Janssen, C.P., Praetorius, L., Borst, J.P. (in press 2024). PREDICTOR: a tool to predict the timing of the take-over response process in semi-automated driving. *Transportation Research Interdisciplinary Perspectives*

PREDICTOR: a tool to predict the timing of the take-over response process in semi-automated driving

Christian P. Janssen[1*]

Leonard Praetorius[2]

Jelmer P. Borst[2]

1:

Experimental Psychology & Helmholtz Institute

Utrecht University

Heidelberglaan 1 , room H0.52

3584 CS Utrecht

The Netherlands

*: Corresponding author: c.p.janssen@uu.nl

2:

Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence

University of Groningen

Bernoulliborg, room 320

Nijenborgh 9

9747 AG Groningen

The Netherlands

Declarations of interest: none

Funding acknowledgement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Credit statement:

Christian Janssen: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Writing – review and editing. **Leonard Praetorius:** Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review and editing. **Jelmer Borst:** Methodology, Software, Supervision, Writing – review and editing.

ABSTRACT

This paper presents PREDICTOR (PREDICting Take-Over Response time): an interactive open-source research software tool to predict the timing of various stages of a transition of control, or take-over, in semi-automated driving. Although previous work has investigated extensively what factors affect the minimum time needed for a successful take-over by the driver, less is known about how specific stages within the take-over process are affected by those factors. PREDICTOR applies a theoretical framework that describes the take-over process as interruption handling through a series of stages. It then ties this theory to a database that summarizes results from previous take-over studies. PREDICTOR can be used to interactively predict through simulation how specific human factors (e.g., alert modality, alert onset time) impact four distinct stages of the take-over response process. The tool simulates and visualizes expected reaction time distributions for each stage of the take-over process. The use of distributions also highlights the likelihood of an accident – as long responses ("outliers") are quantifiable. Moreover, it can help understand at which stage drivers might take relatively longer or shorter, and which stages are most impacted by a specific factor (e.g., alert modality). PREDICTOR also allows users to add their own data, and to define their own dependent variables for analysis. As a tool that allows exploration of various scenarios, PREDICTOR can aid in the prediction and analysis of potential future accidents.

Keyword: Transitions of Control; Take-over Request; Cognitive Model; Interruption Handling

1 Introduction

Although the willingness to adopt automated vehicles is not yet high (e.g., Adnan, 2024; Arowolo et al., 2024), and there are legal (Sever & Contissa, 2024) and other concerns about them (e.g., are these technologies even developed for the right reasons? see, Almlöf, 2024), there is a need to investigate how the use of (semi-) automated vehicles impacts human behavior – the human factors (Gerber et al., 2023). Investigating human factors is particularly urgent, given that the functionality of (semi-) automated vehicles has increased over the last few decades. Current commercial systems can do more than merely assist the driver occasionally and instead are able to control aspects of the drive for prolonged periods of time, during which the human driver monitors the system and the environment (i.e., SAE levels 2 to 3, SAE International, 2014). If either the car requests, or the human desires, a transition of control can take place in which the other party (e.g., human or car) takes over main control (McCall et al., 2019; Mirnig et al., 2017). If the transition of control process is initiated by the car, this is often referred to as a take-over request. Such a take-over request can for example be given in response to bad weather, road construction works, or other events that the car cannot handle on its own.

A meta-review by Zhang, de Winter, Varotto, Happee, and Martens (2019) analyzed 129 take-over studies (i.e., SAE level 2 and up). The results showed that the large majority of studies reported situations with fast response times ($M_{response\ time}$ = 2.7 s; $SD_{response\ time}$ = 1.45s; only one study above 10 s, with 19.79 s). However, as automation improves, and the car is able to take control of the ride for longer segments of time, it is possible that drivers perform other tasks during the drive (e.g., Hancock, 2013; Janssen et al., 2019). Drivers already perform such non-driving activities during non-automated manual driving (e.g., Dingus et al., 2016; Klauer et al., 2014) and they have expressed interest to perform a wide variety of non-driving activities during highly automated driving (e.g., Pfleging et al., 2016; Stevens et al., 2019).

Therefore, one might wonder whether studies that describe take-over as a simple, fast response process are merely a ‘convenience’ instead of a true description of likely human behavior (de Winter et al., 2021). The conception with which engineers have designed the system (often: for an immediate take-over) might not align with actual human behavior (for a broader discussion on misconceptions, see also Walker et al., 2020). Specifically, in the context in which drivers perform other tasks during the automated drive and where automation becomes more reliable, human drivers might not immediately give up on their original task (Janssen et al., 2019), even though current commercially available vehicles are designed for immediate take-over. Various research groups have therefore started to investigate the psychological process of take-over requests in more detail, with an emphasis on understanding how an ‘interruption’ by an alert is handled when one was also attending other (non-driving related) tasks or activities (e.g., Ch, 2023; Ch et al., 2024; Gerber et al., 2020, 2024; Janssen et al., 2019; Nagaraju et al., 2021; Naujoks et al., 2017; Wintersberger et al., 2018). One long-term goal in such efforts is to have accurate, detailed theory- and model-based predictions of human take-over responses that can inform the design of future interfaces.

State-of-the-art computational cognitive modeling tools for human-automated vehicle interaction are not yet that far (e.g., see reviews in Janssen et al., 2024; Lorenz et al., 2024); and current theories and models can not yet be applied in interactive software (which has been done for “regular” driving in the past Salvucci, 2009). Yet, theoretical advances have been made in understanding the psychological underpinnings of take-over requests. Specifically, Janssen, Iqbal, Kun, and Donker (2019) proposed that the process of a take-over should not be considered as a rapid task switch in which the driver always responds to an alert immediately. Instead, they argue that it can be considered a gradual process of interruption handling.

There are many detailed theories of interruption handling from other fields (e.g., Altmann & Trafton, 2002; Boehm-Davis & Remington, 2009; Borst et al., 2015; Couffe & Michael, 2017; Salvucci & Taatgen, 2008, 2011; Sanderson & Grundgeiger, 2015). Although they differ in the details, all these theories propose that when attention is moved between two discrete tasks, there is a process in which specific stages can be identified. In the case of switching between a non-driving related task (e.g., reading e-mails) and driving during a transition of control, at least 10 stages can be identified in the process of going from not driving, to taking over control of a car, to eventually handing control back to the car, and to resume the non-driving related task again (Janssen et al., 2019).

Although the theoretical framework of Janssen and colleagues provides *qualitative* insights about these stages, it does not provide *quantitative* predictions of the timing of these various stages. However, such quantitative predictions are needed to get to the long-term goal of integrated, embedded, predictive computational cognitive models (Janssen et al., 2024). There are two reasons why such quantitative predictions do not yet exist. First, most current studies on take-over response times do not yet explicitly couple their results to stages of the interruption framework. Second, quantitative predictions from studies outside of the driving domain are hard to couple directly to driving scenarios.

The contribution of the current paper is to address this gap and quantify predictions for four stages of the interruption framework. Specifically, the paper provides an open-source interactive tool, called PREDICTOR (PREDICt Take-Over Response time), that can be used to predict how various human factors impact the duration of four of the stages from the interruption model proposed by Janssen et al (2019). These stages describe the process between initial alert onset and the first physical act of initiating control by the human driver.

PREDICTOR allows users to add studies and study results, and to label them for factors that they want to analyze -- aligning with the perspective that visualization is essential to understand human driver behavior (Donmez et al., 2023)

PREDICTOR can aid researchers by allowing them to simulate and visualize (cf. Donmez et al., 2023) expected distributions of response times for each stage of the take-over process. Through visualization of data (cf. Donmez et al., 2023), it allows quick inspection of patterns in the literature. This can then be compared to the results of researchers' own studies, and their data can also be integrated in the tool to test and visualize how both align. In section 5, examples are provided of initial theoretical insights that the tool gives.

PREDICTOR can also aid practitioners and designers of vehicle interfaces. As the consideration of take-overs as interruptions has so far mostly been the focus of theoretical research (Janssen et al., 2019), it is a crucial way to move the underlying ideas forward to practice. In particular, the ability to predict distributions allows consideration of more than mean response, for example, the likelihood of slow responses (which are often treated as "outliers" in empirical studies). In a driving safety context these slow responses are important, as they can help estimate the likelihood that a specific response deadline is not made (e.g., whether people respond within the current conventional 5-7 s after alert onset, Gold et al., 2013). Such delayed responses in turn might be an indicator of accident likelihood. It might help designers in considering how their interface might impact not only the average user, but also these more extremes.

The remainder of this paper is structured as follows. First, the underlying theoretical framework is described in more detail. Second, follows a description of four human factors that might impact the onset of stages on the interruption framework and of which there are sufficient studies to include them in PREDICTOR as predicting variables. Third, the structure of the tool PREDICTOR is introduced. Fourth, PREDICTOR's functionality is illustrated through five critical tests. Finally, implications and limitations are discussed, including how users can further apply PREDICTOR in their own research: by adding studies and by testing other variables of interest.

2 Transitions of control as interruptions

Transitions of control can be considered as a process of multiple stages (Janssen et al., 2019). Figure 1 highlights the 5 stages where a human user takes physical control in response to an alert issued by the vehicle. Below follows a brief description of each stage; for even more detail see (Janssen et al., 2019).

Starting point is the assumption that the car is initially controlling the drive, and that the human is optionally working on a non-driving related task (stage 0). Although in SAE level 3 (and lower) the human's task is to monitor the vehicle and the traffic surroundings, in practice, humans might not always do this. Humans already perform non-driving related tasks during manual, non-automated driving (e.g., Dingus et al., 2016; Klauer et al., 2014), and might reclaim some of their time for work and play if automation levels increase (Kun et al., 2016). Such behavior might particularly be expected if the automation functionality of the car becomes more reliable and advanced and if human intervention is infrequently needed. Many empirical studies of take-over therefore include distracted driving scenarios. For example, in the 520 experiments (129 studies) that Zhang et al. (2019) meta-reviewed, 377 (72.5%) contained a non-driving related task.

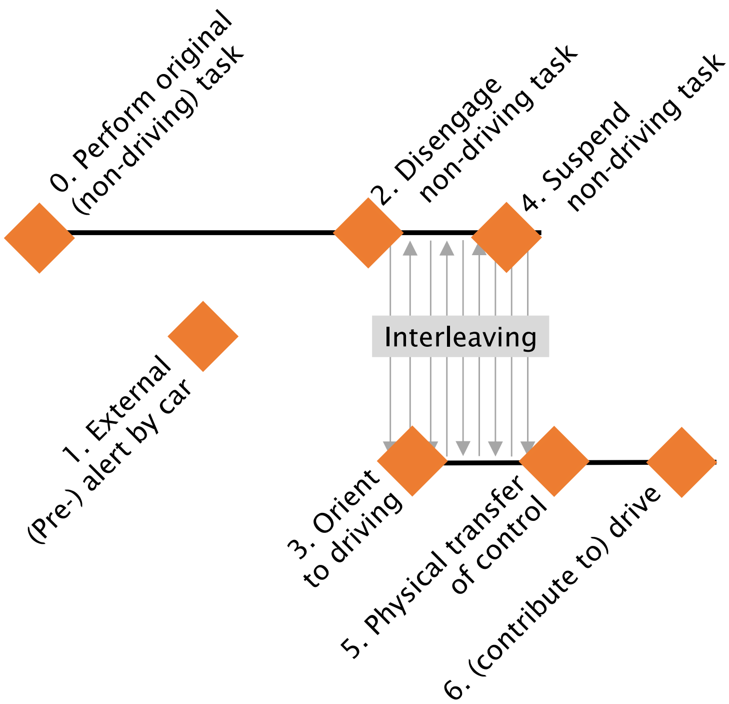

Figure 1: Take-over of control according to the interruption model of Janssen et al. (2019). While the automated vehicle drivers, a human might be working on a non-driving related task (stage 0). After an external alert (stage 1), they then briefly disengage from the non-driving task (stage 2) to orient to driving (stage 3). After optionally going back and forth between non-driving task and orienting to driving (interleaving stage), they fully suspend the original non-driving task (stage 4) to physically initiate the transfer of control (stage 5) and then (contribute to the) drive (stage 6).

The vast majority of studies uses external alerts to warn a driver of a take-over (B. Zhang et al., 2019: 486/520 studies, or 93%). Therefore, the next stage is the issuance of an alert (stage 1). The stages after stage 1 are where the framework from Janssen et al. differs from most of the current literature, as the next stage is not yet the driver’s physical response (e.g., grabbing the wheel; stage 5), but rather: disengaging from the non-driving related task (stage 2). This is defined as the first moment that a person is not visually and/or manually engaged with a non-driving related task. For example, the first moment they look away from their non-driving related activity (e.g., phone, in-car interface, co-driver) in response to an alert.

Stage 3 is the first moment that the driver orients to the driving task and context, but at which time they do not necessarily take control of the wheel. Depending on various factors such as the time that is left before a response by the driver is required, drivers might spend some time on interleaving their attention between further orienting to the driving task and wrapping up their performance on their original, non-driving related task (for recent observations of such interleaving, see e.g., Ch, 2023;Ch et al, 2024; Large et al., 2019; Nagaraju et al., 2021).

Eventually, the driver will suspend working on the non-driving related task (stage 4) and then perform a first physical action to take control of the vehicle (stage 5). This action might take various forms, such as pressing a button to “take control”, pressing the brake with the feet, making a steering correction with the hands, or perhaps even a vocal command. The physical transfer of control stage (stage 5) is then followed by a period during which there is more dedicated time for driving (stage 6).

Although there are some empirical studies that have tested subsets of these detailed steps (e.g., Ch, 2023; Ch et al., 2024; Gerber et al., 2020; Nagaraju et al., 2021; Naujoks et al., 2017), these are few compared to the large bulk of empirical studies on take-over requests (e.g., 520 experiments /129 studies in B. Zhang et al., 2019), which typically measure the interval between stage 1 and 5 without considering substages. The contribution of this paper is to provide a tool, PREDICTOR, that can help to predict possible durations of the intermediate stages. The tool has a database that contains detailed information from many papers (largely based on Zhang et al’s meta-review) and then predicts potential durations of steps.

3 Influencing Human Factors on Stage Onset Times

Multiple Human Factors can impact the onset and duration of the various stages of transitions of control (Janssen et al., 2019). PREDICTOR currently allows users to test the impact of four (human) factors that are reported more frequently in the literature (e.g., B. Zhang et al., 2019), and therefore have sufficient studies to inform simulations: (1) Alert Onset Time, (2) Alert Modality, (3) input modality of non-driving related task, and (4) output modality of non-driving related task. While for each of these factors there have been many studies that report the effect on overall take-over time, the contribution of PREDICTOR is that it allows researchers to simulate and quantify the effect of the factors on all four stages of the take-over process, as well as predicting how factors might interact with each other.

3.1 Alert Onset Time

Alert onset time is the time window between an alert and a future critical event where the user's assistance (take-over) is needed (e.g., aid with navigation roadworks, or in settings where lane markings are missing). Alert onset time is important to consider in PREDICTOR, as it is positively correlated with take-over response time (B. Zhang et al., 2019).

Most studies use a relatively short alert onset (e.g., 5-8 seconds cf., Gold et al., 2013; Mok et al., 2017) and associated fast responses by the user. For example, in Zhang et al.'s meta-review (2019), 93% of the studies report a response time by users shorter than 5 s. More recent studies have introduced earlier onsets to allow drivers more time to finish what they were working on before and to allow sufficient time to orient to driving and gain situational awareness (Borojeni et al., 2018; Ch, 2023; Nagaraju et al., 2021; Van Der Heiden et al., 2017). These studies observed relatively longer response times. During such longer intervals, it is also likely that the other stages of Janssen et al's interruption framework (2019) are observed. For example, a recent longitudinal study observed more frequent interleaving as drivers became more familiar with the vehicle (Large et al., 2019). PREDICTOR allows researchers to explore more thoroughly how variations in alert onset time might affect all four stages of the take-over process.

3.2 Alert Modality

External alerts can be presented in different modalities. The most common alert modalities in automated driving studies are auditory, visual, and bi-modal (audio and visual) alerts (B. Zhang et al., 2019). Although previous studies have already observed that alert modality impacts the moment of physical response (i.e., stage 5 in Figure 1), PREDICTOR allows users to now also explore systematically how modality impacts other stages of the take-over process.

3.3 Input Modality of Non-driving related task

The characteristics of the non-driving related task might also impact take-over performance and timing. As the potential activities that people want to do in automated vehicles varies widely (e.g., from reading e-mails, to playing games, to sleeping, Pfleging et al., 2016; Stevens et al., 2019), there is no single way to cluster all non-driving related tasks.

As a start, PREDICTOR allows clustering of non-driving related tasks based on their general input and output modality. The input modality refers to the modality in which a non-driving related task sends information to the user (e.g., visual, auditory, or tactile presentation). Input modality is a crucial factor in multiple resource models of cognition (Salvucci & Taatgen, 2008, 2011; Wickens, 2002, 2008). The models hypothesize that as two tasks (e.g., driving and non-driving) overlap more in their input modalities, there is more interference and potential performance degradation of at least one task. It has not yet been studied how modality affects each of the four stages.

3.4 Output Modality of Non-driving related task

Output modality refers to the modality in which the user processes information or interacts with the interface of the non-driving related tasks (i.e., manual, cognitive, vocal). Output modality is loosely comparable to the term "processing stages" in Wickens' multiple resource model (2002, 2008). However, output modality places more emphasis on the modality of interacting with a potential interface.

The labeling of input and output modalities is based on how the original studies described the non-driving related task. If the non-driving related task was not explained in more detail, the descriptions from other studies using the same (or a similar) task were considered. For a non-driving related task to be categorized as requiring a cognitive modality (which – one could argue – any activity does to some degree), it had to be introduced explicitly as cognitive in the study. Cognitive was categorized under output (and not input) modality, as, similar to the other output modalities (e.g., manual, vocal), some active processing of the participant is needed.

4 PREDICTOR

4.1 Aims for PREDICTOR and general overview

PREDICTOR can quantitatively simulate predictions of the time interval of each stage of the take-over process. Three research questions motivated the design of PREDICTOR:

(1) How does the driver go through the four stages of the take-over process? PREDICTOR simulates, quantifies, and visualizes likely distributions of the time interval for each stage.

(2) How is the distribution of timing of the four stages affected by the human factors: alert onset, alert modality, and input/output modality of the non-driving related task? PREDICTOR allows users to compare multiple simulations based on different levels of these human factors (e.g., auditory versus visual alerts), and to assess how they affect the distributions of times. Note that PREDICTOR allows users to add new studies and factors to the included database, to explore their impact.

(3) How likely is the transition of control going to succeed in time for the driver to react to a critical event, and how is the success rate affected by different factors? PREDICTOR estimates the cumulative distribution of successful completion of each stage. For stage 5 (i.e., the physical transfer of control) this functionality is particularly interesting, as it can help to estimate the proportion of drivers unable to take back control in time to react appropriately to a critical event.

PREDICTOR was implemented using R Shiny (R version 3.6.1; Shiny version 1.3.2). The software can be used via https://predictor-tool.shinyapps.io/PREDICTOR/ . The code can also be downloaded from the Supplementary Materials, and an additional more detailed description of its structure can be found in the Supplementary Materials ‘Details on Model’. Figure 2 shows the high-level structure of PREDICTOR. It contains an extensive database of results from 265 experimental take-over studies (largely from B. Zhang et al., 2019) (section 4.2). Through the graphical user interface (section 4.3), the user can select specific parameters (e.g., specific alert modality, alert onset time) for which simulations need to be made. Under the hood, PREDICTOR then computes summary statistics (section 4.4), which are then used to simulate 10,000 distributions for each stage (section 4.5). The output is then presented in the user interface in various ways (e.g., distributions of intervals for each stage, success rate of transitions; see section 4.6).

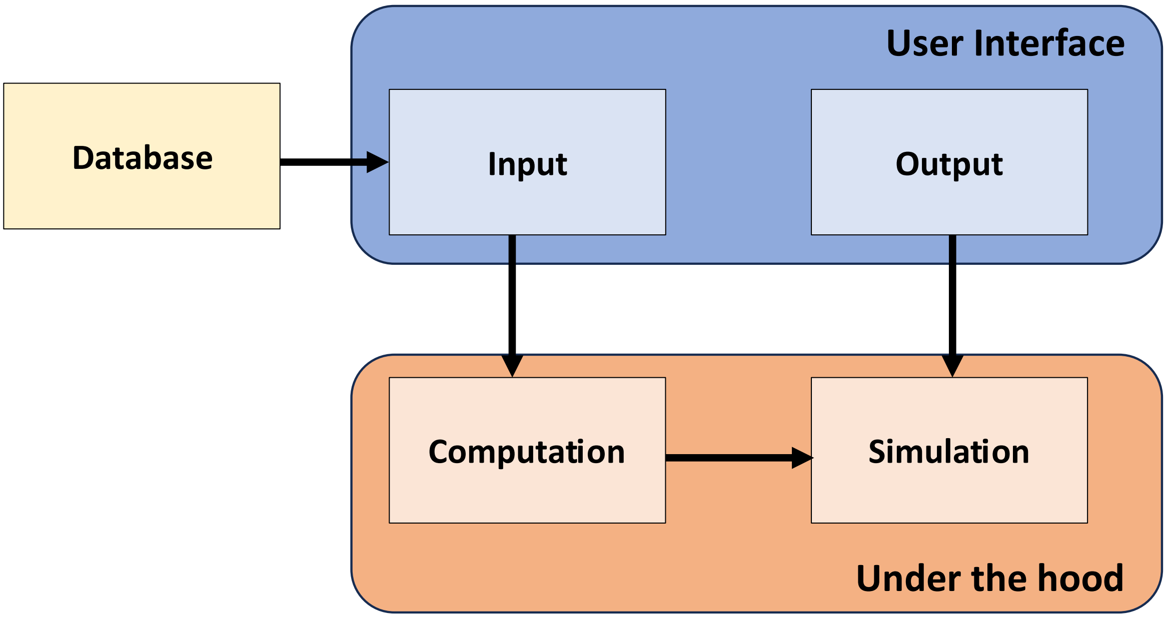

Figure 2: Structure of PREDICTOR. PREDICTOR relies on a database of studies. Within the user interface, a user can select what parameters they want to focus on. Under the hood, the necessary computations are then done, which provide input to simulations of outcomes. These are then presented as output in the user interface.

4.2 Database

PREDICTOR relies on a database that summarizes findings from previous studies of take-over processes. To demonstrate the functionality of PREDICTOR, it was included with data that was summarized in a meta-review by Zhang et al. (2019). As the focus of Zhang and colleagues was not on exploring all stages of the take-over process, all studies from their review (129 studies from 119 records) were re-examined to match data to specific stages (2-

4) from the interruption framework (Janssen et al., 2019). Studies had to fulfill the following criteria to be included:

1. The study had to involve a take-over process: a transition of control from automated driving (or the simulation thereof, e.g. using 'Wizard of Oz' ; Mok et al., 2015) to the human driver (i.e., SAE Level 3 or above).
2. At least one experimental condition had to also involve a non-driving related task while automation was enabled. Otherwise, stage 2 could not be identified.
3. The take-over had to be initiated by an external take-over request (i.e., an alert) to which the subject had to react. That is, no self-interruptions or driver-initiated transitions of control (Gerber et al., 2020).
4. The study had to report the mean and standard deviation (or another metric from which these could be calculated) for at least one of the stages 2-5.
5. The study had to be written in English, German, Dutch, or French (so the authors could understand the study details).

After collecting all the data from each study matching those criteria, experimental groups were further removed if they did not match all criteria (e.g., if a condition did not have a non-driving related task).

This resulted in a database of 265 experimental groups from 67 studies. In total, data from 2,591 participants were retrieved. Out of these studies, 64 studies came from the meta-analysis by Zhang et al. (2019). As among these studies there were few studies that reported stages 2, 3, and 4, three other studies of which the authors knew that these stages were measured were also included in the database (Lotz et al., 2019; Van Der Heiden et al., 2017; Yoon et al., 2019).

While stage 5 (physical transfer of control) is discussed in all but one study, information for other stages is reported less frequently, especially for stage 2 (Disengage), and stage 4 (Suspend non-driving related task). This might be due to methodological reasons: stage 2, 3, and 4 require a measure of visual attention, such as eye-tracking, that not all studies might have available or report. More conceptually, driver distraction research is traditionally mostly focused on the effect that a non-driving related task has on (degradation of) driver performance (or reaction time to the alert: stage 5), and not on performance of the non-driving related task itself (e.g., stages 2 and 4). This contrasts with the interruption framework, where all stages are of interest. Users of PREDICTOR can also add their own data to the database of PREDICTOR.

4.3 User Interface: Input

The interface of PREDICTOR can be used to analyze what the impact is of specific human factors (as defined in section 3: alert onset time, Alert modality, NDRT input/output modality) on the time distribution of each stage of the take-over process. Figure 3 shows the basic interface (for details on output see section 4.6). Depending on the parameter, users either can select from pre-defined values (e.g., for modality: visual, auditory, tactile, or a combination) or define an exact point or range (e.g., for alert onset time: 5s, or 5-8s). Based on the choices for each parameter, the model then filters out entries in the database that do not match the selection.

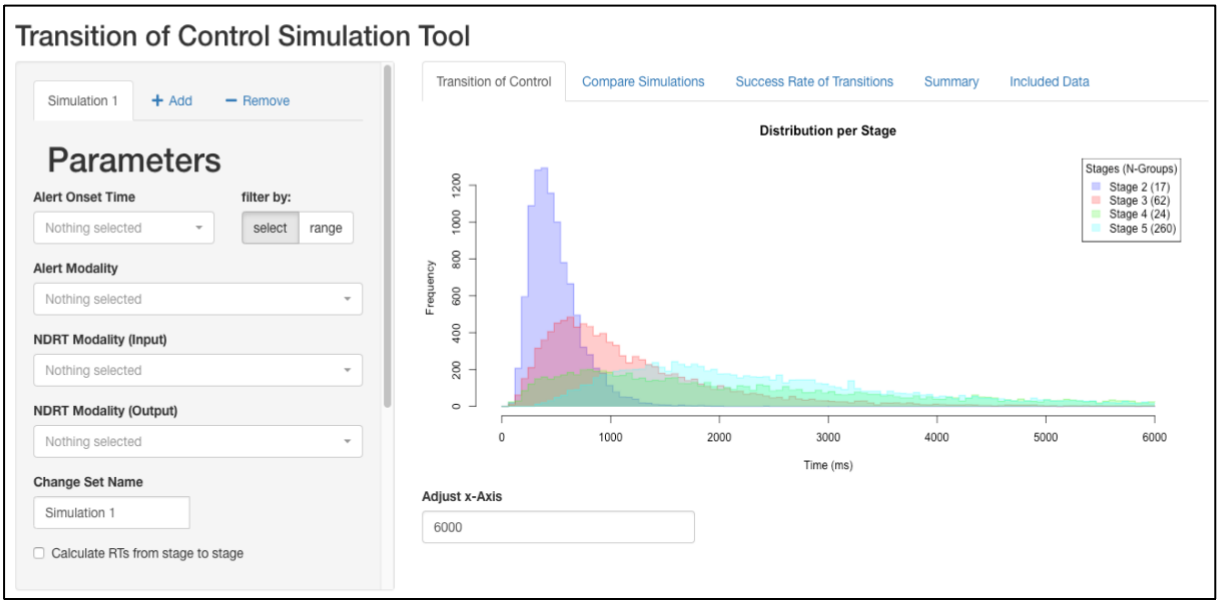

Figure 3: UI at model initiation. The input is presented on the left side of the window and the output on the right side. At initiation, the transition of control considering the entire database is presented.

Users can choose how they want the stage onset times to be determined. In the literature, the stage onset times are typically reported in relation to the alert onset (stage 1) rather than relative to the preceding stage. By default, PREDICTOR uses the values as reported in the literature to simulate each stage of the transition of control, thus simulating each stage independently (i.e., the interval between stage 1 and the other stages). If the "Calculate RTs from stage to stage" tick box is selected, however, for each stage the model also takes into account the results from its preceding stage. For a further discussion of the implications of this option, please see the Supplementary Materials ('Details on Model').

4.4 Computation

Based on the user's input, under the hood, PREDICTOR makes calculations to transform the data from the subset of the database in a way that it can be given as input to the simulation (section 4.5). To estimate variability and compare studies, the simulations use standard deviations. To calculate the means and standard deviations across multiple entries in the database (i.e., multiple samples), it normalizes the values from each study based on the number of participants that took part (see Supplementary Materials file 'Details on Model' for details).

4.5 Simulation

With the calculations done, PREDICTOR can now simulate data based on the selected datasets. For the simulation, it was assumed that the eventual distribution adheres to two assumptions that are common in reaction time studies: (1) The distribution is positively skewed, and (2) there are no negative reaction times.

The first assumption was positive skewness. This is often observed in reaction time data (for a general perspective see Ratcliff, 1993; for observations in take-over studies see B. Zhang et al., 2019). A 'long tail' occurs when a small but significant portion of the reaction times are much larger than the average, while very few, if any, are much shorter than average. In a take-over setting, this translates to a scenario where most drivers have similar, relatively short take-over times, while a small number requires more time to successfully take-over control. Although small in numbers, these outliers can pose a significant risk to road safety, as they might not take back control in time to accurately handle a critical event. Even though these are often treated as outliers and removed before analysis, considering them in PREDICTOR is thus especially important, as they play a major role in the success of the alert.

The second assumption was that there are no negative reaction times. Negative reaction times would suggest that the take-over process was initiated by the driver before the alert onset. While this can happen in practice if the take-over is self-initiated by the driver (e.g., Gerber et al., 2020), PREDICTOR focuses explicitly on the transition of control initiated by an external alert. For simulations that calculate the time distributions as intervals between stages, an additional assumption must be made to exclude negative values, namely that the stages of the transition of control always occur in the same order. While this is perhaps inevitably true between some stages (i.e., after alert onset the driver cannot orient towards the road (Stage 3) before first disengaging from the non-driving related task (Stage2)), exceptions to this rule might still occur. For example, in (Petermeijer et al., 2017) the participants took back control of the vehicle (Stage 5) on average 30ms before orienting towards the road (Stage 3).

To implement these assumptions, a log-normal distribution was used as underlying distribution from which a model can sample (see Supplementary Materials 'Details on Model' for details on how and why exactly). The model uses the log-normal distribution to sample what the data might look like if 10,000 observations were made that come from this log-normal distribution. Specifically, for each stage of the interruption process, 10,000 samples are taken from the estimated log-normal distribution. The resulting values are then visualized in the user interface.

4.6 User Interface: Output

The right side of the user interface (

Figure **3**) presents PREDICTOR's output from the model simulations. Different aspects of the timing of stages in the take-over process are visualized in five panels: Transition of Control, Compare Simulations, Success Rate of Transitions, Summary, and Included Data. This allows scrutinous visual inspection of the data (cf. Donmez et al., 2023). Below we describe these panels. More details can be found in the Supplementary Materials (‘Details on Model’ file).

The transition of control panel (see Figure 3) is presented at model initiation. It visualizes per stage (2-5), what the simulated distribution is of times relative to onset of the alert (stage 1 in Janssen et al., 2019). The legend shows per stage how many studies provided input to the simulations. If no data is available to simulate a stage (given the selected parameter choices), the stage is not visualized. The simulation results in Figure 3 capture two expected patterns: (1) as the stages progress, the time progresses (i.e., the distribution of stage 5 is more to the right compared to that of stage 2), (2) later stages have a wider distribution – as for each stage there might be some variability (or noise), which accumulates over stages.

The ‘Compare Simulations’ panel displays four plots with the distributions of response times per stages (2, 3, 4, 5). If multiple simulations have been created, all simulations can be plotted simultaneously in each panel, with distinct transparent colors.

The ‘Success Rate of Transitions’ panel (Figure 4) displays information about the predicted percentage of successful take-overs in relation to the timing / deadline of a critical event. At the top of the panel, a cumulative plot is displayed, including each stage of the currently selected simulation, along with a vertical line displaying the timing of a fictitious critical event. The user can adjust the time of the critical event, as well as the cut-off point of the x-axis. At the bottom, two tables are shown, one showing the time needed for specific proportions of simulated trials to reach each stage (left), and one showing how many simulated trials have reached each stage at a specific time in relation to the critical event (right). This panel gives the user a clearer picture about the rate of successful transitions of control in relation to the time of the critical event. From a safety perspective, it gives the user valuable insights on the proportion of drivers that are expected to fail to take back control in time to react appropriately to a critical event.

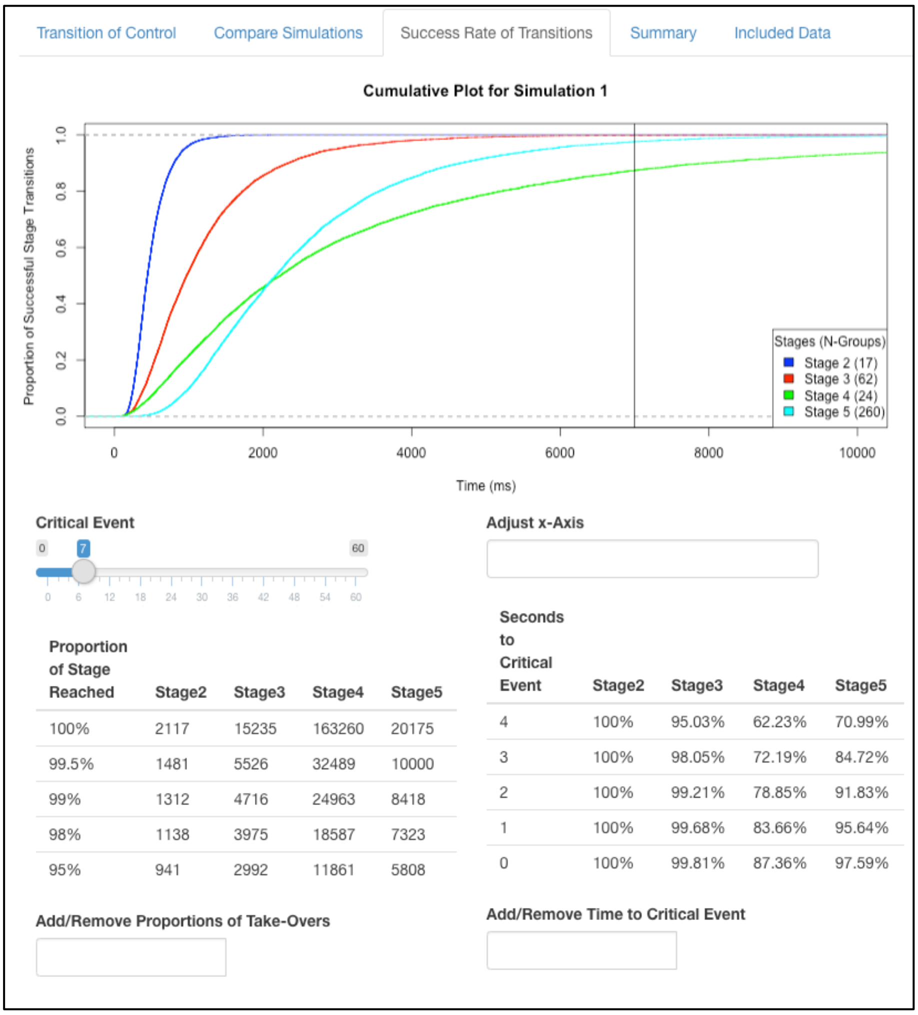

Figure 4: 'Success Rate of Transitions'-Panel. The panel at the top shows a cumulative plot for each stage in relation to the critical event (vertical line). At the bottom, two tables are shown, one showing the times needed for each stage to reach a certain percentage of successful transitions (left), and one showing the percentage of successful transitions by time in relation to the critical event (right).

Finally, the summary panel provides a table with summary statistics for each stage included in the simulation, and the included data panel (not visualized in this paper) shows the subset of the database that has been taken into consideration for the currently selected simulation. These panels can help users to find common traits and discrepancies between the groups to look for new potentially interesting parameter combinations that may be interesting to inspect further. For example: do specific authors investigate specific stages or specific parameters (such as specific modalities)?

5 Illustration of PREDICTOR's functionality through critical tests

The functionality of PREDICTOR is now illustrated through five critical tests. These analyses are not meant to be complete, but rather to (1) further demonstrate PREDICTOR's functionality, and (2) reveal interesting patterns that emerge when results of studies that focus on different stages of the interruption process (Janssen et al., 2019) are combined into a single simulation model.

Unless stated otherwise, the values for each stage were simulated independently from one another, with time intervals representing the interval between alert onset (stage 1) and the respective stage. Due to the expected skewness of the distributions, quartile deviation (QD) (Kokoska & Zwillinger, 2000) is reported as a measure of spread of the data. QD is calculated as half of the distance between the 25th and 75th percentile.

5.1 Test 1 – General Patterns in the Distribution of Stage Onset Times

The goal of the first test was to uncover general patterns in the distribution of stage onset times during the transition of control, when considering all studies in the database (i.e., without splitting by human factors). Figure 3 plots simulation results. There are distinct stage onset time distributions for stage 2 (M = 0.50 s, QD = 0.13 s), stage 3 (M = 1.23 s, QD = 0.47 s), stage 4 (M = 3.84 s, QD = 1.67 s), and stage 5 (M = 2.56 s, QD = 0.90 s).

From a safety and accident perspective, the long tails in the distribution of stage 5 (i.e., physical transfer of control) elicited by the simulation suggest that a small yet significant portion of transitions may result in an unsuccessful transition of control if current literature conventions are applied (e.g., to have a take-over time of 5-8 seconds cf., Gold et al., 2013; Mok et al., 2017). Specifically, after 8 seconds, 1.37% of simulated transitions have not been completed (7.86% had not yet reached completion after 5 s). Considering the negative consequences of a failed take-over, this rate can arguably be considered too high. Thus, while an alert being presented on such short notice might suffice in most cases, the model suggests that it is not sufficient to ensure overall road safety. It should be noted however, that some experimental groups considered for this simulation were exposed to longer alert onset times and may thus distort this result (more systematic analysis of the impact of alert onset is reported in 'Test 5 - Rate of Successful Take-Overs Based on Alert Onset Time').

Perhaps surprisingly, the distribution of stage 4 seems to peak before stage 3, and also seems to have a wider distribution than stage 5 (whereas one might expect that distributions become wider for later stages due to accumulation of variability over time). A closer look at the underlying data reveals an explanation for the unexpected pattern: 5 of the 24 experimental groups with reported results for stage 4 come from one study that used an alert onset time of 21 seconds (in the form of a pre-alert) (Van Der Heiden et al., 2017) all of which elicited a stage onset time for stage 4 that was above 8 seconds. This study only reported values for stage 4, thus not equally affecting the simulation of other stages (a comparatively large stage onset time could have been expected at least for stage 5 in this study). When re-running a simulation that omits these 5 experimental groups, the new estimated distribution of stage 4 (M = 1.43 s, QD = 0.62 s) is now located between stage 3 (M = 1.23 s, QD = 0.47 s) and stage 5 (M = 2.56 s, QD = 0.90 s), however being much closer to stage 3. This pattern now suggests that the onset of stage 4 occurs almost in parallel to the onset of stage 3. This may be due to the generally short alert onset times used in the literature comprising the database (see Supplementary Materials 'Details on Model', Table A2). This might have motivated participants to take control as fast as possible, and a process of interleaving between driving and non-driving related activities might have been skipped (see e.g. Ch, 2023; Ch et al., 2024 for empirical studies that show how shorter alert onset intervals reduce the frequency with which interleaving is applied).

5.2 Test 2 – Effect of Different Alert Modalities on the Transition of Control

Next, the effect of alert modality on the take-over process was investigated. In the literature (B. Zhang et al., 2019), by far the most commonly used alert is a bi-modal visual-auditory alert (in PREDICTOR's database 49 papers; 173 experimental groups), followed by purely auditory alerts (16 papers; 41 experimental groups). Figure 5 plots PREDICTOR's results for three different alert modalities[1].

The results (Figure 5) show that alert modality seems to mostly impact the early stages of the transition of control. The distributions of bi-modal alerts is more narrow (i.e., the red distribution that represents bi-modal alerts has a smaller quartile deviation) and more to the left (i.e., faster) compared to those of the auditory only and visual only alerts for initial disengagement (stage 2), orientation to the driver task (stage 3), and eventual task suspension (stage 4). However, eventual transition time (i.e., physical transfer of control, stage 5) seems hardly affected: the distributions of all three alert modalities overlap substantially.

One interpretation of the results is that the early response to bi-modal alerts allows the user with more time to orient to the driving task (i.e., early onset of stage 3) and therefore more time to prepare for a smooth transition, compared to uni-modal alerts. This can benefit

[1] Note again that the number of studies varies for each stage, with data from the auditory only condition in stage 4 only coming again from one paper (Van Der Heiden et al., 2017)

safety. An implication is that bi-modal alerts should especially be applied when aiming for quick initial reactions from the driver (i.e., fast stage 2, 3).

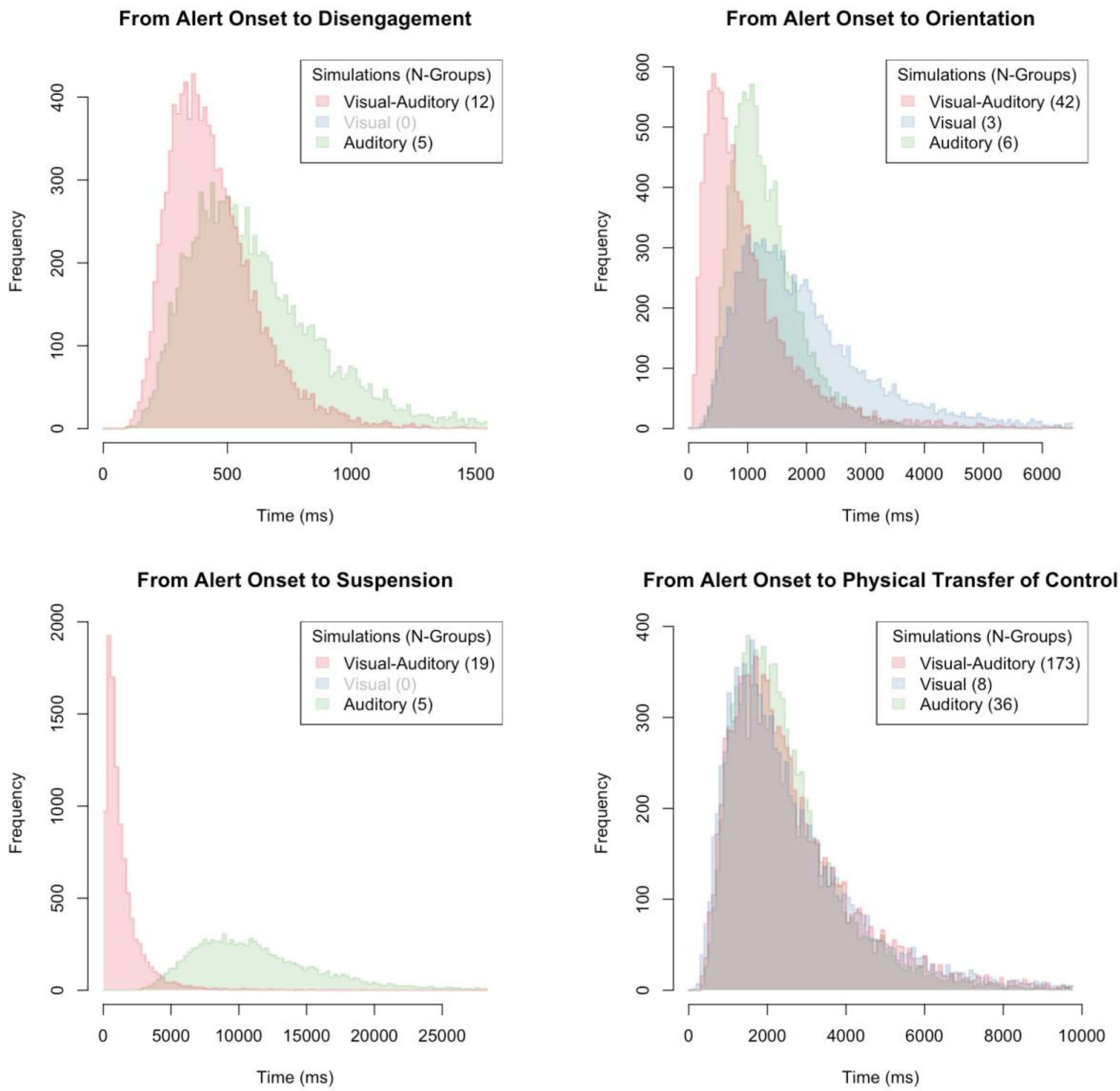

Figure 5: Distribution of the simulated data considering different alert modalities. Overall, the visual-auditory alert (red) resulted in the shortest stage onset times, compared to visual (blue) and auditory (green) alerts. This effect was most prevalent in earlier stages and dissipated in the final stage of the transition of control.

5.3 Test 3 - Effect of NDRT Output Modalities on the Transition of Control

Next, a comparison of the effects of different output modalities for the non-driving related tasks was made. Three groups were distinguished, based on whether the authors of papers reported that their task: (1) had a manual component, (2) had a cognitive component, or (3) did not fall under category 1 or 2 (e.g., no response, or vocal response). Note that in some studies, authors reported that the non-driving related tasks had both a cognitive and a manual component. These were included in both category 1 and 2.

Figure 6 shows the results. For stages 2 and 4, results should not be interpreted due to the low number of studies in the "other" category. For stage 3, the distributions of the manual condition (*M* = 1.01 s, *QD* = 0.46 s) seem to overlap largely with the other condition (*M* = 0.98 s, *QD* = 0.43 s). For stage 5, again the distributions of the three conditions overlap substantially between: manual (*M* = 2.20 s, *QD* = 0.94 s), cognitive (*M* = 2.10 s, *QD* = 0.81 s), and other (*M* = 2.10 s, *QD* = 0.81 s). However, although subtle, the manual condition has a slightly wider distribution, which is visible in the longer tail of the distribution. This in turn resulted in a larger failure rate 8 seconds after onset of the alert for manual (1.94% of the simulations had not yet completed stage 5 at this point) compared to cognitive (0.66%) and other (0.82%) NDRTs.

One interpretation of these results is that despite the common notion that visual-manual tasks can be more distracting in a driver distraction setting than cognitive tasks, cognitive tasks can cause a similar delay in eventual response time. Effects do show up when considering the tail of distributions. However, these simulations should be treated as indicators and predictions of patterns that further empirical work needs to experimentally verify before definitive conclusions are drawn. Note that especially for cognitive tasks few studies report effects on interim stages.

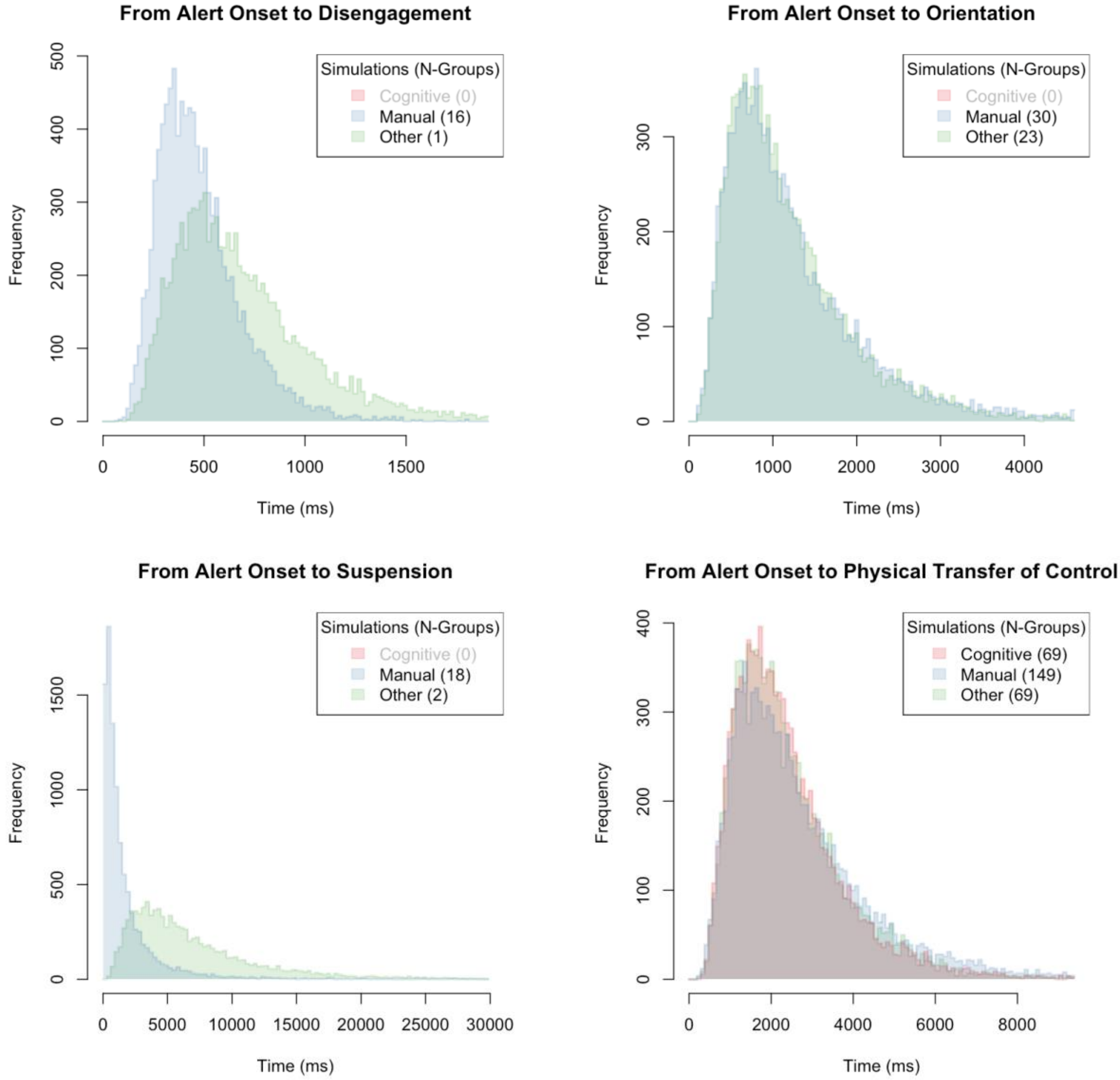

Figure 6: Distribution of the simulated data based on non-driving related task (NDRT) output modality.

5.4 Test 4 – Interaction of Alert Modality and NDRT Input Modality

The next test inspected how PREDICTOR performs when two parameters are manipulated simultaneously: alert modality and input modality of non-driving related tasks (NDRT). Following multiple resource theories (Salvucci & Taatgen, 2008, 2011; Wickens, 2002, 2008), it is hypothesized that when the modality of the alert and the non-driving related task overlap, task performance on at least one task declines. This might result in delays in responses to the alert in such conditions. However, an open question is which of the 4 stages might be affected most.

Nine simulations were performed, one for each combination of alert modality (visual, auditory, visual-auditory) and non-driving related task input modality (visual, auditory, visual-auditory). Figure 7 shows a comparison of the simulated data per stage. Again, the number of observations per condition varies, and simulations that rely on few studies should be interpreted with caution. Nonetheless, the general pattern suggests that for early stages (stages 2, 3 – the first two rows of Figure 7), visual-auditory bi-modal alerts (green distributions) have typically the fastest response distributions, independent of the modality of the NDRT (columns in Figure 7). The effect on eventual physical transition (stage 5; last row) seems less clear. This pattern is consistent with the pattern found in test 2: bi-modal alerts seem to impact early stages of transition, but more empirical work is needed to see if and how it impacts later stages.

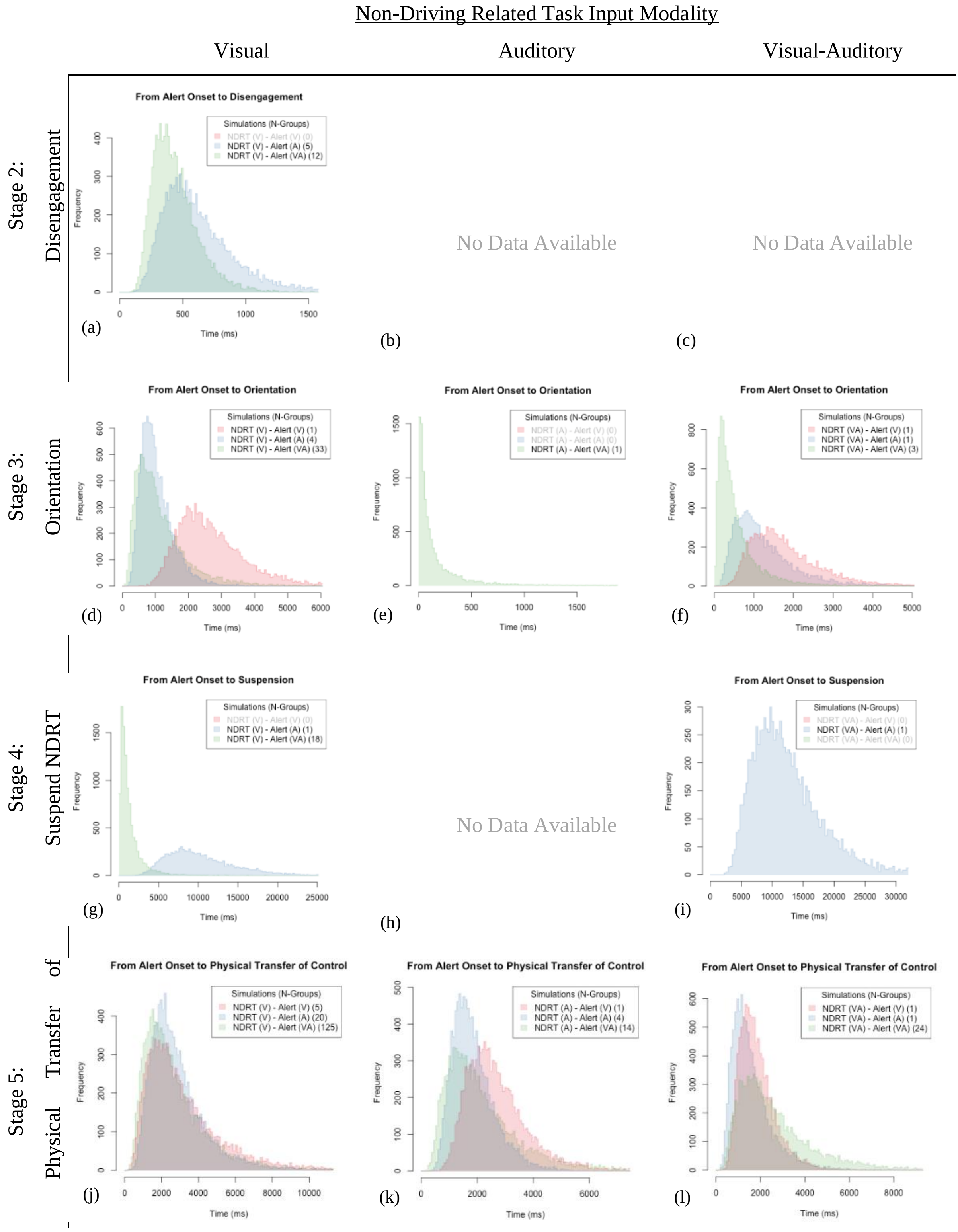

Figure 7: Simulated data per stage for visual (left), auditory (center), and visual-auditory (right) NDRT input-modalities. Each plot contains the distribution from simulations using visual (V), auditory (A), and visual-auditory (VA) alert modalities.

5.5 Test 5 - Rate of Successful Take-Overs Based on Alert Onset Time

For the final test, PREDICTOR was used to investigate the effect of alert onset time on the rate of successful take-overs. A simulation was run for each of the four most common alert onset times: 3.5 seconds (13 experimental groups; 3 studies), 6 seconds (30 experimental groups; 9 studies), 7 seconds (69 experimental groups; 13 studies), and 10 seconds (42 experimental groups; 10 studies). Note that although these studies exist in the literature, PREDICTOR can give novel insights due to its explicit simulation of a distribution of data, including more “extreme” datapoints that some empirical studies might consider ‘outliers’.

Table 1: Percentage of simulated trials having reached stage 5 (physical transfer of control) by alert onset time, in relation to the critical event onset. Generated using PREDICTOR’s ‘Success Rate of Transitions’ panel

	Critical event onset			
	3.5 s	**6 s**	**7 s**	**10 s**
2s before critical event	55.82%	94.97%	99.24%	98.25%
1s before critical event	75.66%	98.24%	99.79%	99.01%
Critical Event	85.16%	99.34%	99.95%	99.45%

Table 1 shows what percentage of drivers reached stage 5 by the time of the critical event (last row), or 1s or 2 s earlier (i.e., with some time to act before the event). These percentages are high in all scenarios for alert onset times of 6, 7, and 10 s. However, for the shortest alert onset time (3.5 s), only 85.16% had reached stage 5 at critical event onset.

These simulation results reinforce the consensus in the literature that giving an alert at least 5-8 s before a critical event is necessary (e.g., Gold et al., 2013; Mok et al., 2017).

However, there are also nuances. Consider for example the simulations when an alert had an onset of 6 s. Here, 5% of simulations had not reached the physical transfer stage 2s before the critical event onset (i.e., after 4 s), suggesting that 1 in 20 take-overs were completed at the last moment. Having such stricter criteria is useful, given that taking control of a vehicle does not always equate to a correct action. For example, remnant distractions from previous tasks might still linger (Strayer et al., 2015), and it might take a while before people have stable control over their vehicle (Merat et al., 2014). For a more detailed discussion, see Janssen et al. (2019). Having at least an onset of 7 s, seems better – as 2 s before the critical incident (i.e., within 5s of alert onset, the lower boundary of the recommended time) the large majority of simulations have completed the transition.

6 General Discussion

This paper introduced PREDICTOR: an interactive open-source software tool to predict how time is spent during various stages of a take-over process. The model simulations predict the distributions of timing of four stages of handling a take-over request, as previously introduced in a theoretical framework (Janssen et al., 2019). The main use of PREDICTOR is to explore how combinations of factors (e.g., alert onset and alert modality) impact these distributions, and how it influences the ability to meet a response time deadline. The combination of visualization and quantification of the predictions allows users to gain detailed insights in an interactive way.

PREDICTOR's current functionality allows exploration of the effects of four human factors that are commonly reported in take-over studies (B. Zhang et al., 2019): alert onset deadline, alert modality, and input and output modality of the non-driving related task.

Although there are many studies that report how these individual factors impact eventual take-over time, PREDICTOR is novel in two ways: (1) it also simulates the effect on intermediate stages of the take-over, and (2) it visualizes and quantifies what the full distribution of times look like, thereby also giving insights into what percentage of simulations (as estimate of drivers) might for example not respond to an alert in time.

The initial analyses (section 5) already gave various insights, of which three are further discussed here. First, PREDICTOR suggests that bi-modal alerts mostly seem to effect early stages of the take-over process, such as how fast someone initially disengages from a non-driving task (stage 2 in Janssen et al., 2019) and how fast they orient to the driving task (stage 3) (see section 5.2). Bi-modal alerts are already the most used alerts in take-over studies (B. Zhang et al., 2019), and multiple resource theories of cognition (Salvucci & Taatgen, 2008, 2011; Wickens, 2002, 2008) also explain why they can be effective: as there is more chance that one modality does not overlap with the modality of the non-driving related task and therefore can reach the user. However, PREDICTOR's results suggest that bi-modal alerts do not necessarily speed-up eventual physical take-over, whereas some alerts might have been designed with that goal.

Second, PREDICTOR's simulations make explicit what percentage of drivers might not react in a timely manner (see section 5.5). In line with the underlying data (B. Zhang et al., 2019), the large majority of simulated responses finishes the take-over within common response and alert guidelines of 5-8 s (e.g., Gold et al., 2013; Mok et al., 2017). That outliers are simulated is in part due to a technical reason: the model that underlies PREDICTOR assumes log-normal distributions, which are skewed. However, this assumption has merit, as skewed distributions are commonly observed in reaction time paradigms (Ratcliff, 1993). Moreover, it is PREDICTOR's ability to combine datasets and then predict performance that gives various nuanced insights about these outliers. For example, that delays due to manual interaction with a non-driving related task might particularly manifest themselves in delays on the eventual physical response (stage 5; see section 5.3).

Third, PREDICTOR makes apparent where there are research gaps in available data and associated understanding of human behavior. Three of those are: (A) few studies explicitly differentiate between the first moment of disengagement (Stage 2) and eventual suspension of a non-driving related task, thereby limiting insights regarding under what conditions drivers might be interleaving; (B) few studies explore how behavior changes when drivers have a longer time to respond (alert onset time), even though this seems to impact the stages significantly; (C) studies that have cognitive non-driving related tasks only have data on the eventual physical transfer of control (stage 5), and not on earlier stages. This last aspect might in large part be due to methodological reasons: whereas with visual and manual tasks it is possible to measure whether someone is looking at or manually manipulating a driving or a non-driving task, it might be harder to assess when they first stopped thinking about a cognitive process (e.g., stage 2) and last stopped thinking about this (e.g., stage 4) (see also Held et al., 2024). There is potential here for recent (model-driven) neuroscience methods that are getting better at estimating cognitive stages that people go through (e.g., Borst & Anderson, 2023).

A limitation of the current work is that it only considered four human factors to use for generating different simulations with (see section 3). These were chosen due to their commonality in studies and the availability of sufficient data to generate simulations. However, other factors (also beyond ‘human’ factors) can also be considered but require further scrutiny of the data and more empirical evidence. For example, the impact of operator/vehicle characteristics, different types of vehicular movements, and varying roadway/environmental conditions. PREDICTOR allows users to do these and other analyses and allows them to add new data to refine insights.

PREDICTOR’s simulations are based on a mix of psychological theory (which identified the stages) and data-driven inference (based on the meta-review from Zhang et al., 2019). To get to a deeper understanding of the mechanisms behind human behavior, even more advanced computational cognitive models of the underlying (cognitive) processes are needed. Recent literature reviews suggest that this is an area that has only just started to develop, and more research is needed before such detailed simulations can be made (Janssen et al., 2020, 2022, 2024; Lorenz et al., 2024).

The current model only makes predictions of time distributions; other metrics such as take-over quality / success, or situational awareness are not modelled. These can currently not be incorporated, as there are few studies that report such metrics and the way these measures are quantified also differs between studies, thereby limiting the ability to combine them in one simulation. Similarly, factors that are even further away from the take-over process itself such as comfort (Peng et al., 2024) and more strategic decisions (Y. Zhang et al., 2021) are not considered, as there is limited data to base simulations on.

Another limitation is that the predictions of PREDICTOR stop at stage 5 of the full transition of control model (Janssen et al., 2019), the stage where the user first physically takes over control. Again, by itself this does not predict how the eventual drive is handled (stage 6), nor how the process of handing back control to the car is handled (stages 7-10). Nonetheless, the hope of the authors is that PREDICTOR is already a valuable tool that can aid researchers, designers, and engineers in their understanding of the take-over process.

7 References

Adnan, N. (2024). Exploring the future: A meta-analysis of autonomous vehicle adoption and its impact on urban life and the healthcare sector. *Transportation Research Interdisciplinary Perspectives*, *26*, 101110. https://doi.org/10.1016/J.TRIP.2024.101110

Almlöf, E. (2024). Beyond the hype: A critical look at the motivations driving automated driving systems research. *Transportation Research Interdisciplinary Perspectives*, *24*, 101075. https://doi.org/10.1016/J.TRIP.2024.101075

Altmann, E., & Trafton, J. G. (2002). Memory for goals: An activation-based model. *Cognitive Science*, *26*(1), 39 – 83.

Arowolo, W., Larsson, M., & Nicolaï, I. (2024). Governance of automated vehicle in the urban transport system: Insight from a willingness-to-use survey and Norwegian cultural context. *Transportation Research Interdisciplinary Perspectives*, *24*, 101040. https://doi.org/10.1016/J.TRIP.2024.101040

Boehm-Davis, D. A., & Remington, R. W. (2009). Reducing the disruptive effects of interruption: a cognitive framework for analysing the costs and benefits of intervention strategies. *Accident Analysis & Prevention*, *41*(5), 1124 – 1129. https://doi.org/10.1016/j.aap.2009.06.029

Borojeni, S. S., Weber, L., Heuten, W., & Boll, S. (2018). From Reading to Driving – Priming Mobile Users for Take-over Situations in Highly Automated. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM Press. https://doi.org/10.1145/3229434.3229464

Borst, J. P., & Anderson, J. R. (2023). Discovering Cognitive Stages in M/EEG Data to inform Cognitive Models. In B. U. Forstmann & B. Turner (Eds.), *An Introduction to Model-Based Cognitive Neuroscience* (2nd ed.). Springer.

Borst, J. P., Taatgen, N. A., & Rijn, H. van. (2015). What Makes Interruptions Disruptive?: A Process-Model Account of the Effects of the Problem State Bottleneck on Task Interruption and Resumption. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM Press. https://doi.org/10.1145/2702123.2702156

Ch, N. A. N. (2023). *Conditionally automated vehicles as a safe and productive workspace* [PhD thesis]. University of New Hampshire.

Ch, N. A. N., Fortier, J., Janssen, C. P., Shaer, O., Mills, C., & Kun, A. L. (2024). Text a Bit Longer or Drive Now? Resuming Driving after Texting in Conditionally Automated Cars. *Proceedings of Automotive User Interfaces and Interactive Vehicular Applications*.

Couffe, C. L., & Michael, G. A. (2017). Failures Due to Interruptions or Distractions: A Review and a New Framework. *American Journal of Psychology*, *130*(2), 163 – 181. https://doi.org/10.5406/amerjpsyc.130.2.0163

de Winter, J., Stanton, N., & Eisma, Y. B. (2021). Is the take-over paradigm a mere convenience? *Transportation Research Interdisciplinary Perspectives*, *10*, 100370. https://doi.org/10.1016/J.TRIP.2021.100370

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences*, *113*(10), 2636 – 2641. https://doi.org/10.1073/pnas.1513271113

Donmez, B., McDonald, A. D., Lee, J. D., & Boyle, L. N. (2023). Road user behavior: Describing, inferring, predicting and beyond. *Transportation Research Interdisciplinary Perspectives*, *22*, 100932. https://doi.org/10.1016/J.TRIP.2023.100932

Gerber, M. A., Schroeter, R., & Ho, B. (2023). A human factors perspective on how to keep SAE Level 3 conditional automated driving safe. *Transportation Research Interdisciplinary Perspectives*, *22*, 100959. https://doi.org/10.1016/J.TRIP.2023.100959

Gerber, M. A., Schroeter, R., Johnson, D., Janssen, C. P., Rakotonirainy, A., Kuo, J., & Lenne, M. G. (2024, May 11). An Eye Gaze Heatmap Analysis of Uncertainty Head-Up Display Designs for Conditional Automated Driving. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3613904.3642219

Gerber, M. A., Schroeter, R., Xiaomeng, L., & Elhenawy, M. (2020). Self-Interruptions of Non-Driving Related Tasks in Automated Vehicles: Mobile vs Head-Up Display. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3313831.3376751

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). "Take over!" How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 1938 – 1942. https://doi.org/10.1177/1541931213571433

Hancock, P. A. (2013). Driven to distraction and back again. In M. A. Regan, J. D. Lee, & T. W. Victor (Eds.), *Driver Distraction and Inattention: Advances in Research and Countermeasures* (pp. 9–25). Ashgate Publishing Ltd. https://www.taylorfrancis.com/chapters/edit/10.1201/9781315578156-2/driven-distraction-back-peter-hancock

Held, M., Minculescu, A., Rieger, J. W., & Borst, J. P. (2024). Preventing mind-wandering during driving: Predictions on potential interventions using a cognitive model. *International Journal of Human-Computer Studies*, *181*.

Janssen, C. P., Baumann, M., & Oulasvirta, A. (2024). Computational models of cognition for human-automated vehicle interaction: State-of-the-art and future directions. *International Journal of Human-Computer Studies*, *185*, 103230. https://doi.org/10.1016/J.IJHCS.2024.103230

Janssen, C. P., Baumann, M., Oulasvirta, A., Iqbal, S. T., & Heinrich, L. (2022). Computational Models of Human-Automated Vehicle Interaction (Dagstuhl Seminar 22102). *Dagstuhl Reports*, *12*(3), 15–81. https://doi.org/10.4230/DAGREP.12.3.15

Janssen, C. P., Boyle, L. N., Ju, W., Riener, A., & Alvarez, I. (2020). Agents, environments, scenarios: A framework for examining models and simulations of human-vehicle interaction. *Transportation Research Interdisciplinary Perspectives*, *8*. https://doi.org/10.1016/j.trip.2020.100214

Janssen, C. P., Iqbal, S. T., Kun, A. L., & Donker, S. F. (2019). Interrupted by my car? Implications of interruption and interleaving research for automated vehicles. *International Journal of Human-Computer Studies*, *130*. https://doi.org/10.1016/j.ijhcs.2019.07.004

Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted Driving and Risk of Road Crashes among Novice and Experienced Drivers. *New England Journal of Medicine*, *370*(1), 54 – 59. https://doi.org/10.1056/nejmsa1204142

Kokoska, S., & Zwillinger, D. (2000). *CRC Standard Probability and Statistics Tables and Formulae, Student Edition*. CRC Press. https://doi.org/10.1201/b16923

Kun, A. L., Boll, S., & Schmidt, A. (2016). Shifting Gears: User Interfaces in the Age of Autonomous Vehicles. *IEEE Pervasive Computing*, 32 – 38.

Large, D. R., Burnett, G., Salanitri, D., Lawson, A., & Box, E. (2019). *A Longitudinal Simulator Study to Explore Drivers' Behaviour in Level 3 Automated Vehicles*. 222–232. https://doi.org/10.1145/3342197.3344519

Lorenz, M., Amorim, T., Dey, D., Sadeghi, M., & Ebel, P. (2024, July 4). Computational Models for In-Vehicle User Interface Design: A Systematic Literature Review. *Proceedings of the International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications*. https://doi.org/10.31219/OSF.IO/KQG6U

Lotz, A., Russwinkel, N., & Wohlfarth, E. (2019). Response times and gaze behavior of truck drivers in time critical conditional automated driving take-overs. *Transportation Research Part F: Traffic Psychology and Behaviour*, *64*, 532–551. https://doi.org/10.1016/J.TRF.2019.06.008

McCall, R., McGee, F., Mirnig, A., Meschtscherjakov, A., Louveton, N., Engel, T., & Tscheligi, M. (2019). A taxonomy of autonomous vehicle handover situations. *Transportation Research Part A: Policy and Practice*, *124*, 507–522. https://doi.org/10.1016/J.TRA.2018.05.005

Merat, N., Jamson, A. H., Lai, F. C. H., & Daly, M. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 274 – 282. https://doi.org/10.1016/j.trf.2014.09.005

Mirnig, A. G., Gärtner, M., Laminger, A., Meschtscherjakov, A., Trösterer, S., Tscheligi, M., McCall, R., & McGee, F. (2017). Control Transition Interfaces in Semiautonomous Vehicles. *Proceedings of the International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 209–220. https://doi.org/10.1145/3122986.3123014

Mok, B., Johns, M., Miller, D., & Ju, W. (2017). Tunneled In: Drivers with Active Secondary Tasks Need More Time to Transition from Automation. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 2840–2844). ACM Press. https://doi.org/10.1145/3025453.3025713

Mok, B., Sirkin, D., Sibi, S., Miller, D. B., & Ju, W. (2015). Understanding Driver-Automated Vehicle Interactions Through Wizard of Oz Design Improvisation. *Driving Assessment Conference*, *8*(2015), 380–386. https://doi.org/10.17077/DRIVINGASSESSMENT.1598

Nagaraju, D., Ansah, A., Al Nahin Ch, N., Mills, C., Janssen, C. P., Shaer, O., & Kun, A. L. (2021). How will drivers take back control in automated vehicles? A driving simulator test of an interleaving framework. In *Proceedings of the International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 20–27). ACM Press. https://doi.org/10.1145/3409118.3475128

Naujoks, F., Wiedemann, K., & Schömig, N. (2017). The Importance of Interruption Management for Usefulness and Acceptance of Automated Driving. *Proceedings of the International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications*. https://doi.org/10.1145/3122986

Peng, C., Horn, S., Madigan, R., Marberger, C., Lee, J. D., Krems, J., Beggiato, M., Romano, R., Wei, C., Wooldridge, E., Happee, R., Hagenzieker, M., & Merat, N. (2024). Conceptualising user comfort in automated driving: Findings from an expert group workshop. *Transportation Research Interdisciplinary Perspectives*, *24*, 101070. https://doi.org/10.1016/J.TRIP.2024.101070

Petermeijer, S., Doubek, F., & de Winter, J. (2017). Driver response times to auditory, visual, and tactile take-over requests: A simulator study with 101 participants. *2017 IEEE*

International Conference on Systems, Man, and Cybernetics (SMC), 1505–1510. https://doi.org/10.1109/SMC.2017.8122827

Pfleging, B., Rang, M., & Broy, N. (2016). Investigating user needs for non-driving-related activities during automated driving. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia* (pp. 91–99). ACM Press. https://doi.org/10.1145/3012709.3012735

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532. https://doi.org/10.1037/0033-2909.114.3.510

SAE International. (2014). *J3016: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems*.

Salvucci, D. D. (2009). Rapid prototyping and evaluation of in-vehicle interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *16*(2), 9. https://doi.org/10.1145/1534903.1534906

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: an integrated theory of concurrent multitasking. *Psychological Review*, *115*(1), 101 – 130. https://doi.org/10.1037/0033-295x.115.1.101

Salvucci, D. D., & Taatgen, N. A. (2011). *The Multitasking Mind*. Oxford University Press.

Sanderson, P. M., & Grundgeiger, T. (2015). How do interruptions affect clinician performance in healthcare? Negotiating fidelity, control, and potential generalizability in the search for answers. *International Journal of Human-Computer Studies*, *79*, 85–96. https://doi.org/10.1016/j.ijhcs.2014.11.003

Sever, T., & Contissa, G. (2024). Automated driving regulations – where are we now? *Transportation Research Interdisciplinary Perspectives*, *24*, 101033. https://doi.org/10.1016/J.TRIP.2024.101033

Stevens, G., Bossauer, P., Vonholdt, S., & Pakusch, C. (2019). Using Time and Space Efficiently in Driverless Cars: Findings of a Co-Design Study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM Press. https://doi.org/10.1145/3290605.3300635

Strayer, D. L., Cooper, J. M., Turrill, J., Coleman, J. R., & Hopman, R. J. (2015). *Measuring Cognitive Distraction in the Automobile III: A Comparison of Ten 2015 In-Vehicle Information Systems*.

Van Der Heiden, R. M. A., Iqbal, S. T., & Janssen, C. P. (2017). Priming drivers before handover in semi-autonomous cars. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, *2017-May*. https://doi.org/10.1145/3025453.3025507

Walker, F., Steinke, J., Martens, M. H., & Verwey, W. B. (2020). Do engineer perceptions about automated vehicles match user trust? Consequences for design. *Transportation Research Interdisciplinary Perspectives*, *8*, 100251. https://doi.org/10.1016/J.TRIP.2020.100251

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159 – 177.

Wickens, C. D. (2008). Multiple Resources and Mental Workload. *Human Factors*, *50*(3), 449 – 455. https://doi.org/10.1518/001872008x288394

Wintersberger, P., Riener, A., Schartmüller, C., Frison, A. K., & Weigl, K. (2018). Let me finish before I take over: Towards attention aware device integration in highly automated vehicles. *Proceedings of the International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 53–65. https://doi.org/10.1145/3239060.3239085

Yoon, S. H., Kim, Y. W., & Ji, Y. G. (2019). The effects of takeover request modalities on highly automated car control transitions. *Accident Analysis & Prevention*, *123*, 150–158. https://doi.org/10.1016/J.AAP.2018.11.018

Zhang, B., Winter, J. de, Varotto, S., Happee, R., & Martens, M. (2019). Determinants of take-over time from automated driving: A meta-analysis of 129 studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *64*, 285–307. https://doi.org/10.1016/j.trf.2019.04.020

Zhang, Y., Angell, L., & Bao, S. (2021). A fallback mechanism or a commander? A discussion about the role and skill needs of future drivers within partially automated vehicles. *Transportation Research Interdisciplinary Perspectives*, *9*, 100337. https://doi.org/10.1016/J.TRIP.2021.100337