SANS

Social & Affective Neuroscience Society

Abstract Book

15th Annual Conference

April 27 - 29, 2023

Santa Barbara, California

1-A-8 Towards human-compatible autonomous car: A study of non-verbal turing test in automated driving with affective transition modelling

Zhaoning Li[1], Qiaoli Jiang[2], Zhengming Wu[3], Anqi Liu[4], Haiyan Wu[1], Miner Huang[2], Kai Huang[2], Yixuan Ku[2]
[1]University of Macau, [2]Sun Yat-sen University, [3]Guangzhou Intelligent Connected Vehicle Pilot Zone Operations Centre, [4]Johns Hopkins University

BACKGROUND AND AIM: Autonomous cars (AC) are indispensable when humans go further down the hands-free route. Although existing literature highlights that the acceptance of the AC will increase if it drives in a human-like manner, sparse research offers the true-to-life ride experience to passengers to examine the human likeness of the AC. Given that it is the key to improving the acceptance of the AC, our main goal was to 1) offer the naturalistic experience from a passenger's seat perspective to measure the human likeness of current ACs; 2) investigate how human passengers ascribe humanness to the AC. **METHODS:** We tested whether the AI driver could create a human-like ride experience for passengers based on 69 participants' feedback in a real-road scenario (Fig. 1). Specifically, we designed a ride experience-based version of the non-verbal Turing test for automated driving. Participants rode in autonomous cars (driven by either human or AI drivers) as a passenger and judged whether the driver was human or AI. Further, based on Lewin's field theory, we advanced a computational model combining signal detection theory with pre-trained language models (PLMs) to predict passengers' humanness rating behaviour (Fig. 2). We used affective transition (AT) between pre-study baseline emotions and corresponding post-stage emotions (collected using the modified Differential Emotions Scale and written description), transformed by PLM, as the signal strength of our model. **RESULTS:** Results of the non-verbal variation of the Turing test showed that the AI driver failed to pass the test because

passengers detected the AI driver above chance. Results of the computational modelling showed that our proposed model could adequately predict passengers' humanness rating behaviour in the test. Further analysis suggested that affective transition (AT), serving as a hypothetical essential part of passengers' subjective ride experience in our model, may play a crucial role in their ascription of humanness. Specifically, we found that the passengers' ascription of humanness would increase with the greater AT. **CONCLUSIONS:** The present study examined whether the current SAE Level 4 AC could create a human-like ride experience for passengers in a real-road scenario for the first time. Though the AI driver failed to pass our test, we offer the first insights into what renders passengers' subjective ride experience truly human-like for future automated driving by using computational modelling. The practical success of basing the computational modelling on Lewin's seemingly abstract and theoretical field theory also speaks directly to his famous maxim that 'there is nothing as practical as a good theory'. Furthermore, our results demonstrate the possibility and feasibility of using NLP techniques (e.g., pre-trained language models) as adjuncts to the interaction between social cognition and artificial intelligence to guide theorising and the generation of conceptual insights.