**ORIGINAL ARTICLE**

# Using driver monitoring to estimate readiness in automation: a conceptual model based on simulator experimental data

Rafael C. Gonçalves[1] · Courtney M. Goodridge[1] · Jonny Kuo[2] · Mike G. Lenné[2] · Natasha Merat[1]

© The Author(s) 2024

## Abstract

This paper provides a theoretical overview of how the concept of driver readiness can be objectively measured, using controlled experimental data. First, a literature review regarding the concept of driver readiness is provided. Then, it highlights challenges for a standardized readiness estimation model. A conceptual readiness estimation model is presented, and a methodology is proposed for defining readiness thresholds for use by Driver State Monitoring (DSM) systems. The paper then explores how this model can be used to estimate readiness thresholds. A proof of concept for the model application is presented, using previously collected experimental involving SAE Level 2 automation. This paper contributes to the state of the art in DSM-development, by providing a methodology for estimating driver readiness, while considering variabilities across individual drivers. The model also allows readiness thresholds to be defined with data from driving simulator experiments, without relying on subjective assessment of readiness as its ground truth.

**Keywords** Vehicle automation · Driver readiness · Computational modelling · Evidence accumulation · Driver state monitoring

## 1 Introduction

One of the current concerns of the Human Factors and Safety community is that SAE Level 2 (SAE 2021) vehicle automation can remove the driver from the decision-action loop of the driving task (Merat et al. 2019), which sometimes impairs their performance during critical scenarios, when compared to manual driving. To mitigate safety–critical situations, and ensure drivers are not "out of the loop"

✉ Rafael C. Gonçalves
  trarg@leeds.ac.uk

  Courtney M. Goodridge
  c.m.goodridge@leeds.ac.uk

  Jonny Kuo
  jonny.kuo@seeingmachines.com

  Mike G. Lenné
  mike.lenne@seeingmachines.com

  Natasha Merat
  n.merat@its.leeds.ac.uk

[1] Faculty of Environment, University of Leeds, Institute for Transport Studies, Leeds LS2 9JT, UK

[2] Seeing Machines, Melbourne, VIC 3000, Australia

when automated driving is in operation, vehicle manufacturers are increasingly implementing systems in the vehicle which monitor driver state and activity. These driver state monitoring (DSM) systems offer a range of capabilities, starting from simple steering-based features (e.g., confirming hands-on wheel), to more advanced (camera-based) models which use drivers' eye movements, facial features, and head position, and in some cases physiological indicators to assess the driver's state, determining level of engagement with the driving task. If drivers' engagement levels fall below a certain threshold, warnings or nudges are used to re-direct their attention, for example to the forward roadway, ensuring they are prepared to respond to a potential takeover request (TOR). This assessment of "driver readiness" by the DSM (also known as driver availability, Marberger et al. 2018) ensures the driver can respond well and in time to the demands imposed by the driving environment. Work conducted by ISO/TR 20195-1 (2020a) and Georg et al. (2017) defines readiness as the likelihood of a successful intervention during a potential takeover situation, based on drivers' capabilities to manage the driving task after resuming control from vehicle automation. Studies from Mioch et al. (2017) and Kim et al. (2022) propose a model that treats readiness as a variable which changes over time, with

driver state, in relation to the minimum amount of resources required by drivers to recover their ability for driving, following resumption of control from automation.

There have been numerous attempts to establish standardised metrics for estimating readiness (e.g., Kim et al. 2022; Mioch et al. 2017; Mariajoseph et al. 2020; Baek et al. 2018; Deo and Trivedi 2020; Du et al. 2020). However, there is currently no consensus about how to measure readiness, nor how to use the measure to define the thresholds required by DSM systems. Mioch et al. (2017) conceptual model of readiness specifies a theoretical relationship between a number of core factors, but is not validated with user data. Mariajoseph et al. (2020) work used physiological metrics to predict alterations in drivers' psychological state, but this work was not related to transitions of control in vehicle automation. Conversely, Kim et al. (2022) used drivers' subjective assessment of their own readiness, and were able to correlate these scores to drivers' takeover performance. However, this approach cannot be used to provide an objective relationship between driver performance and objective driver metrics (e.g., gaze behaviour or posture), which limits its applicability on an actual DSM system.

An alternative approach is data-driven machine learning methods for estimating readiness, which uses gaze position and head pose to establish if a driver's state is appropriate for a transition of control (Baek et al. 2018; Deo and Trivedi 2020; Du et al. 2020). However, the parameter used as the ground truth on the training dataset of this study was a manual annotation of video data, where experts described drivers' state as "good" or "bad", based on their subjective interpretation (Deo and Trivedi 2020; Du et al. 2020). Although machine learning models can provide accurate predictors for the drivers' annotated state, according to Hassija et al. (2024), as the artificial intelligence-based (AI) models become more complex, the relationship between the model's predictors and the predicted variable becomes too complex to be explained. Therefore, machine learning approaches provide limited value as a tool to theoretically define readiness, and the safety thresholds required by DSM systems. Although explainable AI techniques may provide a promising solution for this problem, the state of the art in the field of explainable AI still faces challenges such as balancing transparency and privacy, addressing the diversity of user needs, and creating effective explanations for complex models (Hassija et al. 2024).

The recent regulations set by Euro NCAP for DSM (2022) provide recommendations for the development of advanced driver distraction warnings (ADDW) based on the time drivers spend looking away from the road environment. EURO NCAP (2023) also suggest the use of ocular metrics for the assessment of drivers' fatigue (as in Matthews et al. 2019). However, the guidelines for the

assessment of drivers' cognitive state and an integration of those variables with the driving context are still presented as a future milestone in their research roadmap. Factors such as driver cognitive workload (Louw and Merat 2017), or different types of cognitive distraction which do not necessarily take into account drivers' eyes off road—such as mind wandering (Walker and Trick 2018) may be crucial for the build-up of driver readiness, as previous studies provided evidence of its' effects on drivers' takeover performance.

One of the challenges for defining the readiness threshold is that it is scenario dependent (Marberger et al. 2018; Kim et al. 2018). As different driving scenarios place different demands on drivers' available resources, it is hard to quantify how different elements of a scenario affect the level of driver readiness needed for a safe resumption of control from automation. In addition, readiness is determined by drivers' mental (Kim et al. 2022; Mioch et al. 2017) and physical state (Baek et al. 2018; Deo and Trivedi 2020), which cannot be systematically manipulated and controlled in experimental settings (Spanfelner 2012). This creates a challenge for providing the empirical evidence that is required for defining thresholds. Therefore, there is currently a lack of understanding about how to correlate specific indicators of driver state with the probability of crash avoidance, for a given scenario.

Based on the research gap presented above, this paper proposes a conceptual model that uses evidence accumulation for decision-making (see Ratcliff et al. 2016) and the concept of scenario controllability, as proposed by the ISO 26262 International Standarization Organization (2020b). These concepts are used to develop a method for assessing driver readiness from experimental data, defining safety thresholds for a specific automated driving scenario. Section 2 of this paper summarises the current theoretical definitions of driver readiness and its related factors. In Sect. 3, the theoretical relationship between driver readiness and evidence accumulation (Ratcliff et al. 2016) is outlined, discussing their association with readiness thresholds and scenario controllability as defined by ISO 26262 (International Standardization Organization 2020b). Section 4 introduces the conceptual model for estimating readiness, using evidence accumulation and discusses a methodology for defining safety thresholds of DSM systems, using an experimental setup. To conclude, a model application is presented as a proof of concept, fitting it to experimental data collected from a series of driving simulator experiments assessing driver resumption of control after SAE Level 2 automated drive (Louw et al. 2016, 2017, 2018). It must be noted that the model presented in this work does not aim to provide real-time estimation of driver readiness, but rather provide a way for system

designers to define DSM readiness thresholds, based on empirical evidence from simulator experiments.

## 2 Definition of readiness and related constructs

The term driver readiness (or availability) was first used in the context of assisted driving to define the drivers' capabilities in recovering motor control (or motoric readiness) of the driving task (Zeeb et al. 2015). Readiness is based on the assumption that, once automation is engaged (for SAE levels 2 or higher) the driver is likely to be removed from the cognitive and motor control loops of the driving task (Merat et al. 2019). This removal from the loop may be exacerbated by drivers' engagement in secondary (non-driving related) tasks (Carsten et al. 2012), or simply because drivers are bored when they are not responsible for the moment-to-moment control of the vehicle. Studies report a dispersion of visual attention away from the forward roadway (Louw and Merat 2017; Gershon et al. 2023), and/or a miscalibration of driver visuomotor coordination (Mole et al. 2019) due to a lack of manual engagement. Consequently, drivers may not always be capable of recovering suitable control of the driving task following a TOR. Considering the process described above, ISO/TR 20195-1 defined readiness as a metric associated with drivers' state during the TOR, and the likelihood of them successfully recovering

control from the automated driving task (International Standardization Organization 2020a, b). This definition implies that drivers might need to accumulate resources during the transition process in order to reach the desired state. This definition is similar to that of Georg et al. (2017), who define readiness as: "(…) the fastest ability of the driver to get engaged in the driving task from the Non-Driving Related Task (NDRT) (…)".

Marberger et al. (2018) introduced a temporal aspect to the definition of readiness, by assuming that every TOR has a stipulated time budget for the resumption of control. They propose that a ready driver gathers enough resources to match the demands of the scenario, within the stipulated time budget for the takeover response. Kim et al. (2018) and Marberger et al. (2018) also state that the definition of a ready driver is scenario dependent. Complex scenarios might require drivers to gather more information about the situation in order to achieve a desirable driving state to takeover within the same time budget.

In their theoretical framework, Mioch et al. (2017), later supported by Kim et al. (2022), readiness is defined as an abstract concept that represents the cognitive and motoric resources available to drivers for resuming control of the driving task. According to these authors, this variable varies over time, as the driver becomes more or less engaged with the driving task. Furthermore, the scenario demands for this concept are represented by a threshold of minimum resources needed by drivers to safely resume control. In this sense, the resource (readiness) needs to be accumulated within a given period of time after the TOR is issued, and a ready driver is one that is capable of reaching appropriate levels of readiness within the stipulated time



**Fig. 1** Schematic representation redrawn from Kim et al. (2022) conceptual model of readiness. In this theoretical model, by the time a TOR is issued, the readiness values of the driver can be directly correlated with the likelihood of the driver to recover control of the driving task in time for a given scenario
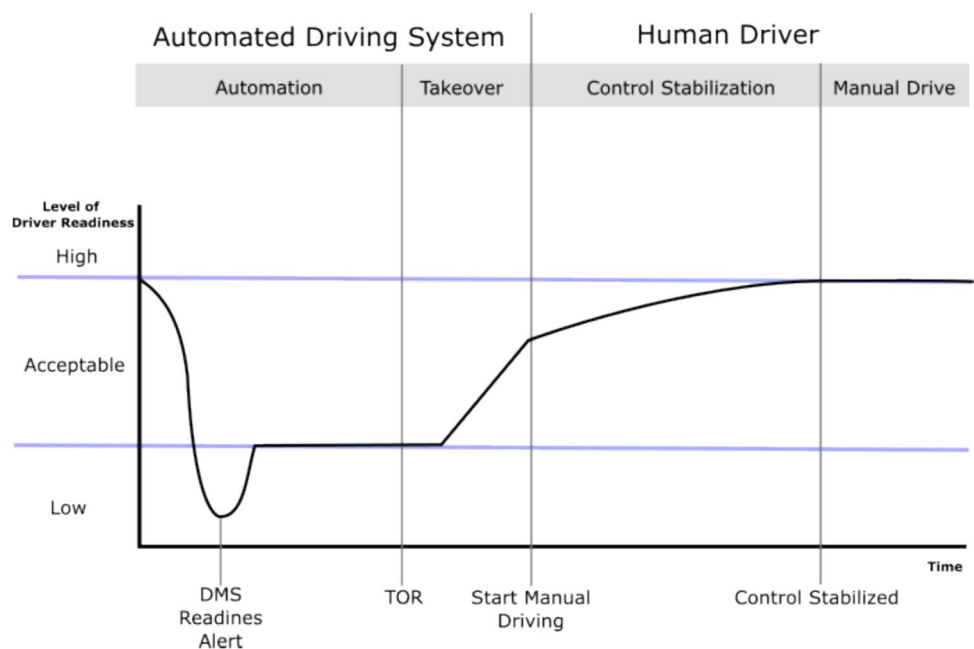
budget, to match with the demands of the scenario (see schematic in Fig. 1).

## 2.1 Correlation of readiness with other theoretical constructs

As readiness is not a metric that can be directly measured, it is generally operationalised as a combination of other constructs that are associated with drivers' capabilities to resume control from vehicle automation (Hecht et al. 2019). Mioch et al. (2017) suggests that the constructs associated with readiness fall into two main categories: motoric/physical, and cognitive/mental readiness. Marberger et al. (2018) introduced the role of individual differences, such as the influence of driver trust, intention, and risk-taking behaviour, into their framework. Although these metrics may not be objectively measurable, they may influence the timing and quality of drivers' takeover from an automated vehicle. Table 1 shows an amalgamation of a number of studies on the subject of readiness and resumptions of control from vehicle automation, including examples of empirical metrics used as proxies for assessing these constructs, in experimental scenarios. It should be noted that this exercise did not involve a systematic literature review of the metrics used for estimating readiness, but includes examples of typical metrics used in this context. Figure 2 summarises these constructs, suggesting the relationships between them. In the next section, more details are provided about the high-level concepts of physical and cognitive resources, and the effect of individual differences on shaping driver readiness.

### 2.1.1 Cognitive readiness

Michon (1985) have described the driving task as an amalgamation of several (parallel or consecutive) small cognitive tasks, requiring the driver to process information (mostly visual), and act accordingly. In that sense, it is to be expected that, in order to be ready to recover control of the driving task, the driver must have enough cognitive resources available to attend to relevant information sources on the environment, and process the information in a timely manner, to make adequate decisions (being in the loop, according to Merat et al. (2019)). This availability of resources has been broadly described as cognitive readiness. Based on the previous description, constructs such as Situation Awareness (Endsley 1995) and visual attention have been commonly associated with drivers' cognitive readiness. Those constructs are associated with one's capabilities to make decisions (Gonçalves et al. 2019a, b), and have been widely considered as a vital part of drivers' capabilities to resume control from vehicle automation (Zeeb et al. 2015; Clark and Feng 2017; Louw et al. 2016, 2017; Louw and Merat 2017; Merat et al. 2019; Gonçalves

et al. 2022). Exposure to automated drive is also known to influence driver state in a range of ways including workload, (Perello-March et al. 2022; Zhou et al. 2021). Drivers are likely to get bored due to the lack of engagement (Körber et al. 2015), or cognitively loaded, as they engage with secondary tasks (de Winter et al. 2014; Merat et al. 2019; Yoon and Ji 2019; Radhakrishnan et al. 2022), therefore, affecting their readiness levels (Hecht et al. 2019).

### 2.1.2 Physical readiness

Physical readiness comes from the idea that even if vigilant, drivers' ability to drive may be impaired by their lack of engagement with the physical controls of the vehicle. Previous studies on safety–critical transitions of control have reported a delay between when drivers place their hands on the steering wheel, and the time taken for a collision-avoidance manoeuvre (see Gold et al. 2013; Louw et al. 2016, 2018). Furthermore, drivers in L2 and L3 automation are likely to be engaged in visual-manual non-driving related tasks (NDRTs). Drivers' hand placement away of the steering wheel will not only be detrimental to their cognitive state (by removing their mind off the driving task) but will also compromise their motoric availability (i.e., their capability to promptly respond to a physical demand from the driving task, such as move the steering wheel or press a button) to intervene.

Drivers' physical readiness can be also correlated with drivers' arousal/fatigue state. As said above, drivers in L2/3 automation are likely to engage NDRTs. This engagement with NDRTs may ultimately increase their demand for cognitive resources, since they need to not only pay attention to the environment, but also to the secondary task. Therefore, drivers' engagement with NDRTs may lead to fatigue, compromising their ability to resume control. On the other hand, the lack of interaction if the physical control of the driving task may lead to boredom. This lack of stimuli may also lead to increased levels of sleepiness and subsequent fatigue, also affecting drivers' ability to resume control of the vehicle (Guo et al. 2021).

### 2.1.3 Individual differences

There is evidence that individual differences between drivers may be associated with the likelihood to safely resume control of the driving task from automation. For example, the adequacy of a drivers' mental model to the actual behaviour of the automated system may be paramount to avoid potential mode errors (Pradhan et al. 2021; Reason 2004). Victor (2009) reported in a Wizard of Oz experiment that drivers were unable to avoid a crash due to a system malfunction even though they had their eyes on the road. The authors argue that even though attentive, they were not

**Table 1** List of theoretical constructs associated with the concept of readiness, and the empirical metrics used as an operational proxy to assess those constructs

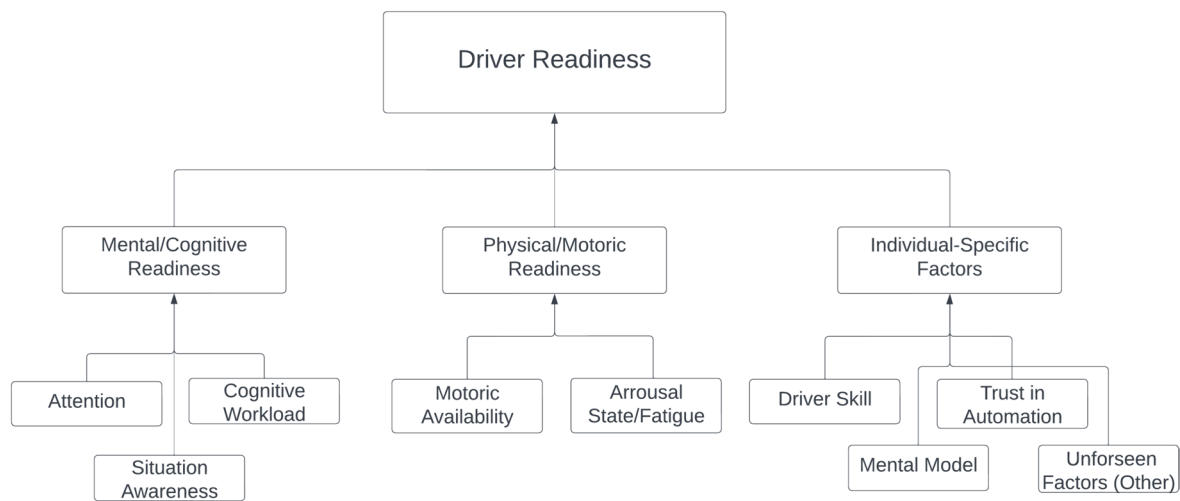| Category of readiness | Theoretical construct associated with the readiness category | Literature evidence of effects on takeover performance, or inclusion on a readiness conceptual model | Operational metrics used as a proxy for the construct | Literature evidence of the metric being used as a proxy for the said construct |
|---|---|---|---|---|
| Cognitive readiness | Situation awareness | Endsley (1995) | Markov chain of gaze transitions | Gonçalves et al. (2019a, b) |
| | | Gonçalves et al. (2019a, b) | | Schnebelen et al. (2020) |
| | | Zeeb et al. (2015) | Gaze dispersion from the road centre | Zeeb et al. (2015, 2016) |
| | | Clark et al. (2017) | | Victor (2009) |
| | | Louw et al. (2016, 2017) | | EURO NCAP (2023) |
| | | Louw and Merat (2017) | | Zhou et al. (2021) |
| | | Merat et al. (2019) | | |
| | Visual attention | Kim et al. (2022) | Gaze dispersion from the road centre | Victor (2009) |
| | | Deo and Trivedi (2020) | | EURO NCAP (2023) |
| | | Georg et al. (2017) | Time-based attention algorithm | Kircher and Ahlström (2017) |
| | | Yoon and Ji (2019) | Computer vision algorithm detection of visual distraction | Baek et al. (2018) |
| | | Zeeb et al. (2016) | | Deo and Trivedi (2020) |
| | | Eriksson and Stanton (2017) | | Du et al. (2020) |
| | | Louw et al. (2019) | Standard deviation of gaze position | Louw and Merat (2017) |
| | | | | Gonçalves et al. (2019a, b) |
| | Cognitive workload | de Winter et al. (2014) | Standard deviation of gaze position | Louw and Merat (2017) |
| | | Merat et al. (2019) | Gaze entropy | Shiferaw et al. (2019) |
| | | Yoon and Ji (2019) | Fixation duration | Shojaeizadeh et al. (2016) |
| | | Louw and Merat (2017) | Driver's physiological state | Radhakrishnan et al. (2022) |
| | | Choi et al. (2020) | Pupil dilation | Wilbanks et al. (2021) |
| | | Hecht et al. (2019) | | Shojaeizadeh et al. (2016) |
| Physical readiness | Drowsiness and arrousal state | Perello-March et al. (2022) | Driver's physiological state | Radhakrishnan et al. (2022) |
| | | Körber et al. (2015) | | Perello-Marc et al. (2022) |
| | | Zhou et al. (2021) | | Zhou et al. (2021) |
| | | Feldhütter et al. (2018) | Pupil dilation | Radhakrishnan et al. (2022) |
| | | Mariajoseph et al. (2020) | | Tseng et al. (2018) |
| | | Guo et al. (2021) | PERCLOS | Kang (2013) |
| | | | | Sun et al. (2020) |
| | | | Amplitude-velocity ratio of eye blinks | Evingerm Manning and Sibony (1991) |
| | | | | Johns (2003) |
| | Motoric availability | Gold et al. (2013) | Camera detection of drivers' body and hand position | Moslemi et al. (2021) |
| | | | | Li et al. (2020) |
| | | Louw et al. (2016) | | Baek et al. (2018) |
| | | SAE (2021) | | Du et al. (2020) |
| | | Bowden et al. (2019) | | Deo and Trivedi (2020) |
| Individual differences | Mental models | Pradhan et al. (2021) | NA | |
| | | Reason (2004) | | |
| | | Victor (2009) | | |
| | Trust in automation | Payre et al. (2014) | | |
| | Experience with automated vehicles | Hergeth et al. (2016) | | |
| | Driving style | Ma and Zhang (2021) | | |

**Fig. 2** Overview of constructs associated with driver readiness, as cited in the current literature (adapted from Mioch et al. 2017 to include results from a more recent literature review)

expecting the system to fail, due to a mismatch between their mental model and the actual system limitations. Other individual-specific factors have been reported by past studies as potential detrimental factors on drivers' performance during transitions of control. Amongst the factors listed there are: trust in automation (Payre et al. 2014), experience with automated vehicles (Hergeth et al. 2016) and driving style (Ma and Zhang 2021).

## 2.2 Measuring readiness

Despite the large body of theoretical definitions of readiness, some issues remain about the objective and practical assessment of readiness, when attempting to create a standardised model for this concept. As stated above, readiness is generally operationalised as a combination of observable metrics of drivers' behaviour as a proxy of the myriad of physical and cognitive aspects that constitute driver readiness. For instance, Euro Ncap recommendations for ADDW (2022) rely on off-road glances to assess the level of drivers' distraction/engagement. The same can be said for the use of eyelid closure and blink-related metrics for the assessment of drivers' drowsiness levels (see Baek et al. 2018). As suggested by the regression analyses from the models presented by Du et al. (2020) and Deo and Trivedi (2020), it is likely that different measures used in readiness estimation models will provide different weight values on drivers' overall readiness build-up. Some measures can also present unexpected correlations between each other, and the overall correlation with a readiness value may be complex/not linear. For this reason, readiness estimation models need

to rely on an objective ground truth of driver readiness, so appropriate regression techniques can be used to account for the complex relationship between metrics.

With respect to objective ground truth, readiness is not directly observable since it is simply a representation of an individual's mental and physical states. Although off-road glances can be used as clear indicators of reduced readiness, there are other instances (e.g., driver not paying attention to the environment, due to high levels of fatigue or cognitive load) where the driver may not be ready to take over, despite their gaze being directed to the road centre. Without directly observable metrics, the development of readiness estimation models becomes a challenge, as it lacks a ground truth for the predicted metric. Most studies in this context have either relied on driver self-assessment of readiness (Kim et al. 2022), or evaluation by expert "safety drivers" (Deo and Trivedi 2020). It can be argued that subjective assessment of a mental state is prone to misinterpretation and cannot be used directly by a DSM system. With no objective ground truth, the validation and parametrization of readiness models is unreliable.

In addition to the issues presented above, the development of a solid ground truth of readiness through experimental research is challenging, since readiness, as a concept, cannot be experimentally controlled. Since readiness is a complex combination of factors, unaccounted intervening variables on experimental manipulations (e.g., the drowsiness state of drivers on an experiment controlling for drivers' cognitive load) may lead to different readiness levels of individuals under the same experimental manipulation. Not to mention that common experimental manipulations of driver state (e.g.,

workload manipulations through secondary task) may present heterogeneous effects between individuals (see Kelley et al. 2023).

## 3 Potential solutions for a standardized readiness estimation and threshold definition

The previous section of the paper presented a literature review on the concept of driver readiness, and the theoretical constructs associated with it, demonstrating two major issues which can impair the development of a standardized readiness estimation model: (1) the lack of an objective ground truth of driver readiness; (2) the challenges to experimentally control readiness, due to inter-participant variability. This section provides some potential theoretical solutions for tackling these issues. After a short introduction, a conceptual model is presented, applying the proposed solutions to a driving simulator experimental scenario.

### 3.1 Assessing driver readiness through scenario controllability

One of the current challenges for readiness assessment is the lack of a ground truth parameter, that can be used to validate a model's prediction. Mariajoseph et al. (2020) suggested that thresholds of driver readiness can be defined through categorisation of hazards within the operational design domain (ODD), for a given automation system, and the corresponding risk assessment procedures (ASIL—Automation System Integrity Level), as defined by ISO 26262 (2018).

ISO 26262 (2018) used three criteria to create a methodology for assessing the level of risk—and therefore the degree of safety countermeasures needed—for a given scenario. This resulted in a final safety score (ASIL) corresponding to the level of intervention needed for a system to be considered safe (See Fig. 3). The three criteria are defined as:

(1) Exposure: Measured as the number of events for every 1000 km driven by a driver. A high exposure event would have an average of 2 occurrences for every 1000 km driven.

(2) Severity: Measures the level of injury caused by the event, in case the accident is not avoided by the



**Fig. 3** ASIL categorisation table (source: ISO 26262)

driver. Here, low severity = light bruises, medium severity = traumas, but with high chances of survival, high severity = life threatening injuries.

(3) Controllability: This which accounts for the probability that average driver avoids the accident, once they face the given event (e.g., low controllability = < 90% chance of crash avoidance, medium controllability = 90–99% of crash avoidance, high controllability = > 99% chance of crash avoidance).

Given the relationship between readiness and the likelihood of accidents following a TOR, Mariajoseph et al. (2020) suggested that the controllability measurement in the ASIL table could be used as an estimate of driver readiness. For this concept, a satisfactory level of driver readiness to a given scenario is related to a high level of controllability (e.g., an over 99% chance of crash avoidance). Therefore, mechanistic models that can predict crash probability/ drivers' response time by using driver behavioural metrics (see McDonald et al. 2019; Markkula et al. 2020) can be employed as ground truth for readiness estimation.

Following the ISO 26262 methodology, and Mariajoseph et al.'s (2020) theoretical consideration, our methodology for assessing readiness thresholds for a given scenario is proposed as follows:

- Given the automated driving system's limitations and ODD, define the characteristics of the safety–critical scenario which the DSM system needs to prepare the driver for (according to ISO 26262 methodology for scenario definition).
- Reconstruct the safety–critical scenario in a controlled experimental environment.
- Using the reconstructed scenario, collect responses from drivers, assessing performance (crash/not crash and response time), as well as collecting behavioural data, to be used as readiness indicators (based on the constructs described previously in this paper) at the time of the TOR.
- Using the experimental data, identify the minimum values of readiness indicators that predicted the probability of avoiding a crash, above the recommended threshold for a high controllability scenario (according to ISO 26262).

## 3.2 Using evidence accumulation models to account for individual variability

Evidence accumulation models (EAMs) have been used successfully to predict human response time and behaviour in a variety of perceptual decision-making tasks (Gold and Shadlen 2007; Ratcliff 1978; Ratcliff and Review 2004; Ratcliff et al. 2016). EAMs consider that every individual

has an internal threshold of the amount of evidence they need to make a decision. Evidence is then accumulated over time as the individual gathers information regarding the scenario in a noisy process until they reach a certain threshold, triggering a response. This method has been successfully used in the past to explain human responses when initiating braking (Durani et al. 2021; Markkula 2014; Markkula et al. 2019), steering (Goodridge et al 2022a; Goodridge et al, 2022b; Markkula et al 2018), as well as takeover responses (see McDonald et al. 2019 for a complete literature review on the topic). Previous research has also shown that evidence accumulation have theoretical correlations with situation awareness acquisition (see Gonçalves et al. 2019b). Situation awareness has often been related to readiness (Mariajoseph et al. 2020), and previous work has considered driver readiness as a construct that fluctuates over time (Kim et al. 2022). Therefore, in the next section, a discussion is provided about the theoretical correlations between evidence accumulation and the concept of readiness, and how EAMs can be used to estimate readiness.

As discussed previously, readiness has been commonly related with the construct of situation awareness (e.g., Mariajoseph et al. 2020), since low levels of situation awareness have been associated with increased crash probability in takeover scenarios (see Louw and Merat 2017) and delayed response to a TOR (see Zeeb et al. 2016). Similarly, Gonçalves et al. (2019b) have demonstrated that the third level of situation awareness—projection of future scenarios of action—is similar to the concept of evidence for choice selection, in a decision-making task. Using this argument, Gonçalves et al. (2019b) proposed that the situation awareness recovery process (as described by Gartenberg et al. 2014) can be directly modelled, using a noisy evidence accumulation process.

Conceptual models have defined readiness as a continuous variable that defines the amount of motoric and cognitive resources at the disposal of the driver to resume control of the driving task (Mioch et al. 2017; Kim et al. 2022). This definition is similar to the description of evidence made by Ratcliff et al. (2004, 2016). Ratcliff et al. (2004, 2016) defined evidence as an amalgamation of perceptual and cognitive resources that is accumulated in the form of neural activity by the human brain. This activity is then used to trigger a behavioural response from an individual. Moreover, a systematic literature review by McDonald et al. (2019) suggests that evidence accumulation models may be a promising tool for predicting drivers' takeover behaviour, suggesting a correlation between evidence accumulation and the decision to takeover. Since readiness is also associated with drivers' decision to takeover, it is believed that evidence accumulation can be used to model drivers' acquisition of motoric and cognitive resources for resuming manual control

of the driving task, as suggested by Kim et al. (2022) conceptual model.

The advantage of this approach is its ability to explain certain internal aspects of human cognition, such as information processing and the noise inherent to neural activity. This allows the creation of mechanistic models, capable of simulating human response output, based on parameters extracted from a real-world dataset, accounting for inter-participant variability. Moreover, EAMs can provide Montecarlo simulations of drivers' response on an individual level, based on a given set of parameters. Considering their capabilities, EAMs may be used to predict response time (and therefore crash probability) of a given individual, based on a series of readiness indicators.

## 4 Conceptual model

### 4.1 Conceptual model description

Considering that readiness is the amount of predisposition of an individual to perform a successful take-over, this metric can be added into a drift–diffusion (see Ratcliff and Review 2004 for a comprehensive description of the model) evidence accumulation model as the initial pre-stored evidence value before the accumulation process (termed "bias", see Ratcliff and Review 2004). As suggested by previous studies (e.g., Mariajoseph et al. 2020; Baek et al. 2018; Mioch et al. 2017; Deo and Trivedi 2020; Kim et al. 2018, 2021), the estimated values for driver readiness would be a product of drivers' cognitive and motoric state. Taking into account the factors mentioned above, we propose the following equation (Eq. 1) for the evidence accumulation process:

$$E_{(T,c)} = \sum_{t=0}^{t=T-1} E_{(t,c)} + X \sim N\left(\mu_c, \sigma_c^2\right)$$

$$E_{(T=ToT,c)} = 1$$

$$E_{(T=0,c)} = f(CR, MR)$$

$$f(CR, MR) = \sum_{n=1}^{n=N} CR_n * i_{CR_n} + \sum_{n=1}^{n=N} MR_n * i_{MR_n}$$

Equation 1 Proposed equation for the conceptual model. Where $E$ is the amount of evidence in time $T$, given the controllability $c$ for the scenario. The model assumes that the evidence accumulation process in each time step follows a gaussian distribution of mean $\mu_c$, and variance of $\sigma^2_c$ for each individual scenario. This approach is in line with previous EAMs reported in the literature (see Ratcliff et al. 2016), and assumes that different takeover scenarios will yield faster or slower evidence accumulation rates, given their complexity (controllability). The evidence threshold
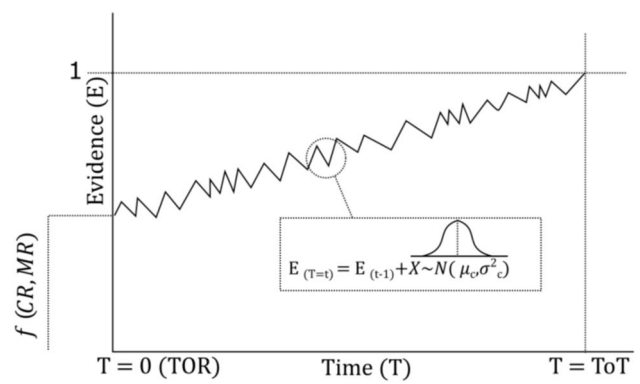


**Fig. 4** Graphical representation of the conceptual model

for the takeover scenario was set to 1, implying that the evidence accumulated by the individual during the observed takeover time ($T = ToT$) will always be 1. The readiness of a driver whenever a TOR is issued ($T = 0$) is represented as a function $f(CR, MR)$, where $CR$ is the combined score for drivers' cognitive readiness, and $MR$ the combined score for drivers' motoric readiness. Figure 4 shows a graphical representation of the conceptual model, and the evidence accumulation process.

The combined scores for $CR$ and $MR$ are calculated as a sum of $N$ distinct metrics, which operationalize the theoretical constructs associated with readiness (e.g., situation awareness, arousal state and motoric availability). For each metric ($n$) is multiplied by a constant $i_n$, which represents the weight for the metric in the equation. This approach assumes that not all theoretical constructs are necessarily of equal importance for the driver state, and therefore assume no a priori value for the metrics. This method proposes that the constants $i_n$ should be selected after a parametrization of the model, fitting the model parameters to real experimental data.

### 4.2 Model assumptions

As said above, the model assumes that "evidence" values, in evidence accumulation models (Ratcliff and Review 2004) can be used as a proxy of drivers' readiness state. This assumption is made based on the definition of readiness by Mioch et al. (2017) and Kim et al. (2022), of a conceptual floating variable that represents the sum of the drivers' motoric and cognitive resources to takeover the vehicle (akin to the definition of evidence for decision-making for evidence accumulation models).

Being a safety–critical decision-making with time pressure as a constraint for the decision, the model assumes that drivers would try to react as soon as they are ready to do so (i.e., whenever their $E_{(T,c)}$ values reach the threshold of 1), maximizing their reaction time. This model also

assumes that the drift rate for evidence accumulation (how fast someone can accumulate evidence) is dependant of the scenario's complexity/controllability ($c$). This assumption implies that more complex scenarios would lead to slower evidence accumulation rates, which translates to distinct values for the gaussian distributions that controls for the drift rate ($X \sim N(\mu_c, \sigma^2_c)$), for every observed scenario.

For a matter of simplicity, the model assumes that the only predictor of drivers' reaction time is their readiness values and accumulation of evidence, disregarding other factors like looming of the incoming obstacle, or an automatized response to the TOR, which may provide potential biases for the model's outcome. The last assumption is that the weight values, measuring the relevance of a readiness predictor in the model (dly provide the publisher name for $i_{CR_n}$) are constant across individuals, regardless of personal characteristics. For example, the importance of the amplitude of the drivers' eyelid openings as a proxy of fatigue would have the same weight value for two distinct subjects, even though the baseline amplitude may vary slightly across individuals (Johns 2003).

## 4.3 Definition of readiness thresholds

Considering the approach in Sect. 3.1, used to estimate readiness thresholds based on scenario controllability, it is possible to use the conceptual model presented above to estimate readiness thresholds for experimental data. According to ISO 26262 International Standarization Organization (2020b), and also supported by Marberger et al. (2018), every scenario $c$ (representing its complexity) should have a maximum time budget for drivers' response (notated in this work as $t_{max}$). According to the standards for controllability defined by ISO 26262 (2018), a highly controllable scenario (therefore considered within safety boundaries), would be the one with a probability for the driver to avoid harm (i.e., crash), to be equal or higher than 99%. In that sense, a driver monitoring system can use predictive models of drivers' performance to estimate the likelihood of avoiding crash, given drivers' readiness levels. Following this logic, a readiness threshold can be defined as the minimum amount of accumulated readiness to ensure that the predicted probability for the driver to avoid risk (crash) is within the desired controllability standards (99% or more). Translating this concept to the model's equation, high controllability level for $c$ (therefore, appropriate readiness levels) is classified in a way that the probability for drivers resuming control before $t_{max}$ is higher than 99%. Applying this definition to the model presented in Sect. 4.1, we can assume the formula:

$$P\left(E_{(t_{max},c)} \geq 1\right) \geq .99$$

Considering the equation for the model presented in Eq. 1, we can assume that the probability distribution for someone to react within the time budget $t_{max}$ is equal to the probability for the readiness function values, plus the sum of the accumulation of evidence in $T = t_{max}$, to be equal or higher than 1, as in:

$$P\left(E_{(t_{max},c)} \geq 1\right) = P\left(f(CR,MR) + \sum_{t=1}^{t=t_{max}} X \sim N(\mu_c, \sigma^2_c) \geq 1\right)$$

To solve the equation above, we can isolate the readiness values (notated as $f(CR,MR)$) from the expression, since they are represented by a constant. Also, the expression for evidence accumulation over time in the model is represented by a sum of $t_{max}$ random variables, drawn from a normal distribution ($N(\mu_c, \sigma^2_c)$). The result of this sum can be represented as a new normal distribution, with $\mu, \sigma^2$ equals the sum of, $\sigma^2$ of the functions that originated the random variables. Following these steps, we have:

$$\sum_{t=1}^{t=t_{max}} X \sim N(\mu_c, \sigma^2_c) = X \sim N(t_{max} * \mu_c, t_{max} * \sigma^2_c)$$

Then

$$P\left(E_{(t_{max},c)} \geq 1\right) = P(f(CR,MR) + X \sim N(t_{max} * \mu_c, t_{max} * \sigma^2_c) \geq 1)$$

Then

$$P\left(E_{(t_{max},c)} \geq 1\right) = P(X \sim N(t_{max} * \mu_c, t_{max} * \sigma^2_c) \geq 1 - f(CR,MR))$$

Then

$$P\left(E_{(t_{max},c)} \geq 1\right) \geq .99 = P(X \sim N(t_{max} * \mu_c, t_{max} * \sigma^2_c) \\ \geq 1 - f(CR,MR)) \geq .99$$

If the objective is to find the minimum readiness value to yield the desired controllability (set to 99% for the purpose of this paper) values for $c$ (notated as $Min(f(CR,MR)_c)$), we can modify the equation as stated below:

$$P\left(E_{(t_{max},c)} \geq 1 | Min(f(CR,MR)_c)\right) \geq .99$$

To solve the problem, it is necessary to find the value for the 99th percentile for the random variable $X$ (which represents the accumulation of evidence over time) that follows a normal distribution $N(t_{max} * \mu_c, t_{max} * \sigma^2_c)$. The resultant of this operation would be the minimum

value of accumulated evidence on the 99th percentile of the distribution generated by the model (noted as $AE$ for accumulated evidence):

$$P\big(X \sim N\big(t_{max} * \mu_c, t_{max} * \sigma^2_c\big) \geq AE\big) \geq .99$$

Assuming that the evidence in $t_{max}$ can be notated as $X \sim N\big(t_{max} * \mu_c, t_{max} * \sigma^2_c\big) + f(CR, MR)$, we can find $Min(f(CR, MR)_c)$ by solving:

$$P\big(X + Min(f(CR, MR)_c) \geq 1 \big| X \geq AE\big) = 1$$

Given that $P(X \geq AE) = .99$, $P\Big(E_{(t_{max}, c)} | Min(f(CR, MR)_c \geq 1\Big) = .99$ and $E_{(t_{max}, c)} = X + f(CR, MR)$ we can say that $AE + Min(f(CR, MR)_c = 1$. Based on the Bayes theorem for conditional probability, we can solve the equation and find the threshold for the minimum readiness values to guarantee an acceptable level of controllability for the scenario $c$ as: $Min(f(CR, MR)_c) = 1 - AE$

Following this proposed methodology, it is possible to estimate readiness thresholds on replicable and controlled experimental scenarios $c$ for DSM systems, based on the parameters estimated from the fitting of the experimental data to the model.

# 5 Proof of concept of the model fitting

To test the applicability of the model in a real scenario, a proof of concept was conducted. This involved fitting data post-hoc from a driving simulator experiment which measured drivers' gaze behaviour and response to a critical takeover, from vehicle automation. It must be noted that this was a preliminary fitting to a simulator experiment that was not designed to validate the proposed model. This is a proof of concept to illustrate how this model can be used as a tool in a real experimental scenario and does not intend to provide evidence for validating the model.

## 5.1 Dataset description

The dataset used to fit the model was collected by the University of Leeds Driving Simulator, as part of the AdaptiVe project (grant agreement number 610428, see Louw et al. 2016, 2017, 2018). These studies were designed to measure the impact the level of engagement with the driving loop (Merat et al. 2019), on drivers' gaze, and their capability to avoid a crash during safety–critical transitions of control.

The dataset consists of 28 participants, with each experiencing two safety–critical trials, resulting in 56 datapoints. Participants were asked to drive a handsfree Level 2 automated vehicle (SAE 2021), resuming manual control whenever requested by the system. For each trial, drivers experienced 100 s of automated driving, before they received the TOR. While the automation was engaged, participants were requested to perform a 2-back task (Mehler et al. 2012), intended to increase workload.

The critical scenarios were preceded by an "uncertainty alert", which was a flashing yellow steering wheel, presented in the instrument cluster. Drivers were required to decide if a takeover was required, in response to a hard braking lead vehicle. The deceleration rate of the lead vehicle was 5 m/$s^2$, resulting in a time budget of 3 s time to collision. Drivers could only avoid a crash by either braking or changing lane. For more details about the methodology of the experiment, please refer to Louw et al. (2016).

## 5.2 Experimental variables

The variables extracted from the dataset to fit the model, included take-over response data and gaze-related data. The response data included crash outcome (crash/no crash), and the reaction time measured from the time the TOR, which was time drivers initiated their first relevant reaction to the scenario (braking/steering/or both, as defined by Louw et al. 2018).

The gaze-related data was extracted during the automated period, before the TOR and was used to generate the variables imputed in the model to estimate drivers' readiness levels. Given the low sample size of the dataset, this study opted to base the readiness levels using only drivers' gaze entropy (as described by Shiferaw et al. 2019), the amplitude-velocity ratio (AVR) of the drivers' eye blinks (as in Johns 2003), and the attention buffer levels of the driver, extracted from the AttenD algorithm (Kircher and Ahlström, 2017).

## 5.3 Model fitting and parameter estimation

The parameter selection for the model variables $\big(\mu_c, \sigma^2_c, i_{entropy}, i_{AVR}, i_{AttenD}\big)$ were selected with a random search technique (as described by Andradóttir 2006), testing a total combination of 100,000 combinations, and selecting the one which yielded the best fit of the model to the dataset. The log-likelihood ratio was used to access the goodness of fit of Monte Carlo simulations, generated by the model to the real experimental data. The fitting process was made individually for each trial, using the same model parameters, in a process described below:

- Every trial $n$ (in the of $N$ trials) presented an estimated readiness level $r_n$, and a response time $ToT_n$.
- For every trial, a total of 10,000 Monte Carlo simulations were generated, using $r_n$, and a set of parameters $\mathbb{P}$, generating a distribution of simulated reaction times $ToT_\mathbb{P}$.
- The log-likelihood ratio for the individual trial was calculated as $\log \mathcal{L}_n = \log(p$ ($\log \mathcal{L}_n = \log\left(p\left(ToT_n \in ToT\right)\right)$).
- The final log-likelihood ratio for the whole distribution for the set of parameters $x$ was calculated as $$log\mathcal{L}_\mathbb{P} = \sum_{n=1}^{n=N} log\mathcal{L}_n.$$

Finally, the model accuracy was calculated as the percentage of cases where the model was able to predict whether a trial resulted in a crash, given drivers' readiness levels. For the prediction of the crash outcomes, the model predicted that a trial $n$ would crash, given their readiness levels ($r_n$), if a simulation $s_{r_n}$ presented a response time ($ToT_s$) bigger the maximum TTC of the experimental scenario (3 s). Therefore, the simulation would correct predict the crash scenario if its outcome matched the observation of the real experimental trial ($ToT_n \bigwedge ToT_s > 3$). The model error was calculated as the average mismatch between the actual reaction time for each individual trial, and the respective model simulations.

## 5.4 Model output

The final $log\mathcal{L}_\mathbb{P}$ score for the model, after the fitting process was − 682.4. Figure 7 shows a histogram comparing the experimental reaction time data (orange) and the simulated reaction times derived from the model.

The fitted model was able to correct predict crash outcomes of all simulations for 45 of the 56 trials. Observing the predictions of each simulation individually, the model yielded a prediction accuracy of 88.5%, however, a relatively low F1 score of 0.452. By observing the confusion matrix (see Table 2), most of the failed predictions were found for the crash cases where the model interpreted that drivers were within a desirable readiness threshold, but crashed regardless (4512 out of 6605). In terms of reaction time prediction, the average prediction error margin for the model simulations

**Table 2** Confusion matrix for the model output

| Experimental trial outcome | | |
| --- | --- | --- |
| Model prediction | | |
| | Crash | No crash |
| Crash | 3721 | 2093 |
| No crash | 4512 | 45,874 |

was 0.534 s., with the RMSE for the average residuals from the simulations of each experimental trial = 0.560 s. Figure 5 shows that the model was able to replicate the behaviour of the original sample with good fidelity, excluding the extreme cases, where a crash occurred. As shown in Fig. 8, the predicted error for each individual trial (i.e., the distribution of simulated response times generated by the model), is evenly balanced between over and under estimations of drivers' real response time (Fig. 6).
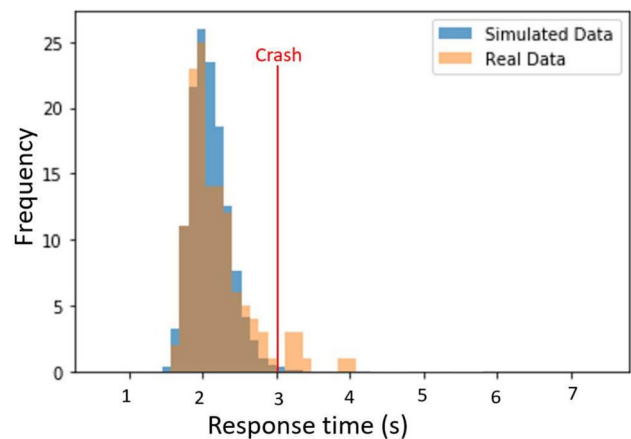


**Fig. 5** Histogram comparing the model output with the real experimental data. The Y axis is described in percentage, since the number of simulations (560,000) is higher than the number of actual data-points (56)
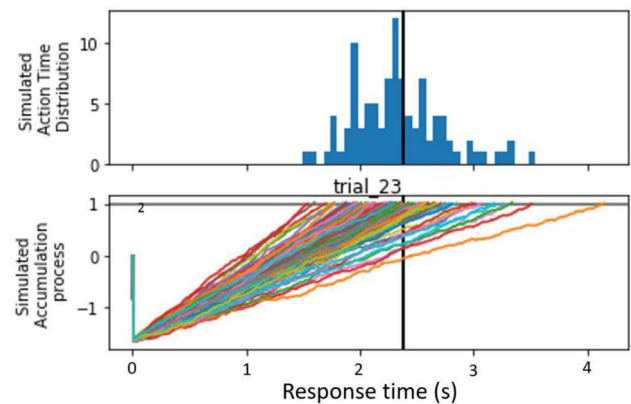


**Fig. 6** Example of the Monte Carlo simulations of one trial. The dark vertical line represents the real response time of the trial that was used to generate the simulation. The histogram on the top axis shows the distribution of simulated response time. The bottom axis shows the simulations of the evidence accumulation process over time, as described on Sect. 4.1

# 6 Discussion

The objective of this paper was to present a theoretical discussion about how the concept of driver readiness can be objectively estimated using controlled experimental data. A conceptual readiness estimation model is presented, and a methodology for defining readiness thresholds for DSM systems is proposed. The proposed model assumes a theoretical relationship between the process of evidence accumulation for decision-making (Ratcliff and Review 2004), and the concept of driver readiness. The paper further explores how the model can be used to estimate readiness thresholds based on the concept of controllability of a critical situation (International Standardization Organization 2020a, b). A proof of concept for the model application was then presented, using previously collected experimental data (Louw et al. 2016, 2017, 2018).

The results observed from the proof-of-concept fitting suggest that the abstract concept of driver readiness can be understood as an accumulation of cognitive and motoric resources over the course of the automated drive, and is directly correlated with the likelihood of crashes, and takeover time. The observed low F1 scores for the model crash prediction may be explained by the low sample size, and the fact that crashes (true positives) were a rare event, inflating the significance of missed predicted crashes (false negatives). It also must be noted that the binary prediction of crash/no crash does not consider near crashes, or how close a driver was to get involved in a collision. In that sense, the prediction of response time shows that the model was able to explain drivers' behaviour with a relative accuracy. This result provides evidence that reinforce the theories proposed by Mioch et al. (2017), Kim et al. (2022) and Marberger et al. (2018), as it was able to explain the reaction time variability in the observed distribution through an evidence accumulation model, based on ocular metrics. The preliminary fitting also suggests that ocular metrics seem to be a promising tool to assess drivers' state in future DSM systems, in line with current standards and recommendations for driver monitoring development (see EURO NCAP 2023).

This work contributes to the state of the art on readiness assessment models (e.g., Kim et al. 2022; Mioch et al. 2017; Mariajoseph et al. 2020; Baek et al. 2018; Deo and Trivedi 2020; Du et al. 2020), by proposing an applicable tool to estimate driver readiness, without relying on subjective assessment of readiness, as a ground truth. Another advantage of the proposed solution is the mechanistic nature of the model, since it allows researchers to directly observe the relationship between driver readiness estimation variables and the behaviour prediction, without relying on black box techniques (Hassija et al. 2024). The last contribution of this work is that the proposed methodology accounts for scenario variability, assuming that different safety–critical scenarios might require different readiness threshold values. This allows the use of experimental data to estimate how driver state is related to crashes during transitions of control from automation.

# 7 Conclusion and future work

The proposed methodology is a promising step forward on the understanding of the factors underlying driver state and readiness estimation in automated vehicles. However, it must be noted that the proposed model is constrained by its assumptions and conceptual limitations, that prevent it to be used as a real-time assessment tool for driver monitoring systems.

The limitations of the proposed model are linked to estimations of readiness using repeated, experimentally controlled data, which cannot be used for applied/real-time predictions. Also, as mentioned in the model assumptions, the estimation of readiness thresholds and drift rate of the evidence accumulation is depended to the scenario complexity. In other words, the model operates under the assumption that more complex scenarios would lead to slower evidence accumulation rate. These limitations imply that the parameters fitted to the model are not generalizable, and a new parameter fitting is needed for every distinct scenario the model is used for. At last, being a mechanistic model, it can provide explicability, but its rigidness lacks the abstraction of machine learn models to find nuanced relationships between variables. For this reason, the proposed model limits itself as a conceptual model, which provides theoretical insights, however, it is not designed as a tool for real-time readiness estimation.

It must be noted that despite having an overall good fitting, the prediction error for crash outcome was unbalanced, favouring false negatives. This paper has provided a list of metrics required for assessing driver readiness, which serve as an input for the model. However, it is acknowledged that this is a constrained list of possible variables, based on previous research. Further studies are required to produce a more comprehensive list of the operational metrics related to driver readiness. It should also be noted that the practical application of this model to the experimental data has some limitations. The accuracy of the model fitting presented in this manuscript is limited by a low sample size, and the lack of a controlled design to account for driver state variability. Moreover, since the dataset used for the model fitting was not designed with the intention to validate the model's applicability, the current output is unable to diagnose whether this solution is precise enough to be used in its current iteration. Despite these limitations, the model provides an improvement on previously presented

concepts in this area, by considering the mechanistic process of driver readiness during critical takeover situations. The next step in this research is to build upon what has been presented here by creating specific experimental scenarios that allow the validation of the proposed model.

## Declarations

## References

Andradóttir S (2006). An overview of simulation optimization via random search. Handbooks in operations research and management science, Elsevier, 13:617–631.

Baek JW, Han BG, Kim KJ, Chung YS, Lee SI (2018) Real-time drowsiness detection algorithm for driver state monitoring systems. International conference on ubiquitous and future networks, ICUFN, 73–75. https://doi.org/10.1109/ICUFN.2018.8436988

Bowden VK, Loft S, Wilson MK, Howard J, & Visser TA (2019) The long road home from distraction: Investigating the time-course of distraction recovery in driving. Accident Analysis & Prevention 124:23–32.

Carsten O, Lai FCH, Barnard Y, Jamson AH, Merat N (2012) Control task substitution in semiautomated driving. Hum Factors: J Hum Factors Ergon Soc 54(5):747–761. https://doi.org/10.1177/0018720812460246

Choi D, Sato T, Ando T, Abe T, Akamatsu M, Kitazaki S (2020) Effects of cognitive and visual loads on driving performance

after take-over request (TOR) in automated driving. Appl Ergon 85:103074. https://doi.org/10.1016/J.APERGO.2020.103074

Clark H, Feng J (2017) Age differences in the takeover of vehicle control and engagement in non-driving-related activities in simulated driving with conditional automation. Accid Anal Prev 106:468–479. https://doi.org/10.1016/J.AAP.2016.08.027

Goodridge CM, Billington J, Markkula G, Wilkie RM (2022). Error accumulation when steering toward curves. Journal of experimental psychology: human perception and performance 49(6): 821. https://doi.org/10.31234/OSF.IO/2E7P4

de Winter JCF, Happee R, Martens MH, Stanton NA (2014) Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence. Transport Res F: Traffic Psychol Behav 27:196–217. https://doi.org/10.1016/j.trf.2014.06.016

Deo N, Trivedi MM (2020) Looking at the driver/rider in autonomous vehicles to predict take-over readiness. IEEE Trans Intell Veh 5(1):41–52. https://doi.org/10.1109/TIV.2019.2955364

Du N, Zhou F, Pulver EM, Tilbury DM, Robert LP, Pradhan AK, Yang XJ (2020) Predicting driver takeover performance in conditionally automated driving. Accid Anal Prev 148:105748

Durrani U, Lee C, Shah D (2021) Predicting driver reaction time and deceleration: Comparison of perception-reaction thresholds and evidence accumulation framework. Accident Analysis & Prevention 149:105889.

Endsley MR (1995) Toward a theory of situation awareness in dynamic systems. Hum Factors: J Hum Factors Ergon Soc 37(1):32–64. https://doi.org/10.1518/001872095779049543

Eriksson A, Stanton NA (2017) Driving performance after self-regulated control transitions in highly automated vehicles. Hum Factors: J Hum Factors Ergon Soc 59(8):1233–1248. https://doi.org/10.1177/0018720817728774

EURO NCAP (2023) Assisted driving highway assist systems. Available in https://www.euroncap.com/media/80151/euro-ncap-ad-test-and-assessment-protocol-v21.pdf. Accessed 14 Feb 2024.

Evinger C, Manning KA, Sibony PA (1991) Eyelid movements. Mechanisms and normal data. Investigative ophthalmology & visual science 32(2): 387–400.

Feldhutter A, Kroll D, Bengler K (2018) Wake up and take over! the effect of fatigue on the take-over performance in conditionally automated driving. IEEE Conference on intelligent transportation systems, proceedings, ITSC, pp 2080–2085. https://doi.org/10.1109/ITSC.2018.8569545

Gartenberg D, Breslow L, McCurry JM, Trafton JG (2014) Situation awareness recovery. Hum Factors: J Hum Factors Ergon Soc 56(4):710–727. https://doi.org/10.1177/0018720813506223

Georg J-M, Hecht T, Diermeyer F, Wang C (2017) Development of a human driver model during highly automated driving for the ASIL controllability classification A novel approach on modeling takeover scenarios for highly automated driving.

Gershon P, Mehler B, Reimer B (2023) Driver response and recovery following automation initiated disengagement in real-world hands-free driving. Traffic Inj Prev 24(4):356–361

Gold JI, Shadlen MN (2007) Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. Neuron 36(2):299–308. https://doi.org/10.1016/S0896-6273(02)00971-6

Gold C, Damböck D, Lorenz L, Bengler K (2013) "Take over!" How long does it take to get the driver back into the loop? Proc Hum Factors Ergon Soci Annu Meeting 57(1):1938–1942. https://doi.org/10.1177/1541931213571433

Gonçalves RC, Louw TL, Madigan R, Quaresma M, Romano R, Merat N (2022) The effect of information from dash-based human-machine interfaces on drivers' gaze patterns and lane-change manoeuvres after conditionally automated driving. Accid Anal Prev. https://doi.org/10.1016/J.AAP.2022.106726

Gonçalves RC, Louw TL, Madigan R, Merat N (2019a) Using Markovian chains to understand drivers' gaze transitions during lane-changes in manual vs. automated driving. Driving Accessment

Gonçalves RC, Louw T, Markkula G, Merat N (2019b) Applicability of risky decision-making theory to understand drivers' behaviour during transitions of control in vehicle automation.

Goodridge CM, Mole CD, Billington J, Markkula G, Wilkie RM (2022) Steering is initiated based on error accumulation. J Exp Psychol Hum Percept Perform 48(1):64–76. https://doi.org/10.1037/XHP0000970

Guo H, Zhang Y, Cai S, Chen X (2021) Effects of level 3 automated vehicle drivers' fatigue on their take-over behaviour: a literature review. J Adv Transp 2021:1–12

Hassija V, Chamola V, Mahapatra A, Singal A, Goel D, Huang K, Hussain A (2024) Interpreting black-box models: a review on explainable artificial intelligence. Cognit Comput 16(1):45–74

Hecht T, Feldhütter A, Radlmayr J, Nakano Y, Miki Y, Henle C, Bengler K (2019) A review of driver state monitoring systems in the context of automated driving. Adv Intell Syst Comput 823:398–408. https://doi.org/10.1007/978-3-319-96074-6_43/COVER

Hergeth S, Lorenz L, Vilimek R, Krems JF (2016) Keep your scanners peeled. Hum Factors 58(3):509–519. https://doi.org/10.1177/0018720815625744

International Standardization Organization (2020a) ISO/TR 21959-1:2020(en), Road vehicles—human performance and state in the context of automated driving. https://www.iso.org/obp/ui/#iso:std:iso:tr:21959:-1:ed-2:v1:en. Accessed 14 Feb 2024.

International Standarization Organization (2020b) ISO 26262 Road vehicles – Functional safety. https://www.iso.org/standard/43464.html. Accessed 14 Feb 2024.

Johns MW (2003). The amplitude-velocity ratio of blinks: a new method for monitoring drowsiness. Sleep, 26(SUPPL).

Kang HB (2013) Various approaches for driver and driving behavior monitoring: a review. In: Proceedings of the IEEE international conference on computer vision workshops, pp 616–623

Kelley A, McAtee A, Duffy M, Feltman K (2023) Evaluation of Inter-subject variability in physiological metrics and workload perception: implications for operator state monitoring

Kim H, Kim W, Kim J, Lee SJ, Yoon D (2018) Design of driver readiness evaluation system in automated driving environment. 9th International conference on information and communication technology convergence: ICT convergence powered by smart intelligence, ICTC 2018:300–302. https://doi.org/10.1109/ICTC.2018.8539408

Kim J, Kim W, Kim HS, Lee SJ, Kwon OC, Yoon D (2022) A novel study on subjective driver readiness in terms of non-driving related tasks and take-over performance. ICT Express 8(1):91–96. https://doi.org/10.1016/J.ICTE.2021.04.008

Kircher K, Ahlström C (2017) The driver distraction detection algorithm AttenD. Driver Distraction Inattention: Adv Res Countermeas 1:327–348. https://doi.org/10.1201/9781315578156-23

Körber M, Cingel A, Zimmermann M, Bengler K (2015) Vigilance decrement and passive fatigue caused by monotony in automated driving. Procedia Manuf 3:2403–2409. https://doi.org/10.1016/J.PROMFG.2015.07.499

Li L, Zhong B, Hutmacher C, Liang Y, Horrey WJ, Xu X (2020) Detection of driver manual distraction via image-based hand and ear recognition. Accid Anal Prev 137:105432. https://doi.org/10.1016/J.AAP.2020.105432

Louw T, Merat N (2017) Are you in the loop? Using gaze dispersion to understand driver visual attention during vehicle automation. Transp Res Part C: Emerg Technol 76:35–50. https://doi.org/10.1016/J.TRC.2017.01.001

Louw T, Madigan R, Carsten O, Merat N (2016) Were they in the loop during automated driving? Links between visual attention and crash potential. Injury Prev: J Int Soc Child Adolesc Injury Prev 23(4):281–286. https://doi.org/10.1136/injuryprev-2016-042155

Louw T, Markkula G, Boer ER, Madigan R, Carsten OMJ, Merat N (2018) Coming back into the loop: drivers' perceptual-motor performance in critical events after automated driving. Accid Anal Prev 108:9–18. https://doi.org/10.1016/J.AAP.2017.08.011

Louw T, Kuo J, Romano R, Radhakrishnan V, Lenné MG, Merat N (2019) Engaging in NDRTs affects drivers' responses and glance patterns after silent automation failures. Transportation research part F: traffic psychology and behaviour 62:870–882.

Ma Z, Zhang Y (2021) Drivers trust, acceptance, and takeover behaviors in fully automated vehicles: effects of automated driving styles and driver's driving styles. Accid Anal Prev 159:106238. https://doi.org/10.1016/J.AAP.2021.106238

Marberger C, Mielenz H, Naujoks F, Radlmayr J, Bengler K, Wandtner B (2018) Understanding and applying the concept of "Driver availability" in automated driving. Adv Intell Syst Comput 597:595–605. https://doi.org/10.1007/978-3-319-60441-1_58/COVER

Mariajoseph M, Gallina B, Carli M, Bibbo D (2020) A physiology-based driver readiness estimation model for tuning ISO 26262 controllability. IEEE vehicular technology conference. https://doi.org/10.1109/VTC2020-SPRING48590.2020.9129132

Markkula G, Boer E, Romano R, Merat N (2018) Sustained sensori-motor control as intermittent decisions about prediction errors: computational framework and application to ground vehicle steering. Biol Cybern 112(3):181–207. https://doi.org/10.1007/s00422-017-0743-9

Markkula G, Romano R, Waldram R, Giles O, Mole C, Wilkie R (2019) Modelling visual-vestibular integration and behavioural adaptation in the driving simulator. Transport Res f: Traffic Psychol Behav 66:310–323. https://doi.org/10.1016/j.trf.2019.07.018

Markkula G, Madigan R, Nathanael D, Portouli E, Lee YM, Dietrich A, Merat N (2020) Defining interactions: a conceptual framework for understanding interactive behaviour in human and automated road traffic. Theor Issues Ergon Sci 21(6):728–752

Markkula G (2014) Modeling driver control behavior in both routine and near-accident driving. In Proceedings of the human factors and ergonomics society annual meeting 58(1): 879–883. Sage CA: Los Angeles, CA: SAGE Publications

Matthews G, Neubauer CE, Saxby DJ, Langheim LK (2019) Driver fatigue: The perils of vehicle automation. In Advances in Traffic Psychology, pp 127–140). CRC Press

McDonald AD, Alambeigi H, Engström J, Markkula G, Vogelpohl T, Dunne J, Yuma N (2019) Toward computational simulations of behavior during automated driving takeovers: a review of the empirical and modeling literatures. Hum Factors 61(4):642–688

Merat N, Seppelt B, Louw T, Engström J, Lee JD, Johansson E, Green CA, Katazaki S, Monk C, Itoh M, McGehee D, Sunda T, Unoura K, Victor T, Schieben A, Keinath A (2019) The "Out-of-the-Loop" concept in automated driving: proposed definition, measures and implications. Cogn Technol Work 21(1):87–98. https://doi.org/10.1007/s10111-018-0525-8

Mehler B, Reimer B, Coughlin JF (2012) Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task: an on-road study across three age groups. Human factors 54(3):396–412

Michon JA (1985) A critical view of driver behavior models: what do we know, what should we do?. In: Evans, L., Schwing, R.C. (eds) Human behavior and traffic safety. Springer, Boston, MA. https://doi.org/10.1007/978-1-4613-2173-6_19

Mioch T, Kroon L, Neerincx M (2017) Driver readiness model for regulating the transfer from automation to human control figure 1. Truck platooning experiment with a CTS agent. Proceedings of

the 22nd International Conference on Intelligent User Interfaces. https://doi.org/10.1145/3025171

Mole CD, Lappi O, Giles O, Markkula G, Mars F, Wilkie RM (2019) Getting back into the loop: the perceptual-motor determinants of successful transitions out of automated driving. Hum Factors: J Hum Factors Ergon Soc. https://doi.org/10.1177/0018720819829594

Moslemi N, Soryani M, Azmi R (2021) Computer vision-based recognition of driver distraction: a review. Concurr Comput: Pract Exp 33(24):e6475

Payre W, Cestac J, Delhomme P (2014) Intention to use a fully automated car: attitudes and a priori acceptability. Transport Res f: Traffic Psychol Behav 27:252–263. https://doi.org/10.1016/j.trf.2014.04.009

Perello-March JR, Burns CG, Birrell SA, Woodman R, Elliott MT (2022) Physiological measures of risk perception in highly automated driving. IEEE Trans Intell Transp Syst 23(5):4811. https://doi.org/10.1109/TITS.2022.3146793

Pradhan AK, Pai G, Knodler MA (2021) Driver's mental models of advanced vehicle technologies: a proposed framework for identifying and predicting operator errors. https://rosap.ntl.bts.gov/view/dot/60329. Accessed 14 Feb 2024

Radhakrishnan V, Merat N, Louw T, Gonçalves RC, Torrao G, Lyu W, Puente Guillen P, Lenné MG (2022) Physiological indicators of driver workload during car-following scenarios and takeovers in highly automated driving. Transport Res f: Traffic Psychol Behav 87:149–163. https://doi.org/10.1016/J.TRF.2022.04.002

Ratcliff R (1978) A theory of memory retrieval. Psychol Rev 85(2):59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff R, Smith PL (2004) A comparison of sequential sampling models for two-choice reaction time. Psychological review, 111(2):333.

Ratcliff R, Smith PL, Brown SD, McKoon G (2016) Diffusion decision model: current issues and history. Trends Cogn Sci 20(4):260–281. https://doi.org/10.1016/J.TICS.2016.01.007

Reason J (2004) Beyond the organisational accident: the need for "error wisdom" on the frontline. BMJ Qual Saf 13(suppl 2):ii28–ii33

SAE (2021) Summary of SAE international's levels of driving automation for on-road vehicles. https://www.sae.org/misc/pdfs/automated_driving.pdf. Accessed 21 Jan 2024

Schnebelen D, Charron C, Mars F (2020) Estimating the out-of-the-loop phenomenon from visual strategies during highly automated driving. Accid Anal Prev 148:105776. https://doi.org/10.1016/J.AAP.2020.105776

Shiferaw B, Downey L, Crewther D (2019) A review of gaze entropy as a measure of visual scanning efficiency. Neurosci Biobehav Rev 96:353–366. https://doi.org/10.1016/J.NEUBIOREV.2018.12.007

Shojaeizadeh M, Djamasbi S, Trapp AC (2016) Density of gaze points within a fixation and information processing behavior. In Universal access in human-computer interaction. Methods, techniques, and best practices: 10th international conference, UAHCI 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17–22, 2016, Proceedings, Part I 10:465–471). Springer International Publishing

Spanfelner B et al (2012) Challenges in applying the ISO 26262 for driver assistance systems. Tagung Fahrerassistenz

Sun Y, Yan P, Li Z, Zou J, Hong D (2020) Driver fatigue detection system based on colored and infrared eye features fusion. Comput Mater Continua 63(3):1563–1574

Tseng VWS, Abdullah S, Costa J, Choudhury T (2018) AlertnessScanner: what do your pupils tell about your alertness. In Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, pp 1–11

Victor T (2009) Method and arrangement for controlling vehicular subsystems based on interpreted driver activity. U.S. Patent No. 7,639,148. Washington, DC: U.S. Patent and Trademark Office.

Walker HEK, Trick LM (2018) Mind-wandering while driving: the impact of fatigue, task length, and sustained attention abilities. Transport Res f: Traffic Psychol Behav 59:81–97. https://doi.org/10.1016/J.TRF.2018.08.009

Wilbanks BA, Aroke E, Dudding KM (2021) Using eye tracking for measuring cognitive workload during clinical simulations: literature review and synthesis. CIN: Comput Inform Nurs 39(9):499–507

Yoon SH, Ji YG (2019) Non-driving-related tasks, workload, and takeover performance in highly automated driving contexts. Transport Res f: Traffic Psychol Behav 60:620–631. https://doi.org/10.1016/J.TRF.2018.11.015

Zeeb K, Buchner A, Schrauf M (2015) What determines the take-over time? An integrated model approach of driver take-over after automated driving. Accid Anal Prev 78:212–221. https://doi.org/10.1016/j.aap.2015.02.023

Zeeb K, Buchner A, Schrauf M (2016) Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. Accid Anal Prev 92:230–239. https://doi.org/10.1016/j.aap.2016.04.002

Zhou F, Yang XJ, de Winter JCF (2021) Using eye-tracking data to predict situation awareness in real time during takeover transitions in conditionally automated driving. IEEE Trans Intell Transp Syst. https://doi.org/10.1109/TITS.2021.3069776