$Conf \sim RT*decision + distance*decision + TTA*decision + (decision|ID)$$

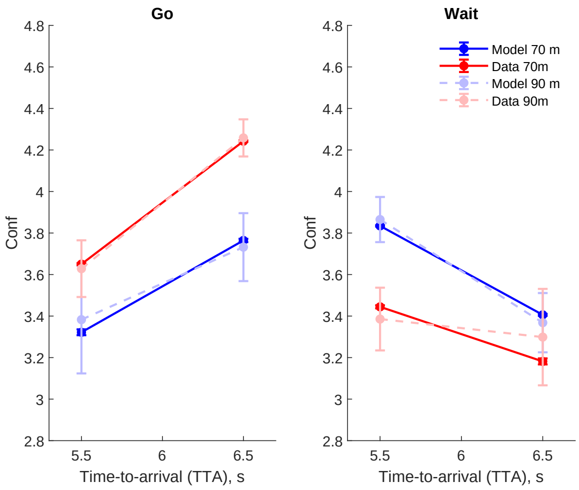

Fig. 12: Linear mixed effects model of confidence based on the decision response time, distance, time-to-arrival and decision

$$Conf \sim Thr_{int}*decision + distance*decision + TTA*decision + (decision|ID)$$

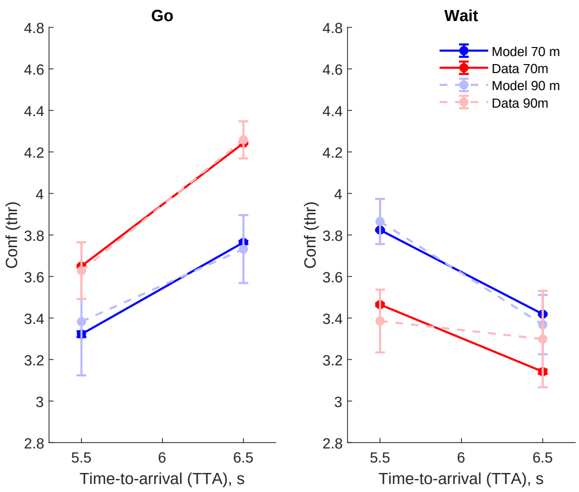

Fig. 13: Linear mixed effects model of confidence based on the initial throttle operation moment, distence, time-to-arrival and decision

Appendix B - Random Effects

In order to account for the effects of individual differences between participants, all regression models have a random effects term (1| ID) or (decision | ID).

Decision behaviour

We investigated the influence of the distance and time-to-arrival conditions by using the following linear mixed effect model: $Pr_{go} \sim distance + TTA + (1|ID)$ The random effects describing the influence of individual differences has a standard deviation 0.3978 of and 95% confidence interval of [0.3839, 0.4122]. For ten of the seventeen participants, an additional (α = 0.05) intercept was found (table 7).

Estimate	pValue
0.13796	0.034472
-0.24048	0.00023872
0.13491	0.037971
-0.21088	0.0013205
0.42177	1.2186e-10
-0.26047	6.7498e-05
-0.15967	0.01408
0.1417	0.033149
-0.42763	1.1548e-09
0.34158	1.7393e-07

TABLE 7: Random effect coefficient of the intercept for the decision probability, capturing individual differences.

Response time

The effect of the time-to-arrival and distance conditions was investigated with the following linear mixed effects model, $RT \sim distance * decision + TTA * decision + (decision|ID)$. The model accounts for individual differences with the random effects, consisting of a random intercept and a random decision slope (additional term for "wait" decision) (Table 8).

	Std	95% CI interval
Intercept	0.25022	0.16431, 0.38105
Wait decision	0.34325	0.22103, 0.53306

TABLE 8: Random effects for the linear mixed effects model of the response time (RT), describing the influence of individual differences for different decision outcomes.

We found a random effect (α = 0.05) of the intercept for six participants (Table 9) and an additional random effect for the wait decision for seven participants (Table 10).

Estimate	pValue
-0.32815	0.0051979
0.37553	0.0020354
-0.27047	0.021912
0.74597	1.1226e-09
-0.29145	0.012199
0.7309	0.0013217

TABLE 9: RT: Random effect coefficient of the intercept, capturing individual differences.

Estimate	pValue
0.51067	1.2307e-09
-0.5062	1.2217e-07
0.20642	0.015385
-0.23606	0.0018559
-0.25012	0.0022344
0.18155	0.032157
-0.25454	0.00091624

TABLE 10: RT: Random effect coefficient of the "wait" decision, describing individual differences

Confidence (RT)

The effects of the response time, time-to-arrival and distance on confidence were investigated with the following linear mixed effects model: $Conf \sim RT * decision + TTA * decision + distance * decision + (decision|ID)$ The model accounts for individual differences with the random effects, consisting of a random intercept and a random decision slope (additional term for "wait" decision) (see table 11).

	Std	95% CI interval
Intercept	0.51158	0.34788, 0.75232
Wait decision	0.48767	0.31956, 0.74422

TABLE 11: Random effects for the linear mixed effects model of confidence (response time), describing the influence of individual differences for different decision outcomes.

We found a random effect (α = 0.05) of the intercept for five participants (table 12) and an additional random effect for the wait decision for nine participants (table 13).

Estimate	pValue
-0.42034	0.038575
-0.56695	0.004159
0.63415	0.0014371
1.0984	6.1472e-08
-0.61043	0.009719

TABLE 12: Confidence (RT): Random effect coefficient of the intercept, capturing individual differences.

Estimate	pValue
0.66059	0.00014618
-0.49468	0.0026842
-0.58171	0.00071363
0.35686	0.039236
0.97672	2.7953e-10
-0.67047	5.7792e-05
-0.51608	0.010064
-0.6033	0.00097385
0.54496	0.0014252

TABLE 13: Confidence (RT): Random effect coefficient of the "wait" decision, describing individual differences

Confidence (RT)

The effects of the response time, time-to-arrival and distance on confidence were investigated with the following linear mixed effects model: $Conf \sim Thr_{int} * decision + TTA * decision + distance * decision + (decision|ID)$ The model accounts for individual differences with the random effects, consisting of a random intercept and a random decision slope (additional term for "wait" decision) (see table 14).

	Std	95% CI interval
Intercept	0.45881	0.31474, 0.66882
Wait decision	0.49196	0.32783, 0.73825

TABLE 14: Random effects for the linear mixed effects model of confidence (throttle), describing the influence of individual differences for different decision outcome.

We found a random effect (α = 0.05) of the intercept for four participants (table 15) and an additional random effect for the wait decision for nine participants (table 16).

Estimate	pValue
-0.6084	0.003291
0.75733	0.00024556
1.1477	3.4713e-08
-0.62541	0.010011

TABLE 15: Confidence (CT): Random effect coefficient of the intercept, capturing individual differences.

Estimate	pValue
0.51859	0.0054811
-0.57188	0.00063699
0.34505	0.039499
0.94403	4.177e-10
-0.54311	0.00075716
-0.58911	0.0027873
-0.4385	0.013682
-0.3781	0.016327
0.53283	0.001304

TABLE 16: Confidence (CT): Random effect coefficient of the "wait" decision, describing individual differences

APPENDIX C - ACTION DYNAMICS, LME

The effect of confidence on the two measures of action dynamics, velocity profile and distance to centre of intersection, was investigated with the linear mixed effects models who is defined in general by: $Metric \sim Conf * decision + (1|ID)$.

For both measures four different metrics were taken into account: maximum/minimum value, deviation from the individual mean, deviation from the group mean and the RMSD.

Velocity

Maximum velocity

	Estimate	Std. Error	t-score	pValue
Intercept	9.9337	0.44954	22.097	3.8434e-94
Confidence	0.097675	0.080369	1.2153	0.22442
Decision wait	-0.26592	0.40842	-0.65111	0.51508
Decision wait: conf.	-0.13321	0.10769	-1.237	0.21629

TABLE 17: Fixed coefficients of the linear mixed effects model describing the relation between confidence and the maximum value of the velocity profile.

	Std	95% CI interval
Intercept	1.3148	0.92872, 1.8612

TABLE 18: Random effects of the linear mixed effects model describing the relation between confidence and the maximum value of the velocity profile.

Deviation from the individual mean

This model has a Hessian matrix with NaNs or Infs, which indicates that the model has more covariance parameters than supported by the data.

	Estimate	Std. Error	t-score	pValue
Intercept	-0.18554	0.16697	-1.1112	0.26666
Confidence	0.047861	0.04165	1.1491	0.25069
Decision wait	0.051122	0.22532	0.22689	0.82054
Decision wait: conf.	-0.01007	0.058309	-0.1727	0.86291

TABLE 19: Fixed coefficients of the linear mixed effects model describing the relation between confidence and the deviation from the individual mean of the distance to the velocity profile.

	Std	95% CI interval
Intercept	2.6211e-16 NaN, NaN	

TABLE 20: Random effects of the linear mixed effects model describing the relation between confidence and the deviation from the individual mean of the distance to the velocity profile.

Deviation from the group mean

	Estimate	Std. Error	t-score	pValue
Intercept	-0.66223	0.29141	-2.2725	0.023197
Confidence	0.10768	0.048922	2.2011	0.027879
Decision wait	0.75535	0.24851	3.0396	0.0024095
Decision wait: conf.	-0.092263	0.065537	-1.4078	0.15939

TABLE 21: Fixed coefficients of the linear mixed effects model describing the relation between confidence and the deviation from the group mean of the velocity profile.

	Std	95% CI interval
Intercept	0.90075	0.63756, 1.2726

TABLE 22: Random effects of the linear mixed effects model describing the relation between confidence and the deviation from the group mean of the velocity profile.

RMSD

	Estimate	Std. Error	t-score	pValue
Intercept	0.98234	0.16598	5.9184	4.0042e-09
Confidence	-0.060233	0.035916	-1.6771	0.093731
Decision wait	0.5175	0.1831	2.8263	0.0047707
Decision wait: conf.	0.0314	0.048229	0.65106	0.51511

TABLE 23: Fixed coefficients of the linear mixed effects model describing the relation between confidence and the RMSD of the distance to the velocity profile.

	Std	95% CI interval
Intercept	Intercept	0.35621 0.24773, 0.5122

TABLE 24: Random effects of the linear mixed effects model describing the relation between confidence and the RMSD of the distance to the velocity profile.

Distance to the centre of the intersection

Minimum distance

	Estimate	Std. Error	t-score	pValue
Intercept	2.9981	0.44352	6.7598	5.1678e-11
Confidence	-0.23771	0.099158	-2.3973	0.016996
Decision wait	-2.0648	0.63656	-3.2437	0.0012833
Decision wait: conf.	0.58373	0.17745	3.2896	0.0010964

TABLE 25: Fixed coefficients of the linear mixed effects model describing the relation between confidence and the minimum value of the distance to the centre of the intersection.

	Std	95% CI interval
Intercept	0.5162	0.33623, 0.79252

TABLE 26: Random effects of the linear mixed effects model describing the relation between confidence and the minimum value of the distance to the centre of the intersection.

Deviation from the individual mean

This model has a Hessian matrix with NaNs or Infs, which indicates that the model has more covariance parameters than supported by the data.

	Estimate	Std. Error	t-score	pValue
Intercept	0.39735	0.2677	1.4843	0.13854
Confidence	-0.093318	0.061783	-1.5104	0.13176
Decision wait	-0.68784	0.42704	-1.6107	0.10807
Decision wait: conf.	0.17391	0.11438	1.5204	0.12922

TABLE 27: Fixed coefficients of the linear mixed effects model describing the relation between confidence and the deviation from the individual mean of the distance to the centre of the intersection.

	Std	95% CI interval
Intercept	0.87233	0.81299, 0.936

TABLE 28: Random effects of the linear mixed effects model describing the relation between confidence and the deviation from the individual mean of the distance to the centre of the intersection.

Deviation from the group mean

	Estimate	Std. Error	t-score	pValue
Intercept	0.6499	0.34919	1.8612	0.063485
Confidence	-0.16022	0.07554	-2.121	0.034561
Decision wait	-1.0952	0.48139	-2.2752	0.023449
Decision wait: conf.	0.31137	0.13551	2.2978	0.022111

TABLE 29: Fixed coefficients of the linear mixed effects model describing the relation between confidence and the deviation from the group mean of the distance to the centre of the intersection.

	Std	95% CI interval
Intercept	0.53949	0.36685, 0.79338

TABLE 30: Random effects of the linear mixed effects model describing the relation between confidence and the deviation from the group mean of the distance to the centre of the intersection.

RMSD

	Estimate	Std. Error	t-score	pValue
Intercept	1.5326	0.23493	6.5236	2.1758e-10
Confidence	-0.099826	0.052983	-1.8841	0.06031
Decision wait	-0.31994	0.34164	-0.93646	0.34963
Decision wait: conf.	0.12392	0.094796	1.3072	0.19193

TABLE 31: Fixed coefficients of the linear mixed effects model describing the relation between confidence and the deviation from the group mean of the distance to the centre of the intersection.

	Std	95% CI interval
Intercept	0.24169	0.15303, 0.38172

TABLE 32: Random effects of the linear mixed effects model describing the relation between confidence and the RMSD of the distance to the centre of the intersection.

Appendix D - Correlations

In the research, several potential correlations were investigated:

1) Decision response time – Initial throttle operation moment
2) Confidence – Decision response time
3) Confidence – Initial throttle operation moment

Decision response time - Initial throttle operation moment

Decision	Correlation coefficient (r)	pValue
All	0.2332	2.0647e-20
Go	0.2818	1.6121e-15
Wait	-0.2409	1.4624e-11

TABLE 33: Correlation coefficients between decision response time (RT) and the initial throttle operation moment.

Confidence - Decision response time

Decision	Correlation coefficient (r)	pValue
All	-0.2729	1.2874e-27
Go	-0.2114	3.1298e-09
Wait	-0.2490	2.8442e-12

TABLE 34: Correlation coefficients between confidence and the decision response time (RT).

TTA\distance	70 m	90 m
5.5 seconds	-0.49 (p=4.13e-06)	-0.34 (p=2.5e-07)
6.5 seconds	-0.22 (p=0.0062)	-0.35 (p=2.5e-10)

TABLE 35: Correlation coefficients between RT and confidence judgements – in each traffic condition for ”Go” decisions.

TTA\distance	70 m	90 m
5.5 seconds	-0.23 (p=6.4e-05)	-0.24 (p=0.0028)
6.5 seconds	-0.19 (p=0.0038)	-0.13 (p=0.27)

TABLE 36: Correlation coefficients between RT and confidence judgements – in each traffic condition for ”wait” decisions.

Confidence - Initial throttle operation moment

Decision	Correlation coefficient (r)	pValue
All	-0.2224	1.1645e-18
Go	-0.2735	1.1195e-14
Wait	-0.1591	9.8448e-06

TABLE 37: Correlation coefficients between confidence and the initial throttle operation moment.

TTA\distance	70 m	90 m
5.5 seconds	-0.44 (p=3.3e-05)	-0.46 (p=4.0e-13)
6.5 seconds	-0.37 (p=2.5e-06)	-0.27 (p=1.5e-06)

TABLE 38: Correlation coefficients between initial throttle operation moment and confidence judgements – in each traffic condition for ”Go” decisions.

TTA\distance	70 m	90 m
5.5 seconds	-0.055 (p=0.35)	-0.089 (p=0.28)
6.5 seconds	-0.21 (p=0.0014)	-0.32 (p=0.0040)

TABLE 39: Correlation coefficients between initial throttle operation moment and confidence judgements – in each traffic condition for ”wait” decisions.

APPENDIX E - OPTIMISATION OF MODEL PARAMETERS

For optimisation of the model parameters, we made use of the “fmincon” function of MATLAB R2020a. The function searches for the set of parameters which result in a (local) minimum of the loss function of the performance of the model.

Lossfunctions

For the loss functions of the performance of the models, we made use of the weight least squares (WLS) and the root mean square error (RMSE) for respectively the decision and confidence models.

Decision models

To train the decision models, we made use of a newly defined loss function build up from the WLS of the model prediction. The WLS was calculated with the use of vincentized distributions [30]. For each of the four different conditions present in the experiment, the WLS was calculated, using two separate terms describing the WLS for “go” and “wait” decisions. The total sum of the WLS over all conditions was used as loss function.

Confidence models

For the confidence models, the root mean square error (RMSE) was used as loss function, which can be described by the following equations:

$$RMSE = \sqrt{\frac{\sum(Conf_{pr} - Conf_{expr})^2}{N_conf}}$$

$$Conf_i = [\mu_{conf,go,i}, \mu_{conf_wait,i}]$$

$\mu_{conf,decision,i}$ contains the mean confidence values for the four conditions in the specified decision of the prediction or the experiment. N_{conf} is the number of measuring points, so $N_{conf} = 4 * 2 = 8$.

Parameters

The “fmincon” function starts searching for an optimal set of parameters round an initial set of parameters, which must be defined in advance.

Decision model

The initial set of parameters used to train the decision models was obtained with the use of the optimisation code “03_fit_model.py” which accompanied our baseline decision model [26]. This optimisation was focused on the prediction of left-turn gap acceptance decision behaviour by a dynamic drift diffusion model only accounting for “go”-decisions. This set of initial parameters was as a result not able to predict wait decisions accurately.

The newly defined loss function was used in combination with the found initial parameters to find the parameter set describing both “go” and “wait” decisions for a dynamic drift-diffusion model and for a race model.

	WLS	α	beta	b_0	k	$\mu_n d$	$\sigma_n d$	θ_{crit}
Initial	3.08	0.985	0.101	1.14	0.357	1.40	0.117	13.7
DDM	1.53	1.12	0.109	1.41	0.396	1.51	0.140	14.0

TABLE 40: Initial trained model (”go”) and DDM models (”go” and ”wait”): a) comparison of performance of the decision models, mean WLS over 10 models, b) parameter values found through optimisation.

For the race model, the drift-rate parameter (α) is assumed to be decision dependent. Besides, we investigated the effect on the performance of the model by decision dependent initial boundaries (b_0) and critical values of the generalised gap (θ_{crit}). The present non-decision time, the sensitivity of the boundary and the definition of the generalised gap were defined as being decision independent, i.e., similar for both evidence accumulators. In order to test the optimisation method used, we also investigated the effect on the performance of the model by defining all model parameters decision dependent. If all model parameters are decision dependent, this should result in the best performance (lowest WLS).

The results of the parameter optimisation showed that the race model with decision dependent drift-rate parameters and the race model with both decision dependent drift-rate parameters as well as critical values of the generalised gap resulted the highest improvement of the performance (lowest WLS-values). These results are based on the initial parameter set of the drift-diffusion model which was not able to predict both “go” and “wait” decisions. Therefore, we additionally investigated whether the use of the optimal parameters for the drift-diffusion model could improve the optimisation of the race model with decision dependent drift-rate parameters and the race model with both decision dependent drift-rate parameters as well as decision dependent critical values of the generalised gap. Thereby, we found that the race model is able to describe the decision behaviour with the decision dependent drift-rate parameter and the critical generalised gap value trained with the use of the drift-diffusion model parameters describing “go” and “wait” decisions (table 41 and table 41, Figure 15).

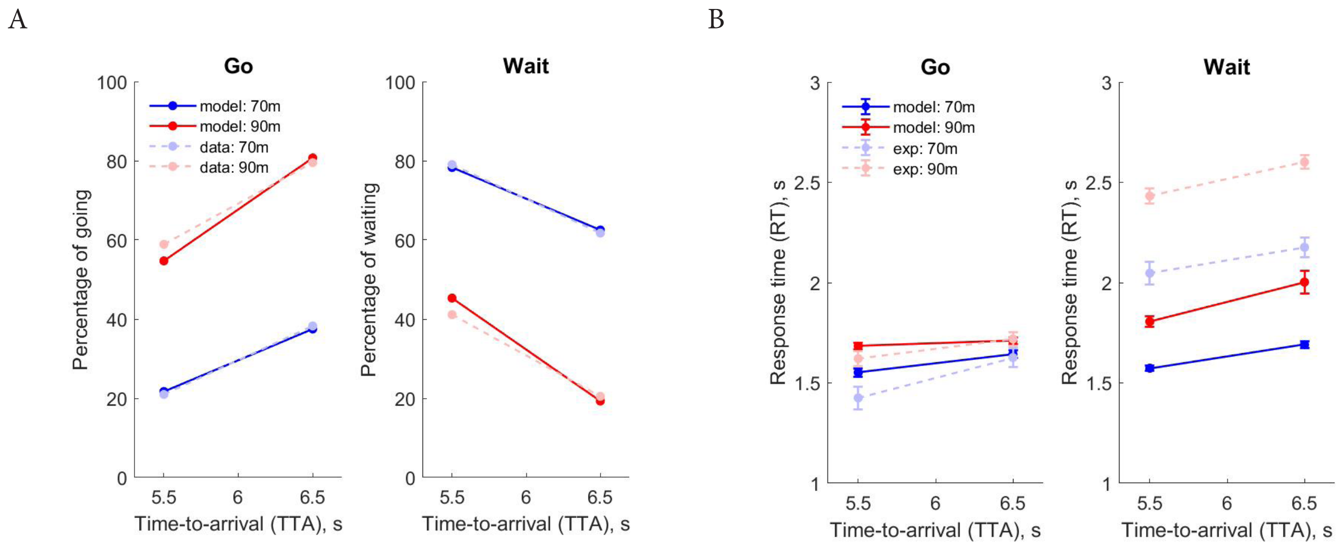

Fig. 14: Performance of drift-diffusion decision model, trained for ”go” decisions (initial parameters).

	WLS	α	β	b_0	k	μ_{ND}	σ_{ND}	θ_{crit}
race (α)	1.67	Go: 0.838 Wait: 0.901	0.117	1.49	0.38	1.50	0.326	14.2
race (α&b_0)	2.62	Go: 1.52 Wait:1.32	0.0938	Go: 0.504 Wait:0.500	0.101	1.85	0.207	13.3
race (α&θ_{crit})	1.73	Go: 1.03 Wait: 1.13	0.112	1.21	0.341	1.51	0.123	Go: 13.8 Wait: 13.7
race (all)	0.48	Go: 0.810 Wait:1.47	Go:0.0988 Wait:0.102	Go:0.966 Wait:1.17	Go:0.394 Wait:0.390	Go:1.35 Wait:1.84	Go:0.131 Wait:0.124	Go:13.8 Wait:13.8

TABLE 41: Race model mean performance (WLS) over 10 models, trained with initial parameters. Influence of different decision dependent parameters.

	WLS	α	β	k	b_0	μ_{ND}	σ_{ND}	θ_{crit}
race (α)	1.73	Go: 1.05 Wait: 1.24	0.105	1.24	0.413	1.53	0.129	13.7
race (α & θ_{crit})	1.33	Go: 1.06 Wait: 1.05	0.108	1.21	0.454	1.54	0.128	Go: 14.0 Wait: 13.5

TABLE 42: Race model mean performance (WLS) over 10 models, trained with initial parameters. Influence of different decision dependent parameters.

Confidence models

The confidence models presented in this paper were built on the decision model by adding four additional free parameters, describing sensitivity and bias for both decisions, and by potentially allowing for additional evidence accumulation.

Sensitivity and bias parameters
The initial values of the additional confidence model parameters describing sensitivity and the bias were found using a linear regression model of the relation between confidence and the input value (V_c). However, the sensitivity parameters can only be positive and linear regression models cannot be constrained. As a result, we trained the confidence models with the use of the “fmincon” function, thereby including the constraint that the sensitivity parameters should be positive.

Inter-judgement times
Two of the four confidence models accounted for additional evidence accumulation after the decision was made. The time evidence continues to accumulate, inter-judgement time (τ), could not be measured during the experiment and thus had to be modelled. In order to do so, the effect of the different values of the inter-judgement time on the RMSE was calculated manually, over a time scale from zero to 2.5 seconds with a time interval of 0.01 seconds (Figure 16). The RMSE of the confidence models with a specific inter-judgment time, used in the comparison of the effect of different inter-judgement times, was defined as the mean RMSE over 5 models. The model parameters of the confidence model were fitted separately for each investigated inter-judgment time with the use of a linear regression model, where all sensitivity parameters smaller or equal to zero were defined as zero due to the constraint that the sensitivity parameters must be positive.

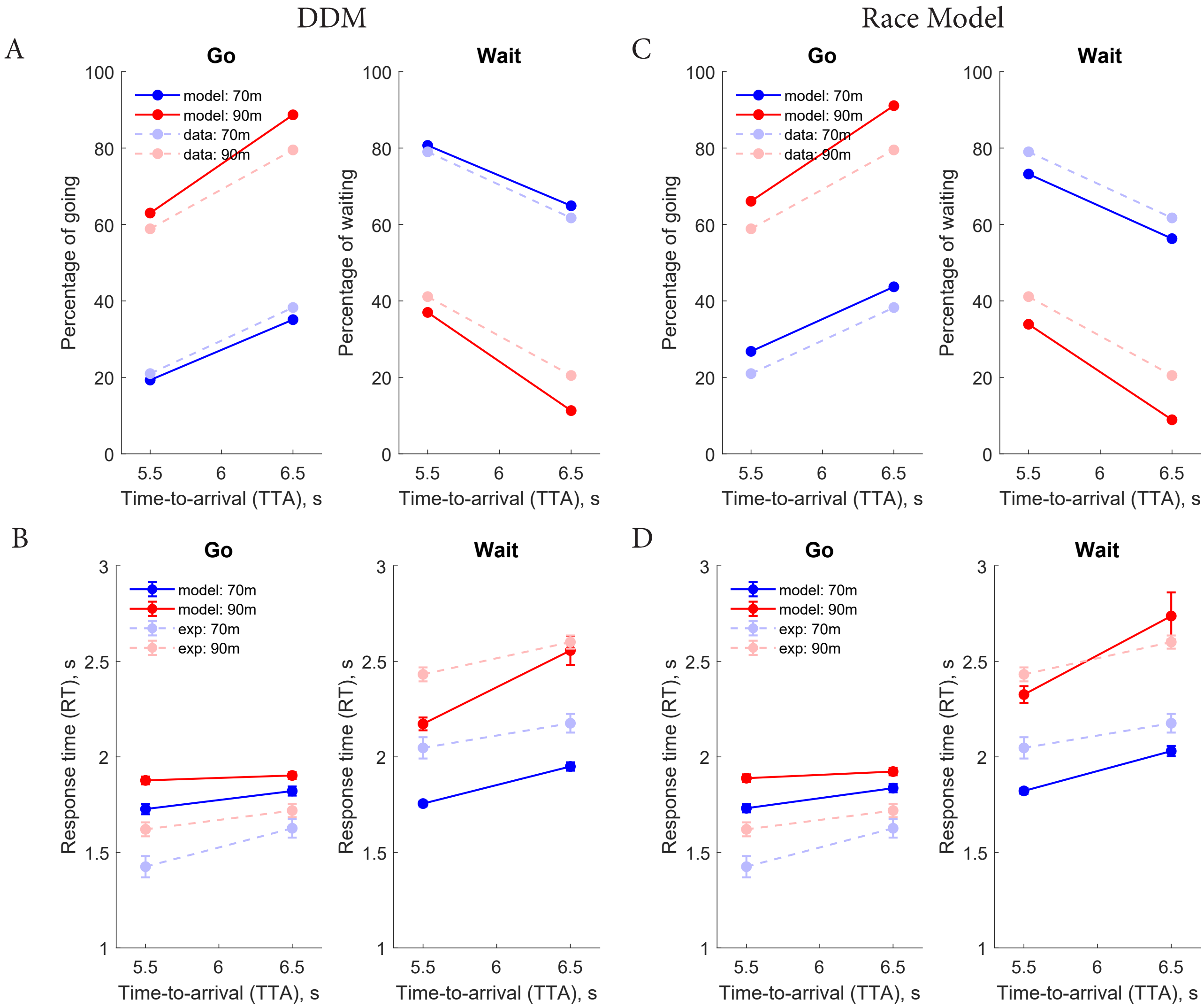

Fig. 15: Performance of drift-diffusion model (DDM) and race model with trained parameters.

The results show that the performance of the models is affected by the inter-judgement time until the inter-judgement time reaches a value of approximately 1.2 seconds after which the performance remains constant (Figure 16). At the inter-judgement time of 1.2 seconds, both models perform optimal. This finding suggests that the model is over parameterised or fails to correctly describe the evidence accumulation process for a longer period of time after the decision is made. The model does for example not account for the moment in time in which the decision is made or in which the oncoming vehicle has passed.

Model parameters
An overview of the finial confidence model parameters used in the confidence models presented in the paper.

	Bias parameter	Sensitivity parameter
Model 1: DDM, $\tau = 0s$	Go: 1.17 Wait: 3.10	Go: 3.42 Wait: 0.707
Model 2: DDM, $\tau = 1.2s$	Go: 4.05 Wait: 2.03	Go: 0.23 Wait: 0.410
Model 3: Race model, $\tau = 0s$	Go: 0.876 Wait: 3.74	Go: 3.74 Wait: 0.552
Model 4: Race model, $\tau = 1.2s$	Go: 3.87 Wait: 1.63	Go: 0.22 Wait: 0.44

TABLE 43: Confidence model parameters, used for the different presented confidence models.

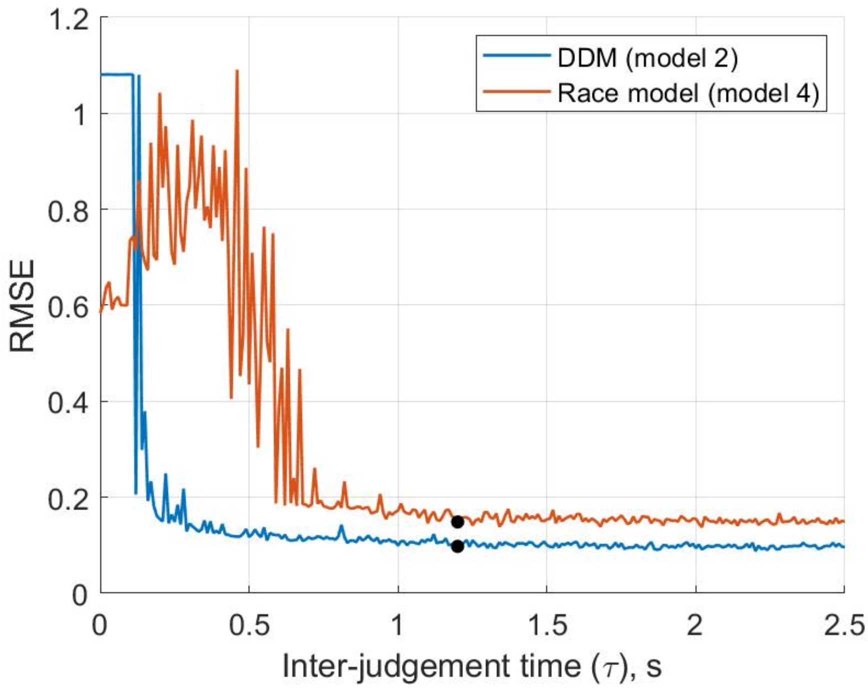

Fig. 16: Effect of different inter-judgement times on the performance (RMSE) of the confidence model based on the drift-diffusion decision model (model 2) and the confidence model based on the race model (model 4). The black point indicates the inter-judgement time used for the final confidence models (1.2 seconds)

APPENDIX F - EXCLUDED LEFT-TURNS

The data analysis in this study was restricted to the left-turn trials in which the indicated decision was actually conducted. We excluded all changes of mind, situations in which the participant carried out a different decision than indicated, as well as the trials in which the participants did not indicate their decision (by failing to press the designated button).

In 3.4 % of all the left-turn decisions, changes of mind were present, of which in 83.9% of the cases the participant indicated a “go” decision and performed a “wait” decision. In 2.2% of all left-turn decisions, no button presses were present. In 83.33% of the cases in which participates failed to press the button, they performed a “go” decision.

5

Form of consent

INFORMED CONSENT

Participant Information

You are invited to participate in a research study titled “Driving behaviour at intersections”, for participation it is required to be in possession of a valid driver's license. This study is carried out by Floor Bontje (Bio-mechanical Engineering Master Student) as part of a master's thesis research conducted at the Delft University of Technology.

The purpose of this research is to get an insight into driving behaviour when making a turn at an unmarked four-way intersection, and it will take approximately two hours to be completed including breaks and some test rounds to get familiar with the driving simulator. The data obtained will be used for a thesis report and potentially also for other scientific publications. In this experiment you will be asked to perform a driving task in a virtual driving simulator. You will drive through a virtual town following navigation prompts. On the route you will have to cross several four-way intersections and other traffic may be present. You are asked to follow the Dutch traffic rules and to drive similarly as you normally would do. In addition to the driving task you will be asked to rate your own driving behaviour.

To the best of our ability the data obtained in this study will remain confidential. We will minimize risks related to identifying participants by anonymising the data; optionally collected email addresses will be decoupled from the rest of the data. All personal data unnecessary for the research (email addresses and names) will be stored on a secure TU Delft storage location. You as participant have the right to request access to and rectify or erase personal data. Your name will not be explicitly linked to a participant number. The experiment software will provide a randomly generated identifier at the beginning. We will ask you to write this identifier down for your own records. Your identifier and name are never stored at the same location. Hence, your data cannot be traced back to you. The identifier that you will keep a record of is needed to remove your data if you decide that you wish to withdraw from the study after the data collection is finished. If you would like to remove all your data from the study, the researchers will ask you to share your identifier, which is then used to remove your data. All research data including deidentified demographic information and deidentified demographic information will be (publicly) shared at the end of the research. The published data will not be traceable to you.

During the experiment mental fatigue and a reduction of concentration could occur. In order to reduce these effects, you can take a break at any time during the experiment. In order to compensate for the time and effort put into participation a €20 VVV-gift card will be handed out at the end experiment. Your participation in this study is entirely voluntary **and you can withdraw at any time.** You are free to omit any questions.

In order to participate in this research, it is necessary that you give your informed written consent. By signing the explicit consent you are indicating that you understand the nature of this research and your role in it and that you agree to participate in the research. The written consent will be stored at a protected location on the departure of cognitive robotics at the TU Delft.

Thank you in advance for your possible cooperation. If there are any questions or complaints about the research, please contact the corresponding researcher (Floor Bontje) or responsible researcher (Arkady Zgonnikov) by email.

Floor Bontje (Corresponding Researcher): F.bontje@**tudelft.nl
Arkady Zgonnikov (Responsible Researcher): A.Zgonnikov@**tudelft.nl

Explicit Consent

PLEASE TICK THE APPROPRIATE BOXES	Yes	No
I have read and understood the study information. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	☐	☐
I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without giving a reason.	☐	☐
I understand that taking part in the study involves: • Participating in a fixed-base driver simulator experiment • Answering questions regarding my performance in the experiment	☐	☐
I understand that I will be compensated for my participation by receiving a gift card.	☐	☐
I understand that personal information collected about me that can be used to identify me, such as e.g. my name and email address, will not be shared beyond the research team.	☐	☐
I understand that the (identifiable) personal data I provide will be destroyed after the research is finished.	☐	☐
I understand that the following steps will be taken to minimise the threat of a data breach, and protect my identity in the event of such a breach: anonymisation of the data and a secure data storage with limited access for identifiable personal data (name, email address).	☐	☐
I understand that the data I produced during the experiment (vehicle telemetry including speed, acceleration, decision reaction time, and demographic information including age, sex, and driving experience) will be anonymised and publicly shared online.	☐	☐
I understand that after the research study the de-identified information I provide will be used for a thesis report and potentially also for other scientific publications.	☐	☐
I give permission for the de-identified data produced during the experiment will be archived in Gitlab and OSF repository so it can be used for future research and learning.	☐	☐
I understand that access to this repository is publicly open.	☐	☐

Signatures

____________________________ ___________________________ _________

Name of participant Signature Date

I, as researcher, have provided the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

Floor Bontje

Researcher name Signature Date

Study contact details for further information: Floor Bontje,