Trusting Automation – Implications for Autonomous Vehicle Design

Andrew Morris
Transport Safety Research Group
Loughborough University
UK

September 2017

What is Trust?

1.reliance on the integrity, strength, ability, surety, etc., of a person or thing; confidence.

2.confident expectation of something; hope

Trust in Automation

- Significant research devoted to the issue of trust
- Complex multi-dimensional construct making it difficult to operationalise, measure and interpret (Simpson, 2007)
- Trust as a concept is so important, it lays the foundations for happy and successful like (Erikson, 1963)
- However, trust in Automation is influenced by a number of factors not seen in interpersonal trust (Hancock, 2011)
- Trust in automation is (conceivably)a new research discipline examining 'Cognitive Agents' (Apple's "Siri"; Amazon's "Alexa"

Trust in Automation

- A significant factor affecting trust is level of feedback received by human from the 'machine'
- Norman (1990) – the chances of failure are high in cases of inappropriate design surrounding automation – inadequate feedback leads to error
- Walker, Stanton and Young (2006) – "vehicle feedback plays a key role in how drivers interact with their environment with the role of spoken feedback particularly prominent
- This follows evidence that spoken feedback has a strong effect on how people drive their vehicles (Horswill and McKenna, 1999)

Trust in Automation

- Verberne et al. (2012) - providing feedback can potentially increase the level of trust in the driver
- Bank and Stanton (2016) - delivering a transparent HMI display can enhance driver trust in automation.
- However investigating HMI that will affect trust in autonomous vehicles is difficult - fully-automated driving still a relatively new concept – and difficult to simulate.
- Important to examine how trust in autonomous vehicles changes with the different level of visual and auditory feedback in order to predict effects that may worsen or promote the acceptance of autonomous vehicle.

Study 1 and 2

Aims and Objectives

The aims of this study are as follows:

- To understand how trust in autonomous vehicles changes with different levels of auditory and visual feedback
- To understand in what circumstances, drivers do not understand what the vehicle is doing and why
- To further develop our understanding of how spoken feedback affects a driver's trust in a level 4 autonomous vehicle.
- To compare different levels of feedback and the amount of trust given.
- To compare how distrust and trust change with time.

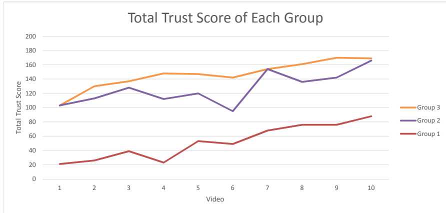
Loughborough University

Study 1 Trust with Auditory Feedback - Method

- **Group 1** received no spoken feedback from the vehicle.
- **Group 2** received moderate spoken feedback – consisting of what the vehicle could "see".
- **Group 3** received a high level of spoken feedback – consisting of what the vehicle could "see" and the action the vehicle would take

Loughborough University

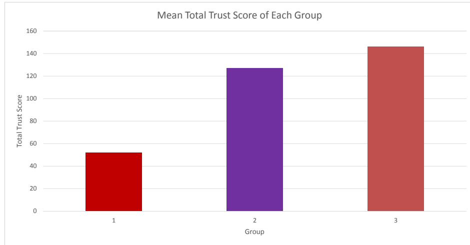
Results



- Group 1 - no spoken feedback from the vehicle.
- Group 2 - moderate spoken feedback – what the vehicle could "see".
- Group 3 - high level of spoken feedback – what vehicle could "see" and what would.

Loughborough University

Mean Trust Score



	GROUP 1	GROUP 2	GROUP 3
MEAN SCORE	51.9	126.9	146.1

Discussion

- Participants who received no spoken feedback were significantly less trusting that those who did: *Group 1 – Group 2 ($t=7.1195$, $d.f=18$, $alpha=0.05$, $p<0.05$, Group 1 – Group 3 ($t=9.4401$, $d.f=18$, $alpha=0.05$, $p<0.0001$)*.
- Conversely, the results also showed that there was not a significant difference in total trust scores between Group 2 and Group 3 *($t=2.0081$, $d.f=18$, $alpha=0.05$, $p=0.06$)*.
- Furthermore, the study also found that people's trust increased in the autonomous vehicle over time, regardless of the amount of feedback.

Discussion

- Provides the framework for more in-depth studies to use a more realistic situation and understand more about how different types of people trust autonomous vehicles.
- To engender trust, AV's should initially provide some level of spoken feedback.
- However, as they become more familiar with the vehicles, they may want the option to reduce the amount of feedback as their trust develops
- People who received spoken feedback experienced falling trust during motorway scenarios, implying that a level of adjustment and customisation may be required for different scenarios.

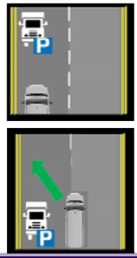
Study 2 - Trust with Visual Feedback - Method

- **Group 1** received no visual feedback from the vehicle.
- **Group 2** received moderate visual feedback – consisting of what the vehicle could "see".
- **Group 3** received a high level of visual feedback – consisting of what the vehicle could "see" and the action the vehicle would take

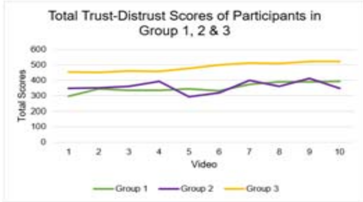
What the vehicle can see and what it intends to do

Vehicle parked in lane (level 2)

Vehicle pulling back into correct lane (level 3)

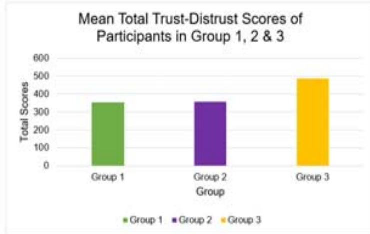


Results



- Group 1 - no visual feedback from the vehicle.
- Group 2 - moderate visual feedback – what the vehicle could “see”.
- Group 3 - high level of visual feedback – what vehicle could “see” and what it would do

Mean Trust Score



	Mean Trust-distrust Scores
Group 1	352.6
Group 2	358.1
Group 3	486.7

Discussion

- High level feedback (G3) was associated with the highest trust ratings compared to the other two groups, with results proposing that there is no significant association between G1 and G2 ($t = 0.360$; $p = 0.723 > \alpha$). but a significant difference between G1 and G3 ($t = 9.744$; $p =<0.0001 < \alpha$) and between G2 and G3 ($t = 8.663$; $p =<0.0001 < \alpha$).
- Additionally, participants' trust level adjusts depending on whether it is a safety or non-safety-critical situations
- It was also found that there is an upward inclination of trust in all groups of feedback due to the idea of familiarity over time.

Study 3 – Trust in Longitudinal Conditions

Study Aims

- To investigate the driver behaviour in a highly automated driving environment (Level 3) when the driver's role shifts from driving to monitoring the system to resuming control during an emergency event.
- To examine driver Trust, Situation Awareness and Behavioural Adaptation during transitions (manual to autonomous mode and vice versa)
- To examine driver engagement in various secondary activities during automated driving mode

Study Methodology

- Qualitative study, multiple qualitative data collection techniques
- Convenience sample
- 6 participants, aged between 29 to 55 years old
- All fully licenced, more than 5 years driving experience
- Participants screened for motion sickness, migraine, epilepsy etc.

Nottingham University Driver Simulator

Methodology – Drive Without Critical Incident

- Each subject drove for ~30 minutes on 1st, 2nd, 3rd and 5th days
- Initial period of manual driving (~5 minutes)
- Followed by Automated Mode (~20 minutes)
- Followed by manual driving (~5 minutes)

Methodology – Drive With Critical Incident

- Each participant drove for ~20 minutes on 4th day.
- Manual drive for ~5 minutes
- Automated drive for ~15 minutes – then system initiated Manual Driving mode through Voice Message because of "unexpected event" (light and heavy fog)

Methodology – Manual to Automated Transition

- Transition from Manual to Automated and vice-versa communicated through voice instruction
- Transition from Manual to Automated initiated by driver voice command (“Start Automated Driving”) – system responds accordingly
- Transition from Automated to Manual imitated by system (“Approaching Takeover in 60 seconds” then “Starting Manual Control”)

Methodology – Questionnaire Measurement

Trust
Situation Awareness (SART)

Methodology – Health Bar

- On dashboard, displaying status of autonomous vehicle
  - Green – autonomous system very effective in handling driving task
  - Amber – minor technical malfunction which it tries to resolve without driver intervention
  - Red – Major technical problem – possibility that it may need driver to take over
- Health bar status constantly changing based on system efficiency and severity of problem

Procedure – Incident Days

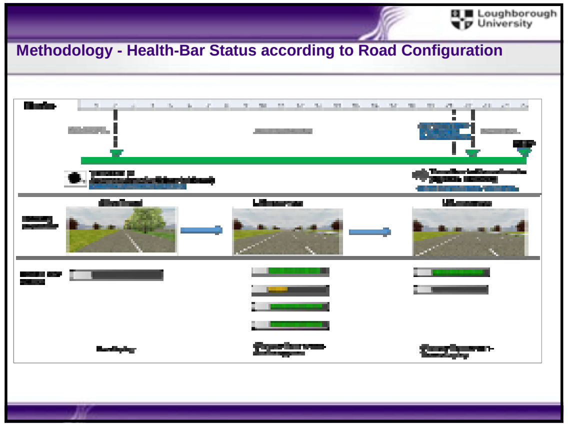
- On successful completion of transfer of authority, participants allowed to drive in manual mode through heavy fog condition for rest of study unless instructed otherwise
- All participants were <u>unaware</u> of impending critical incident event scheduled on DAY 4 prior to their participation.
- This enabled gathering of data (naturalistic responses during critical incidents in automated driving)

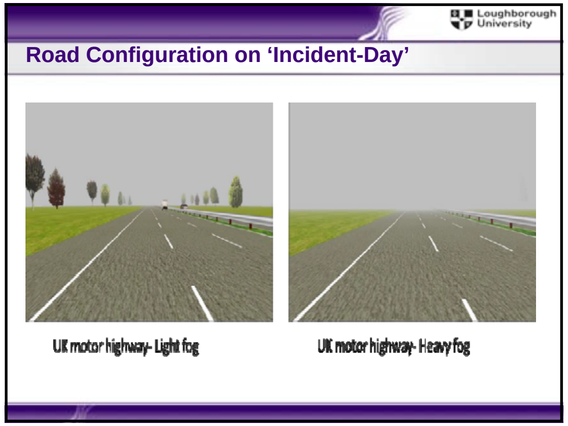
Methodology - Vehicle “Health Bar”

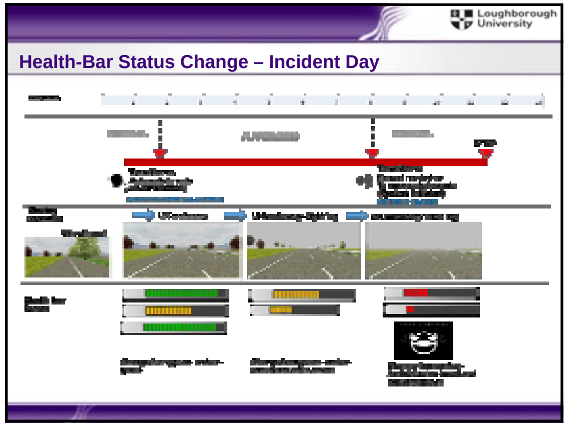
Methodology - Road Layout/Configuration

Slip off road in the UK motor highway

UK motor highway (3 lane)


Loughborough University
Methodology - Health-Bar Status according to Road Configuration


Loughborough University
Road Configuration on 'Incident-Day'
UK motor highway- Light fog
UK motor highway- Heavy fog


Loughborough University
Health-Bar Status Change – Incident Day

Data Analysis

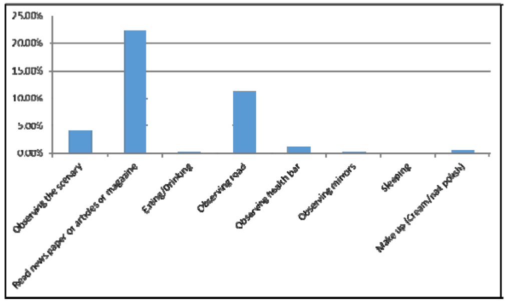
- Video recording of all participants (N=6) gathered during drive
- Qualitative data regarding automated driving experience of participants
- Quantitative data: Questionnaires used in study: Trust, Situation Awareness
- Video data coding taxonomy
  - Focussed on major activities performed during driving task.
  - Also incorporated specific activities that observed during transition phases (M2A mode and vice versa).

Video Coding Taxonomy

(a) using social networking websites
(b) checking email
(c) visiting website
(d) playing games
(e) use camera
(f) download app
(g) watch videos/ download/streamlining videos
(h) view gallery
(g) IM/Chat
(i) use GPS
(j) reading articles or newspapers etc.

List of things brought by participants
1.Mobile
2.IPAD
3.Laptop
4.Mobile and IPAD
5.Mobile and Laptop
6.Journal articles
7.Eye glasses and holder
8. Water bottle
9. Moisturizing creams
10.Nail polish kit
11.Audio head phones
12.Handbags

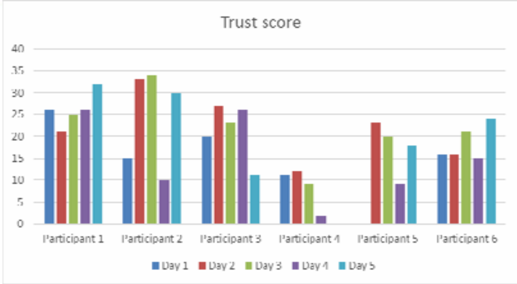
Results – Driver Activities



Activity by Participant

Participant Number	Total Time in Automated Driving Mode	Main Activities
1	90.2 minutes	Reading newspaper (89%)
2	90.8 minutes	Internet (43%) Watching video (29%)
3	91.2 minutes	Watching video (11%) Miscellaneous, fidgets (15%)
4	72.1 minutes	Social networking (36%) Viewing
5	72.2 minutes	Reading newspaper (91%)
6	93.5 minutes	Internet (26%) Observing scenery (22%)

Participant Trust



Effect on Postures

Interaction with Non-vehicle Items

Discussion and Conclusions

- With a highly automated vehicle, people will naturally bring a wide range of objects into the car space
- During automated driving, participants will vary considerably in the range of activities they prefer to carry out
  - Most of these will have a strong visual element
- Trust and situation awareness scores will:
  - Change with exposure
  - Change as a result of an unexpected emergency event
- General trend is for participants to look at health bar less with time
- With current vehicle designs, people will struggle at times to find a comfortable posture in which to engage with their preferred activities

Discussion and Conclusions

- A number of interesting behaviours occurred in the study worthy of further investigation, e.g.
  - Steering wheel wobble on takeover of control
  - Use of reading glasses
  - Postural adjustments for handover/takeover
  - Suspending secondary (now primary) tasks for takeover
- Future work should consider a larger sample and a wider range of driving situations to allow quantitative analysis

General Conclusions Regarding Trust

- Perhaps one of the most important parameters for successful uptake of Autonomous Vehicles
- Changes with time – people initially suspicious but develop trust as familiarity with the system advances
- Feedback is clearly essential for engendering trust –a visual or auditory feedback mechanism in isolation both work well in this regard – a combination of both might be better – although need to consider 'annoyance' to the driver
- However, does the driver actually need visual feedback at all? And does the need for auditory feedback diminish with time?
  - Key thing is that vehicle occupants need to know what the vehicle is doing
  - But VF causes secondary task impediment
- Specific events can cause trust in automation to nearly 'go back to square

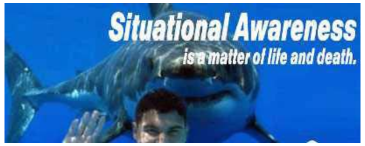
INTRODUCTION

Definition

Situation Awareness

SA 'is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future'

(Endsley 1988 Design and evaluation for situation awareness enhancement Proceedings of the Human Factors Society 32nd Annual Meeting)



44

INTRODUCTION

Definition

Simplified - being aware of what is happening in your vicinity, comprehending the relevance of relevant aspects within the current situation, and predicting the future status of the situation. Apply to driving……..

45

History – OODA Loop - *observe, orient, decide, and act,*

Military theory

1950s/60s Korean and Vietnam wars - USAF SA crucial to combat OODA loop

1980s Aerospace industry – Pilot support

1990s Air Traffic control, process control rooms e.g. nuclear, health, road transportation

Col. John Boyd 'In combat, the winning strategy is to "get inside" your opponent's OODA loop, not just by making your own decisions quicker, but also by having better SA than the opponent, and even changing the situation in ways that the opponent cannot monitor or even comprehend'.

http://www.youtube.com/watch?v=VuBiMM7VGh8

Why is the Relevance of SA in Autonomous Vehicles?

- Partial Autonomy – vehicle part-controlled by operator (e.g. ACC)
  - Where am I? What is around me? What will happen in the future? – needed to avoid other road-users, parked vehicles etc.
- Hand-over – from fully autonomous driving to operator-controlled
  - Where am I? What is around me? What will happen in the future? – needed so driver knows here he/she is within the traffic flow

How to measure Situational Awareness

- Objective Ratings
  - e.g., "real-time probes" presented as open questions embedded as verbal communications during the task (Jones & Endsley, 2000),
  - SA global assessment technique (SAGAT- Endsley, 1995a)
  - WOMBAT situational awareness and stress tolerance test (used in aviation since the late 1980s)

DESIGNING FOR SA

How to measure Situational Awareness

- Subjective Ratings
- Directly assess SA by asking individuals to rate their own SA on an anchored scale
- E.g., the participant situation awareness questionnaire (PSAQ) -Strater, Endsley, Pleban, & Matthews, 2001;
- The situation awareness rating technique (SART)- Taylor, 1989).

31

Participant Situation Awareness