

Conceptual engineering study of moral judgment using virtual reality

Sean Brantley (1) & Veljko Dubljević (1)

Ethical goal functions for autonomous vehicles (AVs) need to reflect the complexity of human morality [1]. Thus it is important to a) develop and test realistic moral situations with vehicle and traffic hazards, operationalized on the basis of the Agent-Deed-Consequence (ADC) model of moral judgment, b) program moral situations in virtual reality (VR) for testing visual rather than textual input in human moral judgment. This will enable research exploring cross-cultural ethical and policy implications for AVs. The ADC model [2] aims to be an empirical and comprehensive modality for studying how people judge morally salient situations—equally incorporating virtue ethics, deontology, and utilitarianism. Work using written vignettes establishes this model's validity [3], a prerequisite permitting the transition from written scenarios to their current virtual form, created with the Unity game engine. This framework shift maintains the conceptual structure that makes the ADC model a good tool for assessing decision-making heuristics while increasing the experiment's resolution, replacing textual stimuli with language-independent auditory and visual stimuli.

The new vignette does not rely on transmitting textual or verbal information (compare [4]). Instead the moral situation shows the driver either acting cruelly (A-) or kindly (A+) towards a stray dog, then enter a vehicle and either run a stop sign (D-) or obey all traffic rules (D+), concluding with a minor vehicle accident (C-) or a safe arrival at a supermarket (C+). The morally unambiguous (i.e., [A-D-C-] and [A+D+C+]) versions are omitted in order to reduce participant fatigue and to avoid ceiling or floor effects.

This is the first VR experimental setup not relying on any linguistic information and so, it is well suited for cross-cultural data experimentation. A transition necessary if the ADC model hopes to inform the judgments of AVs, which will require comparative analysis of human moral decision making in response to such visual stimuli across many countries, as ethical guidance functions for AVs should not be made based solely on the responses of only English-speaking respondents. With Unity, developers can achieve this relevance by manipulating timing, sound, and motion, creating realistic but purely physical situations and behaviors [5].

References:

1. Dubljević, V. (2021): Toward Implementing the ADC Model of Moral Judgment in Autonomous Vehicles. *Sci Eng Ethics* (2020). <https://doi.org/10.1007/s11948-020-00242-0>
2. Dubljević, V. & Racine E. (2014): The ADC of Moral Judgment: Opening the Black Box of Moral Intuitions with Heuristics about Agents, Deeds and Consequences. Target Article: *AJOB – Neuroscience*, 5 (4): 3-20, (2014).
- 3 Dubljević, V., Sattler, S., & Racine E. (2018): Deciphering moral intuition: How agents, deeds and consequences influence moral judgment, *PLOS One*, <https://doi.org/10.1371/journal.pone.0204631>

4. Parsons, T. D., Gaggioli, A., & Riva, G. (2017). Virtual reality for research in social neuroscience. *Brain Sciences*, 7(4), 42. <https://doi.org/10.3390/brainsci7040042>
5. Francis, K. B., Terbeck, S., Briazu, R. A., Haines, A., Gummerum, M., Ganis, G., & Howard, I. S. (2017). Simulating moral actions: An investigation of personal force in virtual moral dilemmas. *Scientific Reports*, 7(1), 13954-11. <https://doi.org/10.1038/s41598-017-13909-9>