

Advisor

Human

AI

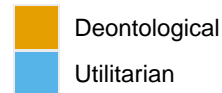
Baby dilemma

Human

AI

Bike dilemma

Advice



-0.2      -0.1      0.0      0.1      0.2

Pre-post shift in judgement after seeing advice (0–1 scale)

Deontological ↔ Utilitarian

