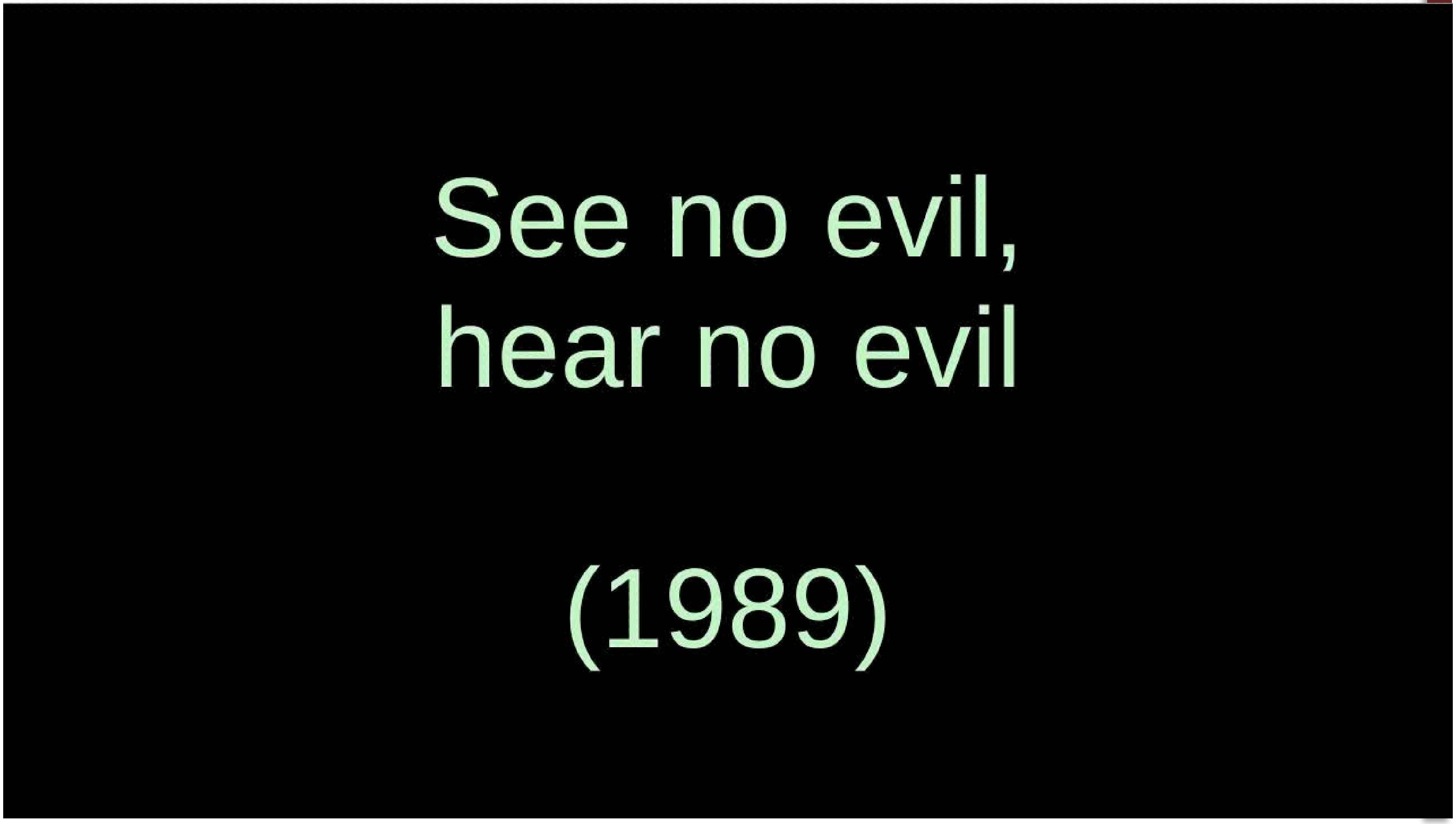
TUTORIAL:
Honest Evaluations of Shared Human-Machine Control Systems

Introduction to Shared Control TC

Tricia Gibo

Postdoctoral Researcher at Delft Haptics Lab
Delft University of Technology

What is shared control?



What is shared control?

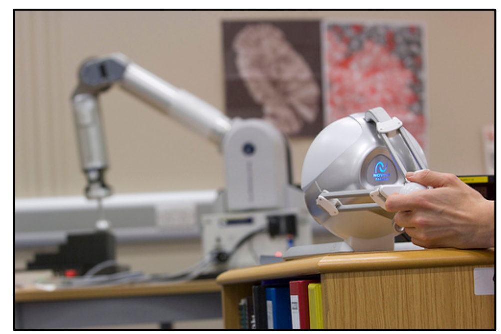
Human(s) and intelligent agent(s) are interacting congruently in a perception-action cycle to perform a dynamic task that can in principle be executed by the human, and where each of the agent(s) may have different capabilities and/or goals.

What is shared control?

Human(s) and intelligent agent(s) are interacting congruently in a perception-action cycle to perform a dynamic task that can in principle be executed by the human, and where each of the agent(s) may have different capabilities and/or goals.

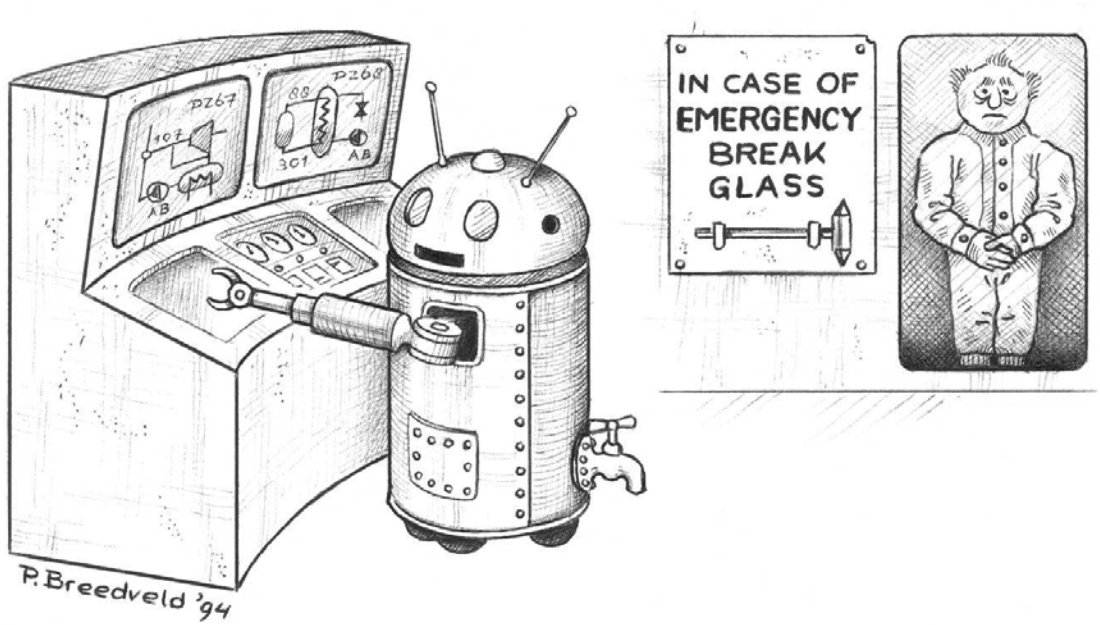
Vehicles

Robots



What is shared control?

Human(s) and intelligent agent(s) are **interacting congruently** in a perception-action cycle to perform a dynamic task that can in principle be executed by the human, and where each of the agent(s) may have different capabilities and/or goals.



What is shared control?

Human(s) and intelligent agent(s) are interacting congruently in a **perception-action cycle** to perform a dynamic task that can in principle be executed by the human, and where each of the agent(s) may have different capabilities and/or goals.

What is shared control?

Human(s) and intelligent agent(s) are interacting congruently in a perception-action cycle to perform a dynamic task that can in principle be executed by the human, and where each of the agent(s) may have **different capabilities and/or goals**.

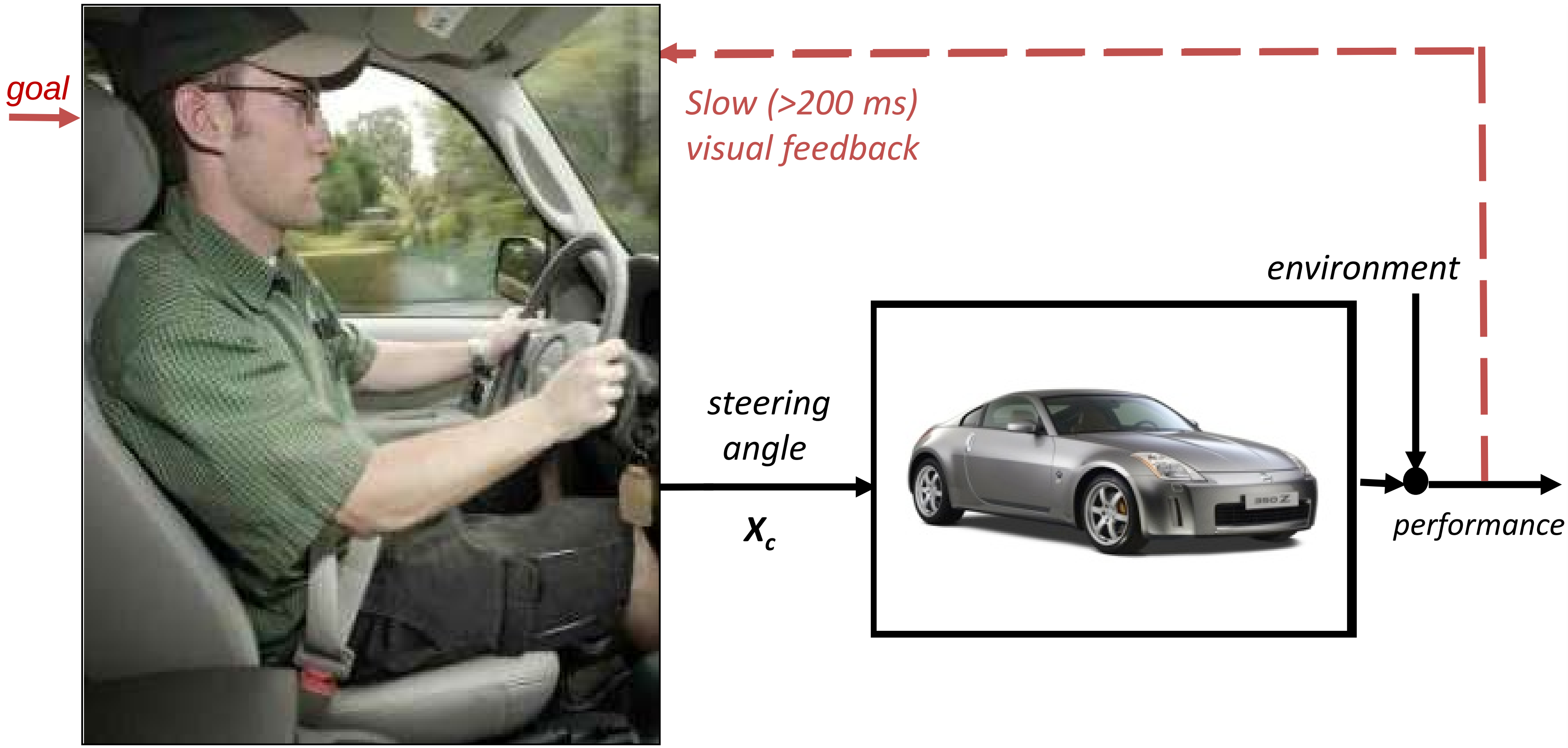
Inventive
Adaptive
Problem solving skills

Accurate
Repeatable
Inexhaustible

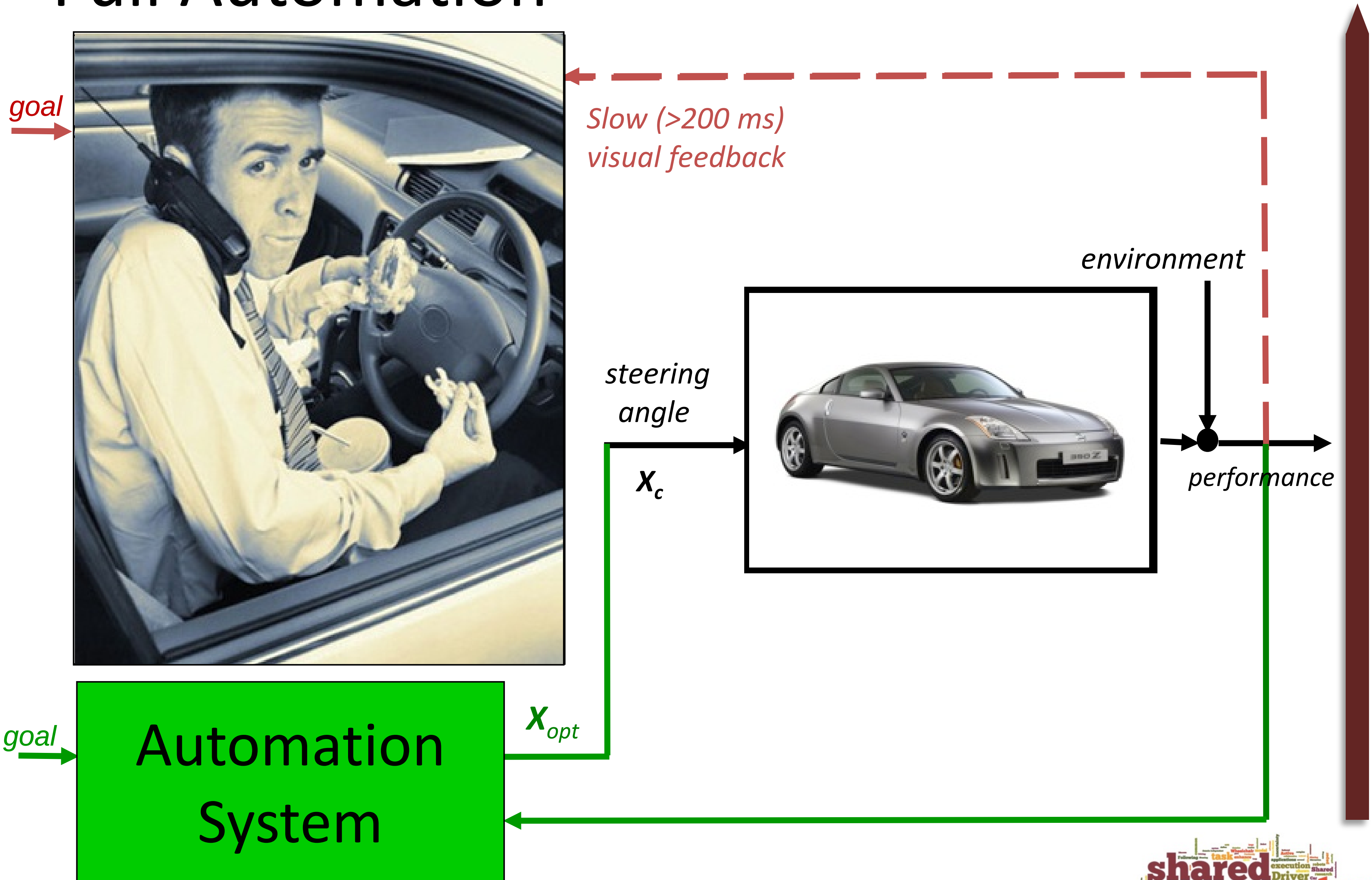
Shared control metaphor

Flemisch et al. (2003)

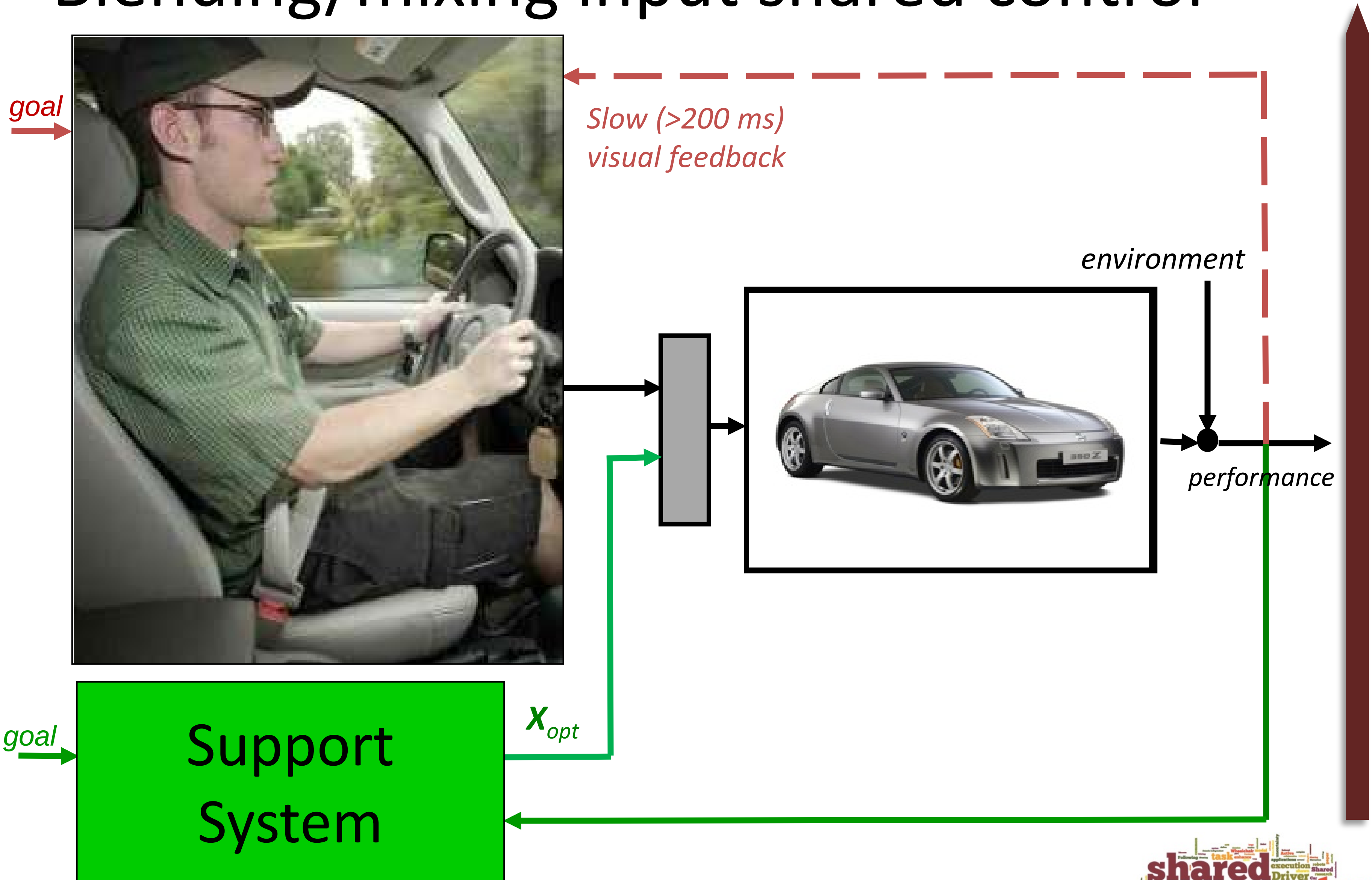
Manual control



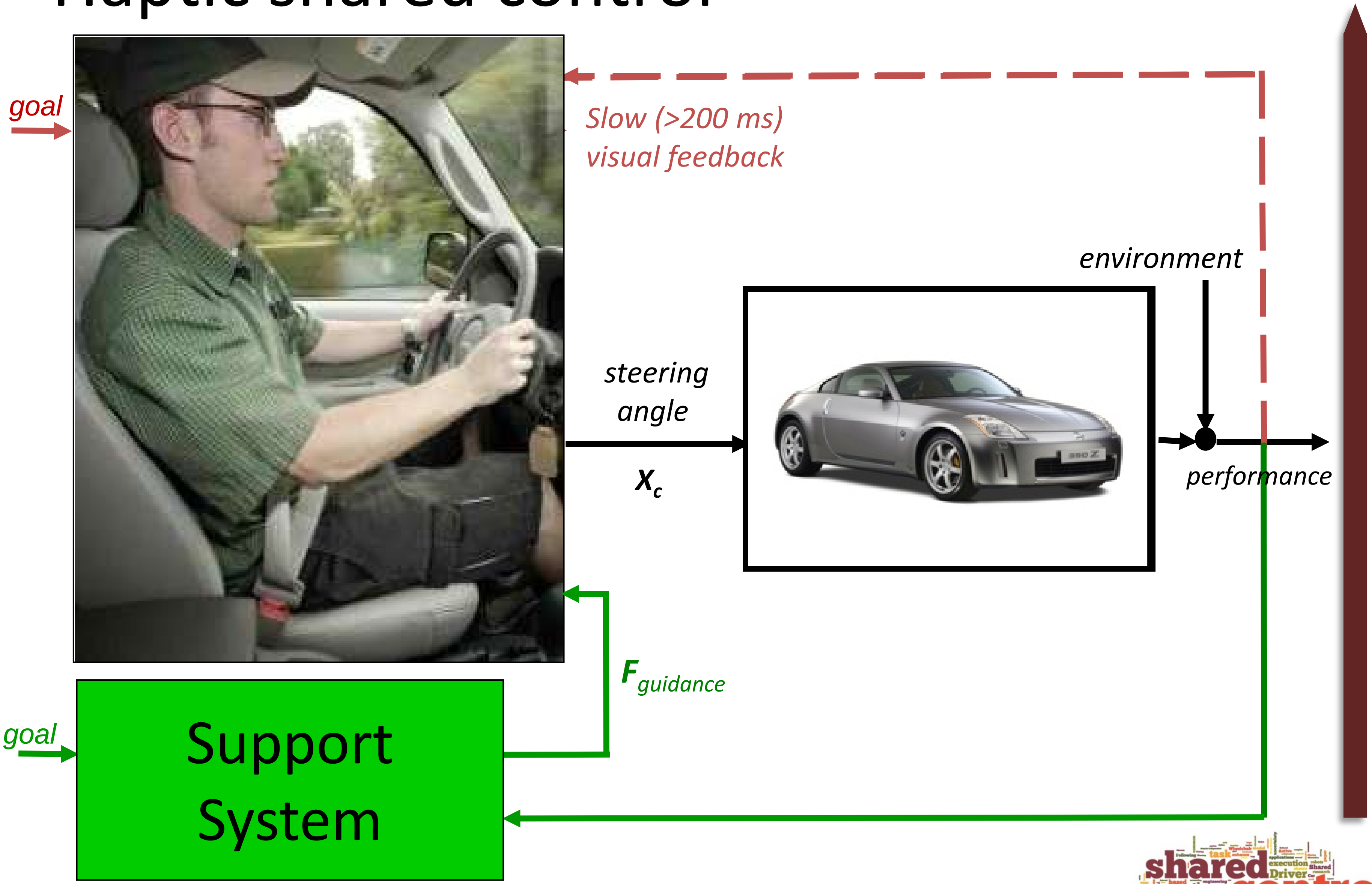
Full Automation



Blending/mixing input shared control



Haptic shared control



Technical Committee: A brief history

Founded in 2012, after 2 years of initial discussions

Mark Mulder

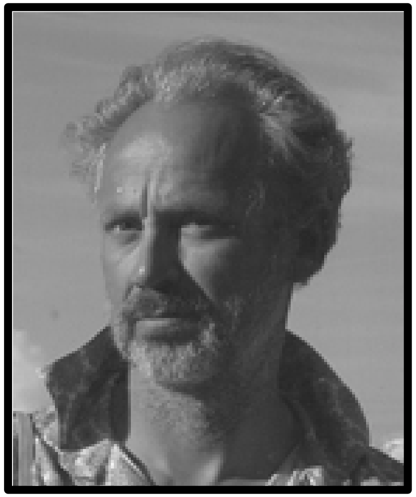
David Abbink

Tom Carlson

Current TC co-chairs

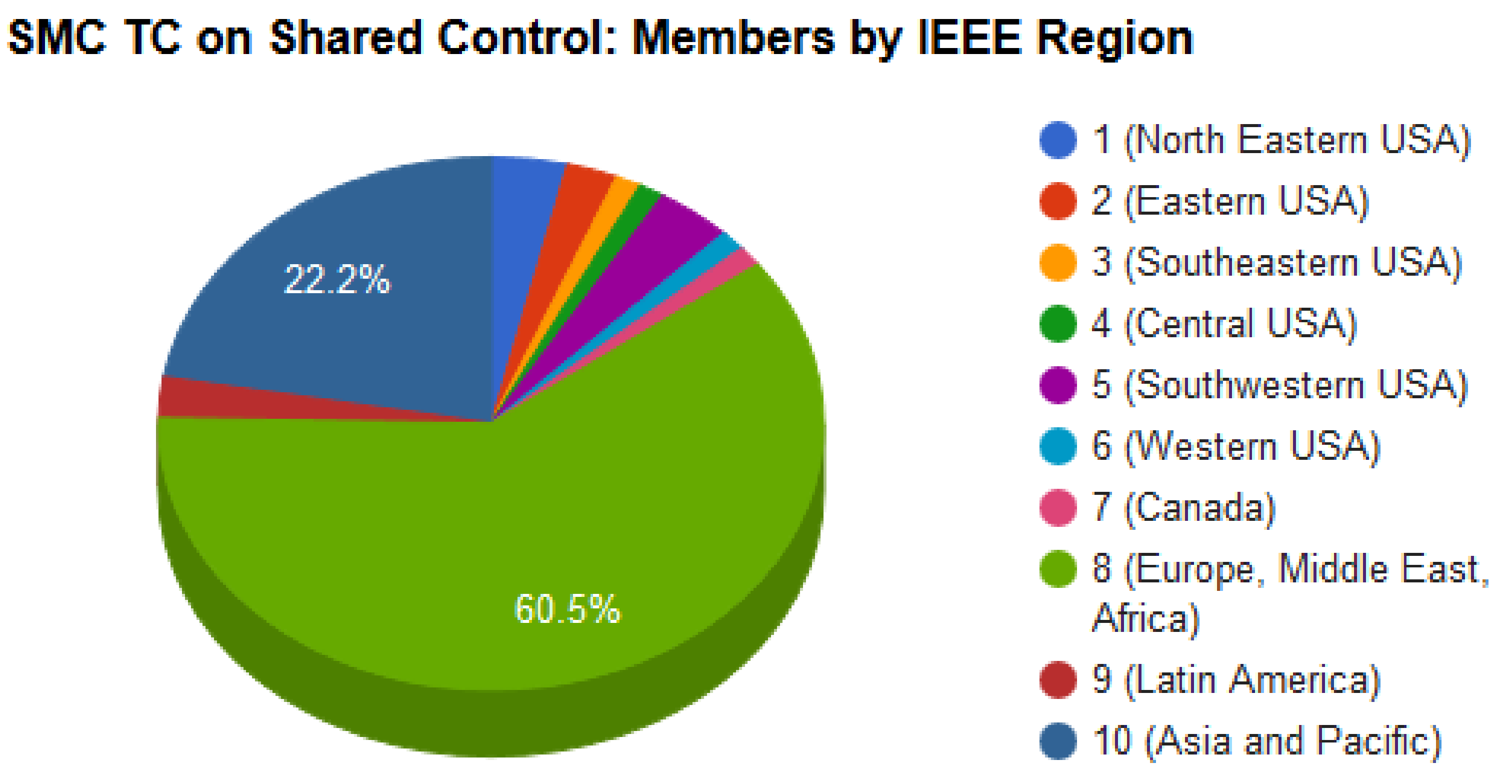
Makoto Itoh

Tricia Gibo



Erwin Boer

Technical committee members



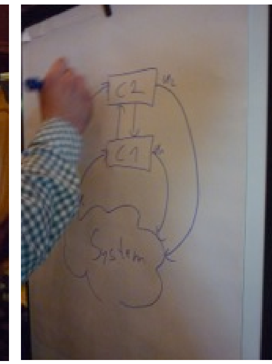
Currently 109 members and growing!

SMC sessions, workshops, & tutorials

2011-2015

Lectures by keynote speakers

Work group discussions



Hands-on demos

2015 SMC Most Active TC

DESIGN factors for shared control

- Human

  Adaptibility, intuitiveness

- Machine

  Position vs. rate control

- Environment

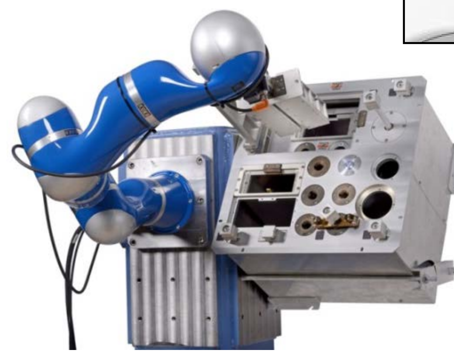
  Variability/uncertainty

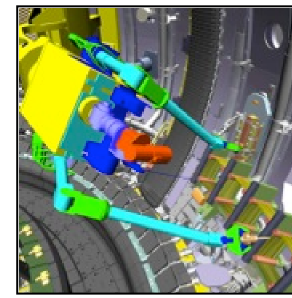
- Task

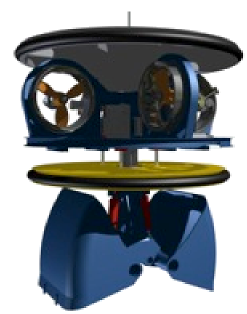
  Guidance vs. avoidance, position vs. force

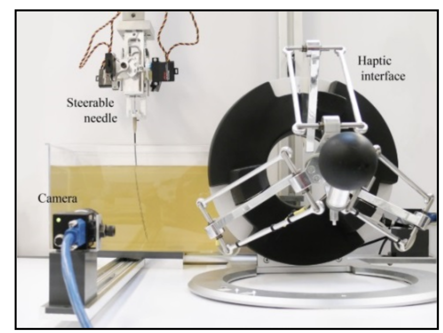
- Conflicts

  Level of authority, errors



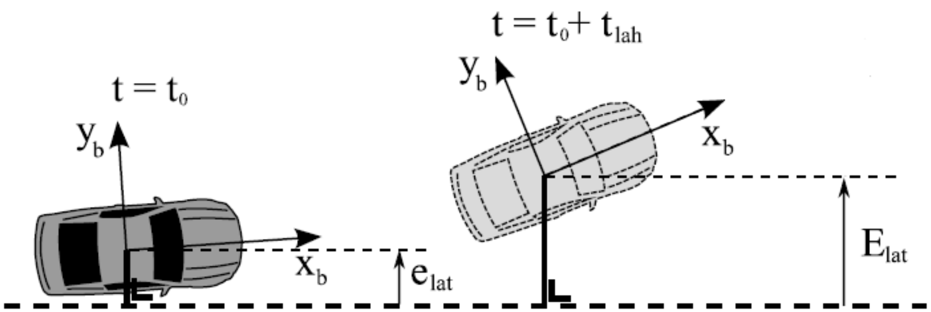






Design example

- Guidance torque on steering wheel
- Follow lane center (assume to be known)
- Trial-and-error parameter tuning



$$T_{HSC} = K_f\,E_{lat} + K_s\,|E_{lat}|\,(x_c - x_{opt})$$

Mulder & Abbink (2010); Mulder et al. (2008)

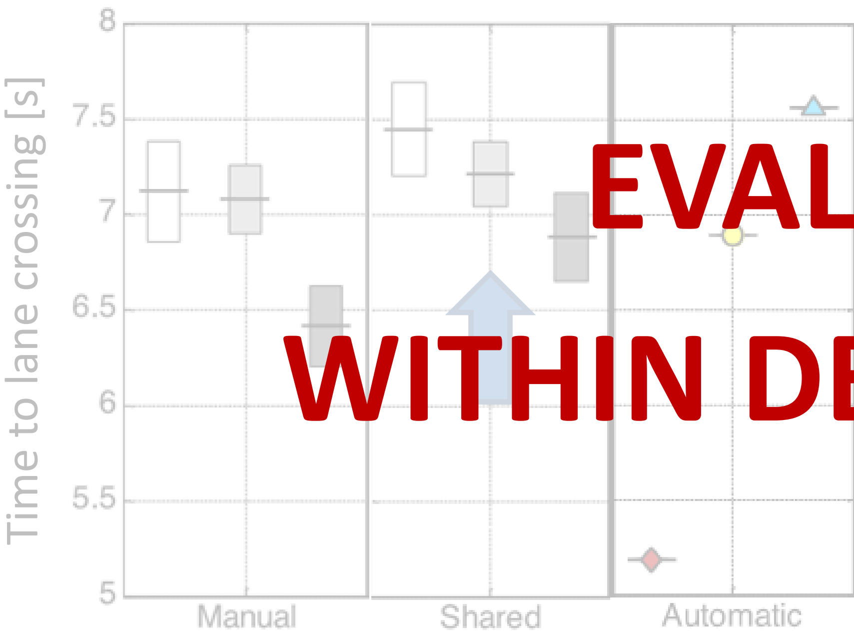
EVALUATION of shared control

- Performance
- Effort
  - Control activity
    - Control input (magnitude, frequency)
    - Physical load (force, EMG)
  - Mental load
    - Dual tasks, eye tracking
- Subjective
  - Questionnaires (NASA TLX, utility/satisfaction)

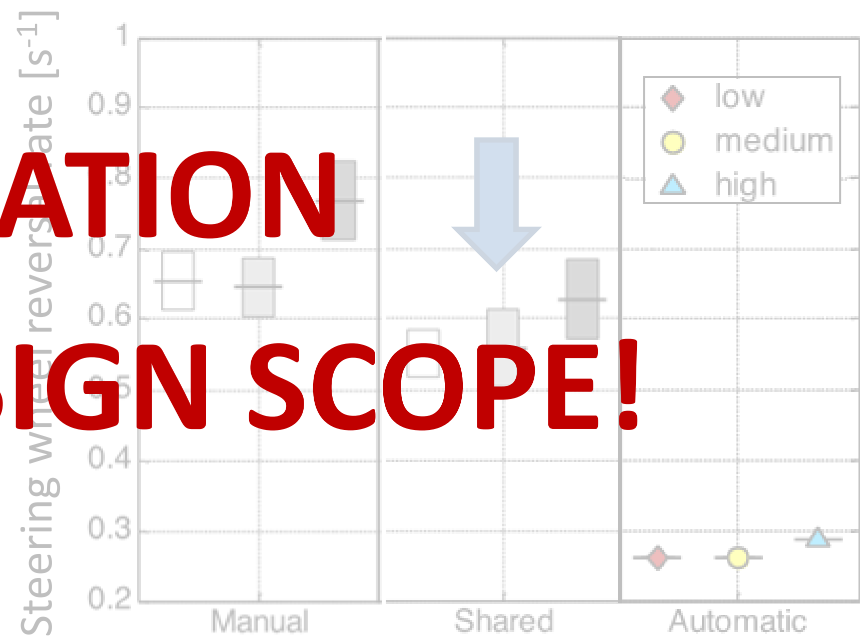
Evaluation example

20-30 yr, inexperienced
20-30 yr, experienced
> 50 yr, experienced

Performance increased



Control effort decreased



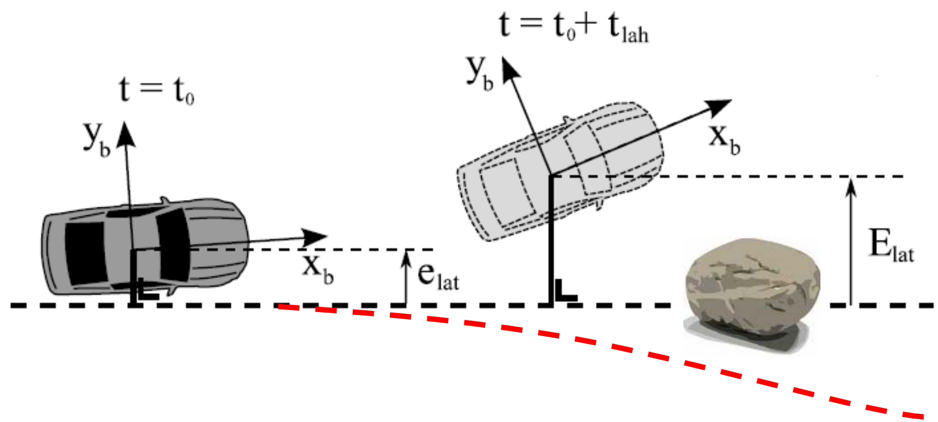
EVALUATION WITHIN DESIGN SCOPE!

Mulder et al. (2012)

Evaluation on fringe of design scope

- Realistically evaluate & compare shared control systems
- Honestly expose limitations

(from human misuse, design assumptions, etc.)



$$T_{HSC} = \cancel{K_f E_{lat} + K_s \, |E_{lat}| \, (x_c - x_{opt})}$$

Today's tutorial

1.	Introduction to Shared Control TC	*Tricia Gibo*	(20m)
2.	Automation (supervisory control)	*Erwin Boer*	(20m)
3.	Manual control	*Rene van Paassen*	(20m)
4.	Shared control	*David Abbink*	(15m)
		Bastiaan Petermeijer	(15m)
Break			(30m)
5.	Build evaluation taxonomy	*Interactive*	(20m)
6.	Evaluation taxonomy + guidelines	*Erwin Boer*	(20m)
7.	Application to other domain	*David Abbink*	(20m)
8.	Q&A	*Interactive*	(10m)

Want more?

Special Session:

Tuesday 11th

(I) 9:30-11:00 (II) 11:00-12:30

Location: Sofitel Budapest Chain Bridge, Academy 3

Sign up now and join our

SMC Technical Committee on Shared Control!

TUTORIAL:
Honest Evaluations of Shared Human-Machine Control Systems

Automation Evaluated

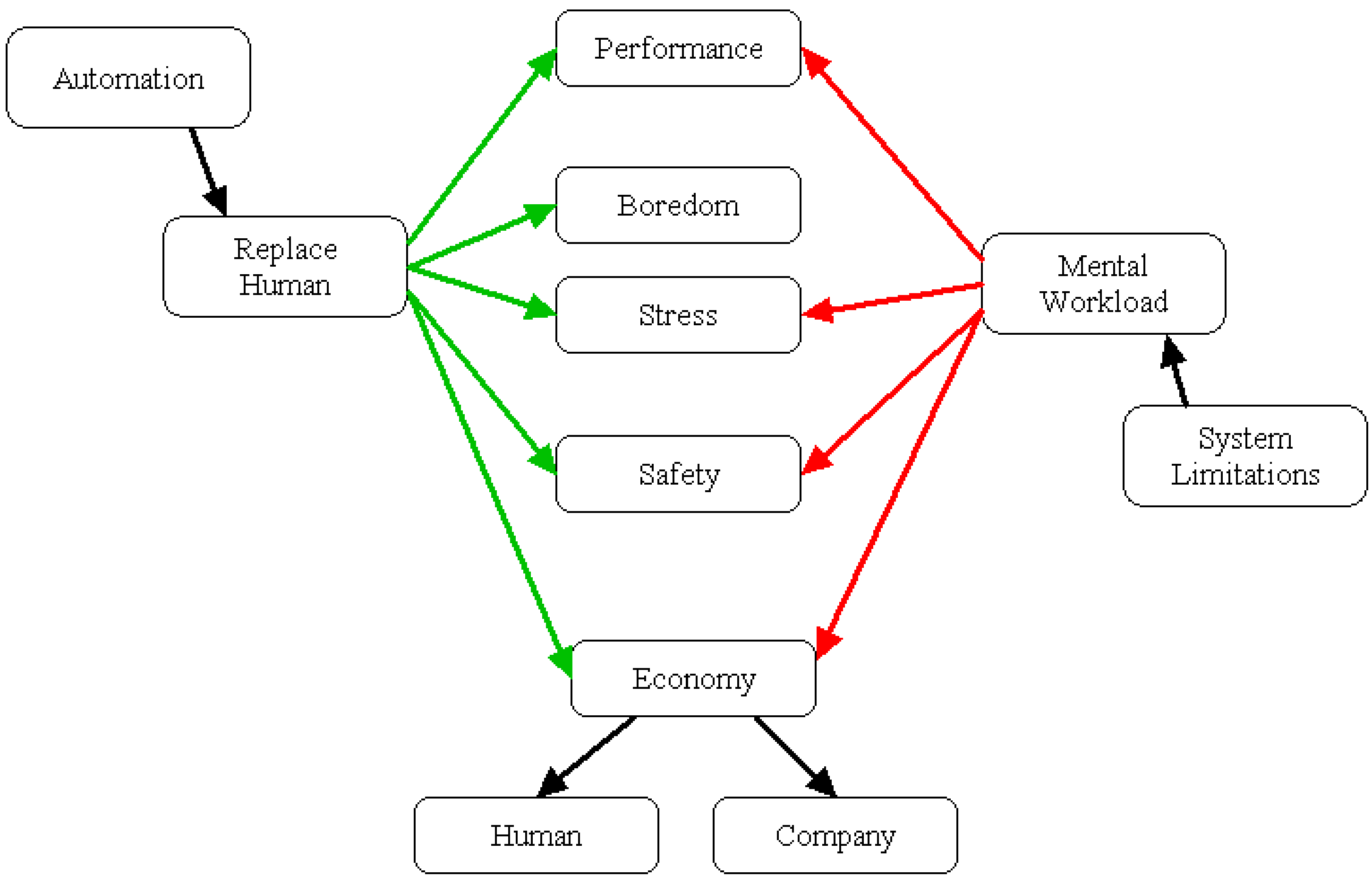
Erwin R. Boer

Associate Professor, Mechanical Engineering
Delft University of Technology

Reasons for Automation: Save the Human

Replace the Human

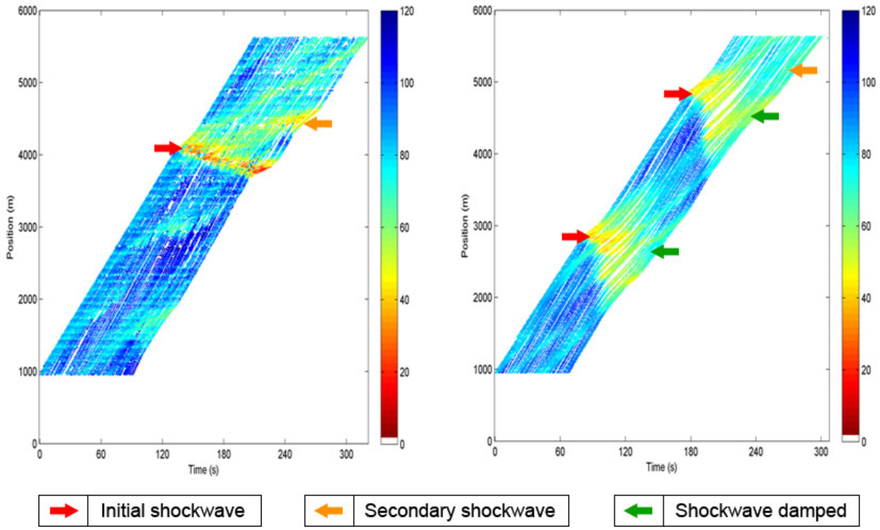
Replace the human to Save the human "Reason" in Asimov's "I Robot"



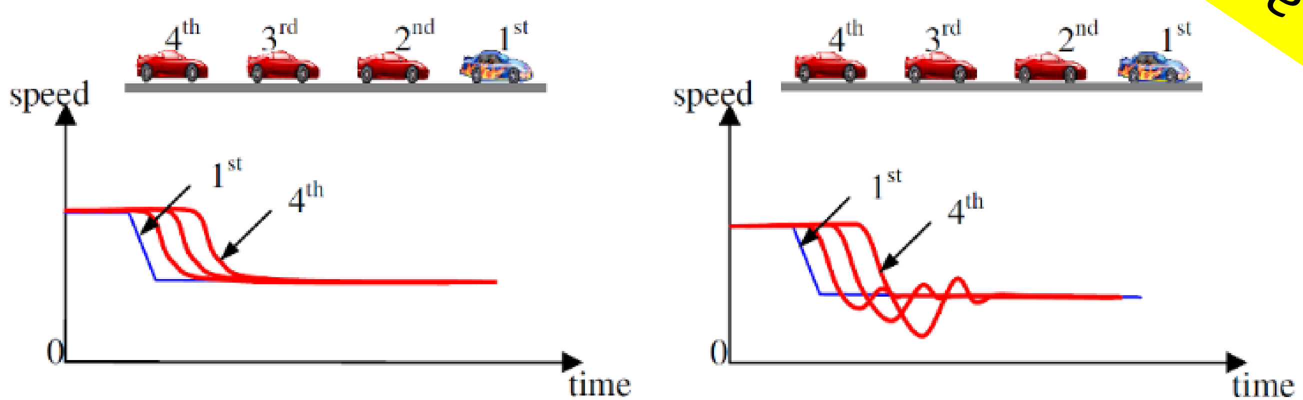
Humans and Today's Machine share Limitations

- System wins in capability:
  - Perception esp. in fog and rain
  - Processing speed
  - Divided attention / Sampling rate
  - Sustained vigilance (not distractible)
  - Perceiving without cognitive bias
  - Quick decision making in complex environments with multiple threats
  - Response rate/strength
- Human wins in capability:
  - Rapid recognition
  - Scene understanding (*new AI Turning test)*
  - Adaptation
  - Learning (*deep learning is beginning to make meaningful steps*)

ACC String Stability

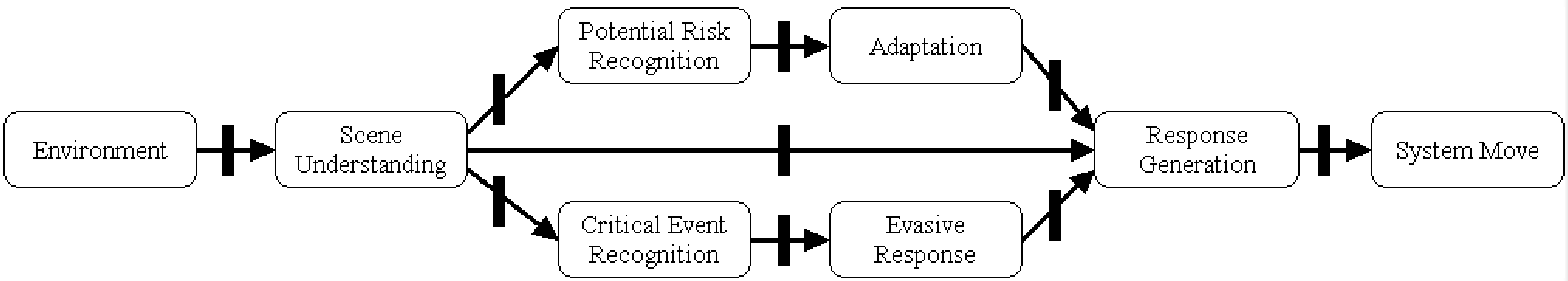


Within Scope Evaluations



(a) (b)

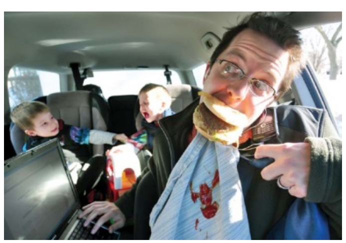
What does Human / Automation need to Know to perform a task like Driving?



Understanding of the operation context - Predicting

Understanding of potential risk - Adapting

Understanding response needs – Interacting w Environment



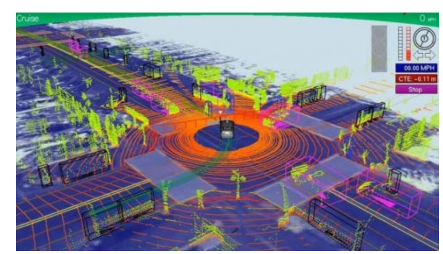
The norm against which to compare automation?

Human Ability Limitations:

- *Youth*
- *Age*
- *Distracted*
- *Disease*
- *Drugs*
- *Drowsy*
- *Stressed*
- *…*

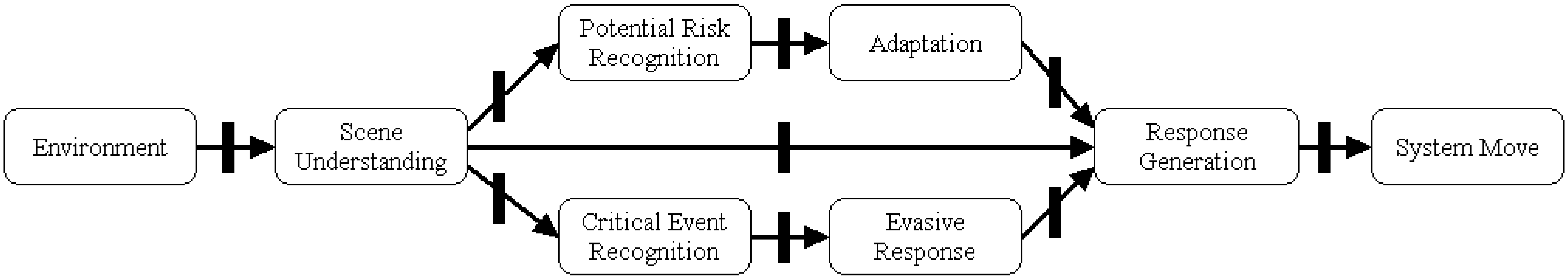
Humans fail in tasks depending on degradation in ability.

System Limitations:



System and human have different limitations; how to compare and contrast?

Reasons why system reliability may be low!



Taxonomy developed later today.

If system ability is **limited** in one or more stages under some conditions in certain situations then reliability drops.

Human needs to be **calibrated** to system situated ability/reliability.

Human needs to develop accurate situated **mental models of system limitations**.

Outside Scope Evaluations

Manual

Types of Crashes – Different Requirements
Exposing Different Limitations

Players

Expectations

<u>Type of crash</u>

- Run off road
- Read end
- Side incursion
- Side swipe
- Head on
- Curve overshoot
- Pedestrian
- Bicyclist

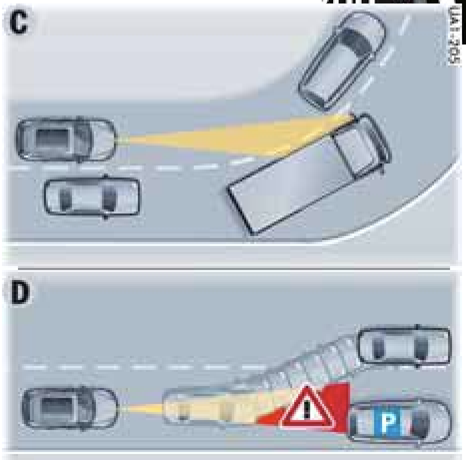
<u>Type of road</u>

- Motorway
- Highway
- Residential
- Urban
- Dirt country

Early Adaptive Cruise Control Systems

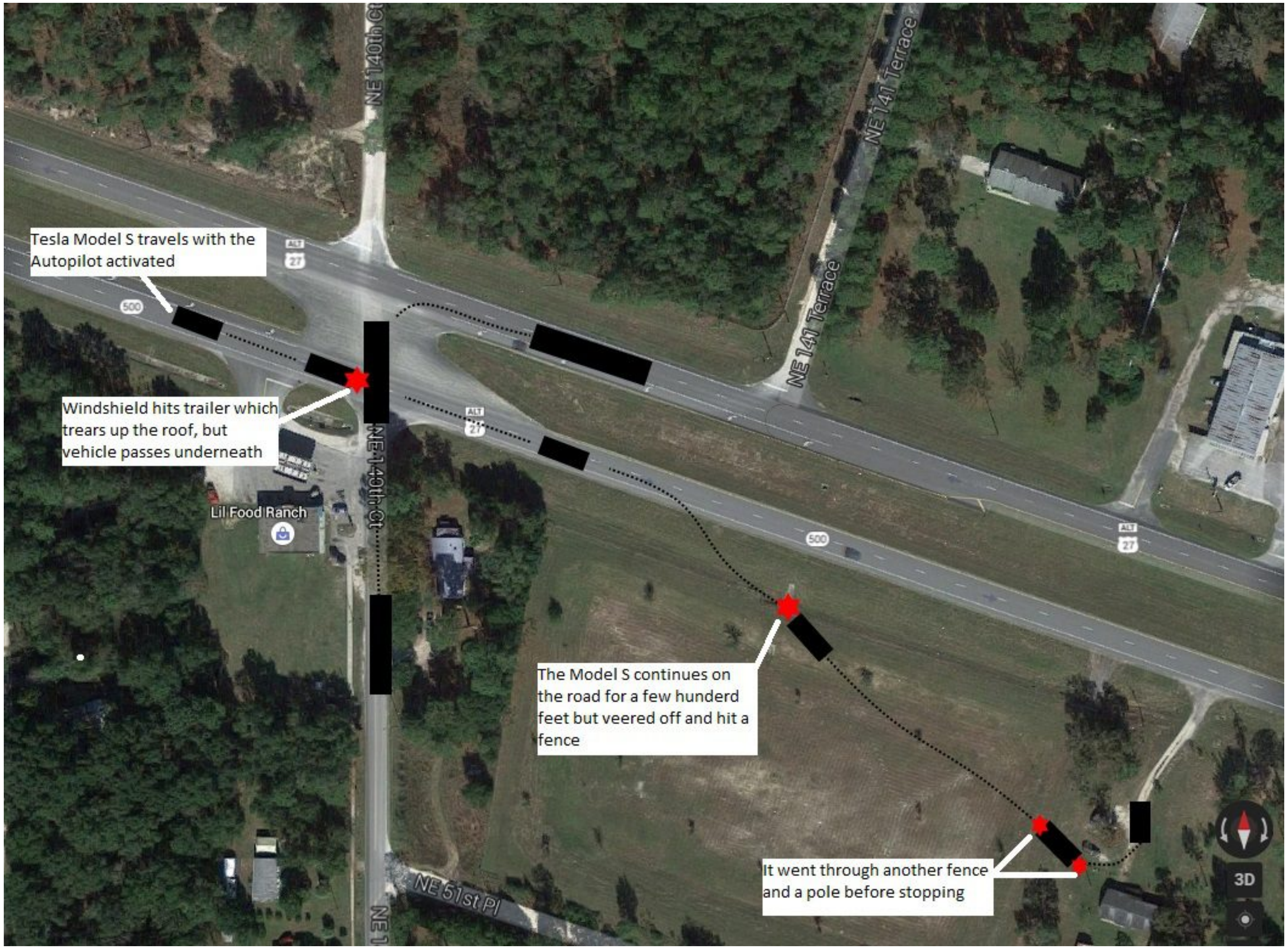
- Hardware limitations
  - Longitudinal range → high speed
  - Lateral azimuth → lose lead car in curves
  - Accuracy and stability degrade in rain → misses and false alarms.
- Imposed limitations to avoid manifestation of hardware limitations
  - Lower speed bound → small chance car in curve (no NAVI)
  - Upper speed bound
  - No dynamic lane changes and cut-ins

Very Limited Feedback



Mental Models often wrong and incomplete → Automation surprise

Tesla Accident – Missed Truck

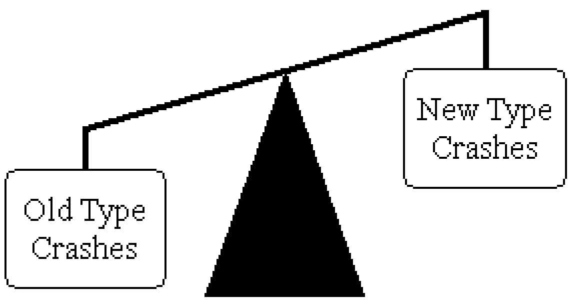


Were these early ACCs Unsafe?

Shift in Crash Types

Global versus Local Risk

- ❑ Global: total number of crashes across society decreases
- ❑ Local: number of new types of crashes that are very unlikely under vigilant manual control increase → Mental Model mismatch



More old types crashes reduced than new ones created → Issue of severity.

Do some types of individuals benefit more than others?

- ❑ Elderly drivers benefit more because relatively more involved in rear end collisions per mile traveled.
- ❑ Skilled attentive anticipatory drivers do not benefit from using ACC themselves but do benefit from its usage in society because the perturbations and thus skilled driver's own risk reduces; unless the ACC behaves in manner unpredictable to skilled driver and causes him to crash because it does not match his mental model of unsupported drivers.

Calculation of System Benefit

Fictional Example

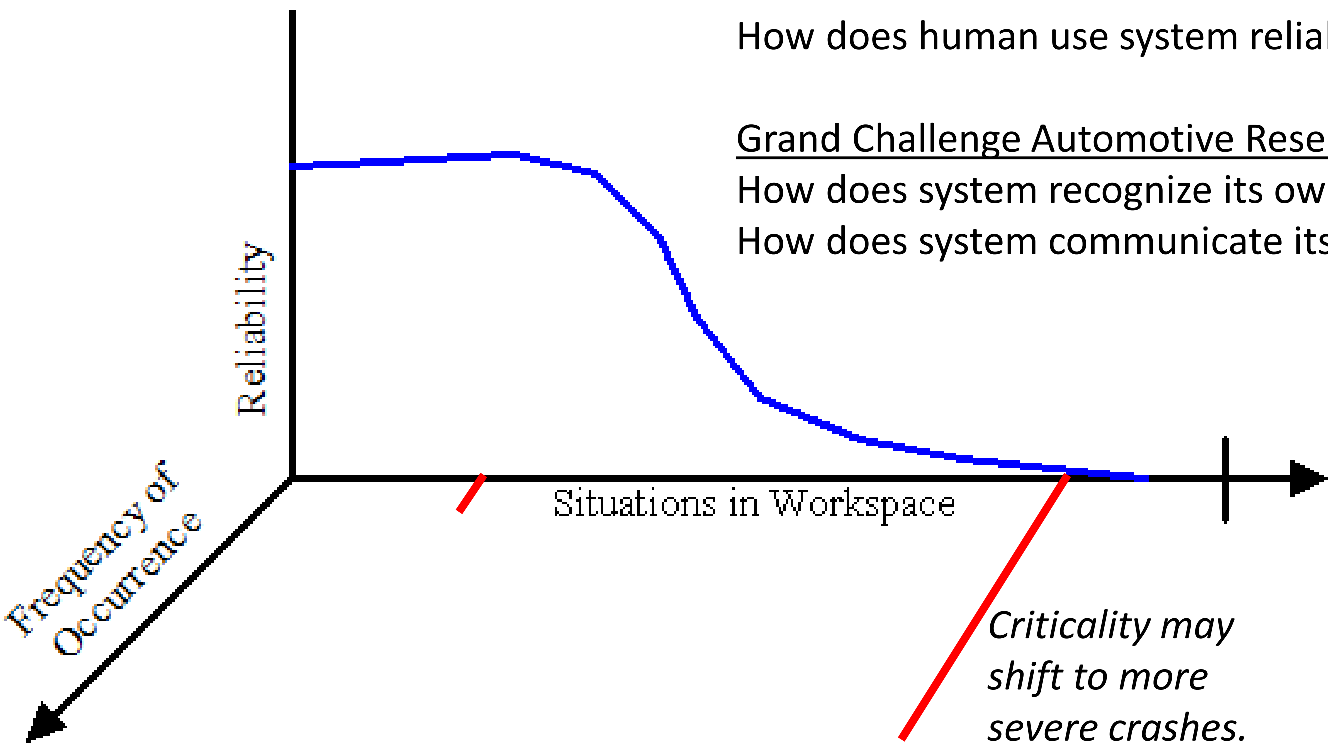
- Systems detect less but better and more vigilantly (limited scene understanding).
- Societal acceptance if total number of accidents and fatalities decreases (shifts in types and number of crashes)

Manual Control		Types of Crashes		
		Complex Intersection Side Incursion	Rear End Collision	Run off Road
Crash Cause	Human Limitation	100	100	30
	Reasonably Unexpected Event	10	10	5
	System Limitation	0	0	0
	Total	110	110	35

Auto Control		Types of Crashes		
		Complex Intersection Side Incursion	Rear End Collision	Run off Road
Crash Cause	Human Limitation	50	50	15
	Reasonably Unexpected Event	5	5	5
	System Limitation	65	5	5
	Total	**120**	60	25

- Full automation not warranted in complex intersections (limited scene understanding – new AI Turing Test)

Key issue in Evaluating Automation



<u>Human Role:</u>
How does human know and understand system reliability?
How does human use system reliability?

<u>Grand Challenge Automotive Research:</u>
How does system recognize its own reliability?
How does system communicate its reliability?

What is the Workspace?

- Situations where system could/would be used!
- Not necessarily situations for which it was intended!

Honest Evaluation → Safety Focus

- Human Needs:
  - Primary Need:
    - Safety has to improve or at least stay the same.
  - Secondary Needs:
    - Comfort, utility, satisfaction
- Limited Safety Footprint:
  - Acceptable if domains/situations/condition under which system is safe are predictable to the human with enough preview to take responsibility;
  - As long as human sufficiently alert to detect need for responsibility transitions.

Acceptance of Automation

- System management and responsibility switching cost versus personal benefit situated benefit on typical drives.

YES

Is consumption
benefit worth
the consumption
cost

NO

- Number and duration of windows of time within which safety is preserved/improved and resources freed-up/maintained drives decision.

MB vs ACC

Evaluating Automation from all Perspectives

- Automation drives differently and interacts differently from normal human drivers
  - Mismatch in mental models across road users and society.
- Implications of automation for
  - Other vehicles
  - Vulnerable road users (e.g. bicyclists)
  - Pedestrians
- Explore in virtual worlds → new use of driving simulators.
  - Keynote address at DSC2015 by Erwin Boer

TUTORIAL:
Honest Evaluations of Shared Human-Machine Control Systems

Manual Control

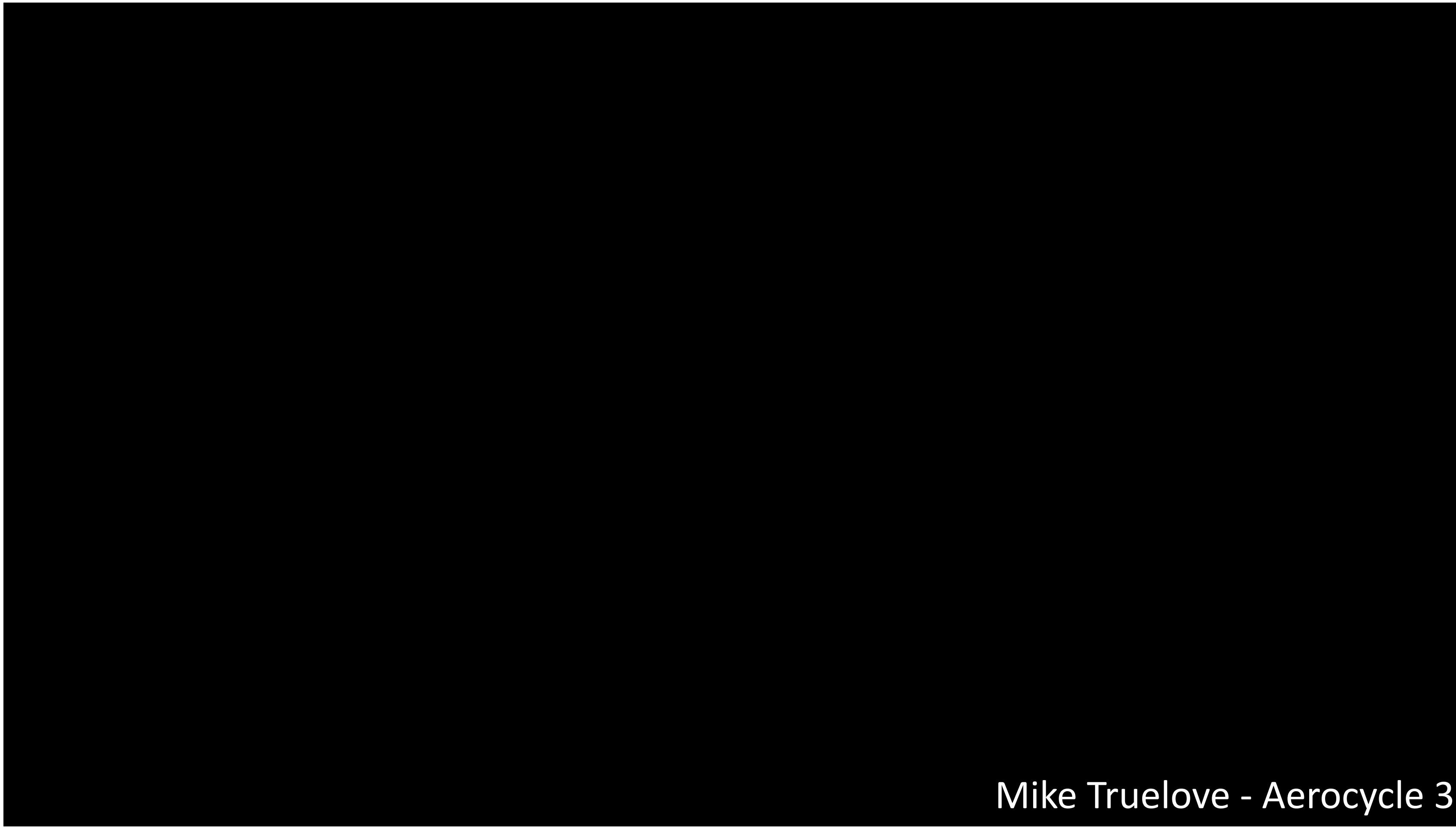
René van Paassen

Associate Professor - Control & Simulation, Aerospace Engineering
Delft University of Technology

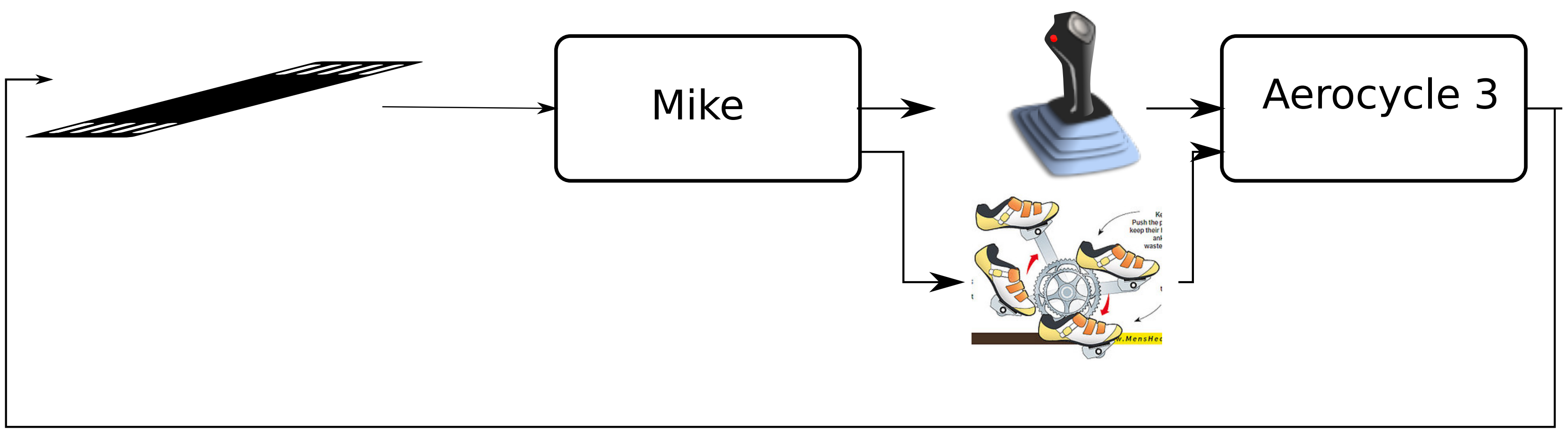
Overview

- What is manual control?
- Is it still relevant? Does it actually still exist?
- How well do we understand it?
- How do we measure & characterise it?
- What is “good” and what is “bad” in manual control?

What is manual control?



Human-controlled *and* human powered



Is manual control still relevant / does it still exist?

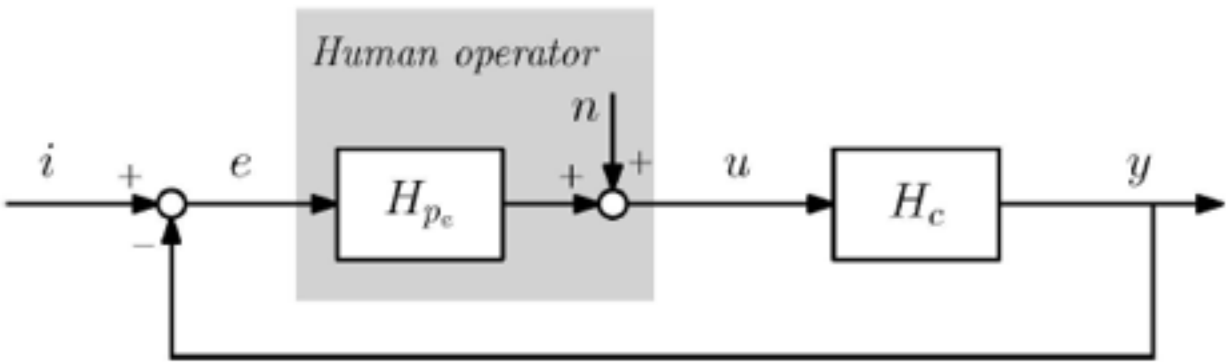
ABS; Anti-lock Braking System, including Brake Assist
ASC; Automatic Stability Control
Brake drying
Brake pre-tensioning
CBC; Cornering Brake Control
DBC; Dynamic Brake Control
DTC; Dynamic Traction Control
Electronic Differential Lock
Hill-start Assistant
Park Distance Control
Automatic headlights

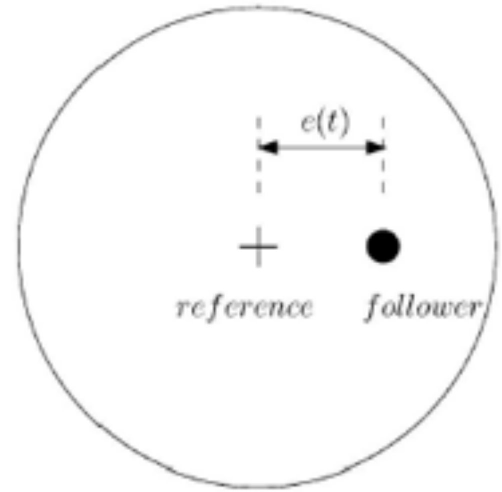
Full manual control is a rarity these days

Models of manual control

- McRuer et al.: cross-over model, simplified precision model, precision model.
- The human is an adaptable, quasi-linear controller
- The models include limitations;
  - time delay,
  - controller order/structure,
  - neuromuscular (motor) system
- Set of rules describe adaptation

What we know best



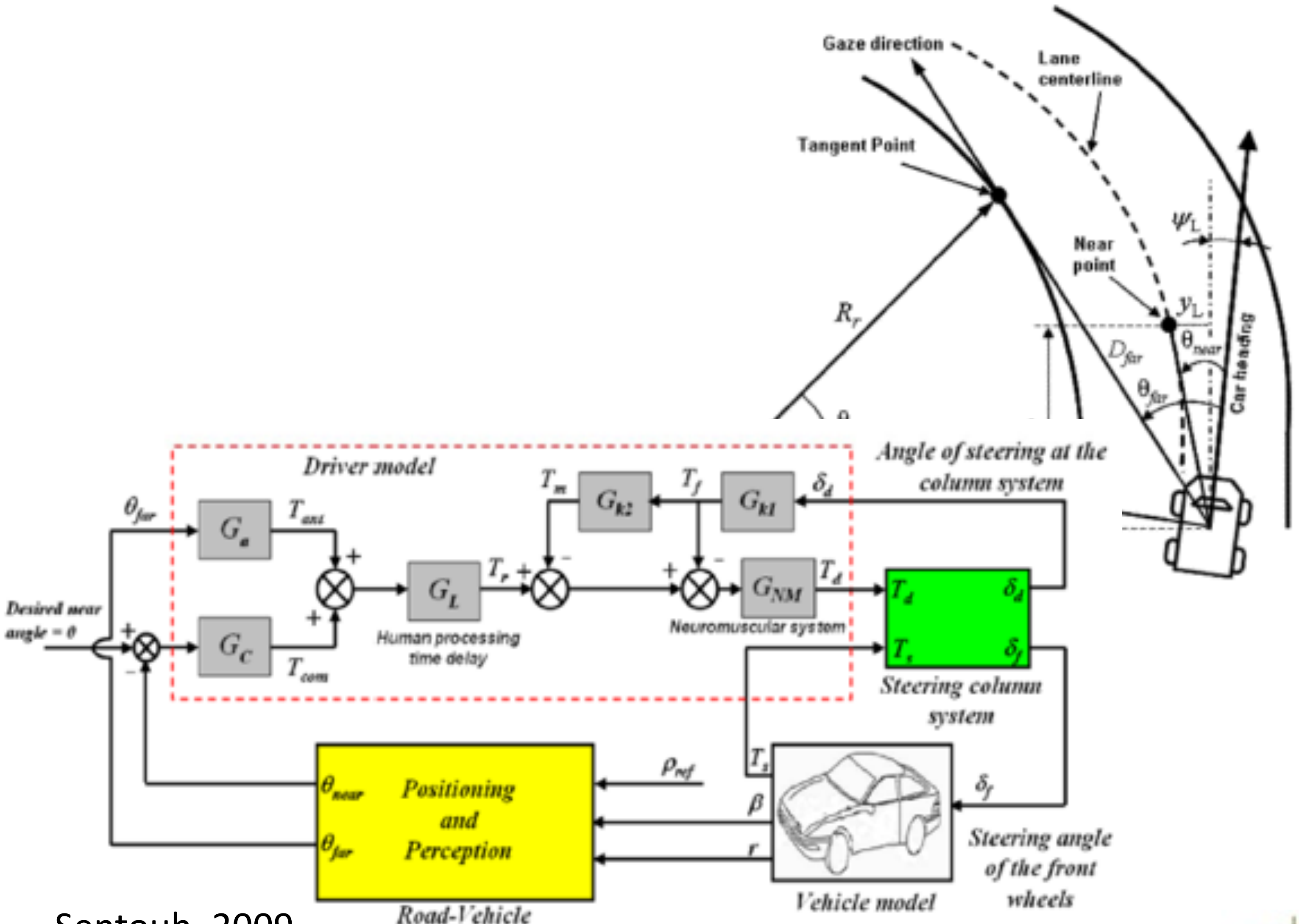


- Compensatory display, only showing error
- Single feedback loop, one signal at a time

What we are working on

multiple signals (motion, peripheral), target displays, preview displays, perspective displays, natural views, changing dynamics

Elaborations for car driving



Sentouh, 2009

Understanding manual control

- cross-over model

$$Y_{OL}(j\omega) = Y_p Y_c \doteq \frac{\omega_c e^{-j\omega\tau_e}}{j\omega}; \qquad \text{near } \omega_c$$

- Simplified precision model

$$Y_p = K_p \frac{(T_L j\omega + 1)}{(T_I j\omega + 1)} e^{-j\omega\tau_e}$$

When the system is…

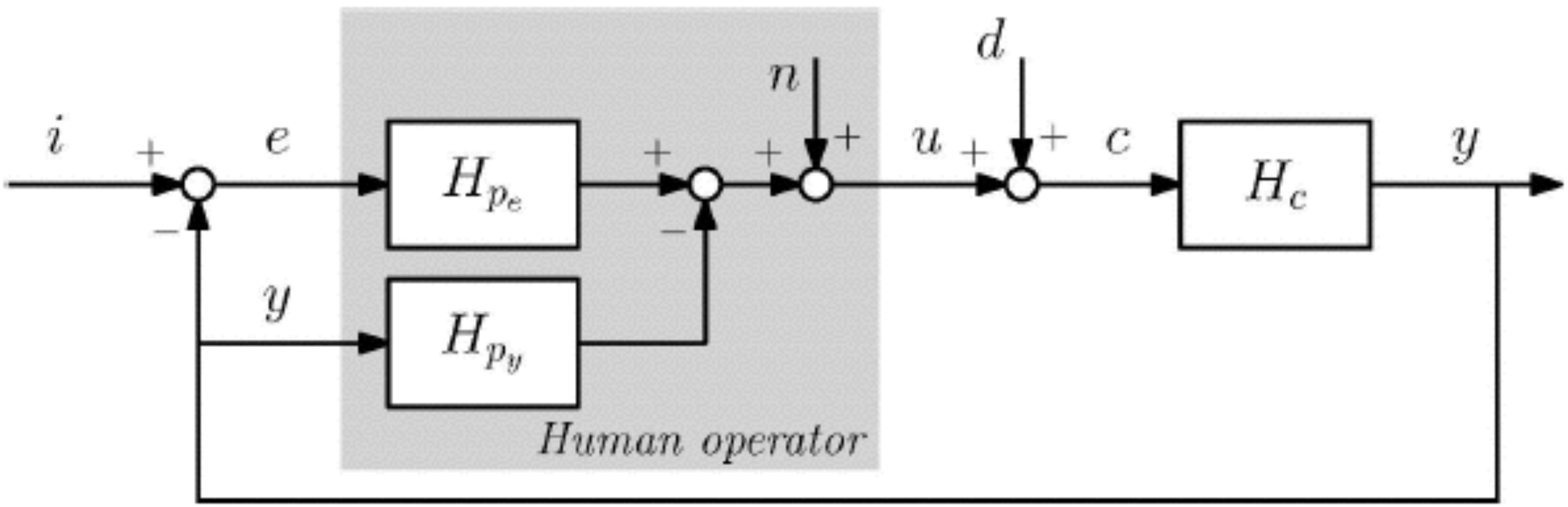
- a mouse pointer  $$Y_p = \frac{K_p}{(1 + T_I j\omega)} e^{j\omega\tau_e}$$
- a bucket  $$Y_p = K_p e^{j\omega\tau_e}$$
- a car  $$Y_p = K_p(1 + T_L j\omega) e^{j\omega\tau_e}$$

but easily visible
in scene

Evaluating manual control

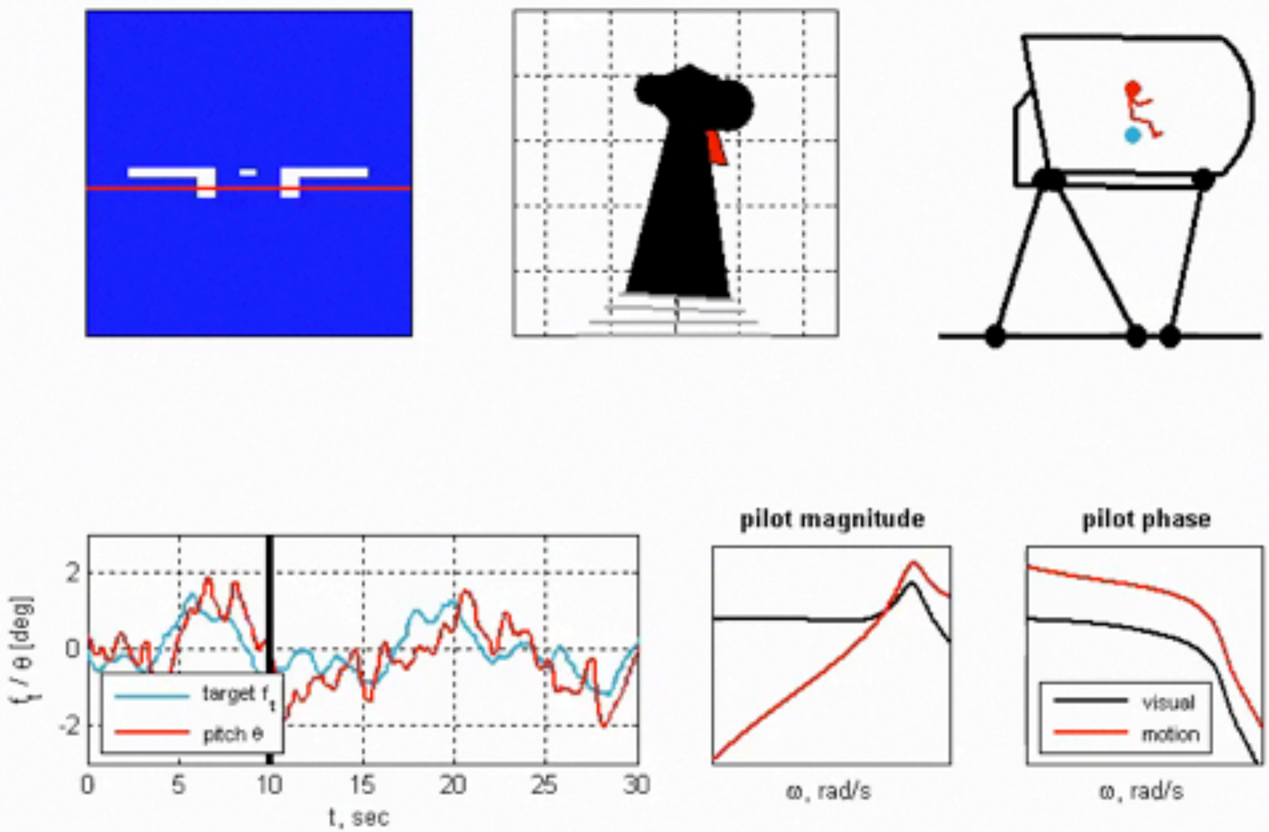
- What can you come up with?
  - Performance; standard deviation lateral error, time to lane crossing, standard deviation heading error
  - Effort; steering wheel excursion, “reversal rate”
  - Workload; subjective questionnaire, secondary task, temporal occlusion
  - Strategy; spatial occlusion, introspection

Example cybernetic approach



- H_{p_y} is here the response to motion
- The interest we have is the effect of motion cueing settings on human operator
- Performance (error e), is an insensitive measure

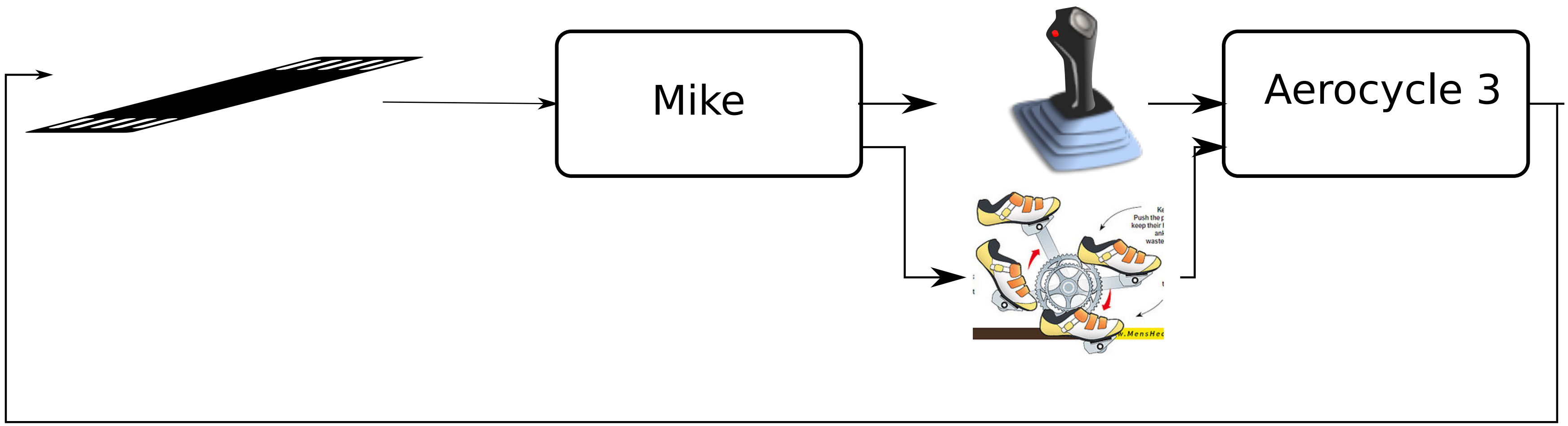
View of the results



Cybernetic approach

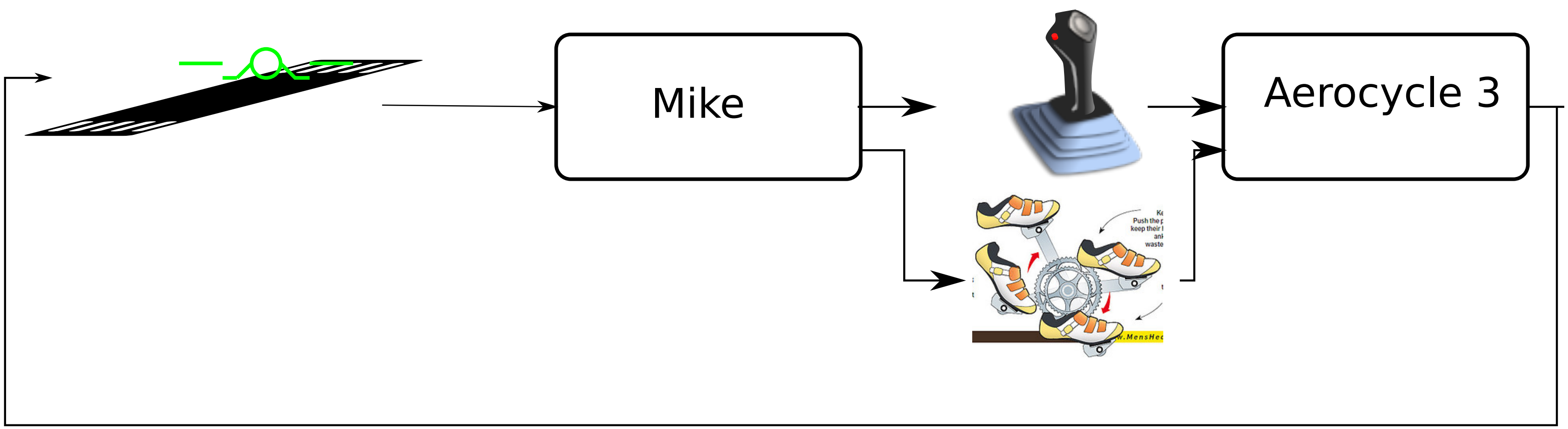
- Global metric (error variance, etc.) are often insensitive
- Identification of human control behaviour (possibly with multiple loops), to **capture** and **characterise** changes in adaptation due to system/display adaptation
- Used to investigate control with Tunnel in the Sky (Mulder, 1999), effects of aircraft flexibility (Damveld, 2009), fidelity of simulator motion (Pool, 2012), motion thresholds (Valente Pais, 2012)

What can we do & still keep it manual?



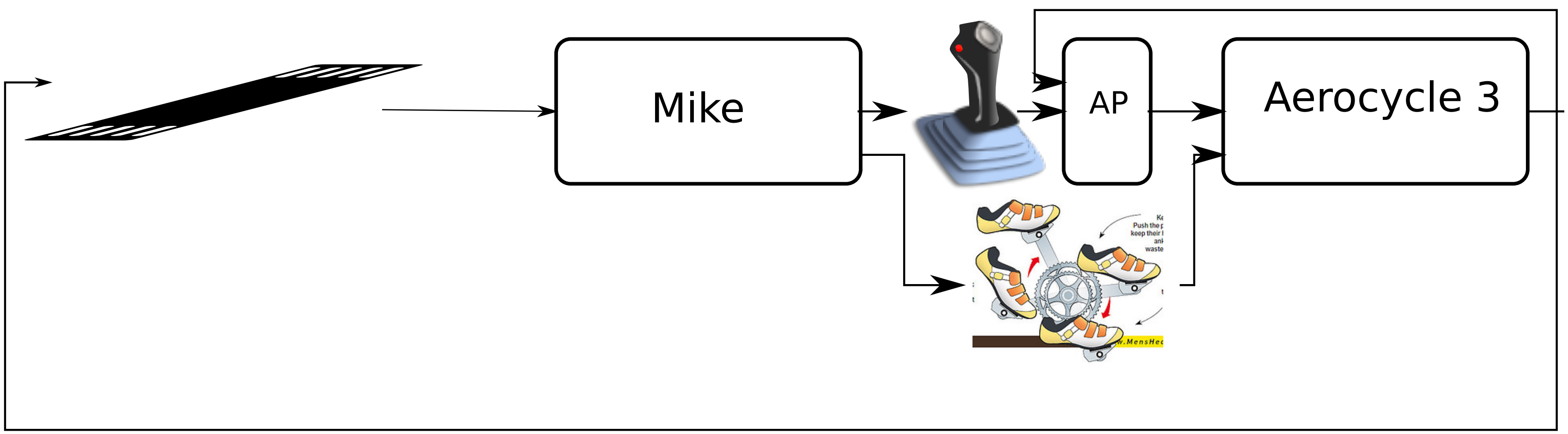
Control augmentation / Display Augmentation

What can we do & still keep it manual?



Control augmentation / Display Augmentation

What can we do & still keep it manual?



Control augmentation / Display Augmentation

Display augmentation

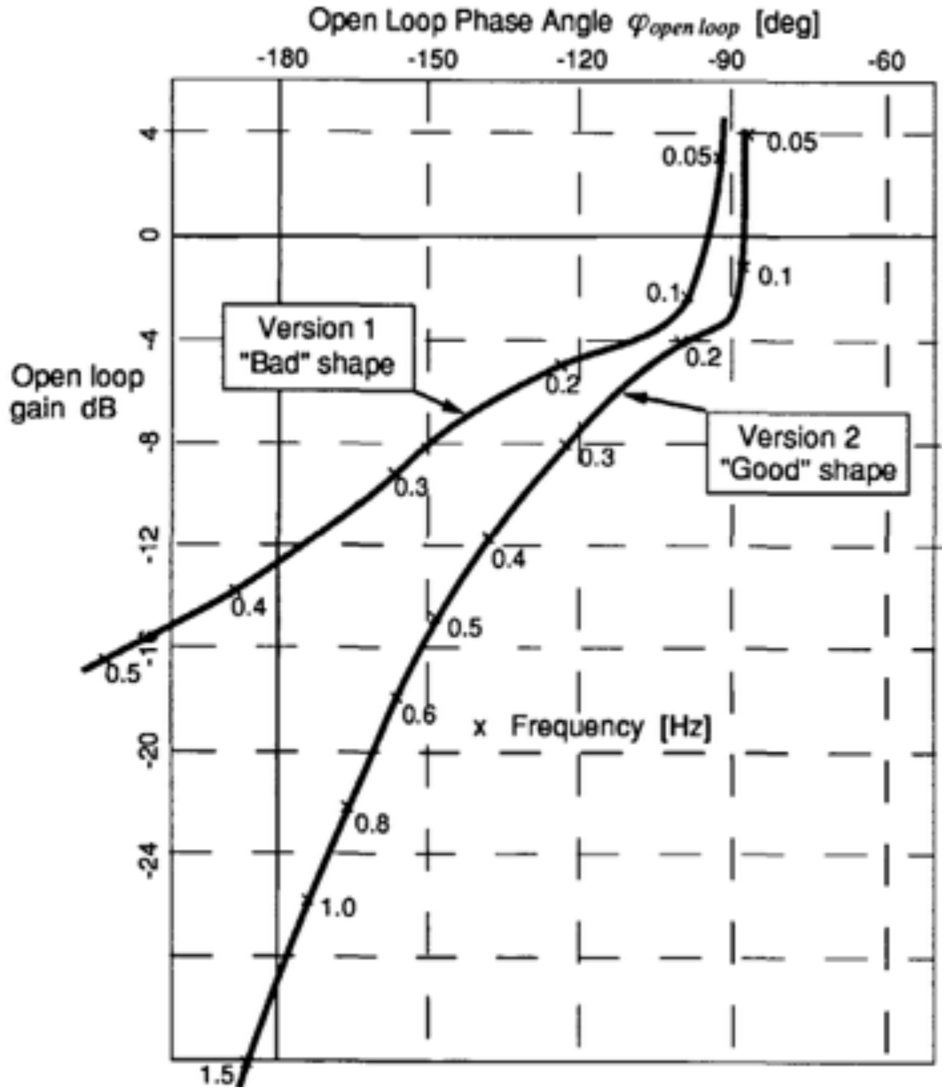
- Visualise prediction, limitations, boundaries or control hints.

Control augmentation

- Remember the BMW 318 with all its abbreviations? Airbus-style, we can make the control of devices easier
  - Inner loop stabilisation
  - Disturbance suppression
  - Control protection
  - Control shaping

Handling qualities

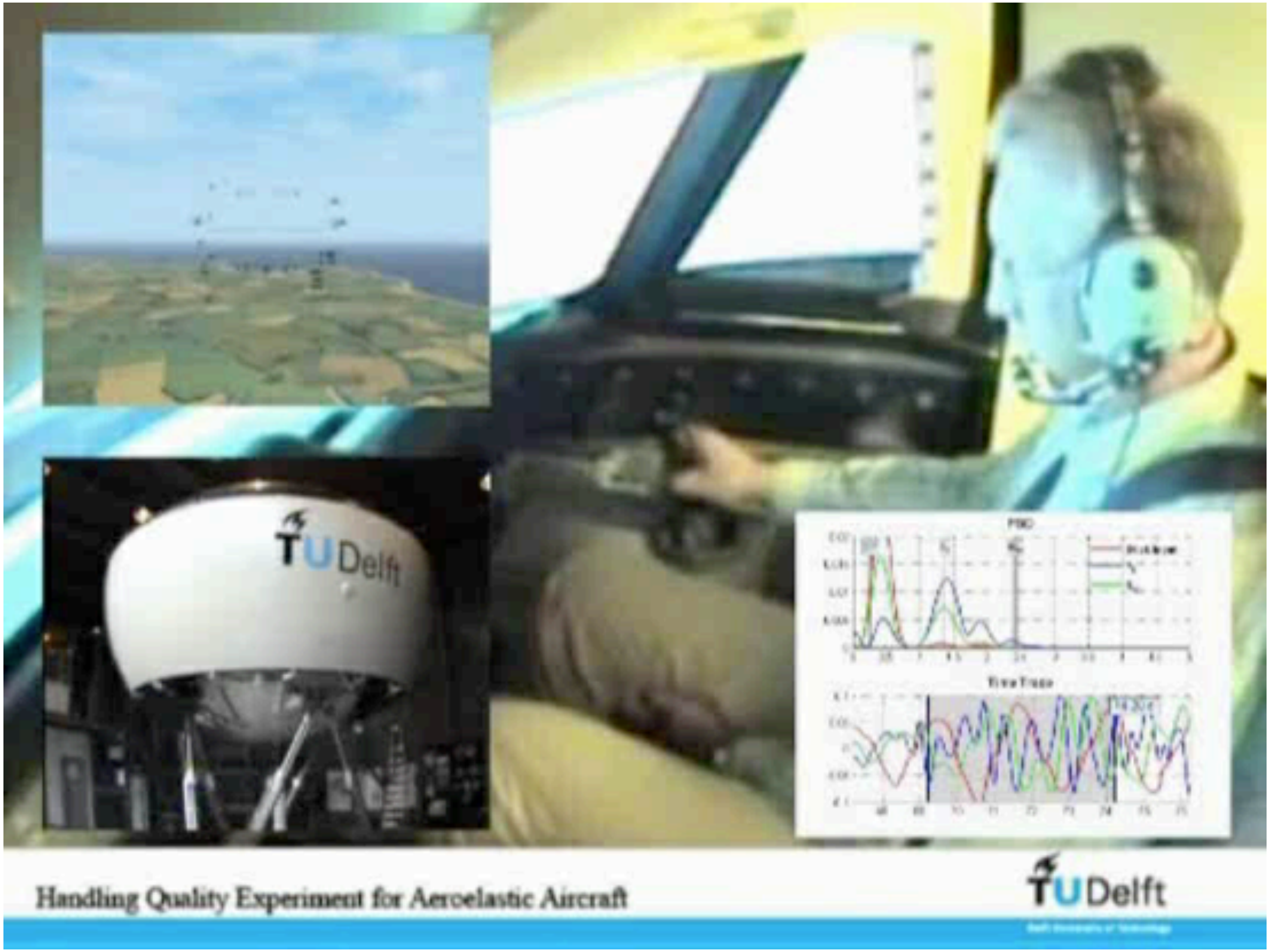
- The resulting device dynamics should be acceptable
  - visible
  - snappy
  - simple & reliable



Conclusion

- Manual control can be enjoyable, adaptive, flexible, and often is the easiest way of performing a non-standard task
- “Pure” manual control is rare; display augmentation and control augmentation lead to a mix of automation and manual control
- Use the “cybernetic approach”, use detected changes in manual control behaviour to evaluate interventions

back-up, cybernetics with a/c flexibility



TUTORIAL: Honest Evaluations of

[Honest evaluations of shared control systems]

Evaluating Shared Control Evaluation

David Abbink

Associate Professor at Delft Haptics Lab
Delft University of Technology

What is a good shared control system?

- Good task performance (what is task performance?)
- Good effort (what is effort?)
- Honest Evaluation: How does this all change in real environments with real people who adapt (including to the support system)

Complicating Factors: influence of support design

Display Augmentation

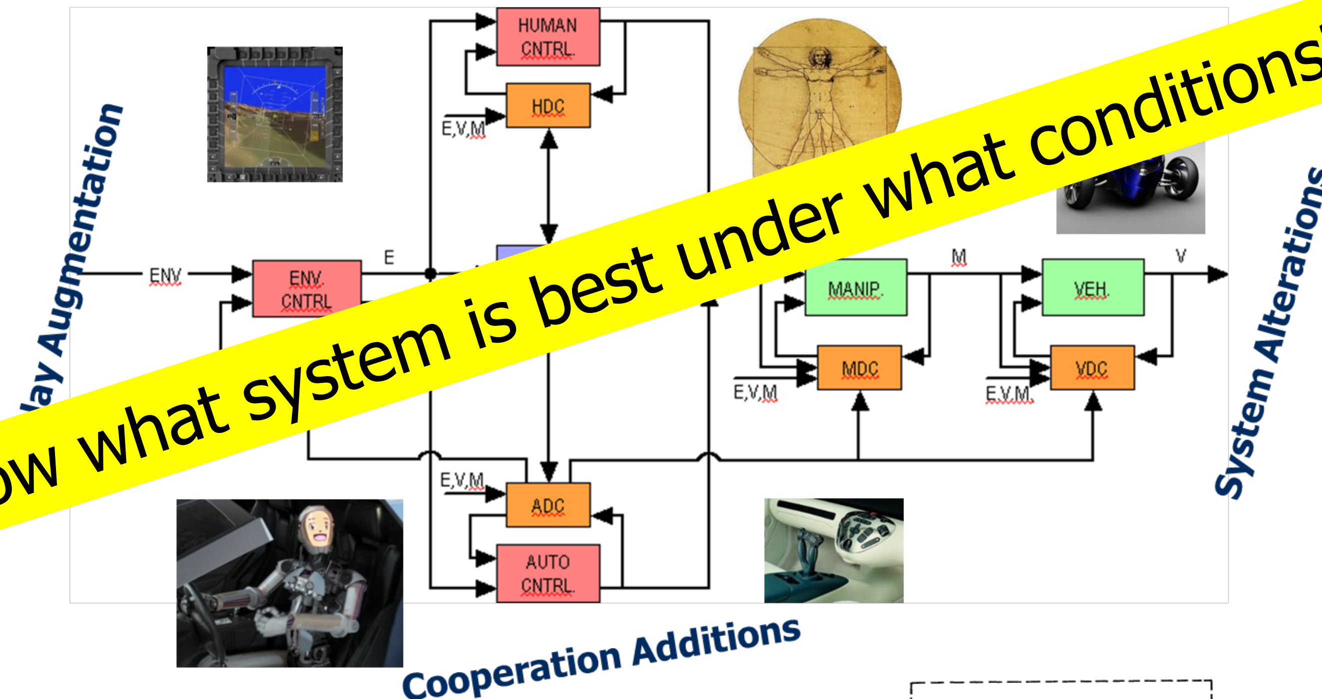
- Aug environment
- Aug state
- Aug system

System Alterations

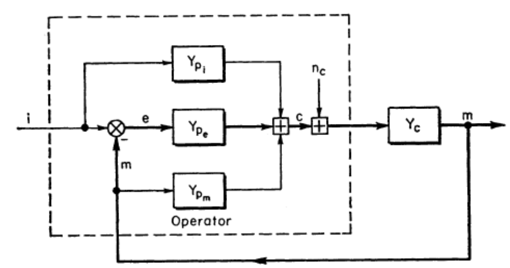
- Alter control
- Alter system
- Alter feedback

C...itions

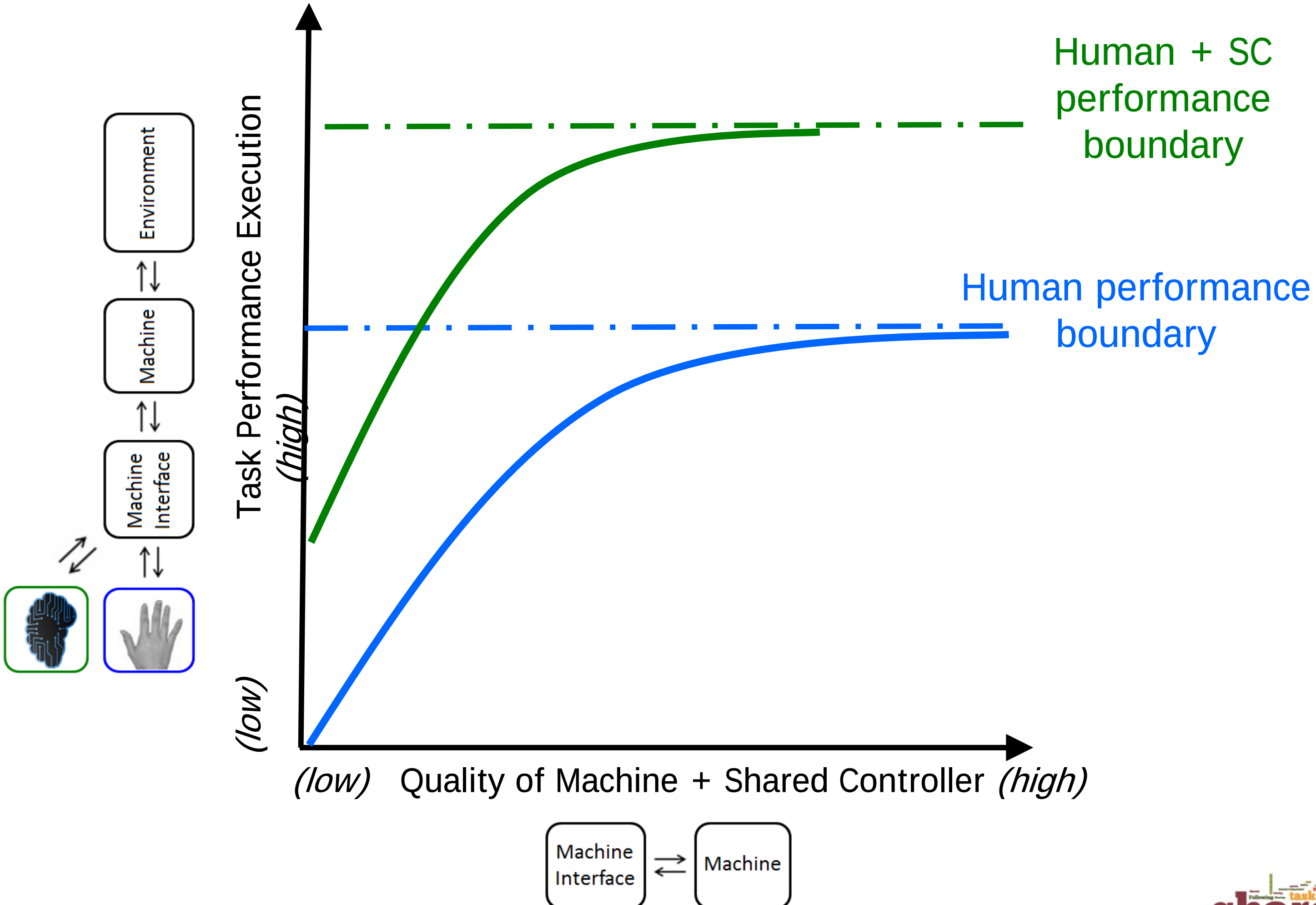
- ...d protection
- Add controller
- Add automation



How do we know what system is best under what conditions?



How to evaluate shared control?



What is relevant for the user?

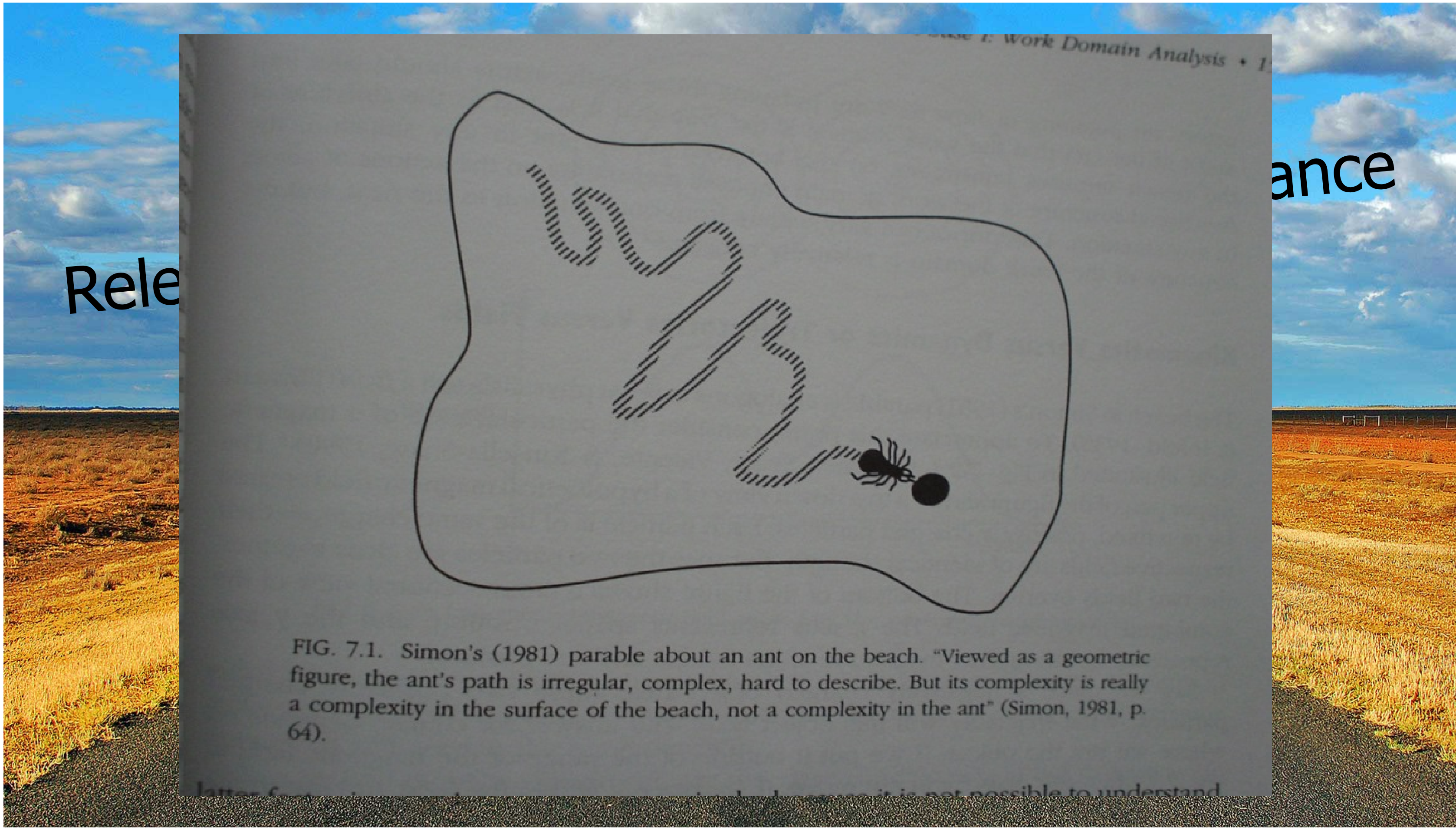


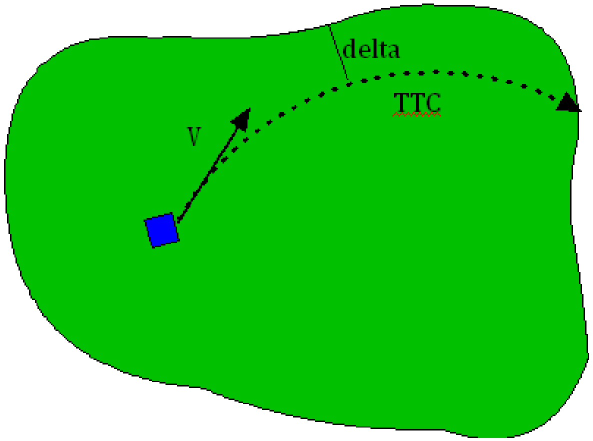
FIG. 7.1. Simon's (1981) parable about an ant on the beach. "Viewed as a geometric figure, the ant's path is irregular, complex, hard to describe. But its complexity is really a complexity in the surface of the beach, not a complexity in the ant" (Simon, 1981, p. 64).

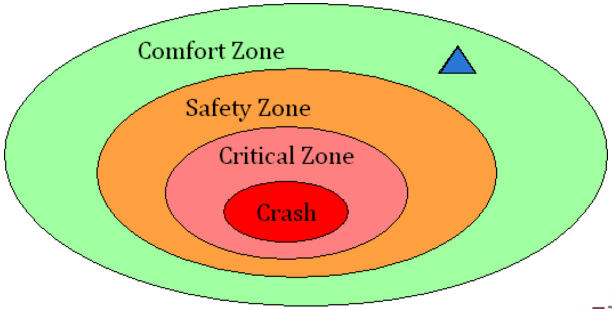
What is performance ??
… tolerance management?

- When task space contains <u>relevant constraints </u>that should not be exceeded
  - Spatio-temporal constraints
  - System dynamic constraints

- Tolerance management performance can be defined as a combination of the
  - Current proximity to these constraints
  - Rate of change in this proximity

- ❑ **Keep state (blue) within established boundaries (in green field).**
- ❑ **Potential Risk and Actual Risk based on V, delta, and TTC.**

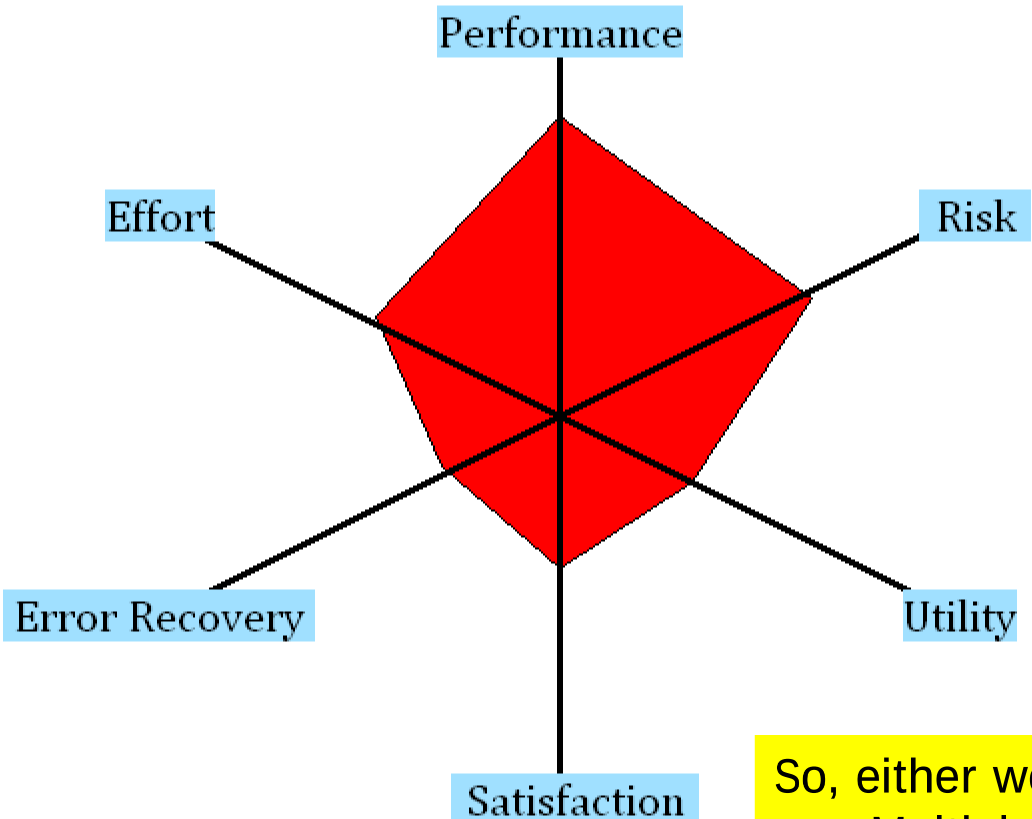


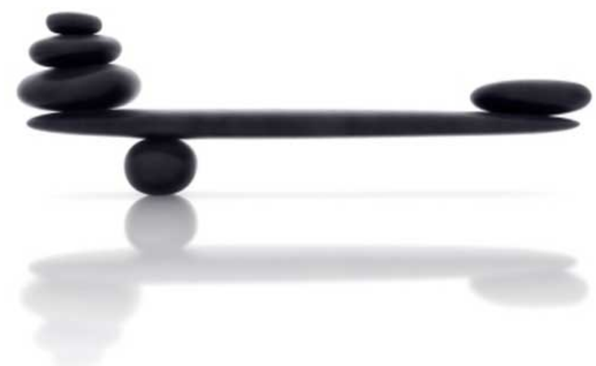


What is effort??

- Control activity
  - **Magnitude, frequency, amount**
    - Statistics of control inputs
  - **Smoothness** (well-matched to dynamics)
    - Steering wheel reversals
  - **Physical Load**
    - Forces on control interface, EMG
- Mental load & visual attention
  - Not too high, not too low
    - Dual tasks
    - Eye-tracking
- Also Subjective!
  - Questionnaires
    - NASA TLX
    - Utility/Satisfaction

Complicating Factors - Trade-offs





So, either we need

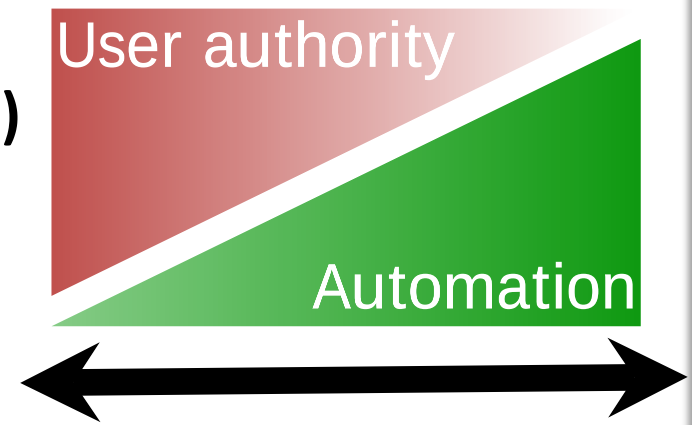
- Multiple good descriptive metrics, or
- a model that ties them together!

Complicating Factors – Variance
(over time & across users and environments)

Tasks change

Spatio-temporal constraints / criticality may change

Shifts in authority may be required



Variance between user (each is different)

- Individual skills and capabilities
- Individual needs and desires
- Individual trade-offs

Variance within individual users

- Individual Tolerances and Trade-offs may change over time
- Attention / motivation / perception may change over time
- Learning / adaptation / skill-loss

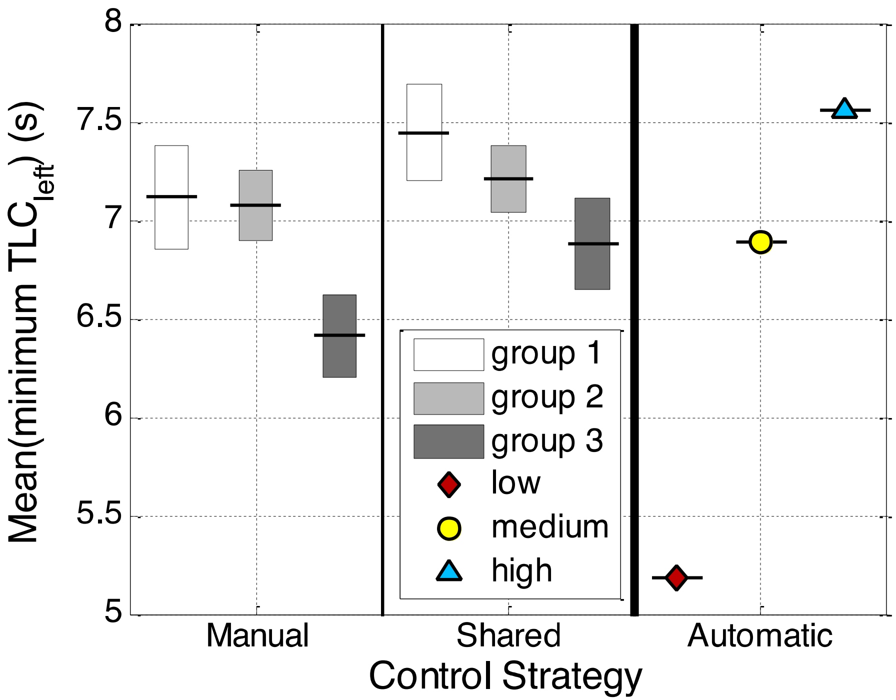
Mulder, Abbink & Boer (2012) - *Sharing Control with Haptics - Seamless Driver Support from Manual to Automatic Control* – Human Factors

Within Design Scope

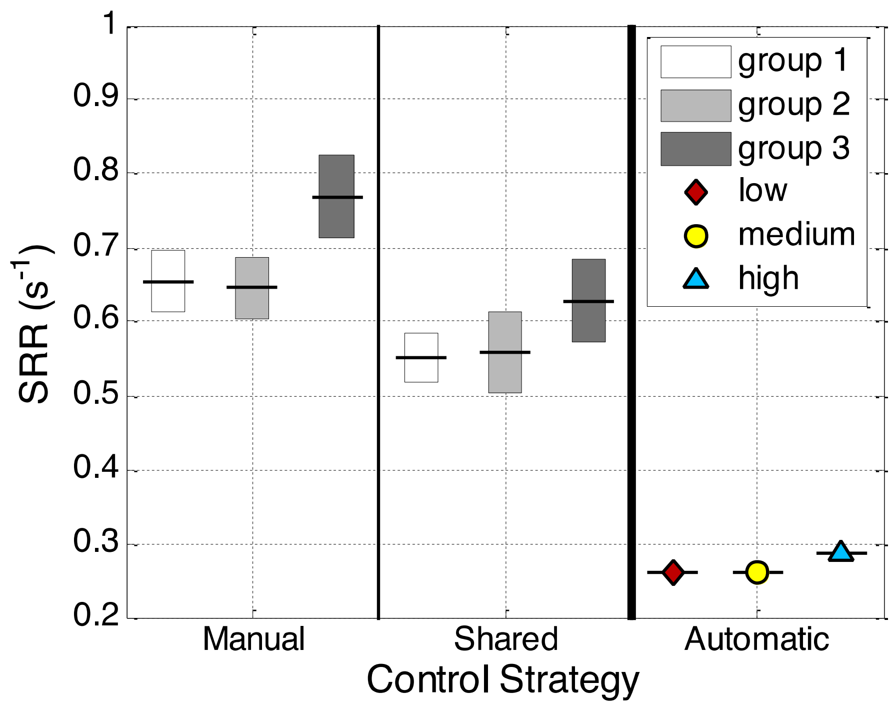
Tested 3 driver groups (from young and unexperienced, to old and experienced), during curve negotiation in a fixed-base driving simulator.

Goal: compare manual control, to shared control, to hands-free driving

Performance increased



Control effort decreased



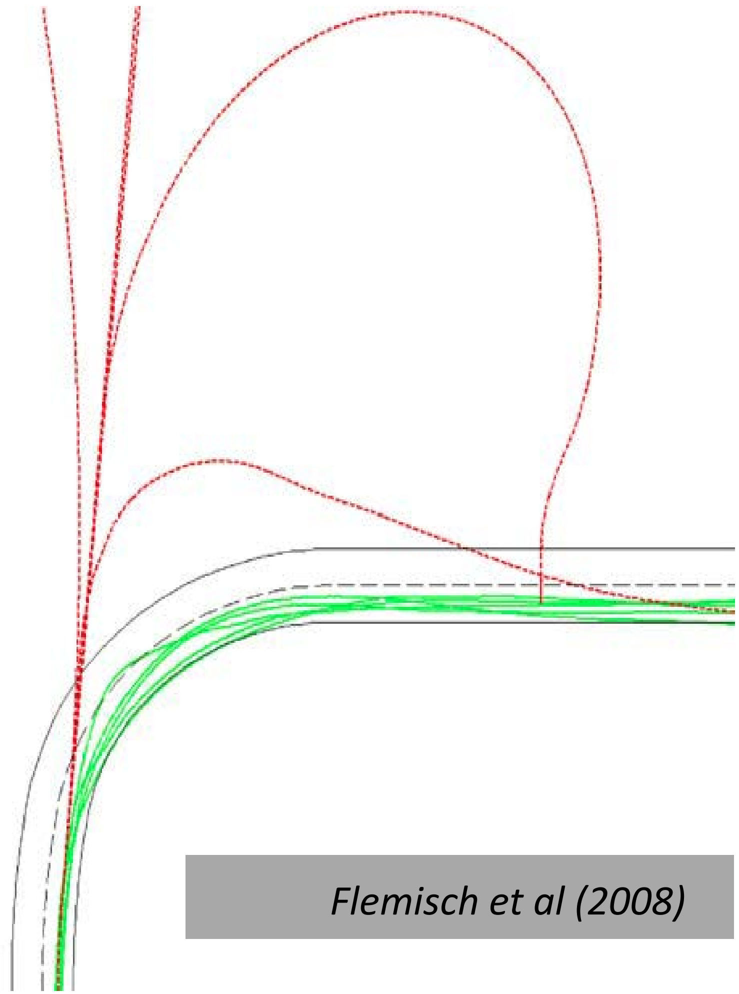
Evaluation outside design scope?

Method: Test automation errors of a curve negotiation support system that would fail just before the onset of a sharp curve

Conditions

with full automation (red lines) that allowed manual override

with haptic shared control (green lines)



Flemisch et al (2008)


Evaluating
the impact of system design
within and outside design scope

So what "fringe of the design scope" was relevant here?

Evaluating the impact of design on behavioral adaptation

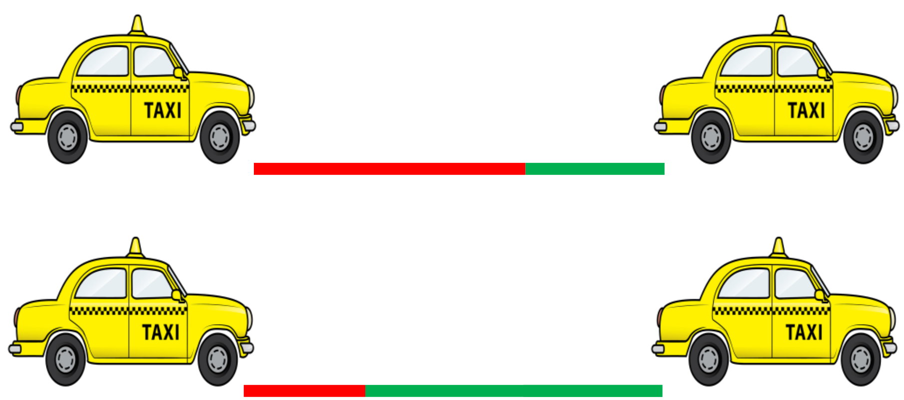
Behavioral Adaptation

Behavioral Adaptation (BA) mitigates the potential safety benefits of an Advanced Driver Assistance System (ADAS).

- Sagberg et al. (1996) showed a reduced headway time among taxi drivers equipped with ABS.

No ABS

With ABS

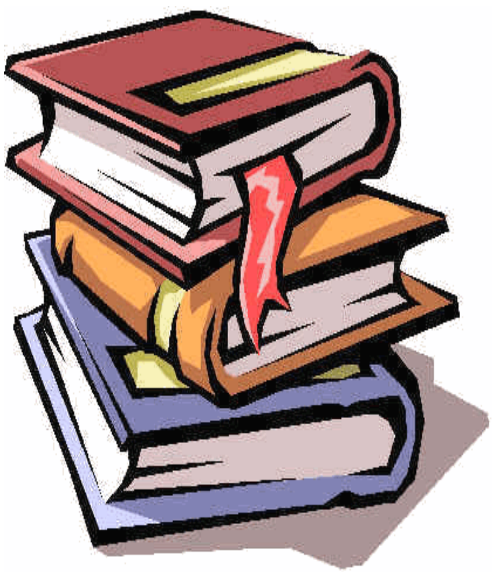


Literature about Behavioral Adaptation

Homeostasis theories in:

- Task difficulty (Fuller, 2005)
- Subjective risk (Naatanen & Summala, 1974; Wilde et al., 1998)
- Time/safety margins (Gibson & Crooks, 1938; van Winsum et al. 1999)

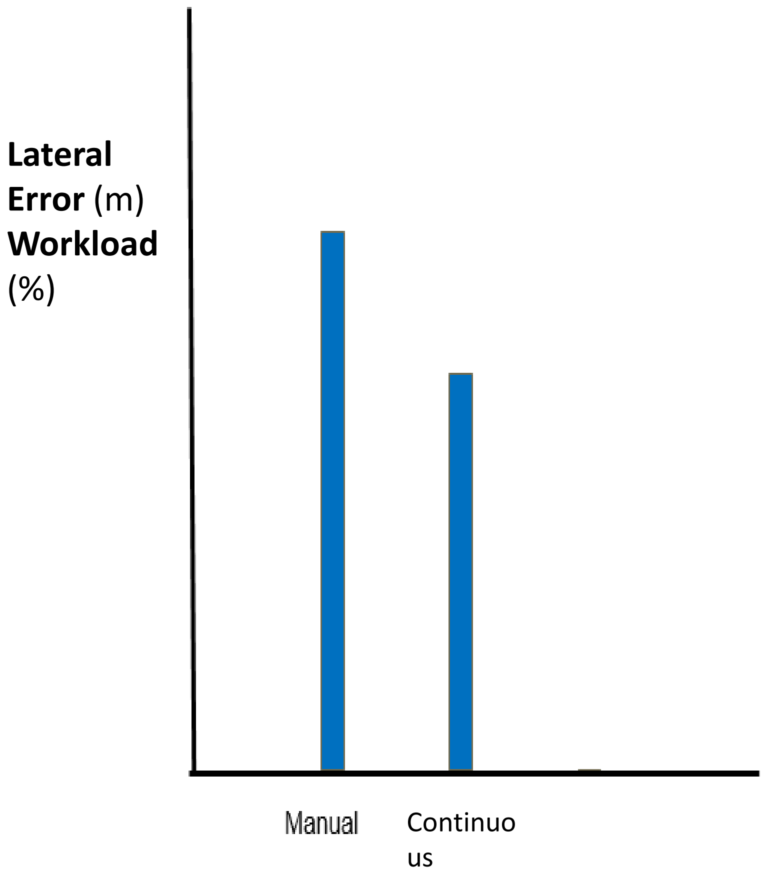
ADAS designers



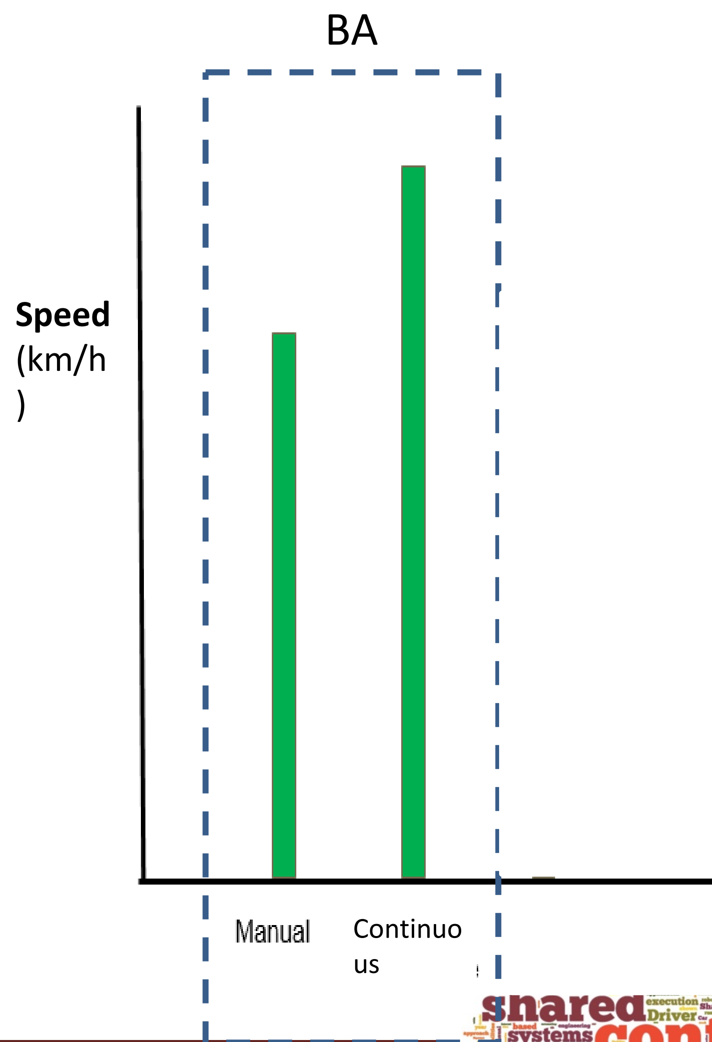
Can we create a more effective support when taking BA into account?

Hypotheses

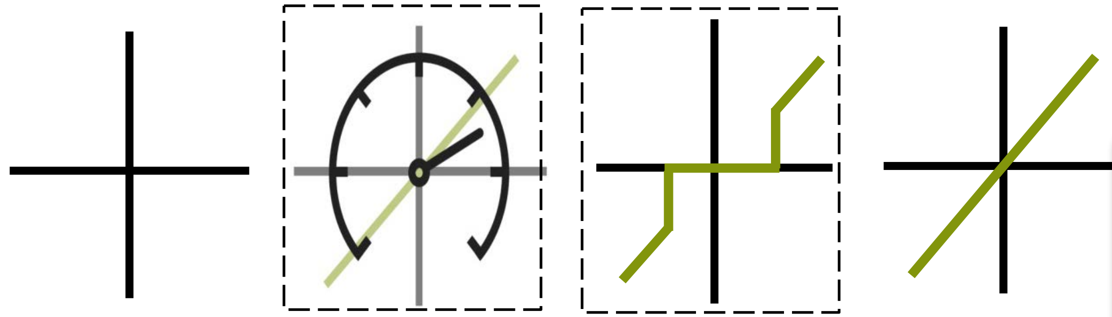
Performance/Workload



Speed



Conditions



Manual　ContRF　Band

Cont

Bandwidth guidance:

- Lateral direction
- Guidance when needed (mostly driving manual)

ContRF guidance:

- Longitudinal direction
- Always guidance except when driving too fast (taking away benefits of speeding)

Results

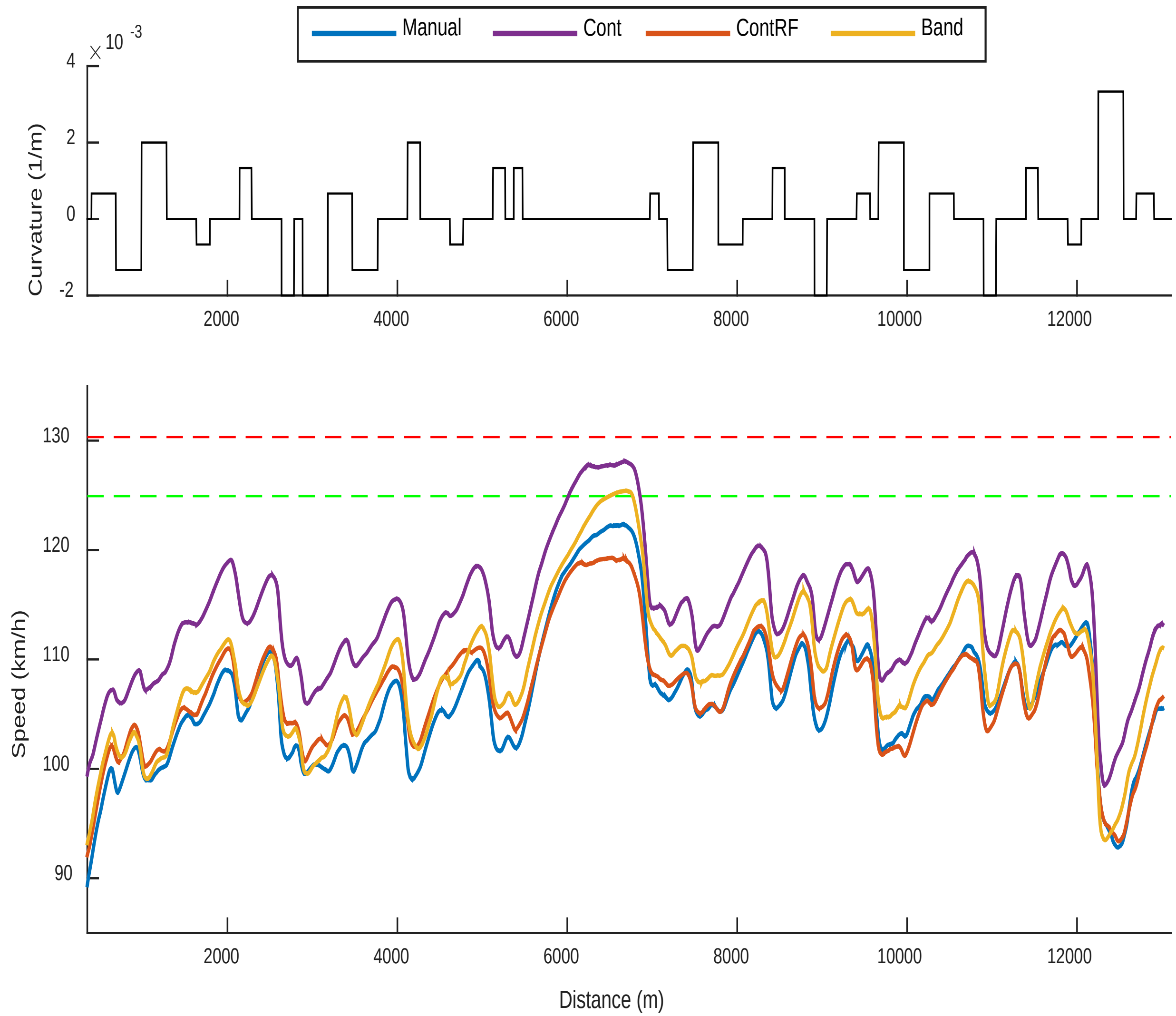
Speed

Performance

Workload

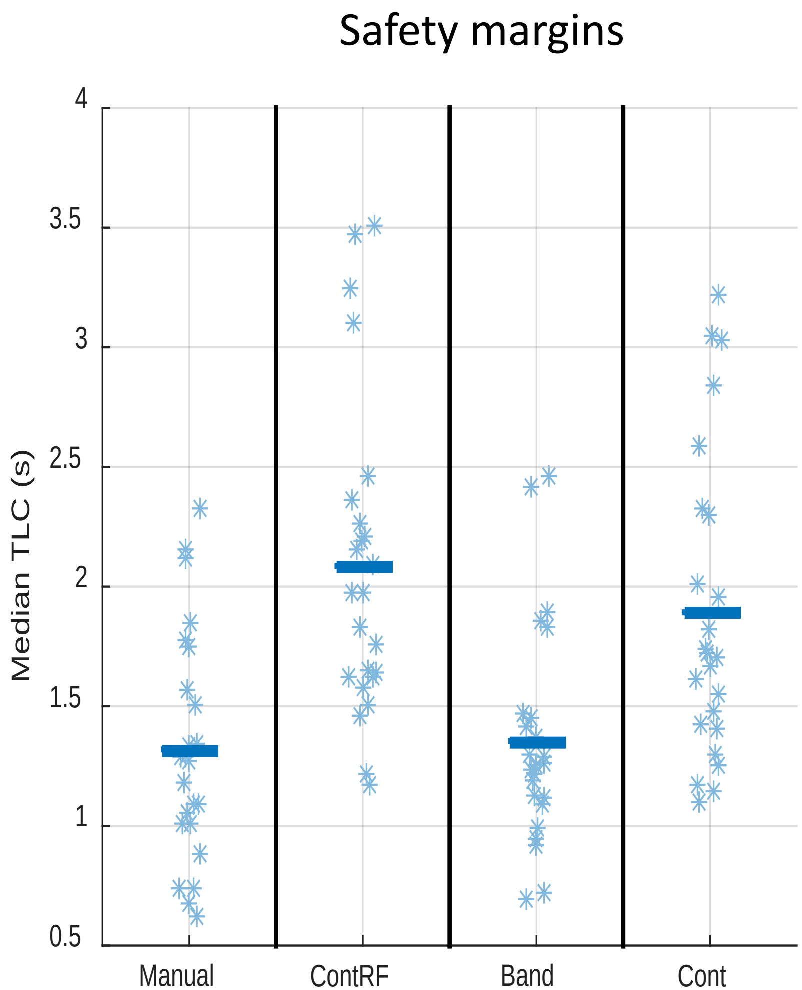
Safety margins

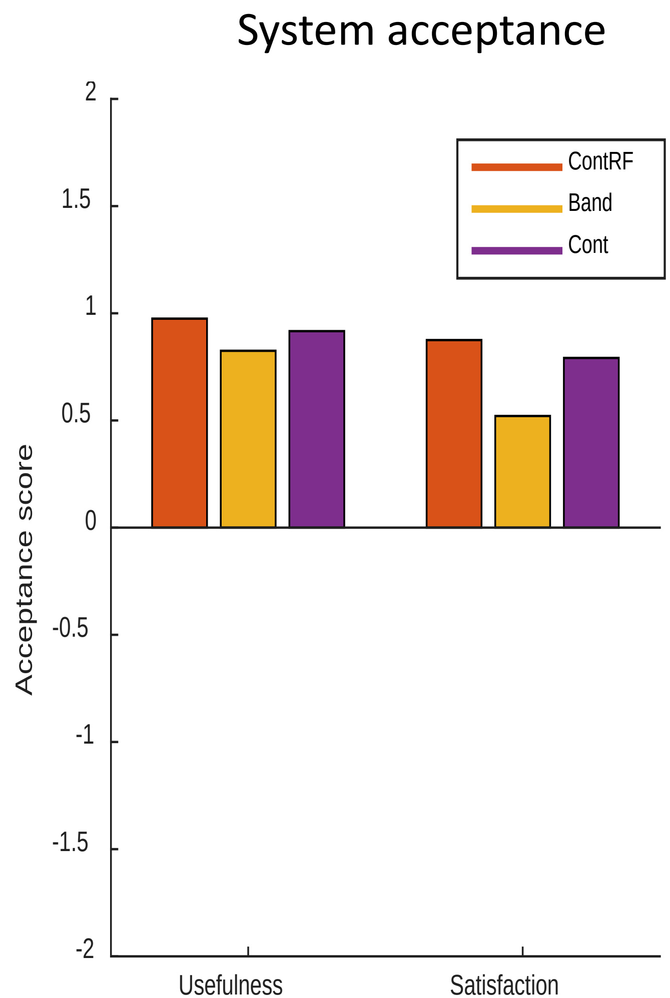
System acceptance

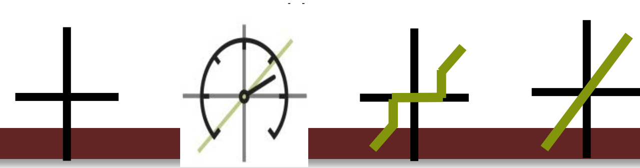


Results

Speed | Performance | Workload | Safety margins | System acceptance







So what fringe of the design scope was relevant here?

Time for interaction!!

TUTORIAL: Designing and Evaluating

Learning from different application domains

Evaluating haptic steering support

Sebastiaan Petermeijer

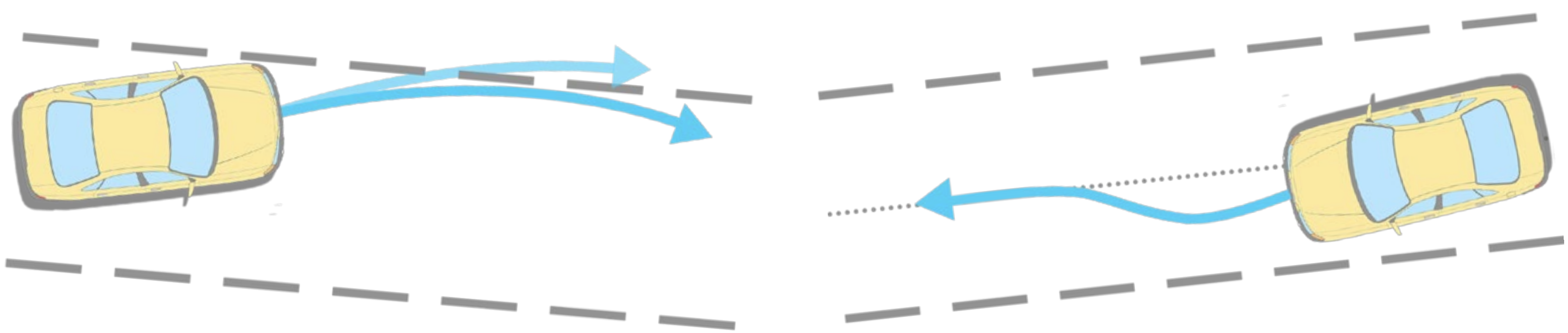
PhD at the Institute of Ergonomics
Technical University Munich

Should drivers be operating within a automation-free bandwidth?

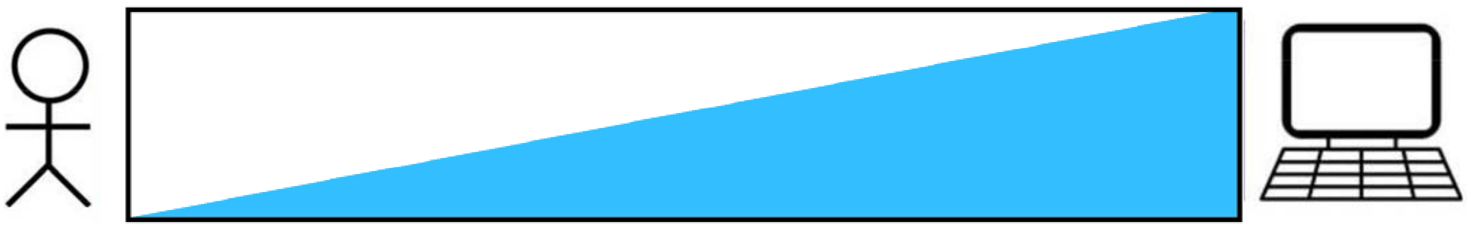
Evaluating haptic steering support systems with different levels of authority

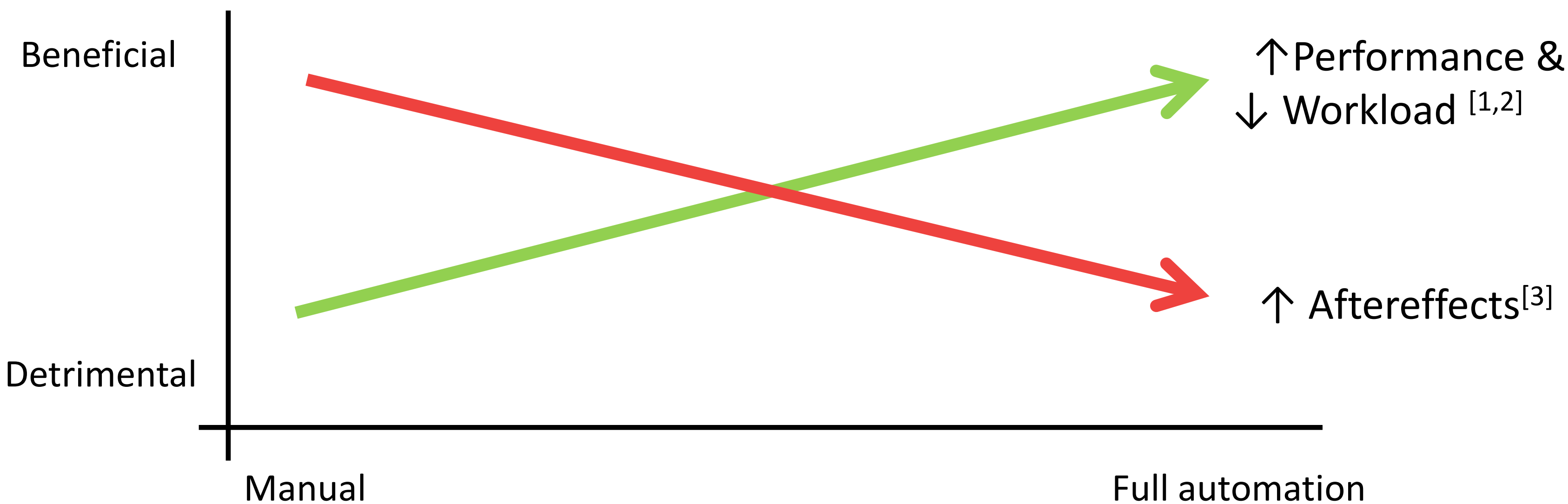
S.M. Petermeijer, D.A. Abbink, & J.C.F de Winter

2014 Human Factors Prize Winner



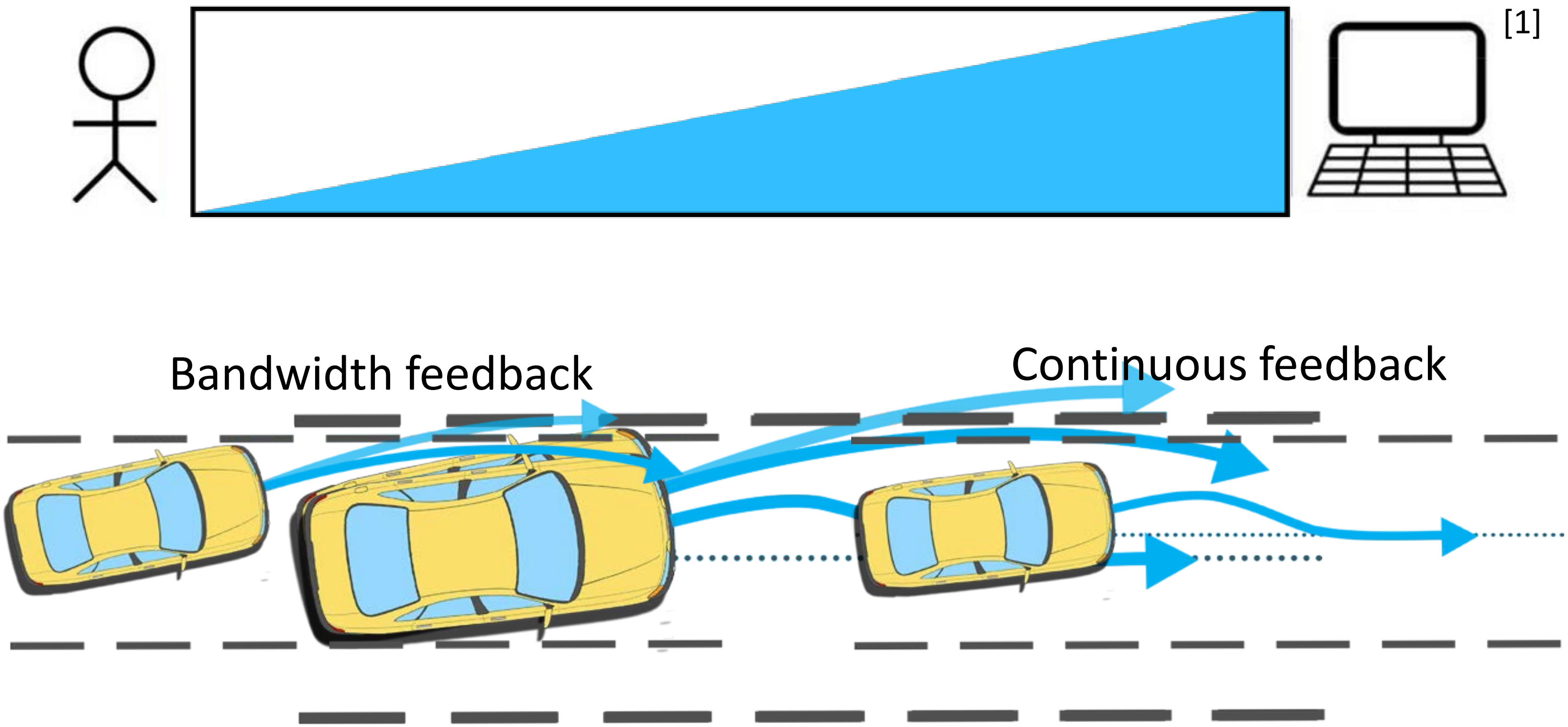
Benefits/cost trade off





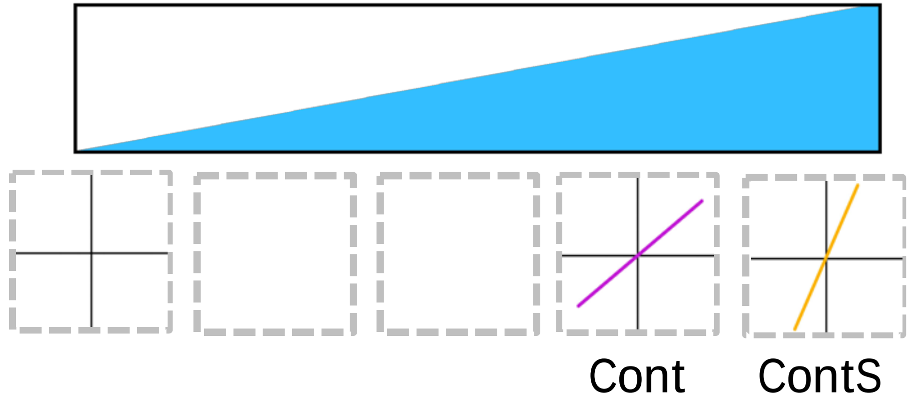
[1] Flemisch et al. 2008, [2] Mulder et al. 2012, [3] De Winter & Dodou 2011

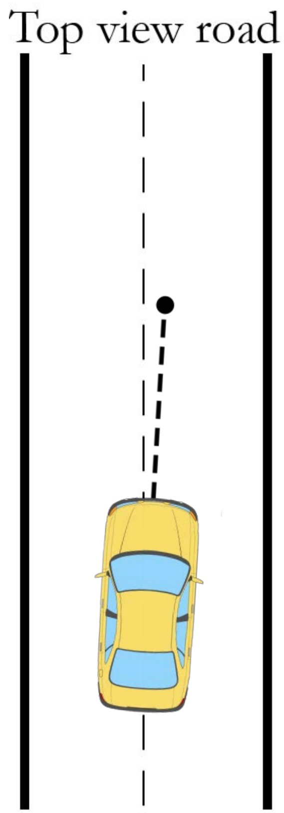
Background



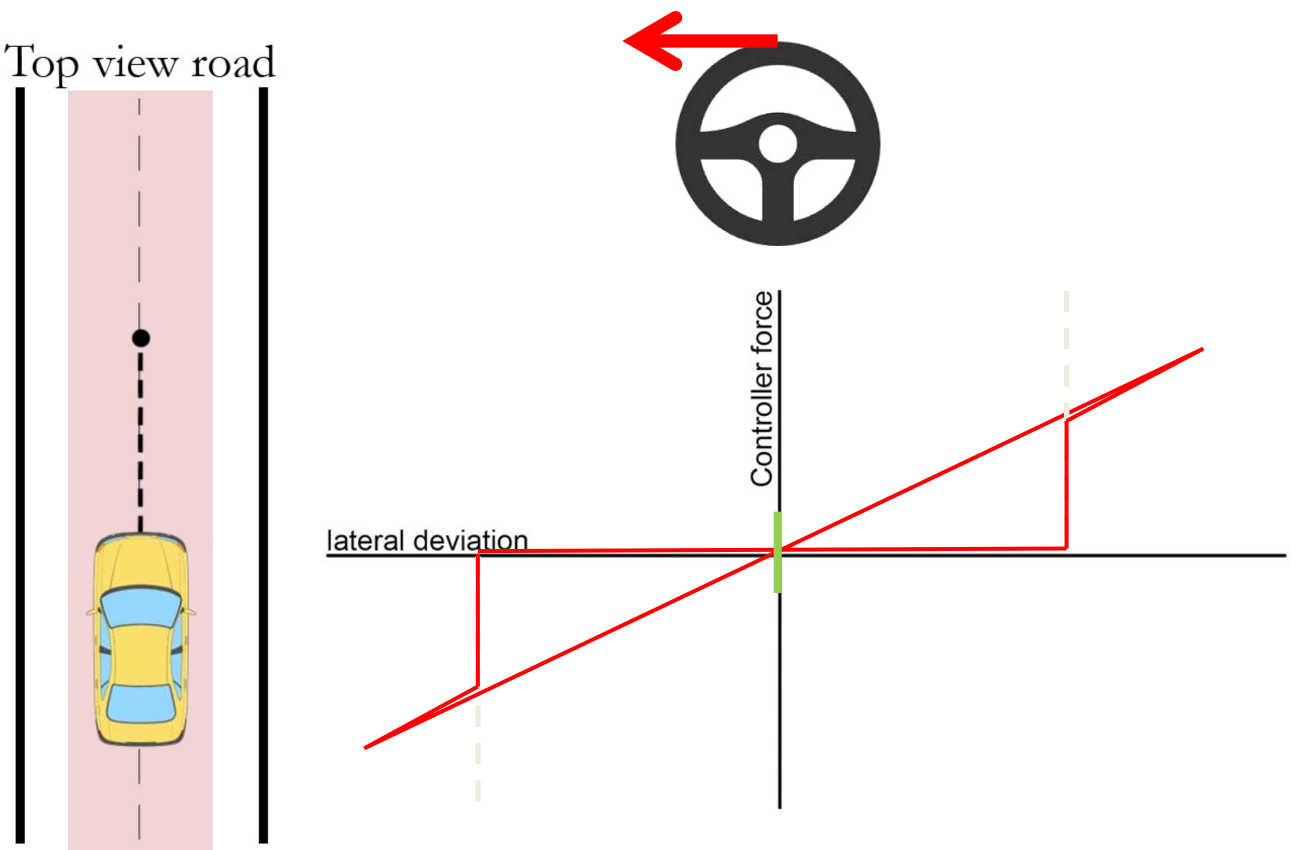
[1] Adapted from Flemisch et al. 2008

Continuous

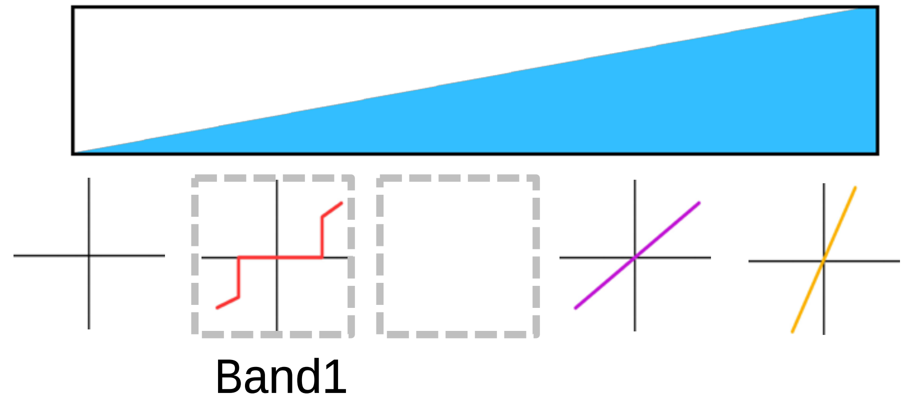


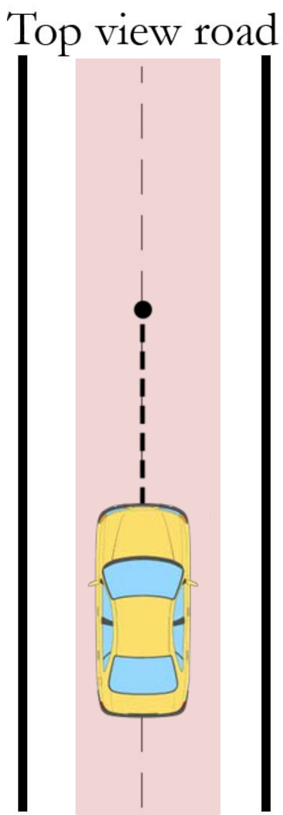


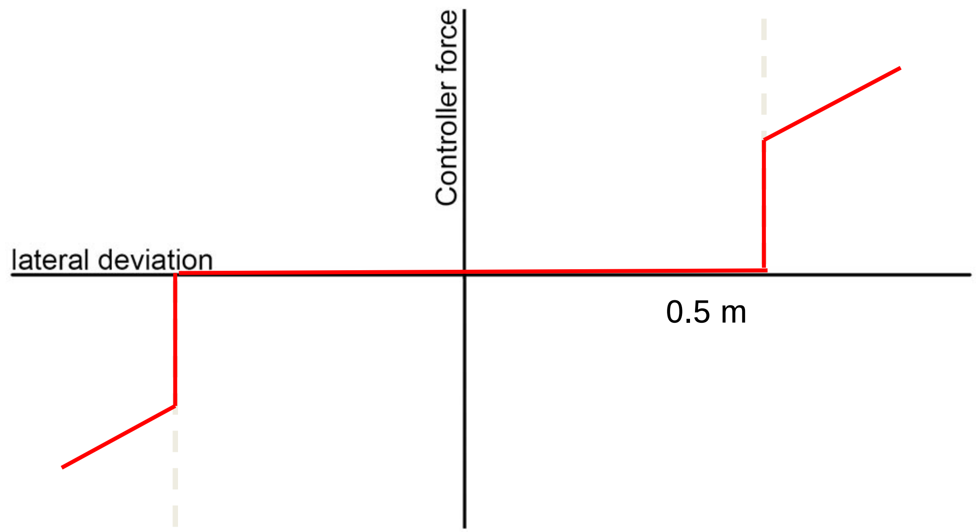
Bandwidth



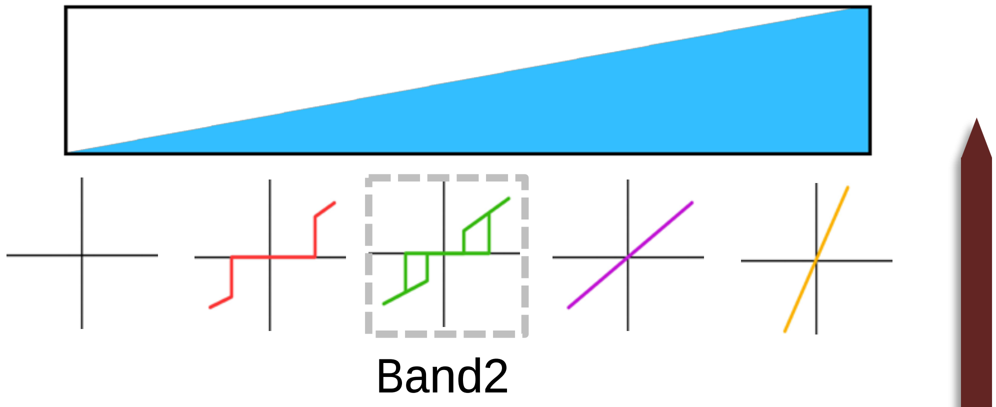
Bandwidth

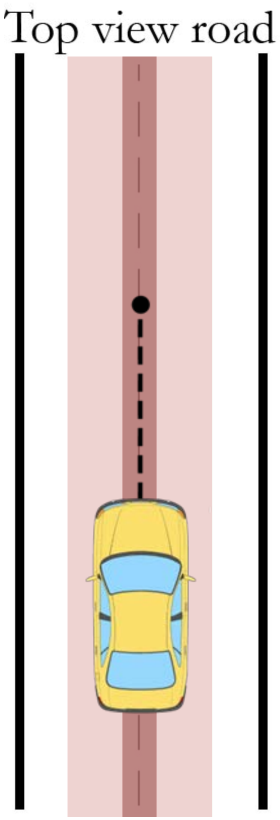

Band1

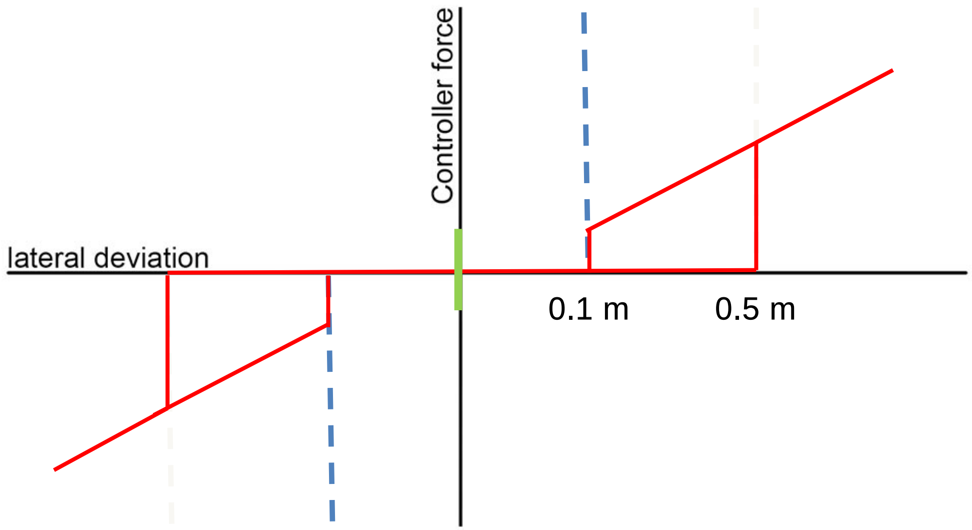

Top view road


Controller force
lateral deviation
0.5 m

Bandwidth

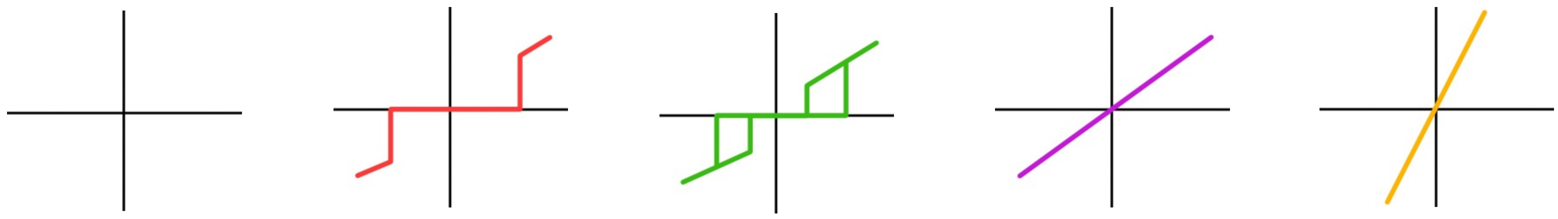

Band2


Top view road

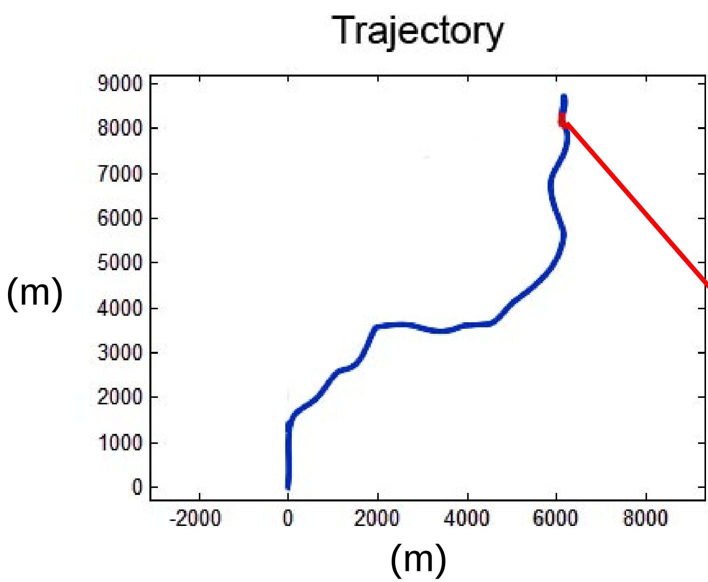

Controller force
lateral deviation
0.1 m
0.5 m

Method

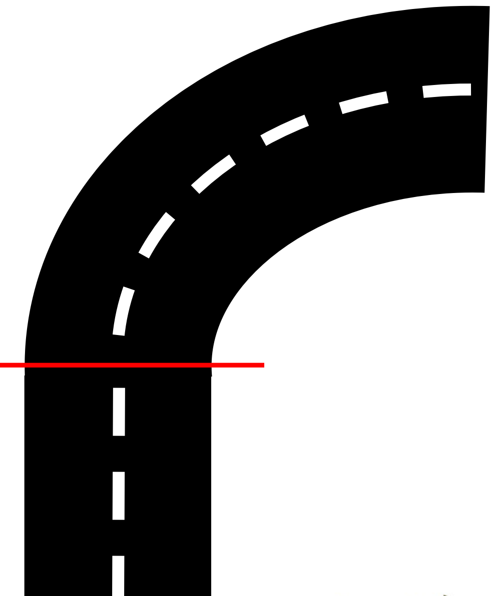
Condition 1	Condition 2	Condition 3	Condition 4	Condition 5



Manual	Band1	Band2	Cont	ContS



System shutdown



Results

Performance (within the design scope)

- Distribution of the lateral position
- Mean absolute lateral position

Workload

- NASA-TLX score
- Reaction time
- Mean absolute driver torque

Performance (on the fringe)

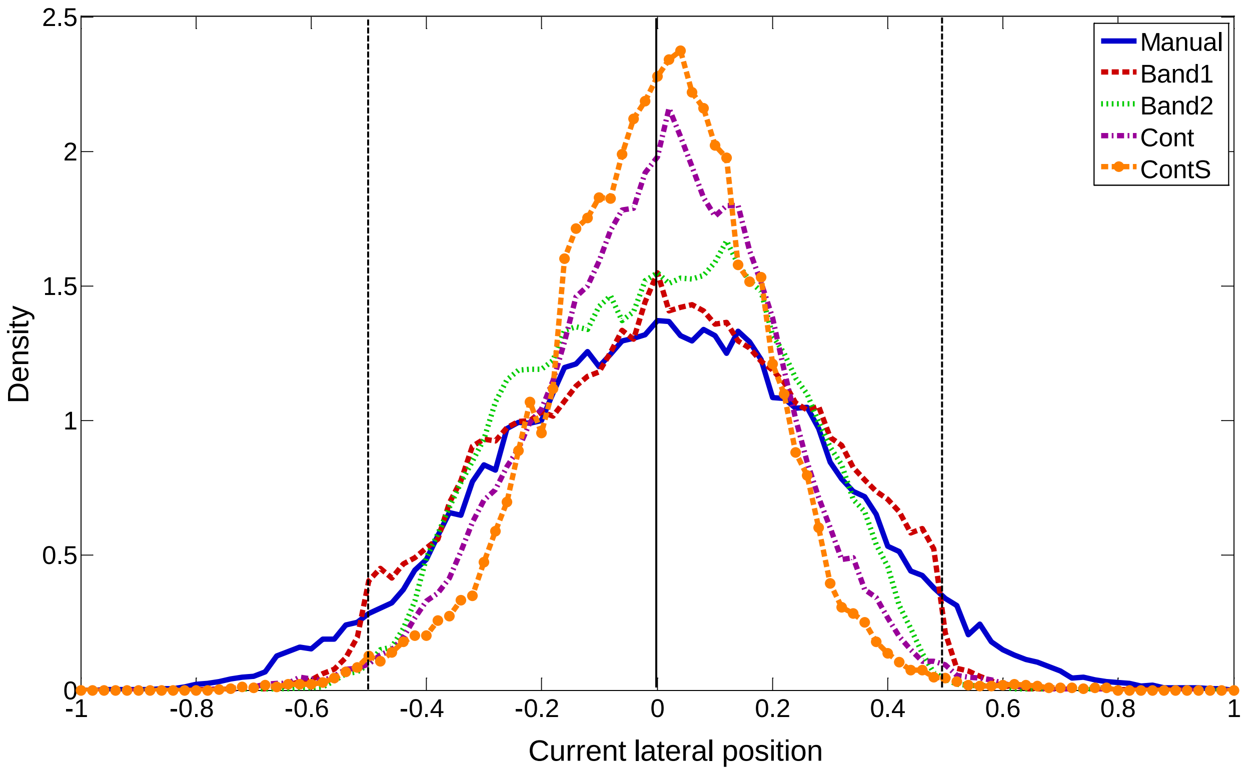
- Max absolute lateral position
- SD of the lateral position

System acceptance

- Satisfaction score

Results

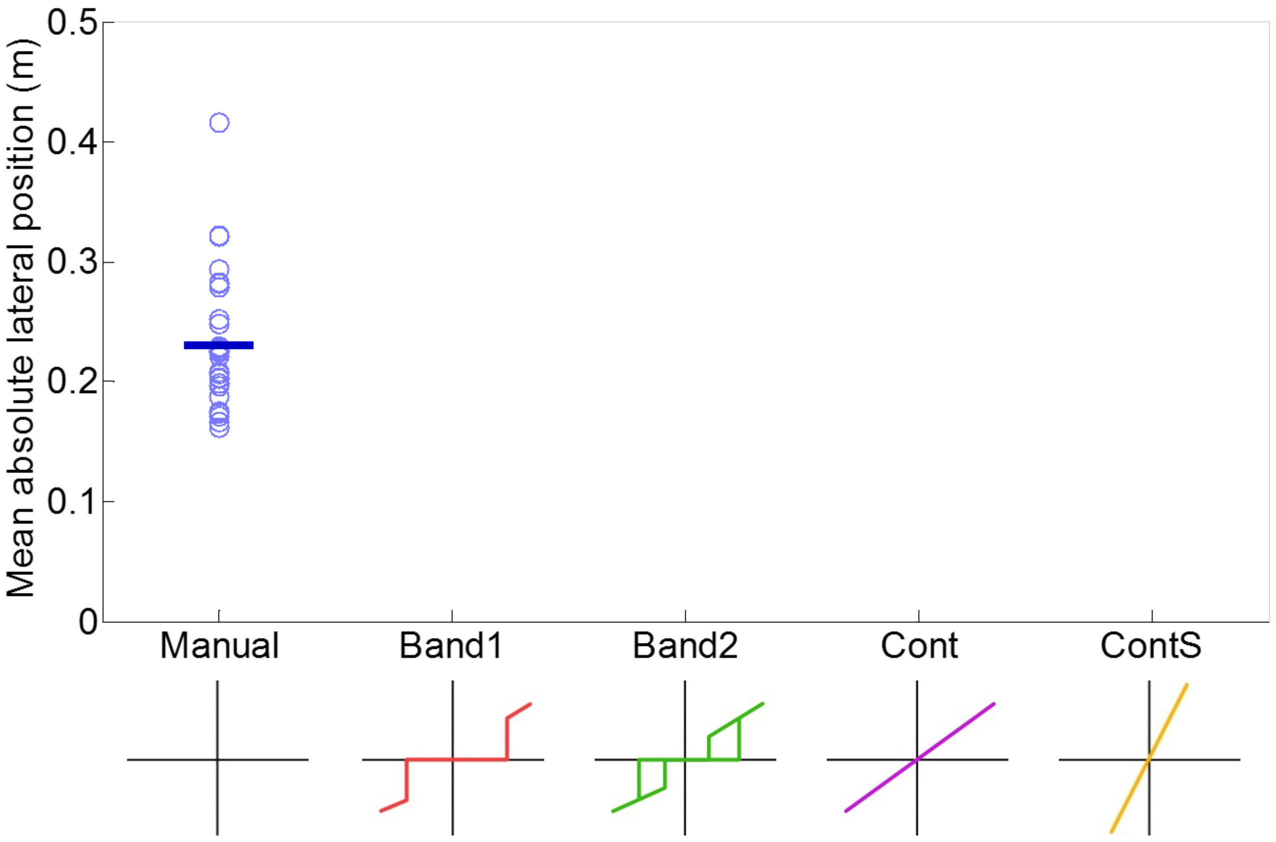
Distribution lateral position



- Bandwidth feedback prevents large lateral errors
- Continuous feedback yields better performance

Results

Mean absolute lateral position



- Performance increases with more feedback
- Confirms results from the literature

$p = 1.71 \cdot 10^{-27}$

Background – Introduction – Method – **Results** – Conclusions

Results

Performance (within the design scope)

- Distribution of the lateral position
- Mean absolute lateral position

Workload

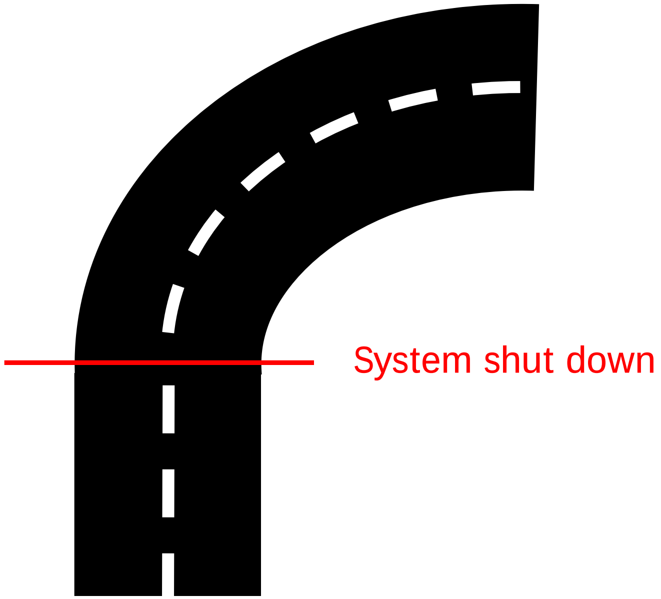
- NASA-TLX score
- Reaction time
- Mean absolute driver torque

Performance (on the fringe)

- Max absolute lateral position
- SD of the lateral position

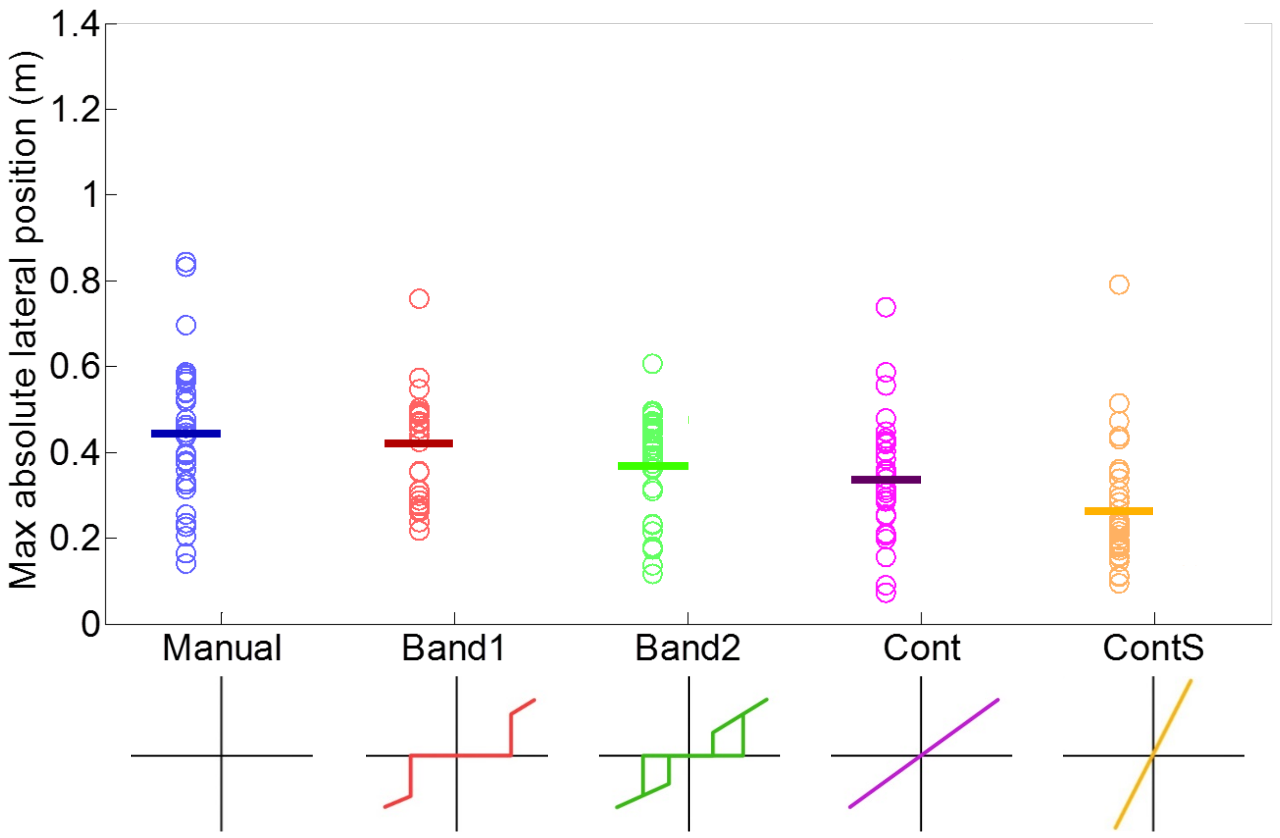
System acceptance

- Satisfaction score



Results

Maximum lateral position – comp/shutdown

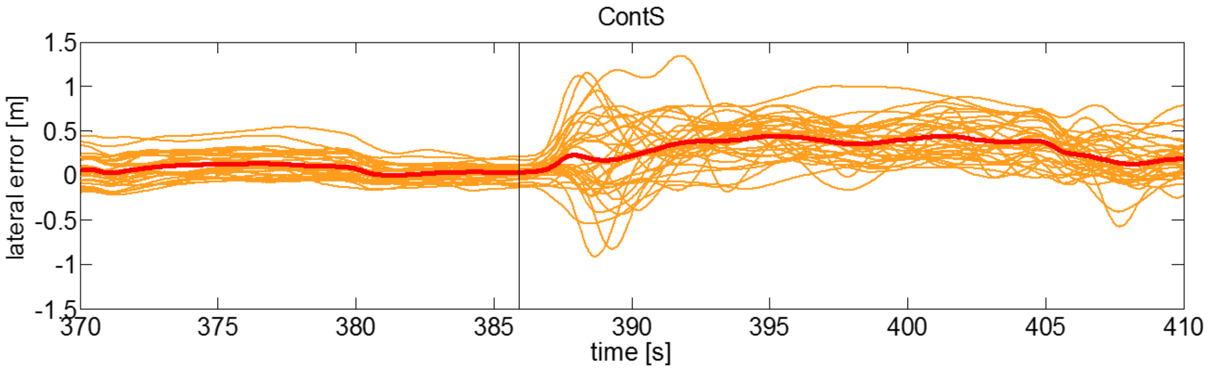


- Continuous feedback yields aftereffects
- ContS is significantly higher than manual and bandwidth

Comp: $p = 1.16*10^{-12}$
Shutdown: $p = 5.04*10^{-7}$

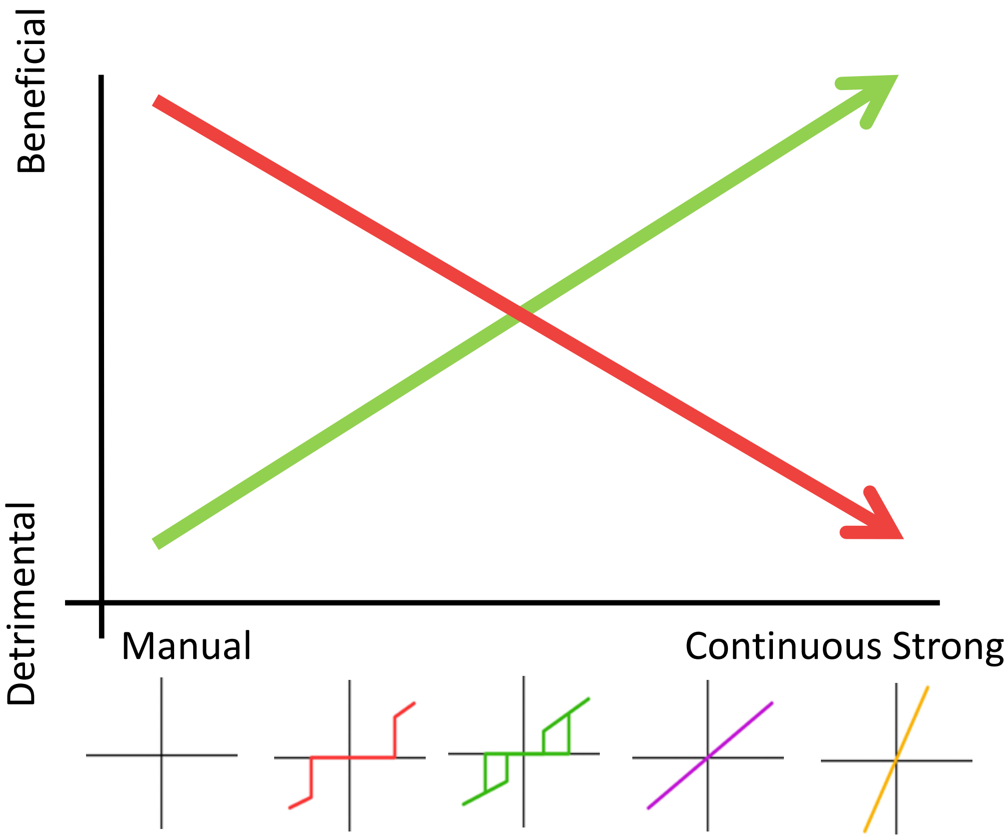
Results

Raw data of the lateral position



- Some driver shoot into the curve
- Some drivers overshoot the curve
- Bandwidth feedback does not show aftereffects

Conclusions



Trade off between continuous and bandwidth feedback:

- \+ Better performance
- \+ More satisfaction
- \+ Less perceived workload

- \- Resulted in aftereffects
- \- Large individual differences

TUTORIAL:
Honest Evaluations of Shared Human-Machine Control Systems

Building Evaluation Taxonomy

Erwin R. Boer

Associate Professor, Mechanical Engineering
Delft University of Technology

Fundamental Differences between Evaluations across Levels of Support

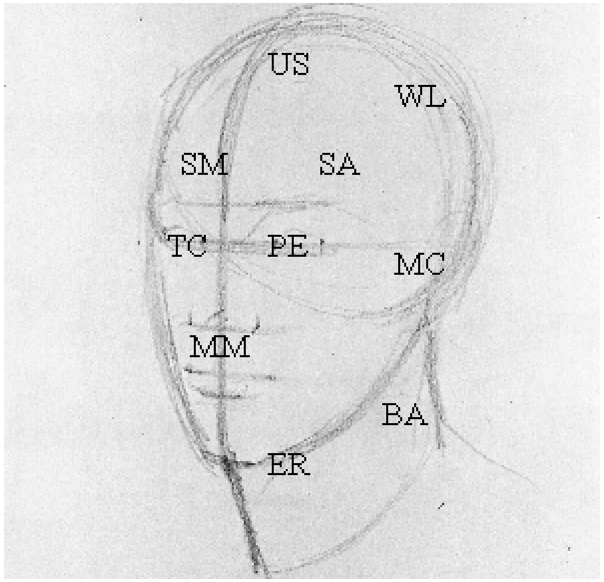
Support Type	Support Domain	Perf. Effort	Event Recovery	Error Recovery	Transition Sys2Hum	Human Operator Role	
						Percep. Motor Control	System Limitation Assessment
Manual		X	X			X	
Shared Control	*Task*	X	X	X	X	X	
	Situation	X	X	X	X	X	
Part Automation	*Task*	X	X	X	X (switching)	X	X
	Situation			X	X (switching)	X	X
Full Automation	*Situation*			X	X (switching)		X
	Always						

Often jerky, odd to road users, or even frightening to driver. Acceptance of control & decision making.

Challenging HMI issues for communicating system reliability and human role.

Experimental Paradigms

WL:	Subjective Workload
PE:	Performance and Effort
MM:	Mental Model
SA:	Situation Awareness
BA:	Behavioral Adaptation
ER:	Error detection and Recovery
TC:	Trust and Complacency
US:	Usability and Satisfaction
MC:	Model Coefficients
SM:	Safety Margins and set points



How are these 10 Human Factors issues intertwined?

What have we heard about – what are we missing?

System <u>Acceptance</u> evaluation on two dimensions
a *Usefulness* scale and an affective *Satisfying* scale

1 Useful |__|__|__|__|__| Useless
2 Pleasant |__|__|__|__|__| Unpleasent
3 Bad |__|__|__|__|__| Good
4 Nice |__|__|__|__|__| Annoying
5 Effective |__|__|__|__|__| Superfluous
6 Irritating |__|__|__|__|__| Likeable
7 Assisting |__|__|__|__|__| Worthless
8 Undesirable |__|__|__|__|__| Desirable
9 Raising Alertness |__|__|__|__|__| Sleep-inducing

Van der Laan, J.D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. Transportation Research - Part C: Emerging Technologies, 5, 1-10.

Honest Evaluation's 3 Pillars DE=F^3

- Fringe Cases
  - Cases that expose system limitations
- Frequency of Occurrence
  - Occurrence of fringe cases across drives
- Fractures
  - Fractures that fringe cases can cause (fatal, harmful to human, damage to property)

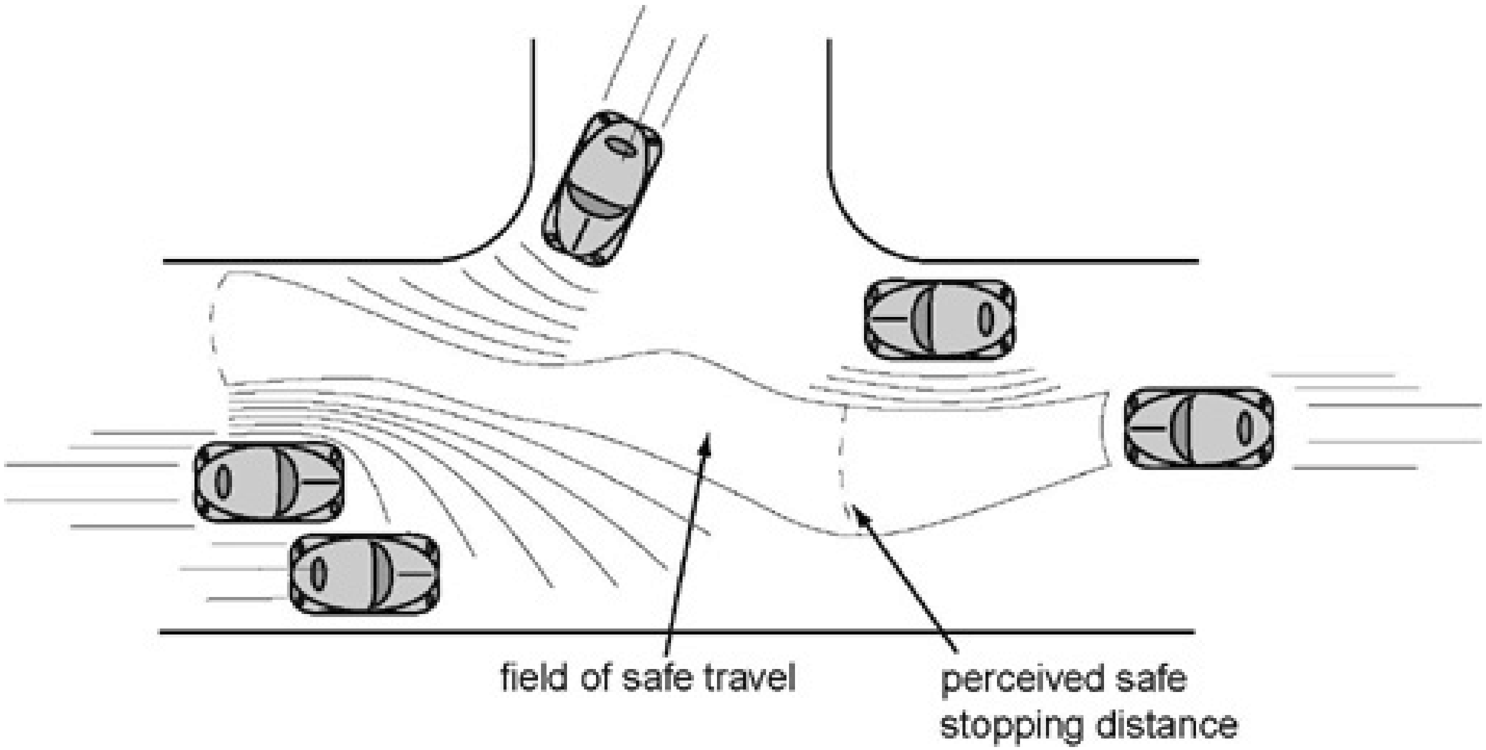
→ Human as a backup.

→ Limit situations.

→ Global safety case/statistics.

HMI Designs

1938 Field of Safe Travel – System Requirement

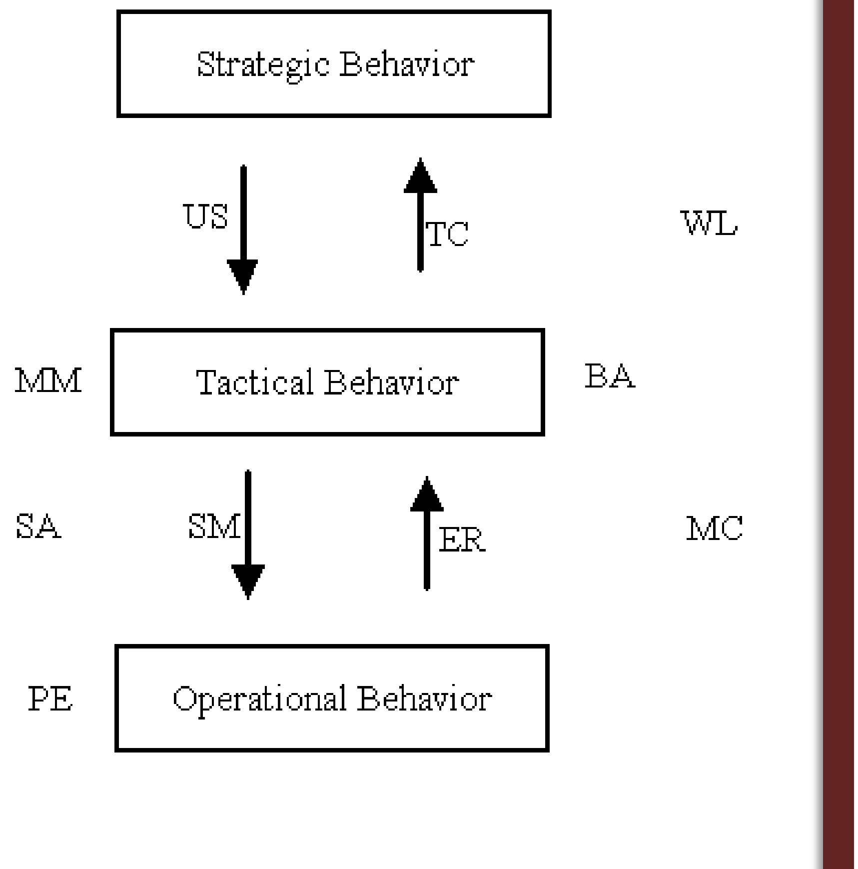


Understand the actions that are possible – Scene interpretation to affordance.

Interaction with Support

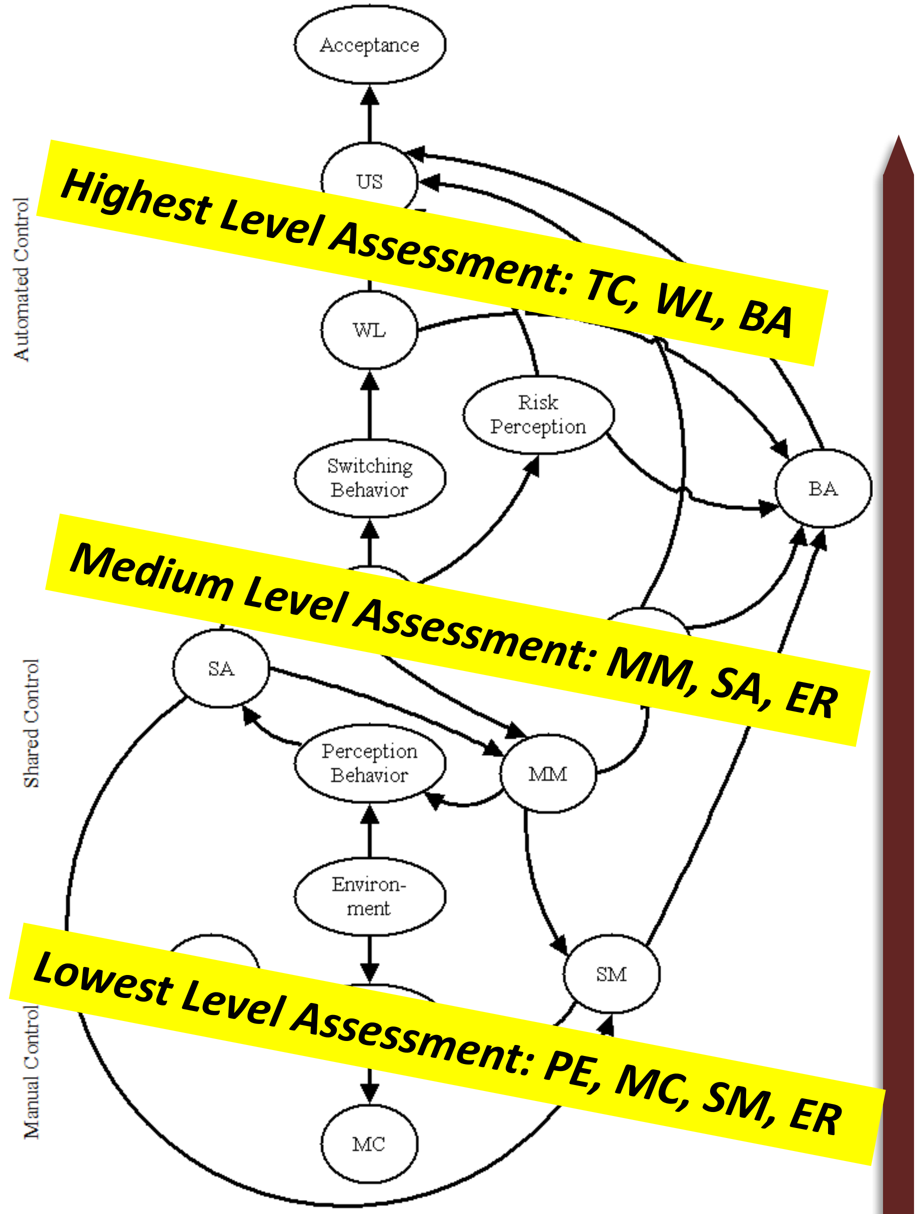
WL: Subjective Workload
PE: Performance and Effort
MM: Mental Model
SA: Situation Awareness
BA: Behavioral Adaptation
ER: Error detection and Recovery
TC: Trust and Complacency
US: Usability and Satisfaction
MC: Model Coefficients
SM: Safety Margins and set points

- At strategic level decide whether to use and interact with system (US).
- This depends on trust (TC) in the system and workload (WL) to interact with is.
- Trust in the system depends on how frequently "errors" or "fringe cases" occur (ER) that are unpredictable or hard to learn (MM).
- These fringe cases require the human operator to maintain situation awareness (SA) and be ready to respond causing an increased workload (WL).
- If human operator decide to use it (US), then depending on performance effort (PE) balance, the human operator may adapt safety margins (SM) and thus engage in behavioral adaptation (BA).
- Behavioral adaptation can also occur at lower levels where it shows up in change in model coefficients (MC) and/or safety margins (SM).



Relating Experimental Paradigms

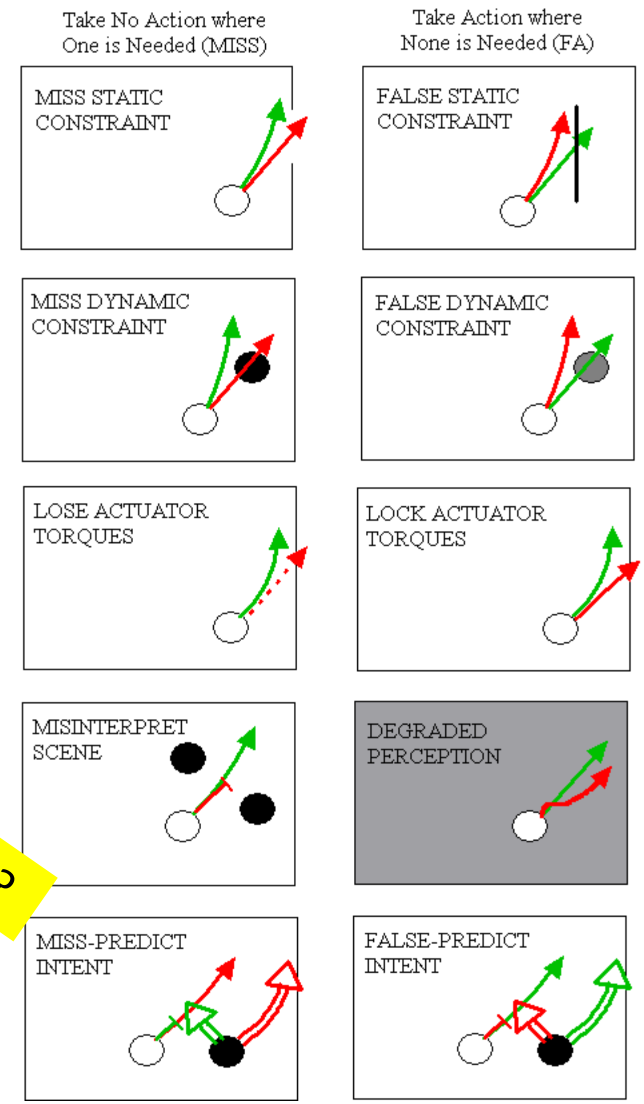
- Usability and satisfaction (US) depend on required workload (WL), experienced risk, whether the system can be trusted (TC) (i.e. is predictable), and whether beneficial adaptations (BA) are possible.
- Trust (TC) depends on the experienced errors (ER) and how predictable the occurrence of the errors is (MM).
- Mental model (MM) of system limitations and functioning depends on how well the errors (ER) can be predicted from the situation (SA).
- Accurate perception of the situation (SA) requires a high level of scene understanding.
- The safety margins (SM) depend on the situation (SA) as well as the mental models (MM) of not just the system but also predictability of the behavior of other vehicles.
- Behavioral adaptation (BA) is only possible if the systems operate predictably (TC) with lower risk at less workload (WL) in which case the safety margins (SM) and set points can be adapted.
- …



Honest Evaluation Taxonomy:
10 Fringe Cases

1. System loses static constraint – *lose lane boundary*
2. ...em misses dynamic constraint – *lose lead* *...cycle*
3. Syste... perturbs – *radar reflections*
4. System fals... strains – *misinterpret lane boundary (e.g. ...ins, oil stripes, off ramp)*
5. System hardware su... tops – *HSC or power steering falls out*
6. System hardware locks – *HSC, ... throttle, steer*
7. System falsely interprets scene – *p... risk of bicyclists swerving into own lane at inte... n.*
8. System falsely interprets intent – *pedestrian crosses in other direction but car slams on brake.*
9. System misses intent – *car backing into roadway*
10. System perceptual ability degrades in some weather / lighting conditions – *noisy sensor returns*

Is this complete and applicable-to/useful-in other domains?



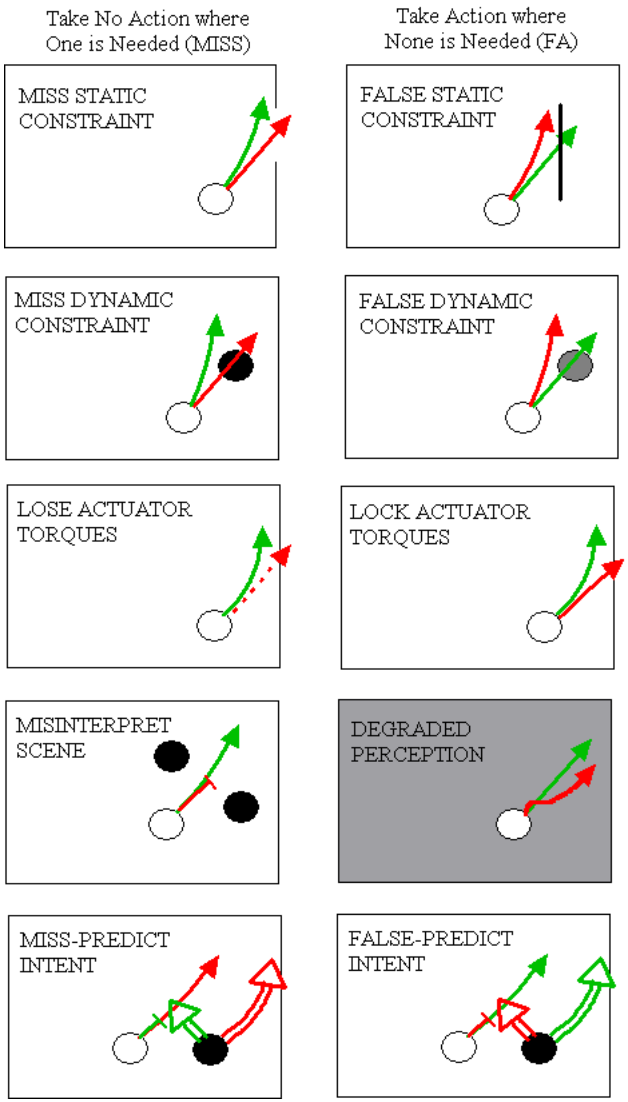
Expanding the fringe cases space dimensionality!

Establishing completeness

1. System loses static constraint – *lose lane boundary*
2. System misses dynamic constraint – *lose lead motor cycle*
3. System falsely perturbs – *radar reflections*
4. System falsely constrains – *misinterpret lane boundary (e.g. Old strips, oil stripes, off ramp)*
5. System hardware support stops – *HSC or power steering falls out*
6. System hardware locks – *HSC, brake, throttle, steer*
7. System falsely interprets scene – *potential risk of bicyclists swerving into own lane at intersection*.
8. System falsely interprets intent – *pedestrian crosses in other direction but car slams on brake.*
9. System misses intent – *car backing into roadway*
10. System perceptual ability degrades in some weather / lighting conditions – *noisy sensor returns*

Expanding the fringe cases space dimensionality!

Establishing completeness



TUTORIAL:
Honest Evaluations of Shared Human-Machine Control Systems

Evaluation Taxonomy

Erwin R. Boer

Associate Professor, Mechanical Engineering
Delft University of Technology

Expanding the fringe cases space dimensionality!

Establishing Utility across Domains

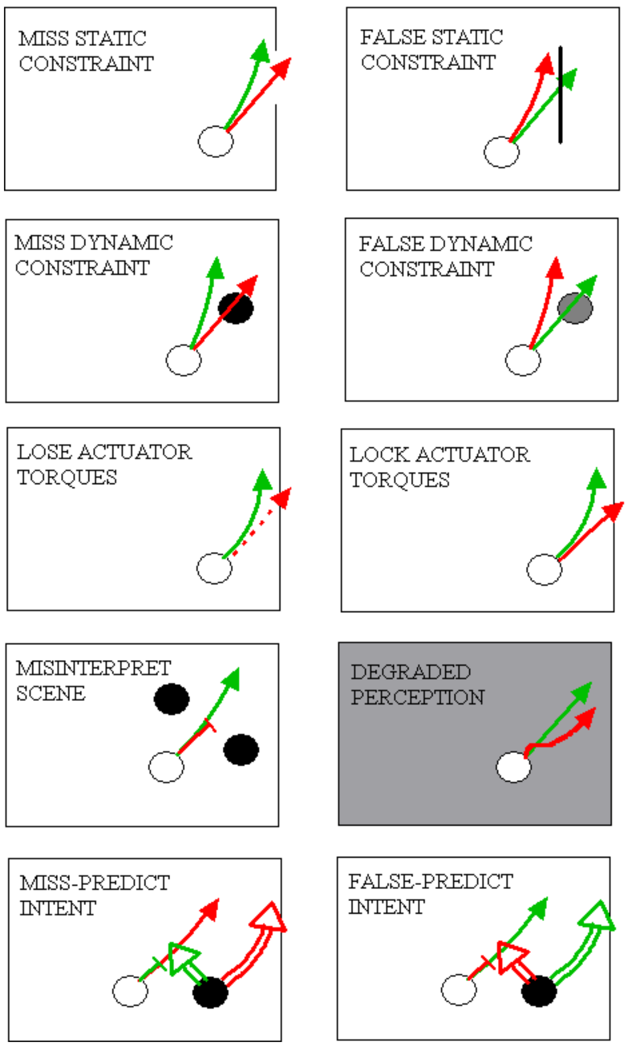
1. System loses static constraint – *lose lane boundary*
2. System misses dynamic constraint – *lose lead motor cycle*
3. System falsely perturbs – *radar reflections*
4. System falsely constrains – *misinterpret lane boundary (e.g. Old strips, oil stripes, off ramp)*
5. System hardware support stops – *HSC or power steering falls out*
6. System hardware locks – *HSC, brake, throttle, steer*
7. System falsely interprets scene – *potential risk of bicyclists swerving into own lane at intersection*.
8. System falsely interprets intent – *pedestrian crosses in other direction but car slams on brake.*
9. System misses intent – *car backing into roadway*
10. System perceptual ability degrades in some weather / lighting conditions – *noisy sensor returns*

Expanding the fringe cases space dimensionality!

Establishing Utility across Domains

Take No Action where One is Needed (MISS)

Take Action where None is Needed (FA)

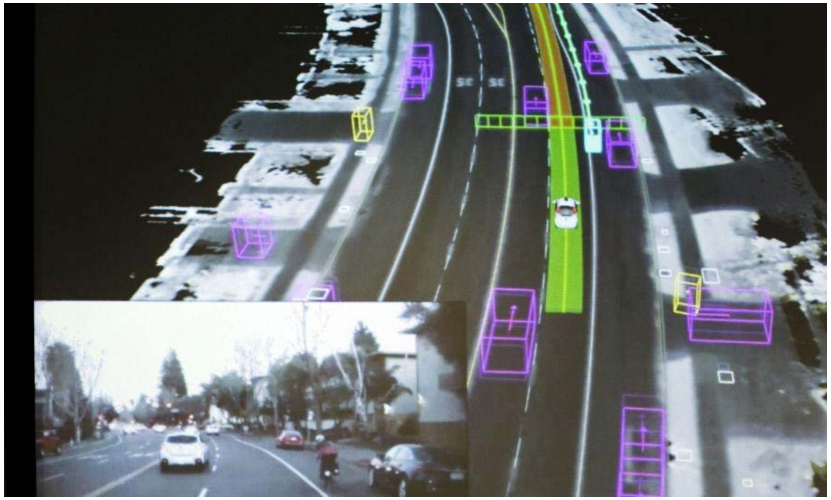


Evaluation - Risk Exposure and Experience

1. Can only surprise a Human Operator (HO) once; after that they expect events even if not known when, where and what type; vigilance will be higher.
2. Use cash based reward system to replace true risk in virtual environment assessments; assures that vigilance remains at realistic levels.

Evaluation - Error Detection and Recovery

1. Explore how long it takes a HO to detect a system failure; e.g. vigilance to system's status icons or awareness of guidance.
2. Error recovery if consequences are safety critical.



Evaluation – Behavioral Adaptation

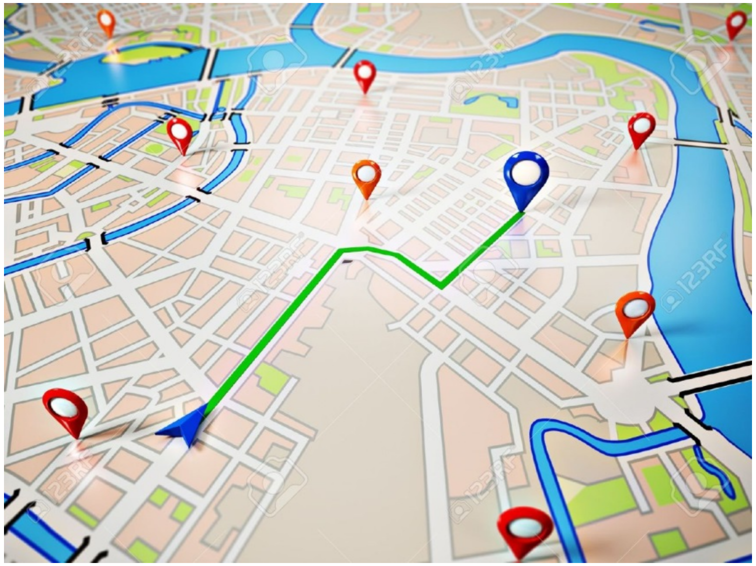
1. Learning the system I order to develop a calibrated trust requires exposure to and experience with situated system limitations; HO develops mental model that drives situated vigilance.
2. If system appears failsafe, HO will decrease vigilance and if system is failsafe under predictable and recognizable conditions, a situated adaptation of vigilance will ensue.
3. If through experience of using the system it is deemed safe and beneficial, behavioral adaptation may occur.
4. Depending on level of support (manual, shared, automated), behavioral adaptation of set points and safety margins may not be possible or system does not adapt with operator; HO may still adapt level of engagement.

Hard to study long enough in lab to present realistic fringe case rates → SHRP2.

Evaluation – Behavioral Adaptation

Humans adapt depending on their experience:

- ❑ Vigilance
  - ❑ Higher to deal with system imperfections
  - ❑ Lower if reliance on system warranted
- ❑ Speed (BA)
  - ❑ Higher if system feels more stable and controllable
  - ❑ Lower if system is unpredictable
- ❑ Safety margins (e.g. proximity to static and dynamic constraints)
  - ❑ Longer if system appears unpredictable
  - ❑ Shorter if system improves controllability.



Trust and Vigilance Dynamics

Evaluation – Interesting Interactions

1. If HO ability degrades (e.g. age, alcohol, distraction) support benefits generally improve.
2. Note that a more demanding environment sometimes improves safety (e.g. lane width Virginia, remove intersections in NYC, switch left to right in Sweden) because people pay more attention and become more vigilant to perturbations and events