Little Road Driving HUD: Heads-Up Display Complexity Influences Drivers' Perceptions of Automated Vehicles

Rebecca Currano
Center for Design Research, Stanford University
bcurrano@stanford.edu

So Yeon Park
Center for Design Research, Stanford University
syjpark@stanford.edu

Dylan Moore
Center for Design Research, Stanford University
djmoore3@alumni.stanford.edu

Kent Lyons
Toyota Research Institute
kent.lyons@tri.global

David Sirkin
Center for Design Research, Stanford University
sirkin@stanford.edu

Figure 1: The study's two scenes. At left, the vehicle stops at an intersection, as pedestrians and cars cross. At right, the vehicle drives through several intersections, passing pedestrians, cars, traffic lights, and speed limit signs along the way. Both scenes show the Complex HUD detailed below.

ABSTRACT

Modern vehicles are using AI and increasingly sophisticated sensor suites to improve Advanced Driving Assistance Systems (ADAS) and support automated driving capabilities. Heads-Up-Displays (HUDs) provide an opportunity to visually inform drivers about vehicle perception and interpretation of the driving environment. One approach to HUD design may be to reveal to drivers the vehicle's full contextual understanding, though it is not clear if the benefits of additional information outweigh the drawbacks of added complexity, or if this balance holds across drivers. We designed and tested an Augmented Reality (AR) HUD in an online study ($N = 298$), focusing on the influence of HUD visualizations on drivers' situation awareness and perceptions. Participants viewed two driving scenes with one of three HUD conditions. Results were nuanced: situation awareness declined with increasing driving context complexity, and contrary to expectation, also declined with the presence of a HUD compared to no HUD. Significant differences were found by varying HUD complexity, which led us to

CHI '21, May 8–13, 2021, Yokohama, Japan

ACM ISBN 978-1-4503-8096-6/21/05…$15.00
https://doi.org/10.1145/3411764.3445575

explore different characterizations of complexity, including counts of scene items, item categories, and illuminated pixels. Our analysis finds that driving style interacts with driving context and HUD complexity, warranting further study.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; *Empirical studies in interaction design*; *Empirical studies in visualization.*

KEYWORDS

augmented reality, heads-up-display, situation awareness, interaction design, vehicle interface, user interface

ACM Reference Format:

Rebecca Currano, So Yeon Park, Dylan Moore, Kent Lyons, and David Sirkin. 2021. Little Road Driving HUD: Heads-Up Display Complexity Influences Drivers' Perceptions of Automated Vehicles. In *CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3411764.3445575

1 INTRODUCTION

Augmented Reality (AR) has become a truly productive, and even life-saving, technology, with applications from medicine [15] to manufacturing [8]. And yet, other than in a few early explorations [69], AR is absent from our modern vehicles, which are among the few places that we spend the most time, and where real-time computer analysis offers great potential to benefit safety and comfort.

The Heads-Up-Display (HUD) is a promising technology to introduce AR into road vehicles. As only smaller optical projection systems can fit within a vehicle's dashboard, current HUDs limit frame size to about a 20° horizontal field of view.[1] Interaction designers are thus challenged to represent ongoing information as well as fleeting (or possibly anticipated) events that occur inside, as well as outside, of this small frame.

An initial design approach might be to have the HUD present all of the information that the vehicle's sensor and analysis systems are aware of. After all, this seems transparent and communicative. However, such an approach gives rise to several concerns: (a) more complex displays could distract from the main task of driving, more so than spartan displays might; (b) it could be difficult for drivers to discern relevant, or critical, events from merely advisory information within complex displays; and (c) drivers' abilities to observe and interpret information are variable—both among people and for the same person across different times or contexts—so a single design approach might not be optimal.

In addition to these perceptual challenges, HUDs may introduce cognitive processing effects. First, redirecting drivers' attention toward a HUD can increase their cognitive load, which in turn can alter typical visual scanning of the driving environment [34, 51]. And second, different formats of HUDs' visual warnings can influence reaction times and detection rates, [32], which correspond to accident involvement [40].

Our contributions are inspired by two gaps in the literature: (1) the scope of prior explanations of HUD complexity, and (2) exploring the interaction between display complexity, driving style, and situation awareness. Prior work has generally considered HUD complexity to be the number of symbols visible in the display [9] or the busyness in the background scene [68]. But as we describe later, HUD complexity can be defined by categories or instances of scene features, or of symbols, or even a count of pixels, and can also be independent of the background. No prior work compares these.

In addition to display complexity, drivers' preferences and driving styles can also impact their behavior and perception of the vehicle [33, 47] and driving environment. For example, easily distracted drivers and attentive drivers likely notice different things in the environment, and would benefit from different interface support. A recent participatory design exercise [3] and follow-up evaluation study [2] by Becerra et al. compared the use of personal HUD display profiles for thrill seeking drivers, and found matches between display and driving style. While the work used still images, focused on a single driving style, and used basic descriptions to represent understanding of components of the display, it suggests a connection between display and one driving style not explored elsewhere.

Research within the CHI community has explored alternative AR visualizations [30], HUDs compared to handheld devices in non-driving task interruption [22], and windshield-displayed applications [25]. In the latter, Haeuslschmid et al. [25] compiled and categorized such applications into entertainment and communication, navigation, vehicle monitoring, and safety—however, only a few relate directly to supporting drivers' situation awareness. Little prior work, within or outside of CHI, has focused on the influence of complexity on drivers' perceptions and awareness from an interaction design perspective.

To better understand the ways that HUDs can support the driving task, and under what circumstances, we developed a study grounded in the preceding issues and structured around three research questions:

- **RQ1**: How does the complexity of a HUD's content influence drivers' situation awareness and perceptions of the vehicle?
- **RQ2**: How do differences in driving style mediate the HUD's influence on situation awareness or perceptions?
- **RQ3**: Do these effects vary across different driving contexts?

To address these questions, we developed an online video-based study (N = 298) focusing on AR HUDs, in which participants completed a driving style questionnaire, viewed two driving scenes (with a single HUD condition), and answered situation awareness questions after each video. This work contributes to building an understanding of how HUD complexity influences drivers' perceptions of the driving environment, the vehicle, and themselves, and of how HUD complexity itself can be evaluated. It also sheds light on the importance of driving style—a critical driver characteristic—to the effectiveness of HUDs in supporting situation awareness in different driving environments. From these insights, we present design recommendations to consider when designing AR HUDs, and suggest areas for further study.

[1]Based on conversations with designers and prototypes within Toyota Research Institute.

2 BACKGROUND

Several areas of research and design bear on our study and analysis, from HUDs, to visual scene processing, to situation awareness, to driving styles.

2.1 Heads-Up Displays

HUDs were introduced to production cars as early as the 1988 Oldsmobile Cutlass Supreme Indianapolis 500 Pace Car Parade Convertible [69]. Since then, while the breadth of content available has increased from vehicle status (such as speed or fuel) to include road and trip details (such as collision warnings or navigation), most standard or add-on HUDs on-the-road still reproduce the information available in existing displays [46].

Betancur suggests that to best support drivers' comprehension of the information displayed in a HUD, its images should appear to be located several meters ahead of the vehicle, within the driving environment [4]. When placed central to drivers' field of vision, HUDs are likely to capture their attention [10] and keep it there[32].

However, despite the apparent advantages of keeping drivers' heads up and their eyes on the road [35], HUDs can also monopolize drivers' attention through *cognitive capture* [70], potentially diverting them from from the broader context and leading to *primary/secondary task inversion* [73], where an operator begins to rely on the backup system (the HUD, in this case) as the primary source of information.

A relevant distinction is thus between two primary modes of visualization used for HUDs [45]. Static visualization refers to the stationary presentation of information on a particular part of the screen or windshield, but not in a location or manner that is registered to the world outside. This format can work well for the kind of content in current vehicle information displays. AR visualization, on the other hand, presents information such that it occurs visually near, and/or in realistic perspective with the objects and events that it refers to, perhaps moving as the scene changes. AR HUDs are therefore well suited to communicating spatially relevant information in context.

2.2 Visual Attention and Scene Parsing

As noted above, the use of HUDs can cause division and repeated shifts in attention between the road environment and the display, making visual attention an important aspect of HUD design.

One aspect of visual processing of scenes relates to focal versus ambient vision tasks [55], although both are interrelated and contribute to situation awareness [23, 72]. This is consistent with findings of Henderson et al., that cognitive factors are the dominant determinants of gaze control and that visual saliency does not, on its own, account for all eye movements in active viewing [27]. Castelhano and Henderson found that even initial saccades direct the gaze toward a reasonable location, based on preliminary understanding of the scene [11], suggesting the role of ambient vision in determining where to fixate focal vision by providing a quick, rough understanding of a scene.

In the context of everyday driving, focal vision supports identifying hazards on the road, while ambient vision relates primarily to keeping the vehicle moving forward within the lane [23]. Both relate to HUD use, in that they address central versus peripheral glancing behavior [41]. And drawing on Castellano and Henderson's findings, ambient vision may also provide support in initial detection of potential hazards while they are still on the periphery. With respect to visual processing in driving, researchers have found that fewer eye movements and longer fixations correlate with better driving performance [5], and that more experienced drivers show a bias to fixate toward the front and center [74], while novices bias toward the dashboard and sides [43]. Active driving in a complex environment also leads to a greater tendency to fixate toward the front of the vehicle [39].

Visual scene parsing in the driving context is therefore a more complex process than previously thought, involving visual and cognitive resources, as well as a meta task of managing the allocation of attention both within and between primary and secondary tasks [23].

Other research has explored the relationship between shifts in attention and cognitive load. Lee et al. found that "cognitive load and short glances away from the road are additive in their tendency to increase the likelihood of drivers missing safety-critical events" [34]. This finding aligns with earlier work that found cognitive interference to be greatest for spatial-imagery tasks, which demand the same resources as when one is driving [51].

2.3 Situation Awareness

Perception and avoidance of potential hazards is fundamental to drivers' ability to drive safely, and is dependent on adequate awareness of the driving environment. In her foundational work, Endsley explained situation awareness by three stages of active, conscious attention and cognitive processing [17, 18]: 1) perception of the elements in the current situation, 2) comprehension of this current situation, and 3) projection of future status. Gugerty makes the case that situation awareness involves different levels of cognitive processing, which he describes as: "1) automatic, pre-attentive processes that occur unconsciously and place almost no demands on cognitive resources; 2) recognition-primed decision processes that may be conscious for brief periods (<1 s) and place few demands on cognitive resources; and 3) conscious, controlled processes that place heavy demands on cognitive resources", and that safe driving requires managing attention both within and between tasks. [23].

Researchers use a variety of methods, including the Situation Awareness Global Assessment Technique (SAGAT) and performance measures, to measure situation awareness, which trade off immediacy, memory, objectivity, and awareness of the test itself [17]. Other methods include eye-tracking [61], subjective reporting methods such as the Situational Awareness Rating Technique (SART) [56, 64], performance measures [65], real-time question probes—exemplified by the Situation Present Awareness Method (SPAM) [13], and post-activity probe questions. Each has its benefits and limitations, and its applicability depends on the type of study being run.

Eye tracking allows researchers to objectively determine what participants are looking at at any given time, and to measure durations and patterns of fixations and saccades. In driving studies, eye tracking is generally used in in-person simulated settings where lighting is more consistent, since the high contrast and changes in natural lighting on the road can interfere with the sensors. The

SAGAT method is also suited to in-person simulated driving settings, since the simulation can be paused and resumed at various points to query the participant about different things. The nature of interrupting the experience allows for in-the-moment assessment of participants' awareness of the situation at any point in time. However, it also encourages participants to be on alert for queries about their situation awareness, which can both artificially heighten alertness toward the kinds of things being queried about (versus a similar experience in a natural setting), and distract them from other aspects of the experience. Together these effects can lead to incorrect conclusions about what they would attend to, and be aware of, in a comparable natural setting. SPAM [13, 14, 14] is another in-the-moment querying method, however it does not involve pausing the experience, and the scene remains in view, but it still intrudes on the experience and heightens participant alertness.

Alternatively, SART is a post-activity assessment of situation awareness and thus it avoids the drawbacks of intruding onto the experience. But it is removed in time from that experience, is highly subjective (asking participants to rate their awareness in more general terms), and does not query about specific events or objects encountered. This can lead to inadvertent biasing toward remembering times of heightened awareness, and forgetting times of lower awareness [54]. Endsely [17] suggested that subjective ratings measure confidence in one's situation awareness rather than their situation awareness itself. In a direct comparison study between SART and SAGAT [19], she found that SART measures of situation awareness correlated highly with subjective measures of situation awareness, confidence, and performance, and did not correlate with objective measures of situation awareness using the SAGAT method. Post activity questionnaires, on the other hand, provide more objective measures, have good validity [28], and have been used to measure differences in situation awareness for automated driving [59], although they lose reliability with longer time lapses between events and questions [44]. Shinohara et al. [59] found that correct responses to the post-activity questions correlated with longer and greater numbers of fixations on the associated scene features (shown from eye tracking data), further supporting the validity of post-questionnaires for measuring situation awareness.

2.4 Driving Style

The seriousness of vehicle accidents, their effects on heath and safety, and their high economic costs have motivated many studies of accident rates and their causes [26, 62, 71]. These have focused on various human factors and personality-related variables, including demographics, skills, behaviors, and attitudes related to driving.

More recent research has focused on discerning driving styles as an important consideration in the application of new technology in the car [20, 53, 62, 63, 66]. Driving styles refer to the ways that drivers drive habitually, and include speed, following distance, assertiveness, and attentiveness. [16, 62]. Driver assistance systems should understand their drivers' stylistic tendencies in deciding when and how to intervene, and in determining what information to provide when transitioning control back to the driver.

Taubman-Ben-Ari et al.[62], developed a self-report scale, the Multidimensional Driving Style Inventory (MDSI), and validated it with participants in Israel. The MDSI uses 44 questions to analyze subjects with respect to four broad driving characteristics, which they derived from a review of prior research scales: 1) reckless and careless [20, 50], 2) anxious [24], 3) angry and hostile [1, 12], and 4) patient and careful [20, 26]. Taubman-Ben-Ari's analysis revealed eight driving style factors: dissociative, anxious, risky, angry, high-velocity, distress reduction, patient, and careful, and showed correlations between driving style and various demographic and personality measures.

However, driving behaviors differ culturally and nationally, and various researchers have analyzed the inventory's validity and stability in different countries [21, 29, 31, 48, 66, 67]. These studies have found the MDSI to be useful, and have validated different versions with modifications for each represented location, however, it has not yet been analyzed and verified for US drivers.

3 METHODOLOGICAL APPROACH

3.1 Building Hypotheses

We formulated several hypotheses as part of addressing our research questions RQ1 and RQ2, which are aimed at better understanding the influence of HUD complexity and driving styles on situation awareness and perception of the vehicle.

- **H1a**: The use of HUDs to move relevant information into drivers' fields of view suggests the goal of supporting situation awareness in the driving task. In a complex driving environment, cognitive load may be reduced with the distillation and presentation of a limited number of potential hazards that might be overlooked. *We hypothesize that less complex HUDs will enhance drivers' situation awareness, especially when situated in a more complex driving environment.*
- **H1b**: Conversely, given the cognitively challenging nature of attending to the potential hazards on the road, while engaged in the driving task, *we hypothesize that more complex HUDs may degrade drivers' situation awareness, particularly when combined with a complex driving environment.*
- **H2a**: Additionally, considering that individual drivers have varying tendencies toward careful attention, distraction, taking risks, and other driving behaviors [58, 62, 66], *we hypothesize that driving style will interact with HUD complexity and driving context to affect situation awareness.*
- **H2b**: For example, *we expect easily distracted drivers to benefit from a HUD that highlights potential hazards they might otherwise overlook,* and in particular, for that HUD to be *simple and minimal*, as a complex design might itself be distracting. In particular, easily distracted drivers should respond in a way that contrasts with careful or attentive drivers, who would already be aware of the driving environment.

3.2 HUD Visual Design

We held two two-hour long design charrettes/workshops, each with 10–12 interaction design researchers as participants, with the agenda to (a) understand what kinds of information about the visible driving context drivers may need, and (b) develop a visual language that could represent this information through an AR display with world-registered features. Regarding process, both design sessions were idea-generative and followed the C-Sketch methodology [57].

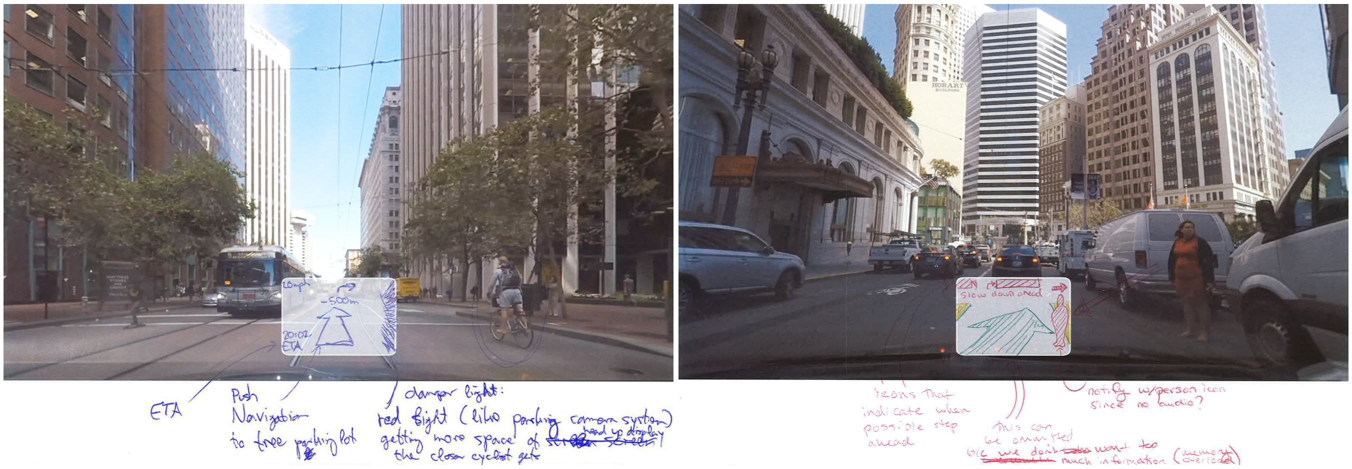

Figure 2: Two example sketches from our design charrettes, in response to the challenge of representing on-road activity and potential dangers, across two different scenes, both within and outside of the HUD's frame.

The first session helped clarify four types of contextual information important to support the safety of vehicle occupants and vulnerable road users, and to support drivers' awareness of both driving and travel contexts:

- *Critical:* Pedestrians, vehicles, and collidable objects
- *Routing:* Navigation symbols and drivable surfaces
- *Traffic:* Road (e.g., stop) signs and intersection lights
- *Vehicle:* Ongoing vehicle status (e.g., speed and fuel)

The second session focused on developing visualizations to represent road activity and potential dangers, and in particular, how to represent things within the HUD's visual area that are visually located outside of its boundary.

For example, Figure 2 shows two sketches that include navigational cues and the presence of a bicyclist or pedestrian on the right-hand side of the road, outside of the HUD's frame. The sketch on the left calls attention to the bicyclist by highlighting the nearest edge of the HUD; it is abstractly suggestive of something to the right, but does not imply an actual bicyclist within the HUD, and therefore is not problematic. The sketch on the right represents the pedestrian using an icon within the HUD; it thus might suggest that a second pedestrian is located at the icon's position, which could confuse drivers.

3.3 Study Design

Revisiting our goal from RQ1 to explore the influence of HUD complexity on situation awareness, we created HUDs with different levels of complexity by varying the number of *categories* of features visible in the driving environment highlighted by the HUD. The four types of information we had identified in the design charrettes served as the initial categories. This approach also indirectly influenced the overall *count* of highlighted features at any time, as people, vehicles, road signs, and traffic lights entered and exited the scene. Most of our analysis separates levels of complexity by the number of categories, but we expand on this concept through further analysis in Section 6.2. To account for RQ2, asking whether

Table 1: Scene features highlighted in the Minimal and Complex HUD conditions. In the Minimal condition, only road signs and intersection lights that required driver action were highlighted.

Information	Scene Feature	Minimal	Complex
Critical	Pedestrians	•	•
	Vehicles	•	•
	Collidable objects		•
Routing	Navigation symbols		•
	Drivable surfaces		•
Traffic	Road signs	◦	•
	Intersection lights	◦	•
Vehicle	Current speed		•
	Fuel remaining		•

driving styles mediate the HUD's effects, we included Taubman-Ben-Ari's driving style inventory. And since the HUD's complexity at any moment depends on the activity within the scene, we included two different driving scenarios as backdrops for the study, allowing us to explore RQ3 about whether the HUD's effects vary across contexts.

3.3.1 HUD Design. The final study design included three levels of HUD complexity:

- **No HUD**, where the driving video was presented without any HUD content added, just as a driver would experience it in a vehicle without a display.
- **Minimal**, which overlaid colored graphics onto a small area of the driving video to highlight pedestrians, vehicles, stop signs, and traffic lights that require action.
- **Complex**, which also included other road signs (such as speed limits), navigation symbols, drivable surfaces, all traffic lights, objects in the roadway, and vehicle status indicators for speed and fuel level.

Figure 3: Detail view of Scene 1, showing all 3 conditions and corresponding HUD overlays (below). The Complex visualization in this frame (left) shows a stop sign, yellow circles highlighting pedestrians, current speed and fuel level readings, a white navigation arrow marking the path ahead, and a sweeping white line indicating a drivable surface. The Minimal visualization (center) shows a stop sign, and yellow circles highlighting pedestrians. No visualizations are used for the No HUD condition. The HUDs' outer frame lines are shown here for clarity, but were not visible to participants, who saw the HUD and the full scene in the left panel of Figure 1.

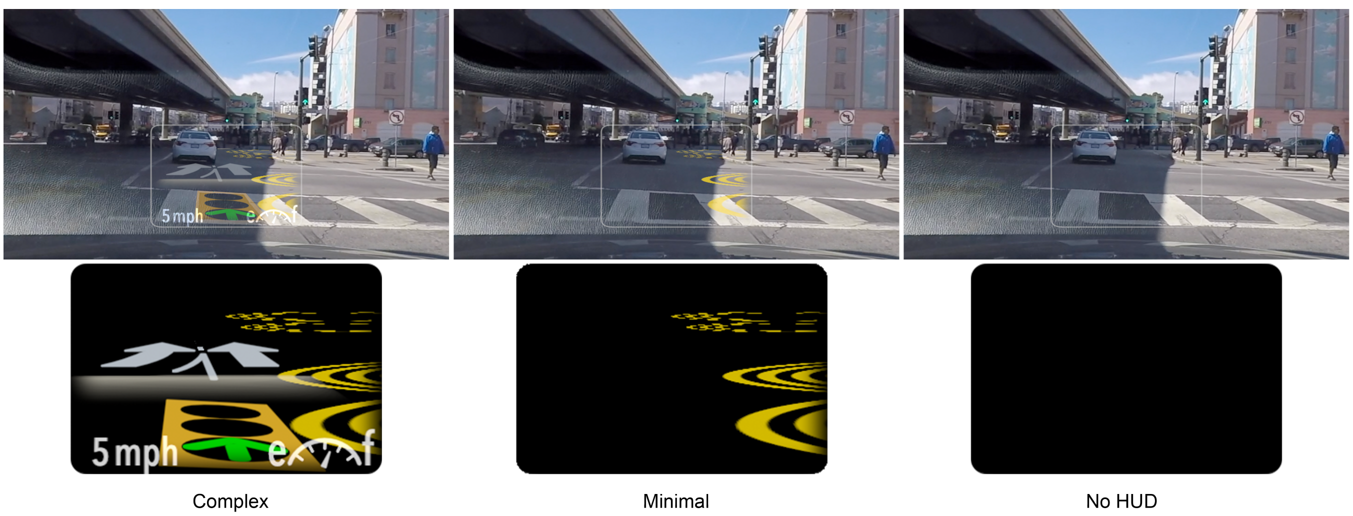

Figure 4: Detail view of Scene 2, showing all 3 HUD conditions and corresponding HUD overlays (below). The Complex visualization in this frame (left) shows a green traffic light, yellow circles highlighting pedestrians, current speed and fuel gauge readings, and white navigation arrows marking the merge area and path ahead. The Minimal visualization (center) shows yellow circles highlighting pedestrians. No visualizations are used for the No HUD condition. The HUDs' outer frame lines are shown here for clarity, but were not visible to participants, who saw the HUD and the full scene in the right panel of Figure 1.

The three conditions are shown for both scenes in Figures 3 and 4. We include outer frame lines in the figures to clarify the HUD's boundary, however these were not included in the study's HUD visualizations, to maintain consistency with the appearance of a real HUD. Also, as our main research goal was to understand the effects of HUD complexity across driving contexts, rather than to compare between alternative visual designs, we used the same visual language across all three HUD conditions and both scenes.

Table 1 distinguishes the scene's features that were highlighted for the Minimal and Complex conditions (the No HUD condition showed no highlighting).

All HUD visual features, aside from the vehicle's status indicators, were registered to the driving environment in AR style (that is, rendered in perspective, at appropriate locations). Given the HUD's small display area, we designed indicator highlights to radially extend out from objects along the ground plane. This caused them to appear in the HUD when the objects came into close proximity of the HUD, even if they never entered the HUD's boundary. Vehicle status indicators, present only in the Complex condition, remained visible and stationary at all times, while other scene features appeared on the roadway as the car approached them, and then disappeared as the car passed them. These features are also shown in Figures 3 and 4.

Our two videos were 16 and 18 seconds in duration. Video length needs to balance several factors: a) participants' ability to remember the objects and events in the video, b) the amount of content that they need to recall, c) their ability to maintain attention to the video, and d) overall study duration. Our scenes were recordings of actual city driving, and were somewhat complex, and we planned to use a post-activity questionnaire to assess situation awareness. It was important, therefore, to keep videos brief, to minimize both time and visual content between when participants viewed objects and events in the driving scene, and when they answered the questions about them. Additionally, as our study was conducted remotely, keeping the videos short also helped reduce the likelihood that distractions in participants' environments would avert their attention from watching the video. Finally, shorter videos allowed participants to view and answer questions about two scenes while minimizing survey fatigue (after each video, they took an average of 1.3 minutes to answer the situation awareness questions).

Prior work has shown brief video clips to be an effective technique for assessing differences in situation awareness in driving studies. Underwood et al. [65] used key presses while viewing driving videos to study situation awareness of novice drivers, experienced drivers, and experienced motorcycle rider-drivers with respect to abrupt and gradual onset road hazards using 20 to 95 sec videos. They found significant differences between road users in hazard detection rates, hazard response latency, and difference ratios between false alarms and hits. Lu et al. [36], and Lu et al. [37] used even shorter, 1 to 20 second, videos of both hazardous and non-hazardous traffic situations to measure differences in situation awareness over various timeframes, using a memory-based scene reproduction task at the end of each video. They found that errors in the number of vehicles placed in the scene, and in the distance and geometry between placed vehicles and actual vehicles decreased with video length, but only up to about 7 to 12 seconds. Our videos are comparable to those in these other studies, and effectively balance the trade-offs mentioned above.

3.3.2 Driving Scenes. The two driving scenes consist of video taken through the windshield of a car driving through streets in San Francisco, using a GoPro Hero4 set to wide-angle (160° field of view), sourced from prior work by Shinohara et al. [59]. The first scene (shown in the left panel of Figure 1) is relatively static, with the car coming to a stop at an intersection and waiting while pedestrians and cars cross, before starting to move again. The second scene (shown in the right panel of Figure 1) is more dynamic, beginning with a brief stop at a traffic light with a single car waiting ahead, then accelerating to 25 mph while passing several intersections, each with traffic lights, nearby vehicles, pedestrians, and road signs, before merging into a multiple-lane roadway.

3.3.3 Survey Development. We designed the study as an online video-based survey administered through Qualtrics, and then recruited participants using Prolific, a method shown to be effective in previous video survey-based research [47, 59].[2] The survey included the following five components (the questions can be found in this paper's Supplementary Material):

- **MDSI**: The six-point Likert scale Multidimensional Driving Style Inventory, comprised of 44 questions querying driving habits and tendencies, administered to all participants.
- **Scenes**: Two videos, each depicting a different driving scene, and overlaid with either No HUD, Minimal, or Complex visuals, depending on the HUD condition.
- **Situation awareness**: A battery of multiple-choice questions (six for the first driving scene and seven for the second) asking drivers to recall elements within the scene they just viewed, and referencing comparable types of items in both scenes. For three questions in each scene, if the participant answered correctly, a follow-up question would probe situation awareness on that item in greater detail. For example, if a question asked whether the participant had noticed a person standing alongside the road, the follow-up question would ask which side of the road.
- **Driver perceptions**: Two Check-All-That-Apply (CATA) questionnaires (shown only in Minimal and Complex HUD conditions) asking participants to select, from a set of descriptive words, those which they associated with the HUD, and a from a separate set, those which reflected how the HUD made them feel as drivers. HUD words included: simple, complex, clear, unclear, pleasant, annoying, familiar, unfamiliar, intuitive, unintuitive, helpful, unhelpful, supportive, and distracting. Driver words included: safe, unsafe, decisive, indecisive, aware, unaware, certain, uncertain, active, passive, calm, and stressed.
- **Demographics**: A concluding questionnaire covering participants' years of driving experience, experience with partially autonomous driving systems, age, and gender.

3.4 Procedure

The survey started with an introduction to the study's context, and asked participants to use a device with a 12 inch or larger screen, to adequately view details of the driving scene and HUD graphics. Participants then viewed a test video to verify that the player worked before advancing to the main survey. Figure 5 illustrates the main survey's process flow.

This study was approved by Stanford University's IRB, under exempt protocol 38139, and all participants viewed an information form before starting the study. Participants were assigned one of the three HUD conditions (No HUD, Minimal, or Complex), and one of two scene orders, determining which scene they saw first. All participants watched a video of both scenes. HUD condition and scene order were both assigned randomly, with HUD type balanced

[2]The Qualitrics online survey platform is at https://www.qualtrics.com and the Prolific participant recruitment platform is at https://www.prolific.co

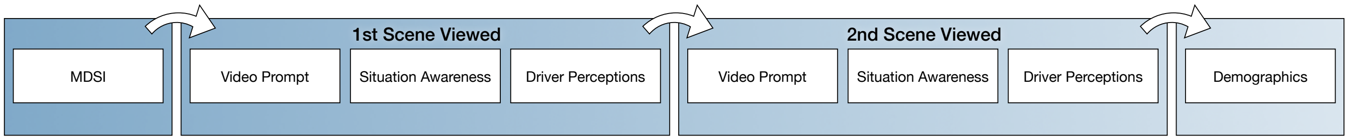

Figure 5: All participants began with the 44 question MDSI. They then watched two scene videos in random order. Each video prompt was immediately followed by a situation awareness test for that scene, and questions eliciting perceptions of the HUD. The study concluded with a brief demographic questionnaire.

overall, and scene order balanced within each HUD condition, to avoid order effects.

Participants first answered the 44 question MDSI, then watched one of the driving scene videos, followed by the situation awareness test. The videos had no player controls and advanced automatically after playing once, to ensure that participants viewed each clip only once before the test. After this, participants in the Minimal and Complex HUD conditions were shown the same video again, to refresh their memories, and then answered the two sets of CATA questions about their perceptions of the HUD. Participants in the No HUD condition saw the video only once, and were not administered the CATA questions (since they hadn't seen any HUD). All participants then repeated the video prompt, situation awareness test, and HUD perceptions questions (if appropriate) for the other driving scene video. The survey finished with a short demographic questionnaire.

3.5 Measures

The study design included two independent variables: scene, which was within subjects, and HUD condition, which was between subjects. It also included a single covariate, driving style, measured by calculating factor scores from the MDSI questionnaire. Dependent variables included situation awareness scores, and responses to the CATA questions on participants' perceptions of the HUD and how it made them feel as drivers.

3.6 Participants

We recruited 300 US-based participants, of whom 298 completed the survey. Of these, $N = 100$ viewed the No HUD condition, $N = 98$ viewed the Minimal HUD, and $N = 100$ viewed the Complex HUD. Regarding demographics, they ranged in age from 18 to 71 ($M = 35.1, SD = 12.4$) years, and in driving experience from 1 to 55 ($M = 17.7, SD = 12.3$) years. There were 156 (52%) males, 139 (47%) females, and 3 (1%) who reported as other or preferred not to answer. Regarding experience with automated systems, $N = 192$ (64%) reported none at all, $N = 92$ (31%) had a little or moderate amount, and $N = 14$ (5%) had a lot or a great deal.

4 ANALYTICAL APPROACH

Our primary interests for the analysis were to (a) identify any relationship between HUD complexity and drivers' situation awareness, (b) explore whether driving style might have a modifying effect on this relationship, and (c) compare whether participants' perceptions of the HUD supported their other responses, or revealed deeper insights about the HUD's influence on them.

We analyzed data using using R Studio 1.3.1093 [52] and the `stats`, `lme4`, and `tidyverse` libraries of R version 4.0.2 [49]. For factor analysis, we used the two functions `fa`, `fa.parallel`, and `factor.scores`; for chi square tests, we used `chisq.test`; for linear and mixed effects models, we used `lm` and `lmer`; and for rendering figures, we used `ggplot`.

4.1 Measuring Situation Awareness

We calculated situation awareness scores by dividing the number of questions each participant answered correctly by the total number of questions that person saw. This approach accounts for the slightly different number of questions in each driving scene, and avoids double-penalizing participants who did not see follow-up questions because they had answered the initial questions incorrectly. The approach yields a nearly linear correlation ($R^2 = .97$ for each driving scene) between the number of correct answers and the calculated score.

4.2 Determining Driving Styles

4.2.1 Factor Analysis. To determine driving styles, we first ran a parallel analysis with scree plot, both of which indicated a 7-factor solution. We then ran a factor analysis on the MDSI responses, applying the principle axis factoring method with oblimin rotation (an oblique rotation, which allows factors to be correlated). Following this, through an iterative process, we removed 13 (of the 44 total) questions which did not load highly on any factors, re-ran the parallel analysis, and re-ran the factor analysis. This resulted in a final factor structure consisting of 31 items distributed over six factors, which explained 46% of the variance.

We derived the following driving style names for these factors, based on the items loading highly (≥ 0.4) on each. Factor 1, *Anxious*, consists of five items which explain 10.4% of the variance (Cronbach's $\alpha = .88$). Factor 2, *Dissociative*, consists of nine items which explain 8.7% of the variance (Cronbach's $\alpha = .81$). Factor 3, *Impatient*, consists of six items, and explains 8.4% of the variance (Cronbach's $\alpha = .80$). Factor 4, *Risky*, consists of four items which explain 6.5% of the variance (Cronbach's α=.81). Factor 5, *Careful*, consists of three items explaining 6.4% of the variance (Cronbach's $\alpha = .81$). And Factor 6, *Stress Reduction*, consists of four items which explain 5.4% of the variance (Cronbach's $\alpha = .69$).

Our driving styles are comparable to those from prior work: Taubmen-Ben-Ari et al. used a varimax (orthogonal) rotation and found a distribution of 44 items over eight factors, explaining 56% of the variance, which they labeled: Dissociative, Anxious, Risky, Angry, High-Velocity, Distress Reduction, Patient, and Careful. Van

Huysduynen et al. used a varimax rotation and found a distribution of 36 items across six factors, explaining 48% of the variance, which they labeled as: Angry, Risky, Anxious, Dissociative, Careful, and De-Stress. They followed this with an oblique Principle Components Analysis, but lost one factor (Careful) due to the cutoff requirement of three high-loading items per factor.

4.2.2 Calculating Factor Scores. To calculate scores for the six different driving styles, we applied the ten Berge factor score method. This approach finds item loading weights that preserve oblique factor analysis solutions. Then, following the methods used by Shahab et al. [58] and van Huysduynen et al. [66], we calculated the mean and standard deviation of each participant's scores on all six factors, and set a threshold for each equal to their individual mean + standard deviation. For each participant, only scores that exceeded this threshold were included in their driving style profile. If a participant's scores did not exceed the threshold for any factor, that indicates that their response patterns were not sufficiently stronger in one factor than another to classify them within a particular driving style.

4.3 Relating HUD Complexity, Situation Awareness, and Driving Style

After calculating participants' driving style scores and their situation awareness for both driving scenes, we searched for relationships between them. Since the experiment had a mixed design—with both between-subjects (HUD complexity and driving style) and within-subjects (scene) conditions—we ran several analyses using linear mixed effects (LME) models. LME models yield similar results as mixed model ANOVAs, but allow for greater flexibility for data form and conditions, including the effects of continuous between and within subjects variables [42]. LME models also support random effects for conditions as well as participants, which ANOVAs do not, potentially lowering the chance of type 1 error [6].

To analyze the effect of HUD condition and scene on situation awareness (hypotheses H1a and H1b), we fit a linear mixed effects model with HUD and scene as fixed effects, and participant identifier as a random effect. Hypothesis H2a anticipated that driving style would interact with HUD complexity and/or scene, so we next ran another linear mixed effects model, testing for interaction effects between HUD condition, scene, and driving style. For hypothesis H2b, we tested significance only for participants whose styles (identified in the factor analysis) corresponded to Distracted and Careful driving.

4.4 Drivers' Perceptions

To examine the relationship between the number of CATA words selected for the Minimal and Complex HUDs, we compared counts of participant responses and ran a chi-square test of independence for each scene.

5 RESULTS

5.1 Driving Styles

Results of the factor analysis verified the effectiveness of the MDSI for identifying driving styles among US participants, further validating the work originally developed by Taubman-Ben-Ari et al. [62] as useful, with some modifications for this population. Our analysis identified six driving styles, closely matching six of the eight found by Taubman-Ben-Ari, and categorized drivers as having zero ($N = 15$), one ($N = 242$), or two ($N = 41$) driving styles in their profiles, which van Huysduynen [66] also found. Drivers who had no identified driving style (that is, whose factor scores did not exceed their individual mean + standard deviation threshold for any of the six styles) were grouped into a separate *No Style* category. Later in the analysis, we assigned drivers who had two driving styles to the group that corresponded to their higher-scored style, to avoid counting any driver twice in the analysis.

The seven identified primary driving style categories, and the number of participants in each were thus: Anxious ($N = 51$), Dissociative ($N = 35$), Impatient ($N = 47$), Risky ($N = 40$), Careful ($N = 69$), Stress Reduction ($N = 41$), and No Style ($N = 15$).

5.2 Effects of Scene, HUD Complexity, and Driving Style on Situation Awareness

5.2.1 Scene vs Situation Awareness. The analysis from our first linear mixed effects model shows that scene had a significant effect on situation awareness, with participants scoring consistently higher for the more static Scene 1 than for the more dynamic Scene 2 ($\beta = -.25, t(298) = -18.38, p < .001$) (see Figure 6).

5.2.2 HUD Condition vs Situation Awareness. HUD condition also showed significant affects on situation awareness, with participants scoring higher in the No HUD condition than in either the Minimal ($\beta = -.04, t(298) = -2.04, p = .04$) or Complex ($\beta = -.04, t(298) = -2.40, p = .02$) conditions (see Figure 6). Situation awareness scores were not significantly different between the Minimal and Complex HUD conditions ($\beta = -.01, t(298) = -.35, p = .73$).

5.2.3 Driving Style vs Situation Awareness. The analysis of the second linear mixed effect model revealed several 2-way and 3-way significant interactions between driving style and HUD, driving style and scene, and driving style, HUD, and scene.

For the 2-way effects, there were significant interactions between Anxious and Complex ($\beta = 0.34, t(277) = 2.67, p < .001$), Anxious and Scene 2 ($\beta = 0.29, t = 2.36, p = .02$), and Impatient and Scene 2 ($\beta = 0.28, t(277) = 2.19, p = .03$).

For the 3-way effects, we found significant interactions between: Anxious, Complex, Scene 2 ($\beta = -0.61, t(277) = -3.62, p < .001$), Careful, Complex, Scene 2 ($\beta = -0.43, t(277) = -2.58, p = .01$), Impatient, Complex, Scene 2 ($\beta = -0.48, t(277) = -2.75, p = .006$), Impatient, Minimal, Scene 2 ($\beta = -0.42, t(277) = -2.48, p = .01$), and Risky, Complex, Scene 2 ($\beta = -0.45, t(277) = -2.54, p = .01$).

5.2.4 Dissociative Driving and the Minimal HUD. We tested the performance of Dissociative drivers with a Minimal HUD (Hypothesis H2b) with two approaches. First we compared Complex versus Minimal HUD for only Dissociative drivers ($\beta = -0.02, t(44) = -0.42, p = .68$), and then we compared Dissociative and Careful driving styles for only the Minimal HUD ($\beta = -0.01, t(70) = -0.16, p = .87$), with both results insignificant.

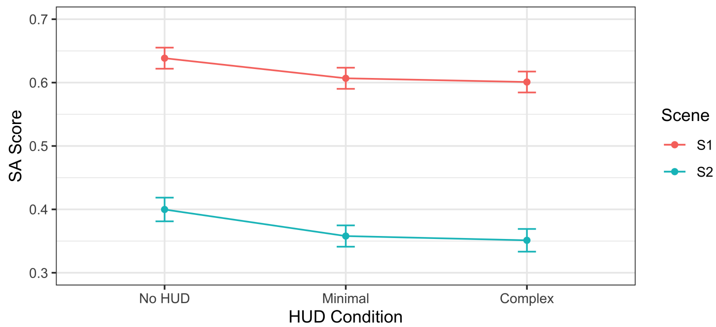

Figure 6: Mean situation awareness (SA) scores, across participants for each of the three HUD conditions, for each driving scene. Individual scores were calculated by summing each participant's count of correct responses to a battery of post-video questions, and then dividing by the total number of questions seen by that participant. Performance in both Minimal and Complex HUD conditions was significantly lower than in the No HUD condition.

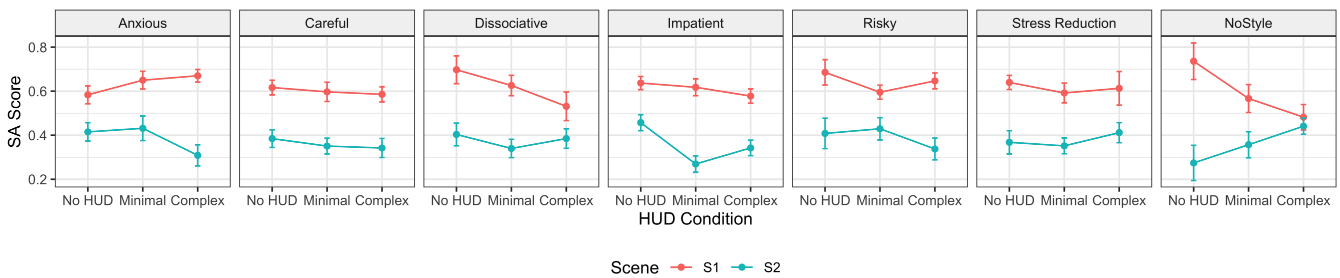

Figure 7: Mean situation awareness (SA) scores for each of the three HUD conditions, across participants, for each driving scene, per driving style. Individual scores were calculated by summing each participant's count of correct responses to a battery of post-video questions, and then dividing by the total number of questions seen by that participant.

5.3 Driver Perceptions

Figure 8 shows participants' choices of words describing their perceptions of the HUD, and Figure 9 shows their choices of words describing how the HUD made them feel as drivers.

The chi-square test comparing Minimal and Complex HUDs was significant in several cases, all for Scene 2. Regarding participants' perceptions of the HUD, more participants in the Complex condition found the HUD to be *helpful* $\chi^2(1, N = 198) = 5.27, p = .021$, and more in the Minimal condition found the HUD to be *unhelpful* $\chi^2(1, N = 198) = 5.53, p = .019$, and *unclear* $\chi^2(1, N = 198) = 4.74, p = .029$. Regarding how the HUD made participants feel as drivers, more participants in the Complex condition found the HUD to make them feel *aware* as drivers $\chi^2(1, N = 198) = 3.95, p = .047$. No other words showed significant differences.

6 DISCUSSION

Our research revealed the importance of visual complexity for drivers' situation awareness, but in a more complicated way than we had predicted. Most likely, various features of both the driving environment and HUD combine to contribute to drivers' experience of visual complexity, and have a compound effect on their situation awareness.

6.1 Scene Complexity and Situation Awareness

While we initially included two scenes as a secondary variable to explore the influence of different HUD designs on situation awareness in more than one context, we found that scene had a much stronger main effect than HUD type on situation awareness. While the more static Scene 1 had fewer *types* of indicators than the more dynamic Scene 2, this was counteracted by a significantly higher number and density of *instances* of pedestrian indicators. Yet in all HUD conditions, participants scored much higher on situation awareness for Scene 1. This finding suggests that visual complexity is a composite of multiple factors (in addition to HUD complexity), such as the number of roadway features and pedestrians in the environment, movement of pedestrians within the environment, and the speed at which the vehicle moves through the environment, and that the factors comprising *scene complexity* may have a greater effect on situation awareness than the more abstract *contextual complexity* of HUD design.

6.2 HUD Complexity and Situation Awareness

Situation awareness scores showed no significant difference between Minimal and Complex HUD conditions, and thus we were not able to support Hypothesis H1a. But as the complexity of the visuals increased from No HUD to either Minimal or Complex HUD

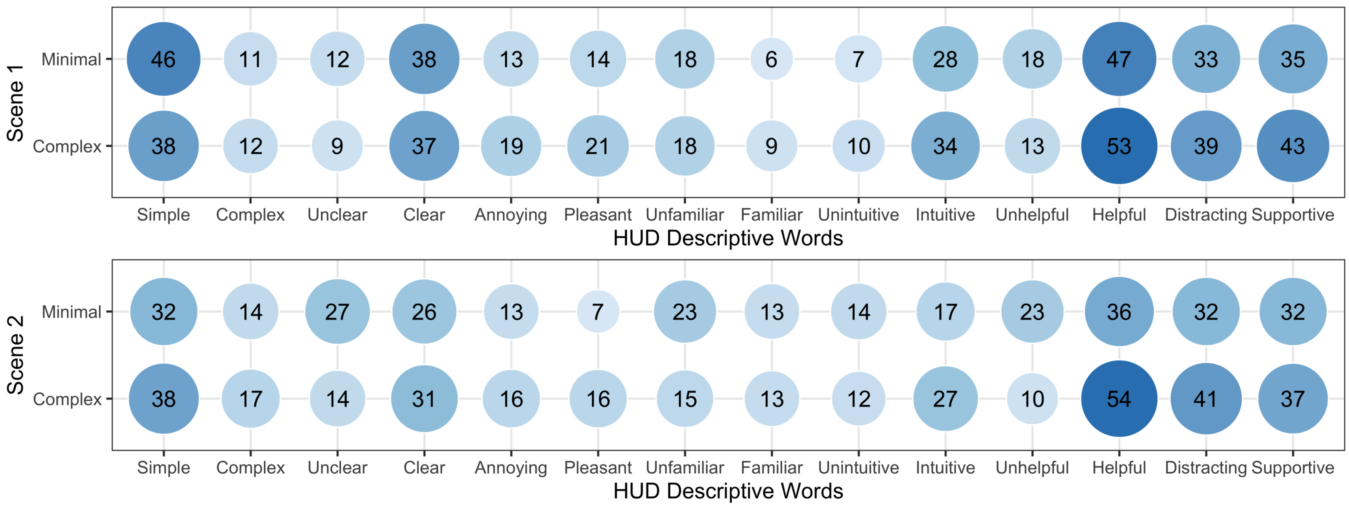

Figure 8: Counts for each CATA word that participants selected to describe their perceptions of the HUD. Size and hue of circles corresponds to count values, to aid visual comparison.

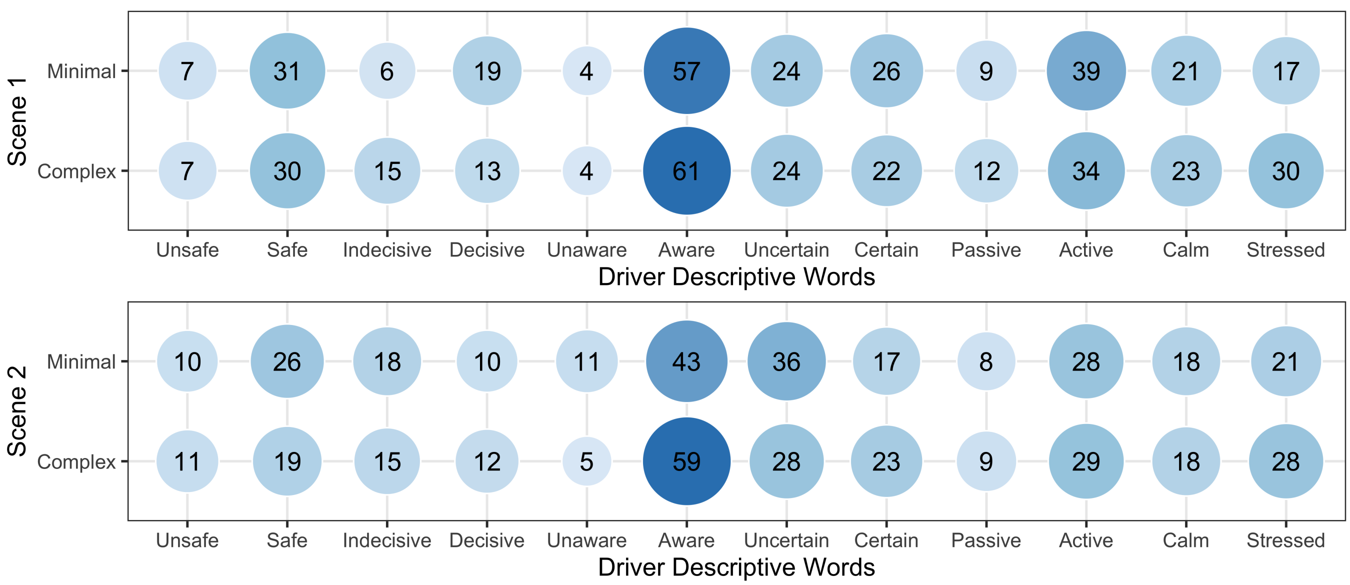

Figure 9: Counts for each CATA word that participants selected to describe how the HUD made them feel as drivers. Size and hue of circles corresponds to count values, to aid visual comparison.

(in both scenes), situation awareness scores decreased significantly, which supports Hypothesis H1b, that increased visual complexity in HUDs can have a negative effect on situation awareness.

These findings suggest that the specific Minimal and Complex HUD designs in our study may not be different enough to be noticed reliably, or that differences in the category complexity designed into the HUDs may be overpowered by differences in overall scene complexity or by other aspects of visual display complexity. In particular, we suspect that the number of categories of indicators used by a particular HUD design may have a smaller impact on situation awareness than the number of instances of indicators displayed at a given time and location. Additionally, it is likely that scene complexity due to the density of objects and people in the environment, and the speed of movement through that environment, may outweigh the display complexity in the the HUD design. Alternatively, there may be a threshold above which increased complexity has no measurable effect.

One potential explanation for drivers' lower performance in the Minimal and Complex HUD conditions compared to the No HUD condition is cognitive capture [7], which makes it more likely that drivers will miss activity that is farther away from the center of their vision. If so, this effect suggests even greater need for HUDs to highlight these features.

Either way, HUDs may need to adjust their behavior in different road environments in order to effectively support, and not impair, situation awareness.

6.3 Measures of HUD Complexity

Our initial concept of HUD complexity implicitly assumed that highlighting more *categories* of features in the driving environment (e.g. pedestrians, cars, traffic lights, road signs) meant greater complexity, but as our analysis evolved, we realized that HUD complexity was more...complex. In particular, a HUD design that highlights fewer categories (our Minimal HUD) might appear more visually crowded during a drive through a busy downtown street than a design that highlights many categories (our Complex HUD) would when stopped at a quiet intersection. Similarly, highlighting a greater *count* of features, regardless of their categories, could show the same reversal of complexity under these same driving conditions. In both cases, the count of illuminated screen *pixels* represents the even more implicit measure of visual, rather than of contextual, complexity.

Figure 10 shows a time series and mean count of highlighted features for both Minimal and Complex HUDs over the duration of both scenes. While the boxplot in the right panel reflects the intent of our current HUD design, where both Complex HUDs generally show more features than both Minimal HUDs, the left panel shows that there can be considerable overlap in which HUD shows more features at any given time.

Figure 11 shows a time series and mean percent of the HUD's pixels that are illuminated for both HUDs across both scenes. With this perspective, both panels show how Scene 1's Minimal HUD (shown in turquoise) almost always lights more pixels than Scene 2's Complex HUD (in green). This makes sense, as Scene 1 was relatively static, with a stop sign filling the lower half of the HUD (due to its proximity), while Scene 2 had various features moving quickly through the HUD's frame. So the visual complexity as measured by illuminated pixel count does not reflect the intended complexity of the current HUD (based on categories).

6.4 Driving Style Effects

The various significant 2-way and 3-way interactions that we found between driving style, HUD, and/or scene, support Hypothesis H2a, that driving style should interact with HUD and scene to influence situation awareness.

Figure 7 helps us visualize and understand some of these interactions. For example, we can see clearly that anxious drivers' situation awareness scores have a very different pattern than No Style drivers' scores in the No HUD and Complex HUD conditions for both Scene 1 and Scene 2. Anxious drivers also score differently than Dissociative drivers in the No HUD and Complex HUD conditions in Scene 1. Likewise, Risky drivers seem to show a different pattern than Impatient drivers, especially for the minimal HUD condition in Scene 2. Therefore, in some driving contexts, a particular HUD design may support situation awareness for some drivers, and degrade it for others.

Because we hypothesized that there would be interactions between driving style and situation awareness, but we did not anticipate their strength or direction, except for the example in H2b, we ran only a preliminary model, setting Scene 1 with No HUD and No Style as the baseline condition for comparison. The significant findings suggest that driving styles are a rich topic for further exploration, and in particular, how HUD designs might be optimized for drivers with different styles.

Regarding Hypothesis H2b, comparing panels 2 (Careful) and 3 (Dissociative) of Figure 7 suggests that Dissociative drivers may perform worse than Careful drivers when viewing the Complex HUD, as expected, but only for Scene 1. It does not show that they experience any benefit in performance (relative to Careful drivers) when viewing the Minimal HUD, and, they may perform worse than having no HUD at all, contrary to expectation. The model also shows no significant difference, and thus does not support the hypothesis.

6.5 Driver Perceptions of the HUD

The significant results (all in Scene 2), reveal that participants found the Complex HUD more helpful and less unhelpful than the Minimal HUD. They also found the Minimal HUD more unclear than the Complex HUD. Yet while participants *felt* that the complex interface served them better than the minimal one (there were no word selections in the No HUD condition), their situation awareness scores overall were not significantly different between the two HUDs.

Considering broader trends in the data, we observe in Figure 8 that response patterns between Minimal and Complex HUDs for most words are quite similar from Scene 1 to Scene 2. The most frequently selected words (in decreasing order) were *helpful*, *supportive*, *distracting*, *simple*, and *clear*. Four of these five most-selected words are positive. The one negative word, *distracting*, suggests that the HUDs may increase demand on drivers and split their visual attention [23, 51].

6.6 Driver Perceptions of Themselves with Different HUDs

The significant results (all in Scene 2), show that participants felt the Complex HUD made them more aware than the Minimal HUD. This agrees with the trend in their perceptions of both HUDs being helpful and supportive. But it also reveals a discrepancy between their perceptions of improved awareness, and the lack of difference in their actual situation awareness scores between HUDs. Therefore, we suggest that designers consider how HUDs affect driver performance as well as perceptions when developing new designs.

Considering broader trends in word selections, we find in Figure 9 that response patterns are similar across scenes, though somewhat variable across HUD conditions, with *aware* the most selected word, followed by *active*. In Scene 1, other frequently selected words included *stressed* and *safe*, while in Scene 2, they included *stressed* and *uncertain*. There is thus a greater mix of positive and negative words about drivers' perceptions of themselves than about the HUDs.

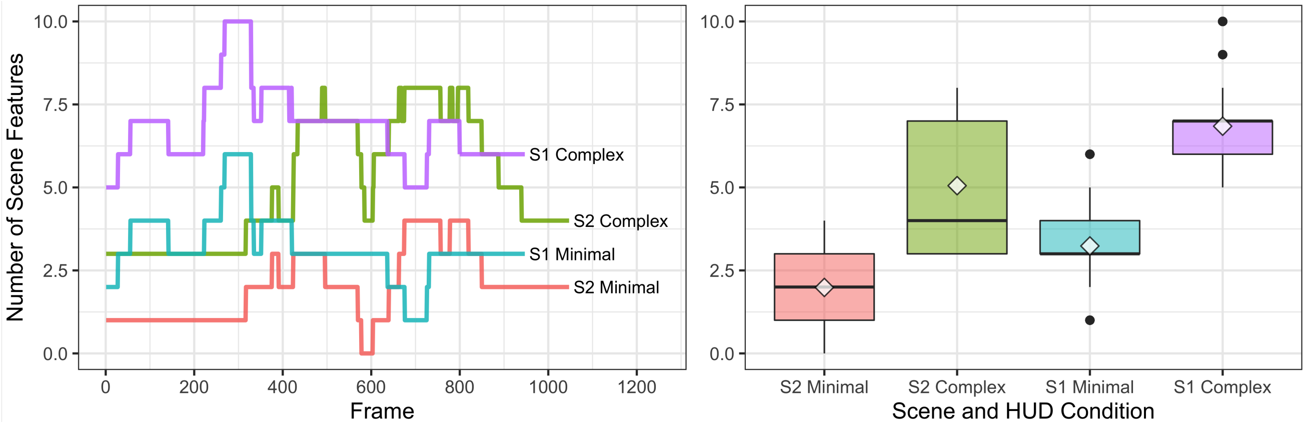

Figure 10: Time series (left) and boxplot (right) representations of HUD complexity based on count of highlighted scene features.

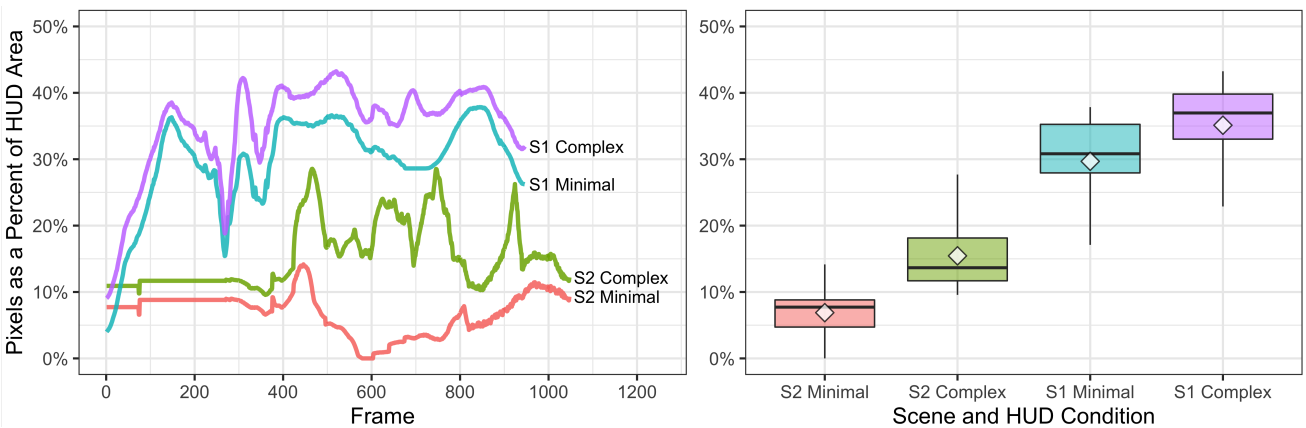

Figure 11: Time series (left) and boxplot (right) representations of HUD complexity based on percent of the HUD's pixels that are illuminated.

7 DESIGN RECOMMENDATIONS

Our results suggest that driving style is a critical factor in the HUD's role of raising drivers' situation awareness in various settings. However, HUDs may not be as effective in enhancing drivers' situation awareness as they are intended to be, given the strong influence of the immediate road environment.

We observed some drivers, including those with no distinct style, performing worse (on situation awareness measurements) with greater HUD complexity in the more static Scene 1, while also performing better with the same HUD in the more dynamic Scene 2. Other drivers, such as those with the Anxious driving style, demonstrated a reverse pattern, performing better with the Complex HUD in Scene 1, and worse with the same HUD in Scene 2. Therefore, designers should account for both driving style and scene complexity when designing the kinds and amount of information to include in HUD visualizations.

Not only do we see these differences in performance patterns in situation awareness metrics, but also in driver perceptions of the vehicle and their own states. HUDs, and in turn, vehicles, may be perceived differently depending on both scene complexity and driver characteristics. Therefore, we propose that HUDs be designed to account for drivers' individual styles, and also to change their presentation behavior (regarding complexity, in particular) based on the immediate driving context.

As the same HUD can be seen as clear or unclear in different situations, more work is needed to understand what information to display, and how to clearly display it in a given context. In effect, there may not be an optimal solution that improves situation awareness across drivers and situations. This insight suggests that adaptive displays, rather than fixed comprehensive designs, may provide the most benefit to a greater number of drivers.

8 LIMITATIONS AND FUTURE WORKS

We conducted our investigation using a survey instrument, which enabled collection of data from a wide range of driving styles across a large and diverse geographic area, but also lacks the degree of ecological validity that a more immersive simulator or on-road study might afford. Further studies of HUD complexity in these contexts will not only verify results, but also allow researchers to examine behavioral responses, including eye movements or physical (e.g., steering or braking) reactions and timing.

Moreover, situation awareness was gauged through question probes *after* participants had completed watching of the scenes presented. This has the advantage of intruding less upon in-the-moment attention than real-time questioning [13, 60] or freeze-frame techniques [17, 38] might, and thus helps to avoid influencing participants' gaze patterns or priming them to be more attentive than they would typically be, but it also introduces uncertainty by conflating participants' situation awareness with their near-term memory. Follow-up work that (also) includes real-time (e.g. eye gaze) measurements could resolve this uncertainty and provide a clearer distinction of situation awareness.

Our study design varied both complexity of the elements within a scene and the scene itself. At different times, any one scene can present relatively unchanging elements (such as traffic lights and road signs) and other, frequently-changing elements (pedestrians, cars, objects in the road). A future study could compare HUD conditions for one scene with different road conditions, separating complexity due to the scene's characteristics from complexity represented by using different scene contexts.

The current study focused on understanding the impact of HUD and scene complexity, with the driving styles analysis as a more exploratory aspect. Therefore, we did not recruit sufficient participants such that the represented drivings styles would be balanced, which also may have required a significantly larger sample. Our results suggest that further study of driving styles is promising, and may be critical for a user-centered approach to HUD design.

It may be noted that driving style seems to have an emotional component (as is reflected in the names of some of the styles). However, driving styles are considered to be influenced by drivers' attitudes and beliefs about driving, and by their broader needs and values [16, 62], rather than the result of transient emotional states. Therefore, drivers may demonstrate behaviors indicative of their driving style regardless of whether they feel a particular emotion in a given situation. Our analysis found that driving style does interact significantly with HUD and scene to affect situation awareness. In-the-moment emotion may have additional effects on situation awareness, independent from, or in addition to driving style effects, however this is not the subject of the current study, and further research would be needed to explore these potential effects.

An earlier design of the HUD, and thus the study, included another condition based on the timeliness of information presented within the HUD, with graphic features appearing or disappearing as corresponding scene items became more or less relevant to the driving task (such as a change in traffic light color), or appearing briefly to announce the presence of a scene item (such as an obstacle in the road), and then disappearing. Presenting time-dependent information, as well as related concepts—such as representing the system's uncertainty over the presence or location of scene items, or of anticipated events that the vehicle may be aware of, but which are not yet visible to the driver—are also promising directions for future HUD design and explorations.

ACKNOWLEDGMENTS

We would like to thank Toyota Research Institute (TRI) for supporting this work. TRI provided funds to assist the authors with their research but this article solely reflects the opinions and conclusions of its authors and not TRI or any other Toyota entity.

REFERENCES

[1] Jeffrey Jensen Arnett, Daniel Offer, and Mark A Fine. 1997. Reckless driving in adolescence:'State'and 'trait'factors. *Accident Analysis & Prevention* 29, 1 (1997), 57–63.
[2] Zoe M Becerra, Brittany Holthausen, and Bruce N Walker. 2019. Bringing the thrill to automated vehicles: an evaluation of thrill-seeking driving displays. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings*. 381–386.
[3] Zoe M Becerra, Brittany E Holthausen, and Bruce N Walker. 2019. Designing displays in automated vehicles for the thrill-seeking driver: Characterizing driving profiles to create personalized driving displays. *Technology, Mind, and Society* (2019).
[4] J Alejandro Betancur. 2011. Physical Variable Analysis Involved in Head-Up Display Systems Applied to Automobiles. *Augmented Reality-Some Emerging Application Areas* (2011), 244–266.
[5] Zheng Bian, Russell Pierce, and George Andersen. 2011. Eye movement patterns and driving performance. (2011).
[6] Matthieu P Boisgontier and Boris Cheval. 2016. The anova to mixed model transition. *Neuroscience & Biobehavioral Reviews* 68 (2016), 1004–1005.
[7] Brittisha N Boston and Curt C Braun. 1996. Clutter and display conformality: Changes in cognitive capture. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 40. SAGE Publications Sage CA: Los Angeles, CA, 57–61.
[8] Eleonora Bottani and Giuseppe Vignali. 2019. Augmented reality technology in the manufacturing industry: A review of the last decade. *IISE Transactions* 51, 3 (2019), 284–310.
[9] GE Burnett and RA Donkor. 2012. Evaluating the impact of Head-Up Display complexity on peripheral detection performance: a driving simulator study. *Advances in transportation studies* 28 (2012), 5–16.
[10] JL Campbell, JL Brown, JS Graving, CM Richard, MG Lichty, T Sanquist, PL Bacon, R Woods, H Li, DN Williams, et al. [n.d.]. "Human factors design guidance for driver-vehicle interfaces,"Report No. *DOT HS* 812 ([n. d.]), 360.
[11] Monica S Castelhano and John M Henderson. 2007. Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance* 33, 4 (2007), 753.
[12] Dennis M Donovan, Robert L Umlauf, and Philip M Salzberg. 1988. Derivation of personality subtypes among high-risk drivers. *Alcohol, Drugs & Driving* (1988).
[13] Frank T Durso, Andrew R Dattel, S Banbury, and Sebastien Tremblay. 2004. SPAM: The real-time assessment of SA. *A cognitive approach to situation awareness: Theory and application* 1 (2004), 137–154.
[14] Francis T Durso, Todd R Truitt, Carla A Hackworth, Jerry M Crutchfield, Danko Nikolic, Peter M Moertl, Daryl Ohrt, and Carol A Manning. 1995. Expertise and chess: A pilot study comparing situation awareness methodologies. *Experimental analysis and measurement of situation awareness* (1995), 295–303.
[15] Martin Eckert, Julia Volmerg, and Christoph Friedrich. 2019. Augmented reality in medicine: Systematic and bibliographic review. *JMIR mHealth and uHealth* 7, 4 (2019), 1–17.
[16] James Elander, Robert West, and Davina French. 1993. Behavioral correlates of individual differences in road-traffic crash risk: An examination of methods and findings. *Psychological bulletin* 113, 2 (1993), 279.
[17] Mica R Endsley. 1995. Measurement of situation awareness in dynamic systems. *Human factors* 37, 1 (1995), 65–84.
[18] Mica R Endsley. 1995. Toward a theory of situation awareness in dynamic systems. *Human factors* 37, 1 (1995), 32–64.
[19] Mica R Endsley, Stephen J Selcon, Thomas D Hardiman, and Darryl G Croft. 1998. A comparative analysis of SAGAT and SART for evaluations of situation awareness. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 42. SAGE Publications Sage CA: Los Angeles, CA, 82–86.
[20] Davina J French, Robert J West, James Elander, and John Martin WILDING. 1993. Decision-making style, driving style, and self-reported involvement in road traffic accidents. *Ergonomics* 36, 6 (1993), 627–644.
[21] F Freuli, G De Cet, M Gastaldi, F Orsini, M Tagliabue, R Rossi, and G Vidotto. 2020. Cross-cultural perspective of driving style in young adults: Psychometric evaluation through the analysis of the Multidimensional Driving Style Inventory. *Transportation research part F: traffic psychology and behaviour* 73 (2020), 425–432.
[22] Michael A Gerber, Ronald Schroeter, Li Xiaomeng, and Mohammed Elhenawy. 2020. Self-Interruptions of Non-Driving Related Tasks in Automated Vehicles: Mobile vs Head-Up Display. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–9.

[23] Leo Gugerty et al. 2011. Situation awareness in driving. *Handbook for driving simulation in engineering, medicine and psychology* 1 (2011).

[24] E Gulian, G Matthews, Aleck Ian Glendon, DR Davies, and LM Debney. 1989. Dimensions of driver stress. *Ergonomics* 32, 6 (1989), 585–602.

[25] Renate Haeuslschmid, Bastian Pfleging, and Florian Alt. 2016. A design space to support the development of windshield applications for the car. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 5076–5091.

[26] Niki Harré. 2000. Risk evaluation, driving, and adolescents: A typology. *Developmental Review* 20, 2 (2000), 206–226.

[27] John M Henderson, James R Brockmole, Monica S Castelhano, and Michael Mack. 2007. Visual saliency does not account for eye movements during visual search in real-world scenes. In *Eye movements*. Elsevier, 537–III.

[28] DJ Herrmann. 1984. Questionnaires about memory. *Everyday memory, actions and absent-mindedness* (1984), 133–151.

[29] Andrei C Holman and Corneliu E Havârneanu. 2015. The Romanian version of the multidimensional driving style inventory: Psychometric properties and cultural specificities. *Transportation research part F: traffic psychology and behaviour* 35 (2015), 45–59.

[30] Richie Jose, Gun A Lee, and Mark Billinghurst. 2016. A comparative study of simulated augmented reality displays for vehicle navigation. In *Proceedings of the 28th Australian conference on computer-human interaction*. 40–48.

[31] Juffrizal Karjanto, Nidzamuddin Md Yusof, Jacques Terken, Muhammad Zahir Hassan, Frank Delbressine, Hanneke Hooft van Huysduynen, and Matthias Rauterberg. 2017. The identification of Malaysian driving styles using the multidimensional driving style inventory. In *MATEC Web of Conferences*, Vol. 90. EDP Sciences, 01004.

[32] Abram J Knarr. 2018. *Do Peripheral HUD Warnings Affect Driving Ability?* California State University, Long Beach.

[33] Jeamin Koo, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry Leifer, and Clifford Nass. 2015. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 9, 4 (2015), 269–275.

[34] Yi-Ching Lee, John D Lee, and Linda Ng Boyle. 2007. Visual attention in driving: The effects of cognitive load and visual disruption. *Human Factors* 49, 4 (2007), 721–733.

[35] Yung-Ching Liu and Ming-Hui Wen. 2004. Comparison of Head-Up Display (HUD) vs. Head-Down Display (HDD): Driving Performance of Commercial Vehicle Operators in Taiwan. *International Journal of Human-Computer Studies* 61, 5 (2004), 679–697.

[36] Zhenji Lu, Xander Coster, and Joost De Winter. 2017. How much time do drivers need to obtain situation awareness? A laboratory-based study of automated driving. *Applied ergonomics* 60 (2017), 293–304.

[37] Zhenji Lu, Riender Happee, and Joost CF de Winter. 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 (2020), 211–225.

[38] Ruiqi Ma and David B Kaber. 2007. Situation awareness and driving performance in a simulated navigation task. *Ergonomics* 50, 8 (2007), 1351–1364.

[39] Andrew K Mackenzie and Julie M Harris. 2015. Eye movements and hazard perception in active and passive driving. *Visual cognition* 23, 6 (2015), 736–757.

[40] William L Mihal and Gerald V Barrett. 1976. Individual differences in perceptual information processing and their relation to automobile accident involvement. *Journal of Applied Psychology* 61, 2 (1976), 229.

[41] Ronald R Mourant and Thomas H Rockwell. 1970. Mapping eye-movement patterns to the visual scene in driving: An exploratory study. *Human factors* 12, 1 (1970), 81–87.

[42] Sohad Murrar and Markus Brauer. 2018. The SAGE encyclopedia of educational research, measurement, and evaluation. *Mixed model analysis of variance* (2018), 1075–1078.

[43] Larry B Nabatilan, Fereydoun Aghazadeh, Ashish D Nimbarte, Craig C Harvey, and Suman K Chowdhury. 2012. Effect of driving experience on visual behavior and driving performance under different driving conditions. *Cognition, technology & work* 14, 4 (2012), 355–363.

[44] Richard E Nisbett and Timothy D Wilson. 1977. Telling more than we can know: verbal reports on mental processes. *Psychological review* 84, 3 (1977), 231.

[45] Hye Sun Park, Min Woo Park, Kwang Hee Won, Kyong-Ho Kim, and Soon Ki Jung. 2013. In-Vehicle AR-HUD System to Provide Driving-Safety Information. *ETRI Journal* 35, 6 (2013), 1038–1047.

[46] Juhee Park and Woojin Park. 2019. Functional requirements of automotive head-up displays: A systematic review of literature from 1994 to present. *Applied Ergonomics* 76 (2019), 130–146.

[47] So Yeon Park, Dylan James Moore, and David Sirkin. 2020. What a Driver Wants: User Preferences in Semi-Autonomous Vehicle Decision-Making. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[48] Fernando M Poó, Orit Taubman-Ben-Ari, Rubén D Ledesma, and Carlos M Díaz-Lázaro. 2013. Reliability and validity of a Spanish-language version of the multidimensional driving style inventory. *Transportation research part F: traffic psychology and behaviour* 17 (2013), 75–87.

[49] R Core Team. 2013. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

[50] James Reason, Antony Manstead, Stephen Stradling, James Baxter, and Karen Campbell. 1990. Errors and violations on the roads: a real distinction? *Ergonomics* 33, 10-11 (1990), 1315–1332.

[51] Miguel A Recarte and Luis M Nunes. 2000. Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of experimental psychology: Applied* 6, 1 (2000), 31.

[52] RStudio Team. 2020. *RStudio: Integrated Development Environment for R*. RStudio, PBC, Boston, MA.

[53] Fridulv Sagberg, Selpi, Giulio Francesco Bianchi Piccinini, and Johan Engström. 2015. A review of research on driving styles and road safety. *Human factors* 57, 7 (2015), 1248–1275.

[54] Paul Salmon, Neville Stanton, Guy Walker, and Damian Green. 2006. Situation awareness measurement: A review of applicability for C4i environments. *Applied ergonomics* 37, 2 (2006), 225–238.

[55] Gerald E Schneider. 1967. Contrasting visuomotor functions of tectum and cortex in the golden hamster. *Psychologische Forschung* 31, 1 (1967), 52–62.

[56] Stephen J Selcon and RM Taylor. 1990. Evaluation of the Situational Awareness Rating Technique(SART) as a tool for aircrew systems design. *AGARD, Situational Awareness in Aerospace Operations 8 p(SEE N 90-28972 23-53)* (1990).

[57] Jami Shah, Noe Vargas-Hernandez, Joshua Summers, and Santosh Kulkarni. 2001. Collaborative Sketching (C-Sketch) An Idea Generation Technique for Engineering Design. *Journal of Creative Behavior* 35, 3 (2001), 168–198.

[58] Qonita Shahab, Jacques Terken, and Berry Eggen. 2013. Development of a questionnaire for identifying driver's personal values in driving. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 202–208.

[59] Yumiko Shinohara, Rebecca Currano, Wendy Ju, and Yukiko Nishizaki. 2017. Visual attention during simulated autonomous driving in the US and Japan. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 144–153.

[60] David Sirkin, Nikolas Martelaro, Mishel Johns, and Wendy Ju. 2017. Toward measurement of situation awareness in autonomous vehicles. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 405–415.

[61] M. W. Smolensky. 1993. Toward the physiological measurement of situation awareness The case for eye movement measurements. In *Proceedings of the Human Factors and Ergonomics Society 37th annual meeting*, Vol. 23.

[62] Orit Taubman-Ben-Ari, Mario Mikulincer, and Omri Gillath. 2004. The multidimensional driving style inventory—scale construct and validation. *Accident Analysis & Prevention* 36, 3 (2004), 323–332.

[63] Orit Taubman-Ben-Ari and Dalia Yehiel. 2012. Driving styles and their associations with personality and motivation. *Accident Analysis & Prevention* 45 (2012), 416–422.

[64] Richard M Taylor. 2017. Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational awareness*. Routledge, 111–128.

[65] Geoffrey Underwood, Athy Ngai, and Jean Underwood. 2013. Driving experience and situation awareness in hazard detection. *Safety science* 56 (2013), 29–35.

[66] Hanneke Hooft van Huysduynen, Jacques Terken, Jean-Bernard Martens, and Berry Eggen. 2015. Measuring driving styles: a validation of the multidimensional driving style inventory. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 257–264.

[67] Yuchen Wang, Weina Qu, Yan Ge, Xianghong Sun, and Kan Zhang. 2018. Effect of personality traits on driving style: Psychometric adaption of the multidimensional driving style inventory in a Chinese sample. *PLoS one* 13, 9 (2018), e0202126.

[68] Nicholas J Ward, Andrew Parkes, and Peter R Crone. 1995. Effect of background scene complexity and field dependence on the legibility of head-up displays for automotive applications. *Human factors* 37, 4 (1995), 735–745.

[69] M. Weihrauch, G. Meloeny, and T. Goesch. 1989. The first head up display introduced by General Motors. In *SAE Technical Paper 890288*. SAE International.

[70] Daniel J Weintraub. 1987. HUDs, HMDs and Common Sense Polishing Virtual Images. *Human Factors Society Bulletin* 30, 10 (1987), 1–3.

[71] Robert West, Davina French, Richard Kemp, and James Elander. 1993. Direct observation of driving, self reports of driver behaviour, and accident involvement. *Ergonomics* 36, 5 (1993), 557–567.

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

[73] Earl L Wiener and Renwick E Curry. 1980. Flight-deck automation Promises and problems. *Ergonomics* 23, 10 (1980), 995–1011.

[74] Jiawei Xu, Shigang Yue, Federica Menchinelli, and Kun Guo. 2017. What has been missed for predicting human attention in viewing driving clips? *PeerJ* 5 (2017), e2946.