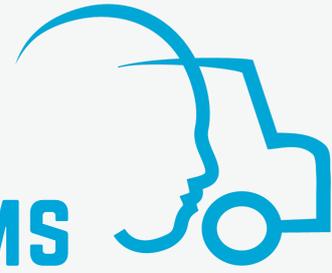


MEANINGFUL HUMAN CONTROL OVER AUTOMATED DRIVING SYSTEMS



MAY 2020

THIS ISSUE'S TOPICS

THE EDITORIAL

An Interview with Jana Price on human factors and MHC
by *Marjan Hagenzieker*

THE JOURNEY

A Trip to Japan
by *Giulio Mecacci*

A WORD ON COVID AND ETHICS

by *Filippo Santoni de Sio*

ACADEMIC LIFE AND COVID-19

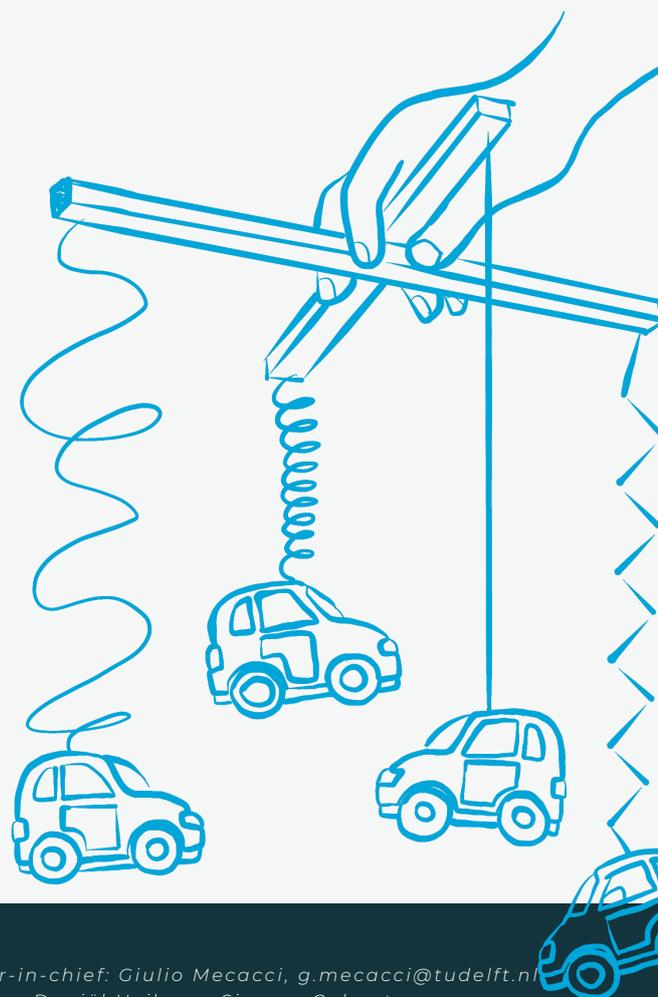
by *Bart van Arem & Daniël Heikoop*

UPDATES FROM OUR TEAM

Our updates on research and things we do in our daily life of hard (?) working researchers

PUBLICATIONS AND DISSEMINATION

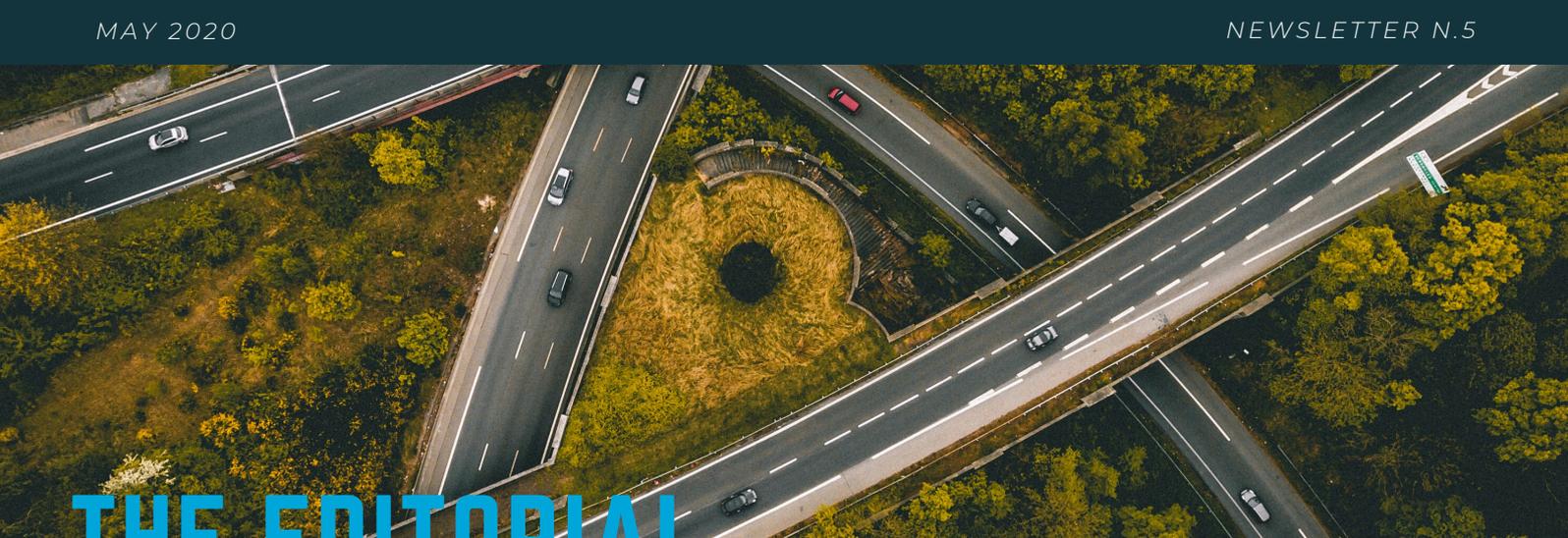
A short list of our most recent publications, and some of our most noteworthy activities



COLOPHON

Meaningful Human Control newsletter
TU Delft

Editor-in-chief: Giulio Mecacci, g.mecacci@tudelft.nl
Editors: Daniël Heikoop, Simeon Calvert



THE EDITORIAL



AN INTERVIEW WITH JANA PRICE ON HUMAN FACTORS AND MEANINGFUL HUMAN CONTROL

by Marjan Hagenzieker

Welcome to our fifth newsletter. You will read about Japan and their renewed interest in ethics of robotics and how COVID-19 is impacting ethics in our research.

In this editorial, I'd like to start with looking back to two accidents that happened in the USA and involved partially automated vehicles. In the past year, the National Transportation Safety Board (NTSB) has published two major reports about these accidents, also known as the "Mountain View" and "Tempe" accidents. I asked Jana Price to reflect on these two accidents. I am particularly curious about her view since she has a background in human factors. Jana has worked at the National Transportation Safety Board for 19 years as a human performance investigator and researcher and she has worked on investigations in every mode of transportation.

While these two accidents are very different, they also appear to share common underlying causes. Is it possible to indicate which were the main factors causing them?

"It is true that in certain ways the two crashes—which occurred within a week of each other in March 2018—were quite different, but they also had similarities. The March 23, 2018, crash in Mountain View, California, involved a Tesla Model X P100D SUV that was being driven by the vehicle's owner from his home to his work. The Tesla had a Level 2 advanced driver assistance system (ADAS) known as "Autopilot", which was active at the time of the crash. About 9:27 a.m., the vehicle approached a paved gore area dividing the main travel lanes of the highway from a left exit ramp. The vehicle moved to the left and continued into the gore area where it struck a previously damaged and nonoperational crash attenuator. The Tesla was involved in subsequent collisions with two additional vehicles and a post-crash fire occurred involving the vehicle's high voltage battery. The Tesla driver died as the result of blunt-force trauma injuries.

Editor's note: In our paper entitled "Gaps in control of automated vehicles on roads", we investigated these and other incidents and how they could be retraced to gaps in control. You can read it [here](#) if you are interested.



Jana Price is a Senior Human Performance Investigator at the USA National Transportation Safety Board

The USA National Transportation Safety Board (NTSB) has investigated accidents involving automated systems in all modes of transportation. Their first investigation of a fatal highway crash involving an automated system was the 2016 crash in Williston, Florida. Since that time, NTSB has launched on five additional highway crashes involving different levels of automation.

The March 18, 2018, crash in Tempe, Arizona, involved a Volvo XC90 SUV that had been modified with a developmental Level 4 automated driving system (ADS) by Uber ATG. The vehicle was operating on an established test loop with a company employee, whose job was to monitor the driving environment and the performance of the ADS, occupying the driver's seat. At about 9:58 p.m., a pedestrian pushing a bicycle by her side crossed the road in front of the test vehicle at a location not marked by a crosswalk. Although the ADS detected the pedestrian, it never accurately classified her as a pedestrian or predicted her path. Onboard video showed that the employee in the driver's seat was looking at the bottom of the SUV center console, where her phone was streaming a video, as the vehicle approached the pedestrian. Although she redirected her gaze to the roadway about 1 second before impact and began steering left, the operator was unable to avoid striking the pedestrian who died as a result of the crash.



These crashes involved vehicles with different levels of automation and different scenarios, yet there were some striking similarities in the underlying factors that contributed to the two crashes:

- In each case, the vehicle driver/operator was not monitoring the driving environment at the time of the crash but was visually distracted by using a cell phone;
- Both cases involved over-reliance on vehicle technology;
- Both cases involved a "novel" situation encountered by the system;
- In each case, NTSB found that the risk assessments conducted by system designers did not adequately address risks in the operating environment;
- In each case, the NTSB discussed limitations of federal and/or state oversight of the systems."

In both accidents' investigation reports it is concluded that human factors contributed to the occurrence of the accidents. Were these 'human factors' of a similar kind in both accidents? And couldn't certain human errors be simply prevented by improving the technology first?

"It is tempting to think that designers will someday be able to improve vehicle automation technology to a degree that will completely eliminate crashes. However, we know from history that although some types of human errors may be prevented, new types of system errors may be introduced by the presence of the technology. Additionally, as we have learned from our investigations, even providing vehicle operators with information about system limitations may not be adequate to ensure that they will continuously monitor the driving environment. The over-reliance on technology in these cases was similar; however, the countermeasures to address issues concerning operator engagement may vary depending on the level of automation and whether vehicles are commercially available or being used in a test environment.

With respect to Level 2 systems that are currently commercially available, the NTSB has recommended verifying that manufacturers incorporate safeguards that limit the use of ADASs to those conditions for which they are designed. The NTSB has also called for the development of standards for driver monitoring systems that will minimize driver disengagement, prevent automation complacency, and account for foreseeable misuse of the automation.

"even providing vehicle operators with information about system limitations may not be adequate to ensure that they will continuously monitor the driving environment"

With respect to higher level automated systems that are currently in development, NTSB has recommended requiring developers to submit safety self-assessments that would be evaluated to determine if they include appropriate safeguards to manage the risk associated with crashes and operator inattentiveness."

What would you consider to be important to take into account when it comes to 'meaningful human control' in relation to safe automation in traffic?

"Costa Samaras has said, 'Robots make excellent backup drivers to humans. Humans make terrible backup drivers to robots'. In my opinion, this quote does a nice job of explaining the challenges in designing a system that can perform safely in all traffic situations. Automated systems have great promise and can likely outperform human drivers in certain respects. However, it is critical, particularly in the context of Level 2 vehicles that are being driven on public roads today, that humans maintain a meaningful level of engagement and control even in the presence of automation. And, as we move toward increased levels of automation, it is important that developers design safeguards into the testing environment, including the management of risk associated with operator inattentiveness."

All in all, it is clear that the human factor will remain crucial in automated driving systems in many ways and for a long time to come. And while our project is coming to an end, I'm already looking forward to future research on Meaningful Human Control in Automated Driving.

Marjan Hagenzieker

THE JOURNEY

A TRIP TO OSAKA

by *Giulio Mecacci*

Last year we were lucky enough to "intercept" the Japanese guru of robotics Minoru Asada during one of his trips to The Netherlands. He's the one who founded, among other things, a very nice annual sport event for robots. You might have seen it around, it's called Robocup (picture to the right). Professor Asada and the robolawyer Tatsuhiro Inatani were around to scout and meet ethicists of technology and possibly establish a network for their novel project on robots & society. They are exploring the societal impact of humanoid robots and the future possibilities of granting them legal personhood. We had a great day in Delft, starting with a morning of short presentations by some of our colleagues, followed by a walking afternoon, when we showed our Japanese colleagues some of our technological marvels made in TU Delft. Beer and Dutch cuisine concluded the visit.

Not long after that day, I and Filippo received an invitation to Osaka, Japan, to participate to the "2nd International Symposium on Symbiotic Intelligent Systems - Challenge to the Edge: How can we design symbiotic society?". We had to bring ethics of technology to the table. And what a big table. The event was hosted by Asada and Hiroshi Ishiguro. I'm sure you've seen him too, like... everywhere in and out the internet. (picture below).



Right before the pandemic exploded, but early enough to make it safely home, we jumped on a plane and flew to Japan. We knew only one thing: Osaka is deemed by very many the world capital of food. Challenge accepted, picky Italians on their way to judge. Leaving the airport we found ourselves in front of a metropolitan panorama that I hardly ever experienced before. The Keihanshin metropolitan area, which includes Osaka, Kyoto and Kobe, hosts almost 20 millions citizens, and its GDP touched a few years ago one trillion dollars.

There was something strange but fascinating in the way people would live their lives. We were immediately attracted and surprised by a few things. One of those was the amount and sheer extension of the indoor areas. We felt like entire cities would fit, and did fit, the Osaka underground. An underground where you could find literally kilometres of dark and cozy alleyways with an infinite succession of restaurants, sometimes with carefully reproduced traditional Japanese facades. Malls that were as big as Delft were fit inside skyscrapers and deep into the ground. Needless to say, we love Blade Runner enough to appreciate the references. And of course, food was amazing, both in quality and quantity. You can tell by the sheer amount of restaurants, sometimes in the order of many tens within a single space, that there is love and dedication, almost obsession for good food. It's probably worth flying there just to eat!





THE JOURNEY

The symposium was interesting. Very many great scholars were there, and it was a highly interdisciplinary event that hosted talks about practically everything related to robotics and intelligent systems, from engineering to cognitive science to ethics of technology. We were happy to see people from all over the world reunited to discuss those topics, and I and Filippo brought Meaningful Human Control on the table since that is one important factor to consider when discussing the relation between society and automation. The discussion quickly stemmed more events and collaborations, and two workshops were organized in Europe. One of them was in Vienna, and was unfortunately cancelled due to the outbreak, but the next one, in Paris, ICRA2020, has just been converted into an online event. I'll be streaming from my room in Delft!

After the successful event, we took a few days off to visit Osaka and Kyoto and, of course, eat more. You can find some pictures I took all across this page.



Sushi, fancy



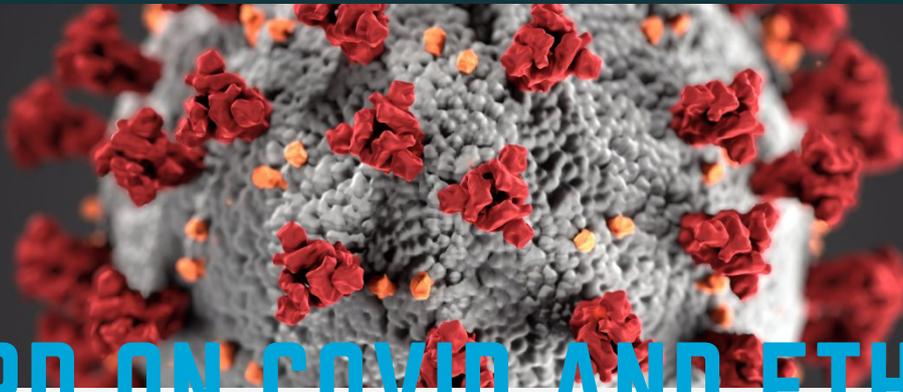
Osaka castle



Kyoto, temples



Kyoto, zen gardens



A WORD ON COVID AND ETHICS



By *Filippo Santoni de Sio*

"The Covid crisis has not only raised new ethical issues, but also made some important general features of ethics dramatically tangible. A recent beautiful article by [Yuval Harari](#) nicely summarises some of these features. First, ethics matters: the ethical choices we will make today (in a time of crisis) will affect the future of society for a longer term, also when society will be back to (a new) normality. I for myself, after the initial shock and struggle, have learnt the benefits of online teaching and home working. Still, how will this affect ourselves, the way we relate to each other, the way we work? Will I be willing to go back to spend time and energies in classrooms and offices, engaging in personal and unplanned conversations with students and colleagues, when the corona restrictions will be lifted? Will there be a way to reap the benefits of home working without compromising the quality of human and social relationships?

Second, and relatedly, ethics often is presented as a matter of moral dilemmas. Think of the debate on the COVID tracing apps: should we protect health OR privacy? In his article, implicitly and possibly unconsciously echoing the motto of TU Delft's Design for Values approach

- by the way, we have created a [working group on the Ethics of COVID crisis](#) - Harari points out that this is a false dilemma: we want both health AND privacy. If mainstream concepts and technology do not allow for this, well too bad for mainstream concepts and technology: let's go for better ones. We have known this well in our MHC-ADS team. Ethics of self-driving cars, we have been repeating since the first lines of our first grant proposal drafts, is not about whom the fully autonomous vehicle of the future should kill or spare, but rather about how to get to a future of more or less automated vehicles while avoiding all avoidable accidents. Also, it is not about choosing between safety and human control, or automation and human control, but rather about how to realise systems that are both safe AND responsive to all (relevant) human reasons (though not necessarily to drivers' actions), highly automated AND human-centred. As Harari effectively writes: this storm will pass. But the choices we make now could change our lives for the years to come. Including the choices about the transition to driving automation."

ACADEMIC LIFE AND COVID-19



By *Bart van Arem & Daniël Heikoop*

"As the spread of Covid 19 virus has been affecting life and work all around the world, I am happy to inform that our research team of Meaningful Human Control over Automated Driving Systems is in good health. The measures taken to limit the spread of the virus are affecting our work and our results. We have been working from home since March. Those of us that are also teaching, were actually launched into the era of on-line teaching. We are making an additional effort in finding new ways of teaching and interacting with students. Regrettably, we had to postpone the series of driving simulator experiments, because of the closing of the facilities. In the near future we hope to find way to finish the preparations and conduct the experiment while safeguarding the health of the participants and the researchers."

Bart

"Instead of running the original simulator experiments, we have been allowed to, as a very exceptional case, perform an N=1-simulator experiment. This means that the two students who were working on this simulator experiment will now run their own experiment by themselves. This way, they will avoid any risk of spreading or contracting the virus, while still being able to do something with all the work they have done to set this experiment up. They will now validate (and improve) their own experiment by manipulating a predetermined set of variables, and run every possible variation of these variables themselves. They both have 5 days each to complete their experiments, and will do so intermittently between each other (one day the one, another day the other). With the results they both gain, we can improve the current setup of the simulator experiment, for future use of the original plan with 100+ participants."

Daniël



UPDATES FROM OUR TEAM



Simeon Calvert, traffic engineer

As we approach the final months of the project, much of the research focus is targeted at concluding on-going activities. For the engineering part of the project, this entails three main efforts. The first is the conclusion of the urban control research that uniquely and effectively highlights a tangible way to operationalise MHC as a concept in transportation for evaluation and design. The ability to be able to develop this approach emphasises the importance of the project to take a philosophical concept like MHC and use it effectively for this purpose. We urge the partners and others alike to take note of this effort and apply it moving forward. A first version of the paper has now been submitted for publication and we are awaiting the first review. The second effort involves a focus on what MHC means for vehicles cooperation and connectivity. These are important trends and developments in transportation and mobility in which MHC is highly applicable. We are addressing the topic by using the case of Truck Platooning as a case example. I hope that during the month of May, we will have submitted this paper for publication. The final effort considers the more philosophical aspects of applying MHC in an urban traffic environment. Although the main approach it taken from the philosophical track in the project, there are key engineering principles that must be carefully considered and applied. Again, this is a research direction that we hope will lead to a submitted paper by the end of May. Looking further ahead, this project has addressed many different areas of automated driving and is really only just opened a jar of research potential and opportunities. Therefore, as we move towards the conclusion of this project, this may be the time to consider how MHC can be carried further forward and find implementation in practice. This is a forthcoming task that is not limited to the core research team, but to all the partners. We hope to hear from you in this regard!



Daniël Heikoop, behavioral scientist

The driving simulator experiment has experienced yet another blow with the spread of the coronavirus blocking all physical activities, thus also experiments with participants. This has been a major let-down for this project, as everything was in place to start the experiment right at the moment where everything suddenly grinded to a halt. Over 100 participants were recruited through yet another round of advertisement, this time through different media targeting different people to get an even richer participant pool. Scenarios were designed, protocols were in place, everything was finished to be piloted that day; until... well... everyone knows by now. Nevertheless, other things happened too: five papers/abstracts were submitted, of which 4 (1 paper, 3 abstracts) got accepted. All 5 of them are accepted for presentation at their respective conferences: the ICTTP and FISTS. Unfortunately, due to the precautionary measures, both conferences may be postponed, possibly until next year. Also the on-road experiments have been cancelled; one student decided to do something completely different; one student is working on an alternative, namely an online questionnaire. The simulator experiment will now be evaluated by the students themselves, rather than be executed with participants. This means that the simulator experiment is not cancelled, but postponed (yet again), possibly for other students to be executed in the near future. To be continued! (?) Meanwhile, an abstract on the literature review on the current state-of-the-art of Human-Machine Interfaces (HMI) in automated driving systems was submitted to a special issue and got accepted for a full paper submission, which' deadline will be 31st of May. Furthermore, an extended elaboration of our truck platoon case study paper is being worked on collaboratively.

As a last remark, personal circumstances kept me from doing more than the abovementioned for this project.



Giulio Mecacci, Philosopher and ethicist

I feel like I don't want to hold any reader any longer after the thorough updates we received from my colleagues. These past months represented for me a very challenging moment. I am currently supervising six students that are writing their thesis on MHC, and teaching two master courses. As Bart said before in this newsletter, education was confronted with the need to salvage as much as possible, and try to push students towards their goals, despite the huge discomfort and loss of motivation that this radical change to their lives brought in. Needless to say, this was very demanding, but the philosophy side of this project made little steps forward as well. As anticipated by Simeon, the objective is clear on what to do in these last months. There are two papers that will be completed and submitted before September. One is about how MHC can help understanding and addressing those value tradeoffs and ethical dilemmas that we face when we decide how to coordinate cars and vulnerable users in urban scenarios. This paper has a wider scope and applicability than it sounds, so stay tuned. A second paper is practically over, and will be submitted within the next few weeks. This one is with Filippo, and it's about the several different kinds of responsibility that exist, and the possible gaps and holes that automation can create for each of them. We also explore how a MHC approach can help filling those numerous gaps. We are also busy organising the final online events with all our partners, we'll update you as soon as possible!



NEWS, PUBLICATIONS AND DISSEMINATION

News

The joint final event of STAD and MHC will be transformed into an on line college tour, starting June 10th. We called it *Automated Driving On-line College tour: spatial and transport impacts and meaningful human control*.

Dissemination

- D. D. Heikoop (2020). *Meaningful human control: Designing safety into automated driving systems*. Research Outreach

Journal articles (submitted)

- S. C. Calvert, G. Mecacci. "A conceptual control system description of Cooperative and Automated Driving in mixed urban traffic with Meaningful Human Control for design and evaluation".

- D. D. Heikoop, S. S. Gürses, R. Happee & M. P. Hagenzieker (2020). Human-Machine Interfaces in Automated Driving Systems: An extensive meta-analysis on the current state-of-the-art. Abstract accepted for full paper submission to the Special Issue for Information, "Test and evaluation methods for Human-Machine Interfaces of Automated Vehicles"

Conference papers

- D. D. Heikoop, M. Adams, J. Baas, & M.P. Hagenzieker, (2020). Big Five personality and ADS: tech-lover stereotype?. To be presented at the International Conference on Traffic and Transport Psychology (ICTTP), 25-27 August, 2020, Goteborg, Sweden.

- D. D. Heikoop, G. K. Srinivasan Ravi Kumar, A. J. van Binsbergen, & M. P. Hagenzieker (2020). Personality and Trust in Automated Cars: A Correlation Study. To be presented at the International Conference on Traffic and Transport Psychology (ICTTP), 25-27 August, 2020, Goteborg, Sweden.

- D. D. Heikoop, A. Rodríguez Sayrol, & M. P. Hagenzieker (2020). Big Five Inventory-based participant selection calculation method. To be presented at the International Conference on Traffic and Transport Psychology (ICTTP), 25-27 August, 2020, Goteborg, Sweden.

- D. D. Heikoop, S. C. Calvert, G. Mecacci, & Hagenzieker, M. P. (2020). A practitioner's view of driver training for automated driving from driving examiners: A focus group discussion. To be presented at the IEEE Forum on Integrated and Sustainable Transportation Systems, 30 June-2 July, 2020, Delft, The Netherlands.

Journal articles (published)

- S. C. Calvert, G. Mecacci, B. van Arem, F. Santoni De Sio, D. D. Heikoop, & M. Hagenzieker, (2020). "Gaps in the control of automated vehicles on roads" IEEE Intelligent Transportation Systems Magazine.

- G. Mecacci & F. Santoni de Sio (2019). "Meaningful human control as reason-responsiveness: the case of dual-mode vehicles". Ethics and Information Technology.

- S. C. Calvert, D. D. Heikoop, G. Mecacci & B van Arem (2019), "A human centric framework for the analysis of automated driving systems based on Meaningful Human Control". Theoretical Issues in Ergonomics Science.

- G. Keeling, K. Evans, S. M. Thornton, G. Mecacci, F. Santoni de Sio (2019). "Four Perspectives on What Matters for the Ethics of Automated Vehicles". In: G. Meyer, S. Beiker (eds) Road Vehicle Automation 6. AVS 2019. Lecture Notes in Mobility. Springer.

Talks and presentations

- F. Santoni de Sio. "Meaningful Human Control over Autonomous Systems". 2nd International Symposium on Symbiotic Intelligent Systems - Challenge to the Edge: How can we design symbiotic society? Osaka, 2020.

- G. Mecacci. "Operationalising Meaningful Human Control into Technical and Institutional Design Requirements". 2nd International Symposium on Symbiotic Intelligent Systems - Challenge to the Edge: How can we design symbiotic society? Osaka, 2020.