A Design Space for Human Sensor and Actuator Focused In-Vehicle Interaction Based on a Systematic Literature Review

PASCAL JANSEN, Institute of Media Informatics, Ulm University, Germany
MARK COLLEY, Institute of Media Informatics, Ulm University, Germany
ENRICO RUKZIO, Institute of Media Informatics, Ulm University, Germany

Automotive user interfaces constantly change due to increasing automation, novel features, additional applications, and user demands. While in-vehicle interaction can utilize numerous promising modalities, no existing overview includes an extensive set of human sensors and actuators and interaction locations throughout the vehicle interior. We conducted a systematic literature review of 327 publications leading to a design space for in-vehicle interaction that outlines existing and lack of work regarding input and output modalities, locations, and multimodal interaction. To investigate user acceptance of possible modalities and locations inferred from existing work and gaps unveiled in our design space, we conducted an online study (N=48). The study revealed users’ general acceptance of novel modalities (e.g., brain or thermal activity) and interaction with locations other than the front (e.g., seat or table). Our work helps practitioners evaluate key design decisions, exploit trends, and explore new areas in the domain of in-vehicle interaction.

CCS Concepts: • General and reference → Surveys and overviews; • Human-centered computing → HCI theory, concepts and models; Empirical studies in HCI.

Additional Key Words and Phrases: systematic literature review; design space; in-vehicle interaction; human sensors and actuators

Authors’ addresses: Pascal Jansen, pascal.jansen@uni-ulm.de, Institute of Media Informatics, Ulm University, Ulm, Germany; Mark Colley, mark.colley@uni-ulm.de, Institute of Media Informatics, Ulm University, Ulm, Germany; Enrico Rukzio, enrico.rukzio@uni-ulm.de, Institute of Media Informatics, Ulm University, Ulm, Germany.
1 INTRODUCTION

With the increasing integration of automation technology into vehicle systems, the scope of in-vehicle interaction is getting broader. According to the Society of Automotive Engineers (SAE) taxonomy J3016 [324], there are six levels of driving automation, ranging from level 0 (no driving automation) to level 5 (full driving automation) in the context of motor vehicles and their operation on roadways. With automated vehicles (AVs) (SAE levels 3-5) changing the role of the driver, automotive user interfaces (UIs) undergo a paradigm shift [80] and the vehicle transforms into a mobile office or living space [167]. Accordingly, the driver can perform non-driving related tasks (NDRTs) [82, 396], such as working, using the smartphone [77], or gaming in virtual reality (VR) [249]. Hence, the interior design of AVs will adapt, and new input and output locations emerge that are anchored (e.g., door, table, or seat) or nomadic (e.g., handheld or wearable), see Figure 1. For example, future AVs may consist of a 4-seat configuration, where passengers face each other and, like passengers of non-AVs, benefit from UIs located throughout the interior, e.g., rear-seat entertainment [78, 129]. As the interior is a closed space surrounding the passengers, all human senses could be stimulated by output modalities and, to some extent, used as input.

We consider input and output from the human perspective (see Figure 2), i.e., human actuators (e.g., mouth or skin) intentionally generate explicit input modalities (e.g., speech) or subconsciously produce implicit input modalities (e.g., electrodermal activity (EDA)) sensed by vehicle sensors (e.g., microphone), and vehicle actuators (e.g., speaker) generate output modalities (e.g., sound) perceived by human sensors (e.g., ears).

Despite the large body of works, concepts, and prototypes regarding in-vehicle input and output, current automotive UI research does not or only partly consider the full range of input and output modalities (which also contains, e.g., vestibular stimuli, EDA, gustatory stimuli, brain, or heart activity) and novel vehicle interior locations (e.g., rear, floor, or ceiling). Thus, a new perspective on the in-vehicle interaction space is required, unconstrained from a front-focused design and including such modalities. Besides, it is partly unknown what modalities and interior locations were already considered in previous works concerning vehicles of any SAE level (0-5). For manual or assisted driving (SAE 0-2), knowledge about possible input and output modalities and their placement may help in designing interactions with minimal driver distraction [299] and workload levels (physical, visual, and mental) [53]. In the context of AVs (SAE 3-5), there exist human factor issues, such as mistrust [107], loss of control [109], or safety concerns [333]. Therefore, it is essential to design in-vehicle interactions that will be accepted [80]. Besides, passengers of AVs can perform NDRTs while interacting with modalities/locations that were previously impractical or dangerous regarding the driving task, e.g., due to sensory overload or reduced takeover readiness. Still, the usability of such modalities/locations (e.g., swivel seats, or VR) in an AV context is underexplored.

To investigate these problems, we defined the following research questions (RQs):

RQ 1: How does current automotive UI research leverage human sensors and actuators for in-vehicle interaction?

RQ 2: What vehicle interior locations can be utilized for in-vehicle interaction?

RQ 3: What is the design space for in-vehicle interaction, including an extensive set of human sensors and actuators?

RQ 4: How do users perceive the usefulness, real-world usage, and comfort of in-vehicle modalities and locations?

To answer RQs 1-3, we conducted a systematic literature review (SLR). We gathered a set of keywords for in-vehicle interaction to define the search query for the SLR, which was based on the PRISMA guidelines [255, 279]. Our SLR considered vehicles of any SAE level (0-5). We selected the databases ACM Digital Library (DL) [20],
IEEE Xplore [161], and ScienceDirect [97], to search for relevant publications, resulting in an initial set of 2534 publications. After an abstract screening and a subsequent full-text screening, we considered 327 publications relevant for the synthesis to answer RQs 1-3.

Our SLR shows which approaches for in-vehicle interaction modalities were used by the included publications. They primarily used visual, auditory, and tactile input and output modalities, while few considered (novel) modalities, such as electrodermal, thermal, olfactory, or cerebral. Furthermore, our proposed combination matrix for multimodal interaction reveals that most publications utilized visual modalities for multimodal input and output, e.g., in combination with auditory, kinesthetic, or tactile modalities. Gaps regarding multimodal input containing olfactory modalities and output containing vestibular, electrodermal, or gustatory modalities highlight future multimodal interaction research opportunities. We then present a design space for in-vehicle interaction that extends [177] regarding the set of input and output modalities, interior locations, and involved human sensors and actuators. The design space reveals little to no utilization of thermal, olfactory, gustatory, cerebral, and cardiac input modalities, only a few approaches for vestibular, kinesthetic, and thermal output, and none for electrodermal and gustatory output. Our design space further shows that publications mainly used the front as output location, e.g., for displays [50] or vibration [376], while other locations are not frequently considered, e.g., table, door, rear, floor, or ceiling. To answer RQ 4, we assessed the feasibility of possible in-vehicle interaction modalities and locations in an online user study (N=48) by presenting concept images deduced from related work and gaps in our design space (see Figure 4 and Figure 5). The study results reveal that input modalities were more accepted regarding usefulness, usage, and comfort than output modalities. Besides, well-established input and output modalities in current vehicles, such as auditory or tactile, were generally perceived as more acceptable. However, novel input and output modalities, e.g., vestibular stimuli, were also perceived as useful. While participants perceived interaction in some nomadic and anchored interior locations as useful, e.g., handheld, wearable, rear, seat, or table, they deemed other locations less useful, e.g., ceiling, floor, door, or VR. The results highlight the importance of considering novel modalities and locations in future in-vehicle interaction design.

Contribution Statement: First, we report the results of an SLR on in-vehicle UI research and the analysis of multimodal interaction and utilization of interior locations, leading to a combination matrix for multimodal in-vehicle interaction and visualization of interaction locations accompanied by a self-developed interactive website¹. Second, we propose a design space for in-vehicle interaction, considering several novel vehicle interior locations and including an extensive set of human sensors and actuators. Third, we provide the results of an image-based online user study (N=48) on perceived usefulness, real-world usage, and comfort of possible in-vehicle interaction approaches deduced from our design space and related work and discuss implications for future interaction design.

2 BACKGROUND AND RELATED WORK

Our work is grounded on (1) a definition of the in-vehicle interaction scheme based on a classification of human sensors and actuators and (2) previous literature reviews on automotive UIs.

2.1 Human Sensors and Actuators

Based on [33, 41, 165, 340], we distinguish seven human sensor/actuator categories: (1) visual, (2) auditory, (3) haptic, (4) olfactory, (5) gustatory, (6) cerebral, and (7) cardiac. We follow the definition of sensors and actuators and the respective interaction scheme shown in Figure 2. The in-vehicle interaction scheme includes two agents: human and vehicle, and describes an input-output feedback loop between both agents [340]. A human/the human body intentionally or subconsciously uses actuators such as fingers, brain, or heart to generate

¹https://in-vehicle-interaction-design-space.onrender.com/ Interactive tool to support investigation of in-vehicle interaction research and design.
input modalities. In this work, intentionally performed input is considered as *explicit* (e.g., touch, or speech) and subconsciously performed input as *implicit* (e.g., brain or heart activity). The input modalities are sensed by specific vehicle sensors (e.g., microphones or touchscreens) at vehicle input locations that are either nomadic (e.g., VR, handheld, or wearable) or anchored throughout the interior (e.g., front, seat, door, or rear). Vehicles use actuator devices, such as screens, speakers, or vibration motors, at nomadic or anchored output locations to generate output modalities (e.g., display, sound, or vibration). The output modalities are sensed by specific human sensors (e.g., eye, ear, or skin). In this work, the sensor/actuator categories are used to categorize input and output modalities.

Fig. 2. In-vehicle interaction scheme including two agents: a human and a vehicle. A feedback loop is created between both agents, which sense the external world using natural (human) and artificial (vehicle) sensors. Both human and vehicle act upon the environment with their actuators. One possible example is shown after each sensor, actuator, input modality, and output modality.

The *visual* category entails the eye as a sensor enabling passengers to perceive any light-based output modalities and as an actuator that produces explicit input modalities, e.g., gaze, pupil dilation, or blink rate [340]. *Auditory* output modalities are sensed by the ears, while auditory actuators are body parts that can explicitly generate sounds, e.g., mouth for voice or hands/fingers for clapping [111]. According to Benyon [33] and the ISO standard 9241-910 [272], we divide *haptic* into kinesthetic, cutaneous, and vestibular. We did not consider proprioception, which is the sense of one’s body position and movement [272] as such sensation is already covered by kinesthetic and vestibular sensors/actuators [165]. *Kinesthetic* sensors in the human joints and muscles detect body motion, while kinesthetic actuators generate explicit output modalities, such as muscle activity or body movement. Such body activity can also be implicitly performed. *Cutaneous* is subdivided into electrodermal, tactile, thermal, and pain, which each have specific skin sensors to perceive output modalities such as pressure, temperature, or pain stimuli. Cutaneous actuators generate skin-related input modalities explicitly via touch and implicitly via EDA or skin temperature. The *vestibular* category is adapted from [33] and describes a sensor that detects balance and general body motion. However, as the vestibular system is a passive sensor and active body motion is a kinesthetic actuator, there is no vestibular actuator. Similar to [33, 340], we include the *olfactory* and *gustatory* categories, which each have a dedicated sensor organ (nose and tongue). However, olfactory actuators are any source of body scent, and gustatory actuators are any source of body flavor, e.g., sweat taste. Olfactory and gustatory actuators implicitly produce input. Besides, we include the *cerebral* and *cardiac* categories, similar to [25, 364]. We consider brain activity measured by, e.g., functional near-infrared
spectroscopy (fNIRS) or electroencephalography (EEG) as implicit input modality and brain stimuli as output modalities "sensed" by the brain. Likewise, heart activity (e.g., heart rate) is an implicit input modality, and heart stimuli (e.g., via a defibrillator) are output modalities "sensed" by the heart.

2.2 Literature Reviews on Automotive User Interfaces

There are several reviews on human-vehicle interaction and considerations for the design of automotive UIs. For example, Lee [216] analyzed 50 years of driving safety research, including an overview of vehicle technology, and Akamatsu et al. [8] presented a detailed description of the history of vehicle UIs and related human factors. Still, these papers give little information on higher automation levels (SAE levels 3, 4, and 5) and future in-vehicle interaction regarding novel modalities. A review including AVs was conducted by Kun et al. [197] who identified problem fields for automotive research regarding the transition to higher automation levels. Similarly, Ayoub et al. [22] identified various trends, e.g., the transition towards AVs or the increasing relevance of NDRTs. They also summarized a broad range of input and output modalities, including (novel) approaches like augmented reality (AR), VR, or emotion recognition. An overview of technologies that are being used or developed to perceive user’s intentions for natural and intuitive in-vehicle interaction was presented by Murali et al. [261]. They found that novel multimodal sensing devices replace legacy display interfaces and haptic devices such as buttons and knobs. However, their overview was not based on an SLR, and they did not consider some sensor/actuator categories, e.g., vestibular, olfactory, and gustatory. Besides, there are reviews regarding human factor-related issues (e.g., distraction, awareness, trust, or acceptance) [76, 164] and technical challenges [157]. The first design space for driver-based automotive UIs was introduced by Kern and Schmidt [177] that describes in-vehicle input and output modalities concerning their location in the interior. However, since driving automation was not yet an omnipresent research topic at the time of publication (2009), their design space focused on a subset of possible in-vehicle modalities (i.e., visual, auditory, and haptic) and locations (i.e., subdivisions of front). In a later work (2021), Detjen et al. [80] discussed the requirements and challenges of interaction with AVs regarding users’ acceptance, namely, security & privacy, trust & transparency, safety & performance, competence & control, and positive experiences. They also classified current in-vehicle interaction literature by their contribution to one of the acceptance challenges and used interaction modality. However, they did not consider novel in-vehicle locations and mainly focused on SAE 3-5 vehicles.

In combination, these works already furnish the direction for future in-vehicle interaction. However, they are limited due to not including AVs (e.g., [8, 216]), novel interior locations (e.g., [80]), or considering a subset of possible human sensors and actuators (e.g., [177, 261]). In this work, we include in our SLR any SAE level (i.e., from manual to highly automated driving) and approaches for input and output modalities while considering an extensive set of human sensors and actuators. Besides, we propose a comprehensive design space that extends previous design spaces (such as [80, 177]) and includes not only the driver but also other passengers as users.

3 SYSTEMATIC LITERATURE REVIEW ON IN-VEHICLE INTERACTION

To answer RQ 1, “How does current automotive UI research leverage human sensors and actuators for in-vehicle interaction?” and RQ 2, “What vehicle interior locations can be utilized for in-vehicle interaction?”, our goal was to elaborate a detailed and comprehensive overview of research on in-vehicle interaction. Therefore, we employed an SLR. The process of this SLR is based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) proposed by Moher et al. [254] and Page et al. [279]. Our multistaged process is depicted in Figure 3 and consists of an identification step, a two-step publication screening part, and a synthesis.
3.1 Identification

To identify publications relevant to our topic, we created a database query using Boolean algebra that combines the keywords of each sensor/actuator category specified in a preparatory step (see Table 4 in Appendix). However, as the used database search engines have keyword limits, we split the query into 12 sub-queries, one for each sensor/actuator category. For readability, the complete list of used queries is shown in the Appendix (see Table 5).

The query for the auditory category is shown as example:

\[(\text{vehicle}^* \text{ OR car}^* \text{ OR driver}^* \text{ OR driving OR "in-vehicle") AND (interaction}^* \text{ OR interface}'^*\text{) AND (auditory OR audio OR ear'}^* \text{ OR speech OR voice OR verbal OR vocal OR "non-speech" OR sound}'^* \text{ OR whistling OR humming OR acoustic OR earcon}'^* \text{ OR conversational OR music)}\]

We adapted each query to suit the respective search syntax required by the databases (see Supplementary material). Although the queries were applied independently, the screening process is uniform for all results to prevent differing analyses. Thus, for example, a publication found in the searches for visual and kinesthetic modalities is coded for both. If queries yielded identical results, the duplicates were omitted.

We selected the databases ACM DL, IEEE Xplore, and ScienceDirect, covering venues relevant for automotive UI and future mobility research, e.g., Conference on Human Factors in Computing Systems (CHI), Conference on Automotive User Interfaces (AutoUI), or Transportation Research Part F. Within these databases, we selected a subset of venues (see Table 6) based on the 16 most cited HCI venues according to Google Scholar [121] and an additional ten venues relevant for automotive research. Besides, we defined three search criteria that were applied directly in the database search: A publication had to be (1) published in the past ten years (2011-2021), (2) written in English, and (3) peer-reviewed. Only publication titles and abstracts were searched, similar to [29, 280, 409], as those parts should contain an overview of all relevant content.

The three databases were queried, and the search results were downloaded on 07/15/2021. The detailed result counts for each sensor/actuator category are shown in Table 5. After removing duplicates with Zotero [74], 2534 publications remained (1244 ACM DL (49.1 %), 1046 IEEE Xplore (41.3 %), 244 ScienceDirect (9.6 %)).
3.2 Screening

For the publications screening, we employed a two-phase process, consisting of an abstract screening and a subsequent full-text screening.

3.2.1 Abstract Screening. We defined a mandatory inclusion criterion: in-vehicle interaction. The criterion was true when the publication described or employed an interaction or interface between a user and an in-vehicle system in the automotive context. This includes the use of any input and output modality. Besides, we defined inclusion criteria that were not explicitly coded to save time but always considered in the decision: (1) The publication should be no literature review, taxonomy, or design space. (2) It considers a vehicle of any SAE J3016 [324] automation level. (3) It describes driver/passenger-vehicle interaction, i.e., no direct passenger-passenger interaction. (4) It considers user interaction inside a vehicle, i.e., no external communication (e.g., [72]), vehicle-pedestrian interaction, or teleoperation.

Sysrev [162] was used for the screening phase, as it is an online platform that offers tools for collaborative screening of large amounts of publications. Besides, it provides automatic randomization of the publications to prevent a biased review order. The Sysrev project is public² to allow researchers to view our work and enable the use in other research projects. Two of the authors conducted the abstract screening. To create a shared understanding of the inclusion criteria, the first 25 publication abstracts were collaboratively reviewed. Both reviewers then reviewed a subset of 1500 publication abstracts without verbal discussion, and the remaining 1034 publications were rated by one reviewer. To assess inter-rater agreement, we calculated Cohen’s Kappa [67] \(\kappa = .83 (p_0 = 94.00\%)\), which indicates an almost perfect agreement [202]. Ninety-four conflicts were discussed and resolved.

In total, 439 publications were included. Based on the results, we defined a code book for the full-text screening, representing our topics of interest for the synthesis. The code book included the vehicle’s SAE level considered in the concept/user study, the sensor/actuator category of the employed input/output modality, multimodal input/output usage, and the interaction location. We recognize an interaction as multimodal when the user is provided with multiple modes for interacting with the system (e.g., speech and touch) [364]. This includes both sequential, where a user will have to switch between modes of interaction and simultaneous multimodal interaction that allows multiple modes at a time [296]. For the interaction locations, we defined two overarching types: (1) nomadic (e.g., head-mounted display (HMD), smartphone, or wearable) and (2) anchored (e.g., dashboard, seat, or door). The abstract screening data set containing all included publications’ metadata is provided as supplementary material, enabling researchers to access our raw data for future projects.

3.2.2 Full-Text Screening. We excluded publications in this screening step if at least one of the following exclusion criteria are met: (1) The publication considers driver monitoring intended only to measure participants’ performance, e.g., cognitive load, drowsiness, or distraction. (2) It considers a haptic torque steering support, e.g., for takeover situations. (3) It does not meet the inclusion criteria defined in the earlier process steps (see Section 3.1 and 3.2.1).

The two reviewers continued with this full-text screening, which was again conducted in Sysrev and made public³. In the beginning, 27 publications were rated together to generate a common understanding and test the applicability of the defined code book. Six of these publications resulted in a conflict and were resolved in a discussion. The inter-rater agreement was qualitatively checked in a discussion similar to [315, 365]. The reviewers rated the remaining 412 publications individually.

²https://sysrev.com/u/5286/p/77887 | This project contains our abstract screening. It provides public access to all 2534 screened publications’ metadata, abstracts, and shows our labels. Labeling statistics can be shown, and filters can be used to select publications.

³https://sysrev.com/u/5286/p/78991 | This project contains our full-text screening. It provides public access to all 439 screened publications’ labels.

3.3 Results Overview

In total, 327 out of 439 publications were marked for inclusion in the synthesis (200 ACM DL (61.1 %), 68 IEEE Xplore (20.8 %), 59 ScienceDirect (18.1 %)). Within the included publications, we found 306 approaches for input modalities, with the following distribution in the sensor/actuator categories: 35 visual (11.4 %), 50 auditory (16.3 %), 101 kinesthetic (33.0 %), 7 electrodermal (2.3 %), 79 tactile (25.8 %), 2 thermal (.7 %), 10 cerebral (3.3 %), and 8 cardiac (2.6 %). While 171 publications used unimodal input, 38 publications considered multimodal input.

Regarding output modalities, we found 416 approaches: 221 visual (53.1 %), 110 auditory (26.4 %), 7 kinesthetic (1.7 %), 52 tactile (12.5 %), 4 thermal (.96 %), and 5 olfactory (1.2 %). Multimodal output was employed by 81 publications, while 207 used unimodal output. Besides, we found 14 publications that utilized a nomadic device (e.g., smartphone or wearable device) for input and 17 publications for output.

4 EXTENDING THE DESIGN SPACE FOR IN-VEHICLE INTERACTION

Although visual, auditory, and haptic modalities dominate current in-vehicle interaction research (see Section 3.3), our SLR reveals various approaches that utilized novel modalities, such as vestibular, thermal, olfactory, cerebral, or cardiac. This highlights that future automotive UI design should consider such (novel) modalities. Especially, regarding increasing automation, shifting the driver’s role to a passenger requiring adaptive intention recognition that considers different driving situations [274] and making modalities using physiological, brain, gaze, or emotion interfaces more useful [80]. The results of our SLR further indicate the relevance of answering RQ 3, “What is the design space for in-vehicle interaction that includes an extensive set of human sensors and actuators?”. Therefore, we propose a design space for in-vehicle interaction extending the design space by Kern and Schmidt [177].

Two dimensions span our design space: (D1) interaction location and (D2) interaction modality. D1 is defined by two parameters that were selected according to the results of the interaction location coding (see Section 4.3): nomadic and anchored. D2 is defined by two parameters: input modality and output modality. The parameter nomadic of D1 has four levels: AR, VR, handheld, and wearable. The parameter anchored of D1 has 12 values: front, windshield, dashboard, head-up display, center stack, steering wheel, seat, table, door, rear, floor, and ceiling. The front location is adopted from Kern and Schmidt [177]. We decided not to differentiate the interaction locations any further to limit complexity. However, the set of values can be expanded in the future, e.g., seat may be differentiated into seat pan, backrest, and headrest. Both parameters of D2 (input and output) have the same 12 values based on the sensor/actuator categories (see Section 2.1): visual, auditory, vestibular, kinesthetic, electrodermal, tactile, thermal, pain, olfactory, gustatory, cerebral, and cardiac.

We followed a morphological analysis and combined the aforementioned values into a multidimensional matrix, also known as Zwicky box [416], which is a well-established tool for the creation of design spaces [27, 145]. Such a matrix contains all combinations of parameters and helps identify promising families of solutions and a possible lack of solutions by the number of solutions in specific cells. The resulting matrix containing all approaches for input modalities found in our SLR is shown in Table 1 and for output modalities in Table 2, where D1 is positioned on the y-axis and D2 on the x-axis. Grayed-out cells indicate that a solution using the respective modality and location does not make sense due to ethical concerns (e.g., pain stimuli) or technical reasons (e.g., no vestibular actuator or impractical cerebral/cardiac output). Our proposed design space can be used: (1) as a tool to classify an existing or one’s own work, (2) as a taxonomy of in-vehicle interaction research to inform a literature search, and (3) as an ideation tool for the creation of novel interaction concepts utilizing interaction locations throughout the vehicle.

In the following, we describe the found input and output modalities mapped to the sensor/actuator categories in detail (RQ 1). As the interaction was often multimodal, we also present a combination matrix for multimodal input and output (RQ 1). We conclude by classifying the found modalities into interaction locations across the vehicle interior and nomadic devices (RQ 2).
Table 1. The matrix obtained by the morphological approach, showing the design space for in-vehicle input modalities. The combinations highlighted in gray make technically no sense or are unethical.

<table>
<thead>
<tr>
<th>Device Type</th>
<th>Visual</th>
<th>Auditory</th>
<th>Tactile</th>
<th>Kinesthetic</th>
<th>Olfactory</th>
<th>Gustatory</th>
<th>Ocular</th>
<th>Cutaneous</th>
<th>Auditory</th>
<th>Vestibular</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR/VR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Headset</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wearable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Windshield</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Head Mount</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arm Band</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steering Wheel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Side Interior</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.1 Input Modalities

Users explicitly or implicitly act using their actuators on in-vehicle systems that are equipped with sensors, e.g., cameras or touchscreens. In the following, we present examples for each human actuator category of the found input modalities. However, we did not find any publications included in our synthesis that considered olfactory or gustatory input modalities.

**Visual:** Our SLR revealed that the eyes were mainly used in gaze-based interaction. For example, Riegler et al. [306] employed gaze-based interaction with windshield displays and evaluated the dwell time for object selection. Roeder et al. [316] investigated gaze-activated voice input to reduce visual distraction, and they enabled drivers to utilize gaze input to perform a secondary task [317]. Gaze was also used implicitly to infer drivers’ interest in objects. For example, by analyzing the gaze direction to determine the referred object in an interaction [382] or to identify referenced objects outside the vehicle [2]. An in-vehicle system could also detect glances, which occur when a user briefly looks at something, e.g., at secondary screens while driving [341]. In this context, Ulahannan et al. [368] used glance behavior to inform adaptive UIs in AVs. A 6-second sequence of glances was used by Fridman et al. [108] to predict the driver state. Likewise, Mahajan et al. [240] analyzed glance behavior (e.g., frequency) regarding situation awareness before takeover requests (TORs). Besides, longer eye movement
sequences can be used, which we found to be mainly implicit input for predictive algorithms. For example, eye movements before a TOR were used to predict takeover behavior to enable driver context adaptive UIs [281], and to predict subsequent driving performance [402]. Other (often implicit) visual input modalities are eye blinks and pupil characteristics. For example, Benedetto et al. [32] measured pupil diameter, blink rate, and blink duration to assess the driver’s mental workload and inform accident prevention systems. Likewise, Mahajan et al. [240] measured pupil diameter and blink frequency to assess the changed driver alertness due to automation and the TOR performance. However, we did not find any driver assistance systems already built into vehicles, such as drowsiness detection using visual input, which could be due to the set time range (2011-2021).

**Auditory:** All sounds that are produced by human actuators, such as the vocal tract, can be used as auditory input for in-vehicle systems. Our synthesis showed that speech-based inputs were mainly used in conversational/dialog systems, which can reply via computer-generated speech. For example, dialog systems by Nawa et al. [267] and Lin et al. [229] retrieve speech input to anticipate the dialog context and response relevant information. Similarly, dialog systems were evaluated regarding usability and driver distraction [147], and design guidelines for dialog UIs were defined to address issues such as mental workload or task-related fatigue [205]. While dialog systems require context, simpler forms of speech-based input are voice commands. For example, Winzer et al. [400] utilized voice commands like “left” or “right” for control of a traffic light assistant. Lee et al. [218] enabled users to change the automated driving state via, e.g., “auto drive” or “manual drive”. We also found approaches that used voice commands for communication with a robot assistant [394] and interaction with an automated cab [210]. However, auditory input is more than using natural language in a dialog or voice command. We found in-vehicle systems that utilized subtle messages conveyed by the human voice. Tischler et al. [360] described emotion recognition (inference of valence and arousal) using changes of pitch, intensity, and energy of the speech signal. Furthermore, non-speech sounds generated by the vocal tract or body parts were used. For example, Funk et al. [111] enabled nonverbal auditory input (humming, snapping, and clapping) for interaction with an assistant while driving.

**Kinesthetic:** Kinesthetic input comprises the position and movement of body parts and activity of muscles and joints [272], to which we also attribute the facial muscle activity needed for emotion and facial expressions. For example, body position classifiers were used to recognize driver distraction [390], and posture tracking allows inference of situation awareness and reaction times [233]. The body pressure applied to the seat was also analyzed to inform optimal seat vibration points [181]. Apart from whole body characteristics, body parts such as the head can provide more specific input. For example, head pose tracking was used to analyze driver behavior and monitor attention [42], and head movement to adapt the rendering of stereoscopic 3D displays [51]. We also found a work that measured the distance between feet and pedals to estimate situation awareness [298]. Besides, facial expressions were used to measure the driver’s stress level [143], to adapt the personality of a voice assistant [46], and facial landmarks were utilized to classify situation awareness [298]. Another commonly found kinesthetic modality was gesture input. For example, Jiang et al. [170] investigated display control via hand gestures, and Ahmad et al. [6] used hand pointing to predict the display item a user intends to select. Some approaches considered finger gestures and pointing, e.g., to control in-vehicle functions such as volume [171] or for contactless operation of a screen using mid-air finger movements [310]. We also found various approaches to recognize kinesthetic input via hardware such as buttons [179], levers [151], switches [376], and rotary controllers [12]. Schartmüller et al. [327] reimagined the steering wheel for NDRTs by attaching a keyboard, while Jakus et al. [166] attached a computer mouse to the steering wheel backside. There were also approaches leveraging shape-changing input, e.g., Heijboer et al. [159] imagined a shape-changing dashboard that users can actively deform.

**Electrodermal:** In contrast to, e.g., visual, auditory, or kinesthetic input, EDA is exclusively implicit input. For example, Wang et al. [384] used body electric conductivity to distinguish driver and passenger touches on screens. Madrid et al. [239] presented a biometric UI for stress and awareness detection using driver’s galvanic
skin response (GSR), which was also used by Pakdamanian et al. [281] to predict takeover performance. Besides, electromyography of the user’s forearm muscles was used to recognize micro-gestures performed on the steering wheel [17].

**Tactile:** We mostly found approaches that recognized tactile input via touch-sensitive surfaces, e.g., [81, 156, 210, 270]. However, Berger et al. [34] investigated an absolute indirect touch interaction concept to enable passenger interaction with an infotainment system. Apart from touch-based input, we also found publications that considered pressure-based inputs. For example, Swette et al. [356] presented pressure-based buttons on a touch surface, and finger pressure input deforms the shape-changing dashboard by Heijboer et al. [139].

**Thermal:** The temperature dispensed via the skin can be utilized as a thermal input. However, as humans do not actively control their body temperature, we only found approaches that used implicit thermal inputs, e.g., to assess driver emotion [386], or stress [143].

**Cerebral:** We found that cerebral input modalities were mainly used for implicit interactions. For example, Zhu et al. [413] employed EEG for truck driver drowsiness detection and early warning system. However, some approaches utilized brain signals for explicit control of vehicle functions. He et al. [137] proposed an interface between drivers and in-vehicle devices using brain signals that can be used while performing primary driving tasks. Likewise, in a concept by Hood et al. [148], a brain-computer interface enabled drivers to control vehicle functions, including acceleration and steering.

**Cardiac:** In contrast to cerebral input, we did not find any approach for explicit cardiac input. Instead, the heart was only used as an actuator for implicit input. For example, Madrid et al. [239] assessed the driver’s heart rate to detect stress and enable real-time feedback so that drivers can take corrective actions on time. Similarly, Hayashi et al. [134] developed an abnormal sign detection system based on driver’s heart pulse measures. Electrocardiogram measures were also used by Wang et al. [386] to improve driver emotion recognition.

### 4.2 Output Modalities

Human sensors perceive any output modality generated by vehicle actuators, e.g., speakers or vibration motors. In the following, we present examples for each human sensor category of found output modalities. However, we did not find any publications that considered electrodermal, pain, gustatory, cerebral, or cardiac output modalities.

**Visual:** The simplest form of visual output is light. For example, Johns et al. [172] used a steering wheel LED strip to create a virtual steering wheel that moves in advance while the steering wheel is locked. In a concept by Zhu et al. [413], a single LED was used to warn in case of driver drowsiness. Besides, we found approaches for peripheral light feedback. For example, peripheral LED strips acting as chase lights to influence the driver’s perception of speed [250] or using peripheral light flashes behind the steering wheel to indicate a TOR [323]. Some approaches utilize ambient light, e.g., to highlight pedestrian trajectories [235]. Most approaches in our synthesis used screens of various sizes and functions, e.g., [189, 193, 354, 393]. However, we also found different visualization approaches, e.g., head-up displays (HUDs) [331], stereoscopic displays [392], active shutter glasses [49], robotic companion displays [406], projections [257], and shape changing devices [257]. Regarding HUDs, Beck and Park [31] investigated the perceived importance of different HUD information items, and Häuslschmid et al. [130] evaluated the recognition of stimuli on a windshield HUD. Wiegand et al. [392] envisioned a stereoscopic 3D display to show information for early takeover preparation. A concept by Broy et al. [49] used shutter glasses to visualize a 3D effect on a dashboard display. Yang et al. [406] presented a robotic companion display that automatically pans and tilts. Mok et al. [257] projected the vehicle’s automation status on the steering wheel while also changing its shape to visualize a TOR. We also found various approaches for visual output that utilized AR. For example, Lauber and Butz [211] displayed driver warnings via AR HMD. There are also approaches without HMDs, e.g., Akash et al. [9] augmented the windshield with AR-highlights of environment objects, or Lindemann et al. [230] explored the use of windshield AR for driver assistance in short-notice takeovers. Similar
Table 2. The matrix obtained by the morphological approach, showing the design space for in-vehicle *output* modalities. The combinations highlighted in gray make technically no sense or are unethical.

<table>
<thead>
<tr>
<th>AR/VR</th>
<th>Handheld</th>
<th>Wearable</th>
<th>Foot</th>
<th>Windshield</th>
<th>Dashboard</th>
<th>HUD</th>
<th>Center Stack</th>
<th>Steering Wheel</th>
<th>Seat</th>
<th>Table</th>
<th>Door</th>
<th>Rear Floor</th>
<th>Ceiling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<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><strong>AR</strong></td>
<td></td>
<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></td>
<td>AR</td>
<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></td>
<td>AR Smartglasses</td>
<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></td>
<td>AR Smartwatch</td>
<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></td>
<td>AR Headset</td>
<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><strong>VR</strong></td>
<td></td>
<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></td>
<td>VR HMD</td>
<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></td>
<td>VR Smartglasses</td>
<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></td>
<td>VR Smartwatch</td>
<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><strong>Handheld</strong></td>
<td></td>
<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></td>
<td>Smartphone</td>
<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></td>
<td>Smartphone</td>
<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></td>
<td>Smartphone</td>
<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><strong>Wearable</strong></td>
<td></td>
<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></td>
<td>Smartwatch</td>
<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><strong>Foot</strong></td>
<td></td>
<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></td>
<td></td>
<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></td>
<td></td>
<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><strong>Windshield</strong></td>
<td></td>
<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></td>
<td></td>
<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></td>
<td></td>
<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><strong>Dashboard</strong></td>
<td></td>
<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></td>
<td></td>
<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></td>
<td></td>
<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><strong>HUD</strong></td>
<td></td>
<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></td>
<td></td>
<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></td>
<td></td>
<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><strong>Center Stack</strong></td>
<td></td>
<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></td>
<td></td>
<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></td>
<td></td>
<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><strong>Steering Wheel</strong></td>
<td></td>
<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></td>
<td></td>
<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></td>
<td></td>
<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><strong>Seat</strong></td>
<td></td>
<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></td>
<td></td>
<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></td>
<td></td>
<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><strong>Table</strong></td>
<td></td>
<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></td>
<td></td>
<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></td>
<td></td>
<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><strong>Door</strong></td>
<td></td>
<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></td>
<td></td>
<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></td>
<td></td>
<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><strong>Rear Floor</strong></td>
<td></td>
<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></td>
<td></td>
<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></td>
<td></td>
<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><strong>Ceiling</strong></td>
<td></td>
<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></td>
<td></td>
<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></td>
<td></td>
<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>

to windshield AR, Häkkilä et al. [158] proposed a concept for AR-enabled door windows. Besides, we found in-vehicle VR concepts. For example, McGill et al. [248] and Li et al. [226] investigated the use of VR HMDs for in-vehicle applications.

**Auditory:** We found various forms of simple auditory output ranging from sounds, e.g., [99] to earcons that are brief, distinctive, often melodic sounds to convey a message similar to an icon, e.g., [113]. Besides, we found approaches that produce artificial speech and interact with the user in a dialog system, e.g., [180]. However, some concepts used text-to-speech systems primarily used in non-conversational interaction, e.g., [126]. While most publications that considered auditory output focused on speech and sound, we found approaches that utilized music. For example, Burnett et al. [54] presented a concept to alter the speed perception of passengers through subliminal music adaptions. Similarly, Hernandez et al. [143] used adaptive music (e.g., relaxing songs) if the driver is stressed, and Chen [60] manipulated background music using blended sonification to communicate reliability in automated driving.
Vestibular: Vestibular output modalities address the human sensor of balance and often occur coupled with kinesthetic output as the body is moved by an exterior force to stimulate the vestibular system in the inner ear [33]. In our synthesis, we found a concept by Winner and Wachenfeld [397] that automatically adapts the seat’s tilt angle when an automated vehicle is driving in a curve to normalize the passenger’s vestibular experience.

Kinesthetic: Kinesthetic output consists of any movement sensed over muscles and joints, making it a versatile way to provide feedback to users. For example, the shape-changing steering wheel by Mok et al. [257] provided kinesthetic stimuli to hands and arms to indicate the automation state and TORs. Bhardwaj et al. [37] employed a system that adapts the steering wheel impedance to ensure a safe transition of control between vehicle and driver. A swivel seat can also be utilized, as shown by, e.g., Sun et al. [352]. Besides, whole-body kinesthetic output was investigated by Gaspar et al. [114] who conveyed forward collision warnings via brake pulses. We also found approaches for haptic pedals that, e.g., provided force feedback to guide drivers toward applying ideal throttle [73] or a pressure point such that the pedal is pressed no further than the recommended angle for ecological driving [141].

Tactile: Most publications in the synthesis employed vibration motors, e.g., [181, 232, 369, 376]. However, some approaches used tactors. For example, in the form of a cutaneous push on the steering wheel for navigation [84], a tactor to convey warning cues for driver inattention [222], a pin array that creates tactile patterns for fingers and hand [173], a silicone touchscreen cover foil to increase tactile feedback [414], or a 3D printed stencil that makes underlying touchscreen controls tangible [66]. Besides, instead of skin contact with the tactile device, Shakeri et al. [338] and Harrington et al. [132] proposed mid-air ultrasonic tactile feedback for gesture interaction.

Thermal: In our synthesis, we only found approaches for thermal output via skin contact. For example, Di Campli San Vito et al. [83] used a Peltier device to apply thermal feedback to the driver’s finger, where the temperature indicated the desired lane change direction. They also attached a Peltier element to the steering wheel to provide thermal feedback indicating navigation cues [85]. Besides, Heijboer et al. [139] envisioned a dashboard that conveys information via temperature changes.

Olfactory: Schartmüller et al. [325] applied essential oils on a piece of felt attached to the center mirror to investigate non-invasive motion sickness mitigation in AVs. Similarly, Dmitrenko et al. [89] utilized essential oils to produce pleasant scents to promote safer driving, better mood, and improved well-being in drivers. They also compared scent output methods and their usability for in-vehicle olfactory interactions regarding distance, volume, and speed [93].

4.3 Interaction Locations

According to our code book (see Section 3.2.1), we distinguish between two types of interaction locations: (1) nomadic and (2) anchored. The locations are visualized in Figure 1 and on our interactive website4.

Nomadic: The nomadic interaction location contains devices unbound within the vehicle interior such as HMDs (e.g., AR or VR), handheld devices (e.g., smartphone or laptop), and wearable devices (e.g., smartwatch or smart ring), see Figure 1. We found that input sensed via nomadic devices was mainly kinesthetic and tactile. Kinesthetic input was performed on handheld devices, such as buttons on a smartphone [224], gaming controller [146], or wearables, e.g., smartwatch [131] and smart ring [118]. Similarly, tactile input was performed on touchscreen handheld devices, e.g., smartphone [260] or tablet [251], and a wearable smartwatch [285]. Besides, physiological input modalities were measured using wearable sensors at nomadic locations, e.g., EDA on the forearm [17], skin temperature on the fingers [386], brain activity on the head[38], and heart activity on ear [281] or torso [319]. Regarding output, we found various approaches that utilized nomadic displays, such as AR [193], VR [248], smartphone [136], tablet [276], or smartwatch [223]. However, our SLR also revealed some approaches that

4https://in-vehicle-interaction-design-space.onrender.com/ Interactive tool to support investigation of in-vehicle interaction research and design.
provided tactile output at wearable locations via, e.g., smart ring [118], vibration wristband [292], vibration belt [195], or tactor belt [293].

**Anchored:** The anchored interaction location contains specific locations throughout the vehicle interior: front, rear, door, ceiling, floor, table, and seat (see Figure 1). To account for the large number of publications focusing on front-based in-vehicle interaction, similar to [177], we further differentiate the front into windshield, dashboard, head-up display, center stack, and steering wheel. Our synthesis showed that input modalities were mainly sensed via vehicle sensors located in the front. For example, cameras for visual input, such as gaze [116], blinking [240], or pupil dilation [240], or kinesthetic input, such as gesture [81], facial expression [153], or head movement [314]. Similarly, microphones were placed in the front to sense auditory input (e.g., speech [245] or humming [111]). Other input modalities often require direct contact with an anchored interior location. For example, kinesthetic input at the center stack [242], steering wheel [221], seat [181], table [163], or ceiling [163]. We also found approaches for tactile input at similar locations, e.g., dashboard [377], center stack [270], steering wheel [289], table [276], or door [158]. Only a few publications considered novel input modalities at anchored locations, such as EDA [143], skin temperature [143], or heart rate [239] (all located on the steering wheel).

Regarding output, we found that out of the 12 anchored locations, only rear and floor were not utilized for visual output. Besides, output was mainly provided using auditory modalities via speakers in the front (e.g., [286]), and via kinesthetic modalities at the center stack (e.g., [139]), steering wheel (e.g., [154]), seat (e.g., [123]), or floor [141]. Similarly, tactile output was provided at the center stack (e.g., [376]), steering wheel (e.g., [329]), or seat (e.g., [114]). Only a few publications considered other output modalities at anchored locations, e.g., vestibular output at the seat [397], scent output in the front [325], or thermal output on the center stack [83] or steering wheel [85].

### 4.4 Multimodal Interaction

Multimodal interfaces provide multiple modes of interaction between passengers and vehicle. We do not distinguish between sequential multimodal, requiring users to switch between modes, and simultaneous multimodal that allows users to use multiple modes at a time. Due to visualization complexity, we do not consider modality changes, e.g., when a passenger uses gaze input and the vehicle responds with speech output. Instead, our interactive website can filter publications that considered such modality changes from our publication data set.

To visualize the found approaches for multimodal interaction, we created a square matrix of order 12 representing combinations between modalities of any two sensor/actuator categories (see Table 3). Both dimensions contain the sensor/actuator categories: visual, auditory, vestibular, kinesthetic, electrodermal, tactile, thermal, pain, olfactory, gustatory, cerebral, and cardiac. We further differentiate between specific modalities within these categories, e.g., speech, sound, and music within the auditory category. As in Section 4.1 and 4.2, we consider input and output from the human perspective, i.e., input is produced by human actuators (e.g., hands), and output is perceived by human sensors (e.g., skin). A matrix cell contains a combination of specific modalities, e.g., display × sound or gesture × touch × lever, and an empty cell indicates that we found no approach for the respective multimodal interaction in our SLR. Gray-colored cells highlight that combinations may not be feasible due to ethical (pain) or technical reasons (vestibular input, gustatory input, brain stimuli, and heart stimuli).

In our synthesis, we found that most publications utilized visual modalities for **multimodal input.** Gaze was combined with auditory inputs, e.g., speech [316] and voice [65], with kinesthetic inputs, e.g., finger pointing [2], hand gestures [120], head pose [13], and body pose [298], or with tactile touch input [187]. Besides, some publications combined eye movement with electrodermal (GSR [87]) or with cardiac modalities (heart rate [281]). Moreover, 30 publications used auditory, kinesthetic, or tactile modalities for multimodal input. Regarding auditory input, speech was used dominantly in combination with, e.g., button [224], gesture [289], emotion [411], touch [285], or heart rate [204]. Kinesthetic modalities were used diversely, e.g., combined with GSR [239].
Table 3. Combination matrix for multimodal in-vehicle interaction. Green-colored cells show combinations of multimodal input, and yellow-colored cells show multimodal output. The combinations highlighted in gray make technically no sense or are unethical.

<table>
<thead>
<tr>
<th>Visual</th>
<th>Auditory</th>
<th>Tactile</th>
<th>Olfactory</th>
<th>Gustatory</th>
<th>Cerebral</th>
<th>Cardiac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech x Head Movements</td>
<td>Display x Air Conditioning</td>
<td>Peltier Element x Shape Changing</td>
<td>Display x Scent</td>
<td>HUD x Scent</td>
<td>Voice x Display</td>
<td>Skin Temp. x HUD</td>
</tr>
<tr>
<td>Speech x Head Movements</td>
<td>Display x Air Conditioning</td>
<td>Peltier Element x Shape Changing</td>
<td>Display x Scent</td>
<td>HUD x Scent</td>
<td>Voice x Display</td>
<td>Skin Temp. x HUD</td>
</tr>
<tr>
<td>Speech x Head Movements</td>
<td>Display x Air Conditioning</td>
<td>Peltier Element x Shape Changing</td>
<td>Display x Scent</td>
<td>HUD x Scent</td>
<td>Voice x Display</td>
<td>Skin Temp. x HUD</td>
</tr>
<tr>
<td>Speech x Head Movements</td>
<td>Display x Air Conditioning</td>
<td>Peltier Element x Shape Changing</td>
<td>Display x Scent</td>
<td>HUD x Scent</td>
<td>Voice x Display</td>
<td>Skin Temp. x HUD</td>
</tr>
</tbody>
</table>

Regarding multimodal output, we found that most publications include visual modalities, mainly because their approaches required a display to visualize a task or an UI. Accordingly, display was combined with various output modalities, e.g., with auditory (speech [359], sound [269], music [16]), kinesthetic (haptic pedal [141], seat movement [289]), or tactile (vibration [369], ultrasound [117], cutaneous push [337]). Other visual output modalities used were LED and HUD. For example, LED was combined with speech [353], sound [323], shape changing [139], vibration [232], ultrasound [338], Peltier element [139], or scent [91]. HUD was combined with, e.g., speech [292], sound [150], vibration [292], cutaneous push [222], or scent [90]. Besides, publications used multimodal output that combined windshield AR with sound [328] or vibration [98]. We also found approaches for multimodal output that did not use visual output. For example, auditory combined with kinesthetic output (sound × steering torque [154] and sound × seat rotation [352]) or with tactile output [173]. We found that only four publications utilized thermal or olfactory modalities for multimodal output. Peltier elements were used along shape-changing output [139], and scent was combined with sound [91] or vibration [91]. We did not find any publications that considered multimodal output containing vestibular, electrodermal, or gustatory modalities.
5 USER ACCEPTANCE OF IN-VEHICLE INTERACTION

The SLR results and our design space highlight many opportunities for novel in-vehicle interaction approaches other than visual, auditory, and tactile. However, it is unclear whether experimental modalities (e.g., electrodrymal, cerebral, or cardiac) are feasible. For instance, passengers might perceive modalities as invasive (e.g., cerebral) or unpleasant (e.g., thermal) and, therefore, reject usage. Moreover, our SLR and design space revealed different interaction locations throughout the interior, with many opportunities for novel interaction locations other than the front. Still, it is unknown whether users perceive interaction at such unfamiliar locations as practical, e.g., rear, floor, or ceiling. To gain initial insights, we designed and conducted a within-subject online study with N=48 participants. The independent variables were interaction type with two levels: input and output, modality with 11 levels: visual, auditory, vestibular (only output), kinesthetic, electrodrymal, tactile, thermal, olfactory, gustatory (only output), cerebral (only input), and cardiac (only input), and interaction location with ten levels: AR, VR, handheld, wearable, seat, door, floor, rear, table, and ceiling. Following our design space, we excluded impractical modalities (e.g., pain) and removed the front location, as user acceptance of this location is already well known to research. The following RQ guided the exploratory study: RQ 4: How do users perceive the (1) usefulness, (2) real-world usage, and (3) comfort of different input/output modalities and interaction locations within a conceptual vehicle?

5.1 Study Procedure and Measurements

We opted for an image-based survey for two reasons. First, we aimed for meta-level insights covering the mere concept of an interaction, disregarding profound functionalities of each modality. Second, it is impractical to develop working prototypes for all 18 input/output modalities, while also using all 10 locations. The study consisted of a questionnaire, starting with an introduction of aim and scenario. In the following, there were two parts: Part 1 showed 18 images (9 input and 9 output) of a passenger on the front seat using possible interaction modalities in a concept vehicle (e.g., see Figure 4). Part 2 showed 20 images (10 input and 10 output) of a passenger on the front seat interacting at possible interaction locations within a concept vehicle (e.g., see Figure 5). The full versions of the study images are visible in the Appendix Figure 12, Figure 13, Figure 14, and Figure 15. For the interaction locations in Part 2, we always showed gesture input, and for output, we used a display. However, this was exemplary, and any other input or output modality from Part 1 may also be feasible. Part 1 was introduced first to ensure that all concepts for input and output modalities were already known to the participants in Part 2. The images per part were presented in randomized order for each participant, and after each part, participants could provide open feedback. Finally, there was a demographic and a concluding questionnaire. On average, a session lasted 17 min and participants were compensated with £2.

After each image in Part 1, a self-developed 7-point scale (1=strongly disagree to 7=strongly agree) was used to assess the participants’ perceived usefulness of the interaction ("I consider the presented interaction to be useful.") whether they could imagine real-world usage ("I would use the presented interaction in a real vehicle.") and the anticipated comfort of use ("I would feel comfortable using the presented interaction."). For Part 2, three adapted questions were shown after each image: "I consider the depicted interaction location to be useful", "I would use the depicted interaction location in a real vehicle", and "I would feel comfortable using the depicted interaction location."

5.2 Results

Before the study, we computed the required sample size via a priori power analysis using G*Power [101]. To achieve a power of 0.95 with \( \alpha = 0.05 \), 37 participants should result in a small effect size (0.15 [110]) in a within-subjects ANOVA. Originally, 62 individuals were recruited via Prolific (a platform to find participants [282], already used in automotive research, e.g., [70]), but 14 had to be excluded due to failed attention checks (e.g.,
Fig. 4. Concept images used in the online study. Input modalities: a) gesture, b) touch, c) gaze, d) electrodermal, e) speech, f) thermal, g) body odor, h) cerebral, and i) cardiac. Output modalities: j) shape changing, k) vibration, l) display, m) electrodermal, n) speech, o) thermal, p) scent, q) seat rotation, and r) gustatory.

Fig. 5. Concept images used in the online study. Nomadic interaction locations: a) AR, b) VR, c) handheld, and d) wearable. Anchored interaction locations: e) seat, f) floor, g) door, h) table, i) rear, and j) ceiling.

"Check the fourth option."). We recruited US citizens only to avoid confounding variables such as driving laws (e.g., left and right traffic). Of the remaining 48 participants, 29 were male, 18 female, and one non-binary. Participants were, on average, \( M=34.25 \) (SD=8.83, between 19 and 50) years old and, in total, 44 participants had a driver license. R in version 4.1.2, RStudio in version 2021.09.0 and, for non-parametric data (checked via normality distribution and homogeneity of variance assumption), the non-parametric ANOVA (NPAV) by Lüpsen [236] was employed. All packages were up-to-date in January 2022.

5.2.1 Input and Output Modalities. Figure 6 shows the average ratings of the perceived usefulness of a modality. The NPAV found a significant main effect of modality on perceived usefulness (\( F(10,470) = 8.45, p<0.001 \)). The NPAV also found a significant interaction effect (IE) of interaction type (input/output) × modality on perceived usefulness (\( F(6,282) = 6.16, p<0.001 \)). Figure 7 shows the average ratings of the anticipated real-world usage of a modality. The NPAV found a significant main effect of modality on real-world usage (\( F(10,470) = 9.03, p<0.001 \)). The NPAV additionally found a significant IE of interaction type × modality on real-world usage (\( F(6,282) = 4.25, p<0.001 \)). Figure 8 shows the average ratings of the anticipated comfort of a modality. The NPAV found a significant main effect of modality on anticipated comfort (\( F(10,470) = 10.62, p<0.001 \)). Besides, the NPAV found a significant IE of interaction type × modality on anticipated comfort (\( F(6,282) = 6.25, p<0.001 \)).
5.2.2 Interaction Locations. Figure 9 shows the average ratings of the perceived usefulness of a location. The NPAV found a significant main effect of location on perceived usefulness ($F(9, 423) = 4.45$, $p<0.001$). Figure 10 shows the average ratings of the perceived usefulness of a location. The NPAV found a significant main effect of location on real-world usage ($F(9, 423) = 3.69$, $p<0.001$). Figure 11 shows the average ratings of the perceived usefulness of a location. The NPAV found a significant main effect of location on anticipated comfort ($F(9, 423) = 2.75$, $p=0.004$).

5.2.3 Open Feedback. Participants perceived the shown input and output modalities as interesting [P32] and comfortable [P19, P24, P25]. However, they also noted that, e.g., some inputs were especially invasive [P5], and the vehicle should not read minds [P11]. Besides, [P38] stated that except for speech output and vibration to acknowledge choices, all modalities are dangerous regardless if the user is a driver or passenger. The shown interaction locations were perceived as interesting [P7], attractive [P25], and comfortable [P19] (e.g., for travel [P24]). Still, [P5] noted that some locations generally seemed not comfortable, and even dangerous to interact with, in a moving vehicle. [P11] thought that looking down for information (e.g., floor) is stressful, and [P38] would feel awkward using the majority of interaction locations in a car except leaning back and watching TV on the ceiling.

6 DISCUSSION
In the following, we will discuss the results of the user study and our SLR and, based on these, outline future research interests in the in-vehicle interaction domain.
6.1 Opportunities for Future In-Vehicle Interaction

Our design space (see Section 4) revealed several underexplored opportunities for future in-vehicle interaction. In this regard, insights from non-automotive HCI research may guide the creation of novel approaches.

Regarding **input modalities**, our design space revealed little to no utilization of thermal, olfactory, gustatory, cerebral, or cardiac modalities. However, such implicit inputs might be helpful, e.g., for monitoring passengers, adapting outputs, or refining explicit interactions. For example, the vehicle could measure the skin temperature employing wearable sensors similar to [243, 389] or without requiring invasive on-body sensors using thermal cameras [1], which might increase comfort and acceptance. Besides, the vehicle may leverage olfactory inputs to distinguish passengers based on body odor [14] or restrict driving when alcohol breath is detected [318]. Gustatory input might also be possible by embedding an electronic tongue (e.g., see [217]) into the vehicle that is capable of sensing flavors, e.g., picked up from the skin. In contrast to gustatory input, which seems impractical, brain and heart activity is valuable in determining the passenger’s state and, in the case of brain input, can even be used for active vehicle control [40]. However, there is more potential for using brain activity via brain-computer interfaces, e.g., demonstrated by [18, 104].

Regarding **output modalities**, our design space revealed only a few approaches for vestibular, kinesthetic, and thermal output and none for electrodermal, gustatory, cerebral, or cardiac. However, such output modalities may be useful as they address a dimension of human perception that differs from visual, auditory, or tactile stimuli, which other tasks might occupy (e.g., the driving task or an NDRT). Moreover, there is a great potential for using them in multimodal interactions combined with visual, auditory, or tactile output. While only two publications in our SLR considered vestibular output [352, 397], there are various techniques to stimulate the vestibular sense, e.g., galvanic vestibular stimulation [103], force simulation [149], vestibular stimuli [71, 305], or motion cues [303]. The vehicle can also apply such cues to provide kinesthetic output, e.g., via electric muscle stimulation [287], exoskeletons and gloves [366], or a moving platform [410]. Electrodermal output can be conveyed by electric stimuli [283], such as slight electric shocks [259] or electric arcs [342] to the skin. While our SLR did not reveal approaches for gustatory output, there are HCI concepts that use, e.g., edible UIs [265], pseudo-gustatory displays [266], or electrical tongue stimulation [297]. For this, the interior might be a suitable environment as edible assets can be stored. Although we only found concepts that consider implicit brain and heart activity, the vehicle could also actively stimulate the brain, e.g., using a brain implant or a defibrillator. However, technology is not yet advanced enough to enable safe and practical usage of such stimuli. Therefore, future research must investigate the usability and user acceptance of experimental output modalities.

In contrast to input modalities, the **interaction location** is often critical for output modalities since the usability may depend on the location, e.g., as users need an unobstructed view on visual output or feel tactile output. Our design space shows that publications mainly used the front as an output location, e.g., for displays [367] or vibration [268], while other interior locations are considered rarely, e.g., table, door, rear, floor, or ceiling. However, those locations may increase in relevance regarding the transition to automated driving and NDRTs [82, 288], which should be comfortably available regardless of whether passengers are seated in front, rear, or swivel seats. Therefore, we argue that automotive UI research should move away from a front-focused design for all input and output modalities in automated driving towards considering the whole vehicle as an interaction space. Our extended design space can thereby guide the ideation process for such research concepts.

6.2 User Acceptance of In-Vehicle Interaction Modalities and Locations

The study results (see Section 5.2) revealed differences between the acceptance of modalities, resembling findings from related work. For example, visual output was perceived as more useful than visual input (gaze), which is in line with Akkil et al. [10] who found that users perceive some gaze interactions as complex, strenuous, and slow. This suggests that implicit visual inputs such as pupil diameter or blink rate should be employed primarily...
instead of straining explicit inputs (e.g., gaze). Similarly, participants would rather use auditory output (sound) than auditory input (speech), which they might have perceived as impractical in some situations, e.g., surrounded by others [24]. Hence, future in-vehicle UIs should provide alternative modalities for such situations. Besides, tactile input via touch was generally more accepted than tactile output via vibration, which we mainly associate with the unfamiliar vibration location on the center stack. This implies that acceptance of input/output modalities depends on their characteristics, e.g., the involved human sensors/actuators, the required effort, mental demand, or interior location. Therefore, future in-vehicle UIs may employ different modalities depending on the context (e.g., touch for accurate selection instead of gaze or speech) or utilize multimodality to mitigate the drawbacks of one modality (e.g., using gesture to improve touch [4], gaze, or speech input [2]). Modalities such as olfactory, gustatory, or cerebral might be perceived as invasive, indicated by open feedback [P11, P38] and low usefulness and comfort ratings. Hence, future vehicles should embed sensors unobtrusively (e.g., seat-embedded heart rate detection [403]), and the sensed information should be made transparent to the respective users. Besides, safety concerns may have influenced the ratings, as some participants perceived modalities [P38] and locations [P5] as dangerous to use in a vehicle regardless of being driver or passenger. The users’ safety needs [80] require future interaction design to overcome such concerns by providing safe interactions in any driving situation, which might render modalities or locations unavailable (e.g., in manual or conditional automated driving).

Regarding nomadic locations, AR was generally more accepted than VR. We assume that participants were concerned about VR motion sickness [248] and perceived the see-through AR as more useful for productivity, which users might prefer over gaming, e.g., see [288]. Also, participants may have expected isolation using an opaque VR HMD and worried about reduced fallback-readiness in a takeover (SAE 3 and 4). Therefore, in-vehicle VR must not induce motion sickness (e.g., using visual motion cues [248]), provide applications other than entertainment, and consider passenger safety and trust for any SAE level. Since handheld and wearable devices are already usable in current vehicles (e.g., Android Auto [15] via smartphone or smartwatch) and have received high usefulness, real-world usage, and comfort ratings, future interaction design should first focus on such locations as acceptance already exists or can be quickly built. Besides, there is potential for modality combinations in nomadic and anchored locations (e.g., using a smartphone while the seat is vibrating [286]), which might assist in introducing unfamiliar anchored interaction locations. The high perceived usefulness of rear, table, and seat indicates that users would accept interaction at locations other than the front. However, as the ratings were lower for anticipated comfort and real-world usage, future design should aim to mitigate discomfort, e.g., by increasing the size of a rear display, using a tiltable table, or ergonomic seat armrest displays. Other anchored locations, e.g., ceiling, door, and floor, were generally less accepted since the participants probably could not imagine any use case, which requires future interaction design to elicit suitable applications to gain user acceptance.

6.3 Future Research Interests on In-Vehicle Interaction
Emerging from the opportunities unveiled in our design space and further pointed out by the user study, achieving user acceptance of novel in-vehicle interaction concepts faces several challenges. We derived the following research interests that can guide practitioners to overcome these.

6.3.1 Technological Advancement. While novel modalities such as electrodermal, olfactory, gustatory, cerebral, and cardiac are interesting from a research perspective, technical challenges may hinder or even prevent their successful introduction in future vehicles. For instance, technologies in early development stages, e.g., for cerebral, cardiac, or electrodermal modalities, may provide unreliable recognition of input [11] and limited actuation of output, thereby negatively impacting usability. Besides, experimental modalities such as olfactory input and gustatory input/output may be challenging to implement or impractical to operate in vehicles. For example, technologies for an artificial tongue [217] or nose [220] exist but have not yet been transferred to a vehicle context. Moreover, it is unclear whether all found modalities in this work can coexist in a vehicle.
We assume that manufacturers will use subsets of the presented possible vehicle sensors and actuators to suit different use cases. The integration of multiple sensors enables the detection of erroneous data from one input modality using another [261]. Likewise, the vehicle can leverage different output modalities best suited to specific situations, e.g., seat vibration for subtle messages or loud sounds alerting the driver. Furthermore, such cross-modal input/output usage can provide redundancy in interactions, increasing the reliability, especially when one vehicle sensor/actuator is unavailable [100]. A safe and fast prototyping testbed for novel in-vehicle interactions are (open-source) VR driving simulators, such as [307] or [334]. Besides, low-cost vehicle motion simulators (e.g., [71]) enable the testing of motion impacts on novel interactions without the need for an actual vehicle.

### 6.3.2 Social Acceptance

Due to our SLR’s scope including passengers other than the driver, e.g., in the rear seat, there are social challenges, which persist for AVs, as they will be most likely used in social settings, such as leisure trips with friends and family [396]. In this regard, the feasibility of in-vehicle modalities and locations may depend on the social context, i.e., sharing the vehicle with friends or strangers. For instance, users are less sensitive to the presence of strangers when on a commute trip compared to a leisure-activity trip [215], and they would feel more confident to privately use gesture [63] or speech input [24]. Likewise, in a social setting (e.g., car-sharing), we argue that users might prefer more subtle interactions via physiological input modalities, such as electrodermal, cerebral, cardiac, or thermal. Besides, the output should be adapted to the social setting, e.g., an output modality (private audio message) might be intended for a specific passenger [184], which excludes modalities (sound or displays) visible for everyone. In this regard, the interaction location may be decisive. We assume that using a touchscreen at the center table would be preferred when riding with friends than with strangers, or interacting with the rear location as a front-seat passenger might disturb rear-seat passengers. Furthermore, there might be different “acceptance personalities” [80] also visible in the open feedback where some participants had a more negative attitude towards interaction modalities or locations (e.g., [P5, P38]). Hence, future in-vehicle interaction research should also investigate the social component of user acceptance.

### 6.3.3 Explicit and Implicit In-Vehicle Interaction

Implicit in-vehicle interaction enables passengers to produce input subconsciously [336] or perceive output from the vehicle implicitly embedded in output modalities. Our SLR revealed various modalities feasible for implicit interactions, such as electrodermal, thermal, cerebral, and cardiac. Other examples include driver fatigue detection [55], emotion recognition [412], or body pose estimation [246]. The advantages of implicit interactions are manifold. For instance, the vehicle can tailor in-time airbag deployment, steering, brake, and crash avoidance if the passengers’ body positions are estimated [261, 263]. AVs can detect driver attention in the driving task via implicit cues (e.g., emotion) to determine if a takeover can be achieved safely [412]. Furthermore, the introduction of novel interaction locations, e.g., seat, door, or table, fosters new options for implicit interaction that utilize physiological vehicle sensors embedded in the seat [403] or door handle [143]. Besides, we found that modalities mainly used in explicit interactions, such as visual or auditory, are potentially usable in implicit interaction. For example, the vehicle can use auditory output via adaptive music to subliminally affect the passengers [54] or alter the color of visual UIs to increase driving performance [133]. In contrast to implicit interaction, passengers can explicitly interact with the vehicle, e.g., in conversations and performing tasks. Examples include voice commands [563], gestures [247], touch [191], and display-based modalities [50]. At times, implicit input can also assist while interacting explicitly, e.g., by guiding the interaction to the most appropriate modality (visual, auditory, or haptic) if the driver’s intentions are recognized [261]. Such a combination of implicit and explicit interactions may be promising for future in-vehicle interaction design, e.g., regarding multimodal interactions.

### 6.3.4 Perspectives on Multimodal In-Vehicle Interaction

Our combination matrix for multimodal interaction (see Table 3) revealed research potential for thermal, electrodermal, and cerebral modalities, e.g., in combination with well-established input modalities, such as visual, auditory, and kinesthetic. Besides, we found little research
regarding multimodal output that considers vestibular or electrodermal output. However, not every modality combination may be feasible, as one modality requires too much mental or physical workload to effectively use the other (e.g., using two explicit input modalities, gesture × touch). In this regard, combining an explicit modality such as touch, gesture, or speech with an implicit modality, e.g., EDA, might overcome this challenge. Still, the location of each modality decisively influences the usability of such multimodal interaction. For instance, sensing the passengers’ EDA requires direct skin contact that might not be guaranteed when they lift (both) arms to perform a gesture. Possible solutions may utilize fallback vehicle sensor locations, e.g., a nomadic wearable sensor (smartwatch), to compensate for an unavailable anchored sensor in the seat armrest. Therefore, we propose a new perspective on multimodal in-vehicle interaction that includes the location to specify the design of possible multimodal interactions, e.g., kinesthetic input at center stack × electrodermal input at seat or visual output at door × visual output at HMD.

6.3.5 Explainable In-Vehicle Artificial Intelligence. Vehicles can leverage artificial intelligence (AI) to enable adaptive or proactive implicit, explicit, and multimodal interactions. In our SLR, we found AI-based approaches, for example, in conversational speech systems [180, 229]. Besides, using physiological input modalities, such as electrodermal, cerebral, or cardiac activity, mostly requires AI to synthesize the data and make correct inferences to provide a suitable output to passengers. However, to achieve user acceptance, passengers should presumably understand the decision-making of AI, especially in case of wrong decisions or faulty output. Similarly, driving automation systems require adaptive UIs to foster trust and understanding [398]. Therefore, future in-vehicle interaction research should consider explainable AI-based implicit and multimodal interactions that may further support the successful introduction of novel modalities (e.g., thermal, cerebral, or cardiac). Solutions may include interfaces that explain the current AI decision or monitoring systems that automatically detect unreasonable outputs and provide explanations (e.g., see [119]).

6.4 Limitations and Future Work

We limited the SLR database search to query only abstracts in a set of selected venues to keep the SLR focused and the resulting set of publications manageable, which had potentially excluded relevant research. However, we assume that our results sufficiently cover the domain of in-vehicle interaction to answer our RQs.

Our user study’s presentation of self-made images might have influenced participants’ understanding of the shown interactions. However, we deem this negligible as we intended to gain meta-level insights on the topic and limit the study complexity. Still, future work should evaluate possible interaction modalities and locations using video or interactive simulations, e.g., VR or a real-world setting. Besides, in the study, we only showed one example for a modality category for internal consistency (see Figure 4). However, there might be examples of the same modality category, which may be more accepted. Therefore, future work may present more than one example per modality, e.g., tactile output via vibration, ultrasound, and cutaneous push. Similarly, the perceived usefulness, usability, and comfort of input at different interaction locations may be influenced by showing only the gesture input modality. Hence, future work should evaluate different modalities for each location to elicit strengths and weaknesses adequately to determine the feasibility of input and output modalities. Another limitation is that the study results refer to US residents only. However, there might be a regional bias, and the user acceptance of information technology may depend on culture (e.g., see [28]). Especially in the EU region, we could imagine more pessimistic results. In the future, surveys should target additional regions and compare cultural differences.

Our design space asks for usage as an ideation tool. Therefore, future work should consider the unaccounted aspects in this design space primarily. The presented research interests can further guide this process.
7 CONCLUSION

This work presents a comprehensive overview of existing literature on in-vehicle interaction regarding SAE levels 0-5 by performing an SLR on 327 publications selected from 2534 candidates. We found that existing literature mainly covers established input and output modalities, such as visual, auditory, kinesthetic, tactile, and vehicle interior locations, such as the front. However, little research was related to other modalities and interior locations. Our SLR revealed multimodal interaction as a relevant topic for future in-vehicle interactions. Therefore, we created a combination matrix for multimodal interaction, enabling practitioners to retrieve existing approaches and elicit novel modality combinations systematically. To classify the existing literature and outline research gaps and future approaches, we presented a design space for in-vehicle interaction with dimensions location and modality. Our design space extends previous design spaces [80, 177] to update in-vehicle interaction research, considering a more extensive set of possible human sensors and actuators and novel interaction locations throughout the vehicle interior. To gain initial insights on the user acceptance of possible concepts for input and output modalities and interaction locations, we conducted an image-based online study (N=48). The shown concepts were deduced from the gaps unveiled in our design space and existing literature. The results revealed general user acceptance of novel input and output modalities and locations while also showing differences between modalities and locations, encouraging future work to consider a modality’s characteristics and interior location to achieve high usability. We argue that the SLR results, the combination matrix of multimodal interactions, the design space, and insights from our user study advise future work to evaluate key design decisions, exploit trends, and explore new areas in the domain of in-vehicle interaction.

ACKNOWLEDGMENTS

The authors thank all study participants. This work was supported by the project ‘SEMULIN’ (selbstunterstützende, multimodale Interaktion) funded by the Federal Ministry for Economic Affairs and Energy (BMWi).

REFERENCES


A Design Space for Human Sensor and Actuator Focused In-Vehicle Interaction Based on a Systematic Literature Review


A Design Space for Human Sensor and Actuator Focused In-Vehicle Interaction Based on a Systematic Literature Review  •  56:39


A Design Space for Human Sensor and Actuator Focused In-Vehicle Interaction Based on a Systematic Literature Review

User Interfaces and Interactive Vehicular Applications (AutomotiveUI '18). Association for Computing Machinery, New York, NY, USA, 137–145. https://doi.org/10.1145/3239060.3239069


A Design Space for Human Sensor and Actuator Focused In-Vehicle Interaction Based on a Systematic Literature Review


Fei Yan, Mark Eilers, Martin Baumann, and Andreas Luedtke. 2016. Development of a Lane Change Assistance System Adapting to Driver’s Uncertainty During Decision-Making. In Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI ’16 Adjunct). Association for Computing Machinery, New York, NY, USA, 93–98. https://doi.org/10.1145/3004323.3004334


Jeong-Yean Yang, Yong-Ho Jo, Jae-Chul Kim, and Dong-Soo Kwon. 2013. Affective interaction with a companion robot in an interactive driving assistant system. In 2013 IEEE Intelligent Vehicles Symposium (IV). IEEE, Gold Coast, QLD, Australia, 1392–1397. https://doi.org/10.1109/IVS.2013.6629661


### A INPUT AND OUTPUT MODALITY KEYWORDS

Table 4. Keywords for input and output modalities that are categorized by the respective human sensors/actuators. **Blue**: keywords obtained from a preliminary literature review in the automotive domain. **Brown**: keywords obtained from a Google Scholar literature search for input/output modalities, including non-automotive research. **Olive**: keywords obtained in a brainstorm session among the authors.

<table>
<thead>
<tr>
<th>Input Modality (generated by Human Actuators)</th>
<th>Output Modality (sensed by Human Sensors)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Visual</strong></td>
<td></td>
</tr>
<tr>
<td>Gaze, Eye Movement</td>
<td>LED, LED Strip, Ambient Light, Peripheral Light Glasses Display, Head-Up-Display, Screen, Monitor</td>
</tr>
<tr>
<td>Eye Blink, Blink Rate [62], Squinting [372], Eyebrow Raises [125], Pupillary Behavior [23], Pupil Diameter [23, 62]</td>
<td></td>
</tr>
<tr>
<td><strong>Auditory</strong></td>
<td></td>
</tr>
<tr>
<td>Speech, Voice, Verbal, Vocal Affect (Rate, Pitch, Intensity, Quality) [61, 415], Natural Language Non-Speech Sounds (Whistling, Humming, ...) [343, 344], Acoustic Gesture (Tone and Pitch) [343, 344]</td>
<td>Auditory Cues, Barcons, Sounds, Sonification [142], Auditory Icons [142], Spatial Sound [345], Speech-based Assistant, Conversational Agent, Voice-Command, Synthetic Speech [61], Natural Language Music</td>
</tr>
<tr>
<td><strong>Vestibular</strong></td>
<td>Vehicle Roll Motion, Vestibular Stimuli [305], Motion Cues [303], Induced Self-Motion [149], Perception of Self-Motion [312], Force Simulation [149], Proxemics [244], Galvanic Vestibular Stimulation [103]</td>
</tr>
<tr>
<td>Proxemics [244], Balance [94], Vestibular Behavior [186], Accelerating Body Motions</td>
<td></td>
</tr>
<tr>
<td><strong>Kinaesthetic</strong></td>
<td></td>
</tr>
<tr>
<td>Lever, Joystick, Gesture, Hand-, Head-, Finger-[75], Arm-[96], Leg-[96], Feet-Movement [379], Body Posture, Body Motion, Walking [36], Jumping [401], Emotion, Facial Expression, Muscle Activity [311]</td>
<td>Electric Muscle Stimulation [287], Functional Electrical Stimulation [271], Kinesthetic Feedback, Exoskeleton force-feedback glove [366], Moving Platform/Plate [410]</td>
</tr>
<tr>
<td><strong>Haptic</strong></td>
<td></td>
</tr>
<tr>
<td>Electrodermal</td>
<td>Electric Shocks [259], Electric Arcs [342], Electric Stimuli [283]</td>
</tr>
<tr>
<td>EDA, Facial Features</td>
<td></td>
</tr>
<tr>
<td><strong>Cutaneous</strong></td>
<td></td>
</tr>
<tr>
<td>Tactile</td>
<td>Shape Changing, Tactile Click Feedback [174], Texture Modulation [174], Pin-array Displays [174], Ultrasound Haptic Vibration, Vibro-tactile Tactons [47], Tactile Feedback Liquid [304], Air Jet [355]</td>
</tr>
<tr>
<td>Touch, Pressure [347]</td>
<td></td>
</tr>
<tr>
<td><strong>Thermal</strong></td>
<td>Changes in Temperature, Thermal Feedback Liquid [304]</td>
</tr>
<tr>
<td>Skin Temperature [389], Body Temperature [243]</td>
<td></td>
</tr>
<tr>
<td><strong>Pain (Nociception)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Olfactory</strong></td>
<td>Scents</td>
</tr>
<tr>
<td>Body Odor [14], Breath Smell [14]</td>
<td></td>
</tr>
<tr>
<td><strong>Gustatory</strong></td>
<td>Edible User Interface, Edible Marker [265], Pseudo-Gustatory Display [266], Electrical stimulation on the tongue [207]</td>
</tr>
<tr>
<td><strong>Cerebral</strong></td>
<td>Brain-Computer Interface, Brain Chip</td>
</tr>
<tr>
<td>Cortical Activity, Slow Cortical Potentials [144], EEG or INIRS [104], Brain Signals [18], Brain-Computer Interface</td>
<td></td>
</tr>
<tr>
<td><strong>Cardiac</strong></td>
<td>Defibrillator</td>
</tr>
<tr>
<td>Heart Rate Variability, Heart Rate, Blood Pressure [389], Carotid Artery</td>
<td></td>
</tr>
<tr>
<td><strong>Nomadic Device</strong></td>
<td>Smartphone, Tablet, Laptop Smartwatch, On-body Device, Wearable</td>
</tr>
<tr>
<td>Smartphone, Tablet, Laptop Smartwatch, On-body Device, Wearable</td>
<td></td>
</tr>
</tbody>
</table>
### Table 5. Search queries based on Boolean algebra used in the SLR covering four contexts relevant to our research topic including the results count for each queried database (ACM DL, IEEE Xplore, and ScienceDirect).

<table>
<thead>
<tr>
<th>(1) Automotive Context</th>
<th>(2) Interaction Context</th>
<th>(3) Sensor/Actuator Context</th>
<th>(4) Input and Output Modalities Context</th>
<th>Resulting Query (1 AND 2 AND (3 OR 4))</th>
<th>ACM</th>
<th>IEEE</th>
<th>ScienceDirect</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Visual&quot; OR &quot;Eye&quot; OR &quot;Pupillary&quot; OR &quot;Pupil&quot; OR &quot;LED&quot; OR &quot;Display&quot; OR &quot;Screen&quot; OR &quot;Monitor&quot; OR &quot;Augmented Reality&quot; OR &quot;Virtual Reality&quot; OR &quot;Haptic Reality&quot; OR &quot;Robotics&quot; OR &quot;Shape&quot; OR &quot;Texture&quot;</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>(visual OR eye OR pasa OR eye-color OR pasa-color OR led OR display OR screen OR monitor OR &quot;augmented reality&quot; OR &quot;virtual reality&quot; OR &quot;haptic reality&quot; OR robotic OR shape OR texture)</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>614</td>
<td>105</td>
<td>123</td>
<td></td>
</tr>
<tr>
<td>&quot;Auditory&quot; OR &quot;Voice&quot; OR &quot;Talk&quot; OR &quot;Ear&quot; OR &quot;Speech&quot; OR &quot;Voice&quot; OR &quot;Verbal&quot; OR &quot;Vocal&quot; OR &quot;Non-Speech&quot; OR &quot;Sound&quot; OR &quot;Whistling&quot; OR &quot;Humming&quot; OR &quot;Acoustic&quot; OR &quot;Tactile&quot; OR &quot;Haptic&quot; OR &quot;Touch&quot;</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>(auditory OR voice OR talk OR ear OR speech OR voice OR verbal OR vocal OR non-speech OR sound OR whistling OR humming OR acoustic OR tactile OR haptic OR touch)</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>205</td>
<td>136</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>&quot;Tactile&quot; OR &quot;Tactile&quot; OR &quot;Touch&quot; OR &quot;Pressure&quot; OR &quot;Shape&quot; OR &quot;Texture&quot; OR &quot;Vibration&quot; OR &quot;Vibrotactile&quot; OR &quot;Liquid&quot; OR &quot;Air&quot; OR &quot;Kinesthetic&quot; OR &quot;Kinesthetic&quot; OR &quot;Lever&quot; OR &quot;Joystick&quot; OR &quot;Button&quot; OR &quot;Gesture&quot; OR &quot;Movement&quot; OR &quot;Posture&quot; OR &quot;Pose&quot;</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>(tactile? OR touch? OR pressure OR shape OR texture OR vibration OR vibrotactile OR liquid OR air OR kinesthetic OR lever OR joystick OR button OR gesture OR movement OR posture OR pose)</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>494</td>
<td>52</td>
<td>123</td>
<td></td>
</tr>
<tr>
<td>&quot;Electrodermal&quot; OR &quot;Electrical&quot; OR &quot;Facial&quot; OR &quot;Facial&quot; OR &quot;Facial&quot; OR &quot;Facial&quot; OR &quot;Facial&quot; OR &quot;Facial&quot; OR &quot;Facial&quot; OR &quot;Face&quot; OR &quot;Head&quot;</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>(electrodermal OR electrical OR facial OR face OR head)</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>97</td>
<td>120</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>&quot;Thermal&quot; OR &quot;Temperature&quot; OR &quot;Heating&quot; OR &quot;Cooling&quot; OR &quot;Heat&quot; OR &quot;Liquid&quot;</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>(thermal OR temperature OR heating OR cooling OR heat OR liquid)</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>119</td>
<td>106</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>&quot;Pair&quot; OR &quot;Pair&quot; OR &quot;Receipt&quot;</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>(pair OR receipt)</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>&quot;Visibility&quot; OR &quot;Visibility&quot; OR &quot;Sight&quot; OR &quot;Sight&quot; OR &quot;Sight&quot; OR &quot;Sight&quot; OR &quot;Sight&quot; OR &quot;Sight&quot; OR &quot;Sight&quot; OR &quot;Vision&quot; OR &quot;Mirror&quot;</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>(visibility OR sight OR mirror)</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>52</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>&quot;Computer&quot; OR &quot;Computer&quot; OR &quot;Interface&quot; OR &quot;Interface&quot; OR &quot;Interface&quot; OR &quot;Interface&quot; OR &quot;Interface&quot; OR &quot;Interface&quot; OR &quot;Interface&quot;</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>(computer OR interface)</td>
<td>(vehicle? OR car? OR driver? OR driving OR &quot;in-vehicle&quot;) AND interaction? OR interface?</td>
<td>18</td>
<td>15</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>
### C RESULT COUNTS OF QUERIED VENUES

Table 6. Selected venues and number of publications found after applying the SLR search queries.

<table>
<thead>
<tr>
<th>Conference / Venue</th>
<th>Number of publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACM Conference on Human Factors in Computing Systems (CHI)</td>
<td>40</td>
</tr>
<tr>
<td>ACM Conference on Computer-Supported Cooperative Work &amp; Social Computing (CSCW)</td>
<td>0</td>
</tr>
<tr>
<td>ACM Conference on Pervasive and Ubiquitous Computing (UbiComp)</td>
<td>2</td>
</tr>
<tr>
<td>ACM/IEEE International Conference on Human Robot Interaction (HRI)</td>
<td>1</td>
</tr>
<tr>
<td>ACM Symposium on User Interface Software and Technology (UIST)</td>
<td>3</td>
</tr>
<tr>
<td>ACM Transactions on Computer-Human Interaction (TOCHI)</td>
<td>1</td>
</tr>
<tr>
<td>ACM International Conference on Multimodal Interaction (ICMI-MLMI)</td>
<td>7</td>
</tr>
<tr>
<td>ACM Designing Interactive Systems Conference (DIS)</td>
<td>6</td>
</tr>
<tr>
<td>ACM International Conference on Intelligent User Interfaces (IUI)</td>
<td>8</td>
</tr>
<tr>
<td>ACM International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI)</td>
<td>9</td>
</tr>
<tr>
<td>ACM International Conference on Tangible, Embedded, and Embodied Interaction (TEI)</td>
<td>1</td>
</tr>
<tr>
<td>ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutoUI)</td>
<td>122</td>
</tr>
<tr>
<td>IEEE Transactions on Affective Computing</td>
<td>0</td>
</tr>
<tr>
<td>IEEE Transactions on Human-Machine Systems</td>
<td>2</td>
</tr>
<tr>
<td>IEEE Transactions on Haptics</td>
<td>2</td>
</tr>
<tr>
<td>IEEE Conference on Virtual Reality and 3D User Interfaces (VR)</td>
<td>1</td>
</tr>
<tr>
<td>IEEE Transactions on Intelligent Transportation Systems (ITS)</td>
<td>12</td>
</tr>
<tr>
<td>IEEE Transactions on Vehicular Technology</td>
<td>1</td>
</tr>
<tr>
<td>IEEE Transactions on Intelligent Vehicles</td>
<td>0</td>
</tr>
<tr>
<td>IEEE Intelligent Vehicles Symposium (IV)</td>
<td>24</td>
</tr>
<tr>
<td>IEEE International Conference on Intelligent Transportation Systems (ITSC)</td>
<td>18</td>
</tr>
<tr>
<td>IEEE Access</td>
<td>8</td>
</tr>
<tr>
<td>International Journal of Human-Computer Studies</td>
<td>5</td>
</tr>
<tr>
<td>Transportation Research Part F: Traffic Psychology and Behaviour</td>
<td>32</td>
</tr>
<tr>
<td>Applied Ergonomics</td>
<td>22</td>
</tr>
<tr>
<td>Robotics and automated Systems</td>
<td>0</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td><strong>327</strong></td>
</tr>
</tbody>
</table>
D  CONCEPT IMAGES OF IN-VEHICLE INTERACTIONS

Fig. 12. Concept images for in-vehicle input modalities used in our online study.

Fig. 13. Concept images for in-vehicle output modalities used in our online study.
Fig. 14. Concept images for input at nomadic (1-4) and anchored (5-10) in-vehicle locations used in our online study.
Fig. 15. Concept images for output at nomadic (1-4) and anchored (5-10) in-vehicle locations used in our online study.