Conceptual engineering study of moral judgment using virtual reality

A methodological shift to facilitate cross-cultural investigations

NC STATE UNIVERSITY
College of Humanities and Social Sciences
Sean Brantley and Veljko Dubljević
Contact Info: veljko_dubljevic@ncsu.edu

Focus of Study

Research Dilemma: Ethical goal functions for autonomous vehicles need to reflect the complexity of human morality.

Research Direction: Utilize virtual reality technologies to develop and test realistic moral situations involving vehicle and traffic hazards, operationalized based on the Agent-Deed-Consequence model of moral judgement.

Agent-Deed-Consequences Model of Moral Judgment

- Utilizes three dominant moral theories to explain heuristic-like intuitive processes that inform moral judgments. [2]
- **Agent(A+/A-) -** Aristotelian virtue ethics – virtues versus vices.[1]
- **Deed(D+/D-) -** deontological ethics – rule-based arguments for morality. [1]
- **Consequences(C+/C-) -** utilitarian ethics – greater good arguments for morality. [1]

Virtual Reality

VR simulations incorporate dynamic, multimodal, and contextually rich stimuli, which can be easily altered between scenarios, enabling assessments of integrative processes over time—likely producing a sweet spot between experimental control and ecological validity. [3]

Research Question: Can virtual non-linguistic moral traffic dilemmas, based on the ADC model, produce similar moral ratings as established textual moral vignettes?

References and Acknowledgments

1. 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).
2. 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
3. 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/brainsci704004
4. Patil, I., Cogoni, C., Zangrando, N., Chittaro, L., & Silani, G. (2014). Affective basis of judgment-behavior discrepancy in virtual experiences of moral dilemmas. Social Neuroscience, 9(1), 94-107. https://doi.org/10.1080/17470919.2013.870091
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
6. Sütfeld, L. R., Ehinger, B. V., König, P., Pipa, G., (2019) How does the method change what we measure? Comparing virtual reality and text-based surveys for the assessment of moral decisions in traffic dilemmas, PLOS ONE, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0223108

We would like to thank for their valuable discussion and feedback the members of the NeuroComputational Ethics Research Group at NC State University. The work on this project is funded by the National Science Foundation CAREER award (#2043612).

Disclosures: None

Methods and Materials

Developed Scenarios in Alphabetical Order

Ambulance Vignette

A scenario following either a nurse agreeing to work a double shift (**A+**) or a drug dealer scheduling a drop (**A-**), portrayed using English voice acting. Both agents face the traffic dilemma of pulling over for an ambulance (**D+**) or accelerating past a construction zone (**D-**). Branching from these decisions, the agent either navigates the scenario successfully (**C+**) or causes an accident, endangering the occupant of the ambulance (**C-**).

Bank Vignette

This scenario starts with a ransom note instructing a father to produce money in exchange for his son. The father either withdraws the money from a bank, even giving some to a homeless woman (**A+**), or steals it from the bank (**A-**). On his way to the swap, he must choose to wait for a train to pass (**D+**) or ignores the warning signals (**D-**). In the end, the son is returned (**C+**), or the money, son, and father's car are taken (**C-**).

Cow Vignette

The cow vignette portrays a couple preparing for their child's delivery. The father is either kind and responsive (**A+**) or abusive (**A-**). On the way to the hospital, the husband either stops for a cow in the road (**D+**) or swerves into the opposite lane (**D-**). The decisions either lead to successful navigation of the situation (**C+**) or cause an accident and a miscarriage (**C-**).

Crowd Vignette

A scenario following either a dedicated fire marshal (**A+**) or an aggressive one (**A-**) is called to assist with a riot. Once on the scene, the marshal sees a distressed woman, and he either blocks the rioters chasing her using his vehicle (**D+**) or drives into the group (**D-**). The woman gets away safely (**C+**), or the marshal hurts several of the rioters (**C-**).

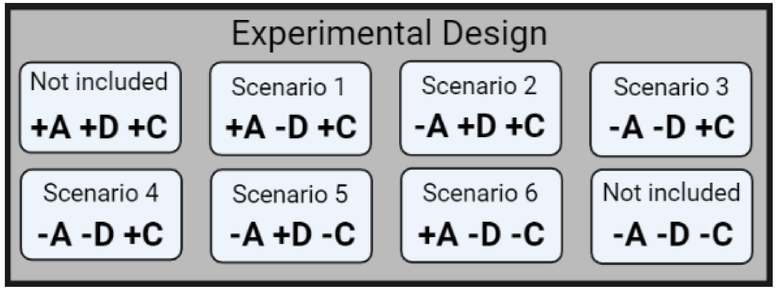
Dog Vignette

This scenario starts with a moral agent on his way to a pet shop. He is either kind to a stray dog (**A+**) or is cruel and abusive to the animal (**A-**). After driving away, he either stops at a stop sign (D+) or runs the stop sign (**D-**). These choices either get him to the shop safely (**C+**) or are the cause of an accident (**C-**). This vignette is the only one that is entirely language-independent.

Intersection Vignette

The intersection vignette follows a mother that must take her son to school after finding out he missed the school bus. She is either understanding and reassuring (**A+**) or abusive and violent (**A-**). During the drive, she either stops at a red light (**D+**) or runs the red light (**D-**). These choices either get the child to school on time (**C+**) or are the cause of an accident (**C-**).

Design



Created with BioRender.com

- Participants will be randomly assigned one of the six versions of a single vignette. They will be asked to rate the moral acceptability of the situation they view on a scale of one (least acceptable) to ten (most acceptable).
- Prior research using ADC-modeled textual vignettes provides legitimacy for not including entirely positive and negative scenarios. These versions have been shown to produce expectedly high and low moral ratings, and inclusion would decrease the number of participants rating the morally conflicting scenarios. [2]

Materials

1. Oculus Quest 2 Head Mounted Display
2. Unity Real-Time Development Platform Version 2020.3.18f1
3. 11th GenIntel® Core™ i7-11700KF
4. NVIDIA GeForce RTX 2080 SUPER

Possible Results

- Increased utilitarian **judgments** (a higher valuation of consequences), aligning with two previous research observations, which indicate that utilitarian **actions** were more likely to occur during Immersive VR experiments paired with physical salience amplifiers.[4,5]
- Contrastingly, level of abstraction (text-based vs. naturalistic VR environments) and modality (Desktop vs. VR) associated with moral situations has been shown to elicit no judgement differences.[6]
- Text-based versions of the ADC model alone predict negative consequences being outweighed by positive agent-deed pairs and negative deeds being outweighed by positive consequences. [2]

Limitations

- There is an inherent U.S. bias in the virtual environments, such as driving side, sign choice, and environment scenery.
- We are assessing moral judgments, not moral actions, which may provide less explanatory power given the goal of informing autonomous vehicle functions due to the lack of the real-time component of actions.