eHMI Design: Theoretical Foundations and Methodological Process

Y. Shmueli and A. Degani
General Motors R&D, Israel

Keywords: External HMI (eHMI), Multiple Resources Theory, Stimulus-Coding-Response Compatibility Principle.

Abstract: In the last decade, substantial efforts have been dedicated to the problem of pedestrian's encounter with driverless autonomous (L-4/5) vehicles. Different communication schemes, involving different design concepts, modalities, and communication formats have been conceived and developed to communicate and interact with pedestrians. It is expected that only a limited subset of these options, perhaps only one, will be selected as an international standard (with some allowance for branding and adaptations to different cultural norms and expectations). Naturally, the selection of the communication scheme has to rely on a valid theoretical foundation, not only to satisfy automotive regulatory agencies, but also as a precursor to a similar communication scheme for robots in the public space. In this paper, we provide an eight-step process which supports the development of an effective communication design. We use Wickens' (1984, 2002) Multiple Resources Theory (MRT), as the theoretical foundation for our work, and the Stimulus Coding Response (S-C-R) compatibility principle (Wickens et al. 1984) as an organizing principle for eHMI design.

1 INTRODUCTION

The transition from manual driving to fully autonomous driving requires a shift in our conceptualization concerning the interaction between vehicles and other road users. Currently, pedestrians primarily use implicit motion-inherent cues such as the vehicle's speed and distance, deceleration rates, and braking profile to anticipate the behavior of vehicles on the road and make the crossing decision (Cohen et al., 1955; Lehsing et al., 2019; Domeyer et al., 2020). In addition to these physical cues, non-verbal explicit cues such as eye contact, head nodes and gestures help pedestrians to interpret the situation and support the establishment of trust between the pedestrian and the driver (Šucha, 2014, Rasouli et al., 2017; Gueguen et al., 2015; Ren et al., 2016; Lagström and Malmsten Lundgren, 2015; Dey, 2021).

Naturally, the absence of a driver in autonomous vehicles precludes the possibility of any communication. As such, autonomous vehicle technology requires additional features that will allow the public to interact with such vehicles and perceive it as safe and accommodating.

Sophisticated interfaces can be devised to substitute for the missing pedestrian-driver interaction and may even achieve higher reliability than the current signaling methods (e.g., use of headlights, hazard lights, and the car horn). These communication methods are naturally idiosyncratic and at times quite ambiguous. For example, does the driver's use of the headlights means that he or she is taking the right of way or giving it to the pedestrian?

The question is how to use this opportunity of the forthcoming need for external HMI (eHMI) signaling communication to not only substitute the driver but to make such communication better. At its most basic, the communication between a robotic agent and a human pedestrian should meet four main requirements: (i) Effectiveness – establish the necessary communication between pedestrians and autonomous vehicles, (ii) Efficiency – be simple, intuitive, and non-intrusive, (iii) Acceptability – form public "trust" in this new technology, (iv) Satisfaction – be elegant, induce comfort, and invoke a rewarding experience.

We propose a step-by-step process for the verification and synthesis of eHMI design solutions to fulfil these requirements: Step 1. Requirements' derivation based on initial conceptual analysis, Step 2. Requirements' derivation based on an empirical needs study, Step 3. Proposal of a generic communication protocol, Step 4. Content selection for eHMI displays, Step 5. Allocation of the selected content to media and modalities, Step 6. Media

Shmueli, Y. and Degani, A.
eHMI Design: Theoretical Foundations and Methodological Process.
DOI: 10.5220/0011686600003417
In *Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2023) - Volume 2: HUCAPP*, pages 201-212
ISBN: 978-989-758-634-7; ISSN: 2184-4321

realization - Representation solutions within different media, Step 7. Verification of design proposals, and Step 8. Validation of design solutions.

2 METHODS

2.1 Initial Conceptual Analysis

We begin by analysing generic interaction patterns. This analysis lists the potential touchpoints between a vehicle (with a driver or without) and a pedestrian in a pedestrian crossing scenario. Framing the interaction between the two as a dialogue will enable us to pinpoint generic user needs. We outline potential actions made by the pedestrian (in aquamarine) and those that can be made by the vehicle (in black). Pedestrian's states are marked by numbers and the vehicle's states are marked by letters:

The vehicle is driving, approaching the scene [a]. The pedestrian is walking (0), facing (or with his back to) the approaching vehicle. The person may look or glance at the vehicle before reaching the curb, communicating the message: "I can see you" (1). If the pedestrian is not planning to cross, his/her body movements, posture and facial expressions will convey the message "I am not crossing" (2). If, on the other hand, there is an intent to cross, the body movements, posture and facial expressions will convey the message "I am about to cross" (3). This is where people seek a confirmation from the vehicle ["I intend to stop (for you)" - b] (Rasouli et al, 2017; Habibovic et al. 2018; Dey 2021). While the vehicle is slowing down, it conveys the message ["I am slowing down and stopping" – c], or ["I am not stopping – d] if it cannot stop on time.

The intent to stop is primarily communicated by the deceleration profile of the vehicle that should be sharp enough to be easily recognizable (Lehsing et al., 2019); a vehicle that stops short of a crosswalk can be interpreted as yielding for a pedestrian and not simply responding to traffic signage (Risto et al., 2017; Domeyer et al., 2020). Leaving the curb, the pedestrian's message progresses into "I am on the pavement, starting to cross" (4). From the stopped vehicle's perspective, the message changes to [I have stopped **for you** – e]. While crossing the road, the pedestrian may portray various messages using implicit and explicit modes of communication. These will vary from a short glance (or no glance at all) to explicit hand gestures and head movements, claiming the space (Rasouli et al, 2017).

Vulnerable pedestrians may have special requirements. Consider for example, people with mobility, cognitive or perception impairments (Xiang et al., 2006), or age-related difficulties expressing vulnerable body language (5). Using eye contact and a reassuring facial expression, the driver in the vehicle may enhance the sense of safety for slow pedestrians crossing the street. The message for the normal population: ["I am giving you the right of way" – f], may add an extra sense of patience and protection to slow pedestrians such as elderly or disabled users: ["I am respecting your space and will not act against you"- g], ["I am also looking around you by being attentive to the surroundings and other vehicles that may infringe this space" - h]. Upon the completion of the crossing act, the transaction ends from the perspective of the pedestrian - "Bye" (6). The vehicle in return, can communicate its intent to leave ["I am leaving"- i]. This is not that important for the pedestrian who has completed the crossing act but may be useful for those that are planning to start crossing the road while the vehicle is in a stop position. Starting to drive ends the transaction ["Bye, I am starting to drive" - j].

There are several variations to the above sequence, however, in all variations the pedestrian, who is more vulnerable, has priority over the autonomous vehicle. The sequence will be similar if the pedestrian is standing at the edge of the curb, waiting for the vehicle to stop and yield. However, when the pedestrian is already crossing the street while the vehicle is approaching the scene, the sense of vulnerability is the greatest and pedestrians tend to establish an eye contact with the driver as the vehicle gets closer (Dey et al., 2019).

2.2 Needs Study

This conceptual analysis is followed by a Wizard of Oz (WoZ) needs and concerns study, which focuses on the encounter between pedestrians and a fake autonomous vehicle. We aim to identify what people actually expect from autonomous vehicles in the public space and what would make them feel more at ease and accepting of this technology.

One of the most interesting questions is how can grounding be established in the absence of a human driver, what elements would be missing? Shmueli & Degani (2019) conducted a naturalistic study to get a clearer understanding of pedestrians' needs and reactions during mundane, non-urgent, crossing scenarios. This study was conducted at General Motors campus where employees experienced an encounter with an autonomous vehicle driven manually. As an alternative to the GhostDriver technique (Rothenbücher et al., 2016), different methods were used to eliminate eye contact with the safety driver and minimize pedestrians' expectations of potential interventions by the safety driver: the driver was wearing a helmet and avoided head movements and hand gestures that could have implied pedestrians that they had been noticed. In addition, the driver was instructed to keep his hands low on the steering wheel and apply a more robotic driving style. People encountered the vehicle and were then stopped by the researchers for a short interview to better understand their experience during the encounter. No specific communication language was used.

30 encounters of 36 pedestrians (7 females and 29 men) were documented. In-depth interviews that were conducted following the encounter with the vehicle provided insights into the participants' intellectual and emotional confusion, fears and wishes. Users' comments were collected and classified into the following categories: (i) Intent to stop & Yield, (ii) Wait & Give way, and (iii) Intent to drive. Consistent with the literature (Rasouli et al, 2017; Habibovic et al. 2018; Merat et al., 2018; Dey, 2021), the autonomous vehicle's "intent to stop & yield" was identified as the main user need with 32 comments: the interviews yielded 10 comments regarding the wish for a general notification, 14 comments specified the need to be personally acknowledged by the vehicle: "*That the vehicle is planning to stop for ME*", 6 comments indicated the wish to get guidance and warnings from the vehicle, and 2 comments described a sophisticated dialogue that can be formed with the vehicle. In contrast, only 6 comments mentioned the need for a continuous indication while the autonomous vehicle is in a stop position. Some participants mentioned that other people, especially slower users (elderly people, people with young children, or people who suffer from some form of disability) may require to be acknowledged also while crossing. Finally, 10 comments cited the need for a dedicated indication when the vehicle resumes driving for the reasons outlined in the Analysis section. Overall, the conceptual analysis and the empirical study suggest that some reassurance is needed once driverless autonomous vehicles are introduced in the public space and that some people are likely to require more reassurance than others.

2.3 Protocol: Representation

Based on the conceptual analysis (in section 2.1), and a better understanding of how people respond to the technology in context (in section 2.2), we derived a set of general communication needs. The set, presented in Figure 1, is more comprehensive than the three main touchpoints discussed in the previous section.

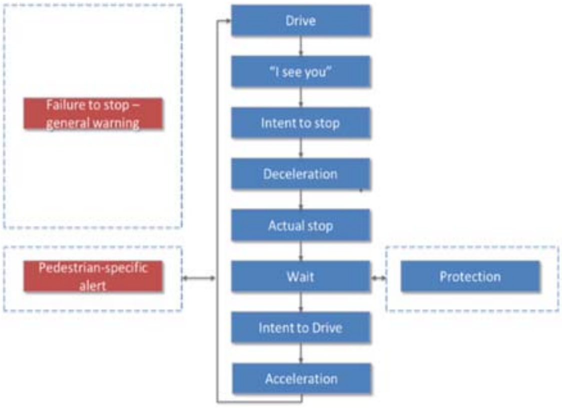

Figure 1: Communication protocol.

We have expanded it to include other needs such as acknowledgement from a distance ("I see you") and deceleration and acceleration that allow for the assessment of a more complete design sequence. The ability to augment the dynamic profile of the vehicle by eHMI design is intuitively promising and will be further discussed in Section 2.6. We also included a "protection" state for pedestrians with special needs and for those who are more vulnerable. These touchpoints define the full set of information content that can be portrayed to the pedestrian.

Now comes the question of what kind of format will each communication touch point employ? The first phase in the design process is **content selection** ("what to convey"); the second is **media allocation** ("which medium – and hence, modality - to say it in") modality; the third, **media realization** ("how to say it in that medium; how to design the content?") and the fourth **media coordination** (``how to coordinate several media") (Maybury, 1993).

2.4 Content Selection for eHMI Displays

The method starts by classifying the spectrum of contents that can be communicated by the autonomous vehicle into 7 different information types that were identified by ISO (ISO/TR 23049): Mode, State, Perception, Recognition & Acknowledgement, Belief state, Intent, Guidance. Car manufacturers may differ in the information types that they wish to include in their communication language. The selected solution may depend on advances in sensing capabilities and on available hardware as these may limit the safety area that one can guarantee. In addition, various policies and international standards may serve as a filter with regards to what to present to guarantee safety. Guidance for example - The autonomous vehicle can act as a semaphore (e.g., like a traffic light) providing supervisory control to road users. Providing full guidance involves great responsibility (and potential liability issues since, at the moment, car manufacturers are not equipped with all traffic information). Therefore, it is recommended to avoid communicating that it is safe to pass, so long that we cannot guarantee the safety of the pedestrians from all sides of the road. Other considerations may also involve planning constraints, the wish to avoid visual clutter, or simply cost.

2.5 Media Allocation: Codes and Modalities

The next consideration is how the information selected should be represented and presented. We propose Wickens' Multiple Resources Theory (MRT) (1984) as a framework for making representation decisions. The MRT asserts that people have a limited set of resources available for mental processes, from sensing to response execution. It consists of four dimensions: (1) stages of attentional processing (2) processing codes, (3) perceptual modalities, and (4) response execution.

1. Resources used for perceptual and cognitive activities, are shared, and are functionally separate from those underlying the selection and execution of responses.
2. Spatial activity uses different resources than verbal/linguistic activity, as evident by working memory studies (Baddeley, 1986) and action studies (e.g., speech vs. manual control; Liu & Wickens, 1992; Wickens & Liu, 1988).
3. Auditory perception uses different resources than visual perception.
4. Manual and vocal reactions rely on separate resources.

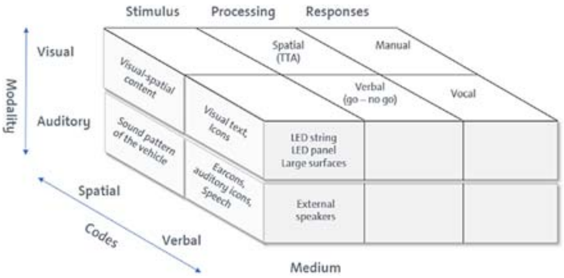

Figure 2: Multimodal Design Space.

In 2002, Wickens added an important qualification regarding foveal and peripheral vision that can take place concurrently: Focal vision primarily, but not exclusively foveal, supports object recognition and in particular high acuity perception such as that involved in reading text and recognizing symbols. Ambient vision distributes across the entire field and (unlike focal vision) can preserve its competency in peripheral vision. Ambient vision is responsible for perception of orientation and movement, for tasks such as those supporting walking upright in targeted directions or lane keeping in the highway (Horrey et al., 2006). This qualification is of a great importance to the discussion of the potential augmentation of the ambient perception of vehicle movement by an external communication language which may involve focal vision.

Figure 2 outlines the model which is based on the theory. The processing stages appear on the X axis, the modalities appear on the Y axis and the codes appear on the Z axis. Following the MRT, media allocation in eHMI design can be framed in terms of Codes (verbal & spatial) and Modalities (visual & auditory). In visual displays, communication may rely on symbolic representations: visual text, icons, and/or on visual-spatial information. In auditory displays, it may rely on speech and/or non-verbal (abstract) earcons and auditory icons (everyday sounds), but also on spatial cues via the engine sound in internal combustion engine vehicles, or the external sound system in electric vehicles.

The method proposes that the assignment of information into codes and modalities should be based on the pedestrians' dynamic needs throughout the interaction with the autonomous vehicle and the available physical platform. Clearly, the presentation of text and icons require a minimal display height, so that the observer will be able to see it from a distance; a single LED strip/band would not be suitable for this purpose. Nevertheless, the compatibility of the design with the processing requirements of the different touchpoints during the encounter with the vehicle is a central aspect and may affect the selected platform. The protocol presented in section 2.3 can be therefore fragmentized into sub-tasks that fall into one of the following processing codes: verbal (conceptual) or spatial, or both. According to the Stimulus-Coding-Response (S-C-R) compatibility principle of the MRT (Wickens et al. 1984), tasks with verbal central-processing demands will be best served by voice input and output channels, but verbal-visual contents such as visual text and icons will also produce an effective resources usage due to their compatibility with the processing demands. Similarly, tasks with spatial demands will be best served by visual-manual channels but may benefit also from spatial auditory information. Hence, code representations for each sub-task determine the optimal stimulus type and response code.

2.6 Media Realization: eHMI Design Requirements

eHMI design should be artistically pleasing but most importantly, be semantically appropriate. Wickens' MRT is highly useful when seeking logic, specifically when considering compatibility of the perceptual content with the processing requirements of the task. Next, we translate the communication protocol into spatial and conceptual sub-tasks.

2.6.1 Driving Mode: Conceptual

Some autonomous vehicles will have a unique appearance which cannot be mistaken with a manually driven vehicle, while other vehicles may maintain a more traditional look. The determination of driving mode is primarily a conceptual "verbal" sub-task that should be made from a distance. In terms of media allocation & realization, this calls for a verbal-conceptual visual solution. To support visibility from a distance, large visual text and icons would require large platforms that may induce visual pollution. A more minimalistic solution can be used to meet the Efficiency criterion. The proposal made by SAE J3134, a marker lamp in cyan that will indicate on a single LED band, or a larger visual panel the autonomous driving status of the autonomous vehicle, could be considered. At the auditory level, a unique sound language for electric autonomous vehicles may promote their recognition by visually impaired users.

2.6.2 "I See You": Conceptual (Potentially Spatial)

Understanding that the vehicle notices the pedestrian is a conceptual task with a spatial component. As noted earlier, pedestrians want to know that the vehicle detects them. As for media allocation & realization, a mild cue which reflects the spatial location of the pedestrian, before an intent to stop is broadcasted, may be a useful cue if it can be perceived from a distance.

2.6.3 Intent to Stop & Yield: Conceptual (Potentially Spatial)

Understanding the vehicle's intent to stop is a conceptual task that currently relies mainly on its deceleration profile. This is not always a sufficient cue (Šucha et al., 2017; Schieben et al., 2018; Klatt et al., 2016; Ren et al., 2016). At any rate, the intent to stop & yield should attract the visual attention of the pedestrians who may be located 40-60 meters away. The visual effect should be carefully planned, it should be clear yet non-distractive to avoid visual capture by pedestrians and other drivers. Another challenge is communicating the intent of the vehicle without providing an explicit guidance.

2.6.4 Deceleration: Spatial

Assessing the autonomous vehicle's deceleration rate is a crucial element in the pedestrian's ability to assess the Time to Arrival (TTA) of the vehicle. The assessment of TTA is primarily a spatial task; people rely on their visual (foveal and peripheral) and auditory perception to perform it. Motion cues and users' trajectories can be used by automated driving systems to promote better integration into the traffic environment (Domeyer, 2020). However, studies have shown that estimates based on motion can be biased by occluding objects and by the vehicle speed. Vehicle size can also affect these estimates, with larger vehicles being estimated to arrive earlier (Caird and Hancock, 1994; Delucia, 1991a, 1991b). In addition, factors associated with capabilities of other road users have also been identified as affecting crossing decisions. For example, older adults and young children suffer from a reduced ability to estimate TTA (Andersen & Enriquez, 2006; Dommes & Cavallo, 2011).

There is a potential to strengthen the understanding of the pedestrian, especially in low-speed situations where vehicle kinematics may be ambiguous (Domeyer et al. 2020) and avoid perceptual mistakes by using an explicit signal. This may assist the ambient vision in poor visibility conditions (poor weather, poor lighting), help pedestrians to correct perceptual biases, and assist pedestrians with restricted visual and hearing capabilities. The analysis of the MRT calls for a visual-spatial media solution rather than a symbolic one. A visual-spatial animation that implies on speed reduction seems advisable.

2.6.5 Actual Stop: Conceptual

This state reflects the situation where the vehicle is approaching 0 MPH, after which it enters the wait state and has a strong conceptual component. This touchpoint can be communicated visually and auditorily by verbal or symbolic means.

2.6.6 Wait: Conceptual (Potentially Spatial)

In this state, the design aims to replace the driver's eye contact with the pedestrian. While some people may trust that the vehicle will stay stopped once it has reached 0 MPH, others may need extra assurance regarding its intent to let them cross the road. The solution must be intuitive and easy to understand across cultures and age groups. Nevertheless, communicating that the vehicle will wait patiently is a conceptual message that should not be confused with an explicit positive guidance to cross.

Naturally, a conceptual solution that communicates patience may involve symbols or an abstract, universal, animated effect. An additional spatial component can be optionally added; a spatial tracking of the pedestrian's reflection may replace the driver's eyes that follow the pedestrian while the vehicle is waiting (See various design implementations of this principle by Nissan, Mercedes Benz: F015 & The cooperative vehicle, AutonoMI, Volvo 360C). The pedestrian's reflection tracks the spatial position of the person/s in real-time. Their perception can be conveyed by an animated effect on the lighting fixture confirming recognition before, at the beginning of the crossing, or during the act of crossing, when they feel the most vulnerable.

2.6.7 Intent to Drive: Conceptual

This message is essential for those who want to start crossing after the agent decided that it is leaving. The 'Intent to drive' precedes the acceleration state (spatial) but it is nonetheless effectively a conceptual message, or task. An 'intent to drive' can be communicated visually using an attention-grabbing visual animation, supplemented by an auditory cue to accommodate visually impaired pedestrians, or by verbal or symbolic means.

2.6.8 Acceleration: Spatial

Similar to deceleration, acceleration is a spatial task, and any visual and auditory spatial representation of speed increment should refer to the solution provided for the deceleration state. In addition, the visual animation could potentially be synchronized with the overall pace of speed increment of the vehicle.

2.6.9 Failure to Stop Warning: Conceptual

Warnings are conceptual in nature and are directed toward all road users in the vicinity, perhaps in an omnidirectional manner. This calls for an auditory solution. There is a possibility to enhance the auditory warning by a compatible visual animation that shares temporal and spectral characteristics with the sound.

2.6.10 Pedestrian-Specific Warning: Spatial

In situations that warrant warning of specific pedestrians, where the information is directed toward a specific area, the auditory warning may have a directional spatial nature.

2.6.11 "Protection": Both Conceptual & Spatial

A sophisticated eHMI can provide special information, in particular to those in need (e.g., children, elderly users, disabled road users) and those that seem apprehensive or reluctant to cross. The vehicle can send them a message about its commitment to yield and protect while they are crossing. Just as in the case of the pedestrian-specific acknowledgement, the realization of this requirement may have a directional, spatial nature and can be coupled with a visual tracking design solution. The animation itself should be designed so as to convey the fact that the vehicle is waiting patiently and will remain stopped until the user completes the task of crossing the road. A very subtle pulsation of the entire display in case of a non-pedestrian-specific design, or the pedestrian/s' reflection in case of a pedestrian-specific design, may boost the pedestrian's confidence.

Having summarized the resources requirements and input of each touchpoint, it is important to note that an elegant design requires some form of continuity. Subtle and sophisticated animated transitions may be required while shifting between different phases.

2.7 Verification

We now wish to verify this proposed method by looking into several design concepts that have been published in the literature and assess if they meet the four criteria listed in section 1: Effectiveness, Efficiency, Acceptability & Satisfaction. One interesting example is the pioneering study conducted by Clamann, Aubert & Cummings (2016) at Duke University. This study aimed to assess the effect of display content on the participants' decision to cross. Two display types were used: (i) an advisory (guidance) display that consisted of a 'Don't Walk' symbol while the vehicle was in motion (this corresponds to all touchpoints except the *Wait* state and the *Intent to drive* in the protocol of section 2.3), and that switched into a 'Walk' symbol when the vehicle came to a stop, and (ii) an information display portraying the vehicle's speed by means of dynamic digits. In both cases, the designs relied on visual-verbal coding that violates the spatial cognitive processing requirements of the vehicle's motion cues in the deceleration assessment task. The two display formats failed to facilitate pedestrians' decision to cross the road. Clamann et al. (2016) study is often cited as a proof of eHMI failure in producing improvements in road users' comprehension of the autonomous vehicle's intention and the proposed account provides an alternative explanation for the results of this study.

The automated vehicle interaction principle (AVIP) project which was developed with the Swedish Victoria ICT (2015) provides a more comprehensive design (Lagström & Malmsten Lundgren, 2015, Habibovic et al., 2018). The visual display consists of a LED band above the windshield. The design represents the vehicle's autonomous *driving mode* using a centralized light in the centre of the screen (a conceptual solution). The vehicle's *Intent to stop* is communicated by the expansion of this light, which continues to expand further while the vehicle is decelerating (a visual-spatial solution that merges the conceptual *Intent to stop* and the spatial *Deceleration* state). This visual expansion of light is compatible with the effect of looming, or the visual enlargement of an object as it approaches the viewer). The light reaches its maximal size when the vehicle is *Stopping*, and subtle pulsation of the full band, imitating human breathing rhythm, communicates the fact the vehicle is *Waiting* for the pedestrian to cross the road (a conceptual solution that does not acknowledge specific pedestrians). When the vehicle *Intends to drive*, the light converges back to the centre in a smooth animation. There is neither a representation of *Acceleration*, nor a representation of *Warning* information. This eHMI design is partially consistent with our method and was successful in inducing a sense of safety and improved confidence in the vehicle's automation technology (Lagström & Malmsten Lundgren, 2015; Habibovic et al., 2018). In 2018, Habibovic et al. mentioned that several pedestrians stated that the pulsating light during the *Wait* state was not contributing to their experience and suggested that it could be removed to make the interface easier to understand.

Interestingly enough, the European project INTERACT used the same breathing light metaphor to communicate the vehicle's *Intent to stop and yield* and the *Wait* state, merging these two conceptual elements with the spatial *Deceleration* state that has no unique representation in this design concept. The only spatial representation in this implementation is facilitated by a separate tracking lamp (by HELLA) to acknowledge specific pedestrians crossing the street. Finally, the vehicle's *Intent to drive* is communicated by the full light band which pulsates quickly a few times, coupled with an auditory cue. Similar to the AVIP concept, this eHMI concept was found to increase participants' comprehension of the vehicle's intention and elevated their level of trust toward the automated driving technology. The two concepts and hence meets the Effectiveness, Efficiency and Acceptability criteria.

To summarize, these three examples show that as long as there is no strict violation in the compatibility between the design and the required processing code, concepts meet the Effectiveness, Efficiency and Acceptability criteria. A beneficial effect of the eHMI design is identified: An increased efficiency, along with increases in perceived safety, comprehension, and trust.

2.8 Validation

After we filter out incongruent solutions, still there remains an abundance of eHMI design solutions that are theoretically valid. It is hard to evaluate in advance which solution will prove better than the others. Therefore, there is a need to contrast them empirically using controlled testing methods (such as video analyses and VR) in laboratory conditions and then, externally, in safe test tracks and natural road context.

Currently, the literature does not provide a full answer regarding the best, complete, eHMI solution. There are quite a few examples of validation attempts of specific concepts (AVIP - Lagström & Malmsten Lundgren, 2015; Habibovic et al., 2018; The Mercedes Benz's Cooperative vehicle - Faas & Baumann, 2019; Ford concept vehicle - Hensch et al., 2019b, to name a few). This line of research sheds light on the intuitiveness and comprehensiveness of specific solutions but does not tell us which representation strategy is superior to others. On the other hand, comparative assessments (Ackermann et al., 2019; de Clercq et al., 2019; Fridman et al. 2017; Dey et al, 2020) have been useful in revealing the relative intuitiveness of specific designs, specifically regarding the *Intent to stop & yield,* and the *Wait* state. However, apart from a partial attempt by Dey et al. (2020), who tested abstract visualizations, no filters have been applied on the Guidance information-type mentioned earlier. These studies contrast abstract lighting solutions with colour-based and icon-based traffic lights solutions, and with textual messages such as "After you", "Go ahead", or

"Safe to cross". As people seek to reduce ambiguity, they show preference towards non-ambiguous solutions that may put them at risk.

A more appropriate contrast should involve states and intents communication designs that are lacking any component of Guidance. Different visual-abstract *Intent to stop & yield* designs, with and without the visual-spatial *Deceleration* component, should be contrasted. In addition, the understanding of the optimal communication solution for the *Wait* state should be based on contrasting several visual-conceptual solutions that communicate that the vehicle will wait patiently, with and without a spatial tracking component that provides acknowledgement of specific pedestrians. A state-informing textual message, such as "Waiting" may be added to the comparison, in spite of the fact that it necessitates reading capabilities, and, as noted by Dey (2021), if the message is partially occluded, for example the "ing" part, both agents will wait for the other to act and traffic-flow would suffer.

As a final note, even when reverting to novel cross-cultural, abstract, visual-conceptual solutions, the issue of guidance emerges. For example, Shmueli & Degani (in preparation) conducted a small, controlled, study to assess the meaning assigned to a central-pulsating light in a fixed position vs. a pulsating tracking-light. Participants perceived the pulsating tracking-light as their reflection: "*the vehicle sees me… it recognizes me… I feel safe*". This did not make them assume that the vehicle is responsible for the road space beyond them. On the other hand, when the light was pulsating in the centre of the lighting display, one participant commented that he would quickly learn to interpret it as a green light that guides to cross the road. We believe that this desire to receive a semaphore signal is common to many people; The tendency to find positive guidance in neutral solutions should be thoroughly assessed.

An additional aspect that our method cannot fully predict is the optimal representation of deceleration information by abstract visual means. Concepts that try to tap the deceleration state of the vehicle vary significantly: light bands diverge (AVIP; Bumper PB eHMI in Dey et al., 2020b), converge (Volvo 360C), descend (Shmueli & Degani in preparation). Each has its own logic for the achievement of the right semantics of the direction of motion of the approaching vehicle:

- The movement compatibility principle (Warrick, 1947), determines that the direction of movement in the display should be consistent with the direction of movement of the approaching vehicle, suggesting some superiority for flow to the centre (converging light band)
- Looming (Lee, 1976), or the visual expansion of an object as it approaches, calls for the expansion of the light band (diverging light band)
- The conceptual compatibility with speed reduction calls for a descending light band

Further research is needed to determine (i) which visualization conveys best speed reduction information and (ii), the effect of synchronicity of the animated effect with the deceleration profile of the vehicle and with the electric vehicle's external sound output.

3 DISCUSSION

An important element in the process discussed above is Wickens' MRT and S-C-R Compatibility principle that are applied to step 5 – Media Allocation – Codes and Modalities. The theory can be used to identify the type of codes that are most appropriate for each type of communication. Designers can optimize their design solution if they meet the nature of the codes that underly various touchpoints between human users and autonomous vehicles. The approach is also useful as a verification method to analyse the potential success of eHMI concepts and can pinpoint specific questions for further research. Evidence suggests that design solutions that compromise the compatibility of the perceptual content with the processing requirements of the task are bound to fail. Additional decisions should be taken to reach a unified design:

1. **Reach Agreement Concerning "Content Selection".** Looking at current concepts, it seems that there is some agreement by car manufacturers and research institutes about the need to represent the vehicle's intents: a substantial consensus regarding the representation of the *Intent to stop & yield*, and partial agreement regarding the need to represent the vehicle's *Intent to drive.* With respect to the state information, there is a great variability between concepts: a few concepts represent the *Wait state* in a unique format and a very small number of concepts try to enhance the deceleration of the vehicle using visual animation. Furthermore, there is some disagreement about the necessity to represent the vehicle's mode, which sets the canvas for subsequent representations.

2. **Reach Agreement About the Optimal "Media Allocation".** The MRT dictates some design requirements from the perspective of effective resources usage: the conceptual processing requirements of the task will be best served by conceptual information, whereas the spatial processing requirements of the task will be best served by spatial content. Further research should be conducted to decide which information should be presented visually, which should be delivered auditorily, which should be presented to both modalities, and in the latter case – how should the information be integrated between modalities to enhance the communicative value of the design.
3. **Reach Agreement About the Optimal Media Realization: Design Decisions.** This discussion concerns the selection of colour, animation design and the interaction between colour and animation

Colour: Several constraints regarding colour currently apply by the Federal Motor Vehicle Safety Standards (FMVSS) in the US and the United Nations Economic Commission for Europe (UNECE) to guarantee that eHMI colours should not interfere with colours already implemented or reserved for specific functions. Traditionally, the allowable colours for any moving vehicle consist of white and amber at the front and sides of the vehicle and red at the rear (FMVSS 108). Additional colours are reserved to traffic devices and emergency vehicles. As specified by SAE J578 Standard, restricted colours are: Red, Yellow (Amber), Selective Yellow, Green, Restricted Blue, Signal Blue and White (Achromatic). Candidate colours to represent states and intents of autonomous vehicles are therefore Cyan (green blue), Selective Yellow, Mint-Green, and Purple/Magenta (Tiesler-Wittig, 2019; Werner, 2018; Dey 2021).

Cyan seems to be a promising choice. It is not prevalent in traditional urban and highway contexts and will therefore stand out in a bright daylight. In addition, there is some precedence for the selection of cyan in the industry; European and Japanese car manufacturers use cyan to mark autonomous vehicles in some design concepts. In addition, the SAE chose this colour to indicate the autonomous driving status of autonomous vehicles (SAE J3134). The suitability of cyan is also supported by several studies (e.g., Faas and Baumann, 2019; Beggiato et al., 2019; Hensch et al., 2019b) which show that pedestrians prefer cyan over white to indicate automation mode. The reliance on cyan can be elaborated further to express additional vehicle's states and intents.

Animation: Another topic that needs to be addressed is animation. As noted earlier, an abstract animated eHMI language is an appealing concept, due to its cross-cultural universal nature and the fact that it does not require reading capabilities, that is - if designed correctly. Clearly, some consensus should be reached regarding the animated effects employed by autonomous vehicles and the messages that they represent. Current regulation allows no animation, except for emergency vehicles (see SAE J845, SAE J595 and SAE J2498, all were written before the advent of autonomous vehicles). Similar to the selection of colour, a substantial body of research should be used to reverse these restrictions and allow the use of animated content by autonomous vehicles.

The interaction between Colour and Animation Colours are loaded with positive and negative meanings when static. For example, Bazilinskyy et al. (2019a) noted the compatibility of green with positive guidance messages such as "please cross" and recommend avoiding green if the eHMI is intended to signal a negative guidance message, such as "please do not cross". This finding was generalized also for aquamarine (sRGB 127, 255, 215) by Bazilinskyy et al. (2021). A pure cyan (sRGB 0, 255, 255) has no associative loading, although Dey et al. (2020b) report that cyan is perceived as "close to green" and is hence suitable for yielding signals.

Interestingly, when coupled with a strong-dynamic animation, cyan (and to a greater extent, magenta) can be associated with emergency vehicles, producing negative valence. This interaction between colour and animation was demonstrated in two studies by Beggiato et al. (2019) and Hensch et al. (2019b) on the eHMI design of Ford which consists of a blinking light to communicate the *Intent to drive* and a vigorous dual-sweep animation format to communicate the *Intent to stop and yield* message; These studies show that vigorous light animations provoke negative valence and a sense of alert, unless a neutral colour such as white is being used. This demonstrates the complexity of coming up with an acceptable design, due to the difficulty of finding what is the relative importance of each element in the design space (colour, animation, and prior loading such as police signals).

4 CONCLUSIONS

This paper provides an eight-step analytical method starting from analysing pedestrians' requirements in their encounters with manual vehicles and subsequently deduces a comprehensive communication protocol with future, driverless, automated vehicles. Maybury's (1993) principles of content selection (sampling information), media allocation (assignment of information into media and modalities), media realization (design methods) and coordination of different media that can be used to materialize this protocol into eHMI design while obeying the recommendations derived from Wickens' (1984) MRT and S-C-R Compatibility Principle. We envision that this method will serve to reduce the almost endless design space to a much more manageable space. It will also allow the industry to use the media-coding verification suggested here to reject poor designs and harmonize toward a more optimal design standard. Finally, the universality of this methodological approach and development process can be used by other industries that will need to develop mechanisms for communication between automated and autonomous machines and humans; whether in an enclosed space (e.g., factories and warehouse) or in the public space. eHMI communication language may also inspire the design of robots that are not necessarily in the public space or in immediate space conflict with humans. Since the automotive industry will be the first to deploy such robots and provide communication systems, we believe that other industries that will deploy robots in the public space (delivery robots, health care and assistive robots, information kiosk robots) will eventually be required to also provide such communication. Naturally, they may look at regulated solutions in the automotive industry to inspire their designs.

REFERENCES

Ackermann, C., Beggiato, M., Schubert, S., Krems, J. F. (2019). An experimental study to investigate design and assessment criteria: What is important for communication between pedestrians and automated vehicles? *Applied Ergonomics, 75,* 272–282.

Andersen, G. J., Enriquez, A. (2006). Aging and the detection of observer and moving object collisions. *Psychology and Aging, 21(1)*, 74–85.

AutonoMI by Leonardo Graziano. AutonoMI Autonomous Mobility Interface. Video.

Baddeley, A.D. (1986). Working Memory. OXFORD: OXFORD UNIVERSITY PRESS.

Bazilinskyy, P., Dodou, D., de Winter, J.C.F., (2019a). Survey on eHMI concepts: the effect of text, color, and perspective. *Transport. Res. F Traffic Psychol. Behav. 67*, 175–194.

Bazilinskyy, P., Kooijman, L., Dodou, D., de Winter, J.C.F. (2021). How should external human-machine interfaces behave? Examining the effects of colour, position, message, activation distance, vehicle yielding, and visual distraction among 1,434 participants. *Applied Ergonomics, Vol 95*.

Beggiato, M., Neumann, I., Hensch, A.C (2019). Light-based communication between automated vehicles and other road users. Chemnitz University of Technology, Department of Psychology. Cognitive and Engineering Psychology. Unpublished presentation Ford_ Color_AVSR-02-17e

Caird, J. K., Hancock, P. A. (1994). The perception of arrival time for different oncoming vehicles at an intersection. *Ecological Psychology, 6(2)*, 83-109.

Clamann, M., Aubert, M., Cummings, ML. (2016). Evaluation of vehicle-to-pedestrian communication displays for autonomous vehicles. Technical Report.

Cohen, J., Dearnaley, E. & Hansel, C. (1955). The Risk Taken in Crossing a Road. *Journal of the Operational Research Society. Vol. 6*, 120–128.

de Clercq, K., Dietrich, A., Núñez Velasco, J.P., De Winter, J., Happee, R., 2019. External human-machine interfaces on automated vehicles: effects on pedestrian crossing decisions. *Human Factors 61*, 1353–1370.

Delucia, P. R. (1991a). Pictorial and motion-based information for depth perception. *Journal of Experimental Psychology: Human Perception & Performance,17*, 738–748.

Delucia, P. R. (1991b). Small near objects can appear farther than large far objects during object motion and self-motion: Judgments of object-self and object-object collisions. In P. J. Beek, R. J. Bootsma, P. C. W. van Wieringen (Eds.). *Studies in perception and action: Posters presented at the Vlth International Conference on Event Perception and Action* (pp. 94–100). AMSTERDAM: RODOPI.

Dey, D., Walker, F., Martens, M., Terken, J. (2019). Gaze patterns in pedestrian interaction with vehicles: Towards effective design of external human-machine interfaces for automated vehicles, *Proceedings - 11th international ACM conference on automotive user interfaces and interactive vehicular applications, automotive ui 2019, Utrecht, The Netherlands*.

Dey, D., Habibovic, A., Pfleging, B., Martens, M., Terken, J. (2020a). Color and Animation Preferences for a Light Band eHMI in Interactions Between Automated Vehicles and Pedestrians, *Chi conference on human factors in computing systems*, Hawai'i, Honolulu, United States.

Dey, D., Holländer, K., Berger, M., P-eging, B., Martens, M., & Terken, J. (2020b). Distance-Dependent eHMIs for the Interaction Between Automated Vehicles and Pedestrians, 12th international conference on automotive user interfaces and interactive vehicular applications (automotiveui '20), Virtual Event, DC, USA.

Dey, D. (2021). External Communication for Self-Driving Cars: Designing for Encounters between Automated Vehicles and Pedestrians. *Doctoral dissertation. Future Everyday*, Industrial Design Eindhoven University of Technology.

Domeyer, J.E., Lee, J.D., Toyoda, H. (2020). Vehicle Automation—Other Road User Communication and

Coordination: Theory and Mechanisms. IEEE Access 8: 19860-19872

Dommes, A., & Cavallo, V. (2011). The role of perceptual, cognitive, and motor abilities in street-crossing decisions of young and older pedestrians. *Ophthalmic and Physiological Optics, 31(3*), 292–301.

DOT HS 813 148. NHTSA (2021). Automated Driving Systems' Communication of Intent With Shared Road Users https://rosap.ntl.bts.gov/view/dot/58325

Faas, S.M., Baumann, M., (2019). Light-based external Human Machine Interface: color evaluation for self-driving vehicle and pedestrian interaction. *Proceeding of the Human Factors and Ergonomics Society Annual Meeting. 63,* 1232–1236.

Federal Motor Vehicle Safety Standard 108 (FMVSS 108)

Ford eHMI design https://www.youtube.com/watch?v=boqG7Ss7chI&t=182s

Fridman, L., Mehler, B., Xia, L., Yang, Y., Yvonne, L., Bryan, F. (2017). To walk or not to walk: Crowdsourced assessment of external vehicle-to-pedestrian displays.

Gueguen, N., Meineri, S., Eyssartier, C. (2015). A pedestrian's stare and drivers' stopping behaviour: A field experiment at the pedestrian crossing, *Safety Science, vol. 75*, pp. 87–89.

Habibovic, A., Andersson, J., Malmsten Lundgren, V., Klingegård, M., Englund, C., Larsson, S. (2018). Communicating intent of automated vehicles to pedestrians, *Frontiers in Psychology., vol. 9*, p. 1336.

Hella-Light. https://www.hella.com/hella-com/en/press/Company-Technology-Products-12-12-2017-16333.html

Hensch, A.C., Neumann, I., Beggiato, M., Halama, J., Krems, J.f. (2019b). Effects of a light-based communication approach as an external HMI for Automated Vehicles — a Wizard-of-Oz Study. *Transactions on Transport Sciences. Vol. 10 (2),* 18-32.

Horrey, W. J., Wickens, C. D., Consalus, K. P. (2006). Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. *Journal of Experimental Psychology: Applied, 12*(2), 67–78.

International Organization for Standardization (ISO)/TR 23049. 2018. TECHNICAL REPORT ISO: Road Vehicles — Ergonomic aspects of external visual communication from automated vehicles to other road users. Technical Report.

Klatt, W. K., Chesham, A., Lobmaier, J. S. (2016) Putting up a big front: car design and size affect road-crossing behaviour. *PLoS One 11(7)*: e0159455.

Lagström, T., Malmsten Lundgren, V. (2015) AVIP—autonomous vehicles interaction with pedestrians: an investigation of pedestrian driver communication and development of a vehicle external interface. Master Thesis. Chalmers University of Technology, Göteborg

Lee, D.N (1976). A theory of visual control of braking based on information about time-to-collision. *Perception 5*, 437-459

Lehsing, C., Jünger, L. & Bengler, K. (2019). Don't Drive Me My Way – Subjective Perception of Autonomous Braking Trajectories for Pedestrian Crossings. In SoICT'19: *The Tenth International Symposium on Information and Communication Technology*, December 4–6, 2019, Hanoi – Ha Long Bay, Viet Nam. ACM, New York, NY, USA, 7 pages

Liu, Y. and Wickens, C. D. (1992). Visual scanning with or without spatial uncertainty and divided and selective attention, *Acta Psychologica, 79*, 131-153.

Maybury, M.T. (1993) *Intelligent Multimedia Interfaces.* Cambridge, MA: AAAI Press/MIT Press

Merat, N., Louw, T., Madigan, R., Dziennus, M., Schieben, A. (2018). Communication between VRUs and fully automated road transport systems: what's important? Accident Anal Prev.

Mercedes Benz F015 - luxury in motion Concept Vehicle. https://www.youtube.com/watch?v=MaGb3570K1U

Mercedes Benz - The cooperative vehicle. https://www.youtube.com/watch?v=VFu2dyqUQPw

Nissan's vision for the future of EVs and autonomous driving. https://www.youtube.com/watch?time_continue=6&v=h-TLo86K7Ck

Rasouli, A., Kotseruba, I., Tsotsos, JK. (2017). Agreeing to cross: How drivers and pedestrians communicate. In: *Proceedings of the 2017 IEEE intelligent vehicles symposium (IV),* pp 264–269.

Ren, Z., Jiang, X. and Wang, W. (2016). Analysis of the influence of pedestrians' eye contact on drivers comfort boundary during the crossing conflict, *Procedia Engineering, vol. 137*, pp. 399–406.

Risto, M., Emmenegger, C., Vinkhuyzen, E., Cefkin, M., Hollan, J. (2017). Human–vehicle interface: the power of vehicle movement gestures in human road user coordination. In: *University of Iowa (ed) Proceedings of the 9th international driving symposium on human factors in driver assessment, training and vehicle design*

Rothenbücher, D., Li, J., Sirkin, D., Mok, B., Ju, W. (2016). Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles. In *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on. IEEE*, 795–802.

Schieben, A., Wilbrink, M., Madigan, C. K. R., Louw, T., Merat, N. (2018). Designing the interaction of automated vehicles with other traffic participants: A design framework based on human needs and expectations. *Cognition, Technology & Work, 0.*

Shmueli, Y., Degani, A. In preparation. Usability evaluation of a bioluminescence-inspired eHMI communication.

Shmueli, Y. & Degani, A. (2019). Vehicle-Pedestrian Communication I. Analysis and Needs Study Internal Report

Shmueli, Y., Degani, A., Zelman, I., Asherov, R., Zande, D. Weiss, J. & Bernard, A. (2013). Toward a Formal Approach to Information Integration: Evaluation of an Automotive Instrument Display. *Proceedings of the 57th Human Factors and Ergonomics Society Meeting.* San Diego, CA.

Society of Automotive Engineers (SAE) (2019) Automated Driving System (ADS) Marker Lamp (standard no. J3134). *SAE International*, Michigan.

SAE Standard J845 - Society of Automotive Engineers (SAE) (2021). Optical Warning Devices for Authorized Emergency, Maintenance, and Service Vehicles. August 2021

SAE Standard J595 - Society of Automotive Engineers (SAE) (2021). Directional Flashing Optical Warning Devices for Authorized Emergency, Maintenance, and Service Vehicles. August 2021

SAE Standard J2498 - Society of Automotive Engineers (SAE) (2016). Minimum Performance of the Warning Light System Used on Emergency Vehicles. November 2016

Šucha, M. (2014), Fit to drive: 8th International Traffic Expert Congress. 8-9 May, 2014, Warsaw

Šucha, M., Dostal, D., & Risser, R. (2017). Pedestrian-driver communication and decision strategies at marked crossings. *Accident Analysis and Prevention, 102*, 41–50.

Tiesler-Wittig, H. (2019). Functional Application, Regulatory Requirements and Their Future Opportunities for Lighting of Automated Driving Systems (tech. rep.).

UNECE regulations ece r-65,' UNECE, accessed December 7, 2011.

Volvo 360C Concept Vehicle https://youtu.be/H5KNPQT72FA Retrieved on Oct 3, 2019

Warrick, M. J. (1947). Direction of movement in the use of control knobs to position visual indicators. In PM Fitts (ed.) *Psychological Research on Equipment Design: Army Air Force Aviation Psychology Program Research Department*. pp.137–46.

Werner, A. (2018). New Colours for Autonomous Driving: An Evaluation of Chromaticities for the External Lighting Equipment of Autonomous Vehicles. *ColourTurn, 10*(3), 183–193.

Wickens, C. D. (1984), Processing resources in attention, in R. Parasuraman and R. Davies (Eds), *Varieties of Attention,* NEW YORK: ACADEMIC PRESS, 63-101.

Wickens, C. D., Vidulich, M., Sandry-Garza, D. (1984), Principles of S-C-R compatibility with spatial and verbal tasks: the role of display-control location and voice interactive display-control interfacing, *Human Factors, 26,* 533-543.

Wickens, C.D., Liu, Y. (1988). Codes and modalities in multiple resources: a success and a qualification, *Human Factors, 30,* 599-616.

Wickens, C., Carswell, C. (1995). The proximity compatibility principle: Its psychological foundations and its relevance to display. *Human Factors, 37*(3), pp. 473-494.

Wickens, C.D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, Vol. 3,* NO. 2, 159-177

Xiang, H., Zhou, M., Sinclair, S.A., Stallones, L., Wilkins, J., Smith, G.A. (2006). Risk of vehicle-pedestrian and vehicle-bicyclist collisions among children with disabilities. *Accident Analysis and Prevention; 38*(6):1064-1070.