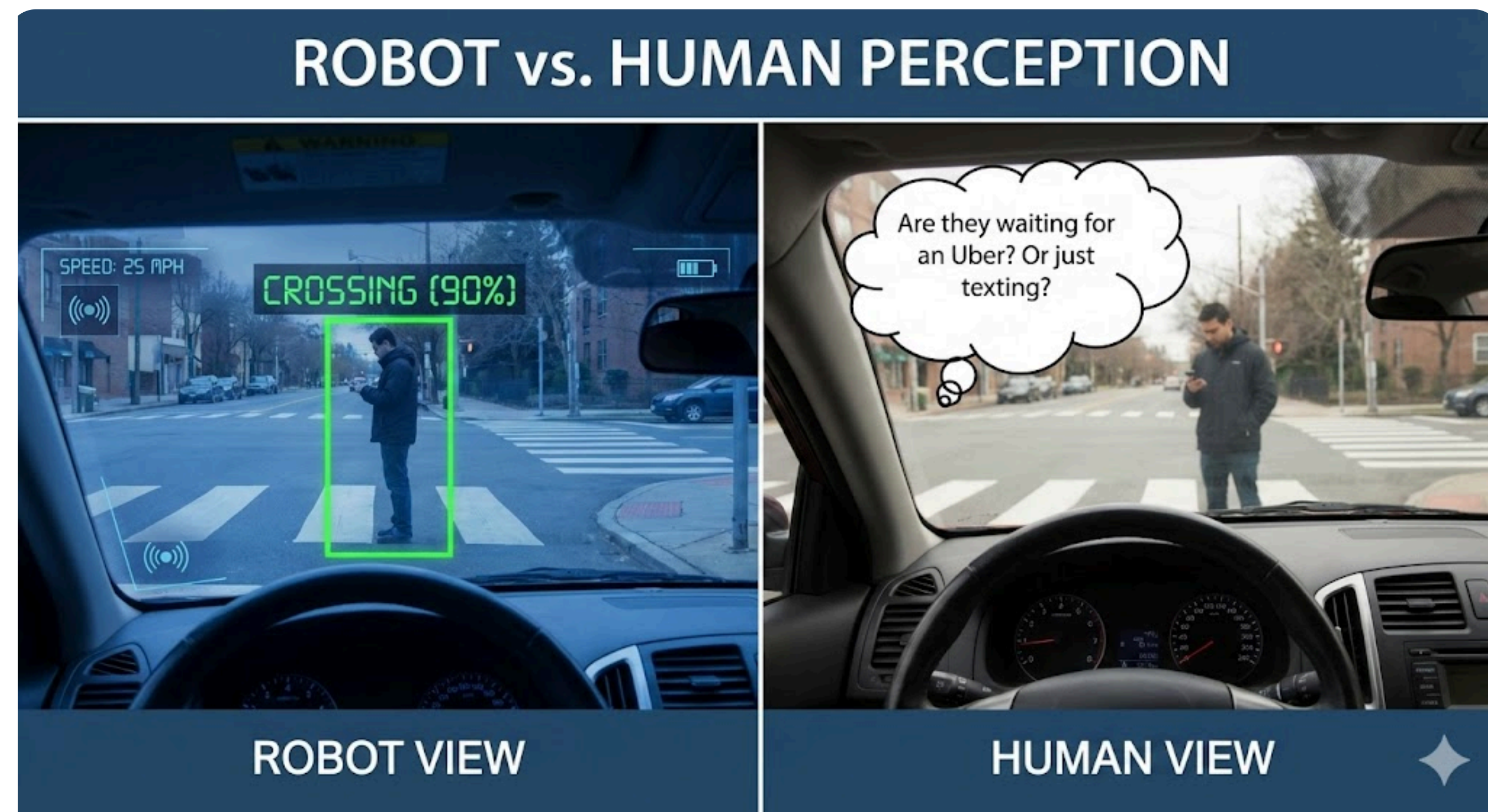


# Investigating Contextual Variations in Explaining Plausible Narratives of Social Intention in Driving

Author: Jang Hun Oh (J.Oh-2@student.tudelft.nl)  
Professors & Supervisors: Hayley Hung, Arthur Mercier, Vitaliy Popov

## 1. Motivation

Most intelligent systems rely on observable kinematics to predict future trajectories rather than underlying intentions. However, reliance on kinematics alone is often insufficient because a single behaviour, like a pedestrian standing near the road, can generate multiple plausible narratives[1].



## 2. Research Questions

How can we systematically map the dimensions of variation in prototypical driving situations to the range of plausible social intention narratives they generate?

Sub Question 1:

What are the primary dimensions of variation that can be found in current literatures on **social perception** and **intelligent systems** that influence the perception of intention in the context of driving?

Sub Question 2:

How can **external scripts** and **internal scripts** be used to model the generation of divergent intention narratives in distinct driving case studies?

Sub Question 3:

What is the relationship between the **open-endedness** or **constraint level** of a driving scenario and the **degree of variation** in plausible intention narratives?

Sub Question 4:

Based on this analysis, what **recommendations** can be provided to intelligent systems designers for a research strategy that scales from constrained to unconstrained scenarios?

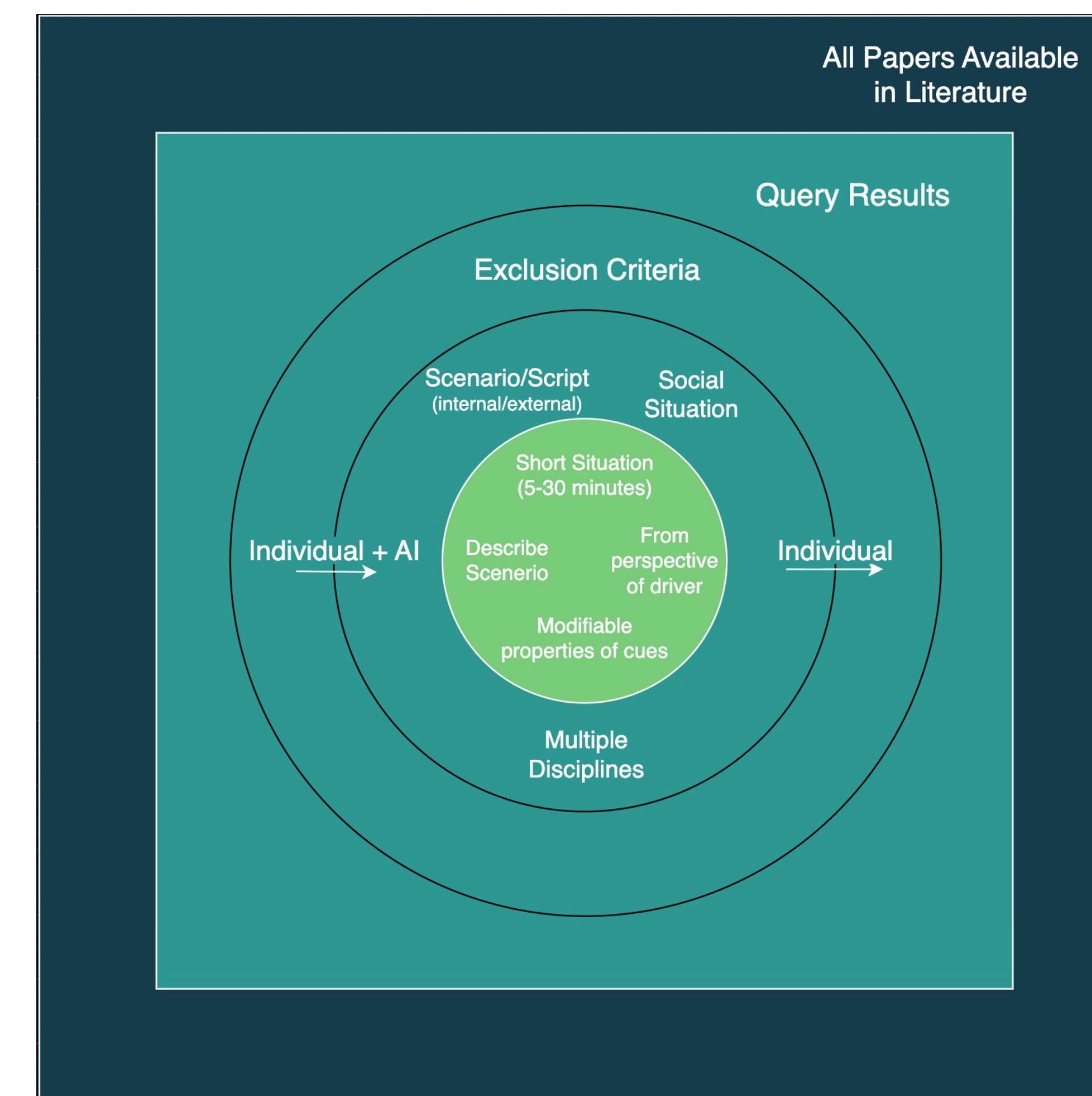
## 3. Methodology

### 3.1 Literature Survey

Searched on Scopus using a query combining concepts of Intention/Behavior, Subjectivity/Context (e.g., script theory, norms), the driving domain and exclusion terms outside this domain.

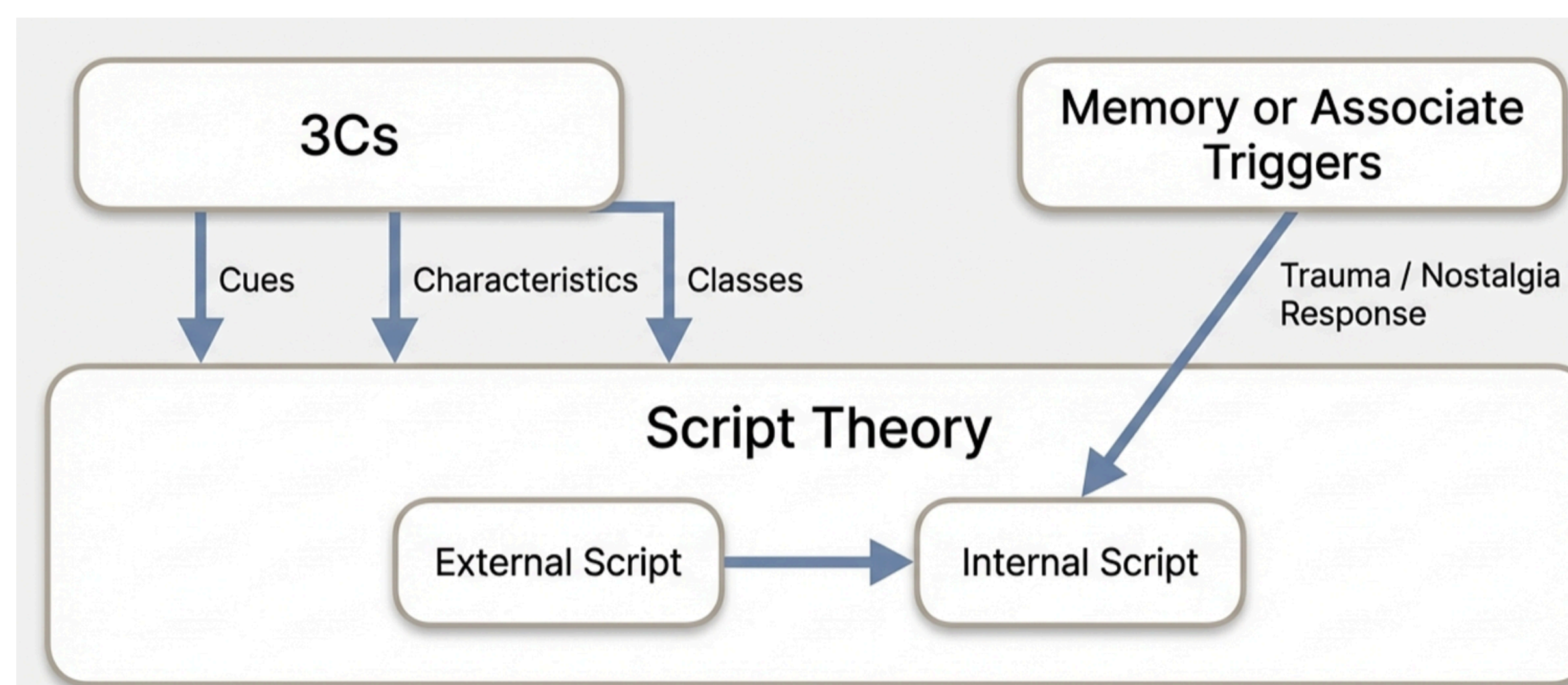
Results were grouped by subject area (progressing from least to most populated) to ensure multidisciplinary diversity and prevent engineering dominance.

Applied a ranking strategy within each subject area, selecting the top 5 most recent and top 5 most cited papers.



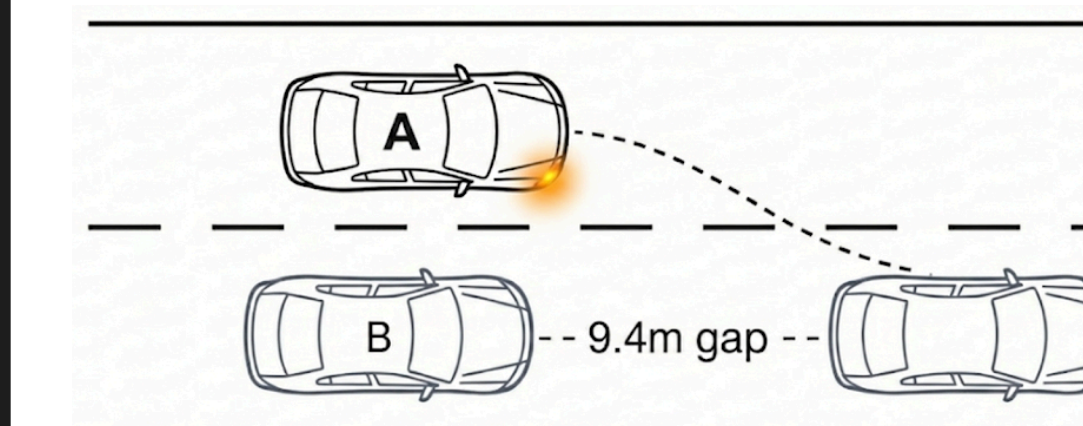
### 4.2 Dimensions Extraction Framework

This framework makes use of the works by Rauthmann et al. [2] for the 3Cs and by Fischer et al. [3] for script theory.



## 4. Case Studies

### Case Study A: Lane Merge



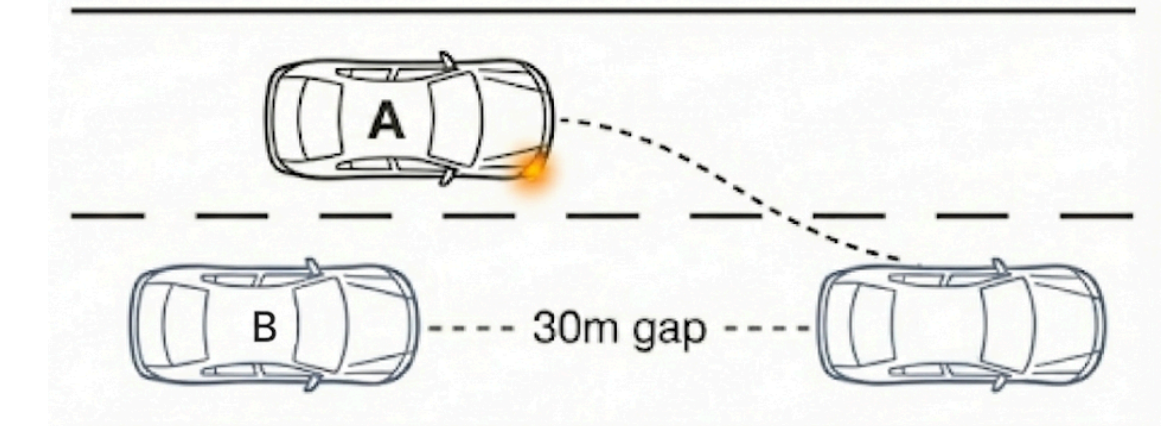
Cues: 9.4m gap size, Late blinker signal timing

Characteristics: Hostile/Aggressive

External Script: Standard lane merge rules

Internal Script: Territorial Defence

Machine Interpretation: Safe to Yield



Cues: 30m gap size, Late blinker signal timing

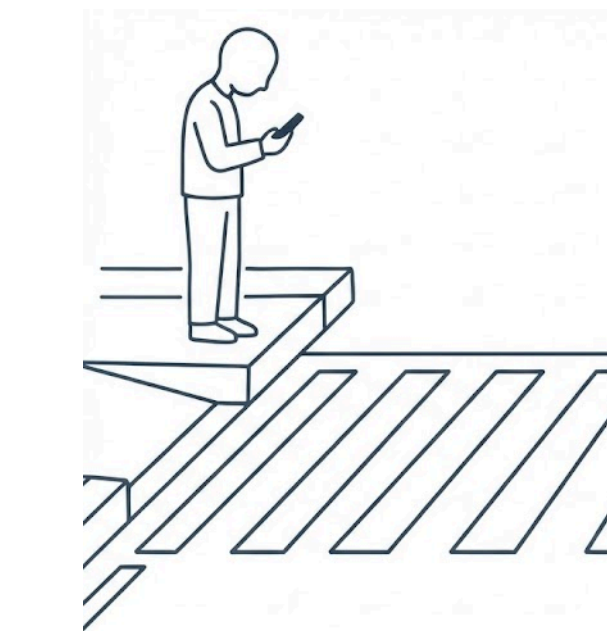
Characteristics: Safe/Polite

External Script: Standard lane merge rules

Internal Script: Willing to yield

Machine Interpretation: Safe to Yield

### Case Study B: Pedestrian in Uncontrolled Zebra Crossing Scenario



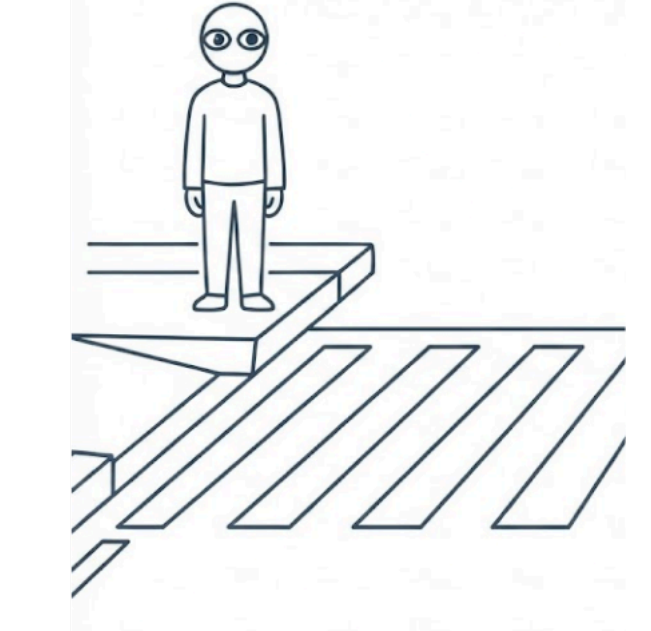
Cues: Stationary, Head down, No eye contact

Characteristics: Distracted/Unaware

External Script: Standard road crossing laws

Internal Script: Caution/Hesitation

Machine Interpretation: Safe to Pass



Cues: Stationary, Head up, Eye contact with vehicle

Characteristics: Attentive/Aware

External Script: Standard road crossing laws

Internal Script: Trust that pedestrian sees vehicle

Machine Interpretation: Stop for pedestrian

## 5. Discussion and Recommendations

Narrative open-endedness is inversely proportional to physical and social constraints; when constraints are weak, human and machine interpretations diverge but when constraints are strong, interpretations converge



## 6. Conclusion

Social driving failures often arise because machines perceive trajectories while humans perceive norm violations. Robust systems must align machine reasoning with human expectations.

Future research requires datasets annotated with socio-psychological labels (e.g., "Hesitation") like the PIE dataset by Rasouli et al.[1].

## References

- [1] Amir Rasouli, Iuliia Kotseruba, Toni Kunic, and John Tsotsos. Pie: A large-scale dataset and models for pedestrian intention estimation and trajectory prediction. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pages 6261–6270, 2019. [2] John F. Rauthmann, Ryne A. Sherman, Chapter 13 - Conceptualizing and measuring the psychological situation, Editor(s): Dustin Wood, Stephen J. Read, P.D. Harms, Andrew Slaughter, Measuring and Modeling Persons and Situations, Academic Press, 2021, Pages 427–463, ISBN 9780128192009, <https://doi.org/10.1016/B978-0-12-819200-9.00009-0>. [3] Frank Fischer, Ingo Kollar, Karsten Stegmann, and Christof Wecker. Toward a script theory of guidance in computer-supported collaborative learning. Educational Psychologist, 48(1):56–66, 2013.