Transportation Research Part F: Psychology and Behaviour 108 (2025) 54–72

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

journal homepage: www.elsevier.com/locate/trf

When is more actually better? expert opinions on assessment of situation awareness in relation to safe driving

Rins de Zwart [a,*], Reinier J. Jansen [a], Cheryl Bolstad [b], Mica R. Endsley [c], Petya Ventsislavova [d], Joost de Winter [e], Mark S. Young [f]

[a] SWOV Institute for Road Safety Research, Henri Faasdreef 312, 2492 JP The Hague, Netherlands
[b] Sandia National Laboratories, Innovation Parkway Office Center (IPOC), 1611 Innovation, Parkway SE, USA
[c] SA Technologies, 5301 S. Superstition Mountain Drive, #104377, Gold Canyon, AZ 85118, USA
[d] Nottingham Trent University, 50 Shakespeare Street, Nottingham NG1 4FQ, UK
[e] Delft University of Technology, Department of Cognitive Robotics, Mekelweg 2, 2628 CD Delft, the Netherlands
[f] Transportation Research Group, University of Southampton, Boldrewood Innovation Campus, Southampton, SO16 7QF, UK

ARTICLE INFO

Keywords:
Situation awareness
Driving safety
Safety criterion
Position paper
Driving performance

ABSTRACT

The use of situation awareness (SA) measures to assess relative safety in driving is common, with higher levels of SA being interpreted as safer. These relative interpretations do not allow researchers to determine whether the level of SA could be considered “safe” or “unsafe”. In contrast to such interpretations based on relative performance, the current position paper explores the potential for a normative interpretation of situation awareness with regard to safety assessment in driving. A series of expert interviews yielded viewpoints on the current relation between SA and safe driving, theoretical underpinnings for a normative approach, and potential actions towards an SA criterion for safe or unsafe driving. Methodological challenges regarding a normative approach are discussed together with considerations towards a weighted criterion-based approach to SA. The selection of SA requirements relevant for safety and the differentiation and weighting of these requirements on high and lower importance is presented. A method towards objective determination of relevance and weight of SA requirements may increase the usefulness of SA measures for assessment of safety in a driving context.

1. Introduction

In order to safely take part in the traffic system, drivers have to be able to observe and interpret their surroundings; in other words, they need to be aware of the situation around them. Situation Awareness (SA) concerns “knowing what is going on” (Endsley, 1995a). Depending on the type of task under investigation, such as air traffic control, aviation, or surface transport, this can entail the current position of the vehicle, the relative position and behavior of others, and a prediction about how these might change in the future. In the automotive domain, the concept of driver situation awareness has been used for example, to measure the effects of different distractions (e.g. Chandrasekaran et al., 2019; Wulf et al., 2013) or interfaces (e.g. Lee et al., 1999; Van den Beukel et al., 2016; Wulf et al.,

* Corresponding author.
E-mail addresses: rins.de.zwart@swov.nl (R. de Zwart), reinier.jansen@swov.nl (R.J. Jansen), cbolstad@sandia.gov (C. Bolstad), mica@satechnologies.com (M.R. Endsley), petya.petrova@ntu.ac.uk (P. Ventsislavova), j.c.f.dewinter@tudelft.nl (J. de Winter), m.s.young@soton.ac.uk (M.S. Young).

https://doi.org/10.1016/j.trf.2024.11.011
Received 28 June 2024; Received in revised form 11 November 2024; Accepted 11 November 2024
Available online 26 November 2024
1369-8478/

2013). The construct of SA has also been applied in studies about the impact of vehicle automation on driver SA (e.g. Ma & Kaber, 2005), and how hazard perception skills can be improved by training (e.g. Van der Kint et al., 2024). Finally, SA has been used in studies on driver monitoring systems (e.g. Hijaz et al., 2019; McKerral et al., 2023). Increased SA has frequently been interpreted in terms of increased safety. This paper explores whether such an interpretation of SA is warranted. Possibilities for a normative interpretation are proposed and discussed with experts in the field of SA, yielding methodological challenges and potential paths towards improved utility of the SA concept.

1.1. Background

There are different definitions of SA, varying on the model or theory being used (Endsley, 1995a; Gugerty, 1998; Salmon et al., 2012). One of the most popular definitions of SA is given by Endsley (1988): “Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future”. The accompanying theory splits SA into three levels: level 1 SA (perception) concerns the perception of the environment and the information contained within, level 2 SA (comprehension) is about the understanding of the information gathered at level 1, and level 3 SA (projection) concerns the future projection of the information (Endsley, 1995b). Amongst other definitions of SA are those addressing the concept of ‘shared awareness’, including the SA that other road users have (Endsley and Jones, 2001; Salmon et al., 2012). The present study views SA from the perspective of an individual driver.

Historically, considerations of SA came from the military and aviation context (Endsley, 1988: Taylor, 1990). Pilots and air traffic controllers are required to process relatively large amounts of information in a short time window and make decisions based on projections of this information. Various methods of measuring SA are used when evaluating how best to utilize different interfaces to support the pilot or controller in maintaining sufficient SA. Such methods can generally be categorized into 4 groups (Endsley, 2021a): process and performance measures (both of which capture SA in real-time), direct measures of SA that probe a person’s SA through direct questioning of the individual, and subjective measures of SA (that ask a participant to rate their own SA). Process and performance measures infer SA based on other measures (such as communications or actions taken). Subjective measures of SA have been found to be more closely related to people’s level of confidence in their knowledge and are not correlated with objective measures of SA (Endsley, 2020a).

Direct, objective measures determine SA by asking the operator to answer multiple questions regarding the state of the world that is relevant to SA for a given task or role. What the operator needs to know is described in a set of SA requirements. SA requirements are determined via an analysis of the information needs of the task or role using techniques such as a Goal-Directed Task Analysis (GDTA) (Endsley, 1993) or Hierarchical Task Analysis (HTA) (Stanton, 2006). A GDTA or HTA commonly results in a list of many information requirements that are needed to make relevant decisions and perform well (Walker et al., 2018). The GDTA or HTA is then analyzed to determine a set of questions that covers a broad range of SA requirements suitable for administering in experimental settings. How this reduction from a full set of SA requirements to a set of questions is done, and if a GDTA or HTA served as the starting point for the full list of requirements, is not always clarified in research papers. Which SA requirements are used may differ between studies and depends in part on their relevance to the domain, task, research topic, and the (practical) ability to adequately test for awareness of those requirements (e.g., Rebensky et al., 2022). However, the process used to define the selection of requirements used for testing is often unclear, resulting in a selection that could differ between studies without explanation. Details on how to conduct a GDTA and develop SA queries are provided in Endsley (2000a, 2021b), along with sets of queries that have been determined for many different domains, including driving. An example of queries selected from a full driving related GDTA is available in Bolstad (2001).

In addition to using SA measures as a way to evaluate differences between interfaces or situations, measures of SA may potentially be utilized as a way of training potential candidates for specialized roles, such as air traffic control personnel or pilots (Certification: Pilots, Flight Instructors, and Ground Instructors, 2023). Both air traffic control personnel and pilots are extremely well-trained professionals in a controlled setting where demands on attentional resources are very high, and tailored programs to train SA using feedback from SA queries have been developed and validated, although the results have been mixed (Kaber et al., 2006; Strater & Bolstad, 2008). While the road traffic environment is different than the air traffic control or aviation domains, SA (or a lack thereof) is equally important for successful performance and could still mean the difference between success (safe arrival at the target destination) and (potentially fatal) failure.

While there is no full test of SA administered during the examination for a driver’s license, there are measures that are similar to SA (Horswill & McKenna, 2004). Hazard perception training or hazard anticipation tests are the closest equivalent of SA testing in a driving context and are already part of the driver licensing system in several countries (e.g., Australia, the United Kingdom). It should be noted that while hazard perception and situation awareness are related, and hazard awareness has been equated to level 3 SA by some (Endsley, 2018, 2020b), there are some important differences between the two concepts. Where hazard perception is limited to information about potential hazards, SA is broader in what information could be included, such as upcoming infrastructure changes, navigational information, and other road users that are not considered hazardous (Evans & Macdonald, 2002). In order to utilize SA for assessment within driver examination or as a criterion for safe driving, the development of a normative approach should be considered.

1.2. From a relative to a normative interpretation of SA

SA has been used as a measure of safety when determining the impact of, for example, secondary tasks and vehicle interfaces (Fisher & Strayer, 2014; Gugerty et al., 2003; Khoda Bakhshi et al., 2021; Schömig & Metz, 2013; Schroeter & Steinberger, 2016). Such studies typically feature comparisons between experimental conditions, yielding insights into relative SA performance (see panel A in Fig. 1).

Fig. 1. Comparison of situation awareness (SA) performance between two hypothetical experimental conditions, based on relative performance (panel A), a reference scale (panel B), a safety criterion (panel C), and a context-dependent criterion (panel D).

However, no conclusions can be drawn on whether to interpret any given level of SA exhibited as 'excellent' versus 'good', or 'fair' versus 'poor', as such interpretation requires a reference scale (panel B). Moreover, a criterion is required if one additionally wishes to interpret SA performance in terms of safety (panel C). Such a criterion could potentially be established as the level of SA relevant to some safety performance indicator (e.g., crash involvement, minimal time-to-collision). Further, such a criterion could be dependent on context (panel D), such that the level of SA needed for any piece of information at any given time is related to the criticality of that information within the current context.

How can an SA criterion for safe driving be established? One prerequisite is an unequivocal positive correlation between SA and a safety performance indicator (SPI). A *meta*-analysis across several domains by Endsley (2021a) looked at the predictiveness and sensitivity of two direct objective measures of SA, the Situation Awareness Global Assessment Technique (SAGAT) and the Situation Present Assessment Method (SPAM). SA measures acquired by using both SAGAT and SPAM correlated positively with performance measures, with mean Pearson's $r = 0.459$ for the SAGAT and $r = 0.411$ for SPAM (Endsley, 2021a). However, these results were achieved by combining studies across a number of domains in addition to driving, such as aviation; submarine management; military; and process control. While there are some comments on the *meta*-analysis (see Bakdash et al., 2020), it is one of the few overviews on the predictiveness of direct SA measures. A study by Salmon et al. (2009) shows similar results with SAGAT correlating with performance measures ($r = 0.662$). Specific SA queries and driving performance measures varied across studies. Consequently, correlations between specific driving-related SPIs and SA cannot be derived from this analysis.

When looking at individual driving simulator studies, several report positive correlations between measures of SA and SPIs. Lee et al. (1999) modeled SA accuracy and confidence as an intervening variable, which helped to explain the influence of in-vehicle messages on driving behavior (e.g., crash involvement, speeding). Likewise, Kass et al. (2007) viewed SA as an intervening variable between cell phone distraction and driving infractions (e.g., crash involvement, stop signs missed, speeding, off-road excursions, centerline crossings). In contrast, Soliman (2010) examined SA by selecting participants showing either low or high SA in a first experiment. In a second experiment, participants with higher SA yielded fewer driving violations (e.g., crash involvement, speeding, stop sign neglect, centerline crossings). Several examples of positive correlations between SA measures and safe driving are further described in a review by Gugerty (2011).

The above studies suggest that SA performance may indeed positively correlate with driving SPIs. Endsley (2000a) stated that several individual factors (e.g., abilities, experience, training) as well as task/system factors (e.g., system capability, interface design, complexity, automation) are involved in the decision-making and task execution process, all of which influence the relation between SA performance and task performance. This makes it difficult to interpret how safe a certain amount of SA will be. Endsley views this relation as a probabilistic link: good SA performance increases the probability of good decisions and performance, but does not guarantee them (Endsley, 2000a, 2021b). Conversely, poor SA performance increases the probability of poor decisions and performance, but does not necessarily result in serious errors. People with poor SA may avoid accidents due to luck, for example.

Further investigation on the relationship between SA and safe driving could help to better understand the strength of this (probabilistic) relationship, as well as which factors influence the relationship. Performing an overarching (meta-)analysis across existing studies may be hindered by at least three inconsistencies in methodology: 1) scenario selection, 2) SA requirement selection, and 3) SA scoring. Regarding inconsistent scenario selection, some studies in the driving domain do not include hazardous or critical scenarios (e.g., Ma & Kaber, 2005), whereas other studies do (Lee et al., 1999; Kass et al., 2011; Soliman, 2010; Clark et al., 2017).

Wickens (1995) has pointed out that the SA needed for normal operations may be very different from that needed for responding to critical, non-normal events. Van den Beukel and Van der Voort (2017) explicitly distinguished between 'hazardous' scenarios (e.g., complex roads, vehicles passing on the right) and 'critical' scenarios (e.g., emergency braking, cutting in), which differed in required mental effort, likely because only critical scenarios required intervention. Evaluating three HMI concepts, they found that the best-performing concept in hazardous scenarios turned out to be the worst-performing concept in critical scenarios. This contradictory finding illustrates the need to either focus on critical scenarios or to include the criticality of a scenario in a prospective *meta*-analysis. If the criticality of a scenario indeed influences the relation between SA and safe driving, it could potentially help to distinguish between safe and unsafe driving performance (as hypothesized in panel D of Fig. 1).

Concerning inconsistent SA requirement selection, seemingly irrelevant queries may be presented to participants to reduce potential artificial cueing (Endsley, 2000b), but they should not be part of the analysis of the association between SA performance and safe driving. Some automotive studies explicitly distinguished between queries probing relevant and irrelevant requirements. For example, Kass et al. (2007) only presented questions pertaining to the driving task. Radlmayer et al. (2018) presented queries related and unrelated to the dynamic driving task, but only analyzed the former. Dattel et al. (2011) separately analyzed relevant queries (e.g., own vehicle speed relative to surrounding traffic) and irrelevant queries (e.g., color of the coat of a pedestrian). However, other studies did not make a distinction based on relevancy, even though it seems a distinction may have been warranted (e.g., Ma & Kaber, 2005; Soliman, 2010).

Finally, inconsistencies in SA scoring are the result of differences in the operationalization of SA measures across studies. Endsley (2021a) argued that SA should ideally be analyzed per query or at least per SA level (e.g., perception, comprehension, projection), as this yields a significant increase in sensitivity compared to using a single combined SA score. Several driving simulator studies have analyzed SA per SA level exclusively (e.g., Van den Beukel et al., 2016) or in addition to a single combined SA score (e.g., Blömacher et al., 2018; Jannat et al., 2018; Tan et al., 2022). However, other studies have reported only a single combined SA score (e.g., Lee et al., 1999; Kass et al., 2007).

When further investigating SA for safe driving, the above inconsistencies may be resolved by excluding certain studies (e.g., studies without explicit attention to safety-relevant queries) and by adding factors (e.g., scenario criticality), with the caveat that countermeasures require an increased sample size to find significant effects of a certain magnitude. For the remainder of this paper, and given the positive correlations found in individual studies, a probabilistic positive correlation between SA performance and safe driving is assumed. When it comes to establishing a safety criterion (Fig. 1), the possibility exists that meeting or failing an SA criterion is, by itself, not a guarantee for safe or unsafe driving performance, respectively. Thus, meeting an SA-based performance criterion will likely be one of multiple prerequisites for safe driving.

1.3. Research aim and approach

The relationship between SA and safe driving remains only partially clear. Results regarding SA are often only relative (comparing one condition to another), reducing their usability merely to comparisons. The usefulness of SA as a metric for safe driving is limited due to this unclear relation and the lack of a normative interpretation of SA results. Without a systematic approach to selection and scoring of relevant SA requirements, inconsistencies between studies likely occur. A structured approach would enable further development of SA measures for use in driving safety assessment.

Therefore, this study aims to explore the potential of a normative interpretation of SA in relation to driving safety. Semi-structured interviews were held with experts in the domain of situation awareness to discuss the theoretical basis for a normative approach, as well as the selection and weighting of different SA requirements in an effort to enhance SA measurements for driving safety. An overview of the viewpoints of the different experts is presented. A discussion comparing and contrasting these viewpoints is presented for the theoretical basis for a normative relation and methodological steps.

2. Method

Semi-structured interviews were held with experts in the field of SA to explore if and how a normative interpretation of SA in relation to safe driving is feasible. The semi-structured approach allowed for comparison between responses, while also creating a degree of flexibility to follow up on responses related to the unique background of each expert involved. The study was approved by the ethical committee of the SWOV Institute for Road Safety Research.

2.1. Participants

Experts on situation awareness were invited to share their views on the relation between SA and safe driving and potential steps forward. Prospective experts were selected based on their publication history in the area of SA and their impact based on citations within the SA domain. Preference was given to experts with a focus on the use of task analysis to determine SA requirements, and with knowledge of SAGAT, as these methods are well established and can be considered the current leading approach to measuring SA. A total of nine experts were contacted via email with a request to participate in an interview regarding SA and safe driving. Five experts responded positively and were available in the allocated time window between October 9th and October 26th, 2023.

2.2. Materials

Input from the experts was gathered using semi-structured interviews to ensure the same topics and terminology were used for all participants. As a preparation, experts were given a survey containing two images of traffic scenarios from the perspective of the driver (see Appendix A). The scenarios originated from a hazard perception test (Vlakveld, 2011), and were accompanied by the request to list a number of SA requirements that seemed relevant to the situation shown. These example requirements were used in the interviews to facilitate discourse.

2.3. Procedure

The preparatory surveys were sent to experts approximately two weeks before each interview and experts were asked to fill out and return it in advance. The interviews were held online using Microsoft Teams, with a duration of approximately one hour.

The interview covered a theoretical interpretation of the construct of SA in relation to safe driving, followed by methodological questions and questions on operationalization. The interview consisted of:

- General questions on the relation between 'SA' and safe driving.
- Questions on the use of SA to assess safe driving.
- Questions on selection of SA requirements.
- Questions on measuring SA.

During the interview, the images from the preparatory survey were shown, e.g., a traffic scenario from the perspective of the driver (see Appendix A). These images served as example situations for the related questions. A short explanation was given on what hypothetical normative interpretations of SA could look like, prior to asking the questions on the use of SA to assess safe driving. This explanation was derived from the information stated in the introduction and was presented together with images showing the same comparisons as in Fig. 1. Appendix B covers a more detailed description of the interview setup. At the end of each interview, a short debriefing was given and the next steps were discussed.

A recording of the meeting was made with consent of the expert, which served to summarize each interview into approximately 800 words. Draft summaries were sent to the corresponding experts for potential adjustment and approval. The lead researchers (RZ, RJJ) independently established a list of recurring themes based on the interview summaries. The lists were merged to ensure complete coverage. For each theme, relevant statements by experts (if available) were compared to determine whether there was consensus or disagreement.

3. Results

3.1. Cheryl Bolstad

3.1.1. Relations between situation awareness and safe driving

Having SA means being aware of the things around you and how they could change. Pure awareness of an element is not enough; a driver also needs to understand how it could change, project how rapidly that change could occur, etc. SA in driving does not only involve the surroundings of the vehicle, but also applies to the interior of the vehicle; e.g., perception of how the gauges are changing and if there are warnings, and comprehension of the state of the vehicle and whether it is functioning correctly.

SA is a prerequisite for safe driving. It is impossible to dynamically do anything safely without SA. This does not mean that perfect SA always results in being a safe driver. There could be elements of a situation that are outside the driver's control (e.g., a car might suddenly pull out in front of you, a tree might fall). Nonetheless, the literature shows a high positive correlation of SA with performance.

3.1.2. Using SA to assess safe driving

SA and safety are correlated, but not 100 %. When I talk about SA, I typically use a relative interpretation, e.g., when studying SA differences between age groups. We never saw participants with a perfect SA score. When two SA scores are far apart, it is pretty clear which is better, but if they are close together, it may be hard to tell.

Establishing a normative interpretation of SA to assess safe driving could be possible in theory but very challenging in practice. If a person runs 10 tests and performs above average on SA, probably this person drives safer than people who perform below average. But where do you put the line between 'poor' and 'good'? Knowing where to draw a line above which all SA scores are declared as 'safe' is susceptible to interpretation.

In order to determine where such a line should be drawn, a better understanding of SA is necessary, including a better understanding of which factors influence SA. You will need to figure out how to get drivers to focus on the correct elements of a situation, which will be situation-dependent. Next, age is likely to be one of those factors. Younger drivers tend to have lower SA but are able to react quicker than older drivers who have more experience and higher SA scores, especially for level 2 and 3 SA. Other known factors influencing SA measures are working memory and processing speed.

3.1.3. Selection of SA requirements

Not all elements of a situation are relevant to assess safety. Vehicle color, for example, could be relevant for a testimony after a crash, but not for safety assessment. In general, SA requirements relating to SA levels 2 (e.g., 'how fast am I going') and 3 (e.g., 'what happens when I press my brakes now versus 2 s later') are more informative about the quality of driver performance than level 1, which serves as a prerequisite for the higher levels. Missing a single SA requirement could still result in safe driving; its impact depends on the importance of the SA requirement missed, which varies across scenarios, and within scenarios over time. For example, a ball rolling on the street in an urban setting is unlikely to appear on a highway. Time of day may not matter much, except during rush hour when one may anticipate more cars or in darkness when certain features are harder to discern. Shifting of importance and attention may occur in rapid succession (e.g., from a lead vehicle to a traffic sign, driving speed gauge, a cyclist, and back to the lead vehicle) within a scenario. Certain elements may be common across scenarios, such as the location of other road users or the speed limit. It should be possible to determine a number of these common denominators. However, using only common denominators could lead to measures that are too superfluous to measure the intricacies of SA.

Using experts with a lot of driving experience could be a viable method of determining the importance of SA requirements. A caveat is that these experts are not always great at explaining why certain elements are important. Using existing databases to look at crash causes could also provide information about the importance of certain SA requirements. While the top-level goals might differ between two different GDTAs, the underlying SA requirements are typically consistent. These could serve as a basis for subsequent prioritization. Finding a consensus on what SA requirements are considered important would add to the usability of SA.

3.1.4. Measuring SA

There are multiple ways to measure SA (e.g., SAGAT, expert rating of participants, post-trial measurements, physiological measures). Regardless of the chosen measure, taking multiple measurements is essential, because SA differs from moment to moment. Repeated measurements across different scenarios give the best insight into performance. When queries are used, it is important to prevent cueing participants, e.g., by selecting a different set of queries at each measurement. It is advisable to separate the three levels of SA during analysis, but this may not always be possible if there are too few questions in certain levels.

3.2. Mica Endsley

3.2.1. Relations between situation awareness and safe driving

SA is fundamental to safe driving. Taking in information and processing is critical to staying safe in traffic. Good SA allows a driver to avoid obstacles, ensure good operational control, and achieve a wide variety of goals.

An increase in SA is likely to result in an increase in safe driving. The results of a *meta*-analysis across several domains showed a positive correlation (Pearson's $r = 0.46$) between SA (obtained through SAGAT) and performance measures (Endsley, 2021a). However, having poor or no SA is not guaranteed to have an unsafe outcome. In theory, a person could avoid issues crossing the street even when blindfolded, either because other drivers take evasive action or due to pure coincidence. And people can have accidents for other reasons, such as choosing to drive too fast. However, having better SA greatly reduces the likelihood of an unwanted safety failure. Therefore, the relation between SA and safe driving should be viewed as a probabilistic relation.

3.2.2. Using SA to assess safe driving

From a practical perspective, dividing or transforming a measure of SA into a reference scale will be difficult. Determining if some level of SA is safe or unsafe is dependent on the context of the situation. Drivers are constantly switching attention, and parts of a scene may be critical only for a small amount of time, depending on what is critical at that time. Drawing a line at some level of SA and calling everything above that line safe seems very difficult.

From a theoretical perspective, it is inappropriate to determine a criterion for safe SA. Maximum SA should be the goal if you are trying to improve safety. Drawing any line under 100 % leaves room for negative outcomes. There is currently no known method of determining what level of SA is enough for safe driving below 100 % because accidents can still occur. This is due to the probabilistic nature of the relation between SA and performance, as mentioned earlier.

3.2.3. Selection of SA requirements

A goal-directed task analysis can be used to lay out SA requirements for a variety of driving goals, which should then be mapped to performance measures. A GDTA showing the SA requirements for driving is provided in Endsley (2021b). Note that certain SA requirements relate to specific aspects of performance. For example, being aware of other road users around the vehicle influences safety performance but has no impact on navigational performance.

The importance of an SA requirement in the GDTA at any given time is dependent on what the current goal of the driver is. People often have to juggle between attention to different SA requirements to meet their goals. In the previous example, looking at the navigation screen could detract from the SA requirements for a safety goal due to a shift in attention that fills an SA requirement for a route choice goal. Yet people frequently need to make these trade-offs between goals (such as avoiding collisions, navigating to the destination, handling vehicle problems, etc.). What goals are most important and which SA requirements are relevant depends in part on the scenario. A certain awareness across the elements within a situation is required to determine what is the highest priority and what could be ignored.

In the case of a truck merging in front of a subject vehicle, noticing the truck should be considered a safety-relevant SA requirement. Depending on the relative speed of the truck, the impact of missing this SA requirement would have more or less impact on safety. If the truck is driving faster than the subject vehicle when it is merging in front of the car, there is a low crash risk, but if the truck is moving slower than the subject vehicle, the crash risk is much higher. The location, trajectory, and speed of the truck are all relevant to SA. So, the actual impact of low SA in many situations is only probabilistically related to performance. But since we want to reduce the risk of bad outcomes, maximizing SA is the goal.

Good drivers are using often incomplete information to anticipate future events and create safe margins. There are, however, some SA elements that can be considered to be always important, such as vehicle warning lights. Furthermore, potential decisions of the driver could influence what elements of a situation should be considered relevant. If someone swerves in front of you, a decision to evade a collision by swerving left requires awareness of vehicles in the left lane. Such information is not strictly necessary if the decision would have been to brake instead. However, people do not always know in advance when an immediate reaction will be needed. Therefore, to maximize safety, maintaining SA about all the vehicles around one's car is optimal. While the relative importance of SA on any given element can vary over time, I have found that people need to maintain at least some SA across their goals in order to determine what actually are the current priorities. For example, they need to at least know that no vehicle warnings are present but may only need to know the details if, in fact, there is a problem.

3.2.4. Measuring SA

Questions that probe SA should be crafted in such a way that they cover a range of requirements across the driver's goals. This provides insight into how attentional resources are traded back and forth between different elements. Asking about elements that are completely irrelevant to the driver's goals serves little purpose. However, questions should also not be too specific because that might lead to cues for participants, altering their SA.

Grouping SA into a single score limits the potential for diagnosis, as it hides trade-offs in attention and has been shown to be less sensitive than examining SA on individual elements (Endsley, 2021a). For example, adding a head-up display may yield more awareness of speed, but at the cost of lower awareness of something else. By adding the scores across SA requirements, you lose that insight. Even grouping into the 3 levels (e.g., perception, comprehension, projection) may not be very useful, because it hides what improved or declined within a level.

3.3. Petya Ventsislavova

3.3.1. Relations between situation awareness and safe driving

Situation awareness (SA) helps drivers to read the road and be aware of what is happening around to anticipate or plan for what will happen next on the driving scene. Good SA allows drivers to prioritize what poses an actual danger (as opposed to prioritizing irrelevant elements of the driving scene). SA is crucial for safe driving. Higher SA leads to better hazard perception and risk calibration, and, therefore, a lower likelihood of being involved in a collision. Nevertheless, while an essential factor, SA alone does not ensure collision-free driving, as drivers may still make incorrect decisions. While SA is a significant factor, other elements come into play, such as the environment and individual personality traits. SA does not necessarily involve decision-making but can contribute towards an adequate decision for safer driving.

An issue with the current body of literature is the number of different methods that are being used to operationalize and measure SA (e.g., different research groups design their own hazard perception methodology and often modify the traditional hazard perception methodology used). While there is no universally accepted framework, the three-level framework by Endsley (1995a) is the most popular definition of SA. The majority of methods used are typically developed to measure one of these levels (e.g., third level hazard prediction). The inconsistent methodologies hinder the ability to quantify and make direct comparisons across studies regarding the relationship between SA and safe driving. These discrepancies do not arise from incorrect interpretations but rather from variations in the aspects of SA measured in different studies. It would be highly beneficial to establish a unified framework for SA and safe driving to address these challenges.

3.3.2. Using SA to assess safe driving

The usefulness of SA to assess safe driving depends on what aspect of SA is being measured and how it is measured. Different performance indicators will impact the relationship between SA and safe driving. An adequate understanding and operationalization of the underlying components of SA is paramount to assess safe driving. To gain a comprehensive understanding of SA as a holistic concept, we must assess its individual components separately. While we cannot definitively predict that someone will never be involved in a crash, a high level of SA can significantly reduce the likelihood of collisions. It is worth noting that even when some drivers show good SA, they may still struggle with accurately estimating risk, particularly in situations involving specific decisions such as driving speed, overtaking, etc. Therefore, it is essential to evaluate all facets of SA.

SA is a broad process which includes a combination of various sub-processes (e.g., prioritizing relevant information, hazard perception and processing, prediction of hazards), which all need to be operationalized individually. Assessing each sub-process separately is the most effective way to understand the complex mechanism of SA. Thus, we will be able to address some of the mixed evidence in relation to the impact of some of the sub-processes of SA on safe driving.

Endsley further suggested that individual and environmental factors can influence the three main components of the SA theory. For example, driving experience, age, and cultural context can impact the way we perceive, understand, and predict dangerous situations. A standardized understanding of SA is feasible by building upon a strong theoretical and empirical knowledge base.

3.3.3. Selection of SA requirements

The model of SA distinguishes three main levels of global projection or, in other words, comprehension of other drivers' behavior on the road: perception, comprehension, and projection (Endsley 1995a). Not all elements in the environment will be prioritized when monitoring for hazards, as they will have different level of importance. Some elements should be prioritized before others to avoid hazardous situations. If drivers are not able to correctly identify which element poses a more immediate hazard, their SA needs to be developed further. Nevertheless, the mere act of observing an element does not inherently imply an understanding of its significance, such as whether it poses a hazard. This comprehension is shaped by context-specific expectations rooted in prior experience and knowledge. An experienced driver is likely to have developed a mental model or representation of different situations and elements that are either more or less likely to be present in a certain driving context. For instance, motorways frequently feature merging vehicles, while urban settings are more likely to include pedestrians. The establishment of broad hazard categories is feasible, as there are common hazards that can be present in different contexts. However, there are also hazards that are culturally specific to the context. Therefore, general categories should be approached with caution and must only function as general guidelines.

Anticipation of hazardous scenarios is essential for safe driving. Experienced drivers typically have good SA, which helps them to anticipate potential hazardous situations in advance and avoid these situations. In order to be able to anticipate hazardous scenarios, drivers should be able to successfully predict what will happen next on the driving scene (as per Endsley's level 3 – projection). Perceiving hazards in a timely manner would only be possible if drivers are able to anticipate them. This allows drivers to plan ahead and adjust their behavior accordingly.

3.3.4. Measuring SA

The third level of SA – prediction – has typically been measured via the ability to anticipate hazards. In order to be able to successfully predict what would happen next in the driving scene, drivers usually make use of precursors to the hazardous situation. For example, a ball on the street may indicate that children might suddenly run into the road. The ball would be a precursor to the hazard (children). Therefore, to grasp the environment and foresee forthcoming events, drivers rely on the context of the situation. In this context, short video clips tend to be more effective than static images. Videos offer a more comprehensive view of the situation, allowing viewers to better gauge factors like the speed of the driving car and that of others, which can be more challenging to discern in still images. To effectively capture this third level of SA, rather than directly asking drivers if they spotted a particular hazard, it would be more effective to ask them if they are able to predict what will happen next on the driving scene (independently of whether they consider the situation hazardous or not).

3.4. Joost de Winter

3.4.1. Relations between situation awareness and safe driving

According to theory, SA is a precursor to decision-making in driving and a vital determinant of safe driving performance. A higher level of SA will usually result in safer driving performance. However, SA is not the same as performance. For example, a driver might have a high level of SA, but the situation could still be unsafe, perhaps because it has become time-critical due to circumstances beyond the driver's control.

3.4.2. Using SA to assess safe driving

I see possibilities for determining the relationship between SA and safety, for example using a psychometric approach (see also Vlakveld, 2011). Just as we see in the Netherlands with the hazard perception test in the theoretical exam for the driver's license, it is possible to administer a large number of hazard perception items and correlate the outcomes of such a test with drivers' recorded accident involvement. The predictive validity of such a test is likely not very high, given the measurement errors that can occur, but a certain distinction between safe and less safe drivers should be achievable in this way.

With such an approach, it should also be possible to select 'good items', meaning traffic situations that differentiate relatively strongly between safe and unsafe drivers. I can imagine that relatively obvious situations, like a salient ball rolling onto the road, are not very discriminating. On the other hand, situations with subtle cues will probably be discriminative between safe and unsafe drivers. Equivalent to a shortened version of an IQ test or personality test, it might be possible to develop a brief SA screening tool.

3.4.3. Selection of SA requirements

When we want to determine SA requirements for a given traffic situation, the psychometric method described above may be useful; that is, compare accident-involved and non-accident-involved drivers, and see how their SA scores differ for this situation. Another possibility I see is the use of expert drivers versus beginner drivers to determine which cues the expert drivers report as being relevant. It should be noted that, based on Stevens' levels of measurement (Stevens, 1946), these approaches only provide information on an ordinal scale.

When aiming to develop SA norms on an interval or ratio scale, for example, for real-time applications like warning systems, other techniques come to mind. One possibility is to use a model of an autonomously driving car that runs in the background. We have previously described these kinds of ideas (De Winter et al., 2019) and are currently working on them in my group. If a driver behaves abnormally, for example, not looking in the mirrors at a time when the predictive model says they should, the SA score would be considered lower. However, a pitfall here is that, ultimately, it is driving performance that is being measured not SA. The circularity that arises is that SA is seen as an explanatory construct, while only the performance of the driver is observable (see also Flach, 1990).

3.4.4. Measuring SA

If, however, we want to measure SA according to theory, with its three levels (perception, comprehension, and anticipation), more introspective methods come to the forefront, such as the think-aloud method or the SAGAT method. Endsley has conducted pioneering work in this area and has published several influential papers (e.g., Endsley, 1995a). The SAGAT method can still be seen as *the* way to measure SA, a way of measuring SA that aligns well with the original meaning of SA.

However, there is little statistical evidence that the three levels of SA are distinct. Although it is conceivable that a driver perceives a relevant element in the environment (e.g., a pedestrian) but does not understand or anticipate what this element is going to do (the pedestrian plans to cross, so I should brake), it is practically difficult to distinguish those information-processing stages; in reality, cognition may not be so linear but highly intertwined instead. Also, there are some disadvantages to using SAGAT because it is mainly suited for virtual worlds (after all, you cannot easily 'freeze' the real world). SAGAT-like methods where participants have to position objects in the environment (Lu et al., 2020), or even answer certain queries while driving, have been described in the literature and can provide good alternatives for researchers.

3.5. Mark Young

3.5.1. Relations between situation awareness and safe driving

Situation awareness as a concept helps us to think about what is necessary for performance and safety. However, it is difficult to use the concept of SA to explain what is happening or has happened. When investigating a crash, it is easy to say that a driver has lost SA, but that is not really helpful. SA is useful as a label, but in a safety system, we need to drill deeper into what is actually meant with SA. SA involves awareness of objects in the near- or medium-term for both space and time. In advanced driving training, drivers are instructed to use this awareness to anticipate what may happen. A driver's knowledge of their own (limited) SA (i.e., *meta*-awareness) could result in increased safety margins by slowing down and increasing distances in response. Additionally, knowledge of other road users' awareness could also be taken into account.

The relation between SA and safe driving is positive: having more SA should result in safer driving. However, if nothing abnormal happens on a familiar route, driving without high SA may be possible without encountering problems. Therefore, lower SA does not necessarily result in unsafe driving.

3.5.2. Using SA to assess safe driving

Making a decision on how much SA is considered "safe" requires knowing what objective criteria for safety need to be met within the context of a measurement. What elements of a situation are given attention is also based on driving experience and recent events. A driver's limited attentional resources are likely diverted to where the driver is going to go. Where that is depends on the context. Thus, boundaries can be determined for a situation within which information can be considered relevant. The proportion of information known within these boundaries could be used to determine what qualifies as "excellent". Information outside these boundaries should be considered less relevant.

3.5.3. Selection of SA requirements

Not all elements of a situation are of equal importance for assessment of safe driving. For example, perceiving the presence of a vehicle merging towards the same lane has more impact on safe driving than noticing the color of a vehicle on another carriageway. Furthermore, the importance of elements in a situation is not static. If a driver decides to take an evasive action towards the left lane, it becomes important to know about other road users on that lane.

A framework is required to structure the importance of SA requirements in a rational and objective way. Time criticality or immediate risk to safety could be useful constructs. Elements that are closer in time or space, or present a bigger risk, should be considered as more important. Akin to Gibson's field of safe travel (Gibson & Crooks, 1938), the dynamic nature of a situation could be addressed by involving the rate of change of elements. Depending on the context, the quicker an element could change and the more often it does, the higher its importance could be. This usually puts other road users above the physical features of a situation, as road users are the more variable elements. Research on visual patterns between novice and expert drivers shows that irrelevant information makes its way more into the minds of novice drivers. Therefore, expert drivers are likely the best population to ask in order to determine what elements of a situation are important.

There might be situations where noticing certain elements could be considered as a distraction. Noticing a vehicle on another carriageway is, at best, a sign of spare attentional capacity and, at worst, a waste of limited resources that should be spent elsewhere. The difficulty with determining which of those is the case is a consequence of the inability to know a driver's attentional capacity.

When incorporating driving automation features such as adaptive cruise control, the situation becomes more complex. It becomes necessary for the driver to also be aware of what part of vehicle control is done by the vehicle and what it might do next. This necessitates some form of shared SA between the driver and the system, almost like a conversation with the technology (see, e.g., Young & Stanton, 2023).

3.5.4. Measuring SA

A constraint of the SAGAT method is the need to define elements of interest beforehand. If one knows up front that there are 10 important elements in a scenario, it is perfectly possible to check how many of the corresponding questions are answered correctly. The first challenge lies in identifying which 10 elements are the most relevant. The second challenge lies in interpretation. If you have correctly answered 7 out of 10 questions, does it mean that you are 70 % aware? The meaning of a certain proportion of awareness becomes more ambiguous if some elements are more relevant than others, and so should be weighted as such.

The three levels of SA as suggested originally have some use. In practice, the foundation is always perception. Considering the practical effect of SA, a fourth level that relates to the effective decision on the basis of the other three levels could be useful.

4. Discussion

The interviews with experts on SA yielded several insights into the feasibility of a normative interpretation of SA in relation to safe driving. These will be discussed first at a theoretical level (e.g., whether such a normative interpretation is valid with the construct of

SA, Section 4.1), followed by a comparison of views on methodology (Section 4.2) and operationalization of SA measures (Section 4.3). Taken together, these insights serve as steps towards a normative interpretation of SA (Section 4.4).

4.1. Comparison of experts' views on a theoretical interpretation of SA

All the experts appeared to have a similar interpretation of SA from the perspective of the driver, involving awareness of the location of surrounding objects, how their location might change over time, and how to anticipate such changes. The experts indicated that while perceiving an object is the first step to good SA, understanding and anticipation are better measures of good SA. Usage of SA measures does differ between the experts, with some indicating use as a diagnostic tool, some researching differences between groups of drivers, and others as a way to think about what aspects of a situation are important. The consensus was that SA is positively correlated with safe driving and that having high SA is important for safe driving. However, all experts indicated that having perfect SA is no guarantee for safe driving, nor does having imperfect SA necessarily result in unsafe driving. The relation between SA and safe driving could, and maybe should, thus be interpreted as a probabilistic one (e.g., Endsley, 2021a). All experts mentioned factors that may influence the relation between SA and safe driving, ranging from driver-related factors (e.g., age, working memory, processing speed, *meta*-awareness) to driving context-related factors (e.g., cultural context, discriminability of the driving context, other road users' SA). Several experts mentioned the need to further investigate the exact impact of these factors and whether they are intervening or mediating the relation between SA and safe driving. The necessity to develop a framework was also mentioned by multiple experts as a solid understanding of SA and its constituent parts would be required to counteract diffusion in interpretation of SA and discrepancies in methodology. The notion that SA as a construct is interpreted similarly across experts provides a basis for framework development. A challenge, however, may be to establish a framework that addresses all specific uses of SA expressed by the experts.

Most experts indicate that it is possible to develop a normative approach to SA in a safe driving context. A major challenge lies in the determination of when some amount of SA that can be considered "safe" or "unsafe". A number of experts indicated that deciding on when something is safe or unsafe requires not only good knowledge of what elements are considered of importance, but also a method to do so rationally and objectively and some decision on what the thresholds should look like.

Some experts expressed ethical concerns with a normative approach, relating to the difficulty of determining what should be considered as enough SA. The probabilistic relation between SA and safe driving was mentioned by several experts as a reason for these difficulties. Some experts indicated that the goal should be to maximize SA as that would likely result in the safest driving, even if some indicated that perfect SA never occurred in their studies and these participants still managed to drive safely.

4.2. Comparison of experts' views on methodology for SA requirements selection

The experts agreed that there are differences in the importance of situational elements with regard to safe driving. Examples of important elements included location and speed of other road users close to the ego vehicle, (latent) hazards such as a ball rolling onto the road, and potential warnings from within the vehicle. Consensus among the experts also indicated that there are elements of a situation that could be missed by drivers without necessarily resulting in unsafe driving. Examples of elements that could be considered less important included other road users on separated carriageways and the color of other road users' vehicles. Experts indicated that the potential decisions of the driver in the here and now could also influence the importance of situational elements in the next instance. Drivers who plan to take evasive action to the left lane should be aware of vehicles in that lane, whereas drivers who plan to make an emergency stop instead have less reason to be aware of those vehicles but more reason to be aware of vehicles behind. Several experts cautioned that while it might be possible to determine some general situational elements or SA requirements, situation-specific requirements should still be used. General requirements are likely to only allow superficial probing of SA, and overreliance on these general requirements could lead to measures that are too superfluous to really understand the intricacies of a situation. The impact of culture or context on what elements are important should also be taken into account. Four experts mentioned the need for a framework and/or method to reach consensus on what elements in a situation are considered important.

Prioritization of elements or SA requirements was mentioned by several experts, suggesting a hierarchy in requirements, with all experts agreeing that the context of a traffic situation plays an important role in prioritizing. The importance of individual SA requirements changes based on this context and multiple experts indicated that the importance of an element could differ from one second to the next. Suggestions of methods to establish a hierarchy of importance differed somewhat between the experts. Multiple experts suggested that experienced or trained drivers could be asked to gather a list of elements that are important. Research on the gaze patterns of novice and experienced drivers shows that experienced drivers are better at selecting important elements (Crundall et al., 2012; Lehtonen et al., 2014; Underwood et al., 2009; Vlakveld, 2011), with the caveat that experienced drivers are not always consciously aware of why certain elements are important. Usage of existing data on common causes of crashes could also provide information about what elements should be considered important. A measure of criticality or risk of an element was suggested as another potential starting point to determine importance, with remaining time or distance to an object as an example of criticality. The suggestion was made to utilize the rate of change of elements as a measure of importance, with quicker and more frequent change resulting in higher importance. This would usually result in other road users being of higher importance than more static objects such as infrastructure elements. The suggested methods are not mutually exclusive, and in fact, the validity of a hierarchy of SA requirements may be increased if alternative methods lead to the same ordering (i.e., triangulation).

In addition to methods for determining the importance of elements, there were suggestions for methods to broadly select elements that could be considered unimportant. This could be done by determining boundaries for a situation within which the information could be considered relevant, and information outside of these boundaries irrelevant. An example of these boundaries could be the

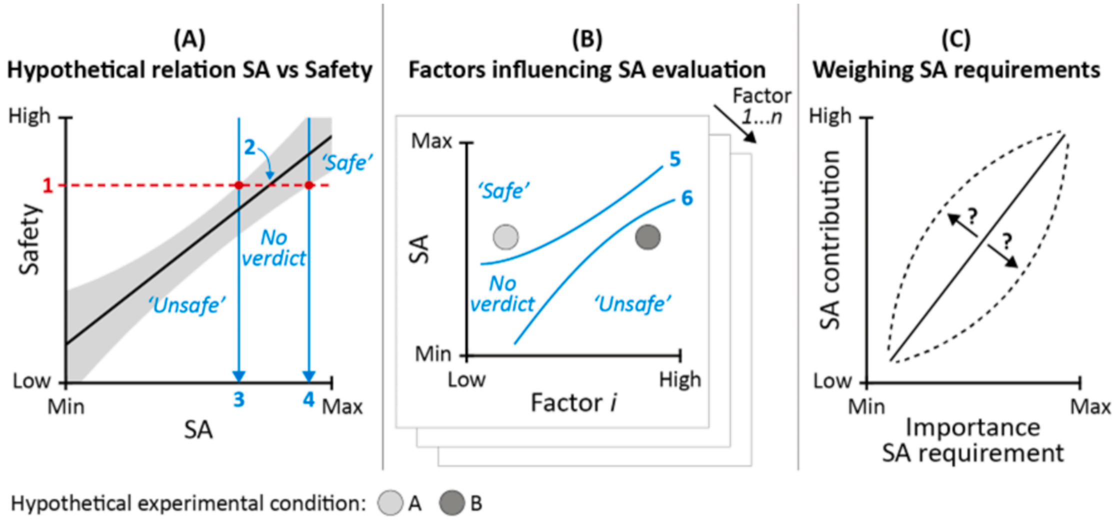

Fig. 2. Considerations for the establishment of a normative interpretation of SA. A hypothetical safety criterion (line 1) is translated into thresholds (lines 3 and 4) for (un)safe levels of a compound or constituent measure of SA based on an established relation between SA and safety (panel A). The evaluation of SA in relation to safety may vary across one of several factors (lines 5 and 6 in panel B). A method to weight individual SA requirements is needed to establish a measure of SA; the relation between importance of SA requirements (e.g., for safety) and SA contribution is not yet known (panel C).

physical separation from opposing traffic on a highway, as vehicles on one side have no realistic way of impacting the safety of the vehicles on the other side.

Comparison of experts' views on operationalization of SA measures.

Most experts agreed that the three levels of SA as proposed by Endsley (1995b) should be differentiated when investigating where potential issues with SA might arise, although ideas have been expressed in which scores are compared without differentiating between the three levels. Most experts also indicated that when scoring a participant's SA the higher levels (e.g., comprehension, projection) are likely to have a bigger impact on safe driving than the lower level of perception. Level 1 should be seen as the precursor to the other levels, but is not regarded as providing as much information about the quality of driving performance, SA, or hazard perception. However, one of the experts pointed out that there is no evidence that the three levels of SA are statistically identifiable, and from a pragmatic viewpoint, one expert noted that there are often too few queries at each level of SA to allow for separate analysis. The three theoretical levels of SA are helpful in determining where attention is spent, and thinking about how to delineate the relevant SA requirements, but the lines between each level are not clear-cut.

A single combined measure of SA could be useful in order to score overall SA; however, this has been found to be less sensitive than examining each query separately (Endsley, 2021a). Other experts agreed that SA needs to be operationalized individually, assessing each SA level separately to understand the complexity of its underlying mechanism, though it should be noted that each level relates to and informs the subsequent one. Another expert indicated that while the different levels have some use, the real impact of having good SA is maybe best observed by what action follows from the awareness, suggesting the potential for an additional "fourth" level that incorporates the effective decision that follows. Endsley (1995b), however, stresses that decisions are not a part of SA, but rather a subsequent stage.

The importance of good probe design was emphasized by all experts. When utilizing questions, these should be crafted in such a way that they do not cue participants on what situational elements to attend to in subsequent queries. It was suggested to ask participants to state their prediction for what happens next rather than asking them if they have spotted an element of the situation, to examine level 3 SA (hazard awareness). Some experts stated that when the goal of measuring SA is to assess safe driving, the questions should be subtle enough to differentiate between safe and unsafe drivers. One expert suggested that questions should cover a wide range of requirements across the different driver goals, in order to provide insight into the trade-offs made in attentional resources. While some of the experts stated that using probes into irrelevant elements or SA requirements serves no purpose for determining SA itself, other experts suggested that these probes could still provide information about where attentional resources are being spent. Use of these questions also depends on what the goal of the measurement is: to determine the state of a driver's SA or to determine where attention is allocated.

4.3. Recommendations

This section synthesizes the expert views into a series of considerations that could lead future research towards a criterion-based assessment of SA for driving safety. Multiple experts characterized the nature of the relation between SA and safe driving as probabilistic: a high level of SA *likely* results in safer driving, and a low level of SA *likely* results in unsafe driving. For the sake of simplicity, the black line in panel A in Fig. 2 describes the relation between these constructs as a linear model, and thus SA is here represented as a single (compound) measure, addressing the question 'Is it safe?'. A representation in which SA is split into its constituent components (e.g., perception, understanding, projection) likely affords more diagnostic power (e.g., '*why* is it (un)safe?'). As another approach, elements may need to be grouped relative to their relevance to individual performance outcomes (e.g., SA needed for avoiding collisions with vehicles, SA needed for avoiding collisions with pedestrians, SA needed for getting to a destination), in order to be able to make meaningful predictions.

One expert noted that an objective criterion for safety is needed if a criterion for SA in relation to safe driving is to be established. Where to draw the red line (line #1 in Fig. 2 panel A) on a safety performance indicator (e.g., crash involvement, time-to-collision) is, in the end, a societal and/or political decision (e.g., European Commission, 2023) and beyond the scope of the present study. However, given that such a decision is made, and given that the relation between a compound or constituent measure of SA and the chosen safety performance indicator is known, it is, in theory, possible to derive a threshold for a safe level of SA (e.g., point #2 at which the black and red lines in panel A of Fig. 2 intersect).

The probabilistic nature of the above relationship implies that a high or low level of SA is no guarantee for safe or unsafe driving, respectively. This notion can be interpreted as an uncertainty margin or a confidence interval, depicted in gray in panel A of Fig. 2. A 'guarantee' for safe or unsafe driving based on a level of SA thus comes down to the degree of uncertainty we are willing to accept in society, analogous to the 95 % confidence interval that is commonly accepted for statistical tests in the scientific arena. Accordingly, the error margin yields a threshold below which one may speak of an 'unsafe' level of SA (line #3 in Fig. 2) and a threshold above which one may speak of a 'safe' level of SA (line #4 in Fig. 2). Between these thresholds lies a range of SA values based on which no verdict on safety can be given (although a conservative interpretation could be that any level of SA that is not guaranteed to be 'safe' should be treated as 'unsafe'). Note that focusing solely on the x-axis yields panel C in Fig. 1, with an additional 'no verdict' zone. Currently, this zone (error margin) is likely too large to state the effects of an SA level even well below 100 % on safety with an acceptable degree of certainty. As knowledge about factors influencing SA increases and subsequent decision making can be corrected for, the error margin can be expected to decrease. In turn, such a smaller margin of error may enable the determination of SA levels that meet a criterion for 'safe' or 'unsafe' driving.

4.4. Factors influencing SA evaluation

Measures of safe driving are frequently expressed as safety performance indicators (SPIs) (e.g., minimum time-to-collision, standard deviation of lateral position). Endsley (1995c) qualitatively modeled SA as a precursor to decision-making, which in turn served as one of the inputs for subsequent performance of actions (i.e., SPIs). Several factors influencing SA are described (e.g., information processing, ability, training, workload, automation). The experts in the present study noted that a better understanding of the (quantitative) contribution of these factors to SA is required to better understand the relation between SA and a measure of safe driving.

The interpretation of SA in relation to safe driving may differ when zooming in on specific factors influencing this relation. Lines 5 and 6 of panel B in Fig. 1 illustrate how variation on a factor (e.g., ability to react, reaction time) may change the evaluation of a certain level of SA in terms of safety. When two drivers score the same imperfect level of SA in the same situation (e.g., vehicle automation issuing a request to intervene in a critical situation), the driver with the physical ability to react faster has a higher likelihood of "safe" driving than the driver with a slower reaction. This hypothetical example illustrates how we may learn about the contribution of an individual factor on SA while keeping all other factors fixed. The thresholds for 'safe' and 'unsafe' levels of SA (e.g., the shape and position of lines 5 and 6 in Panel B) may vary across factors. Applying this principle to all factors {i…n} influencing SA may yield a better understanding of the quantitative impact of individual factors and a more detailed model of SA.

4.5. Weighting SA requirements

A selection of SA requirements is required to establish a measure of SA. Most often, this selection is a subset of all driving requirements, stemming from a task analysis such as a GDTA or an HTA. These individual SA requirements can be seen as constituent parts of an SA measure. It is well accepted that only some elements of a situation can be considered as part(s) of SA requirements while the remaining elements are considered irrelevant for the measured goal. When assessing safety, for example, SA elements of importance are often location and speed of other road users or potential hazards, whereas elements such as the contents of billboards are considered irrelevant.

Hauss and Eyferth (2003) described a method to distinguish between relevant and irrelevant SA requirements, and take this distinction into account when calculating final SA scores. However, not all relevant requirements within a situation have the same direct impact on safe driving, and such impact may vary abruptly over time. Consider a vehicle directly in front performing an emergency stop and a vehicle indicating a desire to merge. While both are relevant to driving safely, the vehicle performing an emergency stop has a much bigger impact in the short term.

The higher the safety relevance or importance of an SA requirement, the more impact it should have on the SA score (see also Fig. 2 panel C). Care should be taken to prevent 'hindsight bias', where a posteriori knowledge of how a situation has developed informs what elements are considered important. For example, drivers need to be aware of all the vehicles around them, and not just the vehicle that happened to swerve into their lane in a particular scenario. The ranking of requirements should only take into account information that would be available to the driver at the moment of measurement. This also applies to requirements whose relevance has a tendency to change rapidly over time.

Fig. A1. First scenario for preparatory survey version A. Caption shown to participants: "Driving in the inner city at 30 km/h". The hazard in this scenario concerns a child located between two parked cars, who may attempt to retrieve a football on the street. . Reprinted with permission.
Source: Vlakveld (2011)

Fig. A2. Second scenario for preparatory survey version A. Caption shown to participants: "Driving on the highway at 100 km/h". The hazard in this scenario concerns a speed differential with a truck intending to merge. In the Netherlands the maximum allowed speed for trucks is 80 km/h. . Reprinted with permission.
Source: Vlakveld (2011)

One method of ranking SA requirements suggested during the interviews is the use of expert drivers. Asking expert drivers to rank the selected requirements based on importance to safety would result in the relative importance of requirements. However, experienced drivers are not always capable of explaining why certain elements should be considered more important than others and show

Fig. A3. First scenario for preparatory survey version B. Caption shown to participants: "Driving in the inner city at 50 km/h". The hazard in this scenario concerns a bicyclist about to overtake a parked car. There is no room for the subject vehicle to overtake the bicyclist, due to oncoming traffic in the opposite direction. . Reprinted with permission.
Source: Vlakveld (2011)

Fig. A4. Second scenario for preparatory survey version B. Caption shown to participants: "Driving with adaptive cruise control active on the highway at 100 km/h". The hazard in this scenario concerns a speed differential with a truck intending to merge, combined with the notion that adaptive cruise control is active. In the Netherlands the maximum allowed speed for trucks is 80 km/h. . Reprinted with permission.
Source: Vlakveld (2011)

bias in their selection (Ventsislavova et al., 2022). Use of eye-tracking data to create a baseline or norm for gaze patterns and ranking requirements based on gaze frequency could provide a similar list of relative importance, potentially probing non-conscious elements of behavior.

Developing a method to more objectively determine the importance of SA requirements could allow for standardized weighting of different requirements. Based on the responses from the experts, the importance of an SA requirement in a particular instance depends on several factors.

A first factor concerns the SA level. Experts noted that the higher levels of SA (levels 2 and 3, comprehension and prediction) have more impact on safety than the lower levels of SA (level 1, perception). Therefore, it could be argued that measures (e.g., SAGAT queries) probing the higher levels are more important and may have a bigger impact on the total SA score compared to those measures probing only the lower levels of SA. An argument could be made that there are cases where a prediction could be made based on previous experience and knowledge (described as mental models in Endsley, 1995b). It is also possible that drivers' implicit awareness (Bellet et al., 2009) yields SA at level 2 and 3 without explicit awareness of the lower level SA. In a study conducted by Lo et al. (2016) among train operators, it was found that level 1 SA scores were remarkably low compared to level 2 and 3 SA scores. This suggests that the train operators may have observed relevant elements in the environment and implicitly processed them but were unable to report them explicitly. Their ability to understand the overall situation and extrapolate it into the future suggests that level 3 SA may be achieved without requiring explicit components of level 1 SA. Alternatively, one could argue that SA level 1 is a prerequisite for higher levels of SA (i.e., one cannot comprehend that which one cannot perceive), and therefore, SA level 1 may be viewed as most important.

Second, a measure of criticality or hazardousness may be included. Elements that pose a greater immediate crash risk should be of higher importance than elements that pose less or no risk. This could be based on trajectory-related measures such as time-to-collision or post-encroachment time, or more generally by taking into account the rate and frequency of change of each element.

Third, driving can also be seen as a dynamic regulation loop focused around SA of the driver (in line with Neisser's perceptual cycle (Neisser, 1976; Plant & Stanton, 2015)). A comparison of the perceptual cycle with the SA concept is also given in Adams, Tenney & Pew (1995). The dynamic loop of the driving tasks means that the importance of elements depends in part on decisions made before as well as changes in the evolving environment. As noted by a number of the experts, the importance of awareness of vehicles in the left lane increases when deciding to swerve left to evade a collision for example.

4.5.1. Future research

Further development of the concepts and recommendations presented so far is required before their feasibility and validity can be tested. Experimental research should consider advancing the development of a method for selecting and weighting SA requirements. A first step could concern comparing and contrasting different methods of determining the importance of SA requirements, by selecting identical situations and applying the methods mentioned above (e.g., expert driver interviews, gaze patterns, criticality of an element) to create rankings. Ideally, asking experts to rank requirements results in similarities in ranking for a particular situation as a ranking based on areas of interest identified through eye tracking techniques, or a ranking established with the use of a theoretical approach based on trajectory measures or frequency and rate of change (see also Bellet et al., 2009). By comparing these differently acquired rankings, insight could be gained on whether a requirement is important or not and help to establish what strengths and weaknesses are for each method. Of key importance in any efforts in this direction is the recognition that the relative importance of different SA requirements will be highly dependent on the decisions and goals being considered. For example, a different set of rankings would be expected for avoiding collisions with other vehicles for example, than for navigating to a particular location, or for avoiding a traffic ticket. Additionally, a focus on individual scores for SA requirements that are deemed important or of high priority provides insight without being occluded by attention shifts that are taking place on lower priority SA requirements.

Another approach could include a method to determine the weight of a specific SA requirement based on the risk of negative consequences when the SA requirement is not met. Being aware of 90 % of information but overlooking a key element for making a safe decision (such as a trajectory conflict) could almost systematically lead to increased crash risk. Conversely, a driver who overlooks 90 % of information but focuses almost exclusively on this key element might prevent a crash. The lack of awareness of key elements may yield a more direct indicator of dangerous driving than the percentage of accuracy of SA as a predictor of safe behavior. However, the quickly changing environment during driving may result in a situation where an element that was not important before may suddenly become critical. A focus on only those elements that are of importance in the moment might leave a driver unable to quickly refocus attention to new elements of importance when required. Care should also be taken to avoid hindsight bias when selecting key elements.

To develop a criterion-based approach of SA and reduce the uncertainty margin between SA and safe driving, various aspects of decision making could be included, such as mental models (Bellet et al., 2009), risk assessment (Bellet et al., 2009), driving schemata (Charlton and Leov, 2021) and differences in types of drivers (Salmon et al., 2014). For example, risk assessment (Bellet et al., 2009) concerns how drivers assess (potentially critical) situations based on their SA and may account for driver-specific variation relating to, e.g. risk-taking.

It should also be noted that other road users play an important role in determining final safety outcomes. When other road users are more aware of the ego vehicle than the other way around (e.g. Salmon et al., 2014), these other road users may be able to compensate for a potential lack of SA from the driver of the ego vehicle in some instances. Technological interventions could also compensate for a lack of SA, influencing the amount of SA that is required (for example, lane change or blind spot warnings and backing warnings that might help to compensate for SA losses). Experimental research that explores the impacts of these factors combined with a method of selecting and weighting SA requirements could result in a better understanding of the impact of SA on safe driving and allow for more focused approaches to tackling road safety issues.

4.6. Limitations

A limitation of the current paper is the number of interviewees. While they were chosen based on their expertise in situation awareness, they only represent a small subset of all SA experts. Another limitation lies in the semi-structured nature of the interviews. Not all follow-up questions were the same for all interviewees due to differences in answers. It can, therefore, not be stated with certainty if experts did not mention something because they did not feel a certain way or merely did not think of it at the time. The saying that "absence of evidence does not equal evidence of absence" applies here. Additionally, the paper does not present new empirical data. The current paper only presents a foundation for further work that should include experimental research. Some suggestions for further research have been presented.

Furthermore, distinct theories and perspectives on SA exist, and this paper adopts the individual driver's perspective when discussing SA. While this limits the perspectives on SA, it allows for a clearer discussion on requirement selection methods and weighting. Projection of the considerations presented here on other theories of SA is recommended to explore if and how, for instance, a level of shared SA could be assessed.

5. Conclusion

In this study, five experts on SA were interviewed to gain insights into the feasibility of, and potential considerations towards, a normative interpretation of SA in relation to safe driving. The experts agreed that situation awareness is associated positively with safe driving but that factors underlying this relation are currently not fully understood. The probabilistic relationship between SA and performance results in an inability to confidently state what amount of SA can be considered safe for driving. While most experts agree that it is, in theory, possible to reach a normative relation between SA and safe driving, there are a number of practical challenges that need to be considered first.

From an academic or scientific perspective, the present study highlights several considerations towards the development of a more complete concept of SA for safe driving. The considerations relate to the selection of safety-relevant SA requirements and ranking or weighting these requirements based on their importance to driving safety. A first approach to the selection and weighting of SA requirements is given in this paper. Deciding on a standardized method of identifying and weighting requirements should be considered as an important goal for future research. Additionally, a criterion based approach to SA is recommended. From a societal or ethical perspective, the present study highlights that there is currently no way to determine what amount of SA below 100 % can be considered 'safe' on relevant elements, or what 'safe' should entail. The discussion about what should be considered an adequate level of safety in driving should be held at a societal level. Until it is feasible to determine what amount of SA is enough, the goal should be to maximize SA, while acknowledging that achieving 100 % SA may not be possible and some trade-offs may need to be made.

The use of relative SA measures to declare one group or interface as safer than another should be done with the utmost care. The current known relation between SA and safety does not allow for the judgment of safety based on relative performance, unless safety-specific measurements are made in addition to SA measurements.

CRediT authorship contribution statement

Rins de Zwart: Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Reinier J. Jansen:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Cheryl Bolstad:** Writing – review & editing. **Mica R. Endsley:** Writing – review & editing. **Petya Ventsislavova:** Writing – review & editing. **Joost de Winter:** Writing – review & editing. **Mark S. Young:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors declare that they have no conflict of interest. This study was funded by the Dutch Ministry of Infrastructure and Water Management. The funder played no role in the study design, collection, analysis, interpretation of data, writing of the report, or in the decision to submit the paper for publication. The funder accepts no responsibility for the contents.

Appendix A:. Preparatory survey

As preparation for the interview, a short survey was sent to the interviewees. The survey consisted of an assignment and two scenarios. The scenarios are shown in Figs. A1 to A4, where the scenarios depicted in Fig. A1 and Fig. A2 were used in survey version A, and those depicted in Fig. A3 and Fig. A4 were used in survey version B. The following assignment was given with the scenarios:

“Please have a look at the images on the next two pages. Each image shows a scenario from the driver’s point of view. *For each scenario think of 3 to 5 potential elements that you would consider requirements for situation awareness. The requirements should be formulated as elements of information that the driver should know about in order to safely take part in traffic.”*

Appendix B:. Questions for interview with experts on situation awareness

The use of situation awareness as a measure of driver engagement or safety is an important part of road safety research. This interview is dedicated to exploring how some of the current limitations of situation awareness in relation to assessing driving safety may be mitigated. Special attention will be given to discussing situation awareness requirements and how to link these to driving safety assessment. The interview starts by probing interpretation of the construct ‘SA’ and its relation to safe driving, followed by if and how measures of ‘SA’ could be operationalized towards assessing driving safety.

0B.1. General questions on the relation between ‘SA’ and safe driving.

1. How would you describe the relation between ‘SA’ and safe driving?
2. Would you describe this relation as uni- or bi-directional?
3. Do you think it is possible to further quantify this relation, and if so, how?
4. Which factors influence the relation?
5. What is keeping us from quantifying the relationship between ‘SA’ and safe driving?

0B.2. Questions on the use of SA to assess safe driving.

1. Do you think it is possible to establish a normative interpretation of SA in relation to safe driving? Why?
2. What would be required to establish a normative interpretation?

0B.3. Questions on selection of SA requirements.

1. Let’s say one of the SA requirements [listed by the expert in the preparatory survey] is not met (e.g., by incorrectly answering the corresponding question). How would you evaluate the driver’s SA in relation to safety?
2. What if not this particular SA requirement is missed, but another one? Would this change your answer?
3. Would you say that even when not meeting a certain SA requirement, SA could still be sufficient for safe driving?
4. If we used a different scenario (such as the one shown earlier), and the driver reaches the same SA level, then would the interpretation in relation to safe driving be the same as well?
5. Looking back at this thought exercise, how would you generally determine what SA requirements are needed to assess driving safety?
6. How would you decide which SA requirements should be met for safe driving? Does this selection change based on research question, research method, traffic scenario?
7. Do you think there are some SA requirements that are always relevant? Which ones?

0B.4. Questions on measuring SA.

1. What do you think is the best approach to translate an SA requirement into a query?
2. We have found differences across studies in terms of scoring answers. Do you think that the 3 levels of SA (perception, understanding, projection) should be grouped to calculate a single combined score? Or should a separate score be calculated for each SA level? Why?

Data availability

Data will be made available on request.

References

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. *Human Factors, 37*(1), 85–104.

Bakdash, J. Z., Marusich, L. R., Kenworthy, J. B., Twedt, E., & Zaroukian, E. G. (2020). Statistical Significance Filtering Overestimates Effects and Impedes Falsification: A Critique of Endsley (2019). *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.609647

Bellet, T., Bailly-Asuni, B., Mayenobe, P., & Banet, A. (2009). A theoretical and methodological framework for studying and modelling drivers’ mental representations. *Safety Science, 47*(9). https://doi.org/10.1016/j.ssci.2009.03.014

Blömacher, K., Nöcker, G., & Huff, M. (2018). The role of system description for conditionally automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour, 54*, 159–170. https://doi.org/10.1016/j.trf.2018.01.010

Bolstad, C. A. (2001). *October). Situation awareness: does it change with age?. In Proceedings of the human factors and ergonomics society annual meeting* (Vol. 45(4,, 272–276.

Certification: Pilots, Flight Instructors, and Ground Instructors, 14 C.F.R. § 61 (2023). https://www.ecfr.gov/on/2023-05-22/title-14/part-61.

Chandrasekaran, L., Crookes, A., & Lansdown, T. C. (2019). Driver situation awareness – Investigating the effect of passenger experience. *Transportation Research Part F: Traffic Psychology and Behaviour, 61*, 152–162. https://doi.org/10.1016/j.trf.2017.12.007

Charlton, S. G., & Leov, J. (2021). Driving without memory: The strength of schema-consistent false memories. *Transportation Research Part F: Traffic Psychology and Behaviour, 83*. https://doi.org/10.1016/j.trf.2021.09.018

Clark, H., McLaughlin, A. C., & Feng, J. (2017). Situational Awareness and Time to Takeover: Exploring an Alternative Method to Measure Engagement with High-Level Automation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61*(1), 1452–1456. https://doi.org/10.1177/1541931213601848

Dattel, A. R., Vogt, J. E., Fratzola, J. K., Dever, D. P., Stefonetti, M., Sheehan, C. C., & Cavanagh, J. A. (2011). The Gorilla's Role in Relevant and Irrelevant Stimuli in Situation Awareness and Driving Hazard Detection. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 55*(1), 924–928. https://doi.org/10.1177/1071181311551192

De Winter, J. C. F., Eisma, Y. B., Cabrall, C. D. D., Hancock, P. A., & Stanton, N. A. (2019). Situation awareness based on eye movements in relation to the task environment. *Cognition, Technology & Work, 21*, 99–111.

Crundall, D., Chapman, P., Trawley, S., Collins, L., van Loon, E., Andrews, B., & Underwood, G. (2012). Some hazards are more attractive than others: Drivers of varying experience respond differently to different types of hazard. *Accident Analysis & Prevention, 45*, 600–609. https://doi.org/10.1016/j.aap.2011.09.049

Endsley, M. R. (1988). In *Design and evaluation for situation awareness enhancement* (pp. 97–101). Santa Monica, CA: Human Factors Society.

Endsley, M. R. (1993). A Survey of Situation Awareness Requirements in Air-to-Air Combat Fighters. *The International Journal of Aviation Psychology, 3*(2), 157–168. https://doi.org/10.1207/s15327108ijap0302_5

Endsley, M. R. (1995a). Measurement of Situation Awareness in Dynamic Systems. *Human Factors, 37*(1), 65–84. https://doi.org/10.1518/001872095779049499

Endsley, M. R. (1995b). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors, 37*(1), 32–64. https://doi.org/10.1518/001872095779049543

Endsley, M. R. (1995c). A taxonomy of situation awareness errors. In R. Fuller, N. Johnston, & N. McDonald (Eds.), *Human Factors in Aviation Operations* (pp. 287–292). Aldershot, England: Avebury Aviation, Ashgate Publishing Ltd.

Endsley, M. R. (2000a). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley, & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 3–32). Mahwah, NJ: Lawrence Erlbaum Associates.

Endsley, M. R. (2000b). Direct measurement of situation awareness: Validity and use of SAGAT. In *Situation awareness analysis and measurement* (pp. 147–173). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Endsley, M. R. (2018). Expertise and situation awareness. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *Cambridge Handbook of Expertise and Expert Performance* (2nd ed., pp. 714–744). Cambridge, UK: Cambridge University Press.

Endsley, M. R. (2020a). The divergence of objective and subjective situation awareness: A meta-analysis. *Journal of Cognitive Engineering and Decision Making, 14*(1), 34–53.

Endsley, M. R. (2020b). Human-automation interaction and the challenge of maintaining situation awareness in future autonomous vehicles. In M. Mouloua, & P. Hancock (Eds.), *Automation and Human Performance: Theory and Applications* ((2nd Edition ed.,, pp. 151–168). CRC Press.

Endsley, M. R. (2021a). A Systematic Review and Meta-Analysis of Direct Objective Measures of Situation Awareness: A Comparison of SAGAT and SPAM. *Human Factors, 63*(1), 124–150. https://doi.org/10.1177/0018720819875376

Endsley, M. R. (2021b). *Situation awareness measurement: How to measure situation awareness in individuals and teams*. Washington, DC: Human Factors and Ergonomics Society.

Endsley, M. R., & Jones, W. M. (2001). A model of inter- and intrateam situation awareness: Implications for design, training and measurement. In M. McNeese, E. Salas, & M. Endsley (Eds.), *New trends in cooperative activities: Understanding system dynamics in complex environments* (pp. 46–67). Santa Monica, CA: Human Factors and Ergonomics Society.

Commission, E. (2023). Road safety in the EU: Fatalities below pre-pandemic levels but progress remains too slow. *European Commission*. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_953.

Evans, T., & Macdonald, W. (2002). Novice driver situation awareness and hazard perception: An exploratory study. *Proceedings of the Australasian Road Safety Research, Policing and Education Conference, 6*(1), 85–90.

Fisher, D. L., & Strayer, D. L. (2014). Modeling situation awareness and crash risk. *Annals of Advances in Automotive Medicine, 58*, 33–39.

Flach, J. M. (1995). Situation awareness: Proceed with caution. *Human Factors, 37*, 149–157.

Gibson, J. J., & Crooks, L. E. (1938). A theoretical field-analysis of automobile-driving. *The American Journal of Psychology, 51*, 453–471. https://doi.org/10.2307/1416145

Gugerty, L., Rando, C., Rakauskas, M., Brooks, J., & Olson, H. (2003). Differences in Remote versus in-Person Communications While Performing a Driving Task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 47*(16), 1855–1859. https://doi.org/10.1177/154193120304701605

Gugerty, L. (2011). Situation awareness in driving. In J. Lee, M. Rizzo, D. Fisher, & J. K. Caird (Eds.), *Handbook for driving simulation in engineering, medicine and psychology* (pp. 1–25). Boca Raton, FL: CRC Press.

Hauss, Y., & Eyferth, K. (2003). Securing future ATM-concepts' safety by measuring situation awareness in ATC. *Aerospace Science and Technology, 7*(6), 417–427. https://doi.org/10.1016/S1270-9638(02)00011-1

Hijaz, A., Louie, W.-Y. G., & Mansour, I. (2019). Towards a Driver Monitoring System for Estimating Driver Situational Awareness. 2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 1–6. Doi: 10.1109/RO-MAN46459.2019.8956378.

Horswill, M., & Mckenna, F. (2004). *Drivers' hazard perception ability: Situation awareness on the road* (pp. 155–175). A Cognitive Approach to Situation Awareness: Theory and Application.

Jannat, M., Hurwitz, D. S., Monsere, C., & Funk, K. H. (2018). The role of driver's situational awareness on right-hook bicycle-motor vehicle crashes. *Safety Science, 110*, 92–101. https://doi.org/10.1016/j.ssci.2018.07.025

Kaber, D. B., Riley, J. M., Sheik-Nainar, M. A., Hyatt, J. R., & Reynolds, J. P. (2006). Assessing infantry soldier situation awareness in virtual environment-based training of urban terrain operations. Proceedings of the International Ergonomics Association. Maastricht, The Netherlands.

Kass, S. J., Cole, K. S., & Stanny, C. J. (2007). Effects of distraction and experience on situation awareness and simulated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 10*(4), 321–329. https://doi.org/10.1016/j.trf.2006.12.002

Khoda Bakhshi, A., Gaweesh, S. M., & Ahmed, M. M. (2021). The safety performance of connected vehicles on slippery horizontal curves through enhancing truck drivers' situational awareness: A driving simulator experiment. *Transportation Research Part F: Traffic Psychology and Behaviour, 79*, 118–138. https://doi.org/10.1016/j.trf.2021.04.017

Lee, J. D., Gore, B. F., & Campbell, J. L. (1999). Display Alternatives for In-Vehicle Warning and Sign Information: Message Style, Location, and Modality. *Transportation Human Factors, 1*(4), 347–375. https://doi.org/10.1207/sthf0104_6

Lehtonen, E., Lappi, O., Koirikivi, I., & Summala, H. (2014). Effect of driving experience on anticipatory look-ahead fixations in real curve driving. *Accident; analysis and prevention, 70*, 195–208. https://doi.org/10.1016/j.aap.2014.04.002

Lo, J. C., Sehic, E., Brookhuis, K. A., & Meijer, S. A. (2016). Explicit or implicit situation awareness? Measuring the situation awareness of train traffic controllers. *Transportation Research Part F: Traffic Psychology and Behaviour, 43*. https://doi.org/10.1016/j.trf.2016.09.006

Lu, Z., Happee, R., & de Winter, J. C. F. (2020). Take over! A video-clip study measuring attention, situation awareness, and decision-making in the face of an impending hazard. *Transportation Research Part F: Traffic Psychology and Behaviour, 72*, 211–225. https://doi.org/10.1016/j.trf.2020.05.013

Ma, R., & Kaber, D. B. (2005). Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics, 35*(10), 939–953. https://doi.org/10.1016/j.ergon.2005.04.002

McKerral, A., Pammer, K., & Gauld, C. (2023). Supervising the self-driving car: Situation awareness and fatigue during highly automated driving. *Accident Analysis and Prevention, 187*. https://doi.org/10.1016/j.aap.2023.107068

Neisser, U. (1976). *Cognition and reality*. San Francisco, CA: W. H. Freeman and Company.

Plant, K. L., & Stanton, N. A. (2015). The process of processing: Exploring the validity of Neisser's perceptual cycle model with accounts from critical decision-making in the cockpit. *Ergonomics, 58*(6). https://doi.org/10.1080/00140139.2014.991765

Radlmayr, J., Brüch, K., Schmidt, K., Solbeck, C., & Wehner, T. (2018). Peripheral Monitoring of Traffic in Conditionally Automated Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 62*(1), 1828–1832. https://doi.org/10.1177/1541931218621416

Rebensky, S., Carroll, M., Bennett, W., & Hu, X. (2022). Impact of Heads-Up Displays on Small Unmanned Aircraft System Operator Situation Awareness and Performance: A Simulated Study. *International Journal of Human-Computer Interaction, 38*(5), 419–431. https://doi.org/10.1080/10447318.2021.1948683

Salmon, P. M., Lenne, M. G., Walker, G. H., Stanton, N. A., & Filtness, A. (2014). Exploring schema-driven differences in situation awareness between road users: An on-road study of driver, cyclist and motorcyclist situation awareness. *Ergonomics, 57*(2). https://doi.org/10.1080/00140139.2013.867077

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics, 39*(3), 490–500. https://doi.org/10.1016/j.ergon.2008.10.010

Schömig, N., & Metz, B. (2013). Three levels of situation awareness in driving with secondary tasks. *Safety Science, 56*, 44–51. https://doi.org/10.1016/j.ssci.2012.05.029

Schroeter, R., & Steinberger, F. (2016). Pokémon DRIVE: Towards Increased Situational Awareness in Semi-Automated Driving. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction* (pp. 25–29). https://doi.org/10.1145/3010915.3010973

Stanton, N. A. (2006). Hierarchical task analysis: Developments, applications, and extensions. *Applied Ergonomics, 37*(1), 55–79. https://doi.org/10.1016/j.apergo.2005.06.003

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*, 677–680.

Strater, L. D., & Bolstad, C. A. (2008). Simulation-based situation awareness training. *Human factors in simulation and training*, 129–148.

Soliman, A. M. (2010). Exploring the Central Executive in Situation Awareness. *Psychological Reports, 106*(1), 105–118. https://doi.org/10.2466/pr0.106.1.105-118

Tan, X., & Zhang, Y. (2022). The effects of takeover request lead time on drivers' situation awareness for manually exiting from freeways: A web-based study on level 3 automated vehicles. *Accident Analysis & Prevention, 168*, Article 106593. https://doi.org/10.1016/j.aap.2022.106593

Taylor R. M. (1990). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In Situational Awareness in Aerospace Operations (AGARD-CP-478) (pp. 3/1–3/17). Neuilly Sur Seine, France: NATO – AGARD.

Underwood, G., Chapman, P., & Crundall, D. (2009). Experience and visual attention in driving. In C. Castro (Ed.), *Human Factors of Visual and Cognitive Performance in Driving* (pp. 89–116). Boca Raton, FL: CRC Press.

van den Beukel, A. P., & van der Voort, M. C. (2017). How to assess driver's interaction with partially automated driving systems – A framework for early concept assessment. *Applied Ergonomics, 59*, 302–312. https://doi.org/10.1016/j.apergo.2016.09.005

van den Beukel, A. P., van der Voort, M. C., & Eger, A. O. (2016). Supporting the changing driver's task: Exploration of interface designs for supervision and intervention in automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 43*, 279–301. https://doi.org/10.1016/j.trf.2016.09.009

van der Kint, S. T., van Schagen, I., Vlakveld, W., Mons, C., de Zwart, R., & Hoekstra, T. (2024). A brief pc-based hazard prediction training program improves young novice drivers' hazard perception skills compared to a control group over time. *Transportation Research Part F: Traffic Psychology and Behaviour, 102*, 64–76. https://doi.org/10.1016/j.trf.2024.02.007

Ventsislavova, P., Rosenbloom, T., Leunissen, J., Spivak, Y., & Crundall, D. (2022). An online hazard prediction test demonstrates differences in the ability to identify hazardous situations between different driving groups. *Ergonomics, 65*(8), 1119–1137. https://doi.org/10.1080/00140139.2021.2016999

Vlakveld, W. (2011). *Hazard anticipation of young novice drivers: Assessing and enhancing the capabilities of young novice drivers to anticipate latent hazards in road and traffic situations*. University of Groningen. Doctoral Dissertation.

Walker, G.H., Stanton, N.A., & Salmon, P.M. (2018). Vehicle Feedback and Driver Situation Awareness (1st ed.). CRC Press. Doi: 10.1201/9781315578163.

Wickens, C. D. (1995). The tradeoff of design for routine and unexpected performance: Implications of situation awareness. In D. J. Garland, & M. R. Endsley (Eds.), *Experimental analysis and measurement of situation awareness* (pp. 57–64). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Wulf, F., Zeeb, K., Rimini-Döring, M., Arnon, M., & Gauterin, F. (2013, 6-9 Oct. 2013). Effects of human-machine interaction mechanisms on situation awareness in partly automated driving. Paper presented at the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013).

Young, M.S., & Stanton, N.A. (2023). Driving Automation: A Human Factors Perspective (1st ed.). CRC Press. Doi: 10.1201/9781003374084.