Improving The Public's Perception Of Autonomous Vehicles By Communicating The Consistency Of Autonomous Vehicle Algorithms

Heather E. K. Walker, Alexander C. Walker, Rafał Muda, Martin Harry Turpin, Lana M. Trick, Jonathan A. Fugelsang, & Michał Białek

AUTONOMOUS VEHICLES

- Involved in *half* the number of collisions per million miles (compared to human drivers)
- Projected to reduce collisions by *up to 90%*

AVERSION TO AUTONOMOUS VEHICLES

- Despite being safer overall, people don't trust AVs, & are averse to their presence on roads

- Would need to appear *5x* safer than human drivers to be considered equally as acceptable

Why Don’t People Like AVs?

Dislike computer algorithms making moral decisions?

Don’t understand how AV algorithms work?

Loss of personal agency over situations?

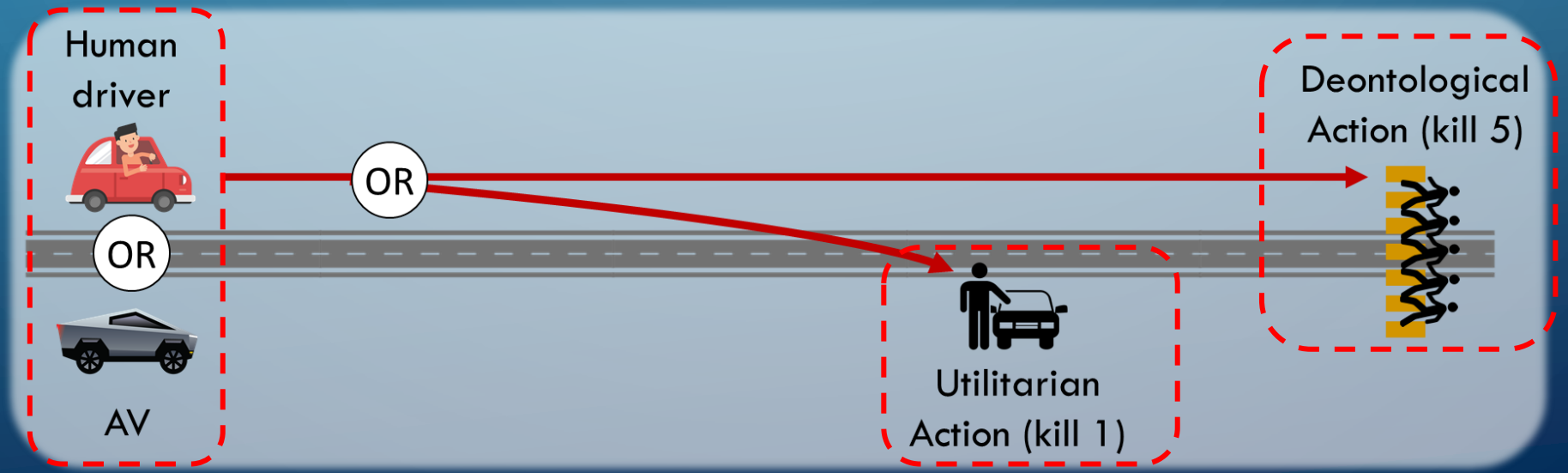
RESEARCH QUESTION

Can explaining the consistency of AV algorithms (in comparison to the inconsistency of human drivers) reduce peoples' aversion to AVs?

CURRENT RESEARCH

- 3 experiments, used moral dilemmas to assess perceptions of AVs and human drivers

- Base Scenario:



CURRENT RESEARCH

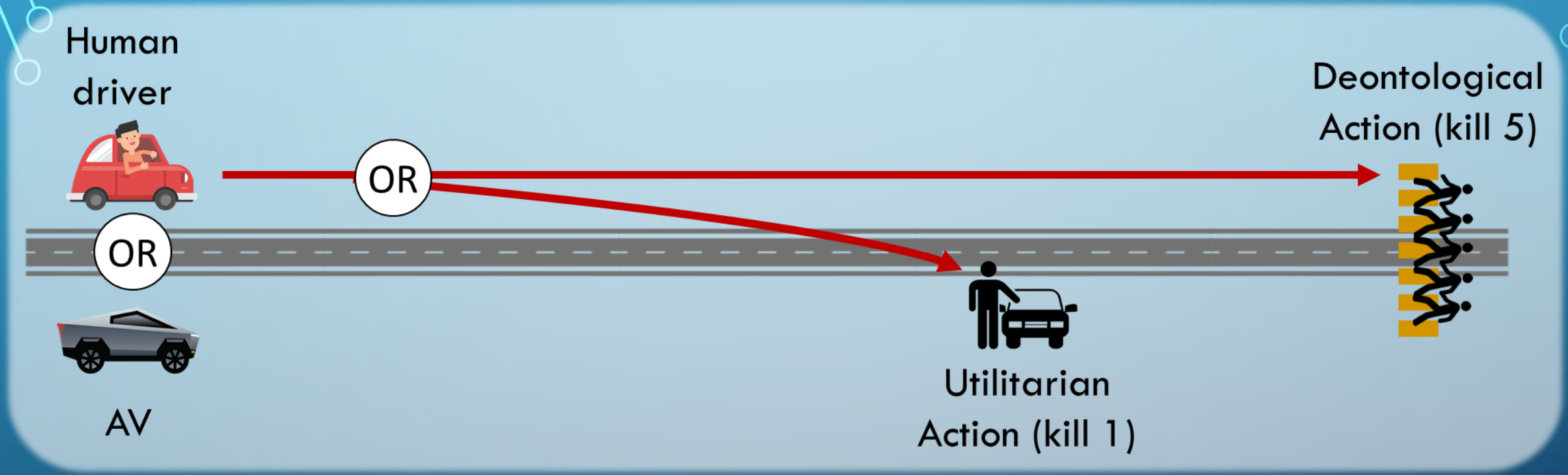
Rating Scales

Bad	○ ○ ○ ○ ○ ○ ○	Good
Immoral	○ ○ ○ ○ ○ ○ ○	Moral
Unpredictable	○ ○ ○ ○ ○ ○ ○	Predictable
Caused a Great Deal of Harm	○ ○ ○ ○ ○ ○ ○	Caused No Harm
Deserves a Great Deal of Blame	○ ○ ○ ○ ○ ○ ○	Deserves No Blame
Actions were Unacceptable	○ ○ ○ ○ ○ ○ ○	Actions Were Acceptable
* Untrustworthy	○ ○ ○ ○ ○ ○ ○	Trustworthy

Moral Perception → Average of Bad/Good and Immoral/Moral ratings

* Study 3 *only*

STUDY 1

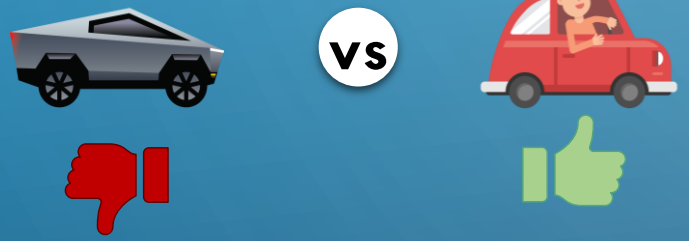


+

Explanation ***VS*** No Explanation

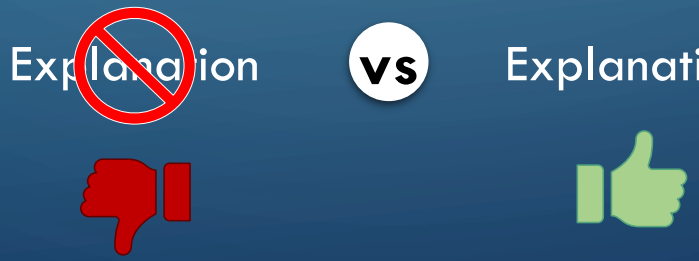
STUDY 1 – Main effects

Pilot

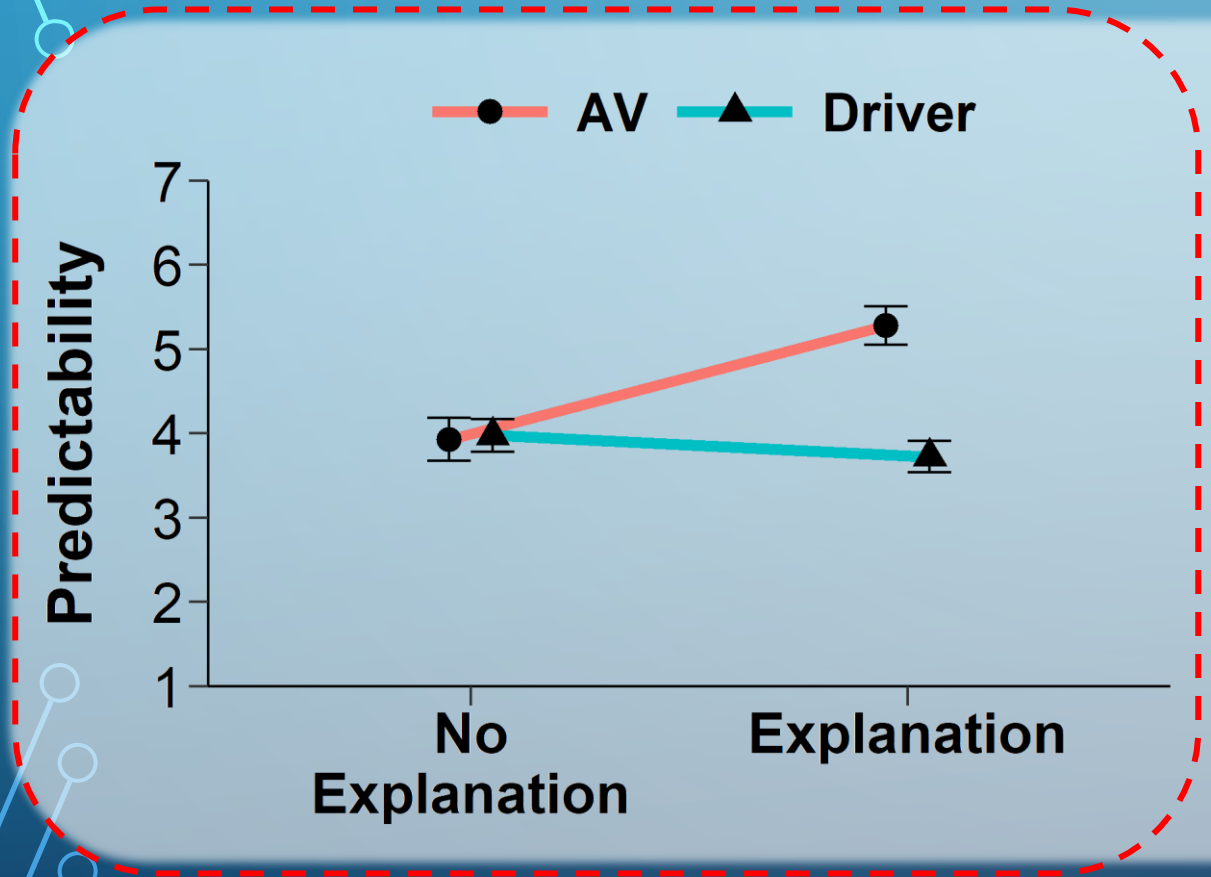


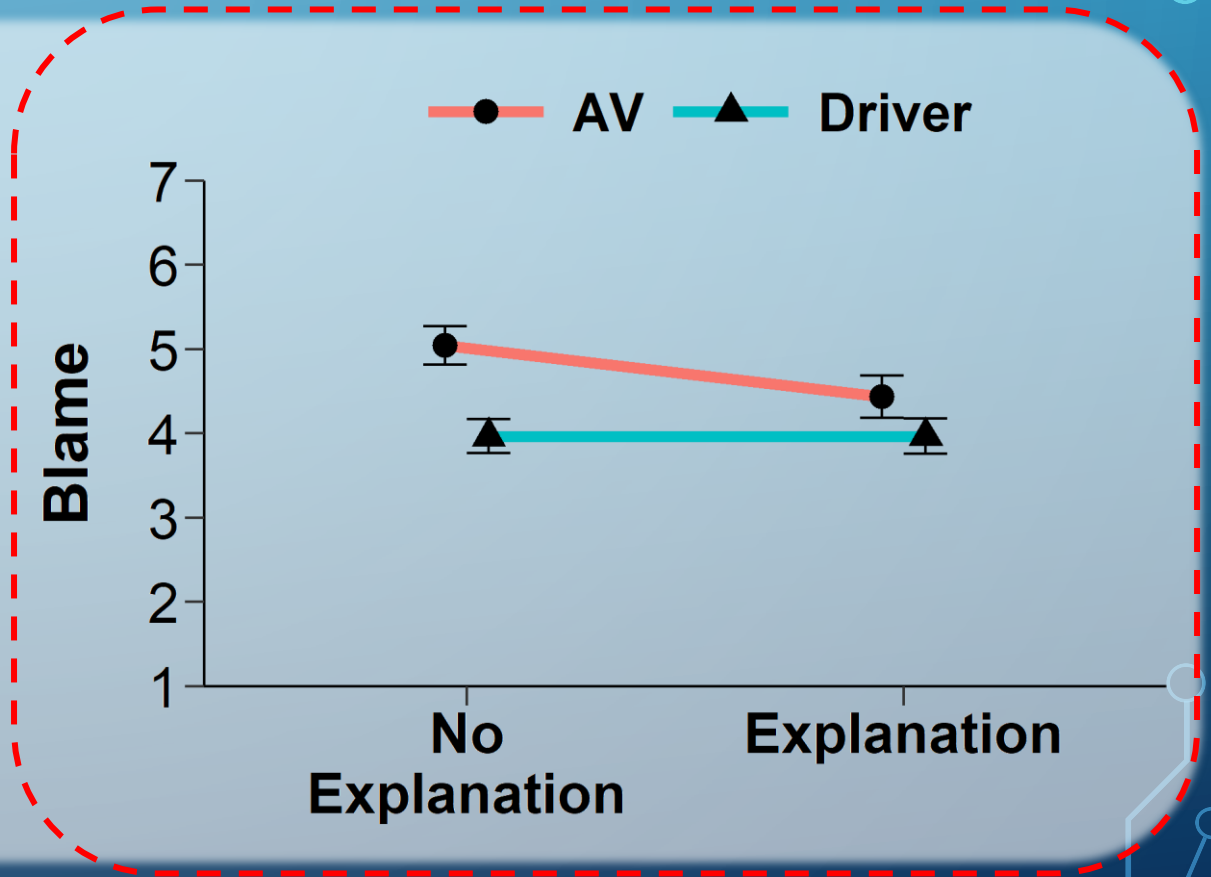
Action

Explanation

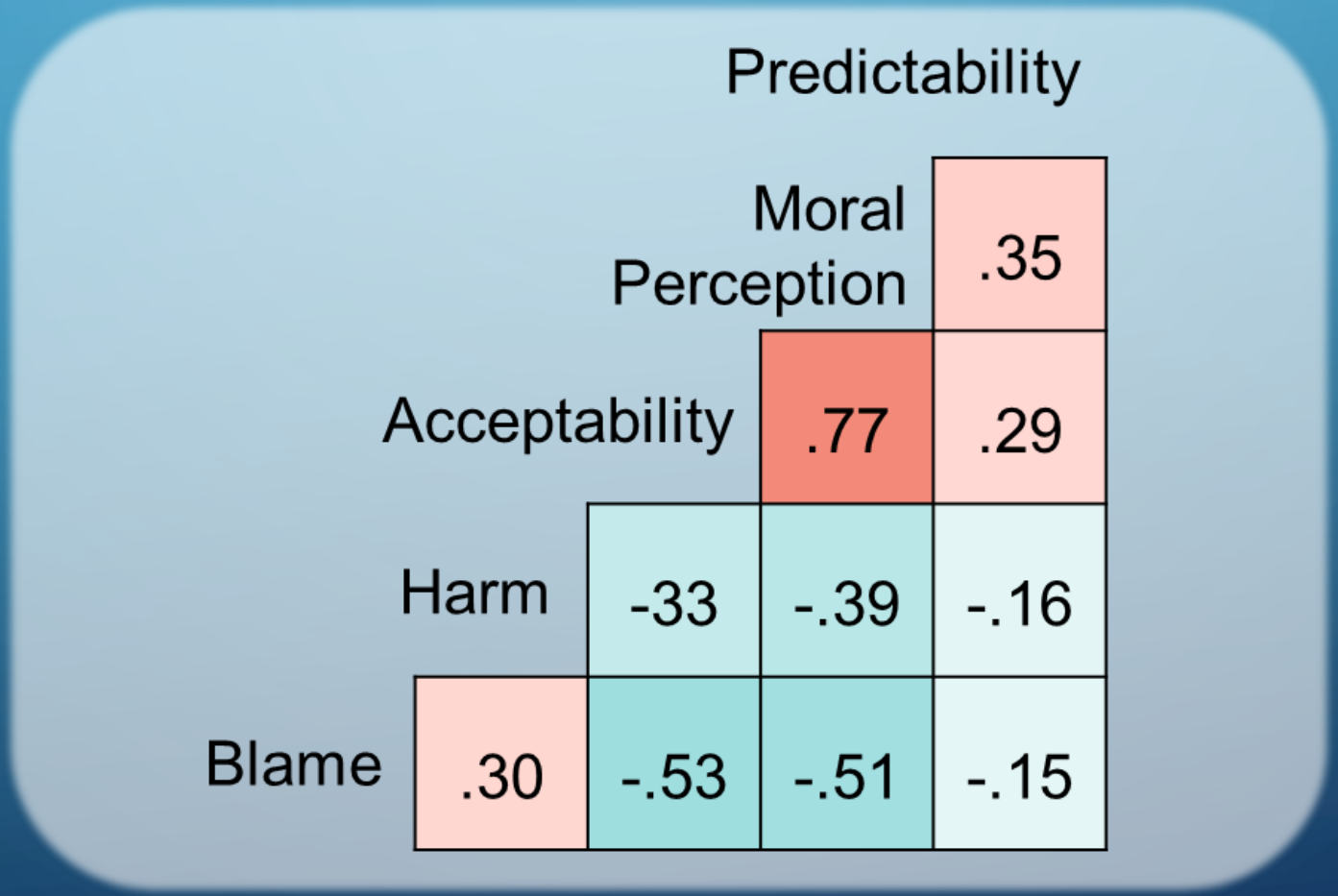


STUDY 1 - Interactions

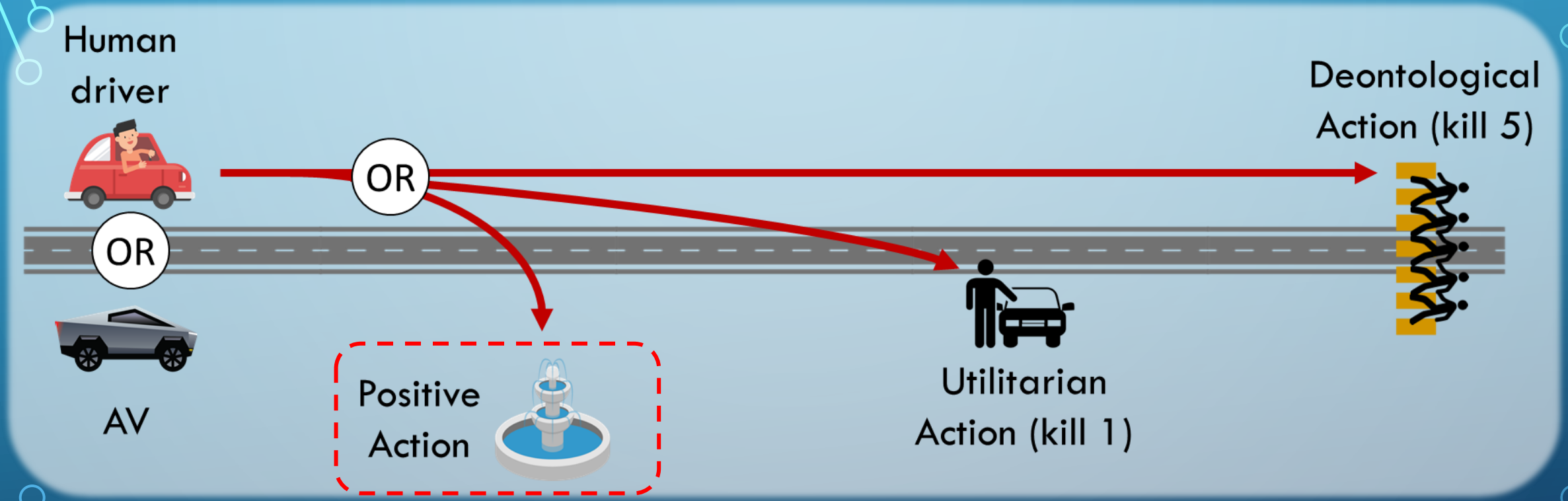




STUDY 1 - Correlations



STUDY 2



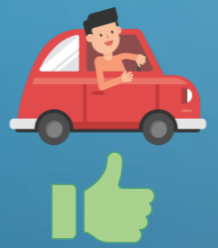
+

Explanation ***VS*** No Explanation

STUDY 2 – Main effects

Pilot

vs



Action

vs

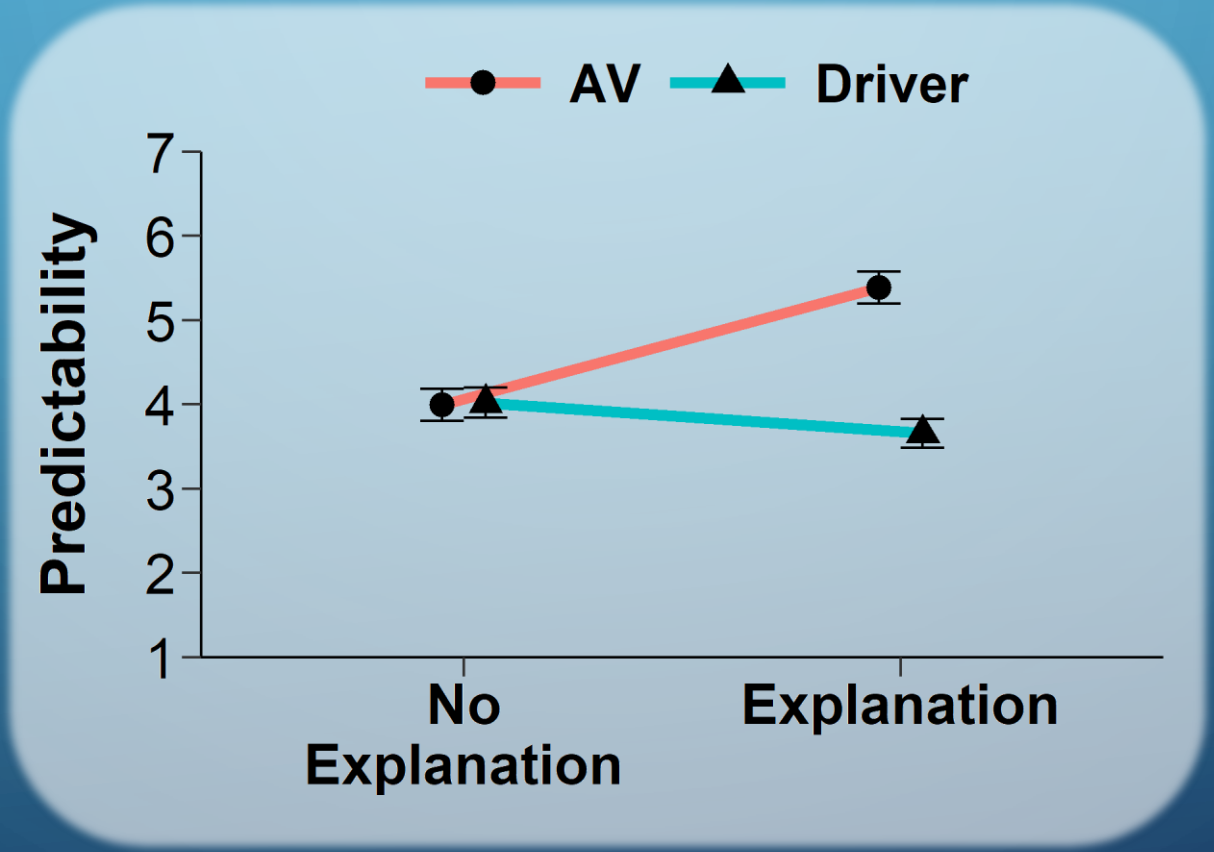
Explanation

Explanation

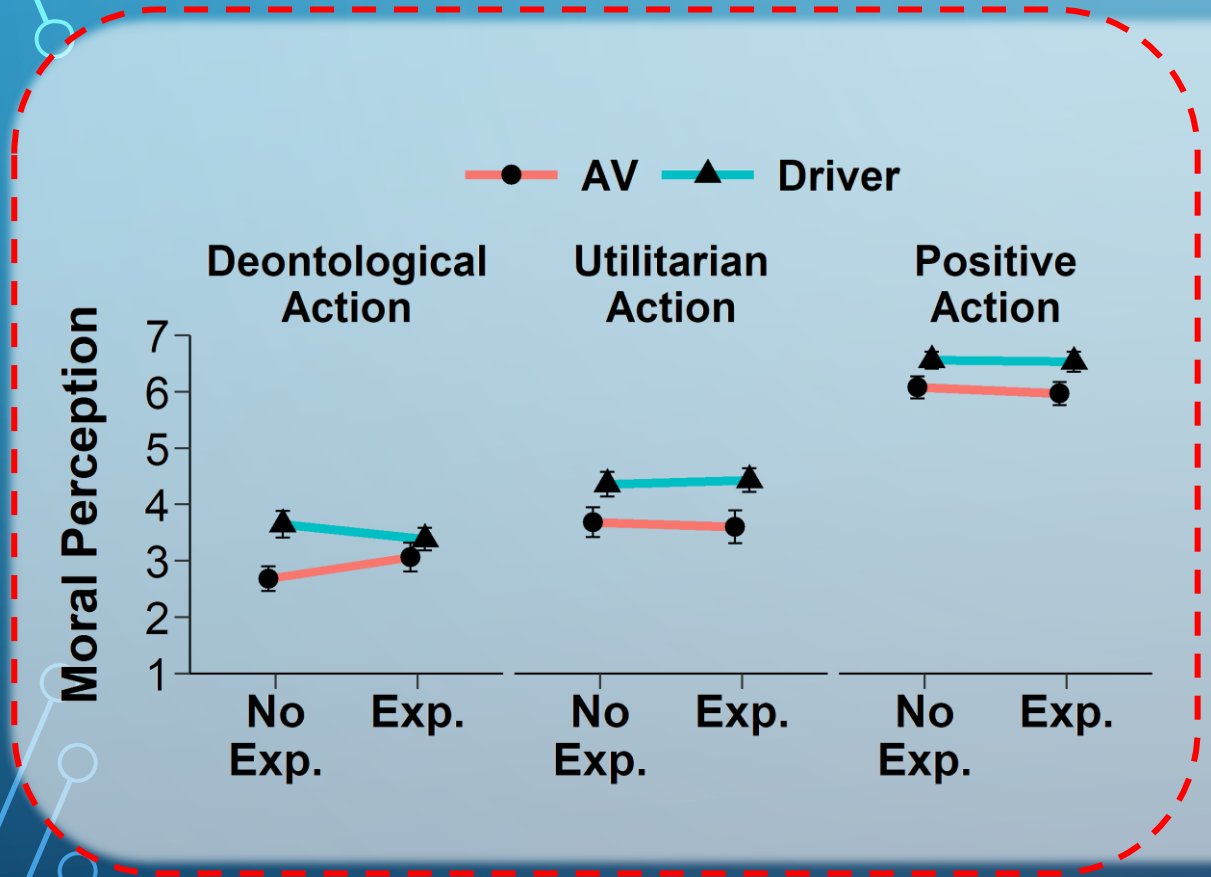
vs

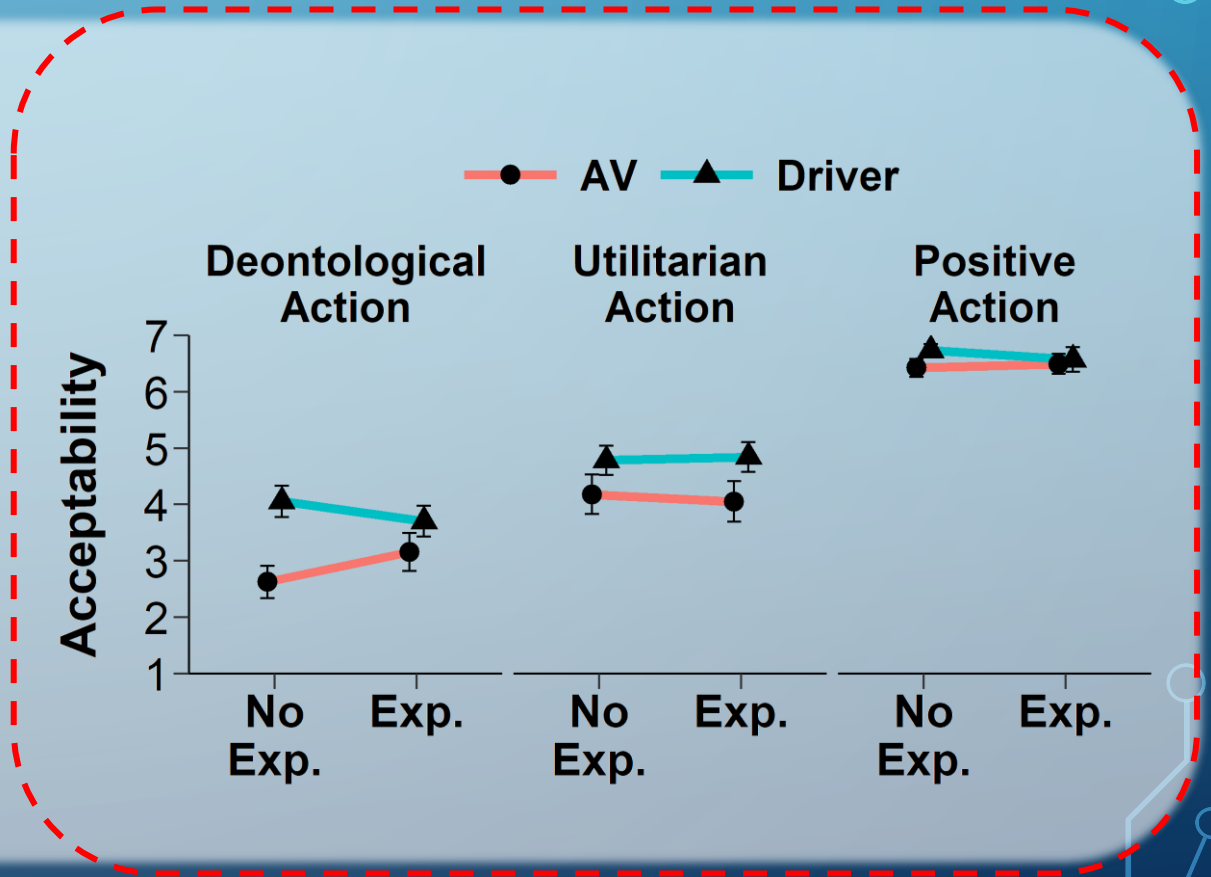
Explanation

STUDY 2 - Interactions

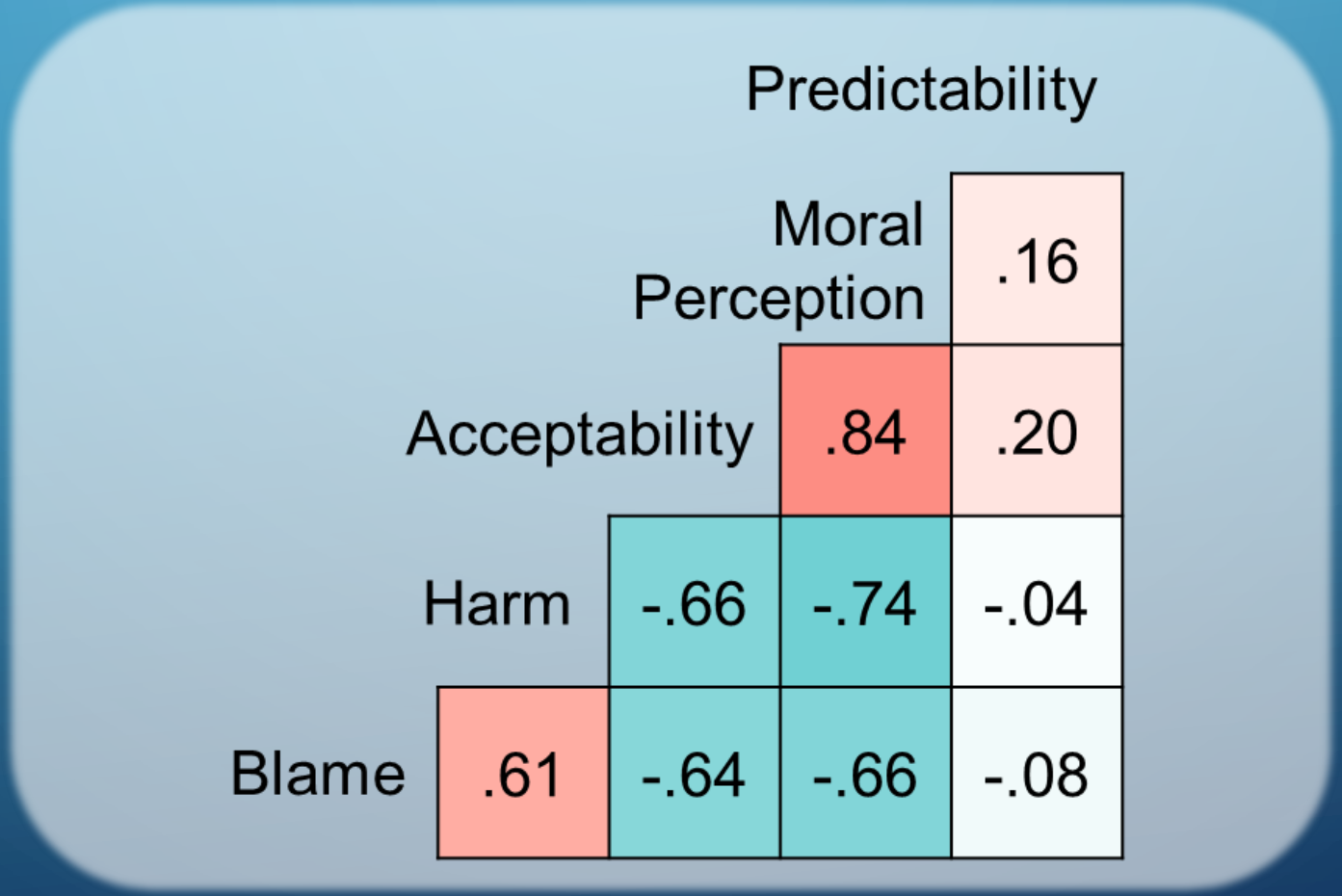


STUDY 2 - Interactions

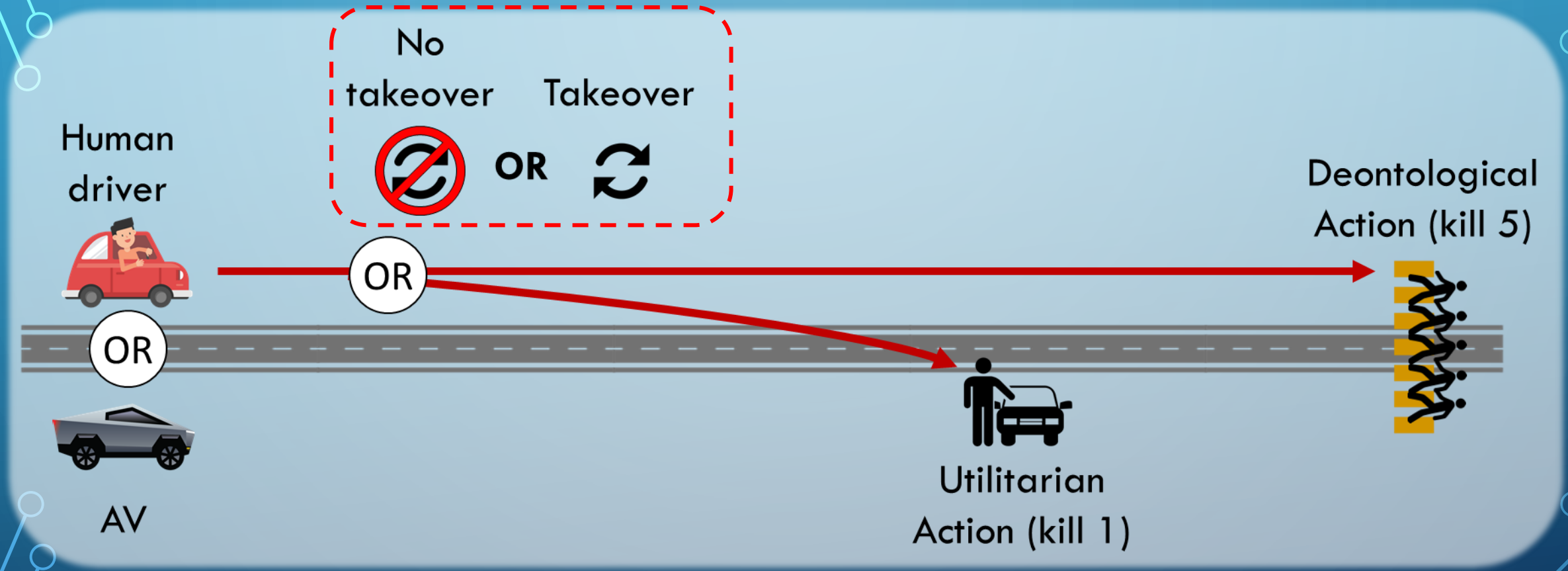




STUDY 2 - Correlations

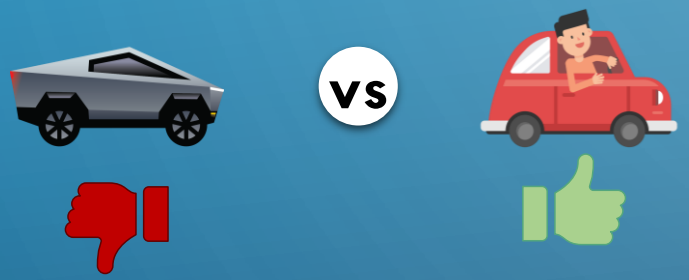


STUDY 3


No takeover
Takeover
OR
Human driver
OR
OR
AV
Deontological Action (kill 5)
Utilitarian Action (kill 1)

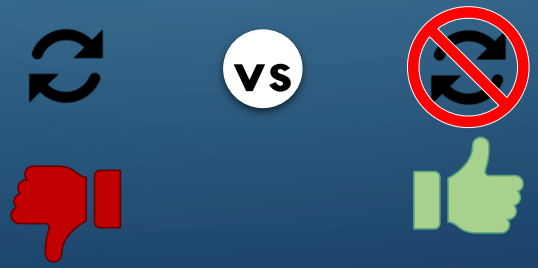
STUDY 3 – Main effects

Pilot

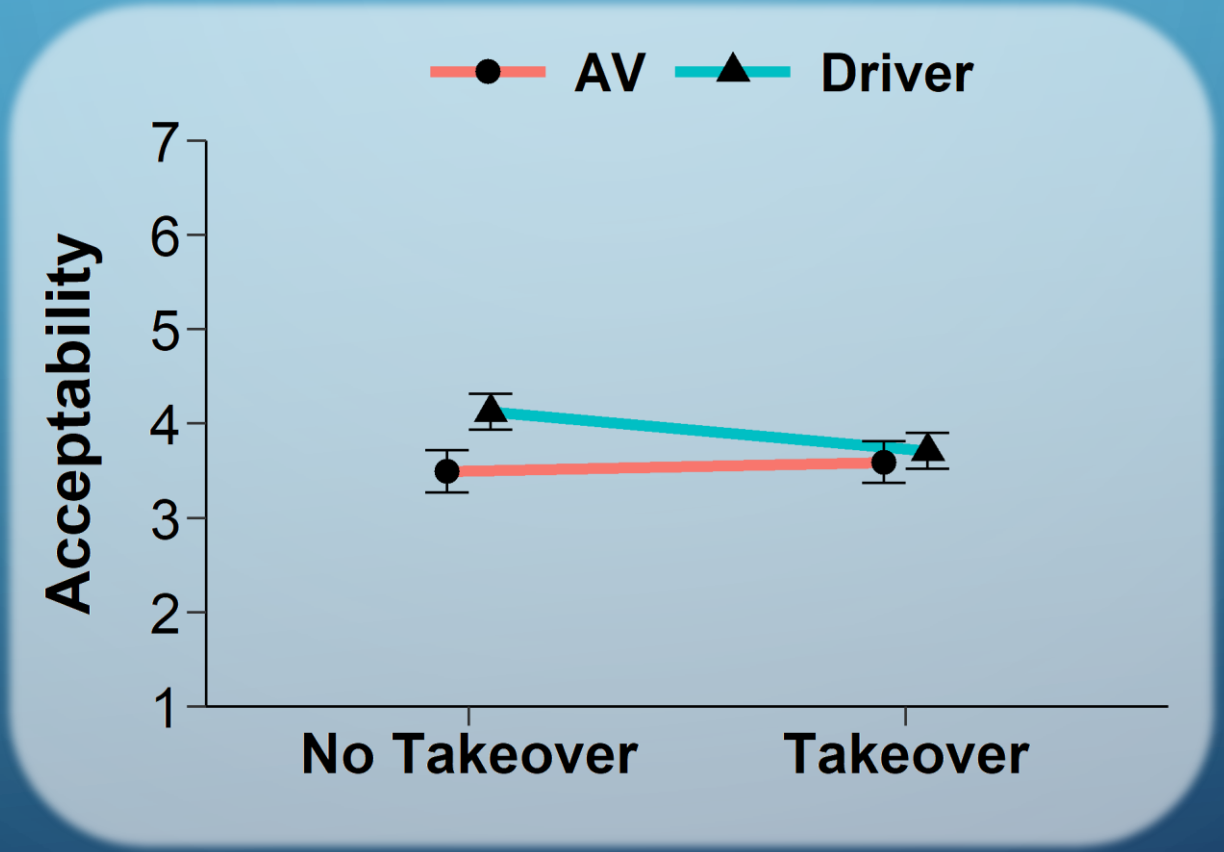


Action

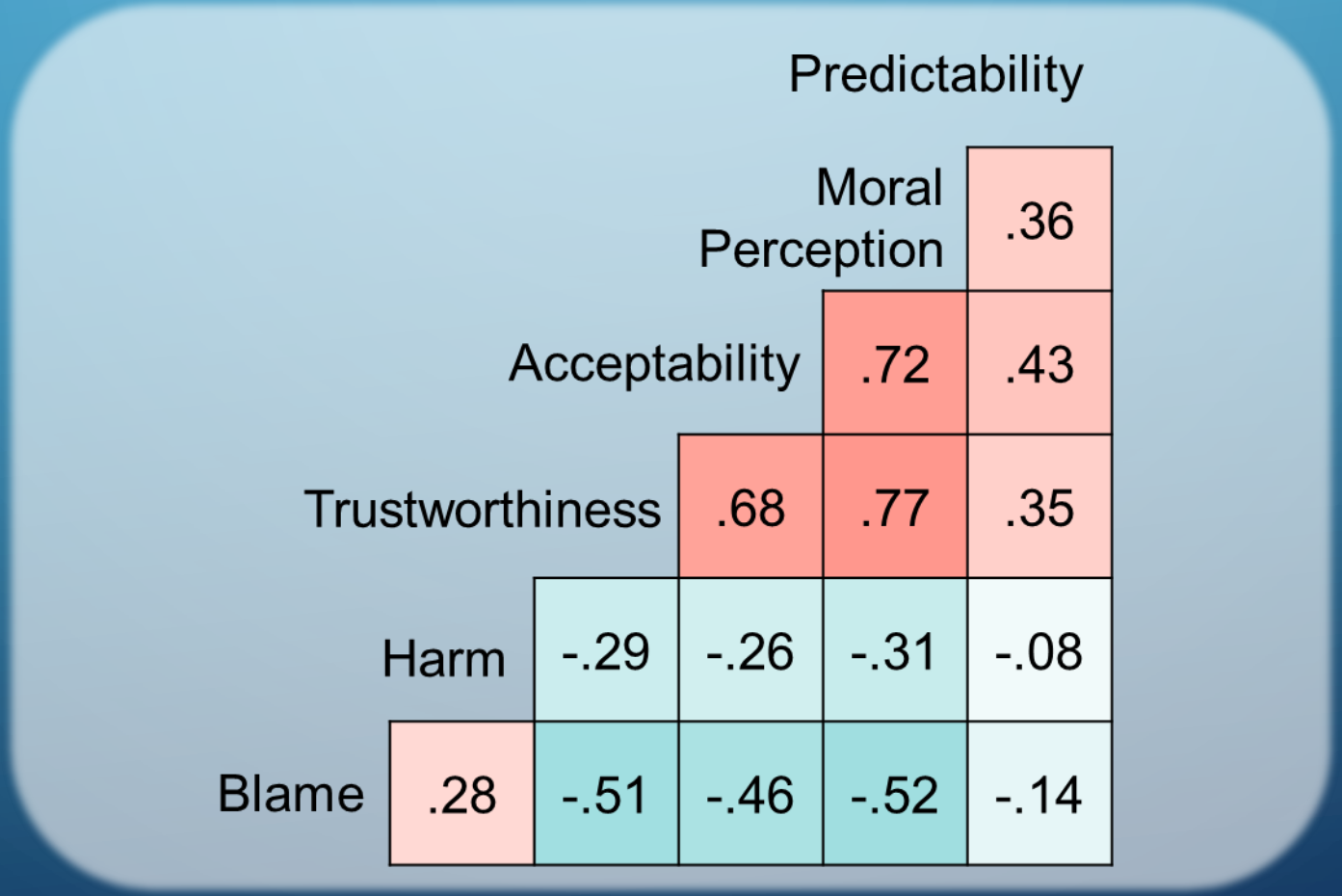
Takeover



STUDY 3 - Interaction



STUDY 3 - Correlations



SUMMARY

AVs are judged more negatively than human drivers

Explaining the consistency of AV actions improves perceived predictability of AVs, reduces AV aversion

AVs making positive actions are judged better than human drivers making negative actions

People are averse to takeover actions, particularly when human drivers take over

CONCLUSION

Highlighting the consistency of AV algorithms could improve perception of autonomous vehicles and reduce barriers to their mass adoption

ACKNOWLEDGMENTS

Heather Walker & Lana Trick

Alexander Walker, Martin Turpin, & Jonathan Fugelsang

Rafał Muda

Michał Białek

QUESTIONS?