

Case Studies 2 – Designing Hybrid Futures

Designing for Interactions with Automated Vehicles: Ethnography at the Boundary of Quantitative-Data-Driven Disciplines

MARKUS ROTHMÜLLER

*School of Architecture, Design and Planning, Aalborg University Copenhagen, Denmark
and Shift Insights & Innovation Consulting*

PERNILLE HOLM RASMUSSEN

School of Architecture, Design and Planning, Aalborg University Copenhagen, Denmark

SIGNE ALEXANDRA VENDELBO-LARSEN

School of Architecture, Design and Planning, Aalborg University Copenhagen, Denmark

This case study presents ethnographic work in the midst of two fields of technological innovation: automated vehicles (AV) and virtual reality (VR). It showcases the work of three M.Sc. Techno-Anthropology students and their collaboration with the EU H2020 project ‘interACT’, sharing the goal to develop external human-machine interfaces (e-HMI) for AVs to cooperate with human road users in urban traffic in the future. The authors reflect on their collaboration with human factor researchers, data scientists, engineers, experimental researchers, VR-developers and HMI-designers, and on experienced challenges between the paradigms of qualitative and quantitative research. Despite the immense value of ethnography and other disciplines to collectively create holistic representations of reality, this case study reveals several tensions and struggles to align multi-disciplinary worldviews. Results show the value of including ethnographers: 1) in the design and piloting of a digital observation app for the creation of large datasets; 2) in the analysis of large amounts of data; 3) in finding the potential of and designing e-HMI concepts; 4) in the representation of real-world context and complexity in VR; 5) in the evaluation of e-HMI prototypes in VR; and finally 6) in critically reflecting on the construction of evidence from multiple disciplines, including ethnography itself.

So (...) perhaps for the first time there is a dataset, where the ontology is with the subjects in it rather than anyone else trying to project things onto that. So that appears to be quite a new thing for data science. We still don't understand what benefits it will have, but it's an unusual with a dataset that has been conceptualized in the subject and terms of ethnographic principles, rather than the experimental way. (...) it seems as if it should be a good thing, but we don't know yet exactly why.

(Data scientist from University of Leeds reflecting on the collaboration with the techno-anthropologists, 2018)

INTRODUCTION AND PROBLEM SPACE

The development of AVs is complex and demanding in regards to technical feasibility and social acceptance among road users. The successful integration of AVs will depend on the ability of the technological system to cope with the social and contextual complexity of mixed-traffic interactions in the future (Vinkhuyzen and Cefkin, 2016). Therefore, research on human negotiations in such interactions will be crucial for AVs to safely navigate through urban spaces and to predict intentions and future maneuvers of human road users. To be

truly cooperative, though, AVs will have to be able to communicate with human road users, which is why the development of e-HMI is one of the currently hottest topics regarding the safe, efficient, and socially accepted use of AVs in urban spaces.

Designing external Human-Machine Interfaces for Autonomous Vehicles

To design human-machine interfaces for future interactions, we need to perceive the interaction through the eyes of the human and through the eyes of the machine and analyze their shared understanding (Suchman 2007, p. 123). Additionally, we have to understand the hybrid relation between the human, the technology and the world they interact in (Verbeek, 2015). To achieve this, research has to investigate traffic interactions from several perspectives: the one of the pedestrian (or other human road user) making sense of driver- and vehicle behavior today; the one of the driver making sense of human road user behavior today; and the one of the AV making sense of human road user behavior in the future. Representing these perspectives and the in real-time interrelated and fluid decision-making processes of each interaction participant, situated in a real-world context, creates the basis for the design of e-HMI concepts.

To start with, research needs to find out what information pedestrians seek for when scanning their surroundings and what information they use when interacting with drivers and vehicles to understand whether it is safe to cross a road or not, today (Schieben et al., 2018). So far, interaction studies have shown that in many situations pedestrians mainly focus on the speed and distance of the approaching cars to inform their crossing decision (Rasouli & Tsotsos, 2018; Dey & Terken 2017; Merat & Madigan, 2017; Clamann et al., 2017; Yannis et al., 2013; Kadali & Perumal, 2012; Cherry et al., 2012). The generalized suspicion is that in many situations pedestrians will not require any additional signals provided by e-HMI, as little to no interaction will take place anyway.

Conversely this means, the first step is to find out when and how additional e-HMI signals might actually add value to the interaction between AVs and pedestrians. One natural starting point is to investigate today's traffic interactions between pedestrians and drivers, the interplay of which might indicate potentials for e-HMI to substitute for the missing driver in the future. The important point is to understand how the perspectives and decision-making processes of drivers and pedestrians inform each other. Once situations where pedestrians make use of additional information from drivers (e.g. where they look, hand gestures, etc.) are identified, the value of e-HMI could be to substitute the information from the missing driver in AVs in the future (Schieben et al., 2018; Nathanael et al., 2018; Wilbrink et al., 2018; Chang et al., 2017; Mahadevan et al., 2017; Merat et al., 2016; Parkin et al., 2016; Langström & Lundgren, 2015). To truly design for interactions with AVs and to provide e-HMI signals in the right moments, however, we need to understand how the AV's artificial intelligence (AI) will make sense of these situations in the future. It is important to know what information the AV detects from a traffic situation and interacting human road users and how it processes this information, as this will be the basis for AI-based decision-making and real-time communication of signals (Drakoulis et al., 2018, p.9; Wilbrink et al., 2018, p.11).

Therefore, it is important to represent the perspective and decision-making process of pedestrians and drivers in interactions today, as well as the perspective and decision-making process of AVs in the future. The real challenge, however, is not to understand these

perspectives and decision-making processes individually. The real challenge is to find proper ways to represent these perspectives and to then align these representations as a form of holistic evidence providing the basis to design and development e-HMI. This is particularly challenging when AVs have neither been fully developed yet nor tested in urban spaces. It makes the investigation of human-AV interactions purely based on today's interactions and speculative best guesses of how these interactions might turn out in a future with AVs (Cefkin & Stayton, 2017).

Evaluating e-HMI Concepts in Virtual Reality

In addition to understanding traffic interactions and designing e-HMI concepts, it is important to evaluate their effects on pedestrians' decision-making processes when interacting with AVs in the future. Since the evaluation in naturalistic city traffic poses a risk to traffic safety and is still restricted by legal frameworks in most places, simulations in VR are an often preferred alternative method of investigation. VR simulations are not only cheaper and safer (Blissing, 2016; Sobhani & Farooq, 2018), but offer an ideal platform for experimental research (Wilson & Soranzo, 2015), which then makes VR an obvious choice to evaluate and measure the effect of e-HMI. Just as any other research method, though, VR experiments have their limitations. As anthropologists we see two profound challenges. The first is that experiments follow a reductionist approach, breaking down the complexity of real-world contexts and social phenomena such as interaction behavior into simpler elements which in turn should enable an understanding of cause and effect. In praxis this means that the complexity of naturalistic traffic situations gets reduced to a considerably simplified representation of reality built in VR. To properly evaluate the effect that e-HMI causes in simulated future traffic interactions, then, all other influence factors that could be found in naturalistic traffic interactions today get reduced to a minimum as they would make the analysis of cause and effect increasingly complex. And even though VR simulations are typically only used as early indicators for whether a design concept works or not, they still do influence decisions in the development of technological products being eventually used in complex real-world contexts. Hence, the challenge here is how to infuse 'important elements' of the complexity and context of naturalistic settings in VR that have been identified as the very basis for why e-HMI might be valuable in the first place.

The second challenge with experimental evaluations of e-HMI in VR is related to cultural and local differences influencing traffic interactions. Even if we manage to infuse features of real-world complexities in VR experiments, the ecological validity of results from these experiments remains highly questionable in terms of applicability to multiple, socioculturally varying real-world environments (Rasouli & Tsotsos, 2018; Schieben, 2018; Stayton, Cefkin & Zhang, 2017; Turkle, 2009). Although some virtual experiments have argued that behavioral differences between virtual- and real-world environments were little to not at all noticeable (e.g. Bhagavathula et al., 2018), one of the most comprehensive surveys on pedestrian behavior studies just recently concluded that contradictions in such generalizations root in "*variations in culture (...) and interrelationships between the factors*" being studied in isolation (Rasouli & Tsotsos, 2018). In simpler words, e-HMI will have to work in several traffic cultures and it is particularly difficult to receive results from VR experiments that can be scaled up and seen as generally valid for different cultures and contexts. Nevertheless, ecological validity and applicability to the real-world complexity and context,

needs to inform technological development just as much as cause and effect studied in experiments under simplified conditions of reality.

Multidisciplinary Representations of Reality

The two previous subsections described the problem space in the prototyping of e-HMI and outlined some of the major challenges in regards to representing reality, both, in the quest of understanding today’s and future traffic interactions as well as in the pursuit of methods for evaluating e-HMI prototypes in VR experiments. This subsection introduces some of the complexity and challenges of multiple research disciplines collaborating in this problem space. As the visualization below shows, we first collaborated with interACT’s human factor researchers, engineers, data scientists, and human behavior modelers in order to investigate today’s traffic interactions in naturalistic settings, and then with experimental researchers, engineers, HMI designers and VR developers to investigate simulated future interactions to evaluate our e-HMI concepts in VR.

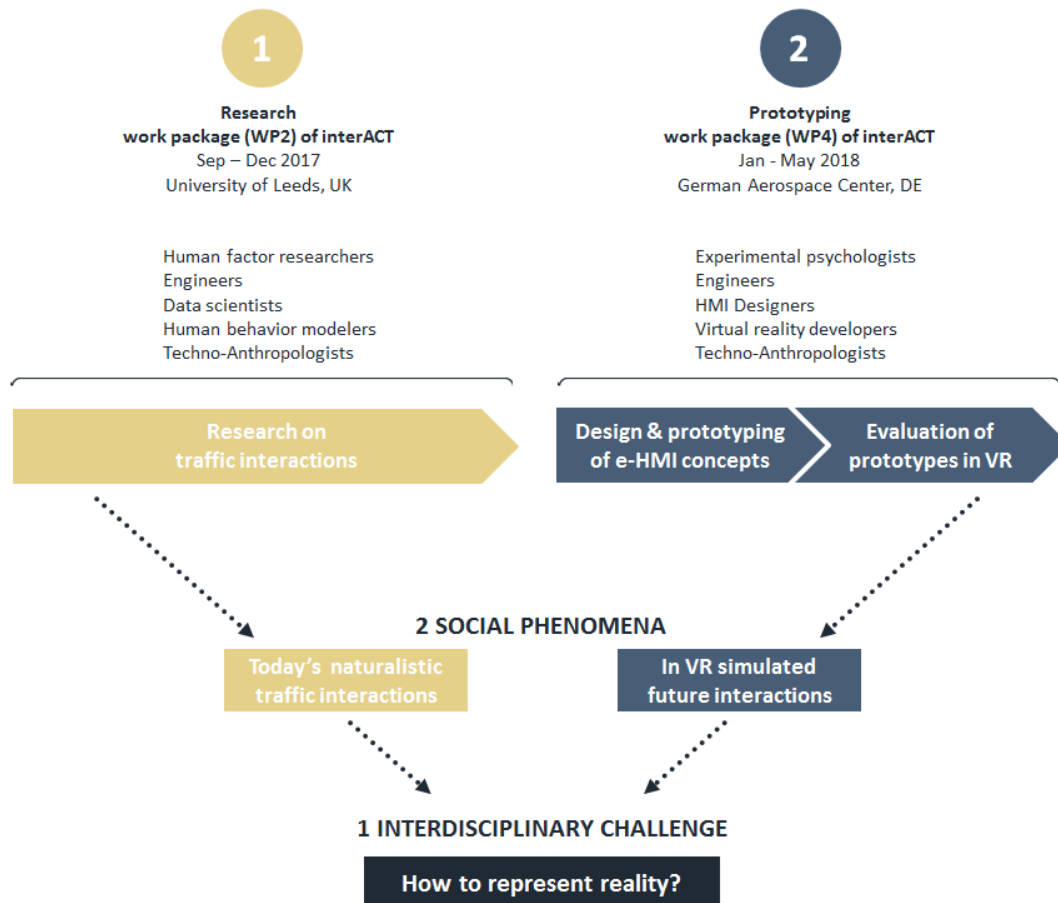


Figure 1. Two collaborations, two phenomena, one central interdisciplinary challenge

Hence, we conducted research on two social phenomena at the core of our engagement with interACT. And even though this research took place at two entirely different locations in different periods with different project teams of interACT, we experienced the same profound interdisciplinary challenge, which was how to represent reality when investigating traffic interactions.

Over the past decades, the often discussed qualitative vs. quantitative dichotomy has opened up for more tolerance between the two paradigms (Sale, Lohfeld & Brazil, 2002). This, however, does not imply that the production of more holistic, interdisciplinary forms of evidence to describe a social phenomenon is an easy task. The human factor researchers, data scientists, engineers, and experimental psychologists that we collaborated with, for example, knew little to nothing about ethnography or anthropological research outside of the realm of adventurous research studying native tribes on tropical islands. This is not necessarily surprising, since anthropologists and ethnographers, are still rarely seen in research directly informing the technical development of AVs¹. In contrast, the fact that psychology-based human factor research has developed strong links to engineering disciplines over the past decades, roots in the compatibility of representing reality through a similar quantitative lens that proves useful to engineering and product design (Stanton et al., 2013). And even though human factor researchers and experimental psychologists do also work qualitatively, the concept of qualitative induction or ethnography is typically not part of their training (McNamara et al., 2015). In our case, we experienced that, although studying the same phenomena, our ways of producing evidence to describe the phenomenon at hand were inherently different to the ones of our collaborators. Of course all of us referred to the same phenomenon (traffic interactions), yet we did so by following different paradigmatic assumptions (Sale, Lohfeld & Brazil, 2002). Eventually, one could say we talked about different things - different representations of reality. This roots in the fact that both paradigms follow two profoundly different worldviews (ontologies) in their attempts to represent reality (ibid.). The quantitative worldview believes in one single objective reality that can be observed and described independent of the perspective one chooses to perceive this reality (ibid.). The qualitative worldview, on the other hand, argues that reality is constructed by the very perspective one chooses to look at reality. Thereby each perspective on reality essentially represents an own (part of) reality. In our case, the real challenge of developing e-HMI concepts, thus, was how to represent reality qualitatively and quantitatively, from the insider's and outsider's perspective².

More broadly, this addresses the challenge that the development of human-machine/human-robot interfaces does not only depend on 1) the investigation of the perspectives of each interaction partner (e.g. pedestrian and driver today changing to pedestrian and AV in the future) and their relation to each other as well as to the context in which the interaction takes place, but also on 2) how research represents the reality from an outsider's and the insider's perspective of each interaction participant in, both, real- and virtual-environments. For example, researchers investigating the reality from an outsider's perspective only, might entirely reject the idea that we need to understand interactions through the insider's perspective of the pedestrian, the driver and the AV, as there is only one objective reality anyway and that is best to be understood from an external perspective. The central question, thus, was: How to represent reality?

HOW TO REPRESENT REALITY

To investigate today's traffic interactions, interACT ran two studies: one using human observations, LiDARs and video recordings in Athens, Leeds and Munich, and one using eye-tracking hardware on drivers in Athens to analyze driver-pedestrian interactions. We, techno-anthropologists, supported the large-scale observational study of interACT in Leeds, and decided to run an in-depth ethnographic study using observations and semi-structured interviews in parallel in the same location. The goal was to understand pedestrians' decision-making processes when interacting with drivers in more detail. As the visualization below shows, each study had its own focus and, thereby, represented reality from a different perspective.

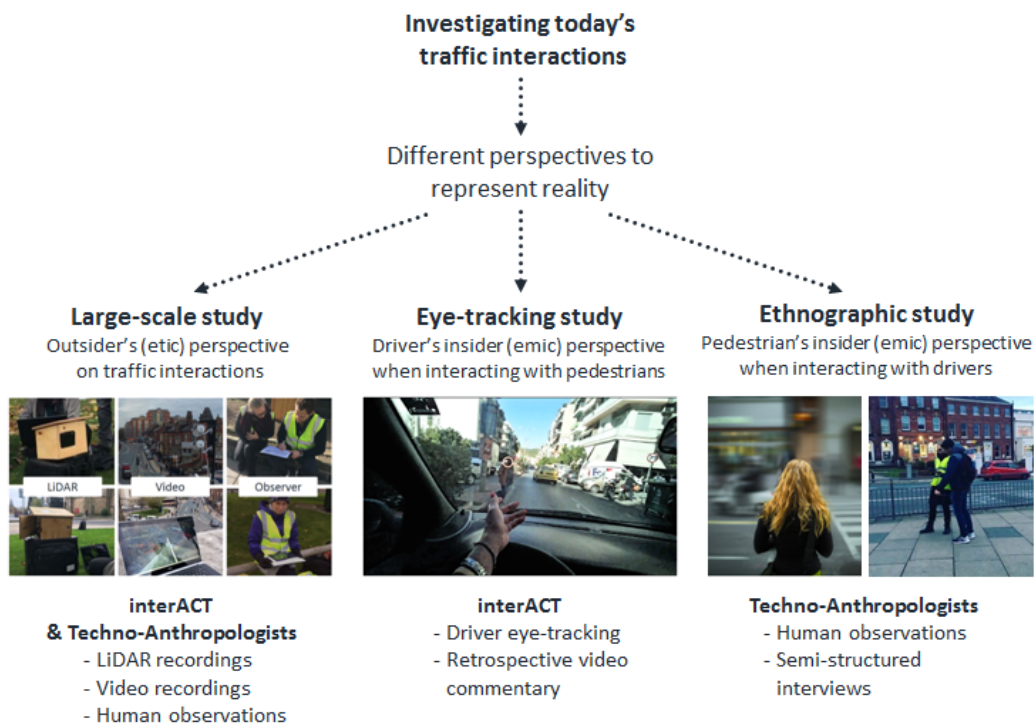


Figure 2. Different perspectives to represent today's traffic interactions (source: 2nd picture on the right: Adobe Spark-Free Pictures; middle: Nathanael et al., 2018; rest: ©interACT2017, used with permission)

The cross-cultural, large-scale study using LiDARs, video cameras and human observations provided an outsider's perspective on traffic interactions. The study applying eye-tracking and retrospective video commentary provided an insider's perspective on driver decision-making when interacting with pedestrians and our ethnographic study offered an insider's perspective on pedestrian decision-making when interacting with drivers. In theory the combination of all these perspectives would be exactly what we need in order to progress the development process of e-HMI, as they provide all the perspectives we described as

being needed to design for future interactions. As can be seen, though, not only did each study view reality from a different perspective, but also did they follow completely different methodologies to represent these perspectives. Before aligning any of the results, it is crucial to reflect on the production of evidence of each approach first. As we will show, the devil lies in the detail.

In the following, we will reverse-engineer the process of producing and utilizing evidence to inform the development of e-HMI concepts and active vehicle controllers. We will present the boundaries and achievements of interdisciplinary work, and through this process address limitations, so we can reflect upon these and offer a baseline for discussions on how ethnographers, data scientists, engineers, human factor researchers and experimental psychologists can collaborate in creating more holistic forms of evidence to inform technological innovation in the future.

Representing the Outsider's Perspective: Data Scientists Using Game Theory Models on a Human Observations Dataset

While interACT is still working on combining the different datasets of each approach, the large-scale observation dataset has already been used for different types of analysis. For example, it currently serves data scientists as naturalistic data to train game theoretic models, which will eventually be implemented in active real-time vehicle controllers of AVs. This is not directly related to the interACT project but runs as a parallel spin-off project led by a data science team at the University of Leeds, UK. They view *“pedestrians as active agents having their own utilities and decisions”* which need to be *“inferred and predicted by AVs in order to control interactions with them and navigation around them”* (Camara et al., 2018). Drawing on game theoretic models, they perceive interactions as a competition for road-space between pedestrians and AVs (ibid.). They used the human observation data *“from real-world human road crossings to determine what features of crossing behaviors are predictive about the level of assertiveness of pedestrians and of the eventual winner of the interactions”* (ibid.). To achieve this, they followed a reductionist approach of *“decomposing pedestrian-vehicle interactions into sequences of independent discrete events”* and applied probabilistic methods (logistic- and decision tree regression) as well as *“sequence analysis (...) to find common patterns of behavior and to predict the winner of the interaction”* (ibid.). How this eventually looked like is visualized in the following pictures.

The analysis process of the data scientists (picture 3) led to the creation of two types of evidence: top-10 n-grams (picture 4) and a decision tree for pedestrian-vehicle interactions (picture 5). N-grams (motifs) are the outcome of a process called motif selection which was used by the data scientists to identify *“common short sequences of events which tend to occur together”* (Camara et al., 2018), meaning they tried to define the ten most often co-occurring combinations of behavioral elements in the observed interactions between pedestrians and vehicles. Decision trees were then used to find out which of the single behavioral elements and motifs are informative to predict the winner for each interaction, while providing a *“visualization helpful for human interpretation, and a fast method for real-time systems such as AVs to make decisions based on a few variables”* (ibid.).

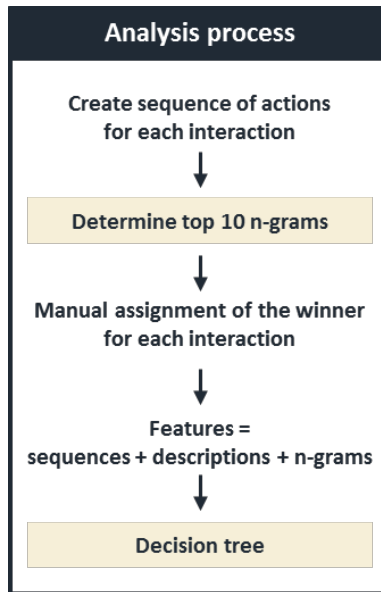


Figure 3. Analysis process of the data scientists (source: Camara et al., 2018)

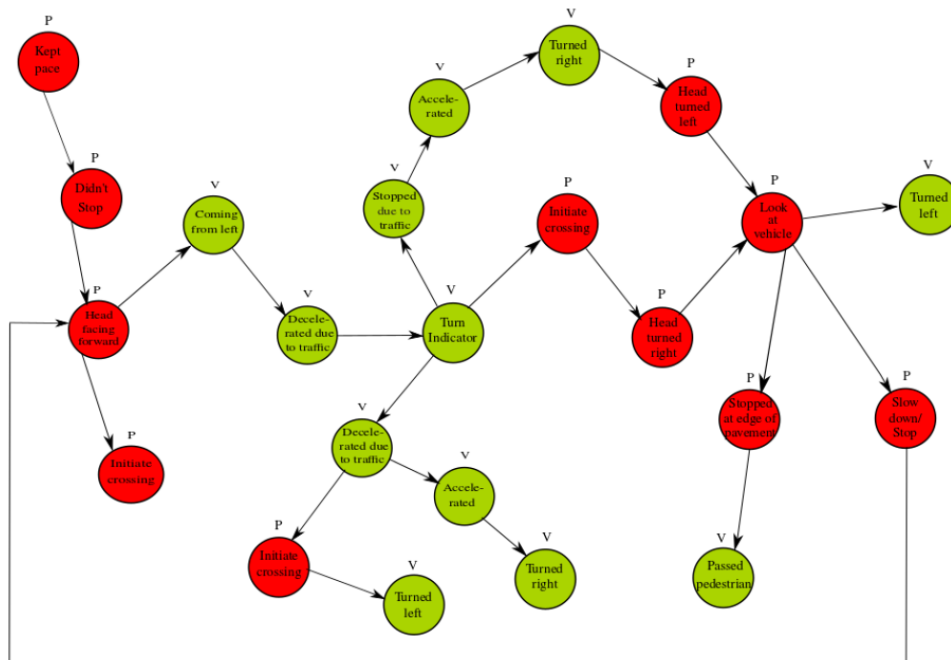


Figure 4. Example evidence of data scientists: N-grams (source: Camara et al., 2018)

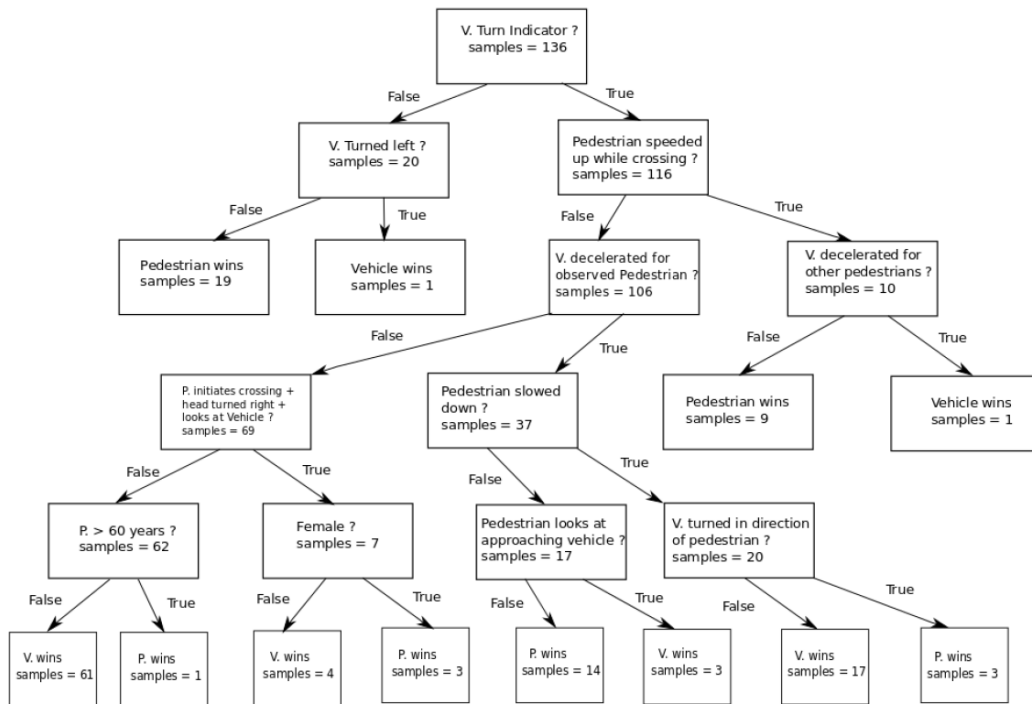


Figure 5. Example evidence of data scientists: Decision Tree (source: Camara et al., 2018)

As stated in the last quote, the described examples of data science evidence built the basis for human interpretation, which led the data scientists to conclude that pedestrians - generally - “(...) seek for cues in the vehicle’s motion [but] not in eye contact with or gestures by the driver” (ibid.) since:

(...) the data collection observers themselves did not record much information about the driver gestures because they were difficult to see. So eye gaze by the pedestrian is important, but eye contact with the driver or AV is not (...). These findings are important for AV design as they suggest that AVs should also be designed to communicate simply via their position on the road (...) but maybe not needing artificial face, eye, or gesture substitutes; and that they do need to detect and process pedestrian faces and eyes in order to inform their interactions. (ibid.)

We present this quote, because it introduces the demand for critical reflections on the construction and interpretation of evidence, which this case study attempts to provide in the first place. The line of argument shows the interpretation of evidence from the data science approach that pedestrians - generally - do not seek cues provided by the driver, since we, observers, could often not see the driver from our position, and therefore noted ‘pedestrian looked at the vehicle’ rather than ‘pedestrian looked at driver’. In the following we will show why this interpretation-based generalization was more of a misconception of what the data could actually tell, rather than any sort of objective representation of reality.

The Construction of Evidence: What Data Tells and What Not

To collect comparable observation data in three different cities in a replicable manner, interACT decided to use a standardized observation protocol. Additionally, interACT decided to develop an app to run through the observation protocol on a tablet, which allowed us to transform the expected large amounts of data into digital formats that could directly be used for different types of analysis later. As trained ethnographers with previous experience in conducting observational studies, our role was to support the development of these digital standard observation protocols. Therefore, interACT provided us with a first draft based on else-where published research and hypotheses on what would be important for the agenda of interACT including the one of the data scientists.



Figure 6. Development of the protocols with pen and paper. Picture 7. Observations with the digital protocols on a tablet. (source: Rasmussen, Rothmüller & Vendelbo-Larsen, 2017)

Our goal was to run a pilot study - a proof of concept of the first protocol draft - following an exploratory observation approach, which should eventually avoid forcing interactions into predefined categories. At first, we struggled to design a 'one-size-fits-all' observation protocol, but ultimately managed to provide the observer with a bit more flexibility by including a space for sketches and additional notes. This enabled the observer to describe interactions beyond the limitations of predefined categories only.

Digital Observation Protocol - App Page 1

Pedestrian - Vehicle - Interaction Observation Protocol

Approaching Phase

Pedestrian Analysis						
Movements while approaching	Slowed down []	Kept pace []	Speeded up []	Stopped at the edge of the pavement []	Stepped on road and stopped []	Did not stop []
Head movements	Turned left []	Turned right []	None / Facing forward []			
Looking at other RU's	Looked at approaching vehicle []	Looked at other pedestrian(s) entering the road []	Others (elaborate in notes) []	None <input type="checkbox"/>	Not observable <input type="checkbox"/>	
Hand movements add meaning of hand movement in notes section	Waved hand []	Raised hand in front []	Raised hand sideways []	Other (elaborate in notes) []	None <input type="checkbox"/>	

Driver / Vehicle Analysis						
Interacting Vehicle	Car []	Motorcycle []	Van []	Bus / Truck []	Other (elaborate in notes) []	None <input type="checkbox"/>
Vehicle approached from...	From left []	From right []	Single <input type="checkbox"/>	Multiple <input type="checkbox"/>		
Vehicle Movement	Decelerated for observed pedestrian []	Decelerated due to other pedestrian []	Decelerated due to traffic []	Accelerated []	Turned left []	Passed the pedestrian []
	Stopped for observed pedestrian []	Stopped due to other pedestrian []	Stopped due to traffic []	Kept pace []	Turned right []	Other (elaborate in notes) []
Used signals add the meaning of used signals in notes section	Honked []	Flashed Lights []	Turn Indicator []	Other []	None <input type="checkbox"/>	
Head movements add meaning of hand movement in notes section	Turned left []	Turned right []	Turned in the direction of the pedestrian []	Other (elaborate in notes) []	None <input type="checkbox"/>	Not observable <input type="checkbox"/>
Hand movements add meaning of hand movement in notes section	Waved hand []	Raised hand in front []	Raised hand sideways []	Other (elaborate in notes) []	None <input type="checkbox"/>	

Figure 7. App-mockup of digital observation protocol – Section for interaction behavior. (©interACT2017, used with permission)

Digital Observation Protocol - App Page 4

Symbols to pull into

<p> = Pedestrian</p> <p> = Group</p> <p> = Movement</p> <p> = Brake</p>	<p> = Vehicle 1</p> <p> = Vehicle 2</p> <p> = Vehicle next to interaction</p> <p> = Intended movement (only if intended action was different to actual action by the pedestrian)</p>
---	--

Figure 8. App-mockup of digital observation protocol – Section for sketching. (©interACT2017, used with permission)

Digital Observation Protocol - App Page 3

Pedestrian - Vehicle - Interaction Observation Protocol

Additional Observation Notes:

General Information

Date: _____
 Intersection: _____

Weather	Sunny <input type="checkbox"/>	Overcast <input type="checkbox"/>	Raining <input type="checkbox"/>	Freezing / icy <input type="checkbox"/>		
Single Pedestrian	Individual <input type="checkbox"/>	m / f <input type="checkbox"/>				
Group	Group: <input type="checkbox"/>	No. of m <input type="text"/>	Was the observed pedestrian the leader of the group? Yes <input type="checkbox"/> No <input type="checkbox"/>			
Age	Child (<13 years) <input type="checkbox"/>	Teenager (13-18 years) <input type="checkbox"/>	Young Adult (18-30 years) <input type="checkbox"/>	Midage Adult (30-60 years) <input type="checkbox"/>	Older adult (>60 years) <input type="checkbox"/>	
Potential Distraction	Listening to headphones <input type="checkbox"/>	Talking on mobile phone <input type="checkbox"/>	Looking at mobile phone / handheld device <input type="checkbox"/>	Was wearing a hoddie <input type="checkbox"/>	Other (elaborate in notes) <input type="checkbox"/>	None <input type="checkbox"/>

Figure 9. App-mockup of digital observation protocol – Section for observation notes. (©interACT2017, used with permission)

This exploratory pilot study led to a final app that allowed us and other observers from interACT to create one of “*the largest and most detailed dataset of its kind*” (Camara et al, 2018), representing nearly 1000 traffic interactions, each through 62 individual behavioral elements (e.g. behavior expressed through motion, signals, head or hand movements) combined as a sequence, and 12 environmental influence factors. Each human observation of interactions in Athens, Leeds and Munich was, therefore, represented by the exact same selection of behavioral elements - which was important for cross-cultural comparability and quantitative analyses such as the ones of the data scientists.

So in that sense, interACT succeeded in following a replicable observation approach, which, nevertheless, did not come without limitations and implications. Even though, the pilot study was successful in enabling the protocol to capture most of the behavior in interactions, it still reduced the complexity of real-world interactions to a simpler representation of reality. First of all, the fact that the protocol primarily focused on one-to-one interactions, considering other actors in the environment as an environmental factor rather than another interaction participant, was one of the first steps to reduce complexity. Secondly, any form of interaction behavior that was not included or sufficiently well described as a category in the protocol had to be added in the notes field, which has not been part of any analysis so far. Thirdly, the fact that these observations only provided an outsider’s perspective on the reality of which the observed pedestrian, vehicle and driver were part of means the dataset does not allow any inference of internal perceptions of these interaction participants. In relation to this we had multiple discussions with our collaborators about what is observable, what is an objective observation, and what a subjective

interpretation. These very discussions defined the way we, observers, described interactions, and thereby predefined the dataset's representation of reality.

It makes the whole observational study a subjective choice of how much of reality should be described in the dataset. As much as one would like to promote objectivity by controlling or excluding subjective interpretations, we argue that any decision on how much of the perceivable reality should be represented in an observation is just as subjective. Because, isn't the choice of representing certain elements of reality in a limited way just as influential for the construction of evidence, as the choice to represent certain elements of reality as an interpretation of trained observers? Eye contact, for example, was genuinely excluded from the protocol, as this was categorized as not observable. And it's true; how could you know that a pedestrian had eye contact with a driver, if you are not either that very pedestrian or driver? Of course we could have asked them whenever we had the suspicion, but we didn't. Instead, it was decided to observe whether the pedestrian looked at the driver, or looked at the car, and whether the driver looked at the pedestrian. In many cases, these observations were based on the direction of head turns of pedestrians, which made it still quite difficult to know whether they actually looked at the driver or the vehicle. This essentially means that we only noted down that pedestrians looked at the driver whenever this was very clear to us, which in turn does not mean that pedestrians who we observed as having looked at the vehicle might not also have look at the driver. Thus, the earlier shown conclusion of the data scientists' publication that eye gaze from the pedestrian is important but from the driver or AV is not, and that eye contact and the substitution of this through e-HMI, therefore, is generally not needed was rather a misinterpretation of what the data could actually tell. We could not know whether eye contact or where the driver looked at was important to pedestrians because we did not include their perception in the data collection. This means that we, ethnographers, could not only really contribute to piloting the very instrument to collect data for data scientists, but also to explain and reflect on what the collected data actually tells.

Understanding the dataset and meaning of single data points is as essential as being able to structure and analyze it. In fact, this seemed to have kick-started a reflection process in the broader research work of our collaborators. The data scientists, for example, told us that our contribution to understanding the dataset correctly, also made them question the representation of reality from other datasets they worked with.

Though the need for objective observations was to some extent met; it is important to be aware that any decision taken in relation to how data gets collected eventually influences how reality gets represented. The fact that the dataset merely represented the observer's perception of reality from an outsider's point of view (classically termed as 'passive observation' within the field of ethnography, leading to an *etic*² representation of the phenomenon), meant that any conclusion being based on constructions of evidence from this dataset, needed to reflect these limitations. Missing awareness of these limitations in the analysis, means the interpreter of constructed evidence acts as subjective as the interpreter of observable behavior.

Nonetheless, the collaboration showed that piecing together a team of ethnographers, human factor researchers, and data scientists can indeed go beyond the limits of either one of the disciplines alone:

- Ethnographers can be useful to make sure that hypothesized representations of traffic interactions can ultimately represent reality in an appropriate and applicable level of detail without forcing the dataset-structure onto reality, but rather shaping the structure in a way that it can (at least partly) represent real-world complexities even in a standardized large-scale data collection.
- The human factor researchers, provided an important link to manage the necessary limits of controlling the complexity of real-world representations in a way that the dataset remains useful for the reductionist approach of them and data scientists when analyzing it.
- And the data scientists were useful to link these representations of reality to the technical domain of AI-based technology and analyze the data in a way that enables it to train models for active vehicle control of AVs.
- Ultimately, this collaboration showed that ethnographers, as the creators of dataset structures or whole datasets, should also be involved in the analysis process as a sparring partner to analyze and to reflect on what data actually means when drawing conclusions from constructed evidence. In the following we will argue why ethnography can yet offer an entirely different worldview to understand the reality of traffic interactions.

Representing the Outsider’s Perspective: Using Sequence Diagrams to Design Future Communication Strategies

In the quest of designing future communication strategies for AVs when interacting with pedestrians, interACT’s work package 4 team worked with so called ‘sequence diagrams’ - another form of evidence based on the human observation dataset. As can be seen below, it is an easy to read visualization of commonly observed sequences of interaction behavior exchanged between the driver and the pedestrian.

Intersection – pedestrian goes first:

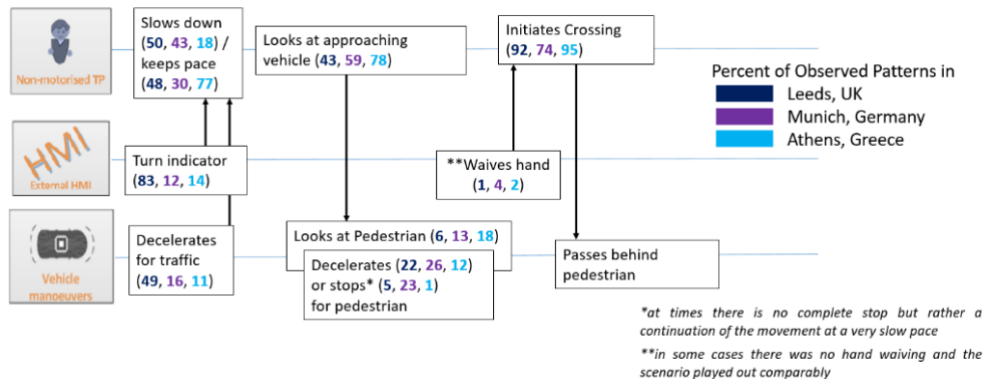


Figure 10. Sequence diagram. (source: Dietrich, 2018; Dietrich et al., 2018; Wilbrink et al., 2018)

These sequence diagrams were used in two ways: 1) to represent one single interaction, 2) to represent a whole dataset of interactions. The upper visualization for example shows a part of the total dataset. It quantitatively visualizes how often (in percentages) a particular

element of an interaction occurred in the given sequence throughout the three different field sites. The power of these sequence diagrams is that they show behavioral elements in combination to former and later occurring behaviors, and thereby produce a great basis to design for similar situations with AVs in the future. This means in turn that they lose their power once this combination is not clear anymore. Presenting a part of the dataset in percentages appears to be a competent foundation upon which to debate e.g. that we do not have to design e-HMI signals for drivers looking at pedestrians because this only occurred in 6% of all cases, or for hand gestures from the driver as this only occurred in 1% of these interactions.

The big question, though, is what does this form of evidence actually tell us, and who interprets it? There are at least two entirely contradicting ways these numbers could be interpreted: First, one researcher could argue since the percentages of drivers looking at pedestrians and given hand gestures are so low, we do not need to include these in the development of future communication strategies at all. The other researcher might argue that it is exactly the other way around; no matter how often pedestrians look for clarifying information from the driver today, if this is the potential of e-HMI, and if pedestrians would actually benefit from having this information all the time, then these are the situations we have to focus on. So even though it might be a minority in the quantitative representation of reality, it might be that very minority that we need to understand deeply. This is where the outsider's perspective reaches its limits to contribute to the design of future communication strategies. What becomes more important to understand then, is what value these signals actually do provide to the pedestrian, and therefore research needs to go beyond what is observable to include the pedestrian's and driver's insider perspective in detail.

Representing Reality from The Outsider's AND Insider's Perspective

To understand when and why e-HMI can be valuable to pedestrians, we, techno-anthropologists, conducted an additional, iterative ethnographic study with two goals: first, to investigate the decision-making process of pedestrians as broad and detailed as possible and second, to investigate why pedestrians sometimes look for more information from the driver and how this information then contributes to the decision-making process.

In contrast to the team of interACT, who selected relevant traffic scenarios based on "*a step-wise process of intensive discussions within the consortium*" (Wilbrink et al., 2017) to come up with relevant scenarios for all three European countries in the initial phase of the project, we followed bottom-up explorations grounded in the context and complexity of our field site. While running the pilot of the observation protocols for interACT, we realized that there was another highly interesting crossing scenario in the same intersection, in which explicit pedestrian-driver interactions occurred more regularly. This was jaywalking in dense and slow-speed traffic situations - a behavior that not only showed to be mundane for pedestrians in this intersection, but is in general not prohibited in the UK as we were informed by pedestrians. Hence, our assumption is that jaywalking might, indeed, be a more common traffic practice for British people, which would increase the relevance of this scenario for other intersections and cities in the UK as well. Thus, to investigate the pedestrian's decision-making process when jaywalking, we conducted semi-structured interviews with pedestrians directly after they were involved in an interaction. This provided us with an understanding of their insider's perspective of the interaction with drivers, which

we then combined with our outsider's perspective as observers. We structured our study in three phases: 1) a *Pilot Phase* for testing and refining our research approach, 2) an *Open Exploration Phase* to find as many influence factors on the pedestrian's decision-making process as possible, and 3) a *Design Problem Orientation Phase* to specifically focus on the relation between pedestrians and drivers as an important basis to identify potentials for valuable e-HMI.

Our approach resembled the process of a developer team using Scrum to iteratively run through 2-4 weeks coding sprints, while embracing a process of daily reflections.

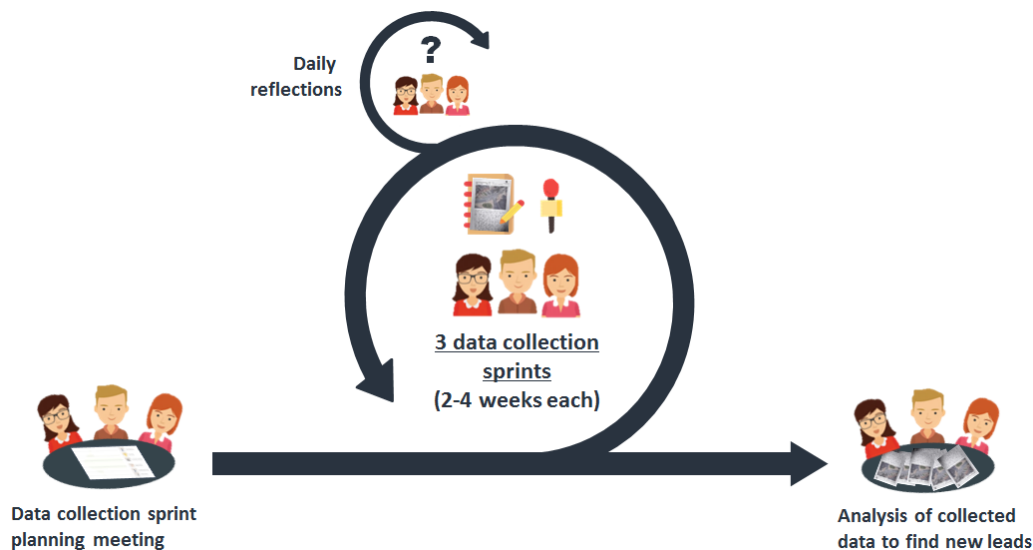


Figure 11. Scrum-like iterative ethnographic research approach.

In between each phase, we preliminarily analyzed the collected data to define the focus of the following phase. We applied Grounded Theory Method (e.g. Glaser & Strauss, 1967) to find emerging patterns in our interview data, which helped us iterate on our research questions and find new leads to explore. In addition to interview recordings, our datasets consisted of observational notes and rough sketches of the situation (see example below).

The combination of sketches and observational notes with the interview data provided the in- and outsider's perspective, and thereby a more holistic representation of the traffic interaction at hand. In comparison to the large-scale observational study of interACT, focusing on what and in which order behavioral elements occurred, our dataset consisted only of 34 interactions but added some depth to *why* and *how* pedestrians made decisions. Each data collection sprint resulted in theoretical models representing an emic² understanding of pedestrian's decision-making processes when interacting with drivers. The framework below, for example, shows influence factors that we found within the first data collection sprint, and how these interrelate with each other (Rasmussen, Rothmüller & Vendelbo-Larsen, 2017). Even though this is only a simplified representation of reality as well, it attempts to address more than just the momentary observable influence on pedestrian crossing decisions, and relates the influence of former experiences in different

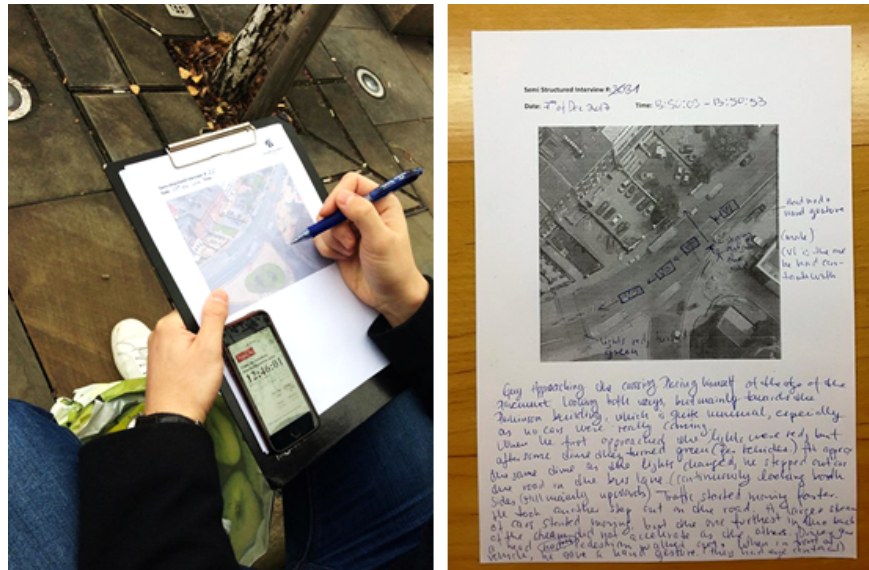


Figure 12. Observer jotting down notes and sketches. Picture 13. Example of a final sketch with observation notes. (source: Rasmussen, Rothmüller & Vendelbo-Larsen, 2017)

traffic cultures and the knowledge about the intersection to tendencies for specific crossing practices, interpretations and expectations when pedestrians assess a given traffic situation.

This example shows of one of the core elements in our process of constructing evidence. This framework built the basis for the second and third data collection sprint, in which we saturated the categories of the framework with then freshly collected interview data. The visualization below exemplifies how we used the model to understand the individual tendency for a certain crossing practice of one of our participants.

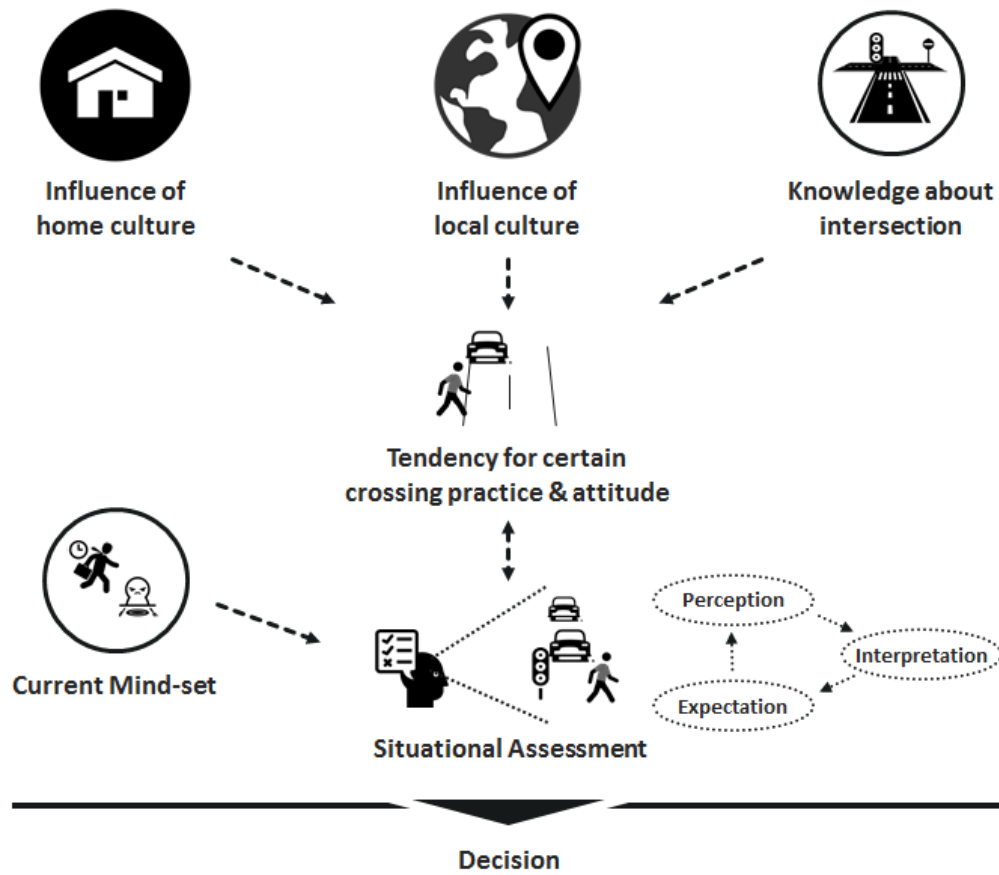
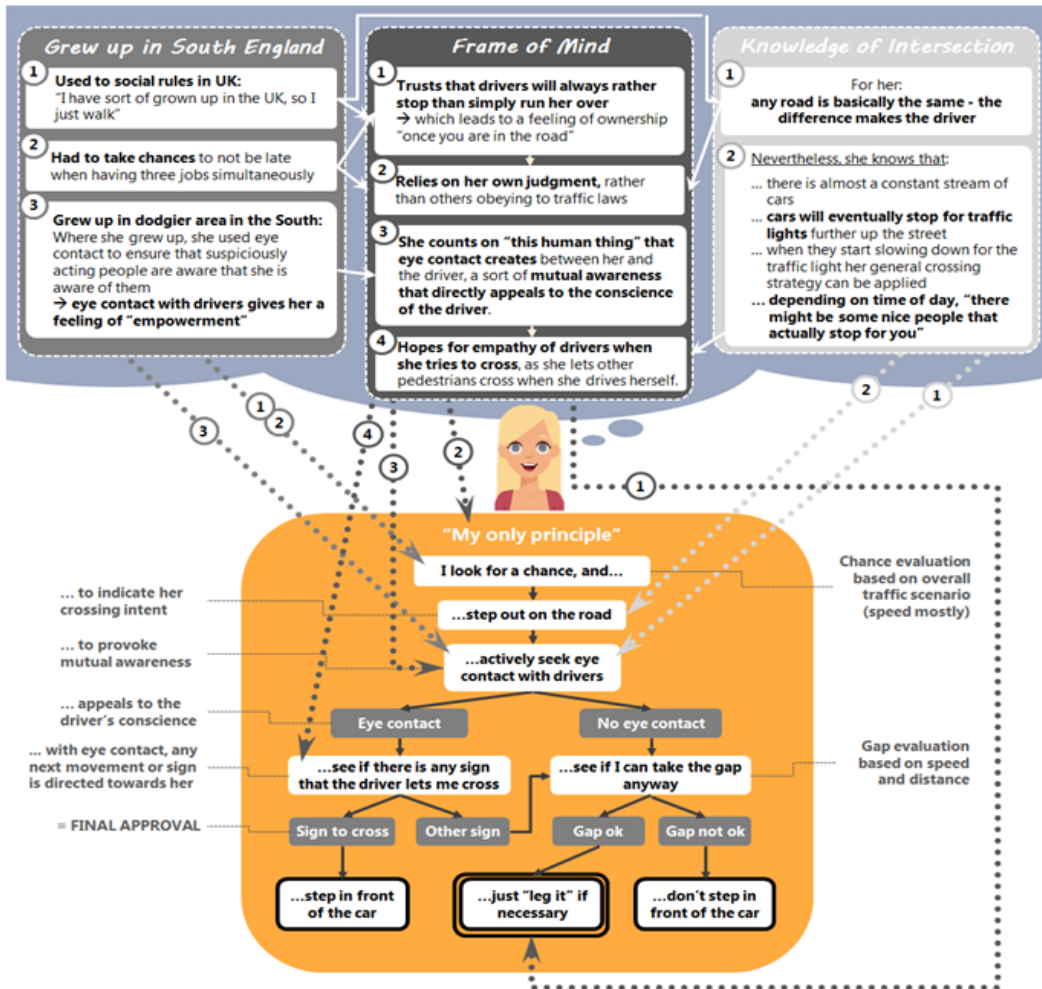


Figure 14. Theoretical framework of pedestrians' decision-making processes. (source: Rasmussen, Rothmüller & Vendelbo-Larsen, 2017)



“ When I cross the road, I sort of hope that someone is gonna make eye contact with me, so that they feel that I need to cross the road (...)
 Are you gonna get that from a driverless car?
 Is there gonna be the empathy there? ”

Figure 15. Example of applied theoretical framework of pedestrians' decision-making process. (source: Rasmussen, Rothmüller & Vendelbo-Larsen, 2017)

The next model shows one of the elements of the upper framework: the situational assessment model (ibid.).

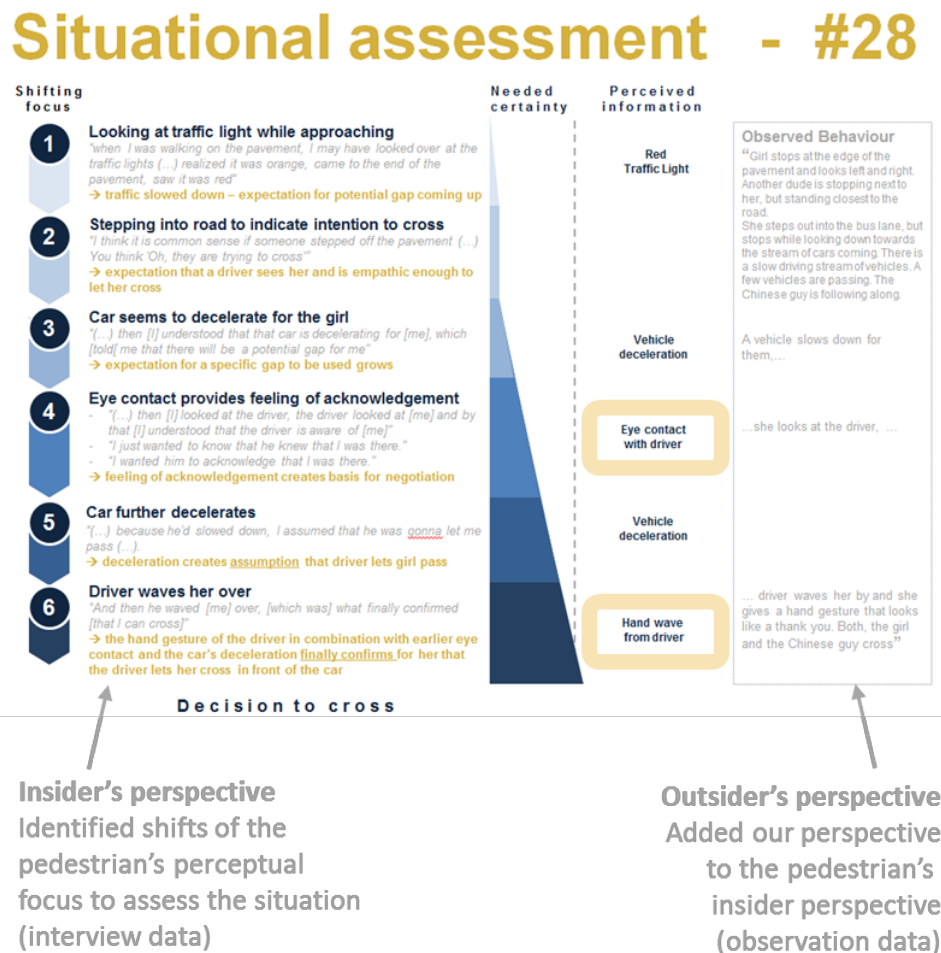


Figure 16. Example of a situational assessment model. (source: Rasmussen, Rothmüller & Vendelbo-Larsen, 2018)

This example of a situational assessment model shows six stages in this pedestrian's decision-making process from one of the interactions of the third data collection sprint. The six stages are essentially six focal points of the pedestrian, which we identified in the interview data (the pedestrian's insider perspective on reality). Each focal point evoked a certain expectation, and whenever an expectation got confirmed, the pedestrian's certainty to cross increased. It visualizes how we combined the insider's perspective to the for us observable outsider's perspective. The yellow highlighted elements 'Eye contact with driver' and 'Hand wave from driver' were the important interaction behaviors that e-HMI could potentially substitute in the future.

Identifying Potentials for Valuable e-HMI Concepts

As shown in the situational assessment model, expectations are central in the decision-making process of pedestrians (Rasmussen, Rothmüller & Vendelbo-Larsen, 2018 & 2017). For example, we investigated a case where a pedestrian nearly had a traffic accident, as he decided to cross in front of a car that he expected to stop. Not surprisingly perhaps, we found out that e-HMI could be of most value in situations where pedestrians are uncertain about what to expect, or need a confirmation of what they expect, to make the decision to cross. In relation to this, we found out that deceleration alone in dense, slow-speed and close-contact interaction scenarios might not always be interpreted as deceleration for the pedestrian. We noticed early on in the study that *awareness* seemed to be one of the central elements in decision-making processes (ibid.). Not only were pedestrians aware of all sorts of human and non-human actors in the environment, such as approaching cars, traffic lights, other pedestrians, cars driving faster or slower than the general speed of the stream, etc. but also were they really good in judging whether drivers were *aware of them* by e.g. seeing where the driver looked in the environment (ibid.). If they did not see the driver, this also provided a piece of information but did not help them much to cross the road. Traffic interaction is not only about adapting movements to avoid colliding with others, it is about the pedestrian's expectation of what drivers might do next (ibid.). It is about the relation between the pedestrian, the driver, the vehicle, and the real-world context. It is about empathy from both interaction participants to anticipate the other's intention, situated in the complexity and context of the given situation (ibid.).

After having focused on the influence of *awareness* in more detail, we learned that the pedestrian's certainty of drivers being aware of them helped them understand whether a driver was decelerating for them or something/someone else. This seemed to be applicable across cultures; due to the intersection's proximity to the University of Leeds, our study participants came from all sorts of different traffic cultures: Asian, African, American, small villages, cities, metropolis, megalopolis, all indicating this need.

Something very similar was found by interACT's eye-tracking study in Athens. Even though they ran an entirely different study, looking at the driver's insider perspective on interactions with pedestrians, they found similar elements in traffic to create the very basis for e-HMI to be valuable. In "*high density unsignalised urban crossing*" scenarios with a high level of "*proximity and low speeds*", the team concluded that body movements of the pedestrian alone are not always sufficient to clarify an interaction (Nathanael et al., 2018). As the visualization below shows, pedestrians looking towards the driver's vehicle, led to significantly more effective clarifications of interactions (ibid.).

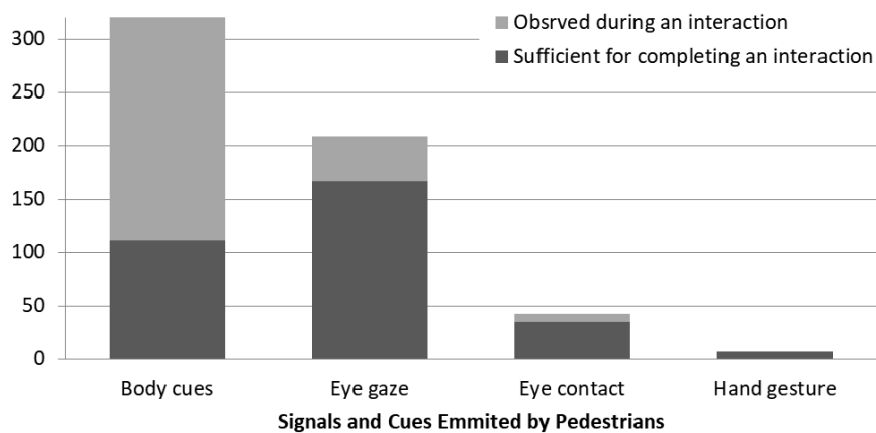


Figure 17. Results on the effectiveness of different cues to resolve an interaction. (source: Nathanael et al., 2018)

Another conclusion from both studies was that in situations where the uncertainty is so high that mutual awareness does not resolve the interaction alone, this awareness then builds the very foundation for further forms of explicit communication such as hand gestures, head nods, or further decelerations to communicate a final approval to the other interaction partner (ibid.; Rasmussen, Rothmüller and Vendelbo-Larsen, 2018). Thus, we identified the potential for valuable e-HMI to communicate a combination of awareness and a final approval, to effectively increase the pedestrian’s certainty in the decision-making process (Rasmussen, Rothmüller & Vendelbo-Larsen, 2018).

The relevance of this finding is not only related to today’s needs of pedestrians or drivers, but also to the problem of public acceptance of AVs. Throughout our ethnographic work on traffic interactions, we realized that there are two types of prejudices about AVs getting implemented in social spaces. First, some of our participants addressed their trust in scientific research to adequately test and evaluate any form of failure of AVs before they are released into urban traffic. They stated that they would trust AVs more than today’s drivers to handle complex traffic situations safely, as they would detect everything around them, and by that would never run someone over. For this group, we see a clear issue of overtrust in future technology. Already today, AVs are tested in urban spaces, which just recently resulted in a lethal situation (Crosley, 2018; Martyn, 2018). Our knowledge about the limitations of sensor technology from prior research confirms this issue (Sangari et al., 2016). At the moment, AV sensor technology simply does not detect everything in the environment (ibid.). For those over-trusting in AV-technology, we expect that e-HMI can effectively reduce risks in future traffic interactions by communicating awareness (Rasmussen, Rothmüller & Vendelbo-Larsen, 2018).

The other prejudice of people was their fear that AV-technology might not work flawlessly in the future, as the process of building this technology is eventually also only a process involving humans. And humans - just as human errors happen in traffic - happen to miss a thing or two in technological innovation (Norman, 2013). This group of people, therefore, would be afraid that the AV might not have detected them. e-HMI communicating awareness would therefore also serve this group of people to know whether

they are detected, and in combination with a final approval inform them whether they can safely cross the road (Rasmussen, Rothmüller & Vendelbo-Larsen, 2018).

In the last two subsections, we presented how we produced and used evidence to represent the pedestrian's insider perspective when interacting with drivers. Surprisingly, our results were very similar to interACT's eye-tracking study representing the driver's insider perspective when interacting with pedestrians, which we consider as an important lead that our findings - even though small in the sample and grounded in only one intersection - could apply to other sociocultural environments as well. Understanding what sort of context creates the basis for potentially valuable e-HMI is just as important as understanding which specific elements of interaction behavior from drivers might need to be substituted by e-HMI. The theoretical framework and models we created through our iterative ethnographic study, did then also help us to understand what additional information from the driver means to pedestrians, how they interpret it, why and in which contexts they need it as well as when this plays a role in their decision-making process.

Ethnography-based Design of e-HMI Concepts

In addition to the former described insights for when e-HMI could potentially be valuable, we learned a lot from being engaged in the design process of work package four at interACT. For example, that it is considered a challenge for e-HMI concepts whether the AV should communicate with one specific pedestrian only, or all the pedestrians around the car (Wilbrink et al., 2018). If the car communicates in a way that all pedestrians could interpret the signal as the allowance to go, this could pose a risk for those pedestrians not having been detected, or worse, for those pedestrians who were detected but actually endangered by other cars not having been considered in the AV's decision-making process.

We decided to focus on side-based communication, meaning that the AV only shows on which side it detected pedestrians, while neither communicating how many were detected, nor addressing only one out of a group for example (which we believe to be not only technologically infeasible for proper prototypes but already proved difficult to simulate in one of interACT's VR studies).

The following describes the first design concept of us techno-anthropologists³. *e-HMI Design 1 - Side-based Detection and Approval* communicates the detection of pedestrians on one or both sides of the vehicle through a pedestrian symbol lighting up on the side of the car. The choice of a pedestrian symbol was due to our intention to design for a cross-cultural application of our design. Pedestrian traffic lights are one of the already existing symbols to communicate an allowance to go for pedestrians and exist in nearly every traffic culture. The symbol, we believe, will be intuitively understood. For cyclists a cyclist-symbol would need to be added to our design. In combination with detection, our e-HMI communicates the final approval to cross via an LED bar lighting up in a motion that indicates the direction a pedestrian can cross in (inspired by a hand gesture. Thereby it shows pedestrians from one or both sides of the car that a cross can be safely performed. (Rasmussen, Rothmüller & Vendelbo-Larsen, 2018)

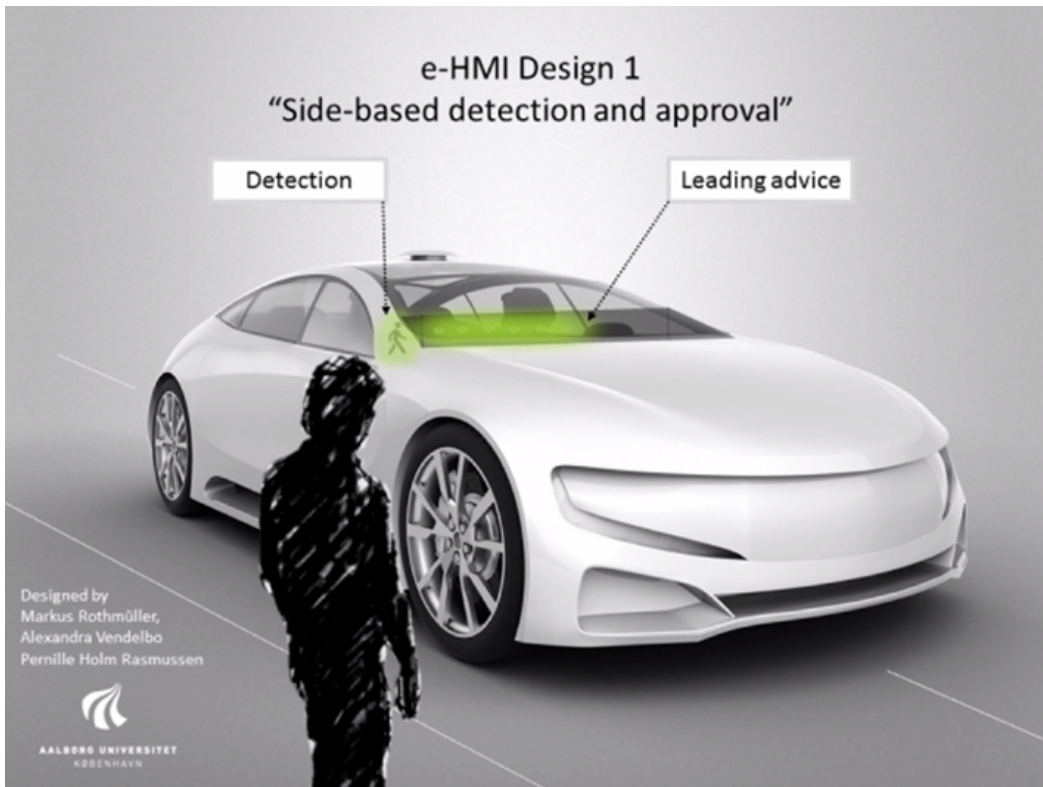


Figure 18. e-HMI Design #1 - ‘Side-based detection and approval to cross’. (source: Rasmussen, Rothmüller & Vendelbo-Larsen, 2018)

While the first holistic design was a very active form of communicating, the second holistic concept was designed in a way to accommodate a more passive form of communication. The basic thought with *e-HMI Design 2 - 360° Awareness and Location-based Tracking* was to empower the pedestrian to decide whether it is safe to cross without the AV communicating an actual advice to cross. In this design, the AV first communicates the readiness to cooperate with the environment through a 360° light band, and then shows the pedestrian(s) around where the car currently detects someone. Once a pedestrian was detected the light band tracks the pedestrian’s movement and communicates continuous awareness of the pedestrian’s location with a differently colored light stripe following the pedestrian. In the picture below, this tracking functionality was designed for the case of pedestrians crossing in front of the car. (Rasmussen, Rothmüller & Vendelbo-Larsen, 2018)

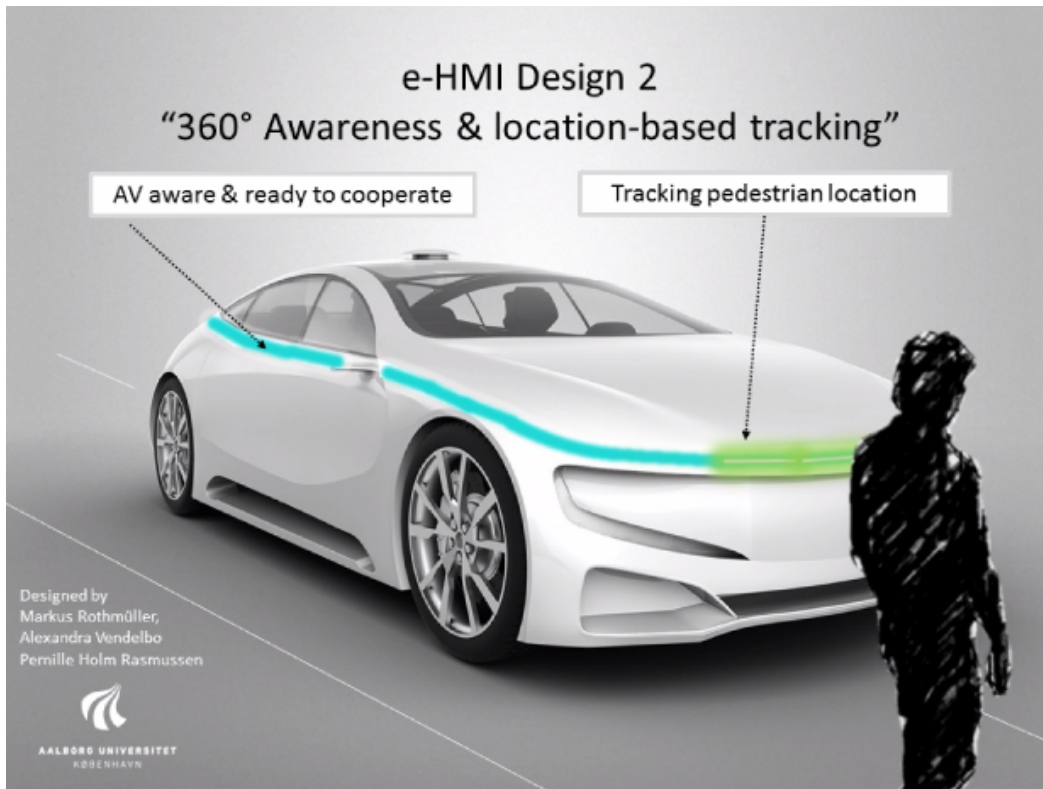


Figure 19. e-HMI Design #2 - '360° Awareness & location-based tracking'. (source: Rasmussen, Rothmüller & Vendelbo-Larsen, 2018)

Our fieldwork-based insights towards the value of e-HMI were, indeed, confirmed by a virtual reality (VR) study in which we evaluated the two e-HMI design concepts. The results of this study suggested that e-HMI communicating the detection of pedestrians and a final approval to cross in situations of high uncertainty for the pedestrian significantly increased the certainty and comfort to cross. In our master thesis *“we argue that e-HMI can show a positive impact on traffic safety, as well as trust and public acceptance of AVs when being integrated in urban traffic”* (Rasmussen, Rothmüller & Vendelbo-Larsen, 2018). A more detailed evaluation of our designs *“in terms of placement, color, motion and the combination of signals and vehicle behavior”*, showed *“the importance of understanding these factors when assessing the effect on decision-making times, the feeling of certainty and comfort to cross when interacting with AVs”* (ibid.). Nevertheless, we experienced yet another clash of disciplines when building the VR scenario to evaluate our e-HMI concepts, which in its core is rooted in different worldviews on how to deal with the representation of reality. This will be the focus of our reflection in this final chapter of our case study. (Rasmussen, Rothmüller & Vendelbo-Larsen, 2018)

Representing Real-World Contexts and Complexity in Virtual Reality & Mixed-Methods Evaluation of e-HMI Prototypes

When using VR to evaluate our e-HMI concepts, our goal was to represent specific elements of the real-world context and complexity which we identified as the very basis for why e-HMI might potentially be valuable (dense, slow-speed, close contact traffic situations with high levels of uncertainty for pedestrians).

As stated earlier in the problem space, VR serves as a perfect platform for experimental research, which in contrast to our goal, strives for controlled environments and simple scenarios to effectively verify hypotheses about cause and effect. Our colleagues at the German Aerospace Center, who thankfully were very patient and open to discuss our different worldviews, suggested to reduce the from us desired complex representation of real-world settings to a scenario with minimal influence factors. We, on the other hand, asked: What use is VR when it is unable to represent naturalistic traffic situations? How could we validate the effects of our e-HMI if we could not test it against the influence factors that set the framework for the need of e-HMI in the first place? Hence, we wanted the full package: multiple other pedestrians, cyclists, traffic streams from both sides, a traffic light making the streams slow down and accelerate again, mixed traffic with vehicles including drivers and AVs, different AVs with e-HMI and without, haptic feedback in the room: a sidewalk to step down from, and possibly the risk of getting hit by a pillow if a car virtually collides with the participants. Each of the previously mentioned elements describes a single influence factor which we would have to test our e-HMI design against; resulting in one thousand and fourteen experiment rounds (with a sample of one). How long would that take, we asked. Decades but this is how experimental psychology works, they answered. To overcome this challenge, we first had to understand the worldview of our collaborators, to then try to merge it with ours. A process of expressing ourselves in hypotheses, dependent and independent variables, and experimental measures helped us getting closer to a statistical controlled experimental design. Eventually, we managed to create a virtual world, which decently satisfied our need for a representation of naturalistic settings, while implementing our colleagues' suggestion for a controlled environment and comparison of e-HMI, without additional influence factors. Therefore, we focused on creating uncertainty for each of our study participants by running through an initial adjustment of gap sizes, setting the baseline for the study, which we hoped would enhance ecological validity (McNamara et al., 2015) - as long as we represented a dense, slow-speed, and close-contact traffic situation in which they had to cross (see picture below).

Our study was a hybrid form of experiment and ethnography-inspired user test. To investigate the outsider's perspective, we jotted down observational notes and measured experimental metrics in video recordings from, both, the participant's in-VR perspective, and a room-perspective recording the behavior of the participant outside of VR (see picture below).



Figure 20. Example of traffic set-up in VR with the goal to evoke a feeling of uncertainty to cross. (source: Rasmussen, Rothmüller & Vendelbo-Larsen, 2018)



Figure 21. Example of measurements in VR – Visibility time of e-HMI signal HD1.1. (source: Rasmussen, Rothmüller & Vendelbo-Larsen, 2018)

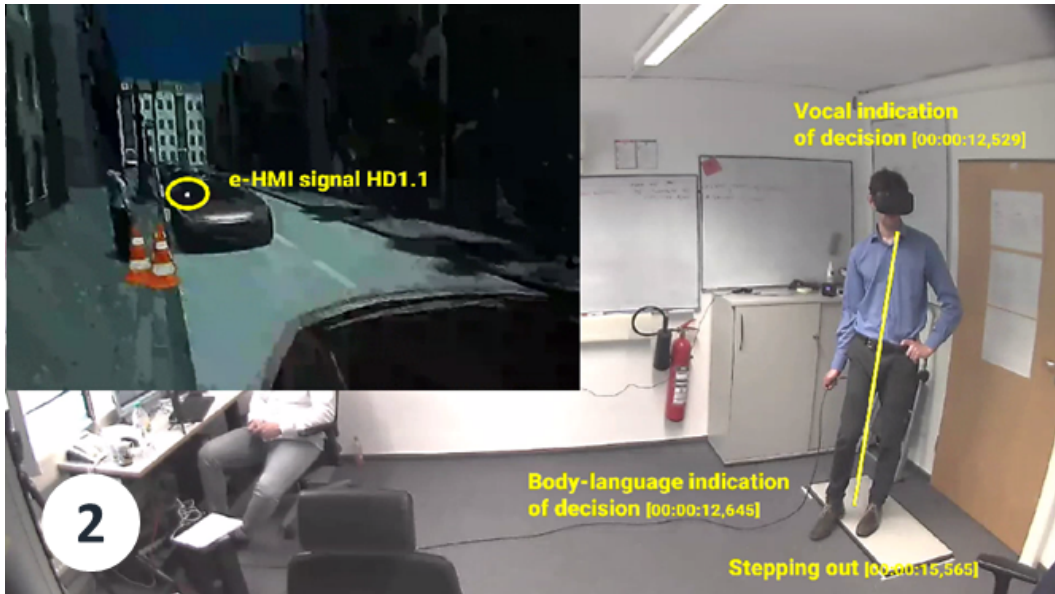


Figure 22. Example of measurements in VR – Indicators of crossing decision. (source: Rasmussen, Rothmüller & Vendelbo-Larsen, 2018)



Figure 23. Example of measurements in VR – Visibility time of e-HMI signal HD1.2. (source: Rasmussen, Rothmüller & Vendelbo-Larsen, 2018)

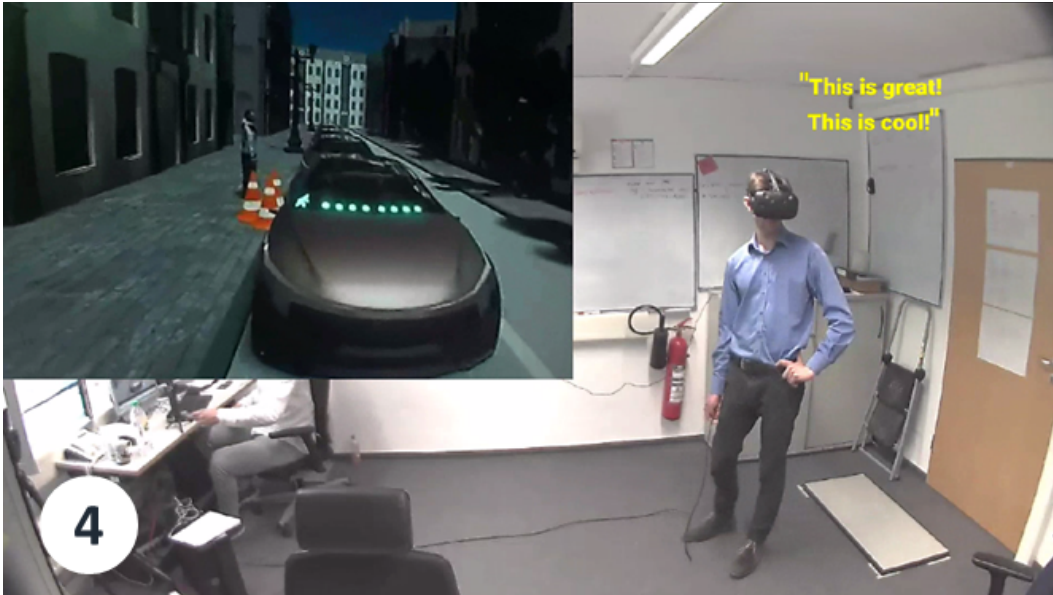


Figure 24. Example of qualitative evaluation in VR – Vocal reaction to e-HMI signal.
 (source: Rasmussen, Rothmüller & Vendelbo-Larsen, 2018)

To investigate the participant’s insider perspective, we combined questionnaire-based ratings of certainty and comfort to cross as well as of the meaning and intuitiveness of our designs with semi-structured interviews, in which we explored a variety of topics such as: the experience of being in a virtual world and its comparability to naturalistic traffic settings; the experience of uncertainty; the experience of receiving an e-HMI versus not receiving any; the meaning of our e-HMI designs; how to improve them; etc.

One of the essential insights from conducting the study in this specific way was that it was the hybrid of merging experimental research and ethnography that provided us with moments of clarity. Multiple times in our analysis we found that the questionnaire based ratings suggested a specific result which was contradicted by the quantitative measures and qualitative insights. Similarly, though, sometimes the ratings of certainty and comfort were really low in alignment with qualitative insights from verbal reactions and interviews, while the rating of intuitiveness and understanding the e-HMI was really high, which could only be described by utilizing the quantitative measures. Similar to the investigation of naturalistic interactions, it was again the combination of representing (virtual) reality from the outsider’s and insider’s perspective through a variety of methods, which allowed us to create different layers of evidence, finally providing us with a more holistic understanding when evaluating our e-HMI designs.

FINAL REFLECTIONS: CONTRIBUTIONS, IMPACT AND CHALLENGES

Our collaboration with experimental researchers at the German Aerospace Center did not only teach us about proper experimental research, but also about the limits of our own discipline. In fact the whole work with interACT taught us a lot about the limits of applying pure ethnographic work to inform technological development. Nevertheless, we also learned how to overcome some of the long discussed but nevertheless still existing challenges in interdisciplinary collaborations between the paradigms of qualitative and quantitative research. Retrospectively, our work with interACT really enabled 1) multiple forms of hybrid representations of today's traffic interactions to design e-HMI concepts for future interactions between pedestrians and AVs, as well as 2) hybrid representations of virtually simulated future interactions to evaluate our e-HMI designs. The collaboration proved to be a very promising attempt to tackle many of the challenges we introduced in the beginning of this case study, and showed that there is indeed space for ethnographers to navigate in technological development:

Contributions

1. Our ethnographic pilot study of the standard observation protocols helped interACT to develop a digital observation app which could represent naturalistic traffic interactions in a structured way; being closer to the complexity and context of reality while still providing a basis of control and comparability to make the dataset useful for the work of data scientists and human factor researchers.
2. This also showed that ethnographers as creators and shapers of datasets can be valuable sparring partners in the quantitative analysis process to reflect on what the collected data can actually tell when being transformed to multiple constructions of evidence.
3. Ethnographers as experts of immersion in different cultures and groups can provide the insider's perspective of interaction partners to represent the reality of interactions from multiple perspectives which need to be aligned and understood from a more holistic viewpoint rather than used individually when designing HMI for future interactions.
4. When using virtual reality simulations to evaluate interaction prototypes, we have presented a promising approach of ethnography to infuse some context and complexity from real-world environments, which was first identified to be the basis for why HMI might be valuable in the first place. This infusion in VR can then not only serve to verify the identified use-context for HMI to be valuable, but also lead to results that are ecologically valid for other sociocultural environments where this use-context exists equally.
5. Last but not least, combining mixed methods from ethnography and experimental research to represent the study participant's insider and the researcher's outsider

perspective on the investigated phenomenon in VR has proved to be just as valuable as the hybrid representation of perspectives in naturalistic interactions.

Impact

1. The impact of our work of infusing our worldview in the environment of interACT has yet to reveal itself entirely, as the analysis of merging the different perspectives is still an ongoing process. As one of our data science colleagues put it, maybe one day we will all find out why a subject-led ontology for the dataset is a good thing, but at first sight it seems to be of value. This very case study attempts to contribute to this understanding.
2. Nevertheless, our worldview has certainly led to critical reflections for some of the researchers at interACT which is the first step to enter discussions on:
 - a) how to align multiple perspectives on the reality of interactions to inform the development of AI-based technological systems enabling valuable means of future human-technology cooperation, and to enter
 - b) discussions on how to achieve ecological validity in the conduction of experimental research-based evaluations of human-machine interfaces to develop technological systems that work in the context and complexity of naturalistic environments.

Remaining Challenges

1. One of the remaining challenges is that ethnography, or anthropology in the broader sense, still lacks visibility and establishment as valuable part of any technological development process. More commonly, ethnography and anthropology should be considered as a must-have rather than a nice to add. To align multiple perspectives of reality when designing for future human-technology cooperation, our suggestion for future collaborations is to involve anthropologists and similar disciplines practicing ethnography from the very beginning in the study design phase to include their perspective in the discussion on creating datasets.
2. Another remaining challenge is to evaluate whether the ethnographic way of creating and shaping a large-scale dataset is actually better in terms of applicability for different sociocultural contexts. Since ethnography is really good in understanding the context and complexity of a certain field-site the question is to what extent using ethnography for the building process of datasets is actually better than directly relying on the ontology of former research and hypotheses.

Markus Rothmüller holds a BSc in International Business & Engineering as well as a MSc in Techno-Anthropology, and focuses on exploring and shaping future human-technology cooperation. His goal is to combine his engineering and anthropology background to innovate product development in collaboration with data science, AI and automation.

Pernille Holm Rasmussen holds a BSc in Technologies & Humanities as well as a MSc in Techno-Anthropology and is passionate about the study of human-machine cooperation. She aims at combining ethnographic methods with other disciplines such as experimental research and data science to find new ways to evaluate human-machine interfaces.

Alexandra Signe Vendelbo-Larsen holds a BSc in Market & Management Anthropology as well as a MSc in Techno-Anthropology. She focuses on a wide range of ethnographic research in both the field of human-technology interaction as well as consumer goods.

NOTES

Acknowledgements - We gratefully thank Natasha Merat and Anna Schieben for enabling and supporting our collaboration with the interACT project funded by the European Union's Horizon 2020 research & innovation program under grant agreement no 723395. Additionally, we would like to thank: Ruth Madigan for her great supervision and an amazing facilitation of our work during the data collection at the University of Leeds; André Dietrich for his great leadership of interACT's data collection and his trust in our skills to design and structure the digital observation app; Charles Fox for his introduction to data science in the realm of AV development; Janki Dodyia for her patience as sparring partner when designing and setting-up the virtual reality study; and Alexander Tesch for his great support in building the virtual world and our e-HMI prototypes in VR. We also want to thank Natasha, Anna, André, Ruth, Charles and Janki for their availability to retrospectively reflect on the collaboration in detail. Another special thanks belongs to Torben Elgaard Jensen, Melissa Cefkin, our families and everyone else who have always supported us in this long and intense journey. Thank you!

1. Next to few anthropologists, like Melissa Cefkin, being directly involved in the technological development of autonomous vehicles, anthropologists are underrepresented in this field of innovation. More generally anthropologists are still just in the uprise of being included in the technological development process, which can be concluded by the fact even latest publications have to point out the value of anthropology for technological innovation (e.g. Hartley, 2017; Madsbjerg, 2017; Roth-Lobo, 2015).

2. Emic/Etic: The emic or insider's perspective explains behavior "*in terms of the actors' self-understanding—terms that are often culturally and historically bound*" (Morris et al., 1999). It is the native's point of view. The etic or outsider's perspective explains behavior in correlation with external factors and is "*more likely to isolate particular components of culture and state hypotheses about their distinct antecedents and consequences*" (ibid.).

3. While having been working for interACT at the German Aerospace Center, we were allowed to build up our own virtual reality study and prototype our own e-HMI designs. These designs and prototypes informed the work of interACT but are not directly part of interACT, and have not been decided by interACT as such.

REFERENCES CITED

- Bhagavathula, R., Williams, B., Owens, J., & Gibbons, R.
2018 *The Reality of Virtual Reality: A Comparison of Pedestrian Behavior in Real and Virtual Environments*. Proceedings of the Human Factors and Ergonomics. Vol 62, Issue 1, pp. 2056 - 2060
- Blissing, B.
2016 *Driving in Virtual Reality: Investigations in Effects of Latency and Level of Virtuality*. Linköping Studies in Science and Technology Licentiate Thesis No. 1759
- Camara, F., Giles, O., Madigan, R., Rothmüller, M., Rasmussen, P.H., Vendelbo-Larsen, A., Markkula, G., Lee, Y.M., Garach, L., Merat, N., & Fox, C.W.
2018 *Predicting pedestrian road-crossing assertiveness for autonomous vehicle control*. White Rose Research Online Available at: <http://eprints.whiterose.ac.uk/135432/>

- Cefkin, M., & Stayton, E.
2017 *Speculating about Autonomous Futures: Is This Ethnographic?* Available at: <https://www.epicpeople.org/speculating-about-autonomous-futures/>
- Chang, C.-M., Toda, K., Sakamoto, D., & Igarashi, I.
2017 *Eyes on a car: An interface design for communication between an autonomous car and a pedestrian*. International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 2017, pp.65-73
- Cherry, C., Donlon, B., Yan, X., Moore, S., & Xiong, J.
2012 *Illegal mid-block pedestrian crossings in China: gap acceptance*. International Journal of Injury Control and Safety Promotion, vol. 19, no. 4, pp. 320 - 330.
- Clamann, M., Aubert, M., & Cummings, M.L.
2016 *Evaluation of vehicle-to-pedestrian communication displays for autonomous vehicles*. Traffic Research Board.
- Crosley Law Firm
2018 *New Insights From The Uber Self-Driving Crash Investigation*. Crosley Law Firm. Available at: <https://crosleylaw.com/blog/new-insights-uber-self-driving-crash-investigation/>
- Dey, D., & Terken, J.
2017 *Pedestrian interaction with vehicles: roles of explicit and implicit communication*. International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 2017, pp. 109–113.
- Dietrich, A.
2018 *Interaction in Urban Traffic—Insights into an Observation of Pedestrian-Vehicle Encounters*, Connected Automated Driving Symposium, https://www.interact-roadautomation.eu/wp-content/uploads/interACT_Webinar_Andre-Dietrich_180509_final-1.pdf
- Dietrich, A., Madigan, R., Portouli, E., Nathanael, D., Ruenz, J., Bengler, K., Merat, N., Giles, O., Lee, Y. M., Markkula, G., Camara, F., Fox, C., Rothmueller, M., Rasmussen, P. H., Vendelbo-Larsen, A., Schieben, A.
2018 *inter.ACT Work Package 2 – How Do Traffic Participants Interact in Current Urban Scenarios and How This Helps when Designing Automated Vehicles*, Poster session presented at the Transport Research Arena 2018, Vienna
- Drakoulis, R., Drainakis, G., Portouli, E., Althoff, M., Magdici, S., Tango, F., & Markowski, R.
2018 *inter.ACT Deliverable 3.1 - Cooperation and Communication Planning Unit Concept*. https://www.interactroadautomation.eu/wp-content/uploads/interACT-WP3_D3.1_CCPU_Concept_v1.1_DRAFTwebsite.pdf
- Glaser, G. B., & Strauss, A.L.
1967 *The Discovery of Grounded Theory - Strategies For Qualitative Research*. (Renewed, 1995, Reprinted 2006) Aldine Transaction, A Division of Transaction Publishers, Rutgers-The State University, 35 Berrue Circle, Piscataway, New Jersey
- Hartley, S.
2017 *The Fuzzy and the Techie: Why the Liberal Arts Will Rule the Digital World*. Houghton Mifflin Harcourt
- Kadali, B., & Perumal, V.
2012 *Pedestrians' gap acceptance behavior at mid-block*. International Journal of Engineering and Technology, vol. 4, no. 2, pp. 158 - 161
- Langström, T., & Lundgren, V.M.
2015 *AVIP – Autonomous Vehicles Interaction with Pedestrians: An Investigation of Pedestrian-Driver Communication and Development of a Vehicle External Interface*. Masters Thesis, Chalmers University of Technology, Gothenborg, Sweden.

- Madsbjerg, C.
2017 *Sensemaking: The Power of Humanities in the Age of the Algorithm*. Hachette Books
- Mahadevan, K., Somanath, S., & Sharlin, E.
2017 *Communicating awareness and intent in autonomous vehicle-pedestrian interaction*. University of Calgary, Tech. Rep., 2017
- Martyn, A.
2018 *Toyota executive says self-driving car technology is overhyped*. Article at Consumer Affairs, 21st of September, 2018 [Accessed 15.10.2018] Available at: <https://www.consumeraffairs.com/news/toyota-executive-says-self-driving-car-technology-is-overhyped-092118.html>
- McNamara, L., Cole, K., Haass, M.J., Matzen, L.E., Morrow, J.D., Steven-Adams, S.M. & McMichael, S.
2015 *Ethnographic Methods for Experimental Design: Case Studies in Visual Search*. In: Schmorow D., Fidopiastis C. (eds) Foundations of Augmented Cognition. AC 2015. Lecture Notes in Computer Science, vol 9183. Springer, Cham
- Merat, N., & Madigan, R.
2017 *Human Factors , User Requirements , and User Acceptance of RideSharing In Automated Vehicles*. Discussion Paper No. 2017-10 in International Transport Forum
- Merat, N., & Madigan, R., Louw, T., Dziennus, M., & Schieben, A.
2016 *What do Vulnerable Road Users think about ARTS?* in CityMobil2 final conference, Donostia / San Sebastián, Spain.
- Morris, M.W., Leung, K., Ames, D. & Lickel Brian
1999 *Views from inside and outside: Integrating emic and etic insights about culture and justice judgment*. Academy of Management Review 1999, Vol. 24. No. 1781-796. Available at: http://www.columbia.edu/~da358/publications/etic_emic.pdf
- Nathanael, D., Portouli, E., Papakostopoulos, V., Gkikas, K., & Amditis, A.
2018 *Naturalistic Observations of Interactions Between Car Drivers and Pedestrians in High Density Urban Settings*, Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018). IEA 2018. Advances in Intelligent Systems and Computing, vol 823. Springer, Cham
- Norman, D.
2013 *The Design of Everyday Things*. Revised and Expanded Edition, Basic Books, a Member of the Perseus Books Group New York.
- Parkin, J., Clark, B., Clayton, W., Ricci, M. and Parkhurst, G.
2016 *Understanding interactions between autonomous vehicles and other road users: A Literature Review*. Project Report. University of the West of England, Bristol. Available from: <http://eprints.uwe.ac.uk/29153>
- Rasmussen, H. P., Rothmüller, M., & Vendelbo-Larsen, A.
2017 *Constructions of Individual Crossing Practices - A Novel Approach To Understanding Pedestrians' Decision Making Processes in Crossing Situations to Inform the Design of e-HMI Solutions in Autonomous Vehicles*. School of Architecture, Design and Planning, Aalborg University Copenhagen
- Rasmussen, H. P., Rothmüller, M., & Vendelbo-Larsen, A.
2018 *Designing Future Interactions between Humans and Autonomous Vehicles - From Techno-Anthropological Field Observations to Prototypes of External Human-Machine Interfaces Tested in Virtual Reality*. Master's thesis, MSc Techno-Anthropology, School of Architecture, Design and Planning, Aalborg University Copenhagen, Denmark (confidential)

- Rasouli, A. & Tsotsos, K. J.
 2018 *Autonomous Vehicles that Interact with Pedestrians: A Survey of Theory and Practice*. IEEE Transactions on Intelligent Transportation Systems
- Roth-Lobo, H.S.
 2015 *What Anthropology Brings to Innovation: John Sherry / A Profile*. Ethnographic Practice in Industry Conference. Available at: <https://www.epicpeople.org/john-sherry-what-anthropology-brings-to-innovation/>
- Sale, J.E.M., Lohfeld, L.H., & Brazil, K.R.
 2002 *Revisiting the Quantitative-Qualitative Debate: Implications for Mixed-Methods Research*. Quality & quantity 36.1 (2002): 43–53. PMC. Web. 6 Oct. 2018.
- Sangari, A.Z., Taniberg, A., Retoft M.F., Rothmüller, M., Rasmussen P.H., Vendelbo-Larsen, S.A.,
 2016 *A Socio-Technical Perspective on Autonomous Vehicles*. School of Architecture, Design and Planning, Aalborg University Copenhagen
- Schieben, A.
 2018 *Management Summary, International eHMI workshop 19th of April 2018, Vienna*, https://www.interactroadautomation.eu/wpcontent/uploads/interACT_ManagementSummary_eHMIWorkshop.pdf
- Schieben, A., Wilbrink, M., Kettwich, C., Madigan, R., Louw, T., Merat, N.
 2018 *Designing the interaction of automated vehicles with other traffic participants: design considerations based on human needs and expectations*, Cogn Tech Work, Springer Online, <https://doi.org/10.1007/s10111-018-0521-z>
- Sobhani, A., & Farooq, B.,
 2018 *Impact of Smartphone Distraction on Pedestrians' Crossing Behaviour: An Application of Head-Mounted Immersive Virtual Reality*. Journal published: Transportation Research Part F: Traffic Psychology and Behaviour, Acceptance date: June 17, 2018 <https://arxiv.org/ftp/arxiv/papers/1806/1806.06454.pdf>
- Stanton N. A., Salmon, P. M. Rafferty L. A., Walker G. H., Baber C., & Jenkins D. P.,
 2013 *Human Factors Methods: A Practical Guide for Engineering and Design*. Human Factors Methods. London: CRC Press. <https://www.taylorfrancis.com/books/9781317120162>
- Stayton, E., Cefkin, M., & Zhang, J.
 2017 *Autonomous Individuals in Autonomous Vehicles: The Multiple Autonomies of Self-Driving Cars*. Ethnographic Praxis in Industry Conference Proceedings, ISSN 1559-8918, [epicpeople.org/intelligences](http://www.epicpeople.org/intelligences)
- Suchman, L.
 2007 *Human-Machine Reconfigurations - Plans and Situated Actions*. [2nd edition] Cambridge University Press
- Turkle, S.
 2009 *Simulation and Its Discontents*. Cambridge, MA: MIT Press.
- Verbeek, P.P.
 2015 *Beyond Interaction: a short introduction to mediation theory*. Available at: <http://interactions.acm.org/archive/view/may-june-2015/beyond-interaction> [Accessed 5 December 2016]
- Vinkhuyzen, E. & Cefkin, M.
 2016 *Developing Socially Acceptable Autonomous Vehicles*, Ethnographic Praxis in Industry Conference Proceedings, p. 522–534, ISSN 1559-8918, <https://www.epicpeople.org>

- Wilbrink, M., Scheiben, A., Markowski, R., Weber, F., Gehb T., Ruenz, J., Tango, F., Kaup, M., Willrodt, J.H., Portouli, V., Merat, N., Madigan, R., Markkula, G., Romano, R., Fox, C., Althoff, M., Söntges, S., Dietrich, A.
2017 *interACT D1.1 Definition of interACT use cases and scenarios*. Available at: <https://elib.dlr.de/116445/>
- Wilbrink, M., Schieben, A., Kaup, M., Willrodt H.J., Weber, F., & Lee M.Y.
2018 *interACT deliverable 4.1 - Preliminary Human Vehicle Interaction Strategies V1*. Available at:
https://www.interact-roadautomation.eu/wpcontent/uploads/interACT_WP4_D4.1_Preliminary_Human_Vehicle_Interaction_Strategies_v1.0_draftWebsite.pdf
- Wilson, C.J., & Soranzo, A.
2015 *The Use of Virtual Reality in Psychology: A Case Study in Visual Perception*. Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine Volume 2015
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4538594/pdf/CMMM2015-151702.pdf>
- Yannis, G., Papadimitriou, E., & Theofilatos, A.
2013 *Pedestrian gap acceptance for midblock street crossing*. Transportation Planning and Technology, vol. 36, no. 5, pp. 450 - 462.