
NOVEMBER 2022

Shashank Mehrotra, PhD
Meng Wang, MS
Nicholas Wong, BS
Jah’inaya Parker, BS
Shannon C. Roberts, PhD
University of Massachusetts-Amherst

Woon Kim, PhD
Alicia Romo, PhD
William J. Horrey, PhD
AAA Foundation for Traffic Safety
Title


(November 2022)

Authors

Shashank Mehrotra, Meng Wang, Nicholas Wong, Jah' inaya Parker, Shannon C. Roberts
University of Massachusetts–Amherst

Woon Kim, Alicia Romo, William J. Horrey
AAA Foundation for Traffic Safety

Author ORCID Information

Mehrotra:  https://orcid.org/0000-0002-6749-3773
Wang:  https://orcid.org/0000-0002-3304-0610
Wong:  https://orcid.org/0000-0002-1707-9769
Parker:  https://orcid.org/0000-0001-7583-5174
Roberts:  https://orcid.org/0000-0002-0052-7801
Kim:  https://orcid.org/0000-0001-9806-1204
Romo:  https://orcid.org/0000-0002-4553-1556
Horrey:  https://orcid.org/0000-0002-9533-4411

©2022, AAA Foundation for Traffic Safety, SAFER-SIM
As vehicle technology and automation continue to advance, the need to keep drivers engaged and informed of system status and system actions is becoming increasingly important. This is especially true in cases where the driver will assume different responsibilities when using automation or when the driver no longer needs to be actively monitoring the driving environment for extended periods of time. The human–machine interface (HMI) is of critical importance in such use cases.

This technical report summarizes a review of the literature concerning the use of different HMIs in the study of driver performance following requests to intervene/resume vehicle control. It also considers existing guidelines regarding the design or implementation in HMI in automated vehicles. In considering both avenues, a new set of guidelines is proposed. The report should be of interest to researchers, safety advocates, the automobile industry, and government entities.

This report is a product of an active cooperative research program between the AAA Foundation for Traffic Safety and the SAFER-SIM University Transportation Center.

C. Y. David Yang, Ph.D.
Executive Director
AAA Foundation for Traffic Safety

Dawn Marshall
SAFER-SIM Director
National Advanced Driving Simulator
The University of Iowa
About the Sponsors

AAA Foundation for Traffic Safety
607 14th Street, NW, Suite 201
Washington, D.C. 20005
202-638-5944
www.aaafoundation.org

Founded in 1947, the AAA Foundation for Traffic Safety in Washington, D.C., is a nonprofit, publicly supported charitable research and education organization dedicated to saving lives by preventing traffic crashes and reducing injuries when crashes occur. Funding for this report was provided by voluntary contributions from AAA/CAA and their affiliated motor clubs, individual members, AAA-affiliated insurance companies, and other organizations or sources.

This publication is distributed by the AAA Foundation for Traffic Safety at no charge, as a public service. It may not be resold or used for commercial purposes without the explicit permission of the foundation. It may, however, be copied in whole or in part and distributed for free via any medium, provided the Foundation is given appropriate credit as the source of the material. The AAA Foundation for Traffic Safety assumes no liability for the use or misuse of any information, opinions, findings, conclusions, or recommendations contained in this report.

If trade or manufacturer’s names are mentioned, it is only because they are considered essential to the object of this report and their mention should not be construed as an endorsement. The AAA Foundation for Traffic Safety does not endorse products or manufacturers.

SAFER-SIM University Transportation Center
Federal Grant No: 69A3551747131

SAFER-SIM is comprised of a multidisciplinary, synergistic team of researchers in human factors, engineering, computer science, and psychology who will use innovative simulation approaches ranging from microsimulation to human-in-the-loop simulation to promote safety. SAFER-SIM sponsors research, outreach activities in STEM areas, and aids workforce development efforts in transportation safety.

The universities comprising SAFER-SIM study how road users, roadway infrastructure, and new vehicle technologies interact and interface with each other using microsimulation and state-of-the-art driving, bicycling, pedestrian simulators.

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation’s University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.
Table of Contents

Abstract ................................................................................................................................. v

Introduction .......................................................................................................................... 1

Method ................................................................................................................................ 2
  Eligibility Criteria ........................................................................................................... 2
  Information Sources ....................................................................................................... 2
  Search .............................................................................................................................. 2
  Study Selection ............................................................................................................... 3
  Data Collection Process for Information Extraction ..................................................... 5
    Top-Down Guidelines .................................................................................................. 5
    Bottom-Up Guidelines ............................................................................................... 6

Results .................................................................................................................................. 7
  Outcomes from Literature Search .................................................................................. 7
    General Patterns ......................................................................................................... 7
    Study Outcomes ......................................................................................................... 8
  Correspondence with Top-Down Interface Guidelines .................................................. 10
  Bottom-up Recommendations ....................................................................................... 13

Discussion ............................................................................................................................ 15
  HMI Design Recommendations ..................................................................................... 16
    Modality ...................................................................................................................... 16
    Information Content and Control ............................................................................... 17
    Timing and Stages ........................................................................................................ 17
  Limitations and Future Research Needs ......................................................................... 17

References ........................................................................................................................... 18
Abstract

The effectiveness of the human–machine interface (HMI) in a driving automation system is based, in part, on how it issues alerts and requests to the driver—it must quickly and sufficiently orient the driver to the driving task. While much research on the design of HMIs exists, the unique associations and meaningful relationships must considered collectively to offer guidance for vehicle automation. The purpose of the current study was (a) to review and synthesize existing research and guidance on HMI design as it related to request to intervene for driving automation systems and (b) to propose a clear and comprehensive set of recommendations that could inform future system development and implementation. The outcome of the review indicates multimodal alerts are prominent, useful, and effective. Many of the HMIs reported in the literature were similar in that they abided by well-established HMI-design guidelines. A small subset of HMIs established evidence for distinct and innovative design principles. Taken together, this study proposed 10 recommendations for HMI design in driving automation systems, ranging from the utility of auditory versus visual alerts to the timing, sensitivity, and location of alerts.
Introduction

Technologies that allow parts of the driving task to be automated have become more widely available over the past several years. Even entry-level vehicles now offer driver support systems, known as Level 1 automation by the Society of Automotive Engineers (SAE), such as adaptive cruise control (ACC) and lane keeping assistance (LKA) (SAE, 2021). When combined, these systems are classified as Level 2, where both lateral and longitudinal control of the vehicle are automated, but still require the driver to be engaged in the driving task and be ready to take over quickly. In the future, Level 3 systems, where the driver is at times no longer responsible for monitoring the road, will be available on production vehicles, further altering the relationship between the driver and their vehicle.

Vehicle automation, at any level, stands to change the role and responsibilities of drivers using the systems. While these technologies offer safety and convenience to motorists, they could pose a risk of being misused. For example, some drivers were found to be 50% more likely to engage in secondary tasks when using Level 2 automation compared to when they drove without the technology engaged (Dunn, Dingus, Soccolich, 2019). Also, drivers using other forms of driver assistance systems (e.g., lane departure warnings or forward collision warnings) sometimes find them annoying or disturbing to the point where driving performance is negatively impacted (Biondi et al., 2014) or drivers completely ignore them (Dijksterhuis et al., 2012). One way to mitigate this risk is using sensor-based alert systems that monitor the driver and road environment. When these systems detect driver inattention or road conditions that the automation cannot handle, they issue alerts to the driver to return their attention to the road or to take over control of the vehicle. These request to intervene (RTI) or request to monitor the driving environment more closely are an important part of the driver–vehicle human–machine interface (HMI), as they need to orient the driver to the driving task without being too startling, distracting, or irritating—lest the driver decide to ignore them or turn off the technology all together. As different levels of driving automation systems have different expectations for the driver, the RTI and corresponding HMI can have a variety of goals. For example, drivers who are in-the-loop (in physical control of the vehicle and monitoring the driving situation) may need a subtle RTI, drivers who are on-the-loop (not physically controlling the vehicle but monitoring the driving situation) may need a more overt RTI, and drivers who are out-of-the-loop (not monitoring the driving situation) may need a highly explicit RTI (Merat et al., 2019).

There is a large body of knowledge concerning the design and implementation of alerts in many different operational domains (e.g., Pritchett, 2009; Wogalter & Laughery, 1996). These efforts have been distilled into various and general human factors design principles (Edworthy, 1994; Laughery & Wogalter, 2006; Rogers et al., 2000; Stanton, 1994; Wogalter et al., 2002). In the field of surface transportation, researchers have also compiled human factors design guidelines for in-vehicle status indicators, telltales, and other alerts such as collision or lane departure warning systems (Campbell et al., 2007; Hoffman et al., 2005; Horowitz & Dingus, 1992; Lee et al., 1999). While there is much to be learned from such general and domain-specific guidance, it is important to underscore that many of these systems and HMIs preceded the advent of vehicle automation. In recognizing this gap, researchers are beginning to fill this void by offering guidance and recommendations concerning the implementation of effective interfaces in the context of vehicle automation.
(e.g., Bazilinskyy & deWinter, 2015; Naujoks et al., 2019; van den Beukel & van der Voort, 2017).

In the context of vehicle automation, much research has been done on the different modalities and design specifications of these RTI alerts and their accompanying HMIs, with findings that are not always congruent or easy to interpret. As such, the current study seeks to:

1. conduct a systematic literature review concerning the implementation and effectiveness of HMIs in SAE Level 1 and above
   a. compile existing guidelines and recommendations relating to HMIs in automated driving applications (“top-down” approach) and document the extent to which past HMIs used in published research studies have considered or implemented these recommendations;
   b. synthesize this body of research to identify new guidelines (“bottom-up” approach);
2. develop a clear and comprehensive set of HMI design recommendations based on information gleaned from both the bottom-up and top-down approaches.

Method

A literature review of HMIs in driving automation systems was conducted, with a focus on RTI. The resulting articles were organized and categorized into top-down and bottom-up guidelines. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines are a highly cited set of procedures that aim to establish a checklist of essential items to report in systematic reviews and meta-analyses (Moher et al., 2009). This paper is formatted in accordance with these guidelines.

Eligibility Criteria

All articles were accepted for publication in conference proceedings, journals, or technical reports and were published between January 2011 until April 2021. All articles were available in English.

Information Sources

The Web of Science (WoS) and the Transportation Research International Documentation (TRID) databases were targeted for queries. WoS provides access to multiple databases for different academic disciplines and is a renowned source for searching and reviewing research articles across different academic disciplines. TRID is an integrated database that combines the records from multiple sources and provides access to more than 1.3 million records related to transportation research.

Search

The databases were explored using an extensive keyword search across a variety of categories. The search terms were combined using OR operators within each category and
using AND operators across different categories. The search terms associated with each category were as shown in Table 1.

**Table 1. Keywords.**

<table>
<thead>
<tr>
<th>Category</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automation</td>
<td>&quot;ADAS&quot;, &quot;ADF&quot;, &quot;SAE levels&quot;</td>
</tr>
<tr>
<td>Driver state</td>
<td>&quot;Driver State Monitoring&quot;, &quot;Driver support features&quot;, &quot;DSF&quot;</td>
</tr>
<tr>
<td>Automation features</td>
<td>&quot;Adaptive cruise control&quot;, &quot;Lane-keeping assistance&quot;, &quot;Autopilot&quot;</td>
</tr>
<tr>
<td>Human machine interactions</td>
<td>&quot;Human–machine interface&quot;, &quot;Human–machine interaction&quot;, &quot;HMI&quot;,</td>
</tr>
<tr>
<td></td>
<td>&quot;Human–computer–interaction&quot;, &quot;Human Automation interactions&quot;,</td>
</tr>
<tr>
<td></td>
<td>&quot;Interface design&quot;, &quot;Multimodal interface&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;Voice I/O&quot;, &quot;Vibrotactile display&quot;, &quot;Haptic I/O&quot;</td>
</tr>
<tr>
<td>Situational awareness</td>
<td>&quot;Situational Awareness&quot;, &quot;Vigilance&quot;, &quot;Monitoring&quot;, &quot;Supervisory control&quot;,</td>
</tr>
<tr>
<td></td>
<td>&quot;Gaze Coordination&quot;</td>
</tr>
<tr>
<td>Driver response</td>
<td>&quot;Reaction Time&quot;, &quot;Take-over&quot;, &quot;Takeover request&quot;, &quot;Time to Collision&quot;,</td>
</tr>
<tr>
<td></td>
<td>&quot;Control transition&quot;, &quot;Transition time&quot;, &quot;Steering&quot;</td>
</tr>
<tr>
<td>Driver impairment</td>
<td>&quot;Non-driving related task&quot;, &quot;Dual-task performance&quot;, &quot;Distraction&quot;, &quot;Fatigue&quot;,</td>
</tr>
<tr>
<td></td>
<td>&quot;Distracting task&quot;, &quot;Distracted driving&quot;, &quot;Secondary task&quot;</td>
</tr>
</tbody>
</table>

While searching the WoS database, the scope of the topics was further narrowed using pre-existing database categories: "Computer Science Interdisciplinary Applications,” “Transportation,” “Ergonomics,” and “Transportation Science Technology.” Similarly, results from the TRID database were further filtered according to the “Safety and Human Factors” category. The categories and search terms were selected after reviewing with the research team, as well as reviewing a small set of relevant articles to check if those terms were included.

**Study Selection**

Figure 1 illustrates the process of gathering and selecting studies. The search terms resulted in an expansive list of articles. A list of 5,645 unique articles was obtained from the WoS database and a list of 8,254 unique articles was obtained from the TRID database, yielding 13,899 articles in the original set. (If any article appeared in both databases, it was attributed to WoS.)

These articles were screened and filtered by the research team who reviewed the title, and sometimes the abstract, for general relevance. Prior to this, to ensure there was good inter-rater agreement and clarity on the selection criteria, each team member reviewed the same set of 200 articles from WoS and 100 articles from TRID. There was an average of 90.8% agreement for articles from WoS and 87.8% agreement for articles from TRID. The title/abstract review resulted in the selection of 883 articles.
Following the quick filtering of articles based on title (and abstract), a more careful round of review was conducted to further filter the articles. Although the original search was more inclusive (see Search section above), the selection criteria for this round of review were more stringent to better align with the current objectives. The inclusion criteria were as follows:

- Involved a road vehicle application, whether for passenger and commercial vehicles
- Implemented Level 1 automation (i.e., ACC or LKA) or higher
- Implemented an HMI or an alert system to notify the driver
- Measured driver performance or behavior (e.g., situation awareness, takeover time, brake reaction time, or glance behavior)

Only articles that met all criteria were assessed further, yielding a total of 194 articles. Articles that met some but not all the desired criteria (N = 405) were logged for future use, but are not discussed further in this report. Next, the full text articles (N = 194) were downloaded and assessed by a second team member. This two-stage process yielded 88 eligible papers. Articles that did not meet all criteria were excluded (N = 106). Finally, each respective team member examined the reference list of each of the selected articles (N = 88) to identify additional relevant publications. This resulted in the addition of 16 new articles to the review, i.e., 16 new articles that were not identified in the original search. During
this same process, 8 articles were excluded due to duplications or other issues. In total, 96 papers met all the criteria and were reviewed in detail.

**Data Collection Process for Information Extraction**

Team members examined the full articles to extract specific information related to the research objectives. This key information included the following:

- **Purpose**—purpose, goal, or objective for the study
- **Participants**—sample size, including participants’ demographic distribution (if available)
- **Automation System**—automation system, as described in the article (e.g., adaptive cruise control or lane keeping assistance)
- **Automation Level**—level of automation, as defined by the SAE, and stated in or inferred from the article
- **Type of Study**—whether the experiment used a simulator study, naturalistic observations, etc.
- **Interface Modality**—HMI modality (i.e., audio, visual, haptic, or a combination)
- **Independent Variables**—the key manipulations in the experiment (e.g., type of alert/HMIs, type of secondary task, or driver demographics)
- **Dependent Variables**—specific measures gathered in the experiment (e.g., driver performance during takeover, situation awareness, or workload)
- **Results**—findings as to the relationship between the independent and dependent variables
- **Conclusion**—implications for HMI design
- **Comments**—researcher comments and recommendations (e.g., feedback on the type of HMIs or the applicability of the study in the context of this research)

**Top-Down Guidelines**

As noted in the Introduction, some researchers or practitioners have offered some guidelines concerning optimal HMI design in driving automation systems. To leverage these existing sources, guidelines from multiple sources were first compiled in a single list, with revision to remove redundancies (referred to, in this report, as the top-down guidelines). This set of pre-existing guidelines was used to inform the later development of the current recommendations. The list was also used to examine the extent to which interfaces identified in the literature search adhered to these already established guidelines.

Three sets of pre-existing guidelines were identified in the literature search (Bazilinskyy & DeWinter, 2015; Naujoks, Hergeth, et al., 2019; van den Beukel & van der Voort, 2017). These three sets of guidelines were combined and paraphrased, as necessary, to yield a comprehensive list of 17 best practices, shown in Table 2.
Table 2. Top-down guidelines synthesized from existing sources.

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alert for smooth transition</td>
<td>Provide alerts that allow drivers to come back into the control-loop in time without causing startle reactions</td>
</tr>
<tr>
<td>Auditory attention</td>
<td>Use auditory signals as a base attention retrieving signal, especially for urgent situations</td>
</tr>
<tr>
<td>Visually informative</td>
<td>Use visual interfaces to enable more content-rich transfer of information and to allow users their own pace of information retrieval</td>
</tr>
<tr>
<td>Multimodal</td>
<td>Use multimodal interfaces together in a complementary fashion, especially for those with impairments and for urgent situations</td>
</tr>
<tr>
<td>Along the line of sight</td>
<td>Present high-priority information close to the driver’s expected line of sight</td>
</tr>
<tr>
<td>Continuous feedback</td>
<td>Provide continuous feedback and feed forward information on system-state (e.g., providing information on activation, deactivation, availability, and malfunction without causing counterproductive effects, like distraction)</td>
</tr>
<tr>
<td>Appropriate alert intensity</td>
<td>Use signal intensity (e.g., frequency, wavelength, pace, and duration) to indicate perceived urgency, but ensure not to annoy the driver</td>
</tr>
<tr>
<td>Clarity in alert message</td>
<td>Use written words to express different levels of urgency, like “Danger” compared to “Notice”</td>
</tr>
<tr>
<td>Involve driver in the control-loop</td>
<td>Keep operators involved in the control-loop (perception, decision-making, and implementation)</td>
</tr>
<tr>
<td>Alert towards source of danger</td>
<td>Use alerts to orient the user towards the source of danger</td>
</tr>
<tr>
<td>No need for continuous alert monitoring</td>
<td>Ensure that time-critical interactions with the system do not require continuous attention</td>
</tr>
<tr>
<td>Continuous mode display</td>
<td>Display system mode continuously</td>
</tr>
<tr>
<td>Unintentional state change</td>
<td>Minimize the potential for unintentional activation and deactivation</td>
</tr>
<tr>
<td>Standardized symbology</td>
<td>Use commonly accepted or standardized symbols to communicate the automation mode; use of non-standard symbols should be supplemented by additional text explanations or vocal phrases</td>
</tr>
<tr>
<td>Tactile cueing</td>
<td>Use tactile interfaces for cueing distracted drivers’ attention back to the road</td>
</tr>
<tr>
<td>Appropriate mode grouping</td>
<td>Group HMI elements together according to their function to support the perception of mode indicators</td>
</tr>
<tr>
<td>System failure contingency</td>
<td>In case of sensor failures, display their consequences and required operator steps</td>
</tr>
</tbody>
</table>

Bottom-Up Guidelines

First, conclusions, limitations, and reviewer comments from the original 96 studies were compiled (see also section below “Study Outcomes”). These data were used to inform a listing of potential bottom-up guidelines. A single mention of an HMI design suggestion was considered sufficient at the outset, provided it was justified and explainable based on the study’s findings. An example can be seen from Lin et al. (2020) who concluded that visual iconography coupled with audio alerts can elicit appropriate responses from drivers. This was included as a bottom-up recommendation: prioritization of pictographic images. Next, the list was compared to the top-down guidelines to ensure that any guidance derived from the studies was novel or mutually exclusive. For example, Borojeni et al. (2016) concluded that visual alerts in the direction of the obstacle that caused automation to disengage can help drivers perform more effective takeovers; this conclusion is consistent with the top-down recommendation that warns the driver of the source of danger and so was not considered as new guidance. Then, the conclusions, limitations, and reviewer comments were reviewed for duplicates, i.e., if two studies pointed towards the same conclusion, they
were combined into a single guideline. For example, Louw et al. (2017) concluded that early avoidance should be the focus of HMI design. However, Merat et al. (2014) also concluded that HMI messages regarding takeover requests need to be timely and predictable (to the extent possible). Taken together, these conclusions imply that the focus of HMI design should be on pre-empting crashes and allowing for early avoidance. Finally, those conclusions, limitations, and reviewer comments that were neither a top-down recommendation, nor were a duplicate of findings from similar studies were compiled into the final list of bottom-up guidelines.

Results

The results are presented in three different parts. First, the overall outcome of the literature search is described, including general patterns across categories and interface characteristics (e.g., interface modality and levels of automation). Then, study outcomes were grouped based on significant findings commonly referenced across study. Next, HMIs from the literature search were evaluated against the 17 top-down guidelines shown in Table 2 (i.e., from best practices in past literature). Finally, a complementary listing of bottom-up guidelines was identified from the literature search. For example, if several articles included a feature or HMI design element that showed favorable results but was not already covered by the top-down guidelines, it was considered in the final list of HMI design guidelines. Note that, in general, each article describes one HMI, hence we refer HMIs as opposed to articles in the sections below.

Outcomes from Literature Search

Overall, the majority of the 96 HMIs focused on RTI requests. Very few HMIs gave drivers information concerning other details of the automation, such as its status. More specifically, the HMIs included in this review were generally subject to one of two different research objectives—(1) to test the effectiveness of RTIs by examining driver performance in different situations or contexts, and (2) to evaluate an HMI that included takeover requests. Some HMIs considered different aspects of takeover requests, driver states (e.g., distraction or time constraints), driver demographics, and driving scenarios. In many cases, non-driving related tasks were introduced to assess the impact of driver state on the effectiveness of the HMI for takeover requests. The majority of HMIs were evaluated in a driving simulator or laboratory; while some were evaluated in the field or on a test track.

General Patterns

**Modality of the Interfaces.** Most HMI employed or examined in the set of studies combined modalities (e.g., visual and auditory). More specifically, most HMIs used a combination of audio-visual, visual-haptic, audio-visual-haptic interfaces (58%) for RTIs. However, some HMIs only used the visual modality (24%), auditory modality (12%), or haptic modality (4%).

**Levels of Automation.** Most HMIs focused on alerts for a Level 3 automation system (59%). A smaller number of HMIs focused on a Level 1 or Level 2 system (15%).
few HMIs focused on Level 4 and above (10%). In certain HMIs, the level of automation was either not specified or was unclear (16%).

**Independent and Dependent Measures.** In several instances, the differences between two or more HMIs were a focus of assessment, including differences in alert types, alert modalities, and presentation of alerts. Other comparisons relating to the HMI referred to features or characteristics of the interface, such as differences in the timing of takeover requests, wording of the requests, and urgency of the requests, as well as other manipulations of interest, such as the scenarios where the RTI is presented. Other commonly implemented independent variables included level of automation and differences in driver characteristics. For example, driving personality measures like aggression, sensation seeking, and driving behavior were explored to further examine the impact of individual traits or driving style on performance.

Almost all HMIs were assessed using dependent variables like response time. Additionally, scenario-specific driving performance measures like speeding, lane change duration, lane deviation, braking behavior, and offset from lane center were used to evaluate driving behavior. In addition to performance measures, eye-glance behavior, such as anticipatory glances, gaze dispersion, glances at latent hazards, glances towards potential hazards, and overall eye glance behavior were used as dependent variables.

**Study Outcomes**

**Multimodal Alerts.** Multimodal alerts are preferred over unimodal visual or auditory alerts (e.g., Bakshi 2019; Gaspar et al., 2018; Huang, Steele, Zhang, & Pitts, 2019; Yang et al., 2018). Trimodal signals, such as visual-auditory-haptic, are the most effective type of takeover transition alert (e.g., Hock et al., 2016; J. Kim et al., 2018; Yun & Yang, 2020). Multimodal alerts with cues of high urgency resulted in quicker reaction times than a single-alert modality, such as auditory, even when the driver is distracted (e.g., Cortens et al., 2019; Huang et al., 2019; Salminen et al., 2019). Multimodal alerts can be effective in capturing drivers' peripheral attention when they are not looking at the alert signal (e.g., Politis et al., 2017). That said, auditory displays for alerts and multimodal displays (visual + auditory + haptic) have been associated with greater perceived annoyance (e.g., Naujoks et al., 2014; van den Beukel et al., 2016), which needs to be considered in their implementation.

**Visual Displays.** Visual alerts mapped spatially in the direction of the obstacle that caused automation to disengage improved driving performance (e.g., Borojeni et al., 2016). In addition, the visual modality can effectively convey critical information. For example, LED strips presented at the bottom of the windshield that give information about the status and intention of the automation have been shown to support more appropriate and effective takeover maneuvers (e.g., Wright et al., 2017; Wulf et al., 2015). As another example, heads up display alerts can decrease cognitive workload during takeovers as compared to alerts displayed on a mobile device (X. Li et al., 2020).

Visual augmented reality (AR) displays can support takeover under certain conditions, but there are situations where it might degrade performance relative to other display locations (Lindemann et al., 2019). The presentation of RTIs in a skeumorphic interface, where display features mimic their real-world counterpart, such as a map with a
combination of automation capability information, is more effective than an abstract interface (e.g., Brandenburg & Chuang, 2019; Gold et al., 2016).

**Specificity of Takeover Requests.** Before and during a takeover control request, drivers generally preferred an alert that also provided the reason for the impending handover and the takeover request (S. Li et al., 2019; Louw et al., 2015). Providing feedback and information related to transition urgency or the situation around the vehicle has been shown to improve reaction times and better maintain the driver in the loop (Koo et al., 2016; S. Li et al., 2019). Continuous and constant alerts that provide information such as the location of a hazard and the environment ahead have been demonstrated to help drivers detect more threats and increased their awareness of the traffic environment and of potential system failures (Rezvani et al., 2016; Schmidt et al., 2017; Shull et al., 2019; Wandtner et al., 2018). Displays promoting information analysis and decision selection have been shown to outperform displays focused on information acquisition in terms of maneuver success rates (Eriksson & Stanton, 2017). However, affording drivers the ability to override decisions of Level 3 driving has been shown to be an effective approach (e.g., Borojeni et al., 2017; Kamezaki et al., 2019).

In Level 3 driving, contextual haptic cues can be located on the driver’s body or on the driver’s seat to assist in decision making (e.g., Borojeni et al., 2017; Kamezaki et al., 2019). Relatedly, haptic displays that help drivers to be more spatially aware of their surroundings during driver takeover led to faster responses, shorter duration lane changes, and more scans to the rearview mirror during lane change (Pradhan et al., 2019; Tijerina et al., 2016).

Providing an alert that includes information about timing or urgency, such as a “takeover soon” and “takeover now,” have been shown to yield better driving performance than messages with lower urgency (Brandenburg & Roche, 2020). Additionally, providing drivers information about the level of confidence of the automation, reason for takeover requests, and consequences for drivers’ non-responses have demonstrated increased driver situational awareness, improving overall driving experience and increasing compliance to system cues, prompts, and alerts (e.g., Llaneras et al., 2017; Friedrichs et al., 2016; S. Li et al., 2019; Louw & Merat, 2017; Telpaz et al., 2015). Pre-alerts can improve performance on takeover and have led drivers to look more at the road before the handover occurred and to disengage from secondary tasks (Van Der Heiden et al., 2017; Vogelpohl et al., 2018). A compatible alert, such as an initial auditory alert and visual alert followed by a takeover request, was an effective means of warning drivers about an upcoming takeover when they were engaged in a non-driving task (Cohen-Lazry et al., 2019; Fitch et al., 2011).

Predictive HMIs, that display anticipatory information (e.g., future location/proximity to objects based on current trajectories), are another approach that can provide drivers with additional support for situation awareness; however, there are limited evaluations of these HMIs, and these have not shown a significant effect on takeover response or glance behavior (Walch et al., 2015). Though relevant for all HMIs, but especially for predictive HMIs, misleading alerts resulted in slowed visual search for hazard detection, slowed overall driver response, and increased inattentiveness (Ruscio et al., 2015).
Timing of Alerts. Timing of HMI alerts have been shown to be an important influence on takeover quality: for example, lateral accelerations were more pronounced in a time-critical scenario and longer time budgets led to smoother control transitions (Cui et al., 2017; Doubek et al., 2020). Others have shown that the available takeover time affected the driver’s takeover performance more than the urgency of the request (Roche & Brandenburg, 2018).

It is also important to consider the duration of displayed alerts; some studies have found that longer displays of an alert resulted in increased takeover times (Louw, Markkula, et al., 2017; Louw, Merat, et al., 2017).

Correspondence with Top-Down Interface Guidelines

HMIs identified in the literature search were evaluated to determine how many adhered to the top-down interface guidelines. As shown in Table 3, there was significant disparity in the percentage of HMIs that subscribed to the 17 top-down guidelines. Alert for smooth transition was the most common feature in HMIs gathered in the review (82%). Use of the auditory channel to capture attention, providing informative visual information, and using multiple modalities were also common design features. Figure 2 provides a high-level visualization of the co-occurrence of design guidelines within each of the 96 HMIs, illustrating the large degree of variability in terms of number and clusters across HMI. (Note that this is intended only as a coarse representation of the general pattern and not a detailed breakdown of each individual study.)

Table 3: Percentage of HMIs in the literature search that comply with top-down design principles presented in decreasing order by percentage of HMIs

<table>
<thead>
<tr>
<th>Design principle</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alert for smooth transition</td>
<td>82.3</td>
</tr>
<tr>
<td>Auditory attention</td>
<td>76.0</td>
</tr>
<tr>
<td>Visually informative</td>
<td>63.5</td>
</tr>
<tr>
<td>Multimodal</td>
<td>62.5</td>
</tr>
<tr>
<td>Along the line of sight</td>
<td>53.1</td>
</tr>
<tr>
<td>Continuous feedback</td>
<td>51.0</td>
</tr>
<tr>
<td>Appropriate alert intensity</td>
<td>50.0</td>
</tr>
<tr>
<td>Clarity in alert message</td>
<td>49.0</td>
</tr>
<tr>
<td>Involve in control-loop</td>
<td>44.8</td>
</tr>
<tr>
<td>Alert towards source of danger</td>
<td>43.8</td>
</tr>
<tr>
<td>No need for continuous alert monitoring</td>
<td>42.7</td>
</tr>
<tr>
<td>Continuous mode display</td>
<td>41.7</td>
</tr>
<tr>
<td>Unintentional state change</td>
<td>28.1</td>
</tr>
<tr>
<td>Standardized symbology</td>
<td>28.1</td>
</tr>
<tr>
<td>Tactile cueing</td>
<td>20.8</td>
</tr>
<tr>
<td>Appropriate mode grouping</td>
<td>20.8</td>
</tr>
<tr>
<td>System failure contingency</td>
<td>2.1</td>
</tr>
</tbody>
</table>
Interaction Between HMI, Automation Level, and Modality.

<table>
<thead>
<tr>
<th>Smooth transition</th>
<th>Auditory attention</th>
<th>Visually informative</th>
<th>Multimodal</th>
<th>Line of sight</th>
<th>Continuous feedback</th>
<th>Appropriate intensity</th>
<th>Clarity in message</th>
<th>Involve in control-loop</th>
<th>Warning towards source</th>
<th>No continuous monitoring</th>
<th>Continuous mode display</th>
<th>Unintentional state change</th>
<th>Standardized symbology</th>
<th>Tactile cueing</th>
<th>Appropriate grouping</th>
<th>Failure contingency</th>
</tr>
</thead>
<tbody>
<tr>
<td>61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>101</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>103</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>105</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>107</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>109</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>111</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>117</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Distribution of HMIs across the 17 top-down design guidelines. Each row in the figure represents an article/HMI (see References for mapping to article numbers); each column represents a guideline. Dark shaded cells indicate HMI adhered to the guideline. Interaction Between HMI, Automation Level, and Modality.
To further examine how HMI design interacted with other study features, a series of heatmaps were created. The aim of these visualizations was to examine whether patterns of implemented HMI guidelines varied by level of automation or the modality of HMI employed.

Figure 3 shows the heatmap between the level of automation employed in the study and the 17 top-down HMI guidelines. Automation levels, from 1 to 4 are depicted (and the multiple category denotes HMIs that were examined for more than one automation level in a given study). Note that the sequencing of the guidelines is determined by the dendrogram tree shown on the far-right side of the graph—a hierarchical clustering analysis that indicates the relationship between guidelines. That is, if two guidelines are connected in the dendrogram, it indicates that the number of papers that abide by each guideline are related (i.e., they belong to the same cluster).

Across all levels of automation, the most common guidelines are “multimodal,” “visually informative,” “auditory attention,” and “alert for smooth transition”. The guideline, “system failure contingency” was only exhibited in Level 3 automation and is most likely a consequence of Level 3 automation being the only automation level wherein the driver must be receptive to RTIs (i.e., they may not have to supervise the automation). Many guidelines did not appear at all in Level 4 automation, e.g., “unintentional state change” and “appropriate mode grouping,” presumably because Level 4 automation is the only automation level wherein the driver becomes a passenger once the automation is engaged. Interestingly, smooth transition and auditory were more prevalent in Level 3 and Levels 2 and 3, respectively, where drivers might be more out-of-the-loop compared to Level 1 (where drivers are responsible for more aspects of driving) and Level 4 (where drivers are less likely to be required to intervene).

![Figure 3: Heatmap showing the relationship between levels of automation (x-axis) and HMI guidelines (y-axis). A darker color represents a greater number of HMIs falling into that category. The combined automation level indicates that the paper examined multiple levels of automation. The figure also displays the result of hierarchical clustering by using dendrograms.](image-url)
Figure 4 shows the relationship between the interface modality and the HMI guidelines. The combined interface category denotes HMIs with at least two modalities, e.g., the visual and auditory modality.

Similar to the automation levels, across all modalities, the most common guidelines are “multimodal,” “visually informative,” “auditory attention,” and “alert for smooth transition.” Comparing the responses across rows, i.e., examining the pattern for one guideline across the modalities, some logical patterns were evident: that auditory interfaces are most likely to abide by the “auditory attention” guideline, visual interfaces are most likely to abide by the “visually informative” guideline, haptic interfaces are most likely to abide by the “tactile cueing” guidelines, and combined interfaces are most likely to abide by the “multimodal” guideline. In addition, visual interfaces were more likely to be “along the line of sight,” which naturally follows from the connection between modality type and the guideline. Auditory and haptic interfaces are more likely to be associated with unintentional state changes, given their propensity to help drivers’ attention (especially when they might be engaged in other visual tasks). Last, haptic interfaces were more likely to have the “appropriate alert intensity,” “alert towards the source of danger,” and “continuous feedback.”

![Figure 4: Heatmap showing the relationship between interface modality (x-axis) and HMI guidelines (y-axis). A darker color represents a greater number of HMIs falling into that category. The combined interface indicates that the paper examined multimodal interfaces. The figure also displays the result of hierarchical clustering using dendrograms.](image)

**Bottom-up Recommendations**

As discussed in the previous sections, pre-defined HMI guidelines can help assess interface design and provide insights into the strengths and limitations of a given HMI. While top-down guidelines are informative and help understand essential aspects of the HMIs, bottom-up guidelines can identify intrinsic limitations of the HMIs. More specifically, the authors often identify overarching principles (or lack thereof) within the article’s conclusion or discussion section. Additionally, the reviewers often identify common
themes across articles. This section presents a set of guidelines compiled from the HMIs to account for these factors.

As noted previously, an initial listing of potential bottom-up guidelines was distilled from a review of the outcomes, conclusions, and reviewer comments from the 96 studies compiled from the literature search. At this stage, to be more inclusive, guidance did not need to achieve any critical mass for inclusion; a single mention of a design suggestion or a favorable outcome was sufficient. Next, this working list was compared against the top-down guidelines to remove any redundancies (i.e., outcomes that already fall under existing guidance). This process yielded a set of 13 potential guidelines. Duplicates or guidance that shared similarities with other entries were also removed or combined into a single item. Finally, those conclusions, limitations, and reviewer comments that were neither a top-down recommendation, nor were a duplicate were compiled into the final list of bottom-up guidelines. This process of down-selecting yielded a final set of 5 bottom-up guidelines (shown in Table 4).

First, where appropriate, visual interfaces should prioritize pictographic information over text-based messages. Such information should be readily understood by drivers and/or is standardized (where standards exist). There are a number of guidelines specific to pictographic information and icons that can support design of such elements (e.g., Korpi & Ahonen-Rainio, 2015; Collaud et al., 2022), as well as guidance related to warnings and display (e.g., Lehto, 1992; Campbell et al., 2007).

Second, the intensity of an alert should increase as the available time (i.e., response window) decreases. Though related to the top-down guideline of applying an appropriate level of intensity, this considers a dynamic component that varies according to urgency. Third, related but separate, staged or gradient alerts should be used to counter driver non-response to earlier alerts. This includes variations in the delivery method (modality) as well as other features.

Fourth, HMI alerts should be augmented or tailored based on input from driver state monitoring systems (Hecht et al., 2018). For example, the HMIs can monitor the state of the driver (i.e., whether they show signs of distraction or fatigue) to ensure that the response is timely. One such approach could provide periodic attention maintenance alerts throughout the drive to increase the situational awareness during unexpected automation failures.

Finally, feedback about external objects and limitations should be emphasized over general information about system confidence. For example, it has been shown that explicit information about a system’s level of awareness of its environment vis-à-vis its limitations can yield more favorable outcomes compared with generalized information regarding uncertainty (e.g., Rezvani et al., 2016).
Table 4. Percentage of HMIs that comply with bottom-up recommendations.

<table>
<thead>
<tr>
<th>Principle</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>When possible, visual interfaces should prioritize pictographic information over text-based messages.</td>
<td>22.9%</td>
</tr>
<tr>
<td>The intensity of an alert should increase as the available time (i.e., response window) decreases.</td>
<td>13.5%</td>
</tr>
<tr>
<td>Staged alerts should be used to counter driver non-response to earlier alerts.</td>
<td>10.4%</td>
</tr>
<tr>
<td>HMI alerts should be augmented or tailored based on input from driver state monitoring systems.</td>
<td>4.2%</td>
</tr>
<tr>
<td>Feedback about external objects and limitations should be emphasized over general information about system confidence.</td>
<td>2.1%</td>
</tr>
</tbody>
</table>

The frequency with which the new guidelines were observed in the set of 96 HMIs was also examined (and shown in the percentages in Table 4), which is likened to the strength of evidence. Fifty-two (over half of) HMIs met at least one of the bottom-up guidelines. Many of the HMIs met two to three of the bottom-up guidelines whereas only one HMI met four guidelines.

Discussion

Technologies that allow parts of the driving task to be automated have become more widely available over the past several years. While these technologies offer safety and convenience to motorists, they could pose a risk of being misused. Sensor-based alert systems that monitor the driver and road environment are one possible approach to help drivers. For example, these systems may detect driver inattention or conditions that the automation cannot handle. In doing so, they issue alerts to the driver to return their attention to the road or to take over control of the vehicle. As part of the driver–vehicle HMI, these RTIs need to quickly orient the driver to the driving task without being too startling, distracting, or irritating—lest the driver decide to ignore them or turn off the technology all together. While alerts have been the focus of much research and guidance from other domains, it is important to consider how the design and implementation of HMIs in the context of vehicle automation is best achieved. The purpose of the current study was to review and synthesize existing research and guidance on HMIs and driver takeovers in the context of vehicle automation and to propose a clear and comprehensive set of recommendations that could inform future system development and implementation.

Based on the literature search, nearly 100 relevant articles were identified and further examined. The majority of HMIs were evaluated in a driving simulator or laboratory setting and focused on Level 3 automated systems. Not surprisingly, the focus or purpose of these studies varied significantly, yielding a wide array of independent variables under scrutiny, including but not limited to differences in alert types, alert modalities, presentation of alerts, timing of takeover requests, request wording, request urgency, use case scenarios, and various driver characteristics. Likewise, there was some variability in the underlying outcome measures; however, given the original search criterion, these were largely grounded in driving performance or behavioral measures (e.g., eye glance metrics).

In terms of the HMIs, the majority used multimodal approaches where visual information was combined with auditory alerts. Collectively, the literature search yielded
useful insight into the design of HMI, types of information or feedback provided, timing and urgency, and other aspects. As found in past studies, multimodal alerts are generally preferred due to their effectiveness in informing the driver about takeover. Additionally, multimodal alerts with cues of high urgency resulted in quicker reaction times, as compared to a single-alert modality, and capture drivers’ peripheral attention. At the same time, in a few instances, multimodal alerts were a source of perceived annoyance.

In evaluating how firmly the HMIs were grounded in some of the existing top-down guidelines, it was noted that there was significant variability in the proportion of HMIs that adhered to different guidance. This is not an indictment of past work as these guidelines were developed in concert with some of the HMIs, not to mention that research in this area is still emerging. In addition, some HMIs were developed for one specific purpose (e.g., to mitigate distracted driving while using a Level 2 vehicle) and hence, guidelines not relevant to that purpose were not germane. Overall, some guidelines were reflected in many HMIs while others have not been widely studied in this domain—possibly representing areas needing future research. For example, the most common guideline—“Alert for smooth transition”—appeared in over 80% of the HMIs whereas “System failure contingency” only appeared in 2% of the HMIs, implying that is an area ripe for future and sustained research. It must be acknowledged that guidance, especially for the less prevalent guidelines, may reflect good practices based on work in other domains or based on fundamental design and human factors principles. It follows that the strength of evidence varies across guidelines, even though many in practice should lead to more effective HMIs.

There was also a lot of variability in the number of different guidelines adhered to in a single study (e.g., see Table 3). More research in this area could help evaluate how combinations of guidance can help improve safety, performance, and user experience. Such work can also help prioritize different features and expose areas or certain display configurations where guidance might not work or interact as effectively (e.g., some design approaches would render the HMI incompatible with some of the guidance).

HMI Design Recommendations

Based on the review, a revised and final set of recommendations was distilled borrowing from past guidance as well as from new information gleaned from studies identified in the review (i.e., a more comprehensive merging of the top-down and bottom-up guidance). In doing so, attempts were made to use clear, concise wording to ease comprehension. The recommendations have been arrayed under broader, non-mutually exclusive categories. In some cases, specific guidelines from other sources have been combined. For example, in leveraging the visual modality to convey information to drivers, several different approaches or sub-items are presented (Bazilinskyy & DeWinter, 2015; Naujoks, Wiedemann, et al., 2019; van den Beukel & van der Voort, 2017).

Modality

1. Systems should be designed to be multimodal.
2. Visual displays should be used to support continuous status information, as well as other content-rich information. Visual interfaces should also prioritize pictographic
information and standardized symbology over text-based messages. Text should be used to supplement non-standard symbols, preferably in non-time critical situations.

3. Auditory and/or tactile displays should complement visual information and be used to help reorient driver attention in critical situations. Sustained attention to HMI should not be required in time critical situations.

**Information Content and Control**

1. System status (e.g., on, off, activation, deactivation/disengagement, availability) should be presented clearly and continuously. In doing so, display elements for a common system should be grouped together.

2. Alerts should help orient the driver to the source of danger or provide some information about the traffic context. High priority information should be presented close to the driver’s line of sight.

3. Feedback about RTIs or failure modes should be provided. For example, providing reasons for a takeover or conveying information about system limitations in situ.

4. Driver decision-making and responses should be supported by providing them with information to support situation awareness and required actions. Systems should be designed to minimize or prevent unintentional actions (e.g., accidental activation or deactivation of the system).

**Timing and Stages**

1. Alert should give sufficient time to drivers to regain control safely and effectively.

2. The intensity of the alert should reflect the urgency of the situation, without being a hindrance, distraction, or annoyance to driver. The intensity of an alert should increase as the available time (i.e., response window) decreases.

3. Gradient or multi-staged alerts (e.g., first visual, then auditory) should be used to help convey urgency and to counter non-responses to earlier alerts.

**Limitations and Future Research Needs**

While the review has highlighted important guidelines for HMI development, there remain areas for future research. First, some HMIs compiled in the review were evaluated with small samples and need to be validated with larger samples to ensure generalizability. Second, very few HMIs were evaluated on different data sources (e.g., physiological data, driving data, and survey data) to understand the efficacy of the HMI in takeover performance. Future research should consider a multi-faceted approach in terms of outcome measures. Relatedly, many HMIs appear to be generated by research teams in support of a research question or utilize features that have been incorporated into driving simulator software packages. As such, there are potential gaps between some HMIs examined in the context of these studies and existing HMIs in OEM production vehicles, which would provide an understanding of how partial or conditional automation can be implemented within the limitations of actual current systems. Third, while HMIs were categorized according to modality and level of automation, another potentially more useful characterization is level of engagement by the driver. Such a categorization could differentiate HMIs based on urgency of actions or cognitive demand (remembering, understanding, evaluating, etc.) and then relate level of engagement to the guidelines (e.g.,
count the number of HMIs that require high driver engagement and abide by the “along the line of sight” guideline. Fourth, a meta-analysis would provide quantitative (and complementary) evidence to bolster the findings of the literature review, subject to data availability. A more precise mapping of different guidelines to safety and performance outcomes would be helpful in prioritizing design elements. Ideally, such an effort would also allow for a more careful delineation of HMI design approaches across different levels of automation. Last, guidance will continue to evolve as technologies and driver responsibilities change. As such, reviews such as these must be conducted on a regular basis and/or be updated as necessary.

Though germane to the current discussion, the current review intentionally did not focus on a large body of research specific to driver state monitoring systems: a system that assesses if a driver is capable of safely completing a task as monitored through their physiological state and driving behavior (Guettas et al., 2019), except in cases where these systems were used in the context of alerting or an HMI. Driver state monitoring is an integral feature of some automation-oriented HMI, though other systems do not necessarily rely on inputs from driver state monitoring. From the review, there were many HMIs focused on driver state monitoring: many were focused on the underlying data elements, such as driver (e.g., physiological information) or vehicle based (e.g., steering/pedal inputs), or on the algorithms that are used to generate predictions about the driver state. Driver state monitoring does not supplant the need for good and thoughtful design of HMI and in-vehicle alerts, but it may have an important role in the implementation, utility, and acceptance of these features. For example, a system designed to reorient the drivers’ attention to the forward roadway might have fewer false positives if it considers information about the driver’s point of regard, compared to a system that uses other (non-driver state monitoring) inputs. Future research should consider the integration of driver state monitoring systems into HMIs and how the thoughtful design of an HMI can lead to a successful driver–automation partnership.

Finally, work is currently underway that leverages the current review of design guidelines. This work examines the effects of different HMI configurations, based on their implementation of different guidelines (i.e., basic HMI, which incorporates few guidelines versus an enhanced HMI, which incorporates many), on various outcome measures.

References

* Denotes article identified in bottom-up literature search. Number in parentheses maps onto identifier in Figure 2.


