

From glances to outcomes

How counterfactual risk estimates relate to real-world crash odds

Jonas Bärghman, Minxiang Zhao, Shweta Joshi, Moli Zhou
Chalmers University of Technology

Monitored, rated and regulated, but not quantified in terms of crash and injury risk

- EU ADDW regulation (2023/2590): distraction warnings are mandatory
- Euro NCAP: substantially extended DMS assessment from 2026
- Driver monitoring in essentially every new European car

What is the impact of ADDWs (and HMIs) in terms of crashes and injuries?
No commonly agreed on method...

Method aim: To enable estimation of crash and injury risk of HMI and ADDWs...

1

Why can't we state the crash and injury risk impact of inattention today?

2

What would a quantitative crash and injury risk metric for inattention look like, and can we trust it?

Traditional assessment vs. estimating crash and injury risk

Traditional assessment

- Guidelines, ISO protocols, Euro NCAP
- **Relatively straightforward**
- Says little about an HMI's or ADDW's impact on crash and injury risk

Methods that address crashes

- Methods estimating crash and injury risk of HMIs and ADDWs
- **Hard to understand: chains of models and assumptions**
- Not validated against real crashes
→ hard to trust

To be used, a crash-risk method has to be understandable, and tested against real crashes

Side note: Four methods for assessment – answering different questions

MiRA

Kircher & Ahlström (VTI)

Is this driver sampling enough, right now?

Definition, hindsight-free

Drive-In Lab

Kujala et al. (Jyväskylä)

What does this task do to car-following margins?

Demand & behavior

AHEAD

Reimer, Lee et al. (MIT)

Does the counter-measure do what it was designed to?

Verification

Counterfactual sim.

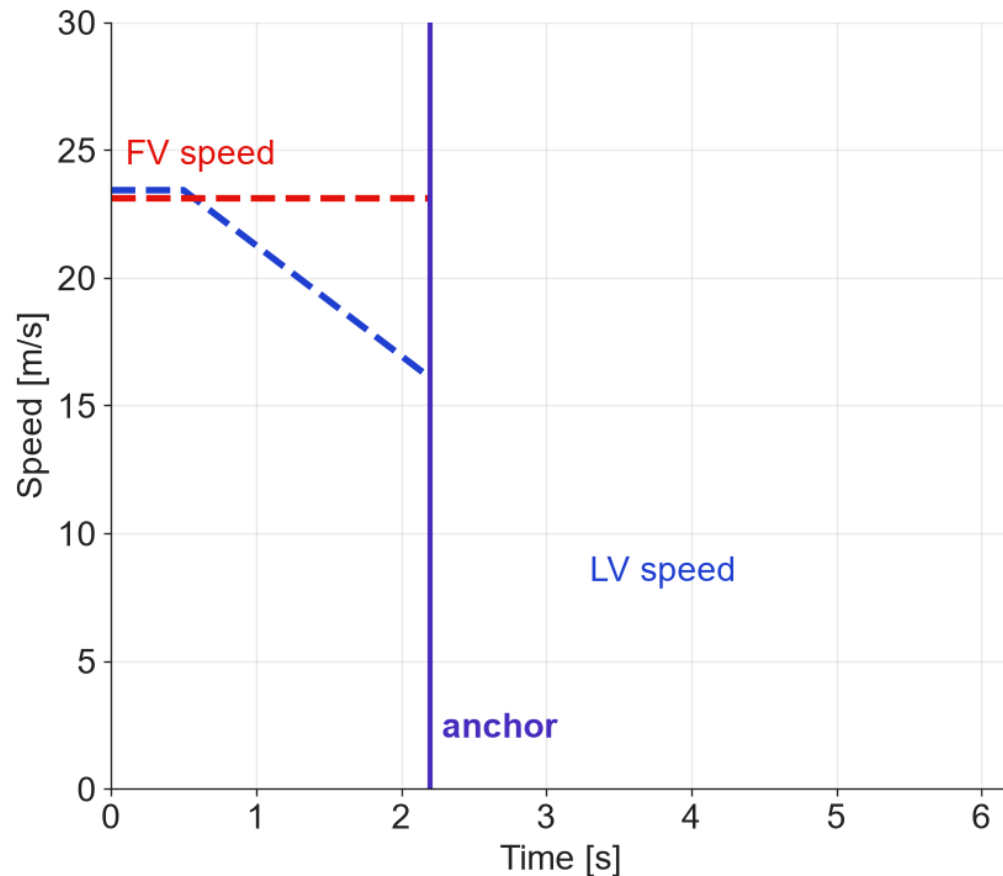
Bärgman et al. (Chalmers)

How does it impact crashes and injuries?

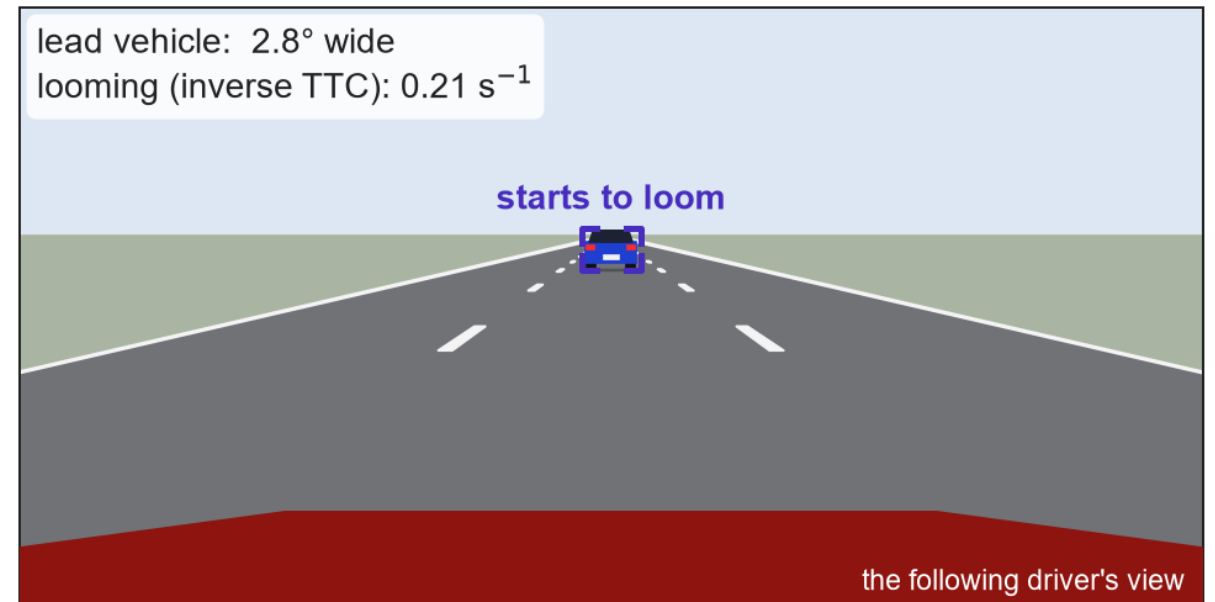
Safety impact

Not alternatives but different questions: ranking them is incorrect

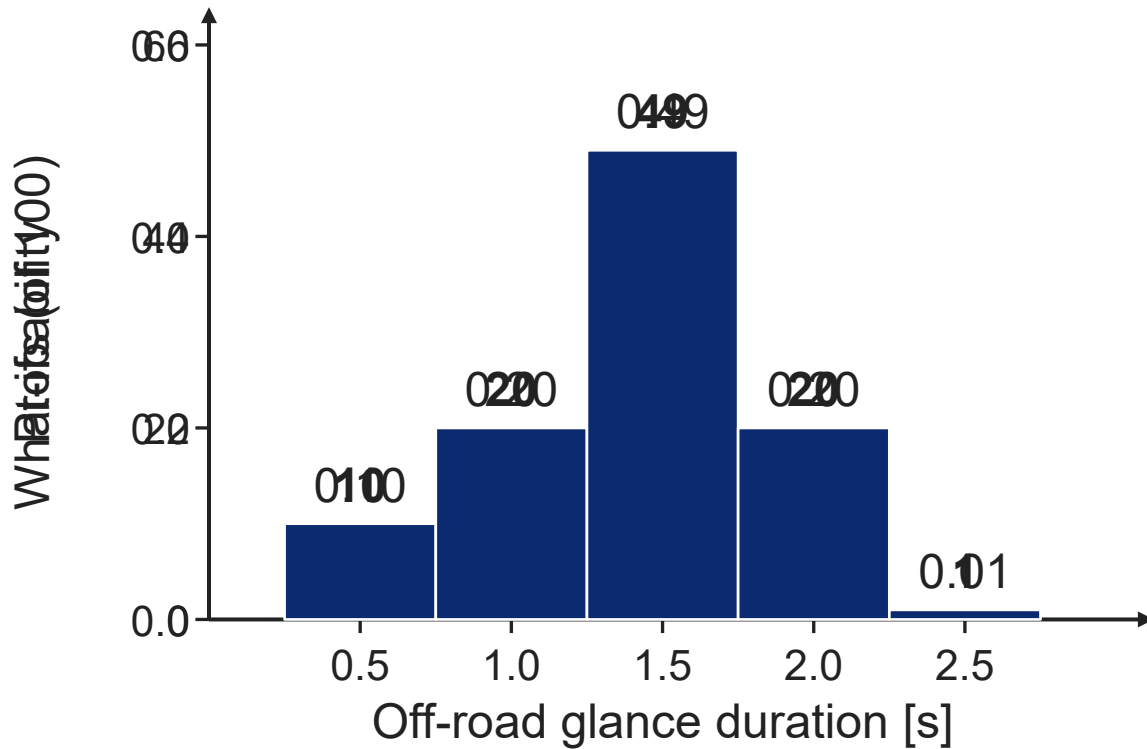
One rear-end crash, seen three ways



It starts to loom: the anchor



A very simple glance distribution: five durations

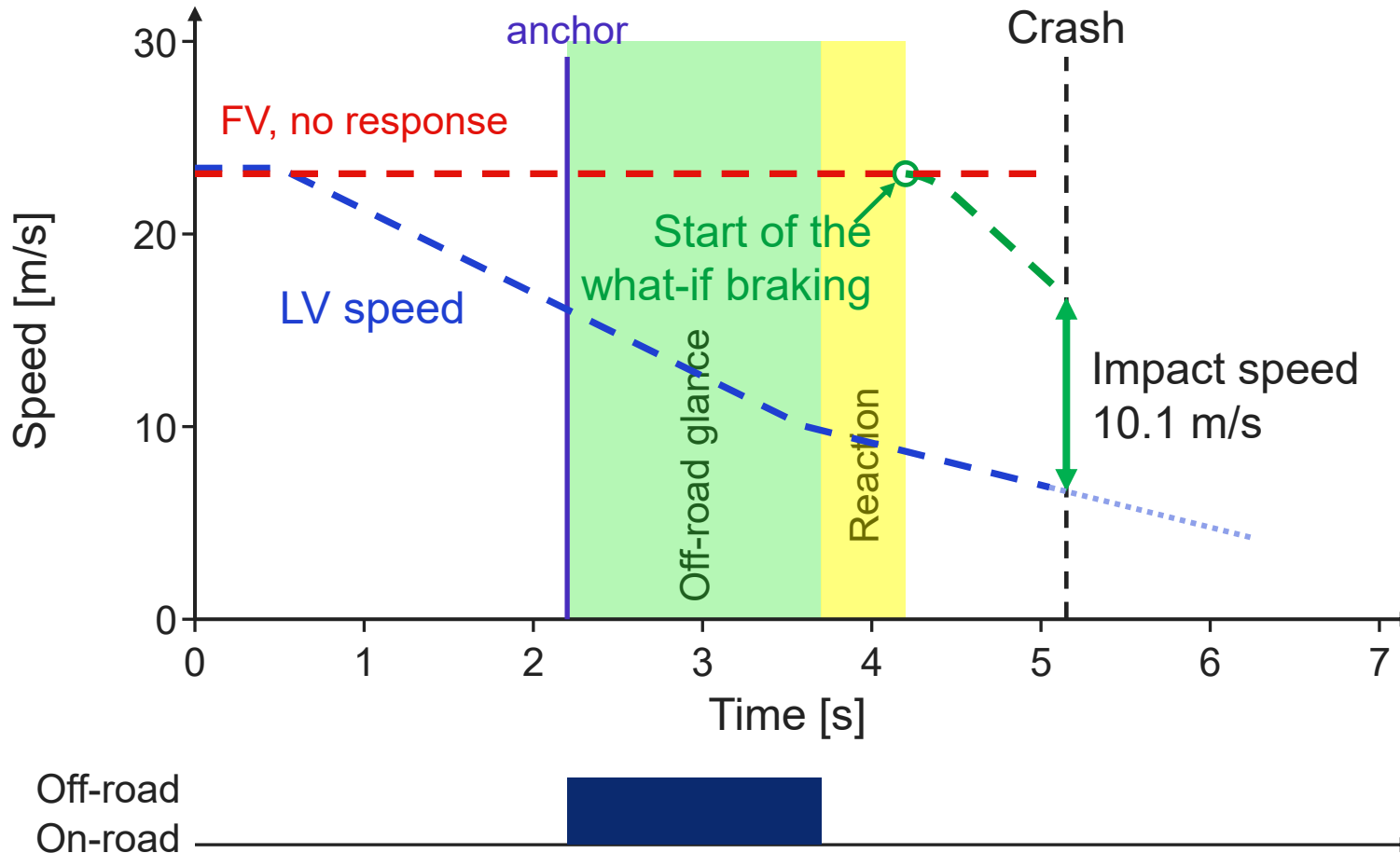


What do the probabilities represent?

Take one crash and run it 100 times, each time with a glance drawn from the distribution: 100 what-ifs

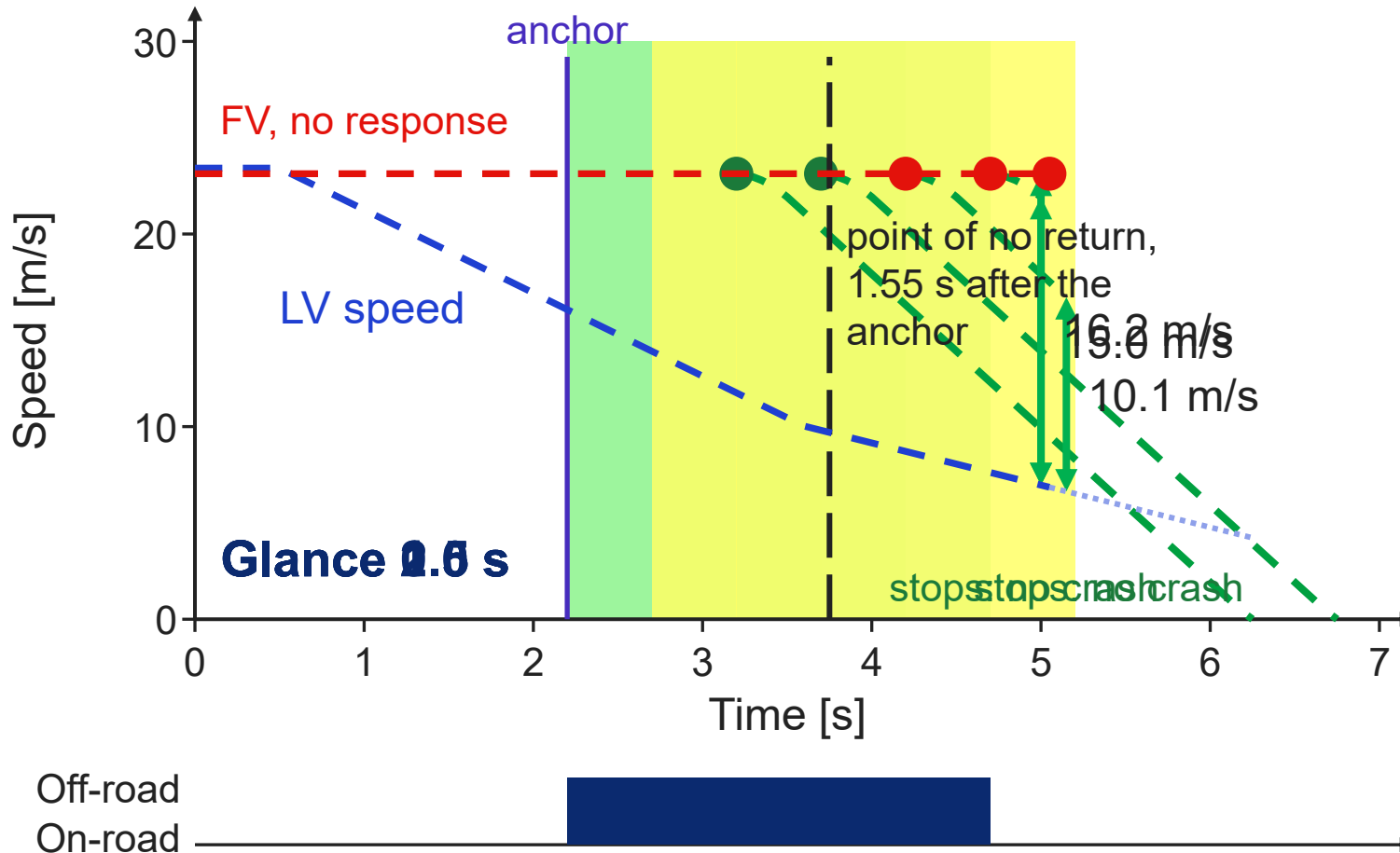


One what-if: a 1.5-s glance away from the road



1. Eyes off 1.5 s, from the anchor
2. Eyes back: 0.5 s to react
3. ➔ What-if braking starts 2.0 s after the anchor
4. Braking builds up to 8 m/s^2
5. Too late to avoid a crash
6. Impact speed 10.1 m/s

All five what-ifs of the same crash



Glance	Of 100	Outcome
0.5 s	10	no crash
1.0 s	20	no crash
1.5 s	49	crash, 10.1 m/s
2.0 s	20	crash, 15.0 m/s
2.5 s	1	crash, 16.2 m/s

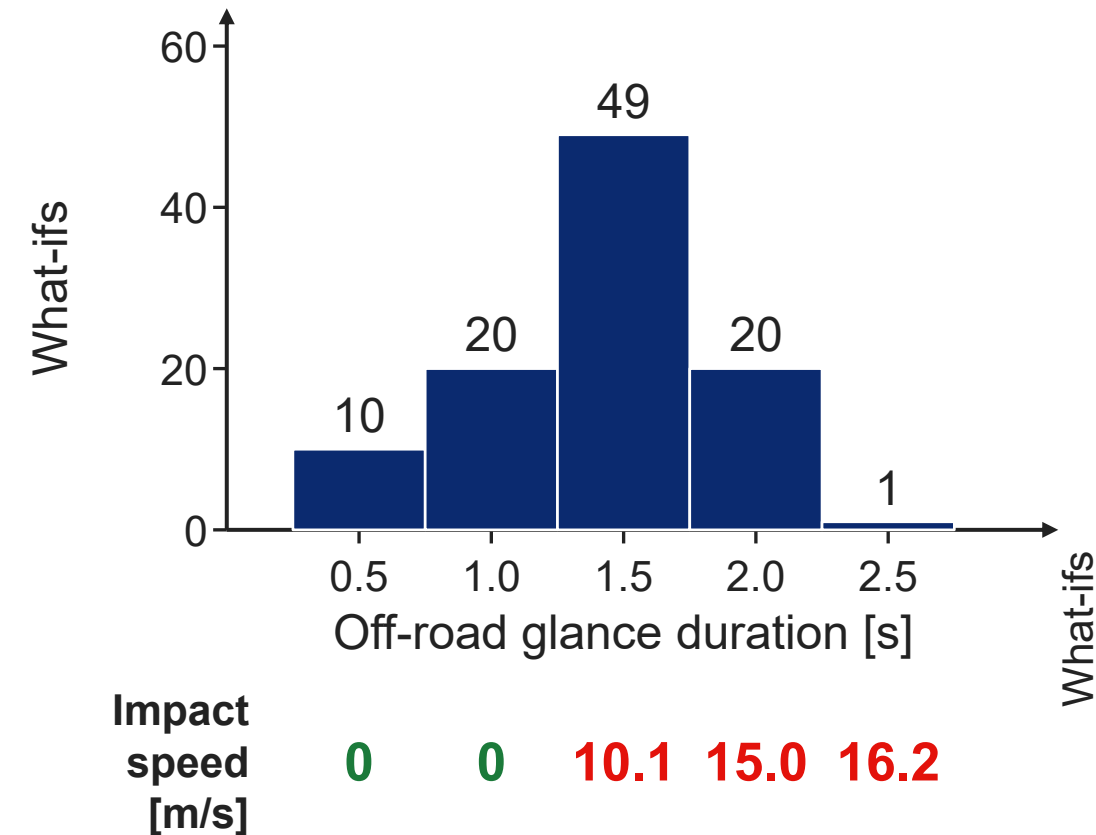
The 2.5-s glance: the crash comes before the driver even starts to brake

**Crash in 49 + 20 + 1
= 70 of 100 what-ifs**

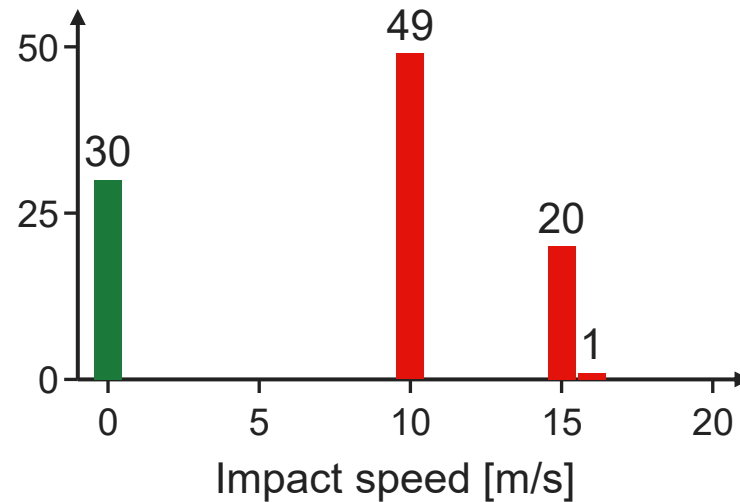
Crash risk for this crash: 0.70

Braking that starts after the point of no return cannot avoid this crash: the dots, one per what-if

One crash, 100 what-ifs: the impact speed distribution



[0 0 0 ... 0] (10 times)
[0 0 0 ... 0] (20 times)
[10.1 10.1 ... 10.1] (49 times)
[15.0 15.0 ... 15.0] (20 times)
[16.2] (1 time)



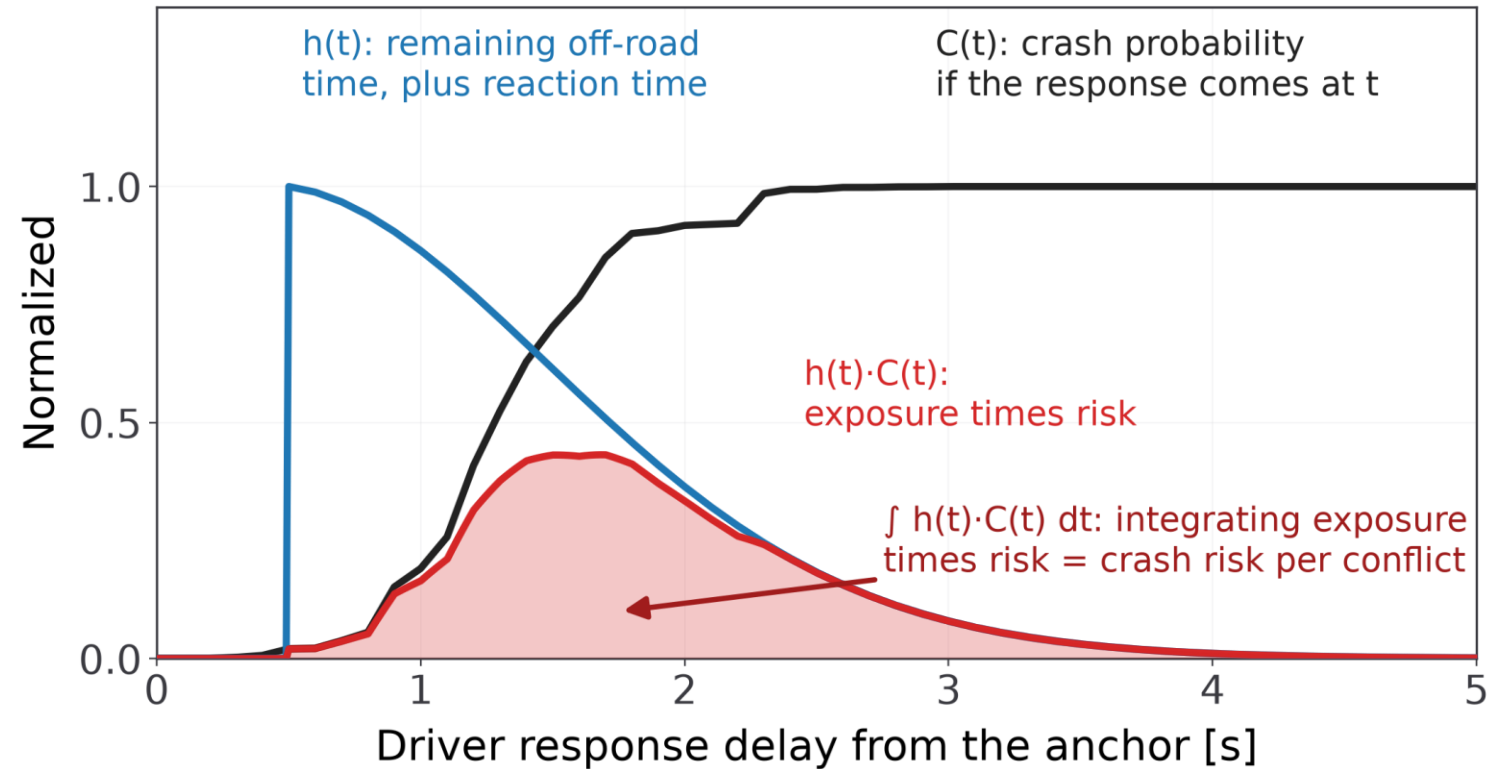
Crash risk
70 of 100 = 0.70

Mean impact speed
8.1 m/s over all 100
11.6 m/s over the 70
crashes

One crash, but 100 what-ifs: its potential severity

Method core: Replay real conflicts with different glances → a “what if” on glance behavior

- Real pre-crash rear-end kinematics; the driver's avoidance maneuver removed
- Glance behavior from studies (naturalistic or experiments) applied
- Risk = where off-road time "meets" rising crash probability (dose-response)
- % on-road glances included as zero second glances
- **Output: crash and injury risk**

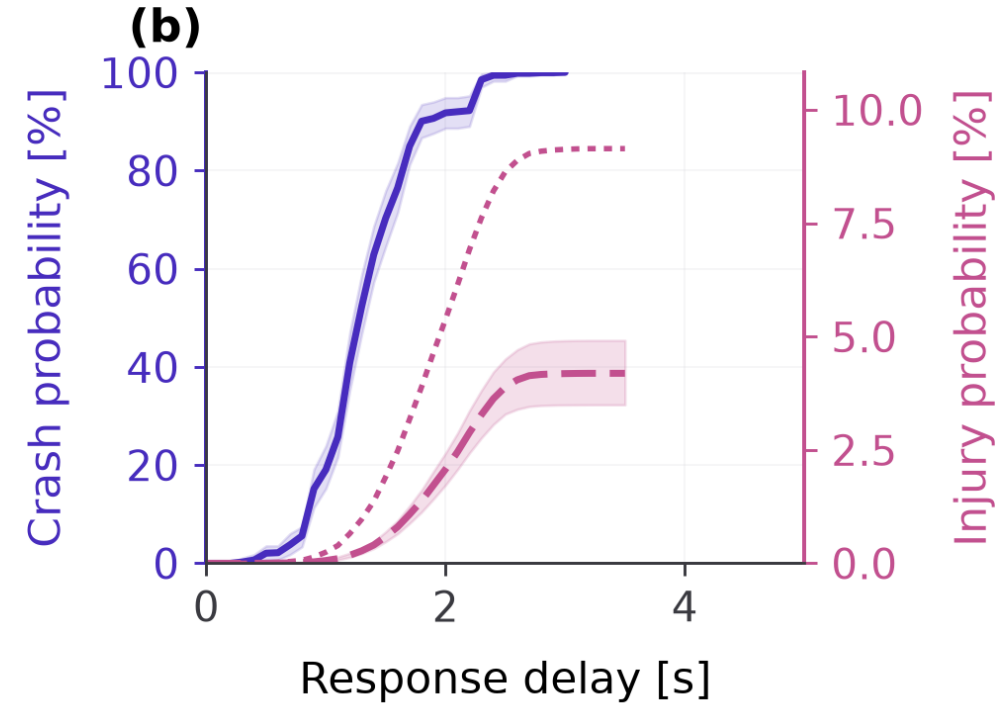
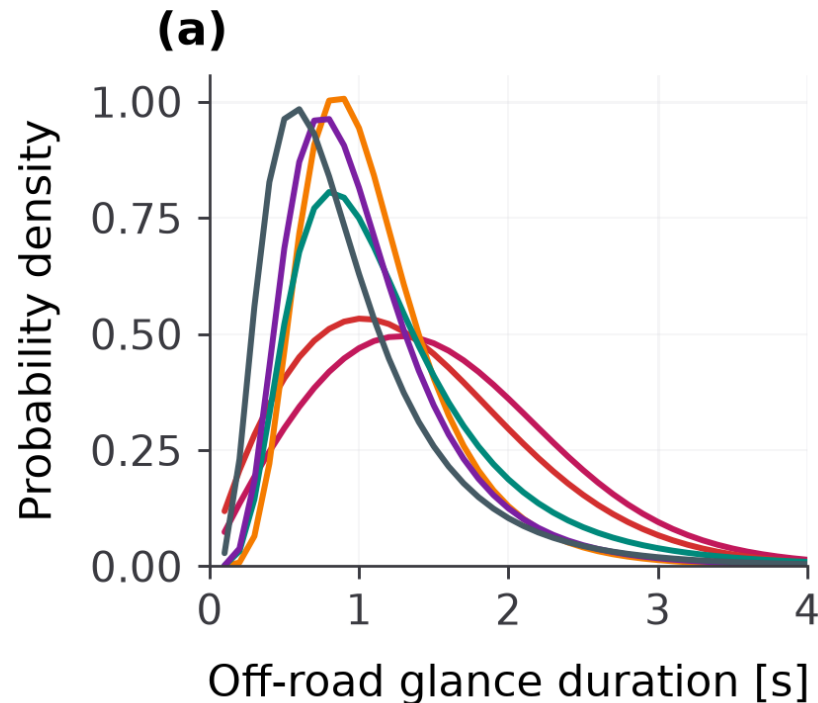


Anchor: the moment the conflict becomes critical, when the inverse time to collision reaches 0.2 s^{-1} (the lead vehicle grows visibly; from Markkula et al. 2016). All response delays are counted from that moment.

Glance behavior and crash/injury risk probabilities

$$\text{Crash risk} = \int h(t) \cdot C(t) dt$$

$$\text{Injury risk} = \int h(t) \cdot I(t) dt$$



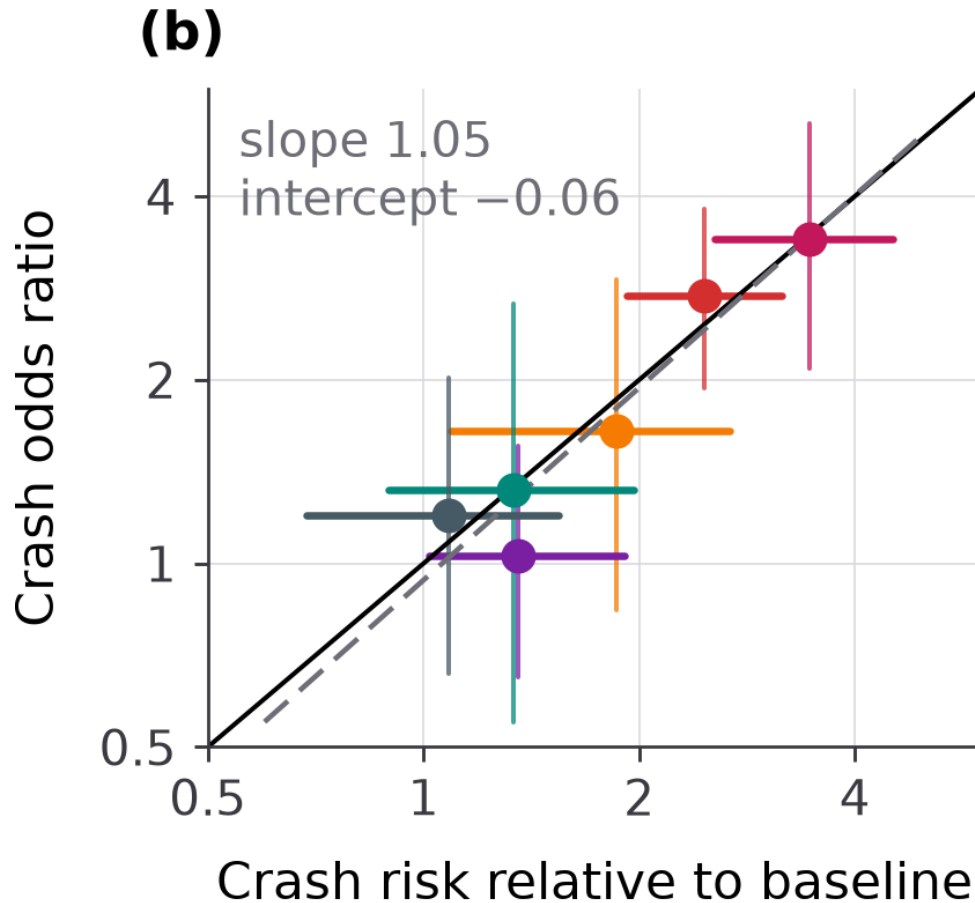
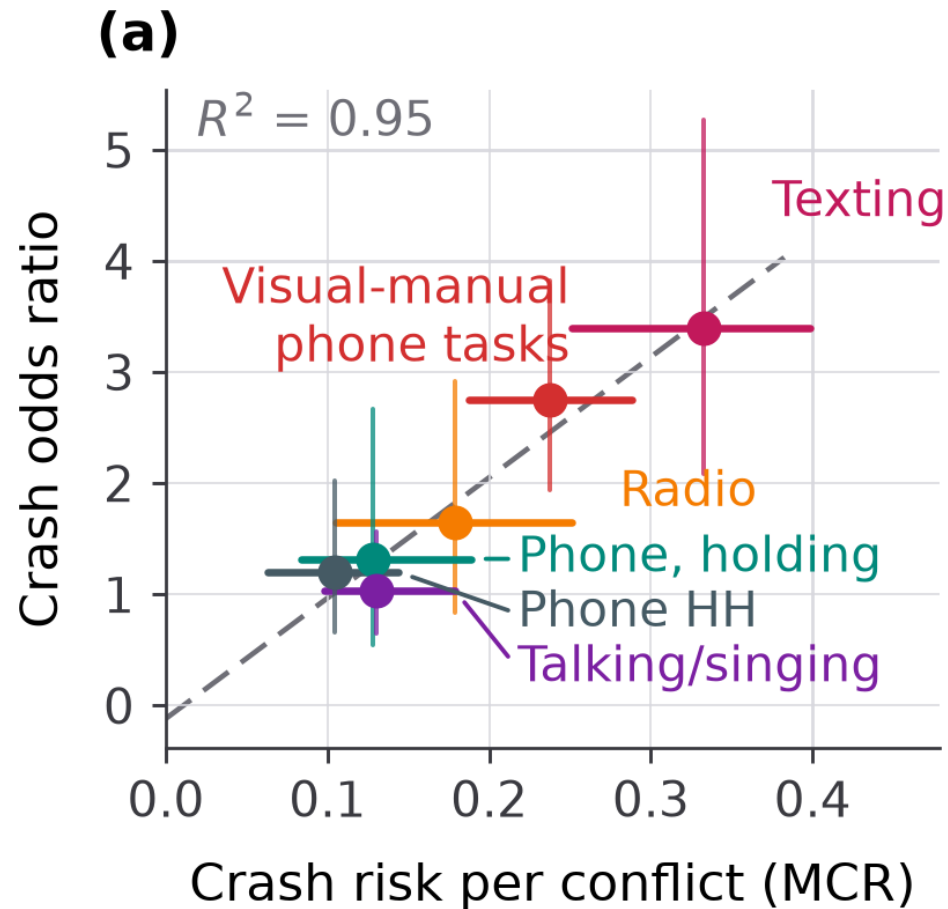
- | | | |
|-----------------------|-------------------|---------------------------------|
| — Texting | — Phone, holding | — Crash, $C(t)$ |
| — Visual-manual phone | — Talking/singing | — — — — — MAIS3+ injury, $I(t)$ |
| — Radio (few cases) | — Phone HH | MAIS2+ injury |

Left: off-road glance durations of six task groups (SHRP2; parametric distribution fit).

Right: crash and MAIS3+ injury probability (dotted: MAIS2+) when braking starts t seconds late

Method validation: The simulated risks agree with real-world crash odds ratios

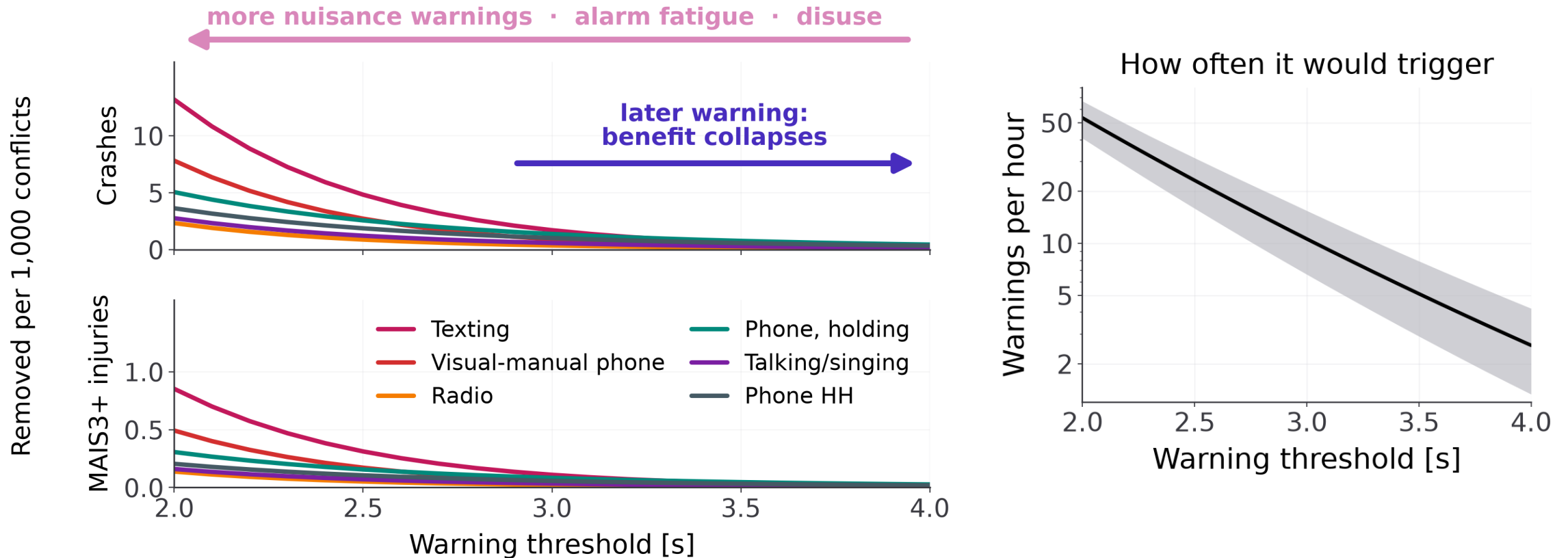
Main results!



Left: simulated risk against crash odds ratios (SHRP2, task alone), $R^2 = 0.95$.

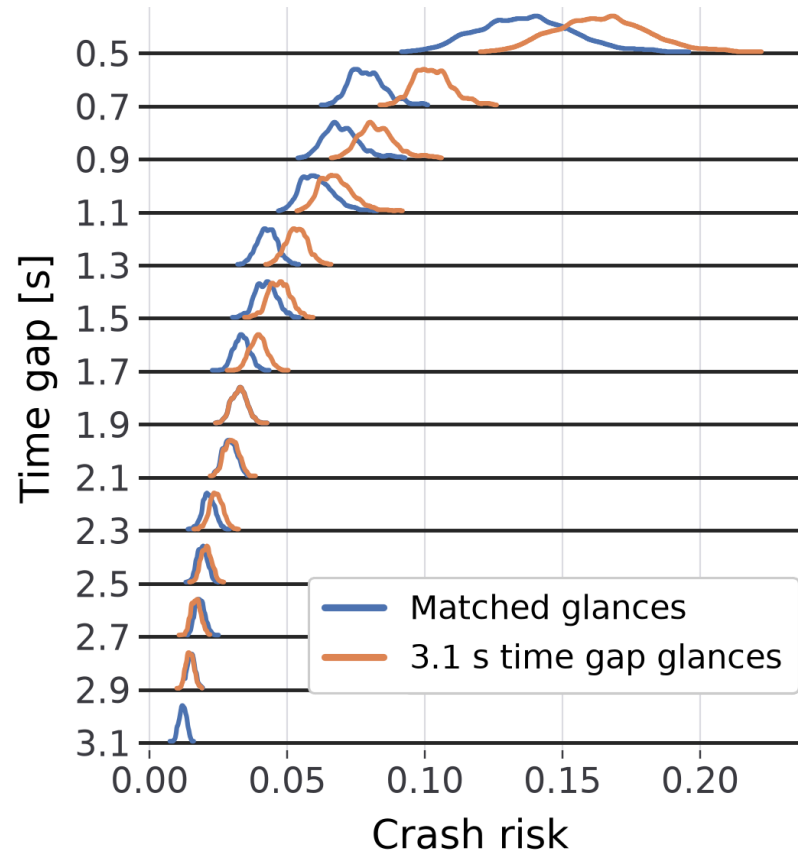
Right: relative to baseline driving, nothing fitted: slope 1.05, the identity line not rejected ($p = 0.96$)

Method predicts that a warning removes few crashes and injuries → but an ADDW that has a real impact triggers often

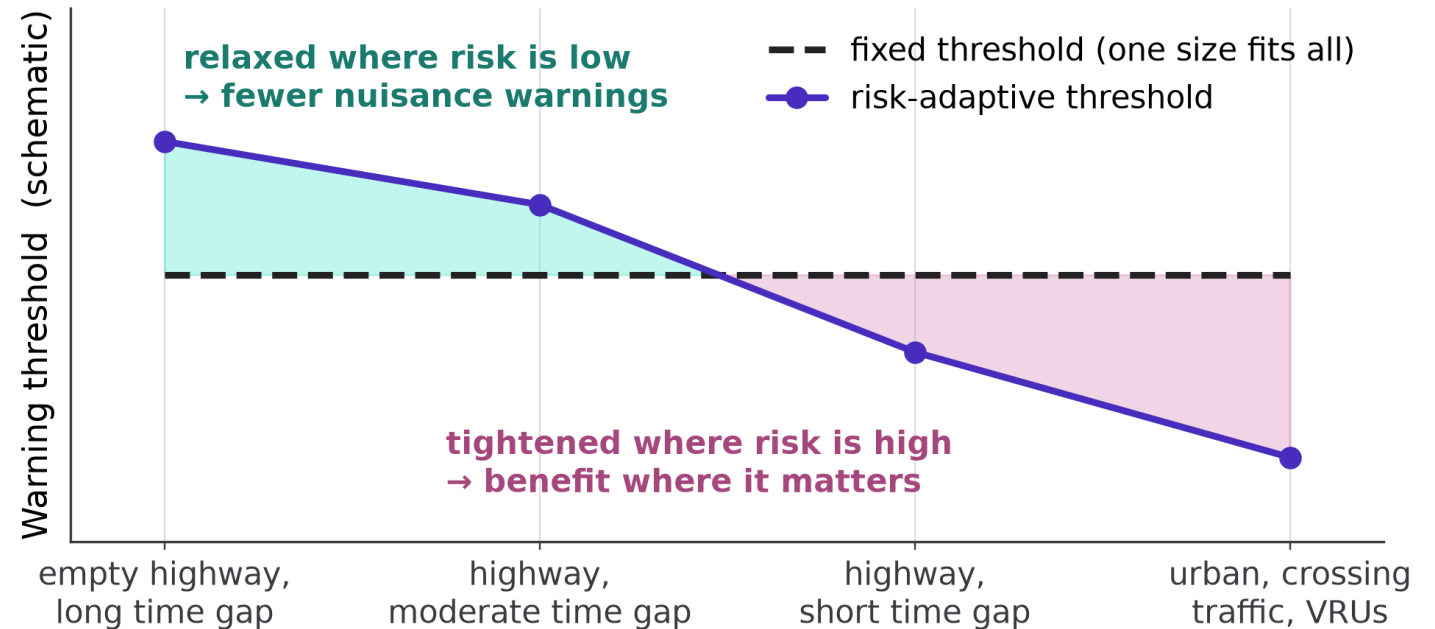


At 2 s: 2-13 crashes and 0.1-0.9 MAIS3+ injuries removed per 1,000 conflicts.
Estimate it would trigger up to 53 times an hour at 2 s and 5.1 at 3.5 s

... and it depends on the context



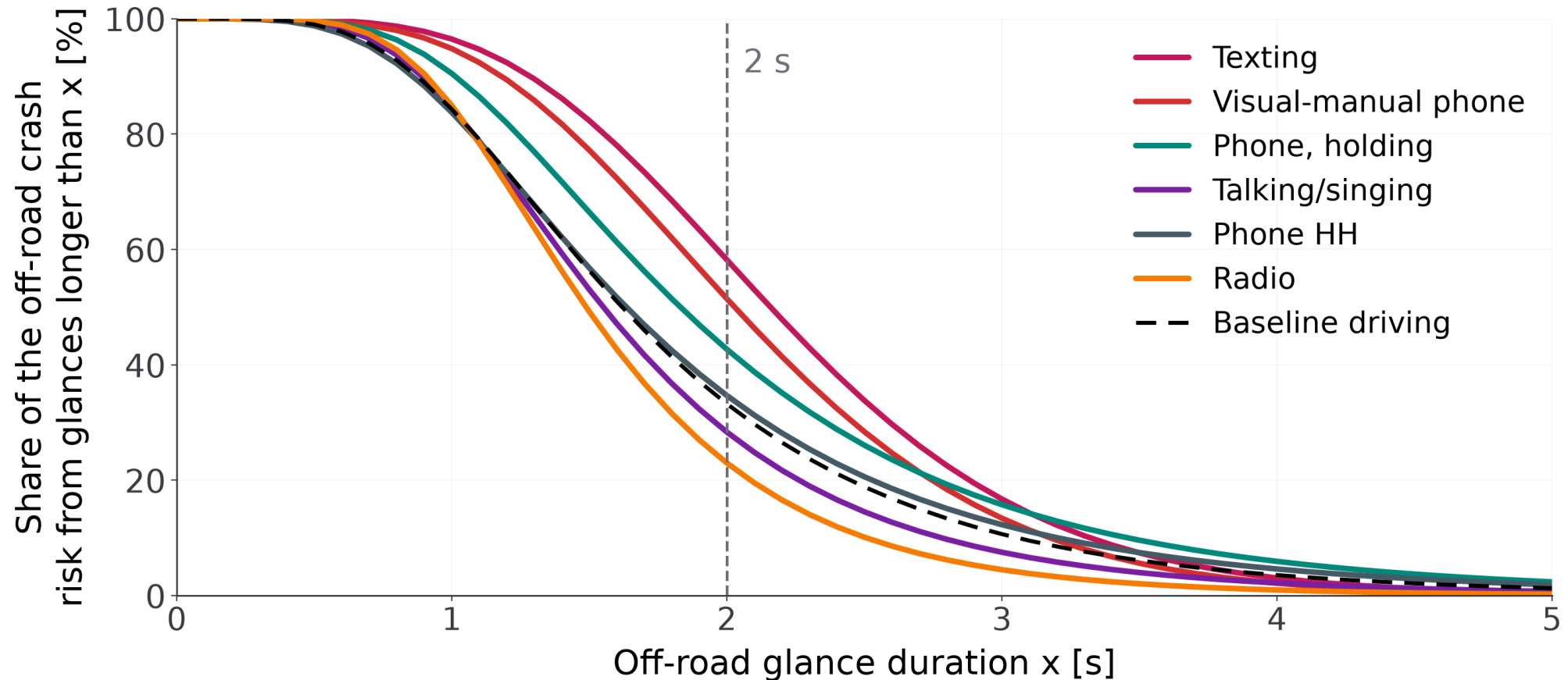
Same glances (orange), $\times 10$ the crash risk at short time gaps, and drivers barely adapt (blue)



- **Design consequence: thresholds that follow the risk, relaxed where little is at stake, strict where much is.**
- **In high-interaction contexts attention means scanning, not eyes locked forward → needs method improvement**

Olleja, Svärd, Zhao & Bårgman (2026),
doi:10.5281/zenodo.19322728

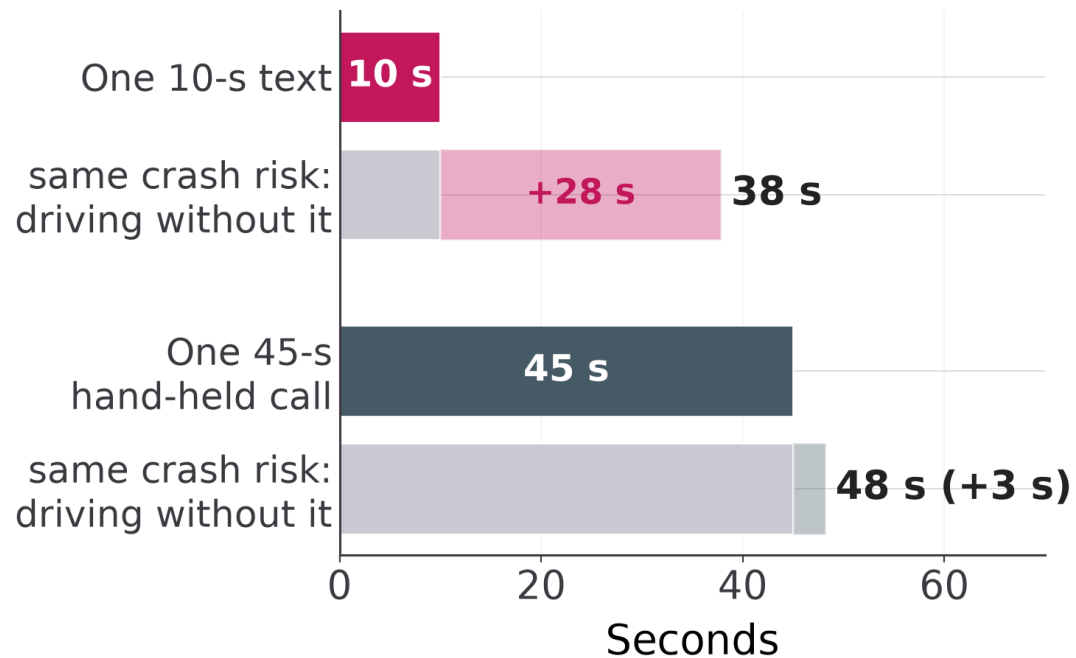
Where is the risk? The risk tail starts earlier than 2 s



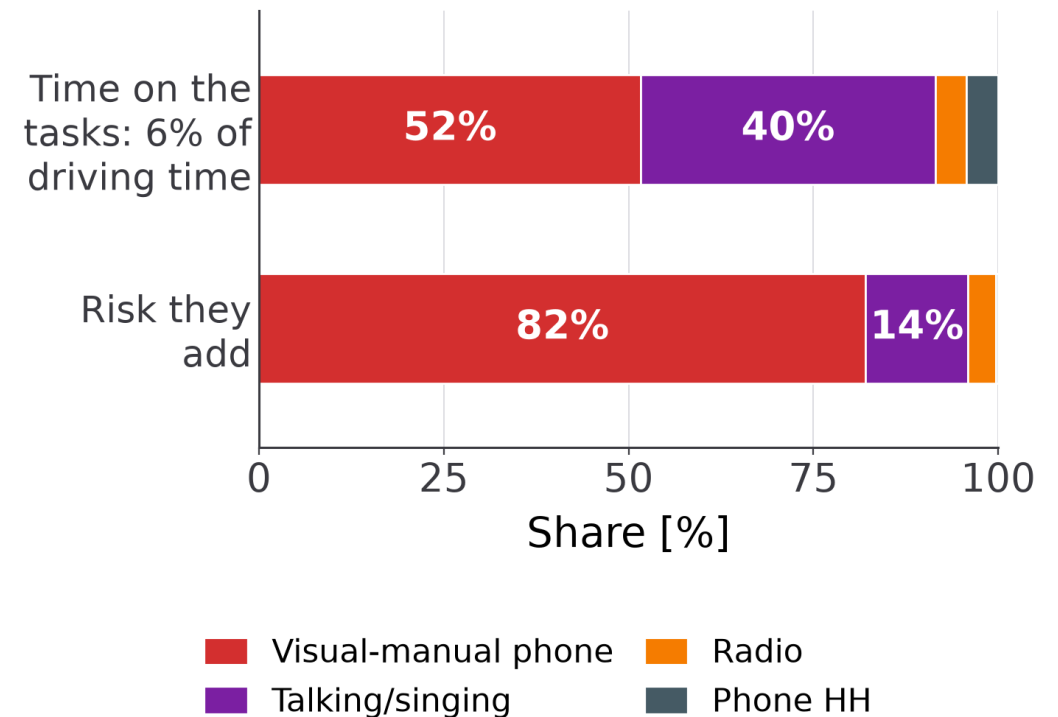
Glances longer than 2 s are 20% of visual-manual phone glances and carry 51% of its crash risk; for Phone HH, 6% of glances and 35% of the risk

Exposure: how long and how often a task is done changes the answer

Per engagement

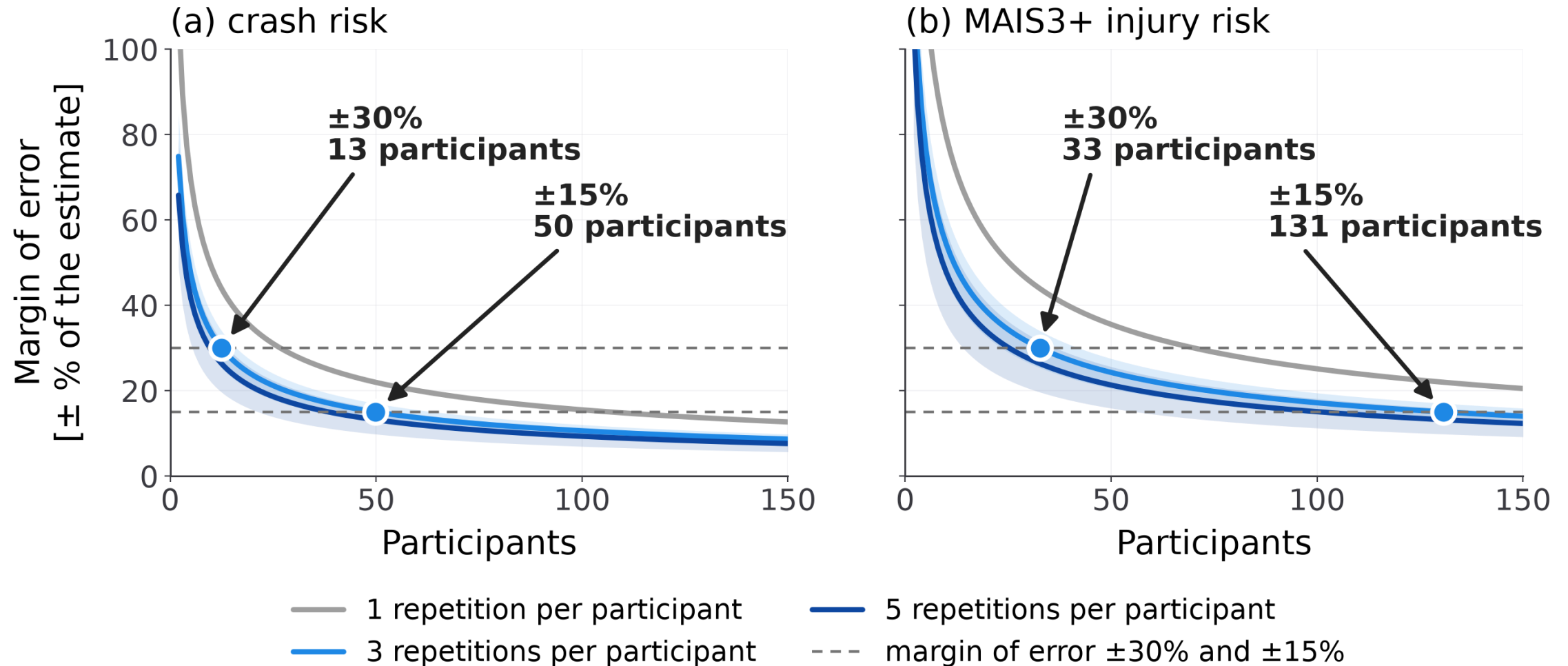


Per hour of driving (UDRIVE)



What a task adds over driving without it: one 10-s text as much as 28 s of driving, one 45-s hand-held call 3 s; per hour (UDRIVE), visual-manual phone use adds most

How much glance data is needed for assessment?



Main limitations

- Only rear-end and off-road glances considered
 - but still agrees with the *overall* crash odds ratios (all crash types)
 - glances do not adapt to the situation: absolute rates too high, the ordering holds
- For proper assessment, exposure is needed: total task time and frequency
 - not in SHRP2's six-second task windows, so not part of the validation
 - enters as the risk a task adds over driving without it

Three take-aways

- **The simulated risks reproduce real-world crash odds:** in rank and, relative to baseline driving, in size
 - Absolute levels depend on assumptions, but the ordering is stable → 10 drivers per task differentiate a 1.4-fold difference (86% of cases)
 - ADDW warnings need trade-offs: even at 2s they remove little crash risk and trigger often, not much better at 3.5s → context-adaptivity likely the way to go
- Paper soon to be submitted and code will be open source

References (1 of 2)

- Bärgman, J., Lisovskaja, V., Victor, T., Flannagan, C., & Dozza, M. (2015). How does glance behavior influence crash and injury risk? A 'what-if' counterfactual simulation using crashes and near-crashes from SHRP2. *Transportation Research Part F: Traffic Psychology and Behaviour*, 35, 152–169. <https://doi.org/10.1016/j.trf.2015.10.011>
- Bärgman, J., Svärd, M., Lundell, S., & Hartelius, E. (2024). Methodological challenges of scenario generation validation: A rear-end crash-causation model for virtual safety assessment. *Transportation Research Part F: Traffic Psychology and Behaviour*, 104, 374–410. <https://doi.org/10.1016/j.trf.2024.04.007>
- Carsten, O., Hibberd, D., Bärgman, J., Kovaceva, J., Pereira Cocron, M. S., Dotzauer, M., Utesch, F., Zhang, M., Stemmler, E., Guyonvarch, L., Sagberg, F., & Forcolin, F. (2017). *Driver distraction and inattention* (UDRIVE Deliverable D43.1). UDRIVE Consortium. https://doi.org/10.26323/UDRIVE_D43.1
- Commission Delegated Regulation (EU) 2023/2590 of 13 July 2023 supplementing Regulation (EU) 2019/2144 of the European Parliament and of the Council by laying down detailed rules concerning the specific test procedures and technical requirements for the type-approval of certain motor vehicles with regard to their advanced driver distraction warning systems and amending that Regulation, OJ L, 2023/2590 (2023). https://eur-lex.europa.eu/eli/reg_del/2023/2590/oj
- Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences*, 113(10), 2636–2641. <https://doi.org/10.1073/pnas.1513271113>
- Euro NCAP. (2025). *Safe driving: Driver task engagement protocol* (Version 1.1). https://cdn.euroncap.com/cars/assets/euro_ncap_protocol_safe_driving_driver_engagement_v11_a30e874152.pdf
- Kircher, K., & Ahlström, C. (2017). Minimum required attention: A human-centered approach to driver inattention. *Human Factors*, 59(3), 471–484. <https://doi.org/10.1177/0018720816672756>
- Krajewski, R., Bock, J., Kloeker, L., & Eckstein, L. (2018). The highD dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)* (pp. 2118–2125). IEEE. <https://doi.org/10.1109/ITSC.2018.8569552>
- Kujala, T., & Sarkar, A. (2025). Evaluating in-car tasks' distraction effects with Drive-In Lab. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. ACM. <https://doi.org/10.1145/3706598.3713590>
- Markkula, G., Engström, J., Lodin, J., Bärgman, J., & Victor, T. (2016). A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies. *Accident Analysis & Prevention*, 95, 209–226. <https://doi.org/10.1016/j.aap.2016.07.007>

References (2 of 2)

- National Highway Traffic Safety Administration. (2013). Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices. *Federal Register*, 78(81), 24818–24890. <https://www.federalregister.gov/documents/2013/04/26/2013-09883/visual-manual-nhtsa-driver-distraction-guidelines-for-in-vehicle-electronic-devices>
- Olleja, P., Svärd, M., Zhao, M., & Bårgman, J. (2026). *Analysis of the safety impact of off-road glances in car-following driving scenarios* [Preprint]. Zenodo. <https://doi.org/10.5281/zenodo.19322728>
- Reimer, B., Angell, L., Noble, A. M., Mehler, B., Skrypchuk, L., Feit, S., & Fitch, G. M. (2026). From distraction to support: Evolving perspectives on driver attention support and the case for a holistic, context-aware framework. *Transportation Research Part F: Psychology and Behaviour*, 121, Article 103658. <https://doi.org/10.1016/j.trf.2026.103658>
- Svärd, M., Bårgman, J., & Victor, T. (2021). Detection and response to critical lead vehicle deceleration events with peripheral vision: Glance response times are independent of visual eccentricity. *Accident Analysis & Prevention*, 150, Article 105853. <https://doi.org/10.1016/j.aap.2020.105853>
- Svärd, M., Olleja, P., & Bårgman, J. (2026). *Drivers' glance-behavior adaptation in response to looming during car following* [Preprint]. Zenodo. <https://doi.org/10.5281/zenodo.19259693>
- Victor, T., Dozza, M., Bårgman, J., Boda, C.-N., Engström, J., Flannagan, C., Lee, J. D., & Markkula, G. (2015). *Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk* (SHRP 2 Report S2-S08A-RW-1). Transportation Research Board. <https://doi.org/10.17226/22297>
- Wu, J., Flannagan, C., Sander, U., & Bårgman, J. (2024). Modeling lead-vehicle kinematics for rear-end crash scenario generation. *IEEE Transactions on Intelligent Transportation Systems*, 25(9), 10866–10884. <https://doi.org/10.1109/TITS.2024.3369097>
- Wu, J., Flannagan, C., Sander, U., & Bårgman, J. (2025). Model-based generation of representative rear-end crash scenarios across the full severity range using pre-crash data. *IEEE Transactions on Intelligent Transportation Systems*, 26(10), 15932–15950. <https://doi.org/10.1109/TITS.2025.3573386>
- Zhao, M., Bårgman, J., & Flannagan, C. (2026, October 20–22). *Conversational cognitive load meets lateral hazard exposure: How environment shapes crash risk* [Extended abstract]. 10th International Conference on Driver Distraction and Inattention (DDI2026), Gothenburg, Sweden.

Thank you

jonas.bargman@chalmers.se

