

International Conference on AI Ethics
in Cross-Cultural Contexts
Capital Normal University
Beijing, China
June 20-22, 2025

MIT's Moral Machine Experiment:
A cautionary tale about
foreseeable societal impact

Zsófia Zvolenszky

zvolenszky.zsafia@btk.elte.hu

zvolenszky@nyu.edu

elte-hu.academia.edu/ZsofiaZvolenszky
phil.elte.hu/zvolenszky

Department of Logic
Institute of Philosophy
Faculty of Humanities
Eötvös Loránd University (ELTE)
Budapest, Hungary



Conclusions

- **Shifting away from the Shared-Responsibility Question** (the restricted as well as the broader version, which included Level2, partial automation) **is a major loss**. For society, for policymaking, for research debates and their impact on AI ethics.
- Was it **foreseeable** that the MM experiment would contribute to such a shift? **Yes**.
- It was baked into its questionnaire design!
- Substantial involvement from philosophers in AI-ethics-related projects, debates is needed. With attention to details of language, interpretation, conceptual clarity, research purpose and consequences so that
 - we can be **more strategic about keeping important questions on the discussion table**.
 - **detect and counter moves that** (intentionally or unintentionally) **can, foreseeably, shift discussion away from the key questions**.