Human Monitoring in Self-Driving Vehicles: An RE Challenge

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

6/20/19

Presented by: Johnathan DiMatteo

CS 846: Requirements Engineering

UNIVERSITY OF
WATERLOO

Outline

1. Motivation
   a. Uber crash
   b. Tesla crash
2. Is monitoring really necessary?
3. The effects of monitoring
   a. Pilots
   b. Driving Simulations
4. Mitigating those effects
5. Conclusion

UNIVERSITY OF
WATERLOO

Problem

────────────────────────

Self-driving vehicles are designed to make driving the responsibility of the software and not the user, making the user less likely to pay attention, causing severe and often fatal accidents.

────────────────────────

Scope

Automation

↓

Boredom, loss of attention, loss of control, overconfidence, etc.

↓

Decreased vigilance, increased complacency, decreased take-over readiness, etc.

↓

Decreased safety

UNIVERSITY OF
WATERLOO

Why is This is Relevant to RE

Introducing a new feature (automation) sometimes has undesirable/unforeseeable effects on the user.

It's hard to think of all necessary safety requirements for system-user interaction.

The role of the user is changing from active to passive.

Motivation - Uber Crash

On a dark night in March, 2018, an Uber Technologies, Inc. test vehicle in autonomous driving mode, struck and killed a pedestrian crossing the street.

A preliminary NTSB report on the crash revealed the backup driver who was responsible for taking control of the vehicle in times of emergency did not have her hands on the steering wheel (as required) and was looking downwards moments before the crash, unable to engage the emergency brakes.

UNIVERSITY OF
WATERLOO

Motivation - Uber Crash

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

UNIVERSITY OF
WATERLOO

Motivation - Tesla Crash

In March 2019, a Tesla Model 3 driving in **AUTOPILOT** struck a truck hauling a semi-trailer. The top half of the Tesla was sheared off, and the driver of the Tesla died as a result of the crash.

Again, an NTSB report indicated the driver did not have his hands on the steering wheel (as required, despite the name).

What do both these crashes have in common?

UNIVERSITY OF
WATERLOO

Motivation - Tesla Crash

Image: post-crash view of the Tesla Model 3.
Source: NSTB preliminary Report HWY19FH008

UNIVERSITY OF
WATERLOO

But … is it really necessary for users to monitor the system?

Is Monitoring Necessary?

Tesla Statement after incident:

"Our data shows that, when used properly with an attentive driver who is prepared to take over at all times, drivers supported by Autopilot are safer than those operating without assistance"

(Yet Elon Musk says a few months earlier: "Very quickly, ... having a human intervene will decrease safety")

Is Monitoring Necessary?

Automation systems are not foolproof.

E.g., RL can game the system (avoid being penalized for getting close to other cars by learning where the sensors are) or negatively affect environment to achieve it's own goals.

Other factors:

- sensors/hardware can degrade
- software malfunctions (bit flips due to radiation…)
- security vulnerabilities

(Czarnecki et al., 2018)

UNIVERSITY OF
WATERLOO

Is Monitoring Necessary?

Even IF the system is foolproof, the user-system experience still may depend on human inputs that are valid but lead to hazards.

E.g., Korean Air Lines Flight 007. The official report attributed the crew's "lack of alertness" as the most plausible cause of the navigational error.

UNIVERSITY OF
WATERLOO
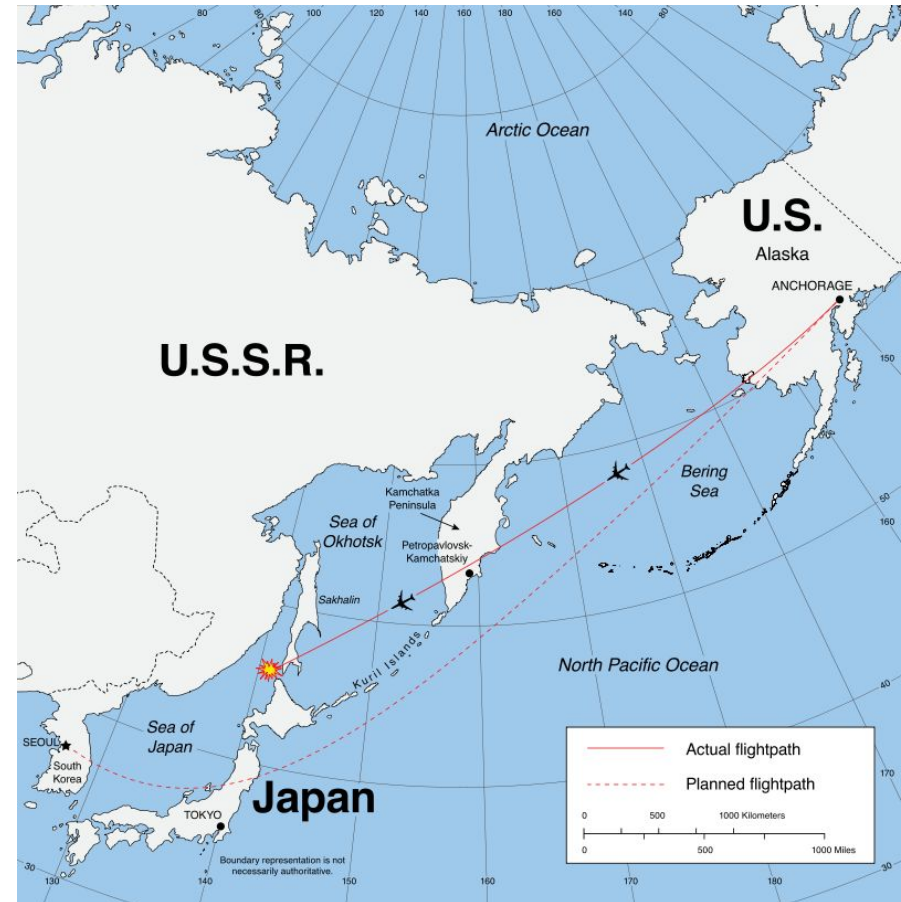
Is Monitoring Necessary?

If is it true that the fatality rate is lower in an autonomous vehicle than it is with humans alone, then why can't we be happy with a few incidents here and there?

In safety-critical domains, like airplanes and cars, we have to do better.

A N.T.S.B. review of thirty-seven major airplane accidents between 1978 and 1990 found that in thirty-one cases faulty or inadequate monitoring were partly to blame.

UNIVERSITY OF
WATERLOO

Is Monitoring Necessary?

To deal with these monitoring issues, pilots decided upon **human-centered automation**, and concluded "the quality and effectiveness of the pilot-automation system is a function to the degree of which the combined system takes advantages of the strengths and compensates for the weaknesses of both elements" (Billings, 1991)

UNIVERSITY OF
WATERLOO

effective automation + vigilant human supervisor = safest system

What are the effects of monitoring automation for long-periods of time?

Effects of Monitoring

You would think if automation rises that people will spend the extra mental resources to look around at traffic or potential hazards, but …

Malleable Attentional Resources Theory (MART):

"Attentional resources shrink to accomodate any demand reduction"

(Young and Stanton, 2002)

UNIVERSITY OF
WATERLOO

Effects of Monitoring

16 pilots were asked to fly in a Boeing 747 simulator, automation levels varied as the flight progressed, anomalies were randomly introduced that forced the pilots to take over.

As automation levels rose, the worse they were at dealing with the anomalies.

(Casner, 2014)

Effects of Monitoring

In a driving simulator study, drivers showed decreased driving performance (increased heading error) on straight road sections but not curved.

Drivers underestimated task demands in the low-workload setting and withdrew necessary focus accordingly.
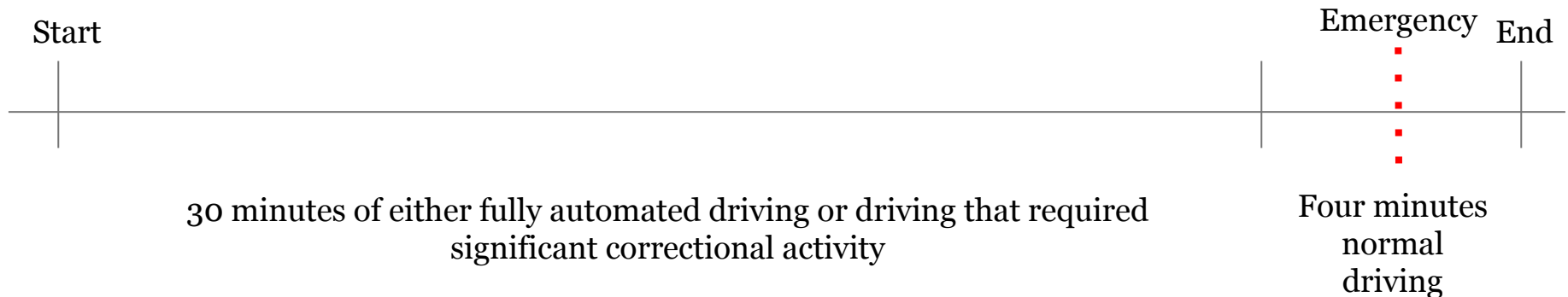
(Matthews and Desmond, 2002)

UNIVERSITY OF
WATERLOO

Effects of Monitoring

2013 study had 168 participants drive for 30 minutes in a simulation either:

- Fully automated, or
- Wind gusts required correctional steering of vehicle

Then they were told to drive for another four minutes and anticipate an emergency event.

Start   Emergency End

30 minutes of either fully automated driving or driving that required significant correctional activity

Four minutes normal driving

(Saxby et al., 2013)

UNIVERSITY OF
WATERLOO

Effects of Monitoring

Drivers who previously had the automated driving experience had slowest steering and braking response to the event, and most likely to crash.

``the loss of safety ... is the combination of low workload, decreased task engagement and low challenge.''

(Saxby et al., 2013)

UNIVERSITY OF
WATERLOO
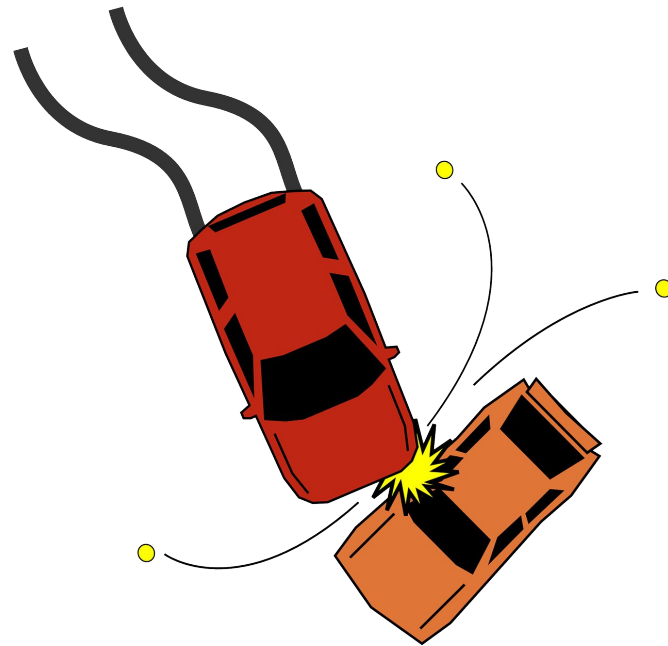
How can we mitigate some of these effects to make it safer for the user?

Mitigating the Effects

ISO 26262 "Road Vehicles: Functional Safety":

Potential hazards including reasonably foreseeable misuse by the operator requires mitigation

So how do we mitigate the effects of inattention, boredom, and passive fatigue caused by supervising automation?

Have we learned anything from pilots?

UNIVERSITY OF
WATERLOO

Mitigating the Effects

Major airline in the US, 301 pilots surveyed: correlation exists between **boredom** and frequency of **attention lapses**.

Pilots who engaged in activities reported lower boredom, and lower self-reported attention lapses:

> admiring the view, doing puzzles, talking to colleagues, paying mental games, fidgeting, looking around, reading training manuals, writing, etc.

"individuals who are better able to relieve boredom through internal sources commit fewer automation complacency errors" (Bhana, 2009)

Mitigating the Effects

Biggest risk for pilots is boredom, and playing sides games, talking to co-pilots, looking at scenery is an effective solution.

These solutions are limited for automobiles on the ground.

Also, cars on the ground need a faster emergency response time (the density of hazards on the ground is much higher than hazards in the air).

UNIVERSITY OF
WATERLOO

Mitigating the Effects

Some papers suggest designing a secondary system to monitor the user, using eye-tracking and head-tracking to determine driver's activity (read, write an email, watch a movie, idle)
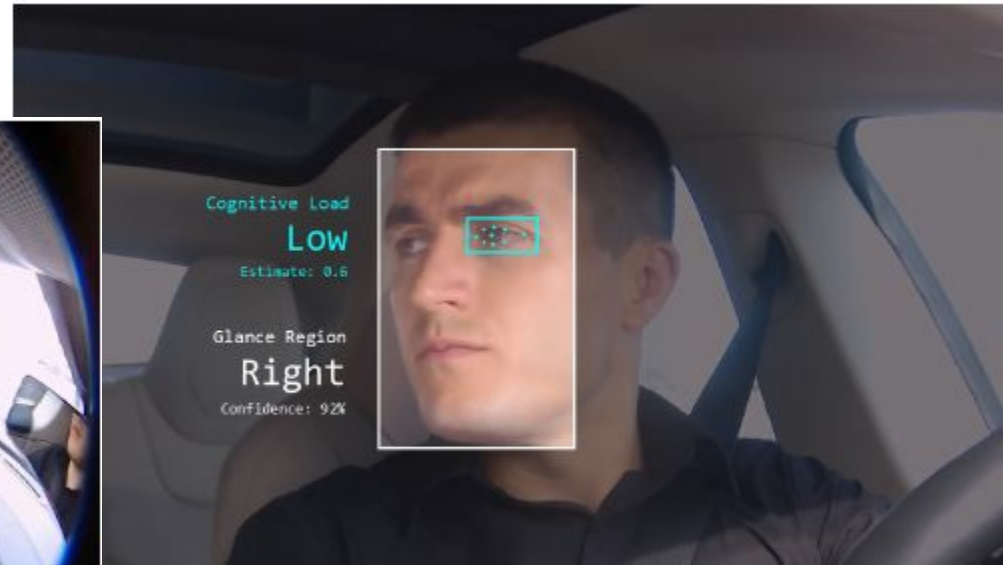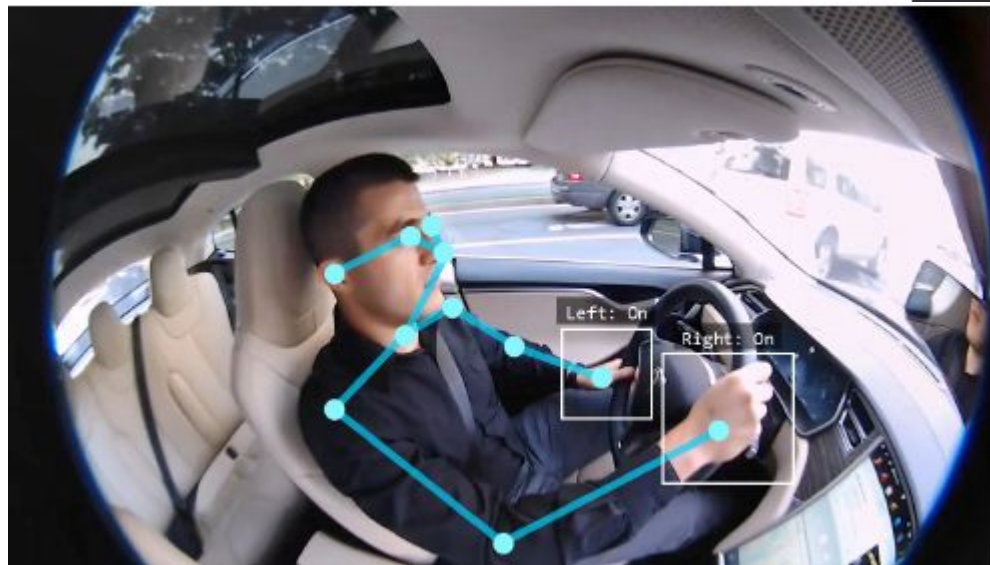
Driving simulator study on 73 participants achieved an average of 70% precision and 76% recall on activity classification.
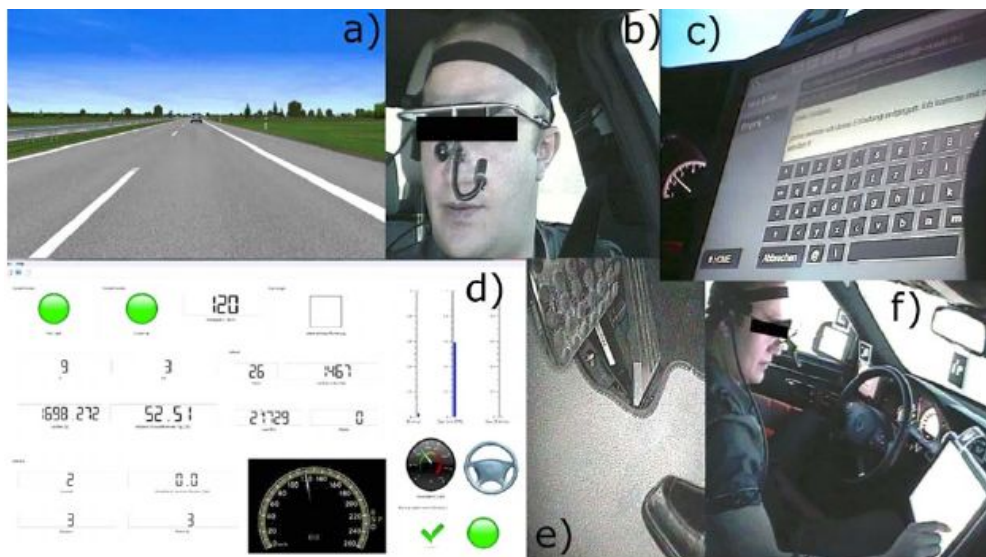
(Braunagel et al., 2015)
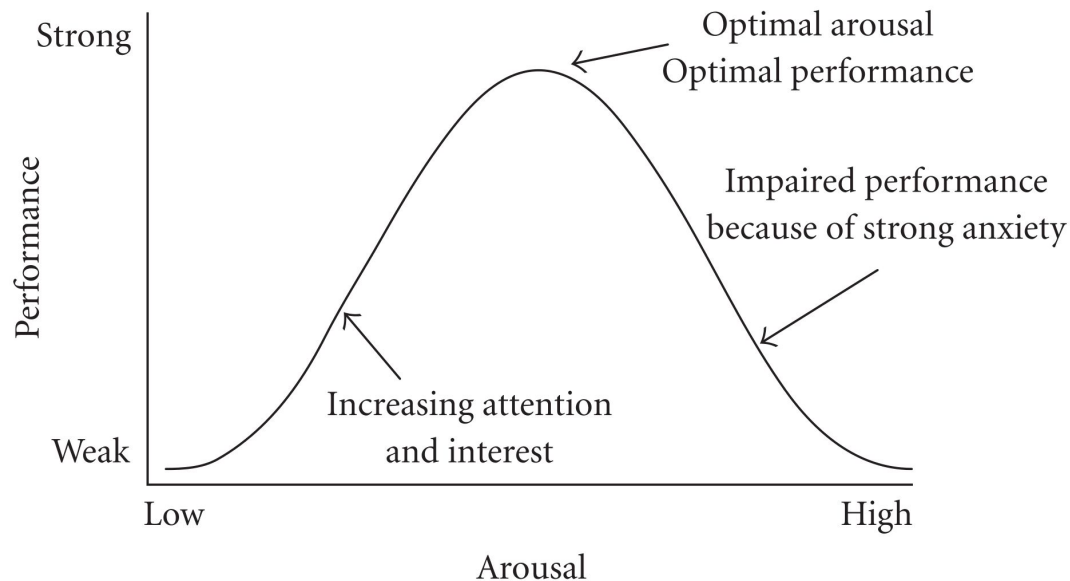
UNIVERSITY OF
WATERLOO

Mitigating the Effects

The human-centered AI research group at MIT has experimented with body/head posture and eye-tracking surveillance (Fridman, 2017).

UNIVERSITY OF
WATERLOO

Mitigating the Effects

System should be providing feedback: Haptic stimulation, visual cues, audio, etc.

Yerkes-Dodson law

Conclusion

- Human monitoring of autonomous systems is necessary and important to fulfill the requirements of ISO 26262.

- Introducing a new feature (automation) sometimes has undesirable/unforeseeable effects on the user.

- An autonomous car should be aware of the user's state, and provide appropriate feedback when necessary (Yerkes-Dodson law).

- Eye-tracking, head movements, biometrics are good features to monitor.

UNIVERSITY OF
WATERLOO

Sources

National Transportation Safety Board. A review of flight crew involved major accidents of U.S. air carriers, 1978 through 1990. Technical Report NTSB/SS-94/01 Notation 6241, Jan 1994.

National Transportation Safety Board.   Preliminary report.   Technical Report HWY18MH010, May 2018.

Charles E. Billings. Human-centered aircraft automation: A concept and guidelines. Technical report, August 1991.

Rick Salay, Rodrigo Queiroz, and Krzysztof Czarnecki.   An Analysis of ISO 26262: Machine Learning and Safety in Automotive Software. Pages 2018–01–1075, April 2018

Mark S Young and Neville A Stanton. Malleable attentional resources theory: a new explanation for the effects of mental underload on performance. Human factors, 44(3):365–375, 2002

Gerald Matthews and Paula A Desmond. Task-induced fatigue states and simulated driving performance. The Quarterly Journal of Experimental Psychology: Section A, 55(2):659–686, 2002

Dyani J Saxby, Gerald Matthews, Joel S Warm, Edward M Hitchcock,and Catherine Neubauer. Active and passive fatigue in simulated driving: discriminating styles of workload regulation and their safety impacts. Journal of experimental psychology: applied, 19(4):287, 2013

Hemant Bhana. Correlating Boredom Proneness With Automation Complacency in Modern Airline Pilots. page 93, 2009.

Christian Braunagel, Enkelejda Kasneci, Wolfgang Stolzmann, and Wolfgang Rosenstiel.   Driver-Activity Recognition in the Context of Conditionally Autonomous Driving.   In 2015 IEEE 18th International Conference on Intelligent Transportation Systems, pages 1652–1657,Gran Canaria, Spain, September 2015. IEEE.

Lex Fridman, Daniel E Brown, Michael Glazer, William Angell, Spencer Dodd, Benedikt Jenik, Jack Terwilliger, Julia Kindelsberger, Li Ding,Sean Seaman, et al. MIT autonomous vehicle technology study: Large-scale deep learning based analysis of driver behavior and interaction with automation. arXiv preprint arXiv:1711.06976, 2017.

UNIVERSITY OF
WATERLOO