# Requirements for Monitoring Inattention of the Responsible Human in an Autonomous Vehicle: The Recall and Precision Tradeoff

Keywords: *automation, self-driving, engagement, monitoring, requirements, engineering, complacency, AI*

6/23/2020

Authors: Johnathan DiMatteo, Daniel M. Berry, Krzysztof Czarnecki

UNIVERSITY OF WATERLOO

# Vocabulary

AV = Autonomous Vehicle

RHV = Responsible Human in Vehicle

recall = *R* = percentage right stuff that is found

precision = *P* = percentage of found stuff that is right

TP, TN, FP, FN = true/false positive/negative

# Motivation - Uber Crash, March 2018


**Left:** location of the crash, showing paths of pedestrian in orange and the Uber vehicle in green.
**Right:** postcrash view of the Uber vehicle.
**Source:** NTSB Preliminary Report HWY18MH010

# Assessment - Uber AV Crash

Despite the AV's software seeing the pedestrian 5.6 seconds in advance, it failed to recognize or predict the path of the pedestrian, and the RHV was not paying attention to the driving.

"Ineffective oversight of the vehicle operators and a lack of adequate mechanisms for addressing operators' automation complacency" (National Transportation Safety Board, HWY18MH010).

Therefore, we believe there should be a system to monitor the driver for signs of inattention in every AV.

# RHV Monitor and Notifier (RMN)

1. the **Monitor**, an AI that somehow monitors the RHV for signs of inattention, and at any time that the Monitor detects that the RHV is inattentive, it informs the Notifier to do its job.

2. the **Notifier**, when informed by the Monitor, somehow notifies the AV, the RHV, or both, that signs of inattention have been detected in the RHV.

# Tradeoffs in the RMN

1. the **Monitor**, monitoring the RHV for signs of inattention:

*R* trades with *P*, i.e., > TPs $\Leftrightarrow$ > FPs; $>R \Longrightarrow <P$, $>P \Longrightarrow <R$:

- fewer failures to detect inattention $\Longrightarrow$ more notifications
- fewer notifications $\Longrightarrow$ more failures to detect inattention $\Longrightarrow$ more deaths

2. the **Notifier**, notifying the AV, the RHV, or both:

The more a human is notified, the more he/she begins to ignore it:

- more notifications $\Longrightarrow$ less effectiveness

# Optimizing and Evaluating the RMN

Too many FPs in Monitor: degradation of Notifier's effectiveness

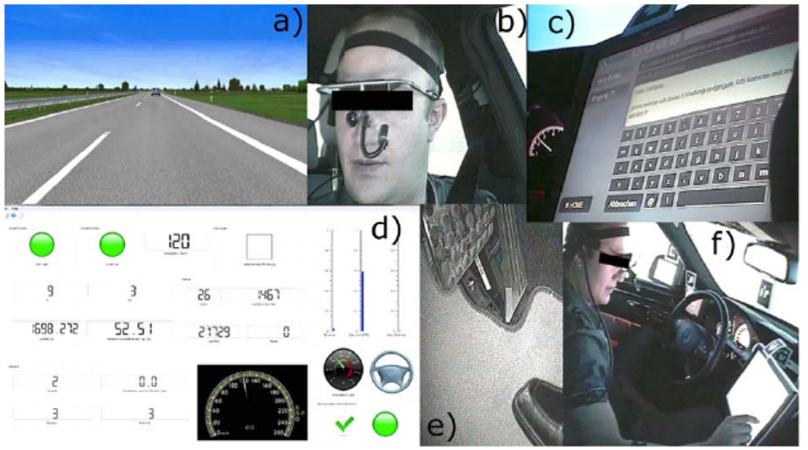
Too many FNs in Monitor: putting driver's and others' lives at risk

So, do we optimize *R* or optimize *P* in the Monitor?

Not clear!

# What we've seen in literature

All 13 items in the literature known to the authors about monitoring algorithms, manage the tradeoff by assuming that FNs and FPs are equally bad.




(Braunagel et al., 2015).

Is this the correct tradeoff?

Let's see what Aviation has learned about notification.

# Aircraft Pilots

Pilots deal with overwhelming notifications and the boring role of supervising automation too.

The FAA and NASA came up with the idea of Human-Centered Automation (HCA) in 1991.

**A few principles of HCA relevant to our discussion:** (Billings, 1991)

1. The human operator must be in command.
2. To command effectively, the human operator must be involved.
3. The automated systems must also be able to monitor the human operator.
4. Each element of the system must have knowledge of the others' intent.

# Human-Centered Automation (HCA)

The FAA took these principles and decided to put the pilot at the ultimate command to supervise the system. To increase a pilot's attentiveness:

- Do puzzles.
- Talk to co-pilots.
- Read training manuals.
- **Decrease automation:** *If during autonomous operation, the vehicle needs assistance that can best be rendered by humans, the human pilot should be called on, even in a non-emergency, if for no other reason than to keep the human pilot engaged.*

So , how do we apply this thinking to AVs and their RHVs?

# Applying HCA to the Notifier

We propose a reduction of automation as a way to keep the RHV engaged, and therefore attentive, gracefully passing responsibility to the RHV.

More specifically, the Notifier will:

- inform the driver about a specific upcoming reduction in automation and
- require some form of acknowledgement from the RHV, before it actually does the reduction (so that the RHV is not dangerously surprised at what is happening).

# Applying HCA to the Monitor

If the effectiveness of such a Notifier can be shown not to degrade with repeated notification, then …

FPs in the Monitor are not so damaging, and …

we can trade lower *P* to achieve higher *R.*

If we have a Notifier whose effectiveness does *not* degrade with repeated notifications, the Monitor should **prioritize *R***, since **FPs just result in the RHV's taking more control of the AV**.

# Conclusion

- An RMN is most effective when
  - its Monitor has 100% recall, and is thus detecting all instances of RHV inattention and
  - the effectiveness of its Notifier's notifications do not degrade when they are repeated.
- The assumption in the literature seems to be that FPs and FNs are equally bad and that *R* and *P* should be weighted equally. However, this assumption may not be true in some circumstances.

# Future Work

- There is a need for future work in experimental testing of high-recall Monitors and low-degradation Notifiers for use in high-effectiveness RMNs for AVs.
- Invent notification techniques that do not degrade with repeated notifications, so that we can reduce automation levels when appropriate and have high-recall Monitors.

Now, go read the paper ☺!

# Sources


D. Berry. Evaluation of tools for hairy requirements and software engineering tasks. In Proceedings of Workshop on Empirical Requirements Engineering (EmpirRE) in IEEE 25th International Requirements Engineering Conference Workshops, pages 284–291, 2017.

H. Bhana. Correlating Boredom Proneness With Automation Complacency in Modern Airline Pilots. PhD thesis, University of North Dakota, Grand Forks, North Dakota, USA, 2009.

C. Billings. Human-centered aircraft automation: A concept and guidelines. Technical Report NASA Technical Memorandum 110381, NASA Ames Research Center, Aug. 1991.

C. Braunagel, E. Kasneci, W. Stolzmann, and W. Rosenstiel. Driver-activity recognition in the context of conditionally autonomous driving. In IEEE 18th International Conference on Intelligent Transportation Systems, pages 1652–1657, Sept. 2015.

J. DiMatteo, D. M. Berry, and K. Czarnecki. Requirements for monitoring inattention of the responsible human in an autonomous vehicle: The recall and precision trade-off. Technical report, University of Waterloo, 2020.

T. Lee. Another Tesla driver apparently fell asleep—here's what Tesla could do. arsTECHNICA, 2019.

National Transportation Safety Board. Highway preliminary report: Hwy18mh010. Technical report, National Transportation Safety Board, May 2018.

On-Road Automated Driving (ORAD) Committee. Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles, revised standard. Technical Report J3016 201806, SAE International, 2018.

T. Saracevic. Evaluation of evaluation in information retrieval. In SIGIR Conf. Res. & Devel. Inform. Retrieval (SIGIR), pages 138–146, 1995.

R. van der Heiden, S. Iqbal, and C. Janssen. Priming drivers before handover in semi-autonomous cars. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI), pages 392–404, 2017.

# Questions