

Queensland University of Technology
Brisbane Australia

This may be the author's version of a work that was submitted/accepted for publication in the following source:

Faramarzian Borojeni, Mo, Pardo Gaytan, Jorge, Mandel, Ilan, Rakotonirainy, Andry, Ju, Wendy, & Schroeter, Ronald
(2025)
Decoding Driver Intention Cues: Exploring Non-verbal Communication for Human-Centered Automotive Interfaces.
In Yamashita, Naomi, Evers, Vanessa, Yatani, Koji, Ding, Xianghua (Sharon), Lee, Bongshin, Chetty, Marshini, et al. (Eds.) *CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems.*
Association for Computing Machinery (ACM), United States of America.

This file was downloaded from: https://eprints.qut.edu.au/256220/

© Consult author(s) regarding copyright matters

This work is covered by copyright. Unless the document is being made available under a Creative Commons Licence, you must assume that re-use is limited to personal use and that permission from the copyright owner must be obtained for all other uses. If the document is available under a Creative Commons License (or other specified license) then refer to the Licence for details of permitted re-use. It is a condition of access that users recognise and abide by the legal requirements associated with these rights. If you believe that this work infringes copyright please provide details by email to qut.copyright@qut.edu.au

License: Creative Commons: Attribution 4.0

Notice: *Please note that this document may not be the Version of Record (i.e. published version) of the work. Author manuscript versions (as Submitted for peer review or as Accepted for publication after peer review) can be identified by an absence of publisher branding and/or typeset appearance. If there is any doubt, please refer to the published source.*

https://doi.org/10.1145/3706598.3713635



Decoding Driver Intention Cues: Exploring Non-verbal Communication for Human-Centered Automotive Interfaces

Mohammad Faramarzian
Centre for Accident Research and Road Safety - Queensland (CARRS-Q)
Queensland University of Technology (QUT)
Brisbane, Australia
Dr.faramarzian@gmail.com

Jorge Pardo
Centre for Accident Research and Road Safety - Queensland (CARRS-Q)
Queensland University of Technology
Brisbane, Australia
jorge.pardogaytan@qut.edu.au

Ilan Mandel
Information Science
Cornell Tech
New York, New York, USA
im334@cornell.edu

Andry Rakotonirainy
Centre for Accident Research and Road Safety - Queensland (CARRS-Q)
Queensland University of Technology (QUT)
Brisbane, Australia
r.andry@qut.edu.au

Wendy Ju
Information Science
Cornell Tech
New York, New York, USA
wendyju@cornell.edu

Ronald Schroeter
Centre for Accident Research and Road Safety - Queensland (CARRS-Q)
Queensland University of Technology (QUT)
Brisbane, Australia
r.schroeter@qut.edu.au

Abstract

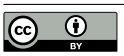
In emerging "driver-less" automated vehicles (AVs), the intuitive communication that exists between human drivers and passengers no longer exists, which can lead to reduced trust and acceptance in passengers if they are unclear about what the AV intends to do. This paper contributes the foundational understanding of how passengers naturally decode drivers' non-verbal cues about their intended action to inform intuitive Human-Machine Interface (HMI) designs that try to emulate those cues. Our study investigates what cues passengers perceive, their saliency, and interpretation through a mixed-method approach combining field observations, experience sampling, and auto-confrontation interviews with 30 driver-passenger pairs. Analysis of posture, head/eye movements, and vestibular sensations revealed four categories of intention cues: awareness, interaction, vestibular, and habitual. These findings provide empirical foundations for designing AV interfaces that mirror natural human communication patterns. We discuss implications for designing anthropomorphic HMIs that could enhance trust, predictability, and user experience in AVs.

CCS Concepts

• **Human-centered computing** → **Empirical studies in interaction design**; **Field studies**; **User studies**.

Keywords

Intention Cues, HMIs, Human-centric Design, Implicit Interaction, Anthropomorphism



This work is licensed under a Creative Commons Attribution 4.0 International License.

CHI '25, Yokohama, Japan

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713635

ACM Reference Format:
Mohammad Faramarzian, Jorge Pardo, Ilan Mandel, Andry Rakotonirainy, Wendy Ju, and Ronald Schroeter. 2025. Decoding Driver Intention Cues: Exploring Non-verbal Communication for Human-Centered Automotive Interfaces. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3706598.3713635

1 Introduction

The transition to AVs represents a fundamental shift in the role of future drivers towards that of a more cooperative driving team with the AV [25]. This represents challenges for how human and machine communicate and collaborate in real-time driving environments. Understanding a team member's *intention* is fundamental to any good collaboration [8], including those cooperations between humans and the AV system [25].

Humans have a unique capacity for inferring intention from the dynamic actions and motions of others [5]. While subtle, these non-verbal cues are essential to the intuitive coordination of effective dyadic interaction [12], including in-vehicle interactions between driver and passenger. Passengers use these cues to anticipate the driver's intentions and future actions, such as turning, braking, or changing lanes [33, 52]. These intuitive dynamics are perhaps most visible when examining interruption behaviours in passenger-driver interactions [10, 22].

In HCI, especially in the field of Human-Robot Interaction (HRI), there has been a large body of research to understand intuitive human intention communication across a wide range of application domains, with the view to inform the design of more intuitive interfaces, that improve trust, and user acceptance and experience [58, 59, 66]. To date, this approach has been underexplored in the context of "driving robots", or AVs.

A gap remains in the literature regarding the fundamental knowledge about how drivers' implicit non-verbal cues communicate their intentions within the vehicle. Addressing this gap may inform the future design of more intuitive HMIs for AVs on the basis of

this fundamental knowledge. Our research aims to address this gap by documenting and analysing these natural intention-sharing patterns between drivers and passengers.

This paper makes a novel contribution to HCI by investigating the implicit communication between drivers and passengers, focusing on the non-verbal cues that passengers recognise, their saliency, and how they are interpreted, through the following research questions:

RQ1: *What implicit, non-verbal cues do passengers perceive as the drivers' intention cue (syntax)?*

RQ2: *What intention cues are passengers particularly aware of and why (saliency)?*

RQ3: *How are these cues interpreted by passengers (semantics)?*

Through a mixed-method approach—combining naturalistic driving observations, experience sampling, and technology-assisted auto-confrontation interviews—this research uncovers the nuanced ways passengers interpret drivers' non-verbal cues. By exploring head and eye movements, posture, and other subtle cues, we contributed empirical findings that offer insights into the design of anthropomorphic HMIs. These findings emphasise the integration of human intuition and implicit communication into AV interfaces, enhancing their capacity to improve user experience in automated systems.

2 Related Work

2.1 Non-verbal Communication in Human Interaction

Non-verbal communication plays a fundamental role in how humans interact and cooperate, particularly in dynamic and high-stakes environments such as driving [39, 52]. A long history of research in human interaction has shown that non-verbal subtle cues, including facial expressions, posture, and body and eye movements, are critical for conveying intentions and emotions[23]).

In human-to-human interactions, these cues provide information that enables humans to estimate others' intentions through their movements [2], often more quickly and intuitively than verbal communication. Research has shown the richness of information conveyed through movement cues alone. For instance, researchers [3] found that arm kinematics were sufficient to discern different intentions (e.g., grasping a bottle for drinking versus pouring) through movement analysis. Similarly, Sartori et al. [50] examined how observers interpreted the reaching and grasping movements of a wooden block under different scenarios (cooperation, competition, or individual action), finding that observers use distinct cues to anticipate others' actions. Further supporting this, Trujillo et al. [57] used stick-light figures to demonstrate that humans can accurately read intentions from movement cues even with limited contextual information. These insights into how humans naturally interpret movement and intention have sparked significant interest in the field of human-robot interaction.

2.2 Non-verbal Communication in Human-Robot Interaction

HRI research identified the value of non-verbal communication, where it informs more natural and effective robot-human interactions. HRI researchers have focused on various aspects of non-verbal cues, such as kinesics and proxemics [51, 56]. Studies have explored how robots' non-verbal behaviours influence human cognition, emotions, and task performance, highlighting the importance of designing robots that can both recognise and generate appropriate non-verbal cues [13, 20, 24, 58, 60]. By incorporating these cues into interaction design, researchers and designers may create systems that foster more natural, inviting, and accessible experiences that foster trust and cooperation [58, 66].

Despite extensive research on non-verbal communication in human-human human-robot interaction, little attention has been paid to how these findings translate in the context of AVs. Existing research on non-verbal communication in this field has mostly surrounded the interaction between AVs and pedestrians [27, 49], rather than between the AV and their human driver/passenger. This leaves a critical gap regarding situations where human passengers may need to infer the intentions of automated systems. While studies have identified drivers' heads and eyes as valuable markers for determining intention, they have generally adopted a data-driven rather than a human-centric perspective [21, 45]. This study bridges this gap by—to the best of our knowledge, for the first time—exploring how passengers decode human drivers' non-verbal cues in real-time driving environments, with the goal of informing more intuitive and effective HMI designs for AVs.

2.3 Contextualising Non-verbal Communication in Automotive HMI Design

While extensive research exists on social interactions in AV HMI design [63], there remains a crucial gap in understanding natural human driver-passenger non-verbal communication patterns. To inform the design of these interfaces, we can look to human driver-passenger interactions [33], particularly the non-verbal communication that occurs in this context to convey intentions. Understanding intuitive human-human communication in current driving tasks is essential for designing more human-centred AV interactions [44, 64].

Prior work has suggested the importance of clear communication between human users and automated systems [25, 30], yet we lack a detailed understanding of how drivers' implicit non-verbal cues unfold—how they are recognised, how salient they are, and how passengers interpret them.

Automotive HMI design has increasingly turned to anthropomorphic depictions of elements and metaphors to create intuitive interactions aligning with human communication patterns. These approaches, such as Ju's [33] "Husband metaphor", provide frameworks for understanding the front-seat passengers' roles in cooperating driving [34]. The metaphor describes how passengers—similar to long-term couples who usually occupy the front seats of the car, develop an intuitive understanding of each other's behaviours without explicit communication—learn to read and interpret drivers'

implicit cues and intentions. This positions the front-seat passenger as an active observer interpreting non-verbal cues and providing feedback, much like the implicit cooperation that develops in long-term relationships. Such approaches provide a conceptual foundation for designing HMIs that facilitate more natural human-machine interactions, particularly relevant in the context of AVs where traditional driver roles are being redefined.

Human-human interaction studies have informed the development of anthropomorphic interfaces in human-robot interaction. For instance, research on human social cues and gaze patterns has led to more intuitive robotic interfaces [13, 24, 36]. In the automotive context, this approach aims to make automated systems more relatable and trustworthy. Anthropomorphic HMIs incorporating human-like gestures, facial expressions, or vocalisations can communicate AV intentions in ways more easily understood by passengers [14, 35].

However, while anthropomorphic elements can improve intuitive understanding of AV actions, there is a lack of foundational research on replicating the implicit non-verbal cues humans naturally decode in driving contexts. Most current automotive HMI designs focus on explicit communication, often increasing cognitive load [40]. Others focusing on anthropomorphic approaches only adopt human likeness in their HMIs in their efforts to create human-centric HMIs [4, 31, 35, 41, 42, 61, 62, 65]. In contrast, implicit communication through non-verbal cues offers more seamless and intuitive interaction.

This study addresses this gap by investigating how front-seat passengers perceive and interpret drivers' non-verbal cues in naturalistic driving, aiming to inform future anthropomorphic HMI designs that leverage implicit communication for more intuitive and less cognitively demanding user experiences.

3 Methodology

This study employed a combination of qualitative methods and tools to explore how front-seat passengers decode drivers' non-verbal cues in real-world driving scenarios. Drawing on Ju's [33] husband metaphor, which emphasises front-seat passenger engagement with driving tasks, all passenger participants were seated in the front passenger seat. The study integrated naturalistic observations, experience sampling, semi-structured interviews, and technology-assisted auto-confrontation interviews, to capture real-time and reflective insights into how non-verbal convey driving intentions between human driver-passenger pairs. This study was extensively rehearsed and iterated through seven pilots (with colleagues outside the project), approved by our university's ethics committee (Ethics Approval Number 2000000959) and conducted in accordance with our national Code for the Responsible Conduct of Research.

The study was conducted in a naturalistic driving environment over the course of two 40-minute drives per participant pair. The drives followed two of eight pre-defined routes, encompassing a range of road conditions, including highways, intersections, and residential streets. These conditions were selected to elicit various driving behaviours and situations, providing opportunities to observe and capture a wide array of non-verbal cues from the driver. The first drive was predefined and the same for all pairs, and the second was chosen by the passengers based on their highest familiarity. This ensured diversity and comparability in the driving contexts and familiarity with the routes of the drives.

3.1 Participant Recruitment and Selection

We recruited 60 participants (aged 18 to 60), 30 women and 30 men, grouped into 30 driver-passenger pairs (13 women and 17 men as *passengers*; 17 women and 13 men as *drivers*). The pairs were selected based on their level of familiarity with each other and grouped into 14 pairs of *familiar* participants who often accompanied each other in driving (e.g., couples, friends) and 16 pairs of *strangers* who had never been driven by each other. This pairing criterion was informed by existing literature's emphasis on non-verbal interpersonal communication and habituation, which suggests the cues and signals humans observe can be different based on familiarity [11]. By including both familiar and stranger pairs, the study ensured that a range of interaction dynamics is captured, though differences between familiarity levels will not be the primary focus of this paper. Within each pair, participants self-selected their roles as drivers or passengers. The sample of 30 pairs exceeded the threshold for both code saturation (9 interviews) and meaning saturation (16-24 interviews) established in qualitative research literature [32]. Our larger sample size allowed us to capture potential outliers and edge cases while ensuring comprehensive coverage of behavioural patterns and their nuanced manifestations.

Participants were required to be 18+ years of age, possess a valid driver's licence, and demonstrate active driving patterns (2+ hours weekly). Additional criteria included familiarity with the inner Brisbane road network and good English language proficiency. For drivers specifically, it was required the absence of injuries affecting driving capability within the previous 6 months along with no at-fault crashes within the preceding 2 years. These criteria were established to ensure the observation and expression of intentional cues through body language during driving.

Recruitment was carried out through social media and university mailing lists, and all participants were fully briefed on the study before their involvement. Participants were given a $100 gift card voucher as compensation for their participation.

3.2 Data Collection Methods and Apparatus

3.2.1 Observational Field Study. An observational field study was the primary method used to capture the naturalistic, unprompted interactions between drivers and passengers. Each vehicle was equipped with an array of four GoPro cameras, which recorded different angles of both the driver and the passenger, as well as the road ahead and the interior context of the car. This setup provided comprehensive coverage of all key visual cues, allowing us to track non-verbal communication throughout the drives. Figure 1 illustrates the apparatus and camera placement. The recorded video data formed the foundation for subsequent stages of the research, such as experience sampling and interviews.

3.2.2 Experience Sampling Method (ESM). During the drive, participants used an inconspicuous wireless clicker to flag moments when they perceived non-verbal cues that conveyed the driver's intentions—unbeknown to the driver. These timestamps were synchronised with the video footage, enabling a detailed analysis of the



Figure 1: The synchronised recorded footage of the natural driving study. On the top left is the view of the driver's entire body. On the top right is the driving context. On the bottom left is the view of both driver and passenger. On the bottom right is the approximate point of view of the passenger. On the timeline below, the various white markers represent the sampled experiences of the passenger.

cues that passengers noticed (see Figure 1). The ESM was critical in capturing the subjective perception and the saliency of these cues in real-time, providing a timeline of the passengers' intention awareness.

3.2.3 Technology Assisted Auto Confrontational Semi-Structured Interviews. Auto-confrontational interviews [47] aim to elicit deep insights into a participant's thoughts, behaviours, or experiences by confronting them with their own actions or situations, e.g, by enabling participants to revisit specific events through recorded footage, thereby facilitating detailed recall and articulation of their lived experiences and perceptions.

In our study, following a 10-minute break after each drive, technology-assisted semi-structured auto-confrontational interviews were immediately conducted with the passengers to gain a deeper understanding of their interpretations of the driver's non-verbal cues.

The interviews focused on moments flagged during the drive (via the clicker) and sought to uncover why those cues were noticed and how they were interpreted (see Figure 1). While all flagged moments (clicks) were reviewed during the interviews, participants occasionally indicated that certain cues were similar to previously discussed ones or were unable to recall specific details of some flagged moments. During these interviews, passengers viewed synchronised video playback of the drive, showing only the right side of the video to match their perspective (see Figure 2a). This approach served two purposes: a) it enabled participants to reflect on their thought processes and the implicit cues they observed, providing valuable insights into their cognitive and emotional responses, and b) it avoided directing participants' attention to specific, irrelevant, or potentially misleading aspects of the video footage. The semi-structured format ensured consistent coverage of key questions such as *"How did you know the driver was going to do that?"* and *"What and how did the driver's body language signal to you?"* while allowing flexibility for participants to elaborate on their experiences. Our approach proved particularly valuable in capturing nuanced insights that might otherwise remain unexpressed in traditional interview settings.

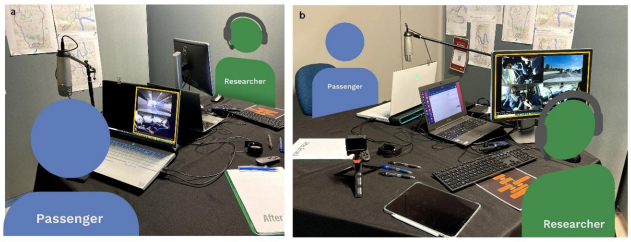


Figure 2: A view of the interview apparatus showing a) what the passenger sees on their monitor, where half the monitor is blacked out, only the right side is visible and b) the researcher sees all the angles recorded on their monitor.

3.3 Procedure

Each session took around 3-4 hours from arrival to the end of the study. This included check-in, study briefing, consent, prep, two approximately 40-minute driving sessions, and the interview. Participants (passengers and drivers) were welcomed and introduced to each other if they did not know each other (stranger pair). They received separate briefings about their roles in the study.

The driver participants were briefed in the car to familiarise themselves with the research vehicle, including a familiarisation drive (accompanied by a researcher). They were instructed to drive safely and as naturally as possible during both the test drive and the study. After the familiarisation drive, they were asked if they felt confident to continue with the study and sign their consent.

The passenger participants were briefed about the use of the clicker and instructed to use it to "flag" driving situations (see 3.4.2 Experience Sampling Method) when they either a) detected/inferred the driver's intention from their body language (cues), or b) when they felt they lacked sufficient information to infer the driver's intentions.

After the briefings, participants began the naturalistic drives. Participants were instructed to refrain from conversations during the drives to a) not distract participants from their experience sampling task (as identified during piloting) and drivers from their driving task (ethics and risk mitigation), b) simulate an environment similar to an automated driving experience in which our driver participants embody and simulate a non-verbal visual "interface", and c) to isolate non-verbal communication to inform non-verbal HMIs. They did two drives with a 10-minute break between them. Upon completion, participants had another short 10-minute break while data was extracted and prepared for the subsequent passenger interview.

Prior to the interview, drivers were debriefed about the passenger's clicker activities and the reason for withholding this information earlier, ensuring their natural behaviour. Drivers then provided written consent to continue. Following the interview and debriefs, participants could ask questions and receive their participation vouchers. During this stage the participants were given the option to give consent to their image release.

4 Data Analysis

The data collected from the ESM and interviews were analysed using a combination of qualitative thematic analysis [9] and Interpretive Phenomenological Analysis (IPA) [53]. For the thematic analysis, we carefully examined the video data, which was synchronised with the clicker events, to identify recurrent patterns (codes) in the driver's non-verbal cues, such as head movements, eye movements, and posture. We paid particular attention to how these non-verbal cues related to the driving context and how passengers interpreted them, allowing us to build a comprehensive understanding of the implicit communication occurring during the drives.

Complementing this approach, IPA explored the subjective experiences of the passengers. It allowed us to understand how passengers made sense of the implicit communication they observed during the drives, and to uncover rich, diverse and nuanced insights into how passengers' interpretation of non-verbal cues contributed to their understanding of the driver's intentions.

To ensure the validity and reliability of our findings, we implemented several strategies. We prioritised reflexivity and transparency throughout the research process and the IPA. The primary researcher engaged in regular reflective practices to identify and document personal assumptions, biases, and preconceptions that might influence data interpretation. This reflexive bracketing [54] was complemented by multiple iterations of peer review and debriefing, where four research fellows and three professors provided external perspectives on the analytic process. These peers reviewed video examples representing proposed themes and offered feedback, challenging assumptions and interpretations to streamline emerging themes. By implementing these measures, we aimed to address potential biases in data collection and analysis, ensuring a rigorous and transparent research process that accurately represents the participants' experiences and interpretations.

4.1 Coding Overview

Thematic analysis of participant reports was conducted using the Driving Intention Grammar (DIG) approach. This approach is inspired by institutional grammar, an analytical tool developed by Elinor Ostrom and Sue Crawford [18] used to systematically dissect and analyse the components of documents. Using DIG, we first broke the reports of passenger observations (syntax) down to subject (generally body parts), predicate (action), object (context), and adverbs (manner, time, place) (see Table 1). This systematically captured the core elements of how passengers reported the driver's non-verbal cues, e.g. *"I saw him turn his head (subject) to look (predicate) in my mirror (object), quickly (adverb)"*. When participants used predicates such as "glancing", it was coded as defined: looking (predicate) quickly (adverb).

The reports about the passenger's interpretation (semantics) were also broken down into their core components. For example, the report *"… and that meant that he is going to merge to the left lane"* was coded as *"lane change, left"*. In other words, DIG helped identify key insights related to what was observed, the identified cue and the driving action that the cue conveyed. The semantics were subsequently iteratively clustered into themes, as discussed in the findings.

Table 1: Driving Intention Grammar Tool (DIG) coding example (simplified).

Subject	Predicate	Object	Adverb
Head & Eyes	Looking	At oncoming traffic	Constantly, quick
Hands	Gripping	Steering wheel	Tense, lose
Head	Surveying	The surrounding	Repeatedly

5 Findings

The reports about the passenger's interpretation (semantics) were also broken down into their core components. For example, the report "… and that meant that he is going to merge to the left lane" was coded as "lane change, left". In other words, DIG helped identify key insights related to what was observed, the identified cue and the driving action that the cue conveyed. The semantics were subsequently iteratively clustered into themes, as discussed in the findings.

These findings are the results of the 30 participant pairs' 22.5 hours of driving and 30 subsequent passenger interviews, each 40-100 minutes long, resulting in around 35 hours of interview data.

Participants showed a diverse range of clicking patterns, ranging from a minimum of eight clicks to a maximum of 421 clicks (Mean=109.6, SD=113.23) recorded across all sessions. On average, familiar pairs recorded more clicks than stranger pairs in both drives (mean=67.1 [SD=68.4, min=4, max=249] vs mean=51.6 [SD=45.0, min=5, max=194] in drive 1, and mean=53.2 [SD=52.0, min=4, max=193] vs mean=51.6 [SD=51.5, min=3, max=180] in drive 2), and participants clicked more during the first drive (59.3) than in the second drive (52.1). In sum, participants collectively logged 1899 clicks in drive 1 and 1720 clicks in drive 2, with a total of 3619 times. Approximately 60% of the clicks were discarded due to reasons such as participants not recalling or accidental clicks. A total of 1451 clicks remained and were discussed in the interviews. They displayed two types of clicking behaviours: a) "atomic" events, which represented individual clicks, and b) multiple events, which involved multiple consecutive clicks. Around 12% of these events involved multiple-clicks instances, while 88% were "atomic" clicks.

The analysis of passenger interviews revealed key insights into how non-verbal cues from drivers are perceived, decoded, and interpreted. These findings emerged through an iterative coding process that organised participant observations into themes, focusing on driver body language and its implicit communication of intent. We focused on the syntax and semantics of driving implicit cues, examining posture and head, eye and hand movements.

5.1 RQ1: Syntax

The coding process revealed that certain body parts and actions were consistently identified by passengers, particularly head, eye and hand movements, as primary indicators of the driver's intentions.

5.1.1 Head and Eye Movements. Head and eye movements were the most commonly reported intention cues (45%), frequently associated with checking mirrors, or surveying the surroundings, and interpreted as indicators of the driver's intent to change lanes. Passengers commonly reported their driver's looking at various markers in the driving environment and associated them with intentions and potential future driving actions, *"...he was looking at the traffic light, so I was sure she was going to take off as soon as the light turned green."* Other times, the observable driver's head and eye movement simply conveyed to the passenger that the driver was aware of the driving environment *"There are subtle head movements. I just know she is taking in everything well. I knew she was scanning everything; I didn't know how, but I know her, and she does that, and I can just tell, with her."* For a list of the discovered codes and their relationship, see Figure 3.

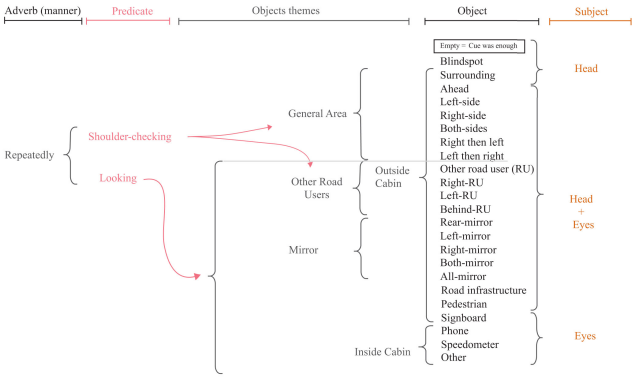


Figure 3: Coding structure of head and eye movements guided by DIG.

5.1.2 Hand and Arm Movements. Participants also noted the driver's hand and arm movements (often in relation to the steering wheel) as critical cues, especially when signalling or preparing for a turn. One participant described the driver's *"tense grip on the steering wheel"* as an indication of heightened focus or decision-making, while another associated a *"loose grip"* with the driver's confidence in knowing the path ahead. Cues such as reaching for the indicator or resting hand on the gear stick were commonly observed and interpreted as preparatory steps for a manoeuvre.

The subtlety of these cues, such as *"hovering fingers over the indicator,"* was also noted, with participants acknowledging the importance of these subtle cues in interpreting the driver's next move. At times, the passenger reported only the way the driver gripped the steering was indicative of the driver's intentions, stating, *"I could tell by how he had his hands on the steering wheel that we were going to turn right"*.

These participant observations provide valuable insights as the cues they observed are almost imperceptible to researchers through video observation. For example, in the situation depicted in Figure 6, both passengers noted the driver's fingers hovering over the indicator, although it is captured, it is hard to see. For a list of the discovered codes and their relationship, see Figure 4.

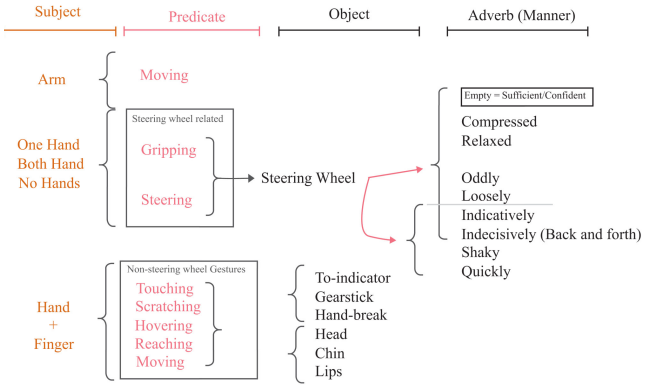


Figure 4: Coding structure of arm, hand and finger observations guided by DIG.

5.1.3 Facial Expression and Postures. Changes in the driver's posture were less frequently noted but were still significant when observed. Research suggests that posture changes usually are related to arousal or anxiety [23, 28]. This aligns with the participants' reports (see Table 2). Participants described instances where the driver's posture suggested discomfort, readiness for action, or indecision. One participant interpreted that when the driver *"sat up straighter"* at a busy intersection, they were preparing for a complex manoeuvre. In addition, a relaxed posture was reported several times to cue the driver's confidence. This posture-based cue provided the passenger with additional context about the driver's attention and SA. However, most of the reports of posture were made in tandem with head and eye movements, making the reported posture movements as kinesics adaptor used to get a better view of the environment. For a list of the discovered codes and their relationship see Figure 5.

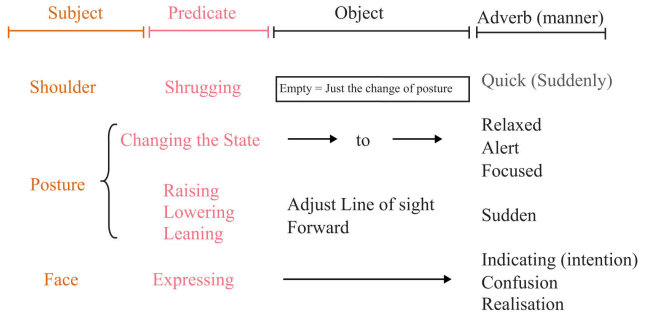


Figure 5: Coding structure of shoulder, posture and face observations guided by DIG.

5.1.4 Leg, Foot, and "Car". The study revealed an unexpected source of driver intention inference, which was predominantly "the car", rather than observing the leg and foot. Passengers consistently reported their ability to glean information about the driver's intentions through *vibrations*, *inertia*, and *sounds* of the car, perceived through touch, vestibular and auditory senses.

Participants demonstrated a remarkable ability to interpret subtle vehicle movements. For instance, passengers noted *"I felt the driver (stopping)"* or *"he slowed down quite a bit when he realised he had to change lanes"*, illustrating how vestibular cues can reveal the driver's immediate actions. Another participant reported, *"Probably just the patting of the brake, it's subtle. I was hearing it in the car's motor"*, and yet another noted, *"I felt he rolled off the accelerator,"* highlighting the close association between vestibular cues and the auditory cue (sound of engine) or the visual cue of the driver's manipulation of the pedals with the foot. This highlights the multisensory nature of this perception, which did not lend itself to in-depth coding using DIG identifying predicates and objects.

The absence of expected vestibular cues was also frequently reported where passengers observed protocol discrepancies. Passengers mentioned instances like *"didn't slow down in time"* or *"going slow, he was unsure of his surroundings, not making sudden movements."* Vestibular cues, far from being mere physical feedback, serve as a rich source of information about the driver's intentions, competence, and cognitive processes when interpreted alongside other contextual cues.

5.2 RQ2: Saliency — Reinforcement Through Repetitions, Timing, and Intensity of Cues

Table 2 presents an overview of the percentages that individual DIG subjects, generally body parts as discussed above, were perceived as providing an intention cue. They are clustered in their semantic themes as discussed further in section 5.3. in relation to RQ3: awareness cues, interaction cues, vestibular cues, and unique or habitual cues. Table 2 quantifies the saliency of DIG subjects in relation to intention. In turn, we discuss factors that further reinforced what passengers were aware of and why.

Amongst the behaviours reported, there were aspects of some of the observed behaviours (coded by *adverb* using DIG) that stood out based on how they influenced the passenger's interpretations. These were repetition and exaggeration which helped the passenger interpret the urgency and the certainty of the driving intentions. Other times, it was the prominence of events unfolding in the driving context that reinforced the interpretation of the driver's Intentions.

Repetition plays a crucial role in this communication process. Passengers often emphasised that the repetition of multiple cues provided certainty of the driver's intentions. As one participant remarked, *"When he checked both mirrors and looked over his shoulder, I knew he was going to change lanes."* For example, one participant explained, *"When he does the double head turn, it's more like he commits to the lane change."* This observation was common, allowing passenger participants to report their interpretations of the driver's intentions with more certainty. One of the most common intentions reported was lane change, which was made up of two or more head turns, identifying the second head turn to be broader (see Figure 6). Even when no immediate action followed, passengers attributed meaning to the cues, recognising them as part of the driver's decision-making process with certainty.

Exaggeration of the observed cues serves as another important attribute. These were the intensity and subtlety of the observed movements of the driver. Participants often used words such as *"kept looking"*, *"many times"*, *"quick"*, or *"checked"* to signify the intensity of the observed cue. When they reported on intense body movements, they were able to infer the urgency of the driver's intention using words such as *"tense"* and *"drastic"*. Usually, the sense of urgency interpreted was due to the observation of intense driver movements. These interpretations of urgency in anticipated action could signify that the driving intentions would be anticipated to take place sooner rather than later.

On the other hand, observation of subtle body movements did not change the urgency of the intentions. Observation of broad, clear movement allowed the passenger to report their interpretations more confidently, thus making them more certain about the anticipated intentions of the driver. Passengers reported *"big"* or

Table 2: Percentages of DIG subjects, generally body parts, perceived as providing an intention cue.

Awareness Cues		Interaction Cues	
Eyes	23%	Hand	16%
Head	22%	Finger	<1%
Posture	5%	Arm	<1%
Total	50%	Total	16%
Vestibular Cues		**Unique and Habitual Cues**	
Vestibular	28%	Face	3%
Leg	2%	Posture	<1%
Foot	1%		
Total	31%	Total	3%



Figure 6: Two passenger pairs' distinct observations of similar repeated head-turns that implicitly communicate a lane change. On the top is a "big" or "broad" head movements, whereas on the bottom, head movements are reported without adverbs.

"broad" movements, which further reinforced the salience of the cue, as seen in Figure 6 (top). The movements are broader and more salient. The various ways the same cues were performed reinforced the passenger's confidence in predicting the driver's intentions.

5.3 RQ3: Semantics

The thematic analysis revealed that passengers rely on a combination of various implicit cues within the driving context to decode the driver's intentions. However, these cues were not always perfectly aligned with the driver's actual behaviour. In some instances, passengers reported cues that seemed to indicate one action but were followed by a different outcome, such as a delayed or cancelled lane change. As one participant reflected, *"Sometimes he ends up not changing lanes, but that doesn't affect me; he just probably changed his mind."* This highlights the interpretative nature of decoding non-verbal cues, where passengers consistently acknowledge the fluidity of intentions during the driving task.

This study uncovered several key themes related to the implicit non-verbal cues that passengers observe and interpret while decoding driver intentions. These cues—categorised as (i) Driver's Situational Awareness (DSA) cues, (ii) interaction cues, (iii) vestibular cues, and (iv) habitual cues—each provide different insights into how passengers understand the driver's body language to decode the driver's intention. This knowledge can inform the design of human-centric automotive interfaces.

5.3.1 Awareness Cues - Driver's Situational Awareness Information. Awareness cues provide passengers with insights into what they perceive as the driver's situation awareness regarding the driving environment. These cues often involve head and eye movements, such as checking mirrors or looking at specific road elements. Passengers frequently interpreted these movements as indicators of the driver's awareness, even if no immediate action followed, from which intention, and sometimes competence, was inferred.

One participant noted how *"he was looking to the left"* as a signal that the driver was aware of other vehicles, even though no lane change occurred. While these cues do not necessarily communicate immediate intention, they offer passengers a sense of the driver's ongoing assessment of the environment and their prioritisation.

Observing these awareness cues contributed to passengers' feelings of security and confidence in the driver's competence, enhancing their overall comfort during the drive. By signalling the driver's focus on relevant elements of the driving context, awareness cues provide valuable feedback and reassurance that the driver is engaged in the task at hand.

5.3.2 Interaction Cues – Feedforward Information. Interaction cues, primarily involving the driver's physical engagement with vehicle controls like the steering wheel, indicators, or gearshift, serve as immediate feedforward information for passengers. These cues allow passengers to instantly interpret the driver's imminent intentions, closely aligning with Ju's [33] concept of "feedforward" regarding interactions.

This immediate interpretation enables passengers to anticipate actions and potentially intervene if necessary. Passengers reported that subtle hand and arm movements, such as *"extended his hand, reaching for the indicator"* or *"where his hands were on the steering wheel"*, were reliable indicators of impending manoeuvres. One participant noted, *"it allowed me to know what's coming before it happens,"* highlighting the predictive nature of these cues. The way the hands gripped the steering wheel also allowed the passengers an insight into the driver's state, *"her hands were tightly gripped on the steering wheel. Good. The driver is ready for any situation."*

The steering wheel emerged as a focal point for these observations, with interaction cues ranking as the third most frequently observed category. Passengers demonstrated a remarkable ability to monitor various vehicle controls in detail, primarily by observing the driver's hands. The saliency of interaction cues emphasises the importance of anticipating the interaction with vehicle controls.

The saliency of these cues underscores the importance of designing interfaces that provide passengers with clear, yet implicit, feedback on vehicle actions—ensuring that they are aware of the vehicle's intended movements without overwhelming them with explicit information.

5.3.3 Multisensory Cues – Multipurpose Feedback and Feedforward Information. Multisensory cues proved to be significant indicators of driver intent in combination with contextual awareness. Vestibular cues often serve as an immediate and tangible form of feedback about the driver's actions, perceived by passengers through the vehicle's physical movements, changes in inertia, including acceleration, deceleration, and directional shifts, and are often augmented in combination with auditory, visual or haptic cues.

The direct and instantaneous nature of these cues provides passengers with real-time information about the driver's control inputs. The physical sensations allowed passengers to interpret the driver's immediate action as *feedback*, and intentions as *feedforward*: *"I felt the car slowing down [feedback], and I knew he was going to stop [feedforward]"*. They are, therefore, multi-purpose, serving as a rich source of information about the driver's intentions, competence, and cognitive processes when interpreted alongside other contextual cues.

Unlike other cue types that may indicate future intentions, multisensory cues offer immediate feedback on the vehicle's current state and dynamics. This direct physical feedback provides passengers with a visceral sense of the driver's control over the vehicle, enhancing their awareness of ongoing manoeuvres and immediate changes in the driving situation.

5.3.4 Unique and Habitual Cues - Driver State and Feedforward Information. Unique cues were typically one-off observations made by passengers, particularly in unfamiliar driving pairs. These cues often involved behaviours or gestures that were not part of the driver's regular routine, such as *"tapping on the steering wheel"* or *"fidgeting."* One participant noted that these unusual gestures made them unsure of the driver's intentions at first but later helped them understand the driver's mood or focus.

In contrast, familiar driving pairs reported habitual cues more commonly by observing their expressions and posture with consistent behaviour they had observed in their drivers over time. Habitual cues are often performed automatically and can be a strong predictor of future behaviour in similar situations [48]. For instance, one participant confidently stated *"There, see! She is thinking. ((biting his lips)) biting her lip, and then she looked, [intonating as if to suggest steps], and then she merged, and then her face relaxed … I*

knew exactly what she was going to do there" referring to the driver's familiar facial expressions and postures. When probed to clarify how he can interpret such resolute interpretations, as the interview apparatus did not allow full view of the driver's face (see Figure 7) he stated, *"… we've been living together for what, 40 years, and we climb together, and we cycle together, so we learn to watch out for everything. It's just the way it is."* In the study of kinesics, these facial expressions and irregular body movements can be categorised as affect displays, which are usually involuntary and can contain emotional meaning [6].



Figure 7: Example of the intricate habitual cue of passenger interaction with her face during the drive (Left). During the interview, Bill mimics the same hand interactions of the intricate habitual cue, which he has learnt to interpret as different levels of the driver's concern with the state of the environment (Right).

As passengers usually understand these cues through experience with the driver and require habituation, we label them habitual cues. These habitual cues often conveyed deeper insights into the driver's cognitive state, such as planning or hesitation, and were recognised quickly by passengers who had prior experience with the driver. This distinction between unique and habitual cues highlights the potential for HMIs to allow users to adapt to HMIs over time based on the user's familiarity with the system.

6 Discussion

6.1 Implications for HMI Design

Our findings present both observed human intention cues (syntax), their saliency, and how a human interprets them (semantics). They provide valuable insights into the complex interplay between driver cues and passenger interpretations. This can inform future human-centric HMI designs for AVs mimicking or abstracting the human driver's behaviour, which we discuss as follows.

6.1.1 Informing Awareness Interfaces. Awareness cues are critical as they convey not only intentions but also information about driving competency, resulting in trust, feeling of safety and comfort. This underscores the importance of visual cues in conveying situational awareness and suggests that future HMI designs should prioritise visual feedback mechanisms. Current vehicle HMIs typically provide a birds-eye view of constant feedback on all driving-related markers without bias [7]. However, our research indicates that a more nuanced, context-aware approach could be more intuitive.

Interfaces could highlight specific elements based on the vehicle's current driving context, filtering out less critical information, such as parked cars in a separate parking lot (see Figure 8). Less prominent but important markers, such as the traffic lights or other road users with actual conflict potential, are emphasised to signify the vehicle's focus and nuanced understanding of what is relevant in the current driving situation. As an abstraction of the passenger's head, a blue field of vision could be animated and informed by the passenger's observation of the driver's head movements. This approach aligns more closely with how human drivers naturally process and prioritise information in specific driving scenarios.

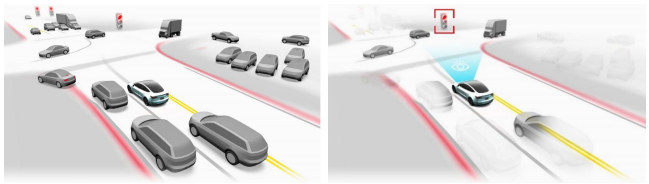


Figure 8: Left: Screenshot of Tesla's situation awareness feedback Photograph by Francesco Barontini. Right: An altered image of Teslas' situation awareness, only emphasising contextual elements highlighted by our participants' reports associated with the intention of turning in a junction (image retrieved and adapted from uk.motor1.com).

A comprehensive HMI design should consider complementary sensory modalities to provide a more holistic representation of the vehicle's awareness and intentions. The design of HMIs could benefit from incorporating visual emphasis or repetition to signal critical information or intentions. For instance, HMIs could use pulsing or repeated but subtle visual cues to indicate important actions or environmental factors.

Implementing these findings in AV HMIs' design could minimise cognitive burdens, improve trust in automated systems, and enhance overall user experience. As AV technology continues to advance, it is crucial to develop interfaces that effectively communicate the system's awareness, mirroring the subtle yet informative awareness cues that human drivers provide.

6.1.2 Feedforward HMIs using Interaction Cues. Interaction cues, primarily involving the driver's hands and arms, serve as crucial feedforward information, allowing passengers to anticipate the driver's intentions before actions are executed. These subtle movements, such as changes in steering wheel grip or reaching for controls, provide implicit yet powerful signals that passengers rely on to interpret upcoming manoeuvres.

In the context of AV interface design, replicating these interaction cues presents an opportunity to enhance user trust and comfort. Visual abstractions of hand movements, similar to those explored in recent animation studies by Gomaa, Zitt, and Reyes [26] and Kang et al. [35], could offer an ambient way to convey feedforward information around the steering wheel. Additionally, both tactile and visual feedback around the steering wheel, inspired by work in human-robot interaction like that of Tinga et al. [55], could provide embodied cues that mimic the subtle changes passengers observe in human drivers. However, while this type of research often relies on animated characters to synthesise interactions, our research allows drawing from natural studies of real-world behaviour.

Our findings suggest that subtle visual and tactile abstractions of the observed cues, such as changes in ambient lighting or haptic feedback on the steering wheel, could effectively replicate the implicit signals that passengers observe and decode driver intentions. While similar interfaces are already incorporated into some steering wheel designs, they often lack grounding in specific human behaviours. The goal of such human-centric interfaces would be to create a feedforward system that communicates impending vehicle control engagements, allowing AV users to anticipate actions and intervene if necessary. This approach aligns with the natural way passengers interpret driver behaviour and could significantly enhance the user experience in AVs.

6.1.3 Feedback HMIs Using Vestibular Cues. Vestibular cues, on the other hand, provide immediate feedback [1] rather than feedforward information. These cues, experienced through subtle changes in longitudinal and lateral accelerations, represent the actual execution of the driver's intentions [16]. Unlike other cue types, vestibular feedback is instantaneous and directly tied to the vehicle's movements [37]. This immediacy makes vestibular cues particularly valuable for validating the observer's previous interpretations.

Vestibular cues are subtle but salient, immediate and implicit. Yet, they are largely under-researched as a complementary HMI modality that does not overwhelm users with explicit information. In the context of AV design, a more intentional design of these vestibular cues through vehicle dynamics (or haptic interfaces) could offer a human-centric approach to conveying vehicle intent. By simulating the subtle physical sensations associated with acceleration, deceleration, and turning, AVs could provide users with an intuitive understanding and early feedback of the vehicle's actions [16, 17].

6.1.4 Habitual cues. The distinction between unique and habitual cues observed in human drivers offers valuable insights for designing future AV HMIs. By leveraging these concepts, designers can create more intuitive and communicative interfaces that enhance passenger trust and comfort. One promising approach to incorporate both unique and habitual cues in AV HMIs is the use of visual cues such as mood lighting to communicate the vehicle's "mood" or operational state, similar to Dandekar et al.'s [19] designs. This design element can communicate the AV's confidence level and stress in various driving scenarios. When navigating complex environments, the AV HMI could employ various colour scheme that indicates heightened alertness or stress. For instance, amber or orange lights could signify caution and increased focus, while pulsing patterns might mimic the "fidgeting" observed in human drivers.

AV HMIs can leverage the principle of habituation by offering customisable interfaces implementing a learning system that adapts its communication style based on user facial expressions and reactions to different contexts. As users become more familiar with the system, the HMI could provide less or more detailed information, mirroring the way passengers learn to interpret subtle cues from human drivers over time.

To bridge the gap between human drivers and AVs, designers can incorporate subtle, human-like behaviours into the HMI. For example, the system could simulate "lip-biting" or "thinking" moments through brief pauses or changes in light patterns before executing complex manoeuvres. Using directional lighting to indicate intended actions could mimic how passengers observe a driver's gaze, while implementing a "relaxation" state after successfully navigating challenging situations could further humanise the AV's behaviour. These examples highlight the necessity in utilisation of our findings to guide designs that use the interior as an HMI to communicate the AV's emotions [38]. The association of emotions to AVs through HMIs have been the focus of external HMI research [43].

By thoughtfully integrating these elements, AV HMIs can create a more natural and intuitive communication channel with passengers. This approach of basing the designs of HMI informed by observable driver behaviour has the potential to increase trust and comfort in automated driving systems by providing familiar cues that passengers can learn to interpret over time, much like they do with human drivers.

6.1.5 Summary. In summary, our findings may offer significant implications for the refinement of anthropomorphic design strategies in AV interfaces. By leveraging body language data from real drivers, researchers may develop HMIs that communicate intentions in a more nuanced and human-like manner, going beyond mere visual likeness to incorporate authentic driver behaviours. This approach addresses the limitations of current efforts in human avatar integration, which often fail to accurately represent driver behaviour [4, 31, 42, 55]. This human-centric approach to HMI design may have the potential to significantly enhance user trust, comfort, and overall experience in automated driving systems, bridging the gap between human drivers and AVs.

6.2 Limitations

While this study provides valuable insights, several limitations should be acknowledged. Participant instruction and recall posed significant challenges. Despite careful briefing, some participants forgot instructions or their sampling task, necessitating reminders before the second drive. More critically, over half of the sampled moments were not remembered during interviews, with participants often unable to explain their clicks. Recall, when present, was often non-linear, requiring careful review with recorded footage.

Click frequency varied widely among participants (mean=109.6, SD=113.23), affecting interview depth. Some participants used single clicks for multiple cues, while others used multiple clicks for single observations. Although these variations complicated analysis, the clicks remained valuable for identifying common themes and cues. Furthermore, maintaining participant engagement during the lengthy auto-confrontational interviews (often exceeding an hour) was challenging. We employed a mix of question types and positive feedback to keep participants focused and motivated.

Our recruitment method may have introduced generalisability concerns, including self-selection bias, sample size and seating positions. All passengers were seated in the front passenger seat. This may limit the generalisability in terms of the visual perspective from which participants made their driver intention observation. Lastly, instructing people to refrain from talking is another limitation. While this constraint may have introduced an artificial element to the interaction context, particularly for familiar pairs, it was methodologically necessary to observe non-verbal cues without verbal communication interference. Furthermore, the study's

sample size, self-selection bias, and cultural context may influence the interpretation and expression of non-verbal cues, potentially limiting the applicability of findings to other cultural settings.

However, we still observed numerous instances of passengers identifying implicit intention. The value of qualitative fieldwork lies not in the quantity of observations but in uncovering repeatable patterns of different phenomena [15]. In our case, the repeatable patterns reveal how passengers decode drivers' non-verbal cues. These patterns represent fundamental social practices in driver-passenger interactions and may offer generalisable insights into how passengers interpret driver intentions, regardless of specific contextual constraints.

Despite these limitations and acknowledging the methodological challenges, our diverse sample provided rich insights into driver-passenger interactions and intention cues. We acknowledged challenges in participant recall, click frequency variability, and potential biases. While our sample size limits generalisability, the depth and variety of data gathered offer a valuable foundation for understanding the complexity of non-verbal communication in driving contexts. These limitations highlight areas for methodological refinement in future studies of driver-passenger interactions.

6.3 Future Work

Our study lays the groundwork for further exploration of non-verbal communication in driving contexts and its implications for AV design. Future research should address limitations in real-time data collection and conduct comprehensive validation studies of human-centric cues. These studies could employ trained actors in simulators, to investigate implicit understanding across demographics, determine optimal levels of abstraction for information presentation, and evaluate impacts on user trust, transparency, and predictability. Our methodology could inform the development and evaluation of interactive, cooperative driving mechanisms, with a focus on their effects on driving safety and user experience. We propose leveraging our dataset and participant elaborations to enhance AI training in AVs, particularly in natural language processing and large language models, for sophisticated interpretation, understanding and contextualisation of human cues and emotions. These directions aim to advance the development of more intuitive, user-friendly, and culturally adaptable automated driving systems.

7 Conclusion

This study investigated how passengers decode drivers' non-verbal cues during real-time driving interactions, providing insights into the implicit communication that occurs between drivers and passengers. By examining body language cues such as posture, head, eye, and hand movements, we uncovered patterns in how these cues are recognised, decoded and interpreted by passengers.

These findings suggest that subtle implicit non-verbal communication plays a significant role in cooperative driving. We highlight the importance of different cue types, including awareness cues, interaction cues, vestibular cues, and habitual cues. We show how each cue type contributes to how passengers perceive the driver's intentions and provide insights into how they may affect predictability and trust, feelings of safety, and confidence in driving competence. We discuss how these findings about human implicit communication in vehicles may inform the design of future human-centric HMIs for AVs, and guide the design of visual, haptics, auditory, and anthropomorphic approaches. By leveraging these observations, future AVs' HMIs could incorporate more natural, human-like interactions, offering intuitive feedback to drivers without overwhelming them.

Our study's implications extend beyond immediate AV interface design, encompassing broader applications in transportation systems. The precise timing and contextual execution of driver intention cues, crucial for passenger experiences, can inform communication strategies in human-centric navigational systems and automated public transport. These insights have potential applications in ride-sharing services, offering valuable guidance for passenger care and driver training programs [29, 46].

In conclusion, this study lays the groundwork for creating more intuitive, human-centred AV interfaces by understanding how passengers decode and interpret a human driver's implicit non-verbal cues. Just as the turn signal originated from explicit hand gestures incorporating these findings into future HMI designs, we can move closer to developing automated systems that foster trust, safety, and comfort through abstractions of implicit, human-like interactions.

Acknowledgments

We thank our colleagues who provided valuable feedback (Andrea McCrindle, David Rodwell, Michael Pascale, Sonali Nandavar, Xiaomeng Li) during the piloting of this study. We thank Melinda McDonald for her help in annotation of the large dataset. This research was supported partially by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (projects DP180103491 and DP220102598). The views expressed herein are those of the authors and are not necessarily those of the Australian Government or Australian Research Council. Large Language Models were utilised to help edit the writing of this work.

References

[1] Dora E Angelaki and Kathleen E Cullen. 2008. Vestibular system: the many facets of a multimodal sense. *Annu. Rev. Neurosci.* 31, 1 (July 2008), 125–150.
[2] Caterina Ansuini, Andrea Cavallo, Cesare Bertone, and Cristina Becchio. 2015. Intentions in the brain: the unveiling of Mister Hyde. *Neuroscientist* 21, 2 (April 2015), 126–135.
[3] Caterina Ansuini, Livia Giosa, Luca Turella, Gianmarco Altoè, and Umberto Castiello. 2008. An object for an action, the same object for other actions: effects on hand shaping. *Exp. Brain Res.* 185, 1 (Feb. 2008), 111–119.
[4] Erik Aremyr, Martin Jönsson, and Helena Strömberg. 2019. Anthropomorphism: An Investigation of Its Effect on Trust in Human-Machine Interfaces for Highly Automated Vehicles. In *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)*. Springer International Publishing, 343–352.
[5] Dare A Baldwin and Jodie A Baird. 2001. Discerning intentions in dynamic human action. *Trends in cognitive sciences* 5, 4 (2001), 171–178.
[6] Ray L Birdwhistell. 1952. Introduction to Kinesics. Louisville, Ky.
[7] Avishag Boker and Joel Lanir. 2023. Bird's Eye View Effect on Situational Awareness in Remote Driving. In *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ingolstadt Germany). ACM, New York, NY, USA, 36–41.
[8] Michael E Bratman. 1992. Shared Cooperative Activity. *The Philosophical Review* 101, 2 (April 1992), 327.
[9] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 2 (Jan. 2006), 77–101.
[10] Alexandra WD Bremers, Natalie Friedman, Sam Lee, Tong Wu, Eric Laurier, Malte F Jung, Jorge Ortiz, and Wendy Ju. 2024. (Social) Trouble on the Road: Understanding and Addressing Social Discomfort in Shared Car Trips. In *Proceedings of the 6th ACM Conference on Conversational User Interfaces*. 1–13.

[11] Judee K Burgoon, Valerie Manusov, and Laura K Guerrero. 2016. *Nonverbal Communication.* Routledge.

[12] Judee K Burgoon, Valerie Manusov, and Laura K Guerrero. 2021. *Coordinating Interaction.* Routledge, Chapter 12.

[13] Giulio Campagna and Matthias Rehm. 2024. A systematic review of trust assessments in human-Robot Interaction. *ACM Trans. Hum. Robot Interact.* (Nov. 2024).

[14] Chia-Ming Chang, Koki Toda, Daisuke Sakamoto, and Takeo Igarashi. 2017. Eyes on a car: An interface design for communication between an autonomous car and a pedestrian. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Oldenburg Germany). ACM, New York, NY, USA.

[15] Andy Crabtree, Peter Tolmie, and Mark Rouncefield. 2013. "how many bloody examples do you want?" fieldwork and generalisation. In *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus*. Springer London, London, 1–20.

[16] Stephanie Cramer and Jana Klohr. 2019. Announcing automated Lane changes: Active vehicle roll motions as feedback for the driver. *Int. J. Hum. Comput. Interact.* 35, 11 (July 2019), 980–995.

[17] Stephanie Cramer, Benjamin Miller, Karl-Heinz Siedersberger, and Klaus Bengler. 2017. Perceive the difference: Vehicle pitch motions as feedback for the driver. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (Banff, AB). IEEE.

[18] Sue E S Crawford and E Ostrom. 1995. A Grammar of Institutions. *Am. Polit. Sci. Rev.* 89 (Sept. 1995), 582–600.

[19] Aditya Dandekar, Lesley-Ann Mathis, Melanie Berger, and Bastian Pfleging. 2022. How to display vehicle information to users of automated vehicles when conducting non-driving-related activities. *Proc. ACM Hum. Comput. Interact.* 6, MHCI (Sept. 2022), 1–22.

[20] Amol Deshmukh, Bart Craenen, Alessandro Vinciarelli, and Mary Ellen Foster. 2018. Shaping robot gestures to shape users' perception: The effect of amplitude and speed on godspeed ratings. In *Proceedings of the 6th International Conference on Human-Agent Interaction*. ACM, New York, NY, USA.

[21] Anup Doshi and Mohan M Trivedi. 2011. Tactical driver behavior prediction and intent inference: A review. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. 1892–1897.

[22] Jens Edlund, Fredrik Edelstam, and Joakim Gustafson. 2014. Human pause and resume behaviours for unobtrusive humanlike in-car spoken dialogue systems. In *Proceedings of the EACL 2014 Workshop on Dialogue in Motion*. 73–77.

[23] P Ekman and W V Friesen. 1969. Nonverbal leakage and clues to deception. *Psychiatry* 32, 1 (Feb. 1969), 88–106.

[24] Stephen M Fiore, Travis J Wiltshire, Emilio J C Lobato, Florian G Jentsch, Wesley H Huang, and Benjamin Axelrod. 2013. Toward understanding social cues and signals in human-robot interaction: effects of robot gaze and proxemic behavior. *Front. Psychol.* 4 (Nov. 2013), 859.

[25] F Flemisch, D Abbink, M Itoh, M-P Pacaux-Lemoine, and G Weßel. 2016. Shared control is the sharp end of cooperation: Towards a common framework of joint action, shared control and human machine cooperation. *IFAC-PapersOnLine* 49, 19 (Jan. 2016), 72–77.

[26] Amr Gomaa, Robin Zitt, and Guillermo Reyes. 2023. Advancing dynamic hand gesture recognition in driving scenarios with synthetic data. In *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ingolstadt Germany). ACM, New York, NY, USA.

[27] Azra Habibovic, Victor Malmsten Lundgren, Jonas Andersson, Maria Klingegård, Tobias Lagström, Anna Sirkka, Johan Fagerlönn, Claes Edgren, Rikard Fredriksson, Stas Krupenia, Dennis Saluäär, and Pontus Larsson. 2018. Communicating Intent of Automated Vehicles to Pedestrians. *Front. Psychol.* 9 (Aug. 2018), 1336.

[28] Karlee J Hall, Karen Van Ooteghem, and William E McIlroy. 2023. Emotional state as a modulator of autonomic and somatic nervous system activity in postural control: a review. *Front. Neurol.* 14 (Aug. 2023), 1188799.

[29] Madeline J Hallewell, David R Large, Catherine Harvey, Leah Briars, Jenna Evans, Madelaine Coffey, and Gary Burnett. 2023. Deriving UX Dimensions for Future Autonomous Taxi Interface Design. *J. User Exper.* 17, 4 (June 2023), 140–163.

[30] The Anh Han. 2013. *Intention Recognition, Commitment and Their Roles in the Evolution of Cooperation: From Artificial Intelligence Techniques to Evolutionary Game Theory Models.* Springer, Berlin, Heidelberg.

[31] Renate Häuslschmid, Max von Bülow, Bastian Pfleging, and Andreas Butz. 2017. SupportingTrust in Autonomous Driving. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces* (Limassol, Cyprus) *(IUI '17)*. Association for Computing Machinery, New York, NY, USA, 319–329.

[32] Monique M Hennink, Bonnie N Kaiser, and Vincent C Marconi. 2017. Code saturation versus meaning saturation: How many interviews are enough? *Qual. Health Res.* 27, 4 (March 2017), 591–608.

[33] Wendy Ju. 2015. The Design of Implicit Interactions. *Synthesis Lectures on Human-Centered Informatics* 8, 2 (March 2015), 1–93.

[34] Wendy Ju. 2015. Driving into the Future, Together. In *The Design of Implicit Interactions*. Springer, 63–67.

[35] Yumin Kang, Seonga Choi, Eunsol An, Seokhyun Hwang, and Seungjun Kim. 2023. Designing Virtual Agent Human–Machine Interfaces Depending on the Communication and Anthropomorphism Levels in Augmented Reality. In *Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ingolstadt, Germany) *(AutomotiveUI '23)*. Association for Computing Machinery, New York, NY, USA, 191–201.

[36] Nihan Karatas, Shintaro Tamura, Momoko Fushiki, and Michio Okada. 2018. The Effects of Driving Agent Gaze Following Behaviors on Human-Autonomous Car Interaction. In *Social Robotics*. Springer International Publishing, 541–550.

[37] Andras Kemeny and Francesco Panerai. 2003. Evaluating perception in driving simulation experiments. *Trends Cogn. Sci.* 7, 1 (Jan. 2003), 31–37.

[38] Taesu Kim, Gyunpyo Lee, Jiwoo Hong, and Hyeon-Jeong Suk. 2023. Affective role of the future autonomous vehicle interior. In *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ingolstadt Germany). ACM, New York, NY, USA.

[39] Satoshi Kitazaki and Matthias J Myhre. 2015. Effects of non-verbal communication cues on decisions and confidence of drivers at an uncontrolled intersection. In *Driving Assessment Conference*, Vol. 8. University of Iowa, University of Iowa, Iowa City, Iowa.

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

[41] Johannes Maria Kraus, Florian Nothdurft, Philipp Hock, David Scholz, Wolfgang Minker, and Martin Baumann. 2016. Human After All: Effects of Mere Presence and Social Interaction of a Humanoid Robot as a Co-Driver in Automated Driving. In *Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ann Arbor, MI, USA) *(AutomotiveUI '16 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 129–134.

[42] Johannes Maria Kraus, Jessica Sturn, Julian Elias Reiser, and Martin Baumann. 2015. Anthropomorphic agents, transparent automation and driver personality: towards an integrative multi-level model of determinants for effective driver-vehicle cooperation in highly automated vehicles. In *Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Nottingham, United Kingdom) *(AutomotiveUI '15)*. Association for Computing Machinery, New York, NY, USA, 8–13.

[43] David R Large, Madeline Hallewell, Xuekun Li, Catherine Harvey, and Gary Burnett. 2023. I, AV: A ghost driver field study exploring the application of anthropomorphism in AV-pedestrian communication. In *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ingolstadt Germany). ACM, New York, NY, USA.

[44] Zheng Ma and Yiqi Zhang. 2024. Driver-automated vehicle interaction in mixed traffic: Types of interaction and drivers' driving styles. *Human factors* 66, 2 (2024), 544–561.

[45] Sujitha Martin, Ashish Tawari, Erik Murphy-Chutorian, Shinko Y Cheng, and Mohan Trivedi. 2012. On the Design and Evaluation of Robust Head Pose for Visual User Interfaces: Algorithms, Databases, and Comparisons. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Portsmouth, New Hampshire) *(AutomotiveUI '12)*. ACM, New York, NY, USA, 149–154.

[46] Johanna Meurer, Christina Pakusch, Gunnar Stevens, Dave Randall, and Volker Wulf. 2020. A Wizard of Oz Study on Passengers' Experiences of a Robo-Taxi Service in Real-Life Settings. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 1365–1377.

[47] Vanina Mollo and Pierre Falzon. 2004. Auto- and allo-confrontation as tools for reflective activities. *Appl. Ergon.* 35, 6 (Nov. 2004), 531–540.

[48] Judith A Ouellette and Wendy Wood. 1998. Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychol. Bull.* 124, 1 (July 1998), 54–74.

[49] Alexandros Rouchitsas and Håkan Alm. 2019. External human-machine interfaces for autonomous vehicle-to-pedestrian communication: A review of empirical work. *Front. Psychol.* 10 (Dec. 2019), 2757.

[50] Luisa Sartori, Cristina Becchio, and Umberto Castiello. 2011. Cues to intention: the role of movement information. *Cognition* 119, 2 (May 2011), 242–252.

[51] Shane Saunderson and Goldie Nejat. 2019. How robots influence humans: A survey of nonverbal communication in social human–robot interaction. *Int. J. Soc. Robot.* 11, 4 (Aug. 2019), 575–608.

[52] Bridie Scott-Parker. 2017. Nonverbal communication during the learner lesson with a professional driving instructor: a novel investigation. *Transportation research part F: traffic psychology and behaviour* 47 (2017), 1–12.

[53] Jonathan A Smith and Megumi Fieldsend. 2021. Interpretative phenomenological analysis. In *Qualitative research in psychology: Expanding perspectives in methodology and design (2nd ed.)*. American Psychological Association, Washington, 147–166.

[54] Vicki Squires. 2023. Reflexive Bracketing. In *Springer Texts in Education*. Springer International Publishing, Cham, 425–430.
[55] Angelica M Tinga, Ilse M van Zeumeren, Michiel Christoph, Elmer van Grondelle, Diane Cleij, Anna Aldea, and Nicole van Nes. 2023. Development and evaluation of a human machine interface to support mode awareness in different automated driving modes. *Transp. Res. Part F Traffic Psychol. Behav.* 92 (Jan. 2023), 238–254.
[56] Claude Toussaint, Philipp T Schwarz, and Markus Petermann. 2023. Navel - a social robot with verbal and nonverbal communication skills. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–4.
[57] James P Trujillo, Irina Simanova, Harold Bekkering, and Asli Özyürek. 2020. The communicative advantage: how kinematic signaling supports semantic comprehension. *Psychol. Res.* 84, 7 (Oct. 2020), 1897–1911.
[58] Jacqueline Urakami and Katie Seaborn. 2023. Nonverbal cues in human–robot interaction: A communication studies perspective. *ACM Transactions on Human-Robot Interaction* 12, 2 (2023), 1–21.
[59] Valeria Villani, Fabio Pini, Francesco Leali, and Cristian Secchi. 2018. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics (Oxf.)* 55 (Nov. 2018), 248–266.
[60] Kai Vogeley and Gary Bente. 2010. "Artificial humans": Psychology and neuroscience perspectives on embodiment and nonverbal communication. *Neural Netw.* 23, 8-9 (Oct. 2010), 1077–1090.
[61] Jianmin Wang, Yujia Liu, Tianyang Yue, Chengji Wang, Jinjing Mao, Yuxi Wang, and Fang You. 2021. Robot Transparency and Anthropomorphic Attribute Effects on Human-Robot Interactions. *Sensors* 21, 17 (Aug. 2021).
[62] Jianmin Wang, Chengji Wang, Yujia Liu, Tianyang Yue, Yuxi Wang, and Fang You. 2022. Multimodal Warnings Design for In-Vehicle Robots under Driving Safety Scenarios. *Sensors* 23, 1 (Dec. 2022).
[63] Meng Wang, Shashank Mehrotra, Nicholas Wong, Shannon C Roberts, Woon Kim, Alicia Romo, William J Horrey, et al. 2024. Human-Machine interfaces and vehicle Automation: A review of the literature and recommendations for system Design, Feedback, and alerts. *Transportation Research Part F: Traffic Psychology and Behaviour* 107 (2024), 549–561.
[64] Wenshuo Wang, Letian Wang, Chengyuan Zhang, Changliu Liu, Lijun Sun, et al. 2022. Social interactions for autonomous driving: A review and perspectives. *Foundations and Trends® in Robotics* 10, 3-4 (2022), 198–376.
[65] Kenton Williams, José Acevedo Flores, and Joshua Peters. 2014. Affective Robot Influence on Driver Adherence to Safety, Cognitive Load Reduction and Sociability. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Seattle, WA, USA) *(AutomotiveUI '14)*. Association for Computing Machinery, New York, NY, USA, 1–8.
[66] Kun Xu, Mo Chen, and Leping You. 2023. The Hitchhiker's Guide to a Credible and Socially Present Robot: Two Meta-Analyses of the Power of Social Cues in Human–Robot Interaction. *Int J of Soc Robotics* 15, 2 (Feb. 2023), 269–295.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009