The Human Controller

Class 4. Design & Evaluation

“The Red and Blue Chair”
Gerrit Rietveld, 1917

From Novelty to Utility

From Aesthetics to Performance

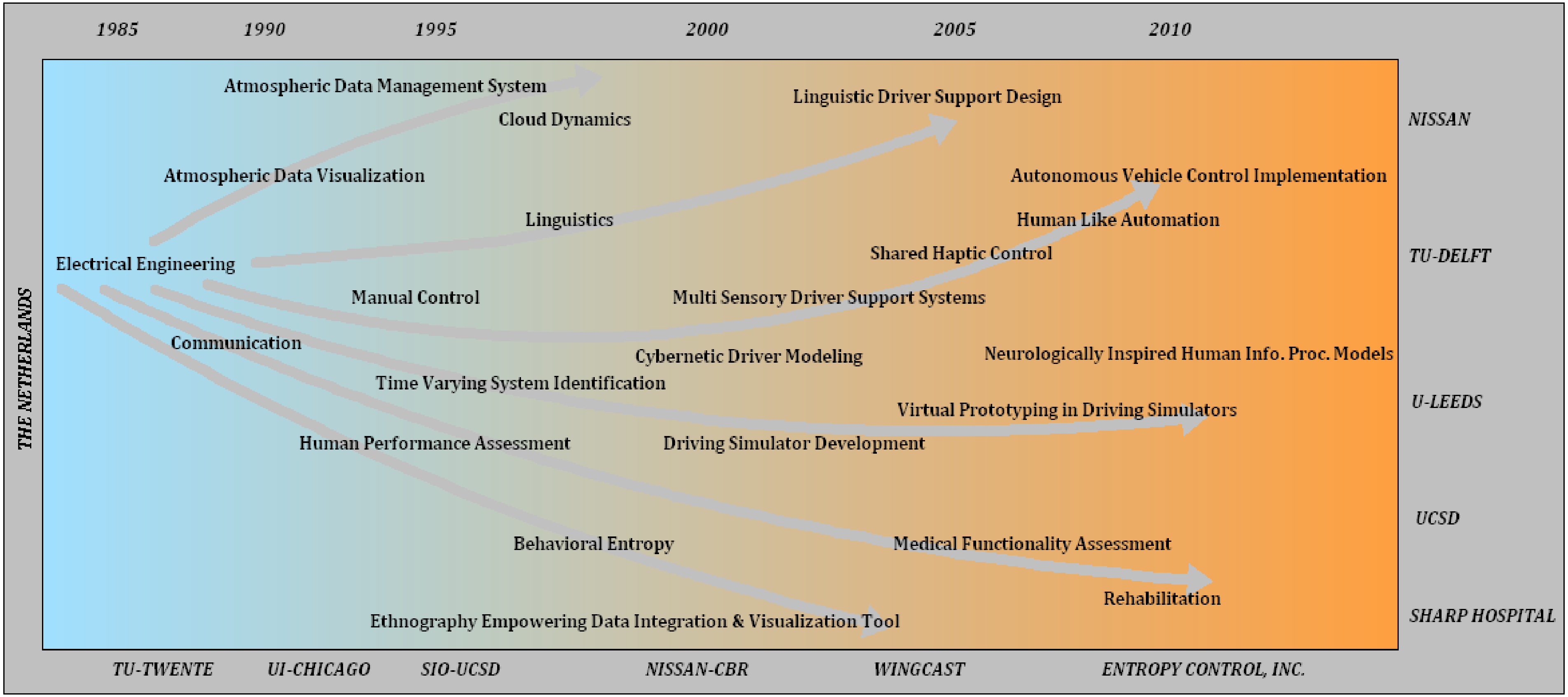
Teacher:

Erwin R. Boer

BioMechanical Engineering, Delft University of Technology, The Netherlands

Dr. Erwin R. Boer

President of Entropy Control, Inc.



1985
1990
1995
2000
2005
2010

THE NETHERLANDS

Electrical Engineering

Atmospheric Data Management System

Cloud Dynamics

Atmospheric Data Visualization

Linguistics

Linguistic Driver Support Design

Manual Control

Multi Sensory Driver Support Systems

Shared Haptic Control

Human Like Automation

Autonomous Vehicle Control Implementation

Communication

Cybernetic Driver Modeling

Neurologically Inspired Human Info. Proc. Models

Time Varying System Identification

Virtual Prototyping in Driving Simulators

Human Performance Assessment

Driving Simulator Development

Behavioral Entropy

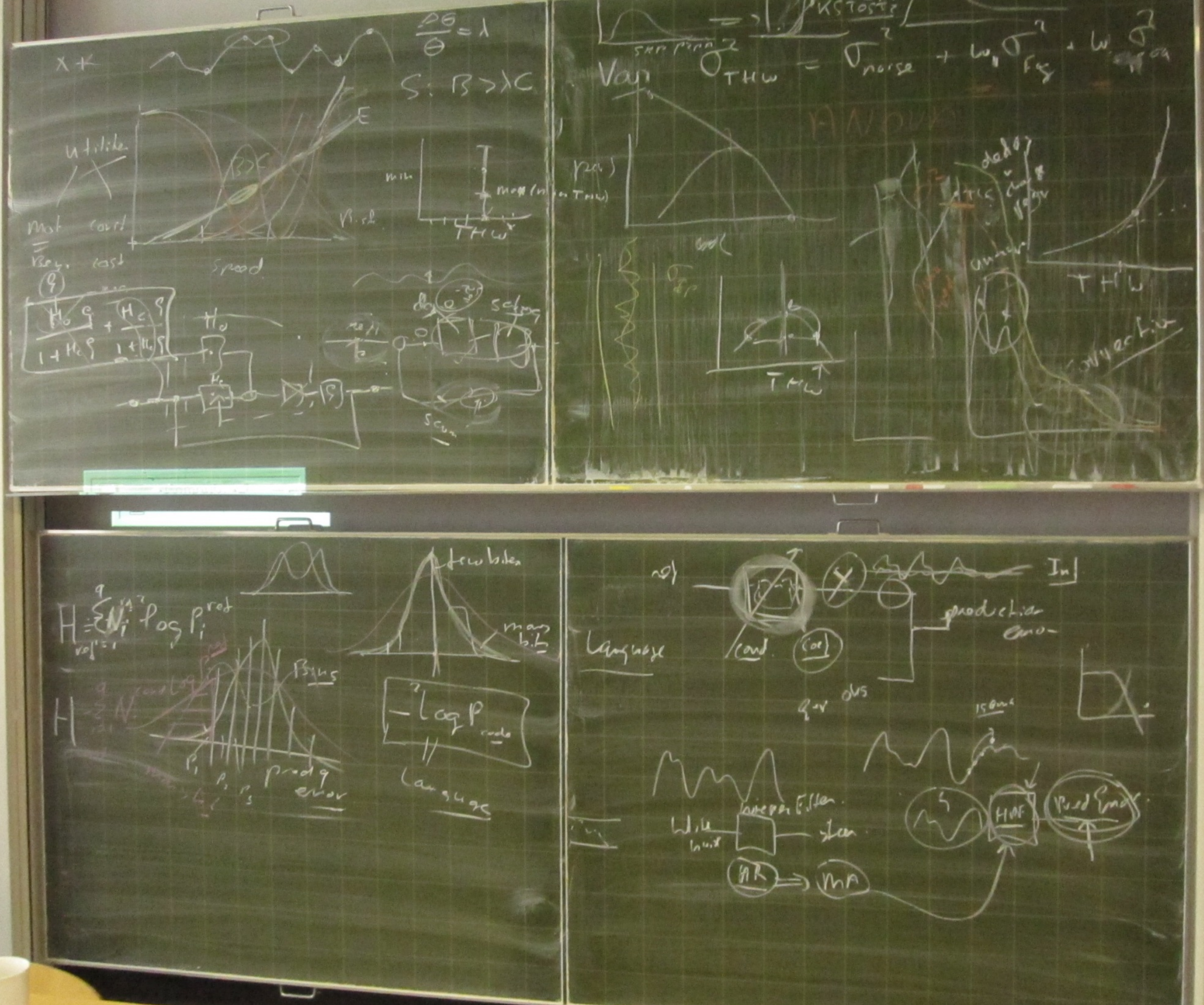
Medical Functionality Assessment

Rehabilitation

Ethnography Empowering Data Integration & Visualization Tool

NISSAN
TU-DELFT
U-LEEDS
UCSD
SHARP HOSPITAL

TU-TWENTE
UI-CHICAGO
SIO-UCSD
NISSAN-CBR
WINGCAST
ENTROPY CONTROL, INC.


ΔG/θ = λ
S: B > λC
utilit.
Mat cost
=
Beg. cost
Speed
Risk
E
min
max (min THW)
THW
H₀
H_c
selog
scan
Var
σ²_THW = σ²_noise + w σ²_Fg + w σ²
ANOVA
p(x)
cond
THW
convection
H = Σ N_i log p_i
Bayes
pred error
few bits
max bits
− log p_code
Language
Language
In
production error
cond
AR → MA
White noise
linear filter
HMF
Pred Error

Learning Goals Lecture 4

After this lecture, you will be able to:

1. Reproduce:
    a. Progression in our understanding of human functioning through time.
    b. Progression in design of systems and human machine interaction through time.
    c. Key types of designs that enhance human machine interaction.
    d. Classical ways to assess human machine interaction performance.
    e. Ways to characterize and measure human behavior.
    f. Statistical means to compare systems or conditions in human machine interaction
2. Apply:
    a. Methods to characterize human machine interaction behavior.
    b. Methods of evaluation to determine what system or interface is better.
3. Be critical of:
    a. The ways in which a human machine interaction task can be supported.
    b. The limitations in evaluating human machine interaction from a single narrow perspective.

The Human Controller
Human Machine Interaction

Focus:

- Human control of mechanical things.

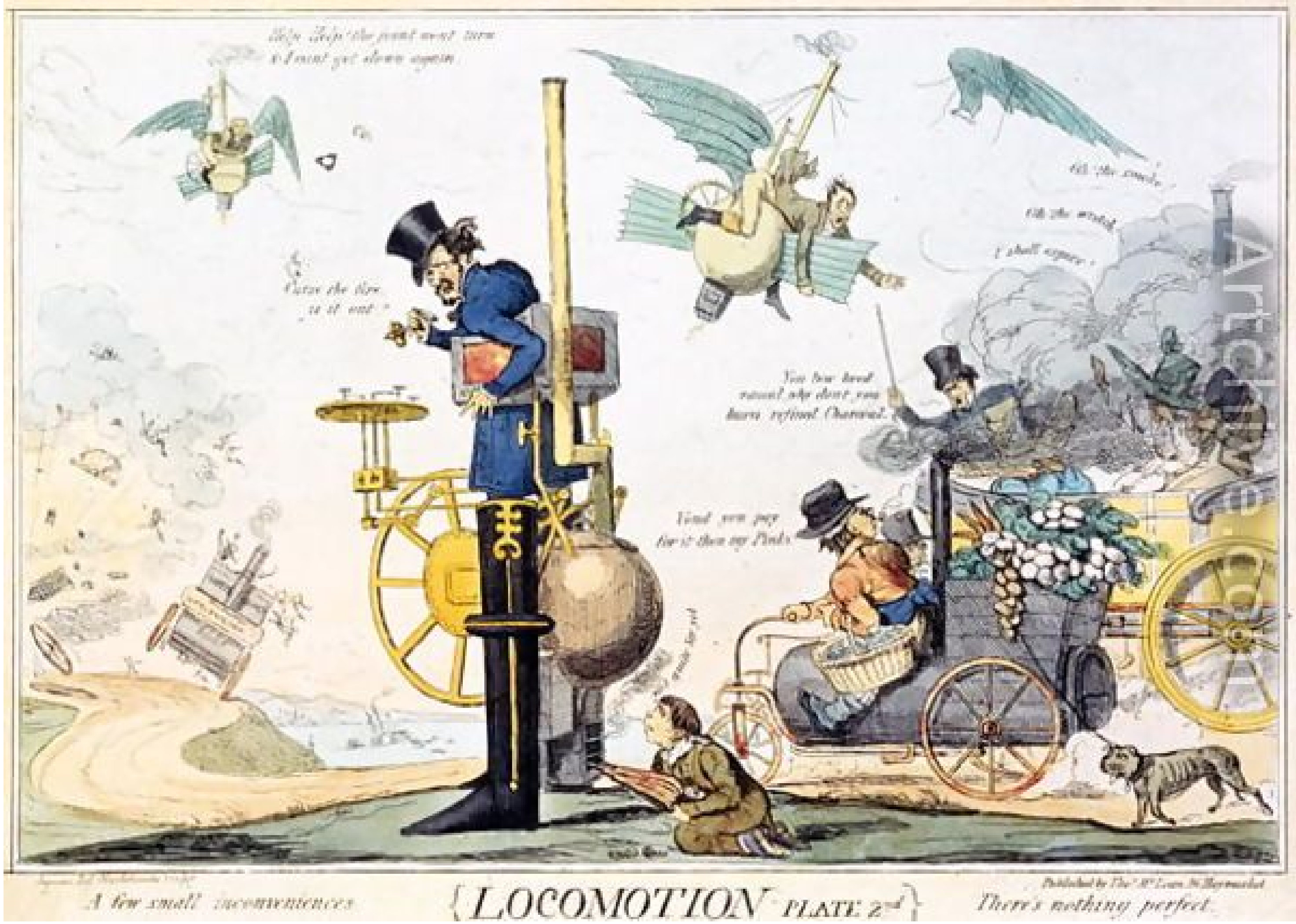
Examples:

- Hand tools
- Controlling the mechanical movements of own limbs
- Extensions of their limbs such as prostheses
- Tele-manipulators

- Controlling the mechanical movements of vehicles
  - Aircraft, automobiles, and trains.

- Movements of discrete products through manufacturing plants, or chemicals and other fluids through process plants such as refineries or nuclear power stations.

Early Human-Machine Interaction



'Locomotion', London, c1820.

Designing for and by Humans Shaping our Environment through Time

Human Creation

- Natural environment
- Change natural environm
- Design powerfu
- Adapt work
  machine co
- Adapt machi
  manipulators
- Adapt interacti
  cooperation wit machines
- Design reliable automation

Hu nowledge

e tools

y tasks

hysical limitations

man factors) (1950

controller (1960 –

nological advances in sensing and actuation (1990 – 2010)

- Learn human adaptation (2010 - ...)

Create new environments for humans to work and play in.

Support

Information Acquisition

Knowledge Building

Controllability

Supporting Humans

Display Augmentation

- Aug environment
- Aug state
- Aug system

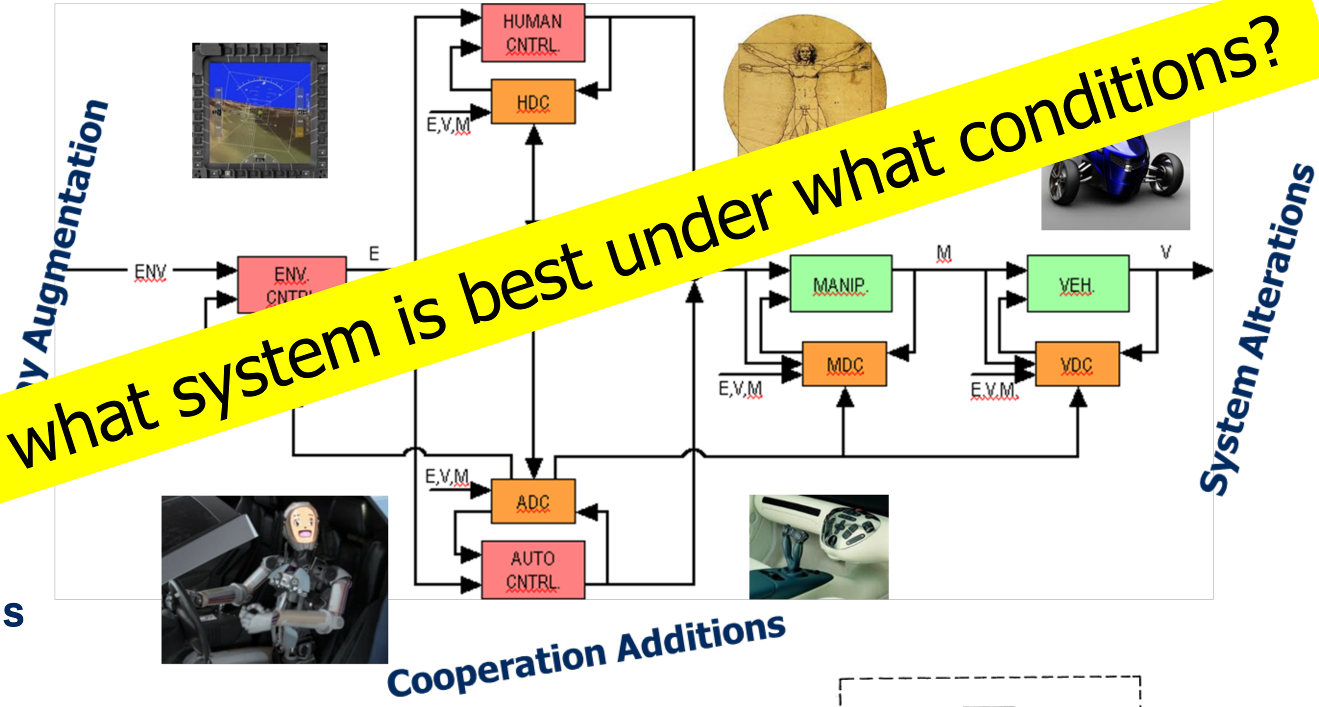
System Alterations

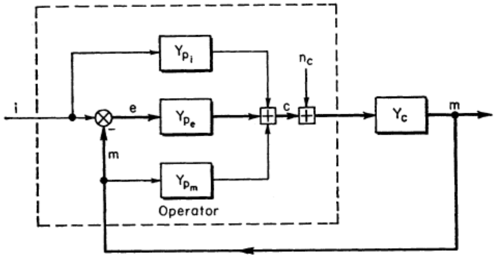
- Alter control
- Alter system
- Alter fe

peration Additions

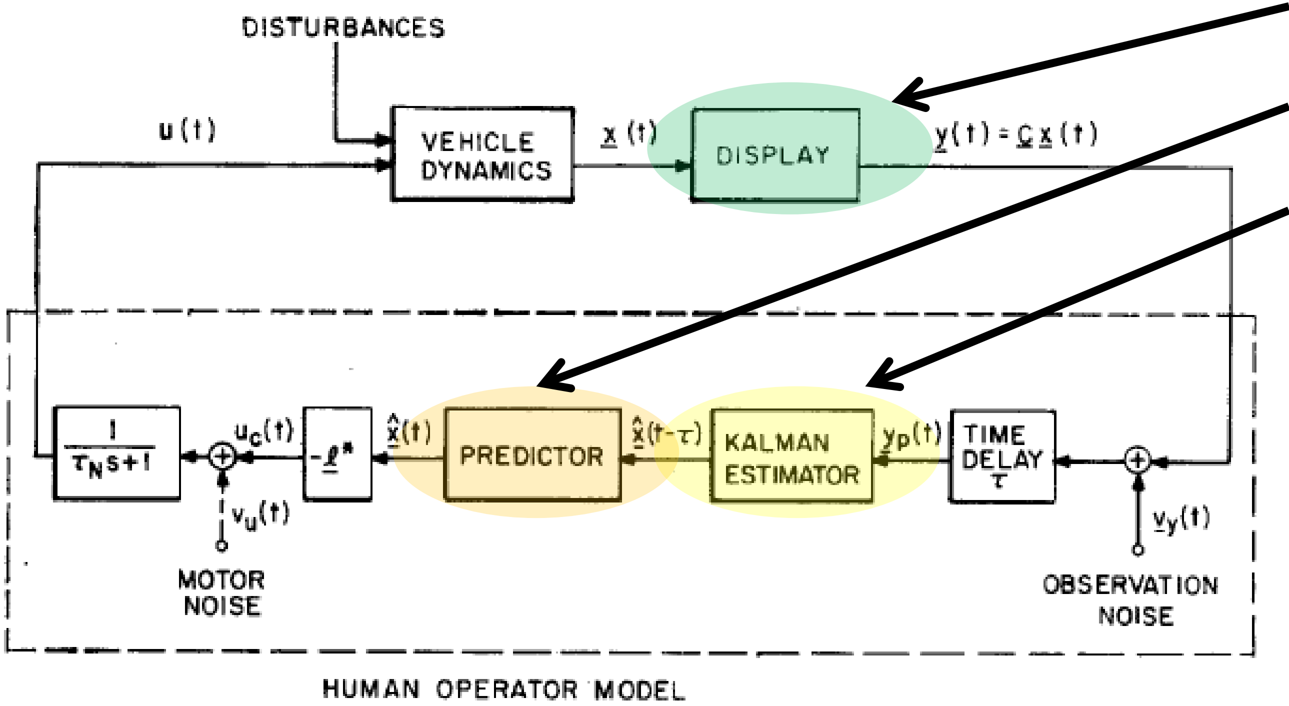
- Add protection
- Add controller
- Add automation

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





Supporting the Human Perception and Prediction



Augmentation

Quickened

Filtering

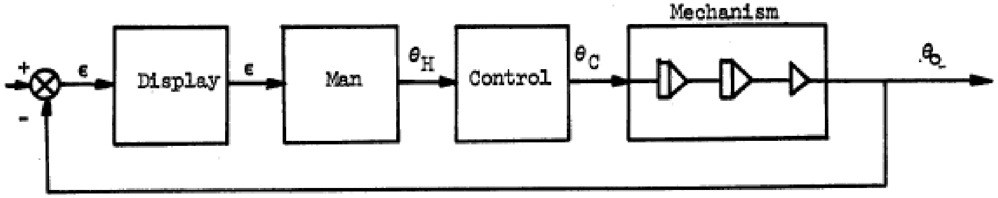
FIG. 2. Control-theoretic model of optimal human behavior.

D.L. Kleinman, S. Baron, and W.H. Levison, "An Optimal Control Model of Human Response. Part 1: Theory and Validation," *Automatica*, vol. 6, no. 3, 1970, pp. 357-369.

Quickening – Since 1958

(a) Unquickened System

$$\epsilon = \theta_i - \theta_o$$



(b) Quickened System

$$\epsilon = \theta_i - \left[k_1\theta_o + k_2\dot{\theta}_o + k_3\ddot{\theta}_o\right]$$

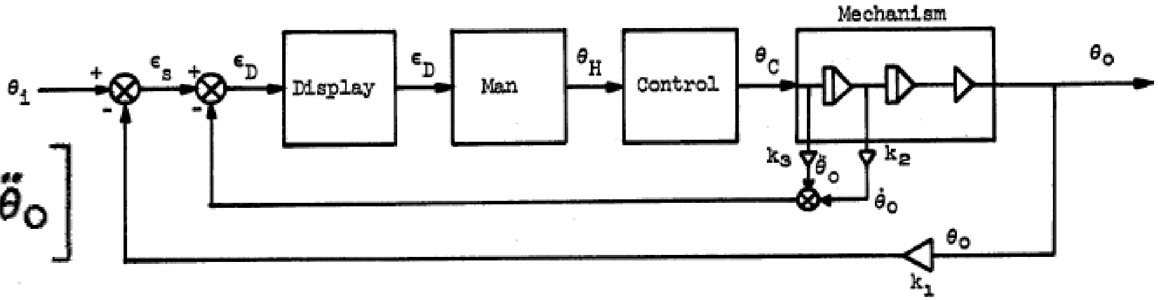
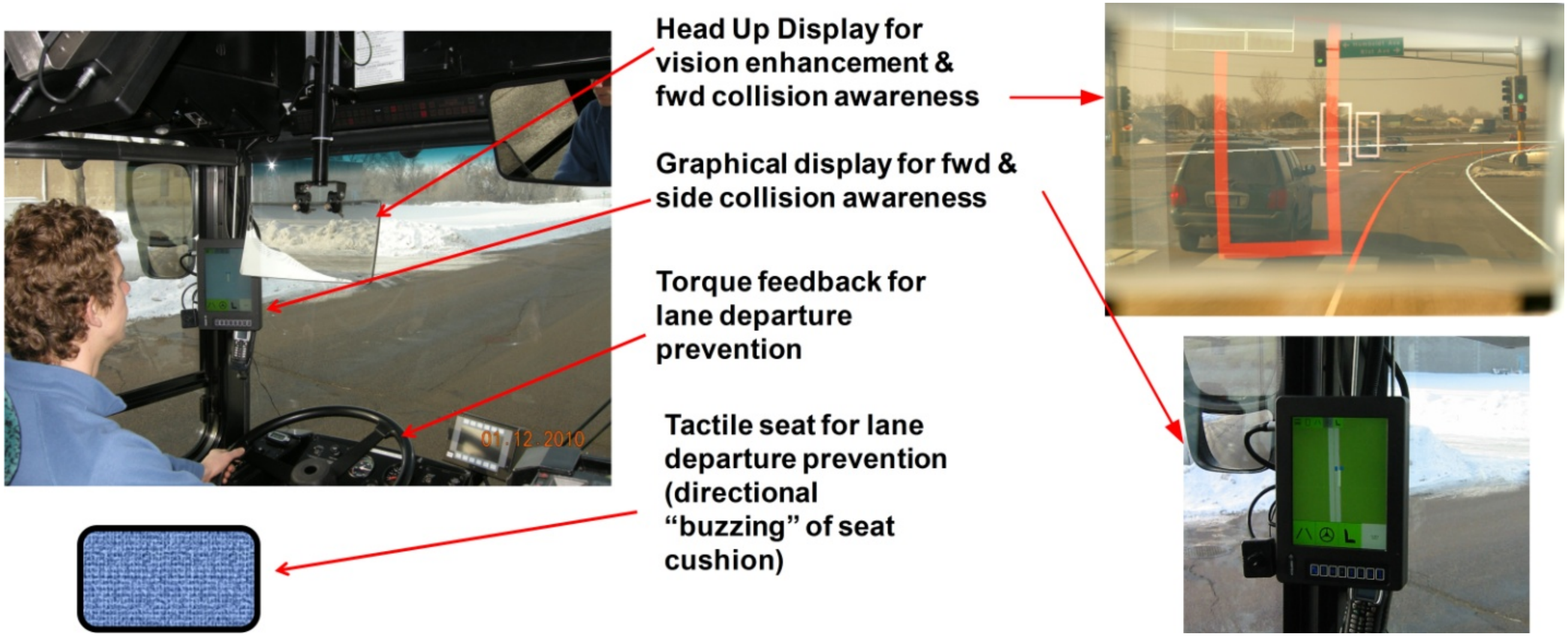


Figure 2. Example of an unquickened and a quickened system

Forst, George, 1965: Effect of display quickening on human transfer functions during a dual-axis compensatory tracking task. U S Air Force Tech Doc Rep Amrl Tr4: 1-207

Augmentation



Head Up Display for vision enhancement & fwd collision awareness

Graphical display for fwd & side collision awareness

Torque feedback for lane departure prevention

Tactile seat for lane departure prevention (directional "buzzing" of seat cushion)

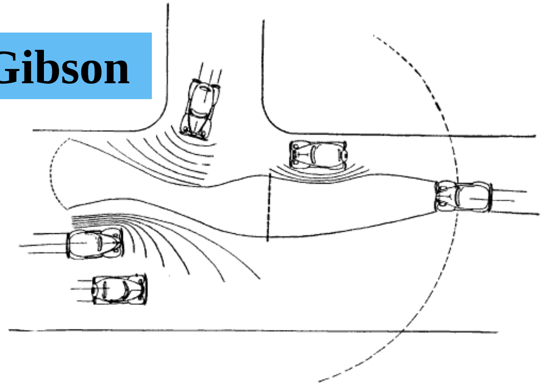
http://www.bus2.me.umn.edu/system.html

Three Key People for the Human Controller

A Theoretical Field-Analysis of Automobile-Driving
James J. Gibson and Laurence E. Crooks
The American Journal of Psychology, Vol. 51, No. 3 (Jul., 1938), pp. 453-471

HOW PEOPLE PERCEIVE AND USE THE ENVIRONMENT: DIRECT MANIPULATION OF PERCEPTS

J.J. Gibson



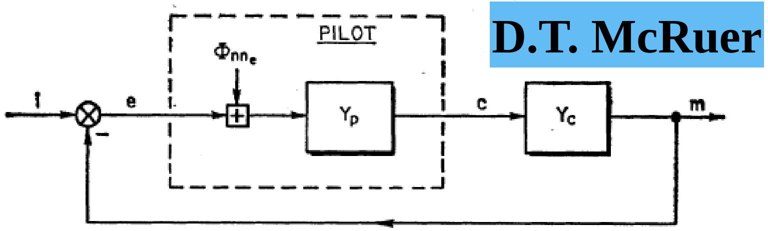
McRuer, Duane T.; Jex, Henry R., **A Review of Quasi-Linear Pilot Models**,“
Human Factors in Electronics, IEEE Transactions on ,
vol.HFE-8, no.3, pp.231,249, Sept. 1967

The *crossover model* is

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

HOW PEOPLE VIEW A CONTROL TASK AND ADAPT THEIR BEHAVIOR TO ACHIEVE DESIRED PERFORMANCE

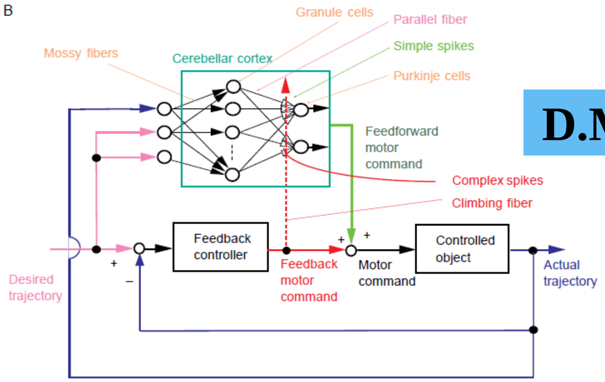
D.T. McRuer



Wolpert DM, Miall RC & Kawato M (1998)
Internal models in the cerebellum.
Trends in Cognitive Sciences 2:338-347

HOW DO PEOPLE INTEGRATE SENSORY INFORMATION AND WHAT INTERNAL MODELS DO THEY BUILD?

D.M. Wolpert



Similarity Across all Control Tasks

How to Quantify Risk across all Possible Situations & Conditions?

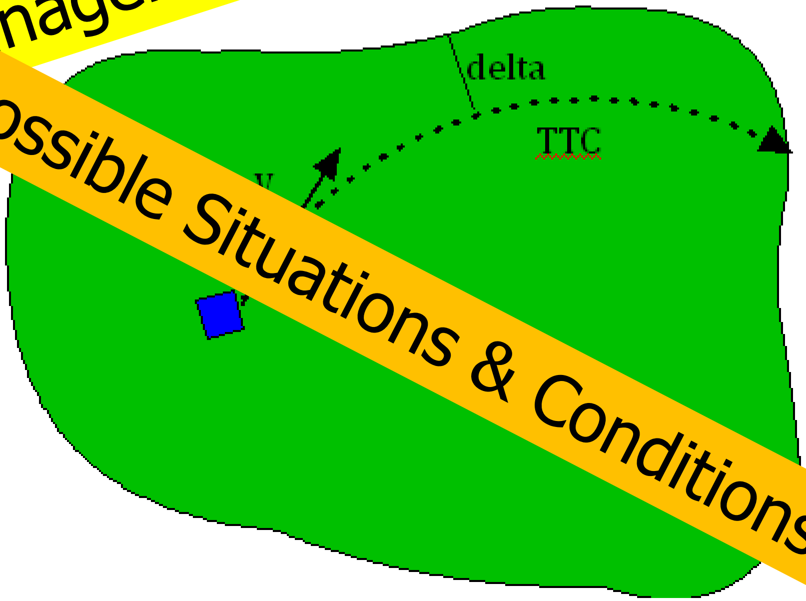
Those that support toleranc anagement under all conditions!

lerance management

- ndaries in the task space uld not be exceeded
  - Spatial constrai
  - System dynamic cons
- Risk can be defined combinatio
  - these boundaries
  - rse at high parallel speed along boundary)
  - Approach to these boundaries

❑ **Keep state (blue) within established boundaries (in green field).**

❑ **Potentia sk h ta, and TTC.**



Three Main HMI Design Perspectives

Interior Layout → Distributed Cognition & Ergonomics

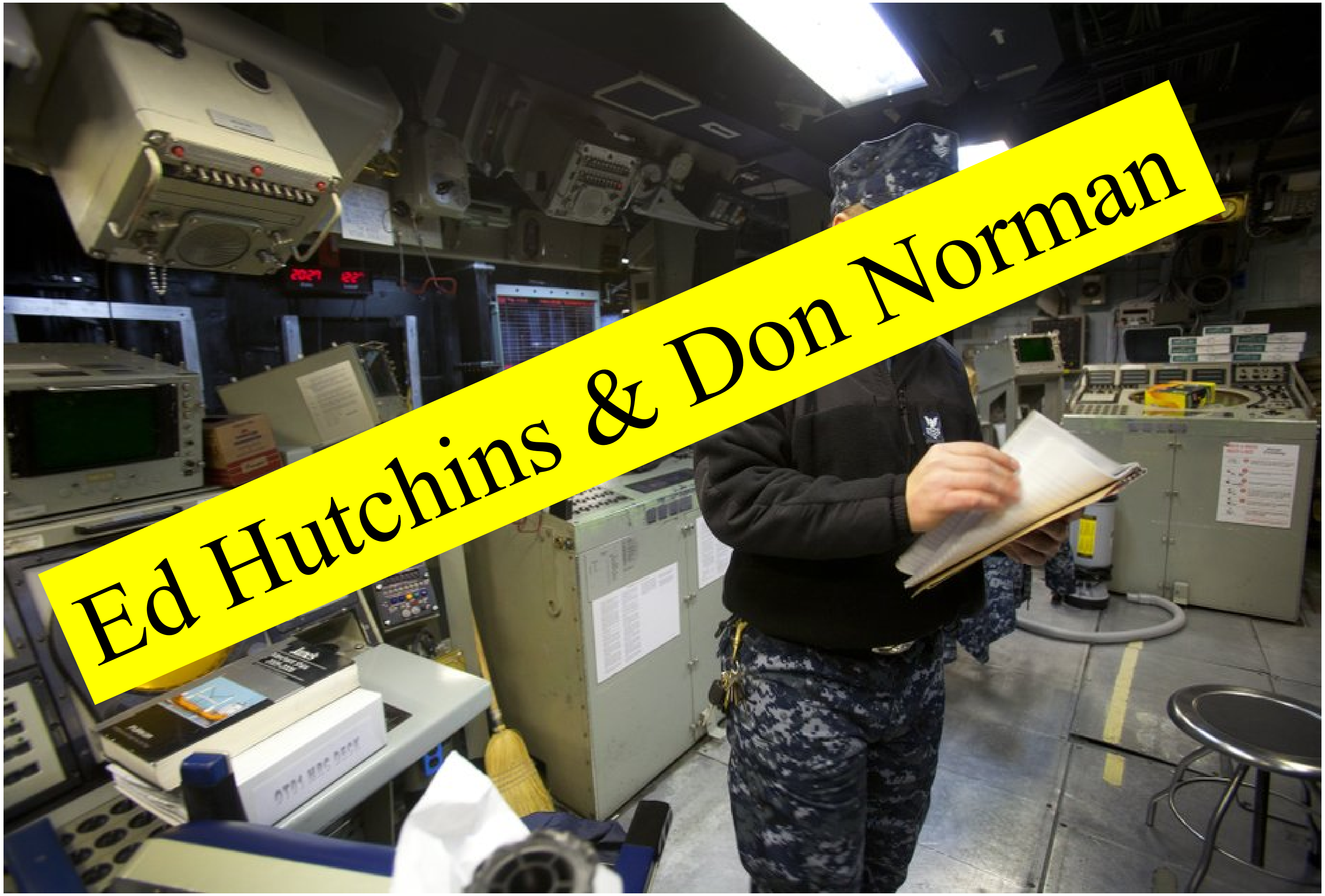
Interface Layout → Ecological Interface Design & Usability

Interaction Layout → Interaction & Automation

MATCH HUMAN AND MACHINE

Perceptual, Cognitive, Control

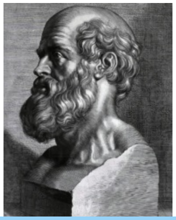
How to Design the Space Layout



Ed Hutchins & Don Norman

History of Design for Humans – Ergonomics

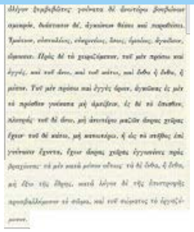
A good deal of evidence indicates that Greek civilization in the 5th century BC used ergonomic



e

4

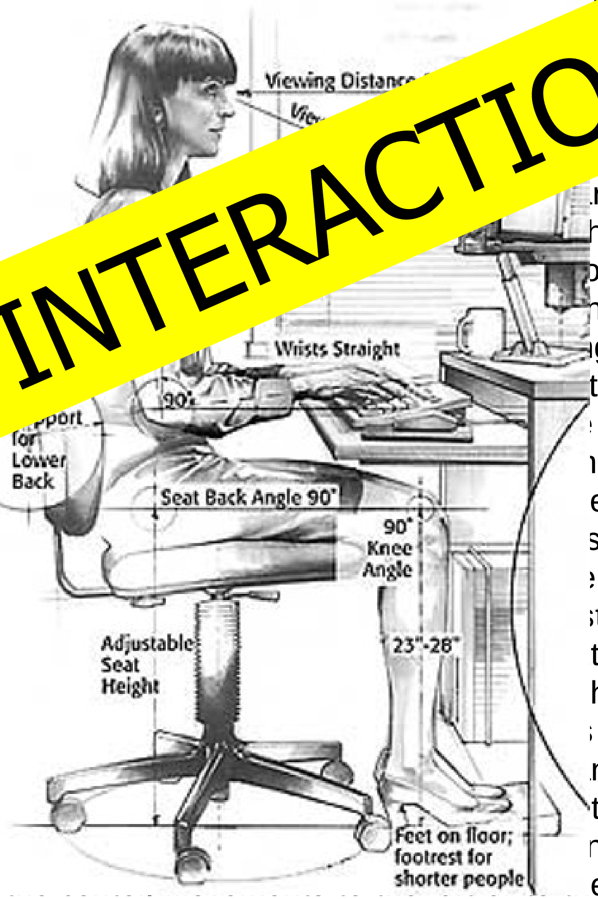
kplace
be designed and how the tools he uses should be arranged.



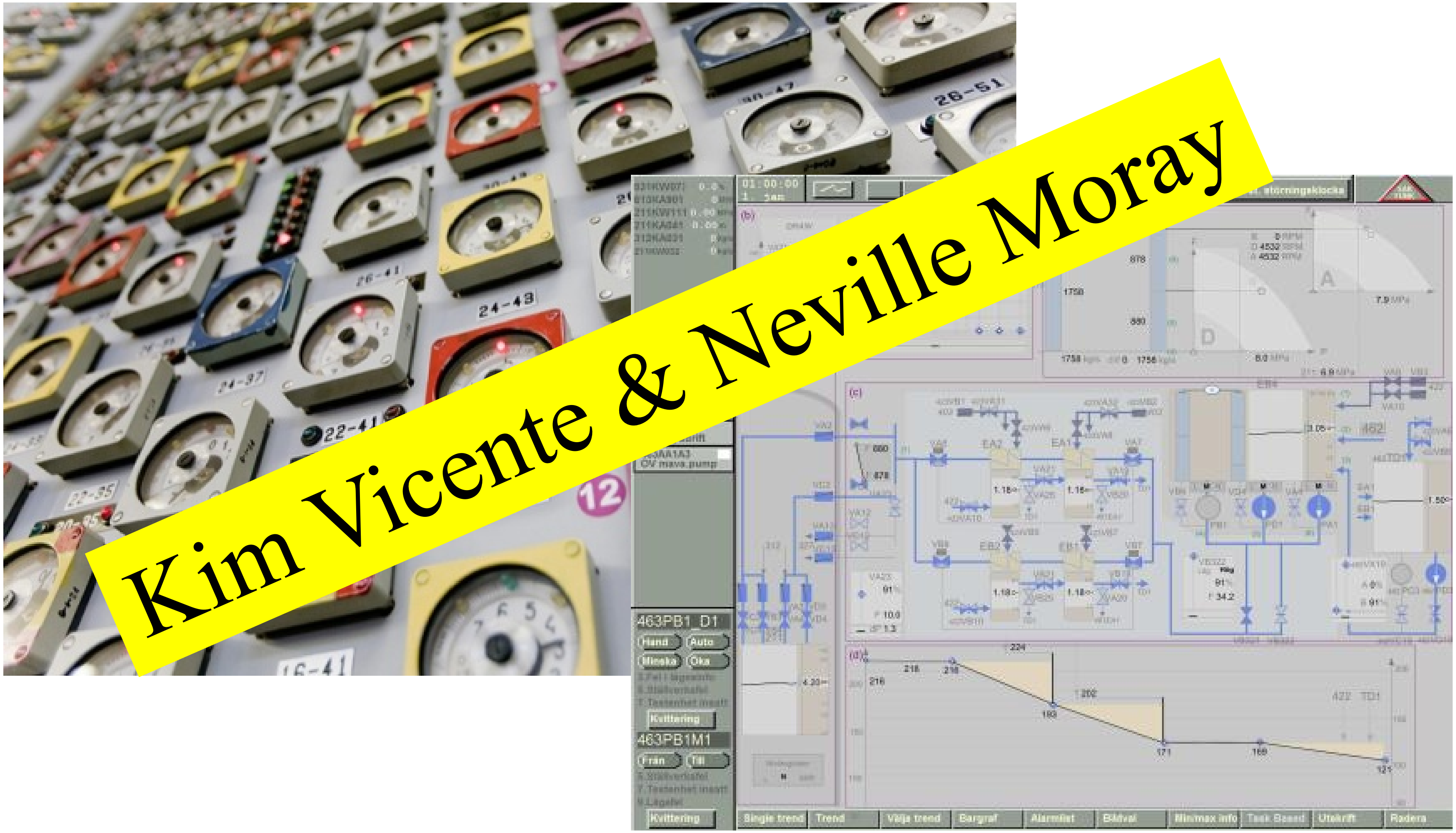
1: Body and working height
2: Work area
3: Reach zones
4: Parts presentat
5: Range of
6: Li
t of work equipment

BUT WHAT ABOUT INTERACTION?

3. The surgeon may stand or be seated, in a posture comfortable for hi
and dependent on th
s
of lighting: natural ar
artificial, however, is
ways namely directly
and
it is obvious to which
r direct
lighting, the point of
rea, of
he limbs in
common
nt and the
g him. This
tient will not
seated,
ust be a
e elbows
surgeon's
no folds
t also
tion, that is
ht or at the
are behind
center. The surgeon
nds must
his knees and in fron
t must not
not be higher than hi
n 900. The
touch his knees and
same rule applies for
e right must
not cause him to leave his seat. If, however, he needs to turn, the patient's body and the area of operation must be repositioned.



How to Design Interface Symbology



Kim Vicente & Neville Moray

History of Design for Humans – Usability

10 Usability Heuristics for User Interface Design by JAKOB NIELSEN on Januar

Visibility of system status The system should always keep users informed about what is going on, thr within reasonable time.
Match between system and the real world The system should speak the users' langua familiar to the user, rather than system-oriented terms. Follow real-world conventions logical order.
User control and freedom Users often choose system functions by mis ency exit" to leave the unwanted state without having to go through an extended
Consistency and standards Users should not have to wo ctions mean the same thing. Follow platform conventions.
Error prevention Even better than good error a problem from occurring in the first place. Either eliminate error-prone conditi a confirmation option before they commit to the action.
Recognition rather than re king objects, actions, and options visible. The user should not have to remember another. Instructions for use of the system should be visible or easily retrievable
Flexibility en by the novice user -- may often speed up the interaction for the expert user such ced and experienced users. Allow users to tailor frequent actions.
ges should not contain information which is irrelevant or rarely needed. Every extra unit of ith the relevant units of information and diminishes their relative visibility.
ose, and recover from errors Error messages should be expressed in plain language (no codes), blem, and constructively suggest a solution.
entation Even though it is better if the system can be used without documentation, it may be necessary to provide umentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried nd not be too large.

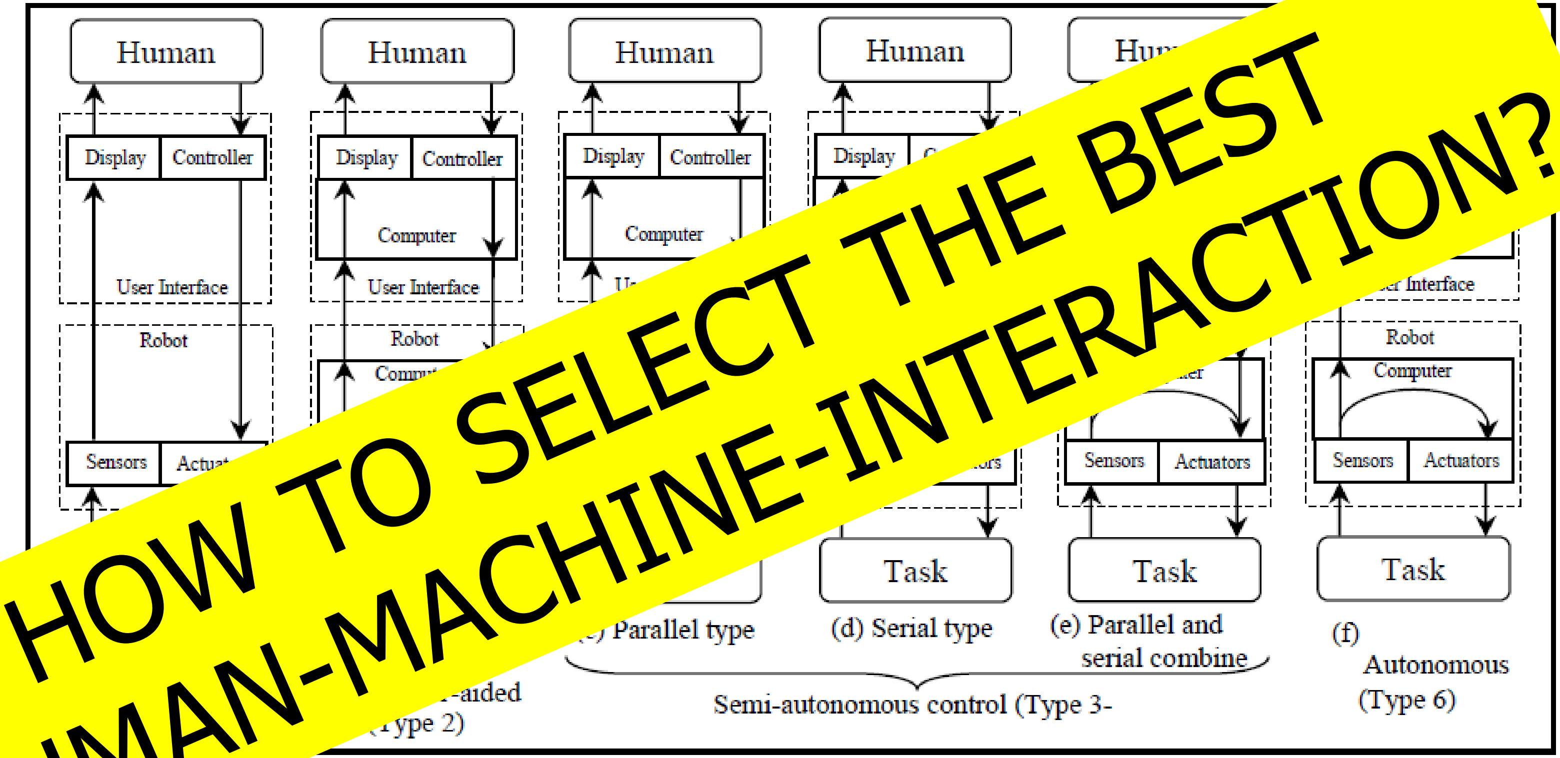
USABILITY: EASE OF USE AND LEARNABILITY OF A HUMAN-MADE OBJECT OR INTERACTION

- Nielsen, J., and Molich, R. (1990). Heuristic evaluation of user interfaces, *Proc. ACM CHI'90 Conf.* (Seattle, WA, 1-5 April), 249-256.

How to Design Human Machine Interaction & Responsibility

Tom Sheridan & Toshi Inagaki

History of Design for Humans – Interactive Control with Machines



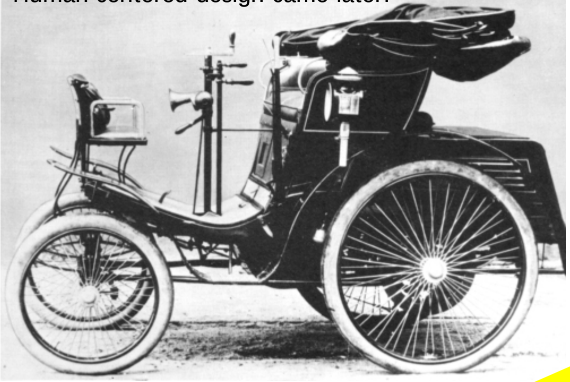
spectrum of control modes ((Fig. (a), (b), (e) & (f) are adapted and modified from Sheridan (1992) (Fig. (c) & (d)) are adapted and modified from Yasuyoshi et al. (1993))

HOW TO SELECT THE BEST HUMAN-MACHINE-INTERACTION?

Kai Wei Ong, Gerald Seet and Siang Kok Sim (2005). Sharing and Trading in a Human-Robot System, Cutting Edge Robotics, Vedran Kordic, Aleksandar Lazinica and Munir Merdan (Ed.), ISBN: 3-86611-038-3, InTech, DOI: 10.5772/4665.

HMI Designs Gone Wrong

Human centered design came later.



1886 is regarded the year of birth o
automobile - with the Benz P

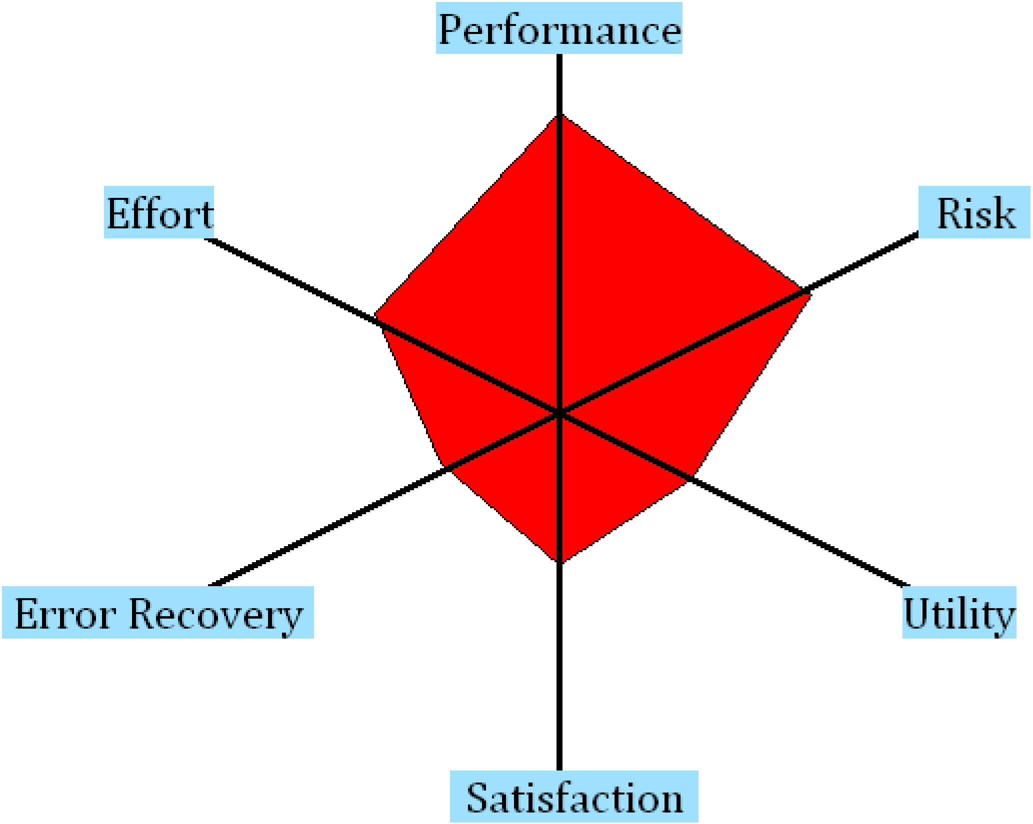
menu
gnition was
cuts.

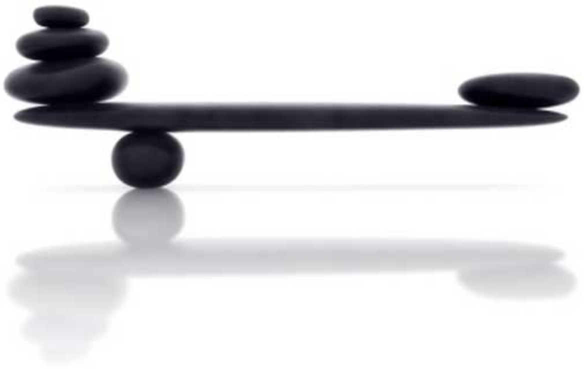
rd cars dropped from 5th to 23rd in
Power's customer satisfaction survey
primarily due to a complex and hard to use in-car digital UI.

100 years since the first car,
we still produce bad designs!

Evaluation of HMI Designs

Dimensions of HMI Evaluation

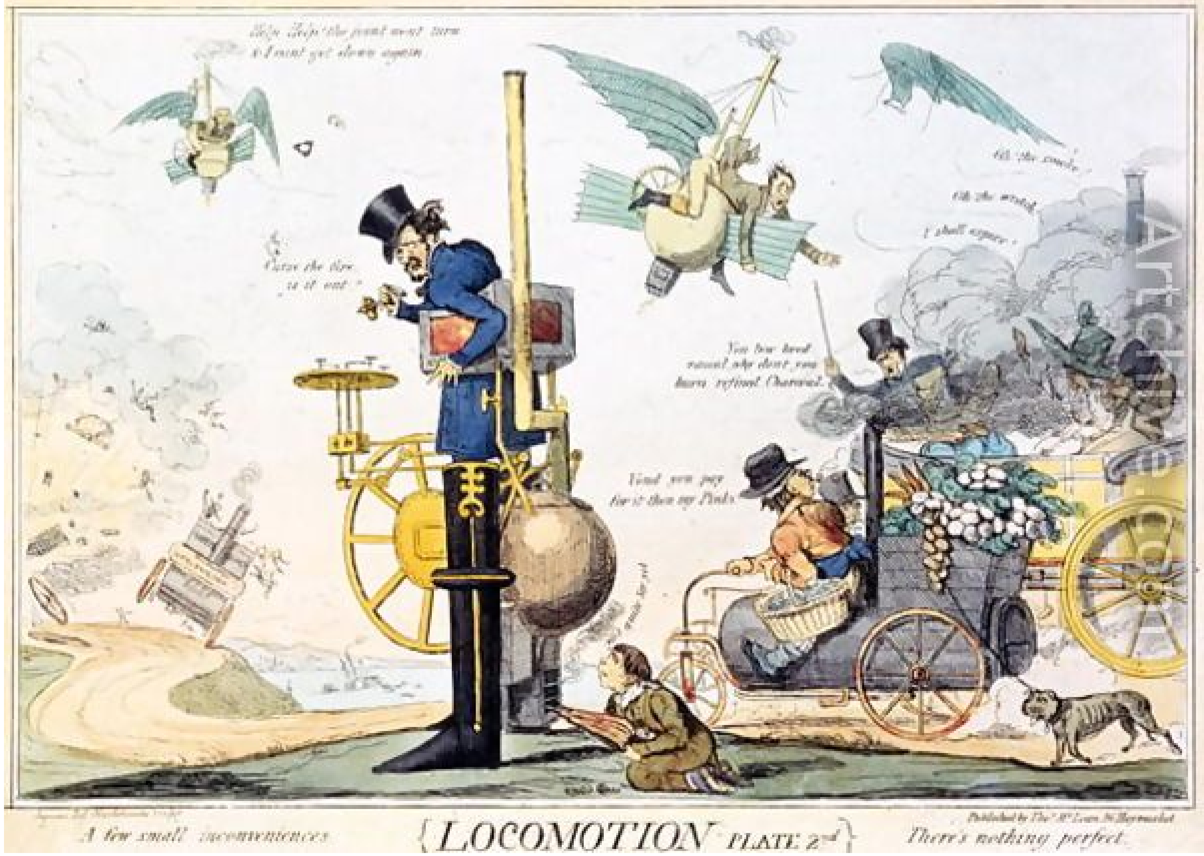




Clasical Evaluation Methods

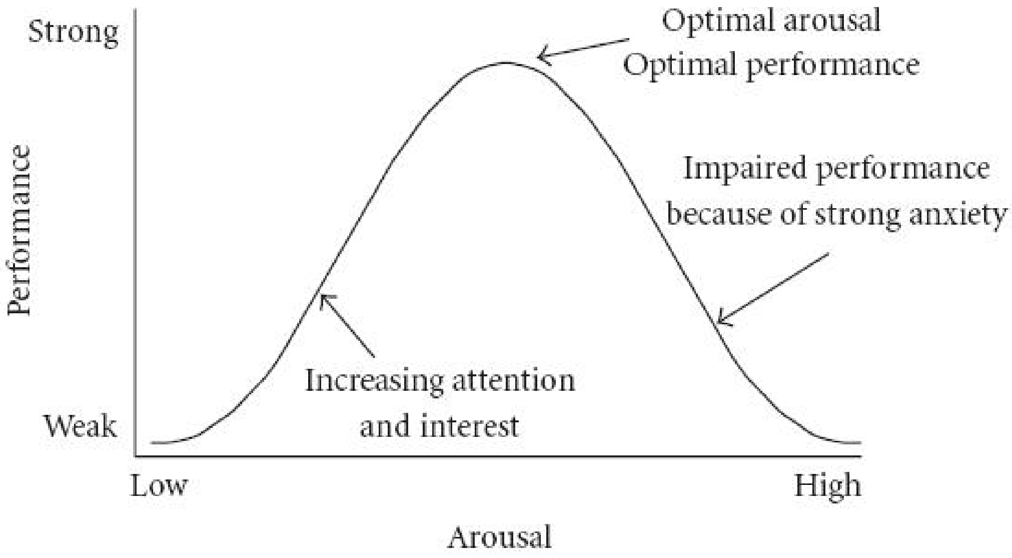
Objective & Subjective
Quantitative & Qualitative

- Inverted U
- NASA TLX
- Utility-Satisfaction
- Micro-Worlds
- GOMS Models
- Safety Margins
- *Whose Performance?*
- Reversal Rates
- Steering Entropy



Inverted-U Quantitative Performance-Arousal Balance

Yerkes–Dodson law



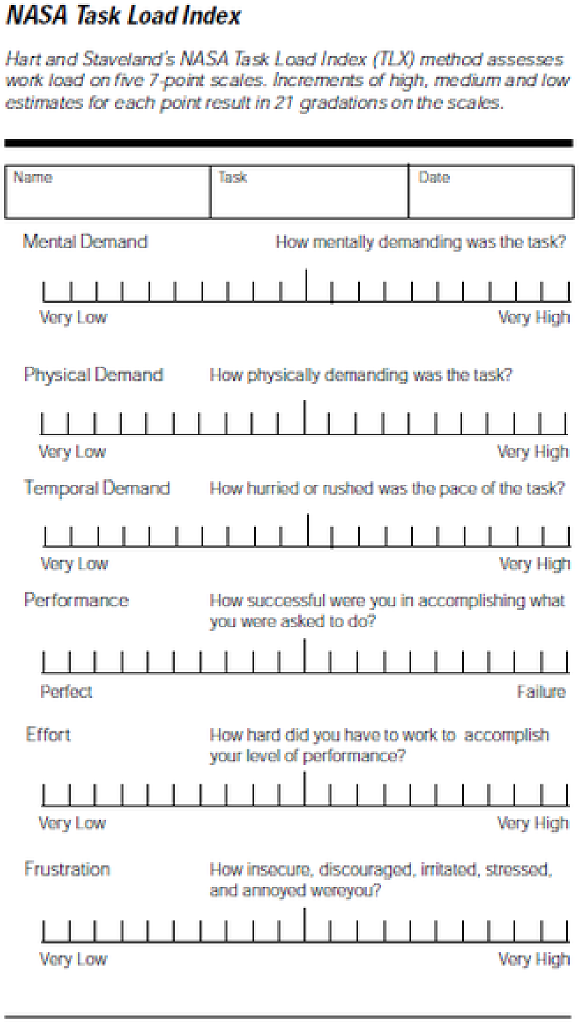
Yerkes RM, Dodson JD (1908). "The relation of strength of stimulus to rapidity of habit-formation". Journal of Comparative Neurology and Psychology 18: 459–482.

Qualitative Subjective Evaluation

User Perspective on Interaction with Systems

Quantify personal task execution experience along different assessment dimensions.

Hart, S., & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research



NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Utility – Satisfaction

Personal Assessment – Buy?

<table>
<tr><td>1</td><td>Useful</td><td>|__|__|__|__|__|</td><td>Useless</td></tr>
<tr><td>2</td><td>Pleasant</td><td>|__|__|__|__|__|</td><td>Unpleasant</td></tr>
<tr><td>3</td><td>Bad</td><td>|__|__|__|__|__|</td><td>Good</td></tr>
<tr><td>4</td><td>Nice</td><td>|__|__|__|__|__|</td><td>Annoying</td></tr>
<tr><td>5</td><td>Effective</td><td>|__|__|__|__|__|</td><td>Superfluous</td></tr>
<tr><td>6</td><td>Irritating</td><td>|__|__|__|__|__|</td><td>Likeable</td></tr>
<tr><td>7</td><td>Assisting</td><td>|__|__|__|__|__|</td><td>Worthless</td></tr>
<tr><td>8</td><td>Undesirable</td><td>|__|__|__|__|__|</td><td>Desirable</td></tr>
<tr><td>9</td><td>Raising Alertness</td><td>|__|__|__|__|__|</td><td>Sleep-inducing</td></tr>
</table>

Principal Component Analysis → 2D → **Satisfaction** & **Usability**

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.

Micro-Worlds

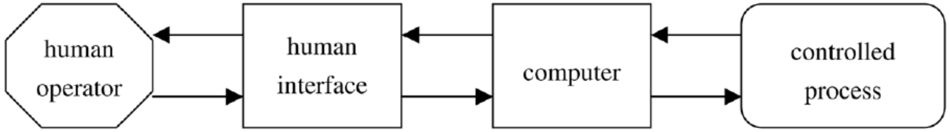


FIGURE 1 Supervisory control.

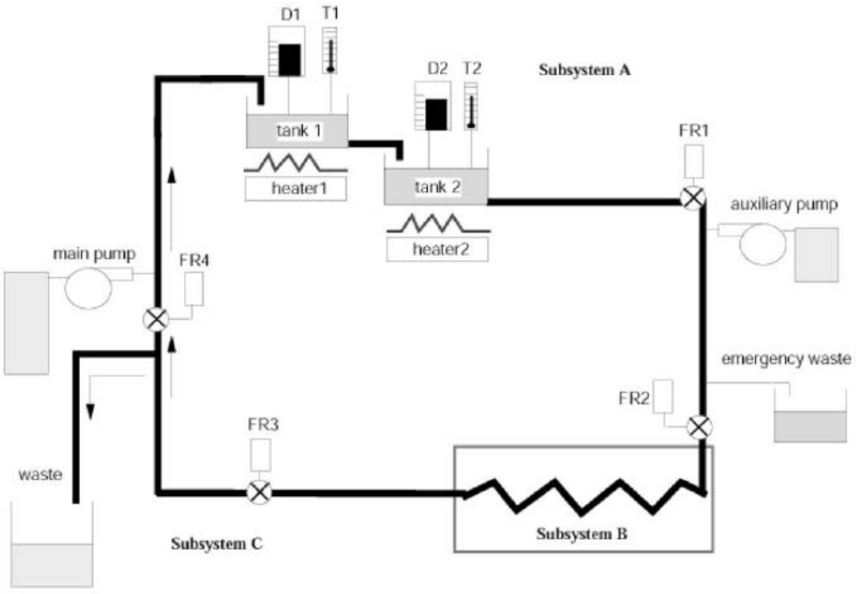
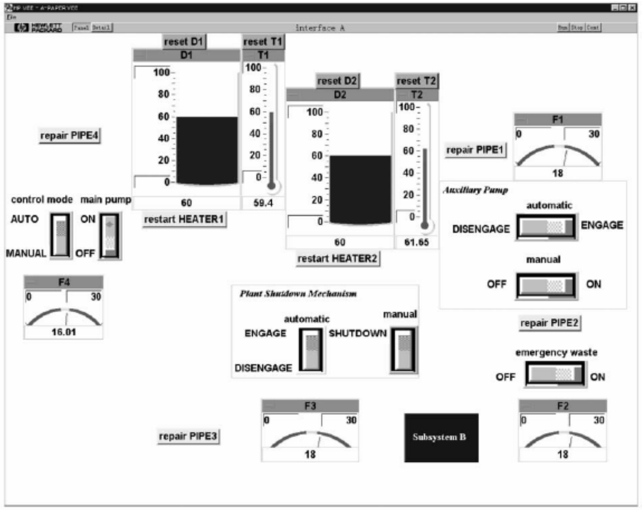
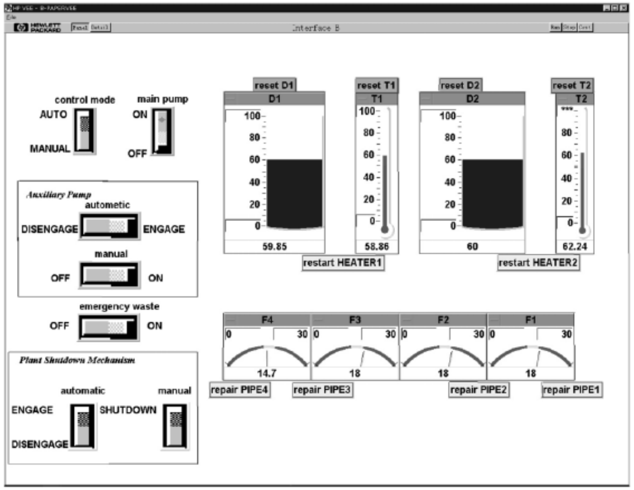


FIGURE 2 Fluid circulation microworld.

Makoto Itoh, Toshiyuki Inagaki: "A Microworld Approach to Identifying Issues of Human-Automation Systems Design for Supporting Operator's Situation Awareness", International Journal of Human-Computer Interaction, 17(1), pp. 3-24, 2004(3).



Interface A.

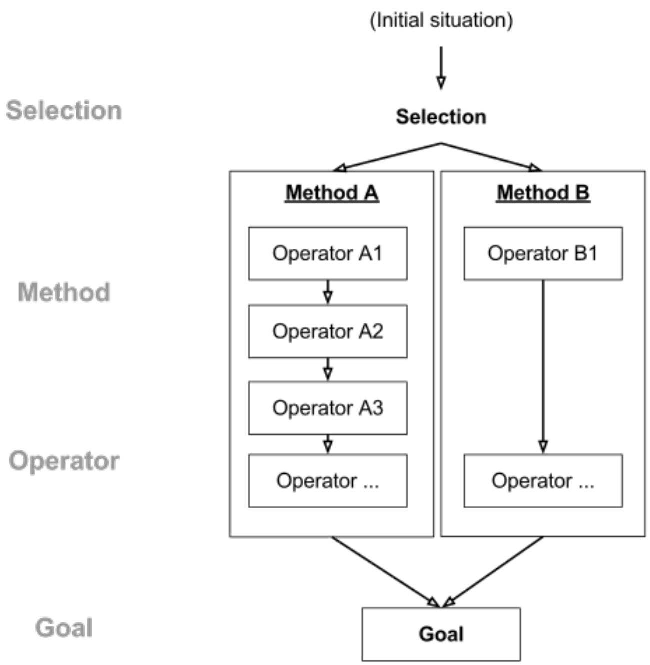


Interface B.

Human Processor Model – Interface Design and Evaluation

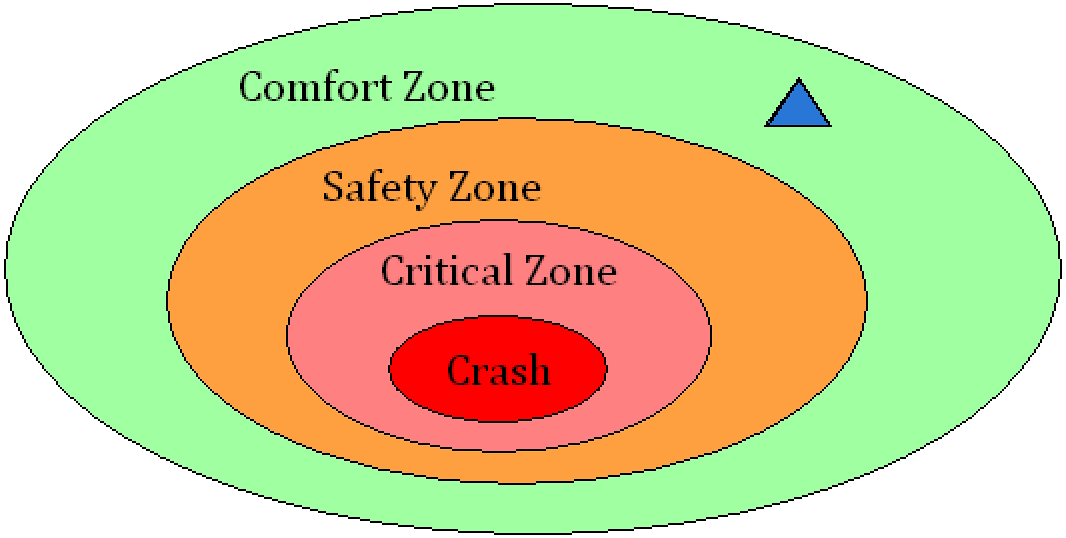
GOMS (Goals, Operators, Methods, and Selection rules)

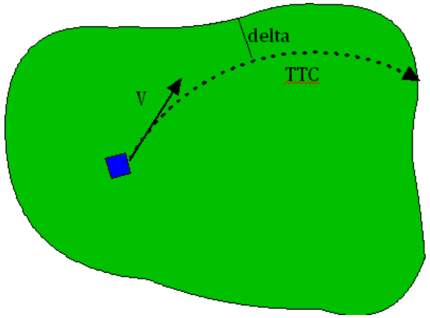
Specialized human information processor model for human-computer interaction observation. Developed in 1983 by Stuart Card, Thomas P. Moran and Allen Newell.



Parameter	Mean	Range
Eye movement time	230 ms	70-700 ms
Decay half-life of visual image storage	200 ms	90-1000 ms
Visual Capacity	17 letters	7-17 letters
Decay half-life of auditory storage	1500 ms	90-3500 ms
Auditory Capacity	5 letters	4.4-6.2 letters
Perceptual processor cycle time	100 ms	50-200 ms
Cognitive processor cycle time	70 ms	25-170 ms
Motor processor cycle time	70 ms	30-100 ms
Effective working memory capacity	7 chunks	5-9 chunks
Pure working memory capacity	3 chunks	2.5-4.2 chunks
Decay half-life of working memory	7 sec	5-226 sec
Decay half-life of 1 chunk working memory	73 sec	73-226 sec
Decay half-life of 3 chunks working memory	7 sec	5-34 sec

Safety Margin





Tolerance & Risk Management

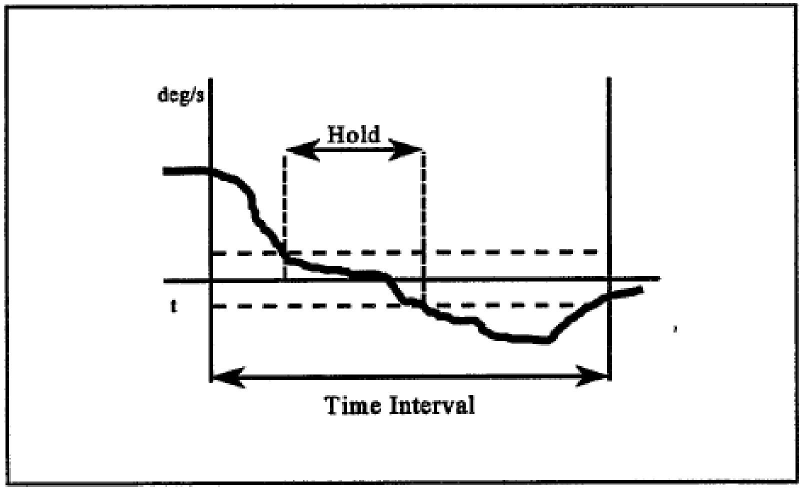
The general idea for the **safety margin concept** is to divide the time before the crash into 3 stages. In the **Comfort Zone** the system has to inform the driver, but the reaction needed (to avoid a possible accident, or to cope properly and safely with the given scenario for a specific application) is very comfortable. In the **Safety Zone** the situation is already relevant for safety and the driver has to react in a significant timeliness to safely comply with the road scenario. The **Critical Zone** is the zone just before a possible collision. In this zone, the driver has to react immediately and with the correct manoeuvre in order to avoid the accident.

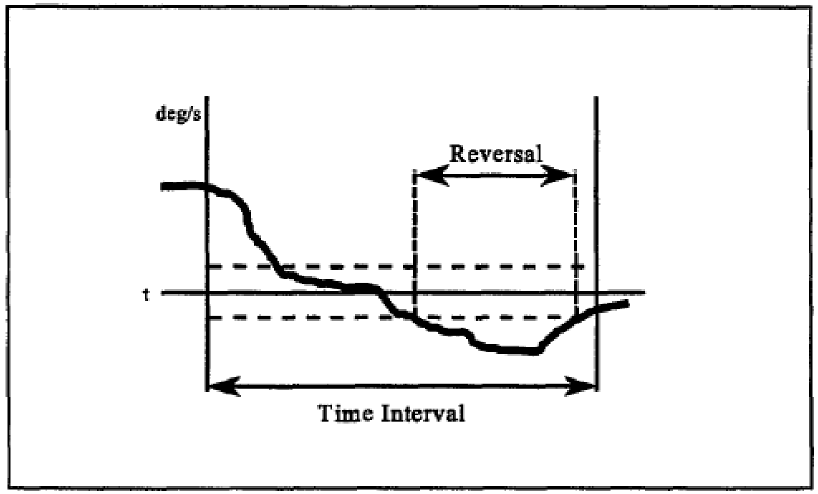
(1988). Risk Control Is Not Risk Adjustment: The Zero-Risk Theory of Driver Behaviour and Its Implications. Ergonomics, 31(4), 491-501.

Whose Performance?
Experimentor or Participant?


Relevance & Risk versus Arbitrary Performance

Steering Reversal Rates

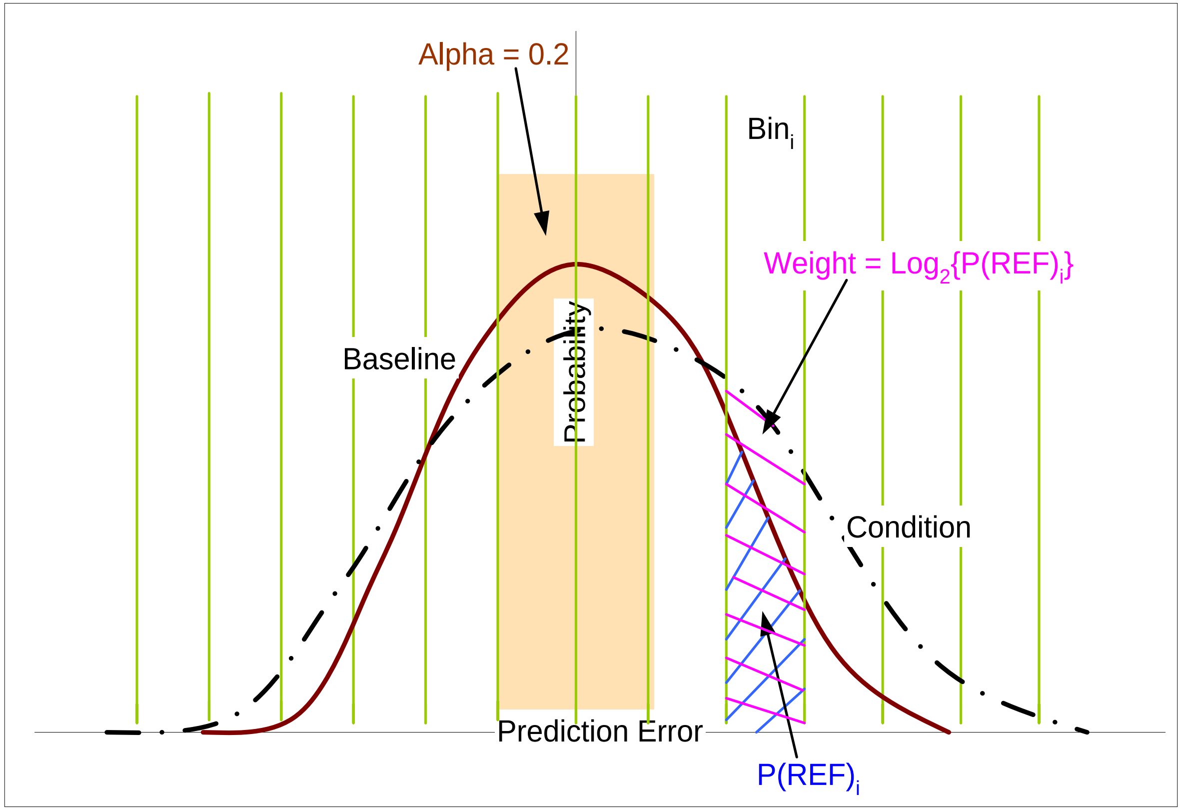




McLean, J.R., Hoffmann, R.: Steering Reversals as a Measure of Driver Performance and Steering Task Difficulty. Human Factors 17(3), 248–256 (1975)

Steering Entropy

Quantifying Driver's Own Risk Response



Nakayama, O., Futami, T., Nakamura, T., and Boer, E.R. (1999). Development of a Steering Entropy Method for Evaluating Driver Workload, SAE Technical Paper Series: #1999-01-0892.

$$H^{cond_m} = \sum_{k=1}^{K} \left\{ -P_k^{cond_m} \log_2\left(P_k^{bas_m^{ref}}\right) \right\}$$

$i=1$

Information Content

Observed Behavior: “this is an example of a huffman tree”

Encoding Behavior: 135bits

Char	Freq	Code
space	7	111
a	4	010
e	4	000
f	3	1101
h	2	1010
i	2	1000
m	2	0111
n	2	0010
s	2	1011
t	2	0110
l	1	11001
o	1	00110
p	1	10011
r	1	11000
u	1	00111
x	1	10010

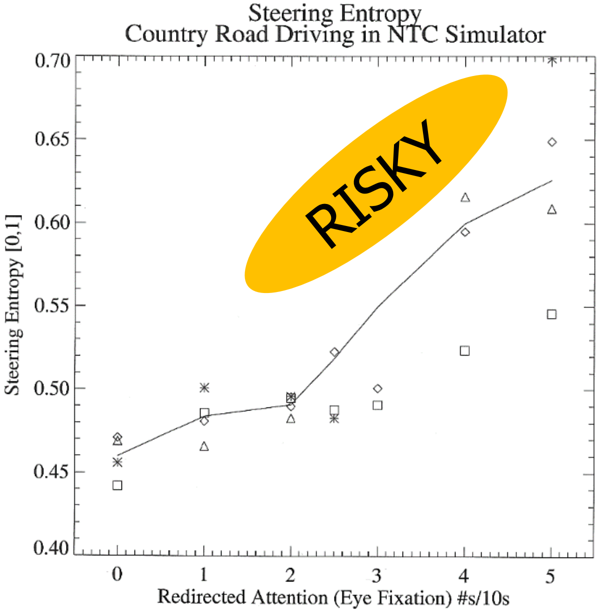
Predictability Impacts Model Complexity!

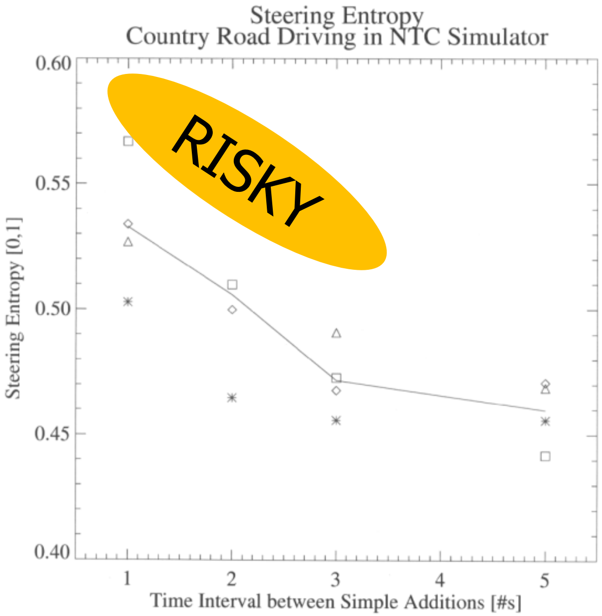
36bits “aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa”

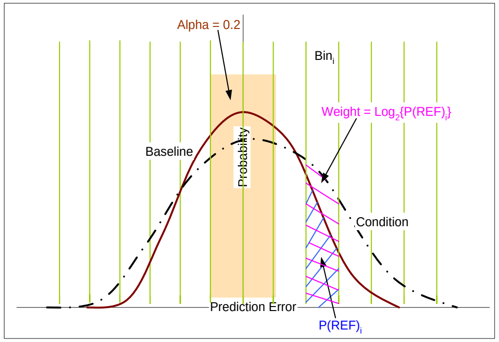
Char	Freq	Code
a	36	0

Steering Entropy

Quantifying Driver’s Own Risk Response



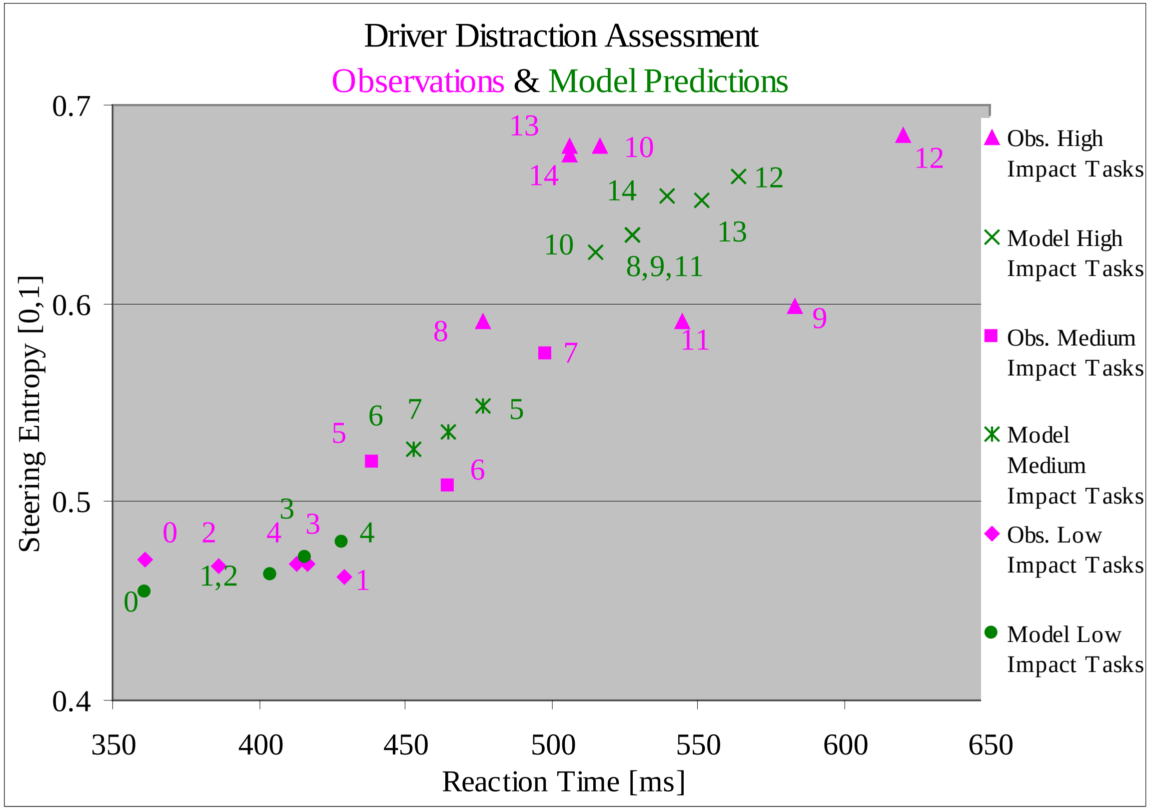




Nakayama, O., Futami, T., Nakamura, T., and Boer, E.R. (1999). Development of a Steering Entropy Method for Evaluating Driver Workload, SAE Technical Paper Series: #1999-01-0892.

$$H^{cond_m} = \sum_{k=1}^{K} \left\{ -P_k^{cond_m} \log_2\left(P_k^{bas_m^{ref}} \right) \right\}$$

Meaningful Assessment of Impact of X on Driving: Risk Impact



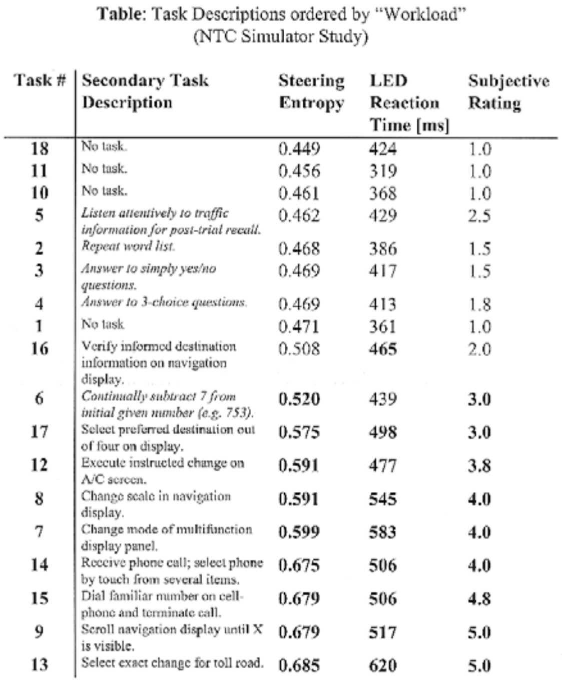


Table: Task Descriptions ordered by "Workload"
(NTC Simulator Study)

Task #	Secondary Task Description	Steering Entropy	LED Reaction Time [ms]	Subjective Rating
18	No task.	0.449	424	1.0
11	No task.	0.456	319	1.0
10	No task.	0.461	368	1.0
5	*Listen attentively to traffic information for post-trial recall.*	0.462	429	2.5
2	*Repeat word list.*	0.468	386	1.5
3	*Answer to simply yes/no questions.*	0.469	417	1.5
4	*Answer to 3-choice questions.*	0.469	413	1.8
1	No task	0.471	361	1.0
16	Verify informed destination information on navigation display.	0.508	**465**	2.0
6	*Continually subtract 7 from initial given number (e.g. 753).*	**0.520**	439	**3.0**
17	Select preferred destination out of four on display.	**0.575**	**498**	**3.0**
12	Execute instructed change on A/C screen.	**0.591**	**477**	**3.8**
8	Change scale in navigation display.	**0.591**	**545**	**4.0**
7	Change mode of multifunction display panel.	**0.599**	**583**	**4.0**
14	Receive phone call; select phone by touch from several items.	**0.675**	**506**	**4.0**
15	Dial familiar number on cell-phone and terminate call.	**0.679**	**506**	**4.8**
9	Scroll navigation display until X is visible.	**0.679**	**517**	**5.0**
13	Select exact change for toll road.	**0.685**	**620**	**5.0**

Types of “Models” to Characterize Human Operator

Task & Driver Characteristics

Optimizing / Satisficing
Performance / Effort / Risk

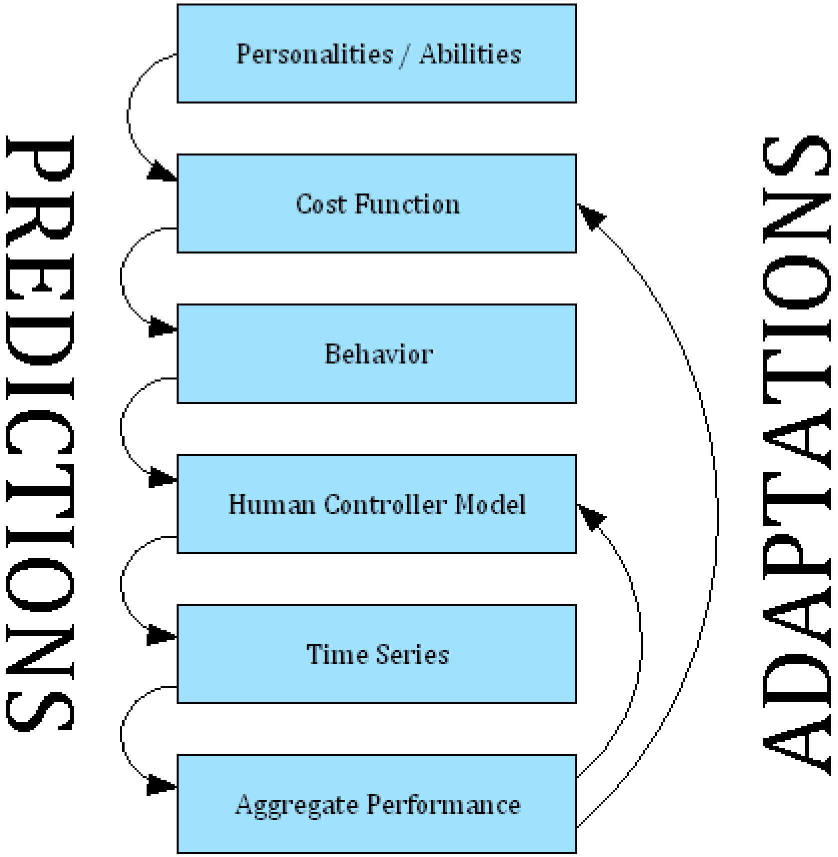
Continuous / Intermittent / Event

Compensatory / Pursuit / Preview
Closed Loop / Open Loop
Time Invariant / Adaptive

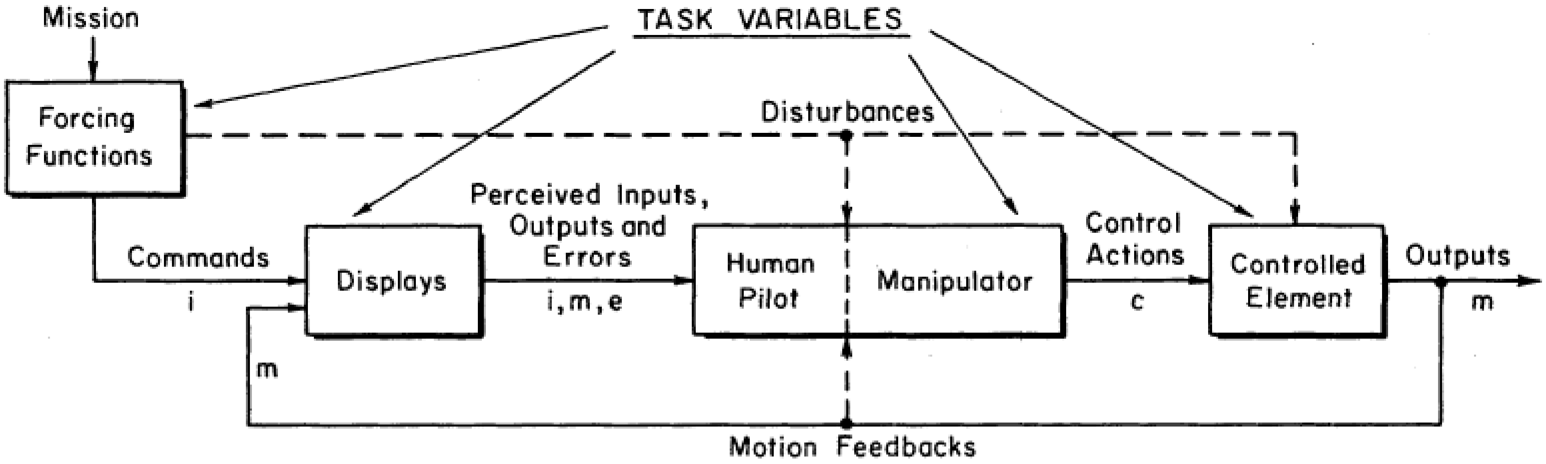
Frequency / Patterns

Statistical / Frequency

Driver Representations


Personalities / Abilities
Cost Function
Behavior
Human Controller Model
Time Series
Aggregate Performance

Factors Influencing the Human Controller



ENVIRONMENTAL VARIABLES:

- In-Flight vs. Fixed-Base
- Vibration
- G-Level
- Temperature
- Atmospheric Conditions
- Etc.

OPERATOR-CENTERED VARIABLES:

- Motivation
- Stress
- Workload
- Training
- Fatigue
- Etc.

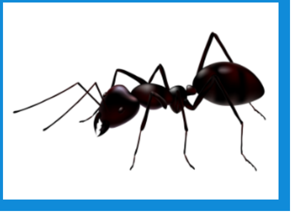
PROCEDURAL VARIABLES:

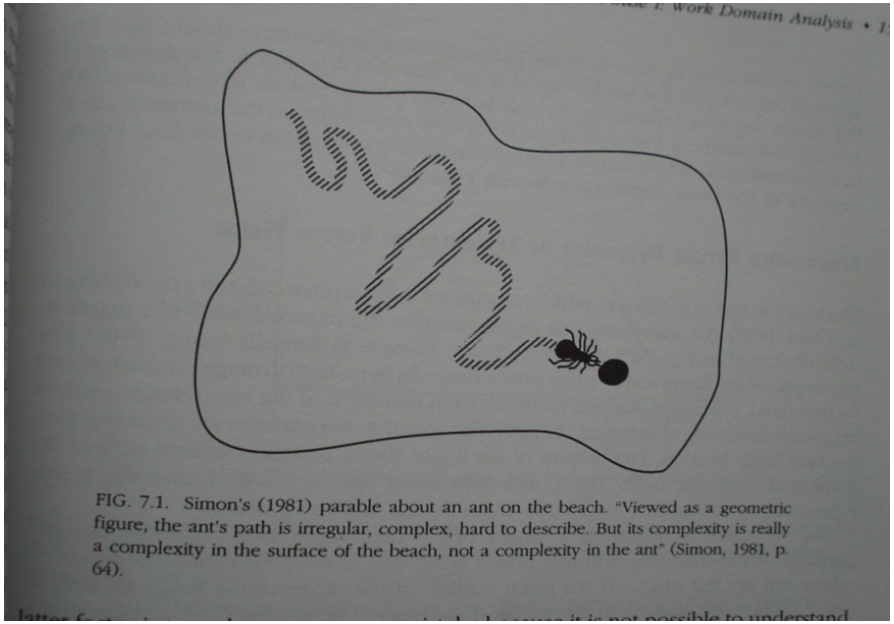
- Instructions
- Practice
- Experimental Design
- Order of Presentation
- Etc.

Fig. 1. Variables affecting the pilot-vehicle system.

D.T. McRuer, R. Jex, **A Review of Quasi-linear Pilot Models.**
IEEE Trans. Hum. Factors Electron., HFE-8 (3) (1967), pp. 231–249

Contextual Relevance
Frame Problem



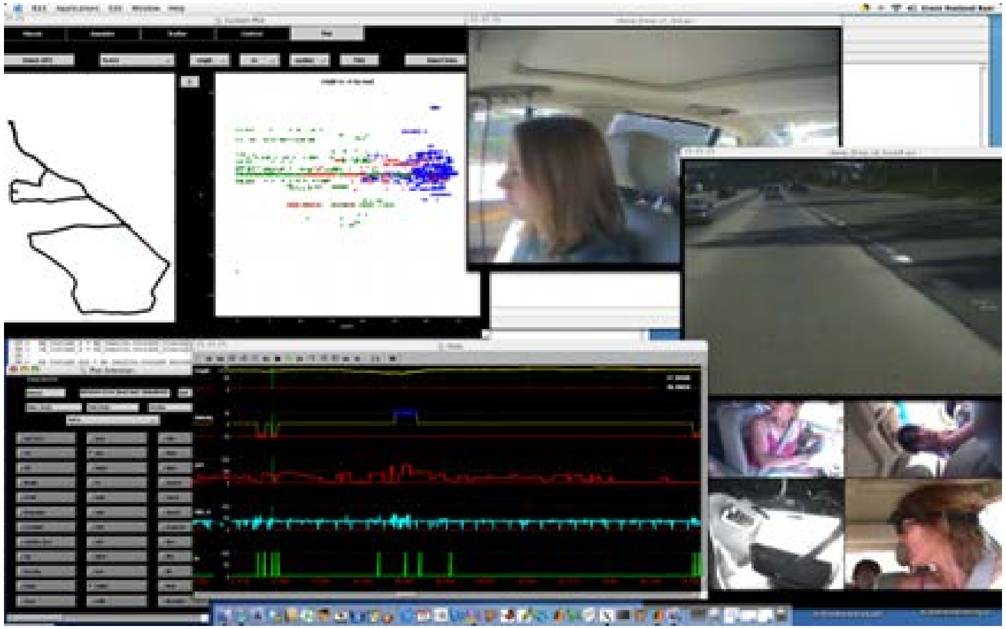


...e I: Work Domain Analysis • 1

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).

CoBeX: Contextual Behavior eXploration

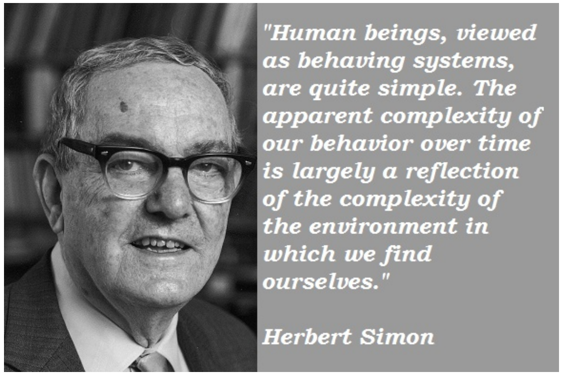
DCOG Ethnographic approach



Satisficing Decision Making

Satisficing: Suffice + Satisfy

Bounded Rationality because of Limited Cognitive Resources to Optimize



"Human beings, viewed as behaving systems, are quite simple. The apparent complexity of our behavior over time is largely a reflection of the complexity of the environment in which we find ourselves."

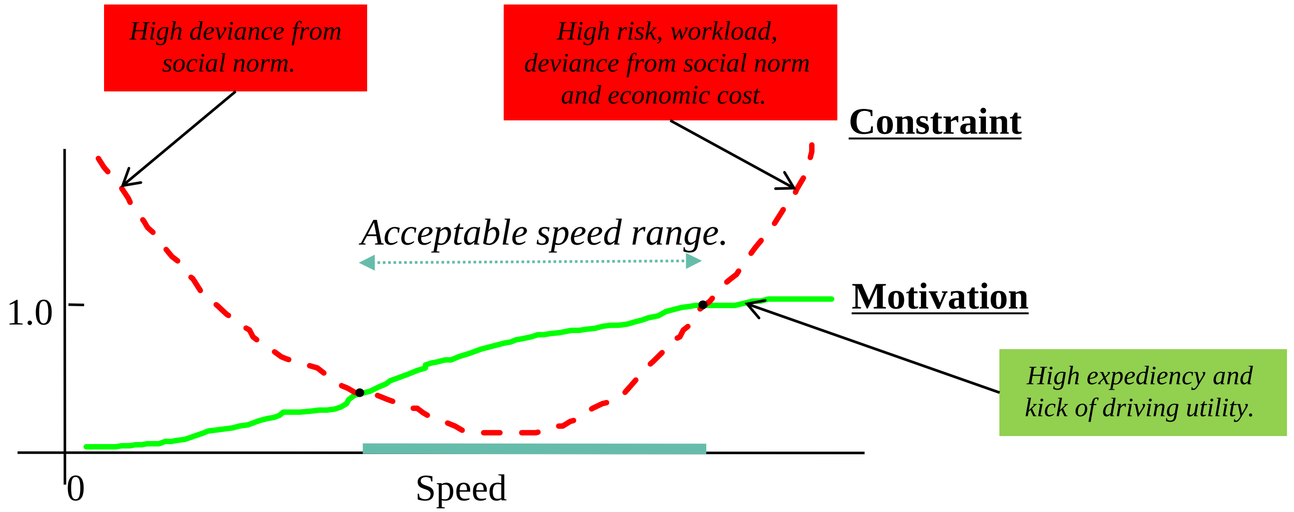
Herbert Simon

Alternatives (a_k) that are *good enough*

$$Benefit(a_k) > b\ Cost(a_k)$$

are *<u>not</u>* acted on (e.g. same behavior maintained).

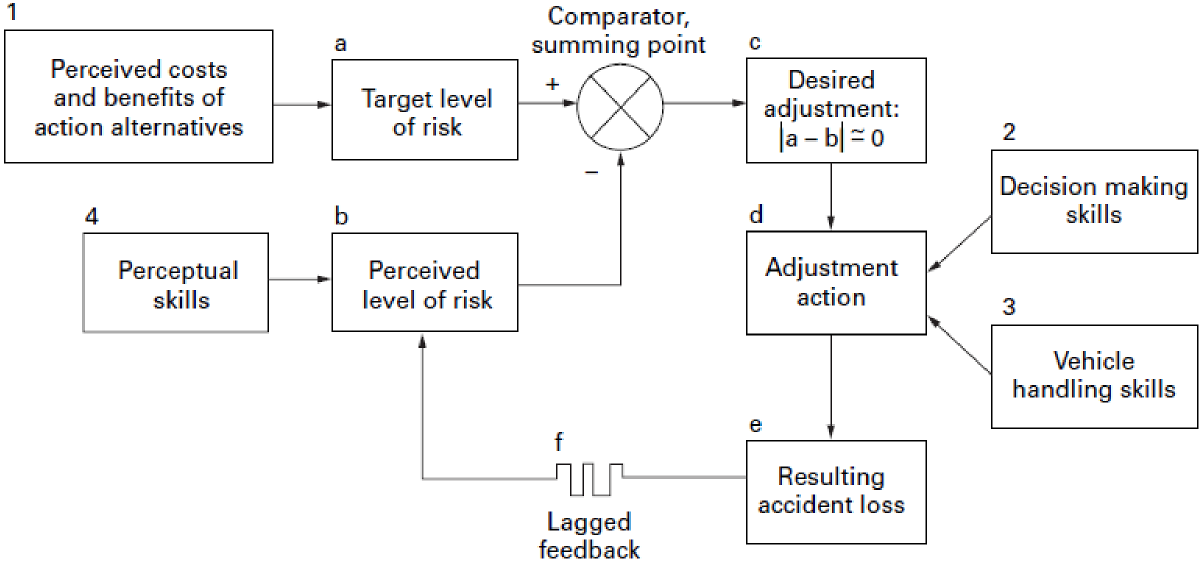
Motivational Speed Choice



NOT all VARIABILITY is NOISE!
SOME VARIABILITY is ACCEPTED!

Is this an Optimal Driving Task?

Risk Homeostasis



1 Perceived costs and benefits of action alternatives

a Target level of risk

Comparator, summing point

+

−

c Desired adjustment: $|a - b| \simeq 0$

2 Decision making skills

3 Vehicle handling skills

d Adjustment action

4 Perceptual skills

b Perceived level of risk

e Resulting accident loss

f Lagged feedback

Additional support does not always have the expected effect because people adapt.

Wilde GJS. The theory of risk homeostasis: implications for safety and health. Risk Analysis 1982;2:209–25.

Driving is incredibly dangerous and incredibly safe

- Never more than a few seconds from a potential crash.
- Generally more than 30 years away from an actual non-fatal crash

- 41,717 crashes with fatalities in the US per year
- 6,242,000 reported non-fatal crashes per year
- 2,691,335,000,000 miles traveled per year

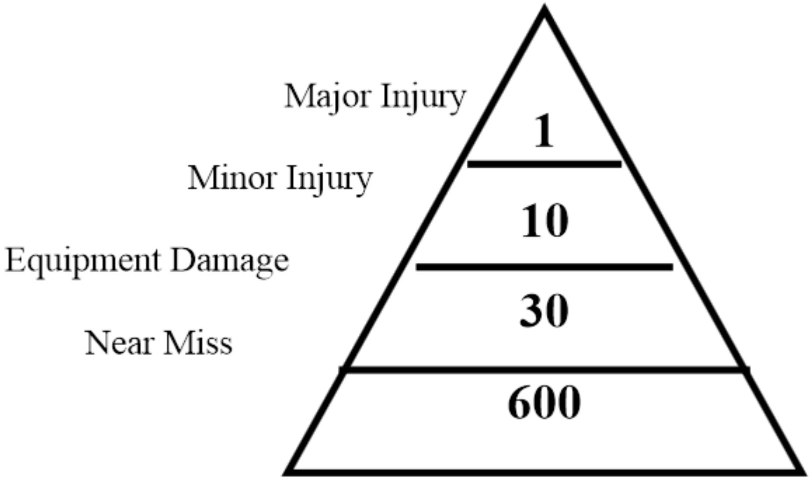
- 431,165 miles traveled between non-fatal crashes
- 10,779 hours between nonfatal crashes at 40mph
- 29.53 years between nonfatal crashes at 1 hour of driving per day.

Crashes generally are Result of Culmination of Unexpected Events

THE SAME IS TRUE FOR SUPPORT SYSTEMS

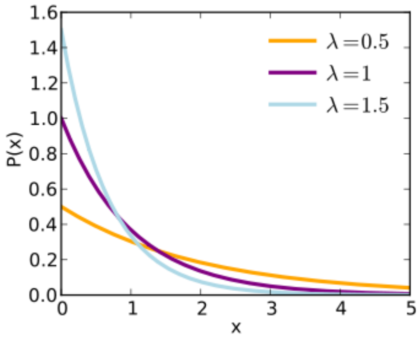
We cannot wait for crashes to make predictions about system acceptance

- Use near missed and the well known accident ratio triangles or pyramids.
    - Interpretation, if 600 near misses are observed over a period of N km of driving, then one would also expect 1 major injury.



- Propose to define near miss with TTC, then probability of a crash Exp(-beta*TTC) at the TTC definition of a near miss needs to be 1/600 given the triangle.
- With near miss TTC of 0.5s, Exp(-beta*TTC) = 1/600 yields a beta of 12.8.
- A different beta is obtained for each crash type because the pyramid is different (e.g. from naturalistic data under normal driving conditions – much data available).
- If a TTC of 0.25 is observed the crash risk becomes Exp(-12.8*0.25) = 0.04 or 1/25.
- This offers a principled way to compare risk exposure in different systems.

$$e^{-\beta\tau}$$



$$e^{-\lambda\tau}$$

How do we define TTC in different driving tasks?

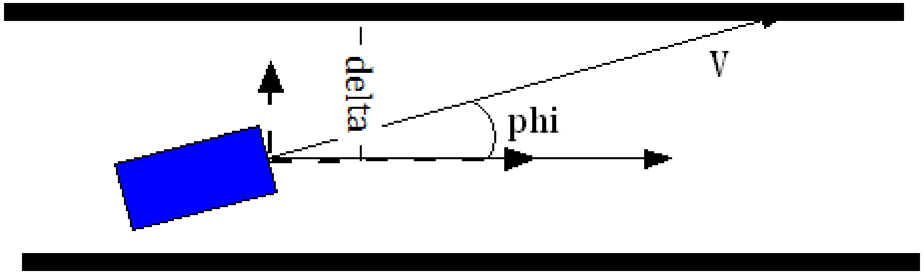
Driver Risk -- *TTC Calculation*

TTC against static behavioral constraints.

$$\tau_{pot} = \frac{\delta}{V\cos(\phi)}$$

$$\tau_{act} = \frac{\delta}{V\sin(\phi)}$$

Use Percentiles (e.g. 95% or 99% of TTC CDF)



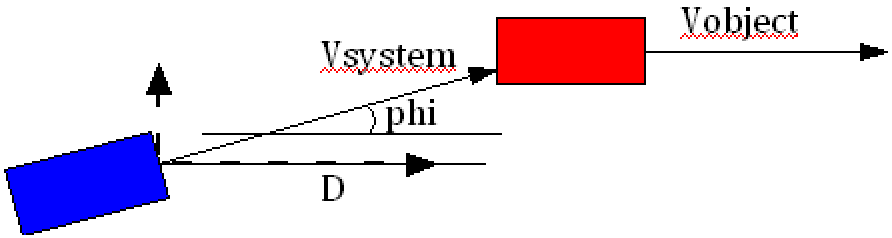
e.g. Lane Keeping

- TTC against dynamic behavioral constraints.

$$\tau_{pot} = \frac{\delta}{V_r\cos(\phi)}$$

$$\tau_{act} = \frac{\delta}{V_r\sin(\phi)}$$

$$V_r = V_{system} - V_{object}$$



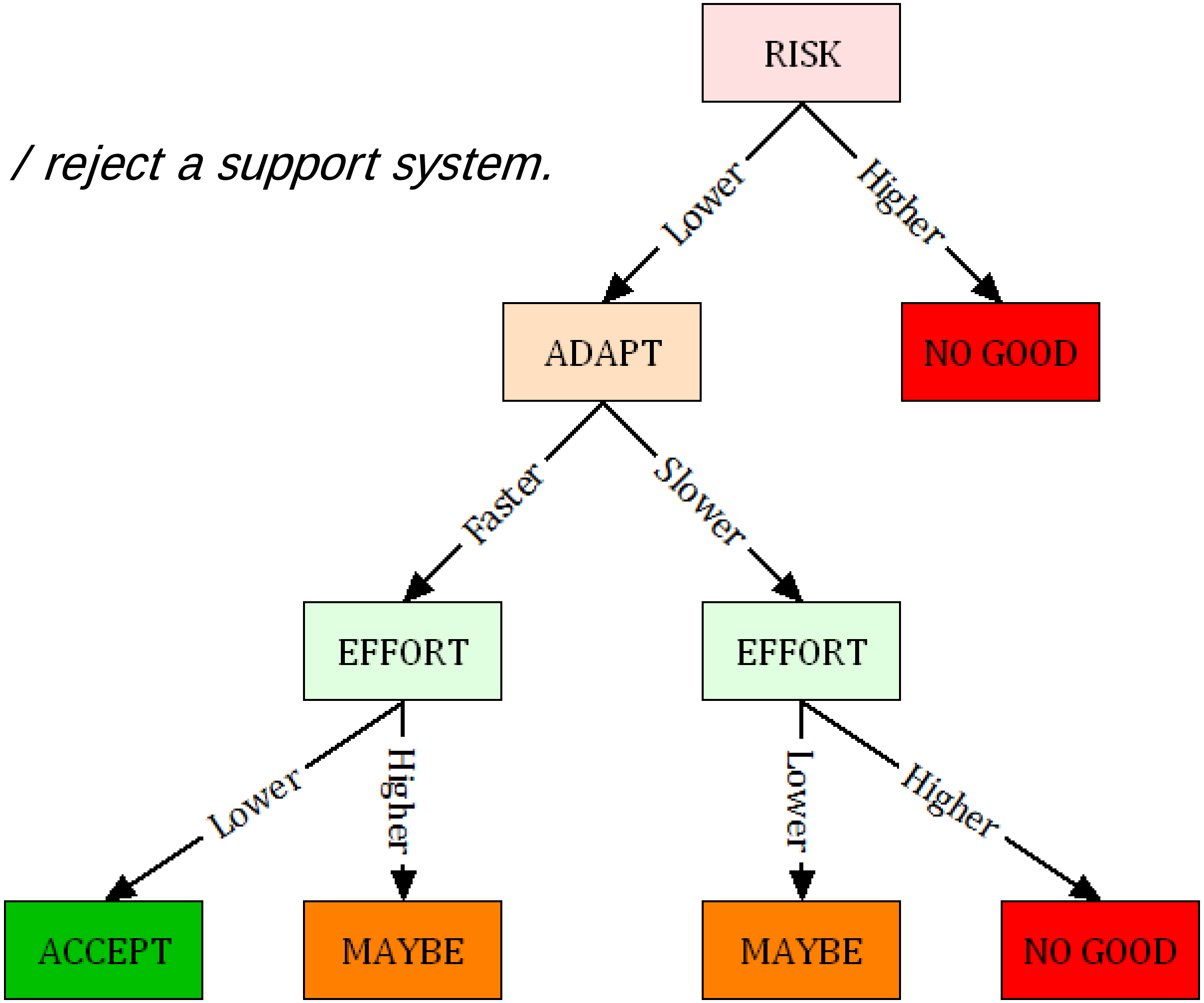
e.g. Car Following

$$Risk_{sit} = \gamma_{pot} e^{-\beta_{pot}\tau_{pot}} + \gamma_{act} e^{-\beta_{act}\tau_{act}}$$

e.g. Curve Technology

Three Stage Evaluation Tree

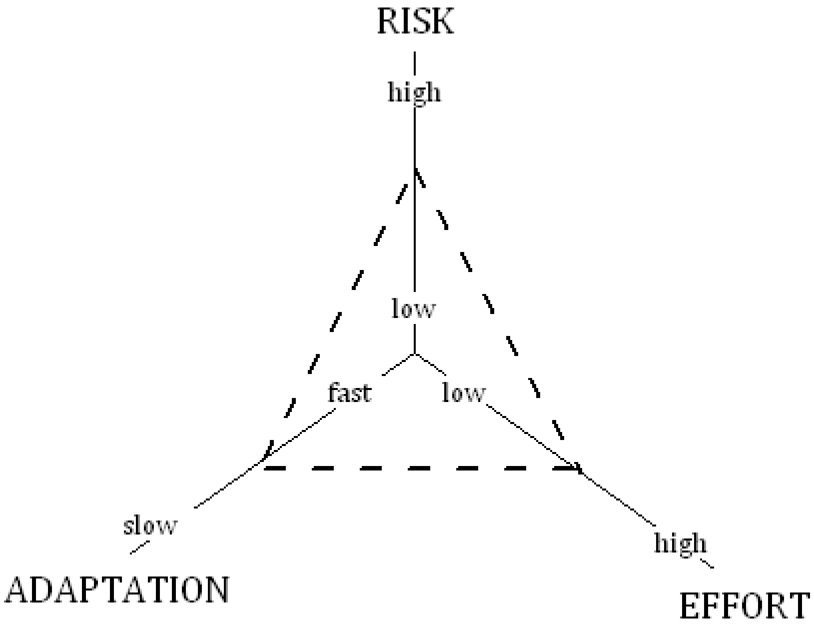
Process to accept / reject a support system.



Evaluation Triangle on Human Interaction

- **Geometric transformations in evaluation triangle from manual to supported**
  - **Only shrinkage along any or all axes is most desirable**
  - **Only expansion along any or all is most undesirable**
  - **Shift down is undesirable because effort is higher and productivity is lower (adapt to slower progression) even if the risk is lower.**
  - **Shift up is undesirable because risk is up even though effort is lower and productivity is higher.**

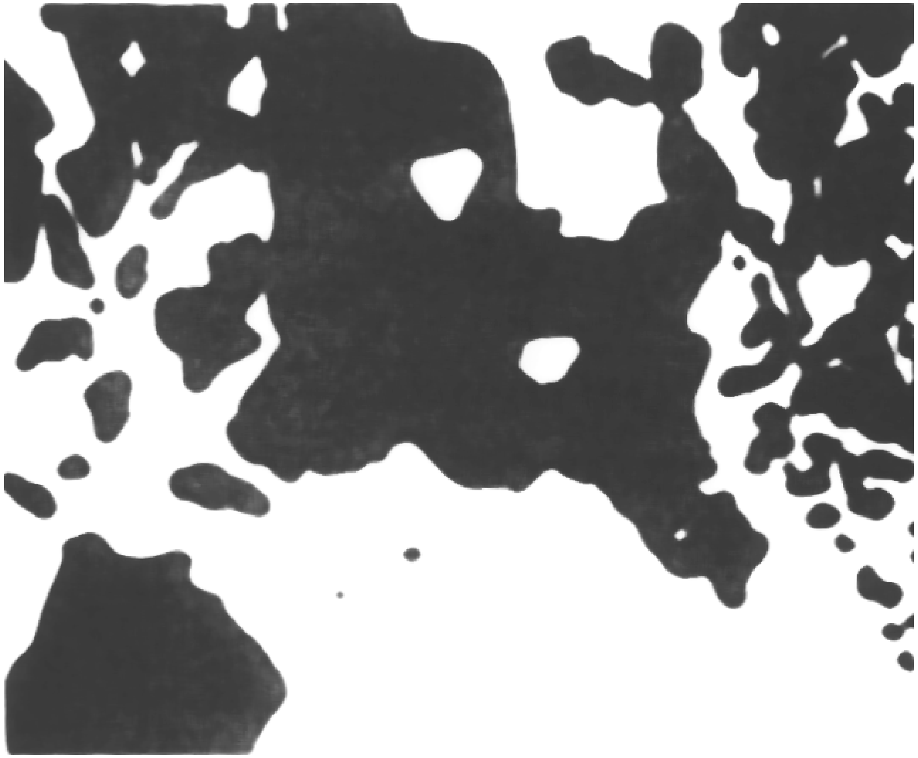
- **If axes are scaled to satisficing boundaries, then ANY shrinkage in area is treated as better as long as all three values are within the satisficing set.**
  - **An increase in risk will need to go paired with a large decrease in effort and adaptation to faster progression (see interaction risk and effort).**

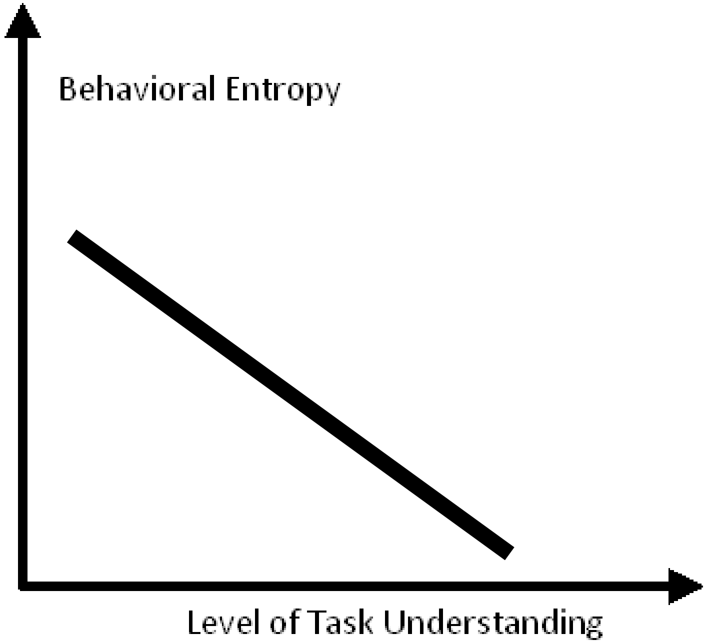


Shift down may indicate that: human is fighting with system and possibly does not trust the system or mismatch in control strategy between human and system.

Shift up may indicate that: human may trust the system too much.

How do People manage Risk
Behavioral Entropy

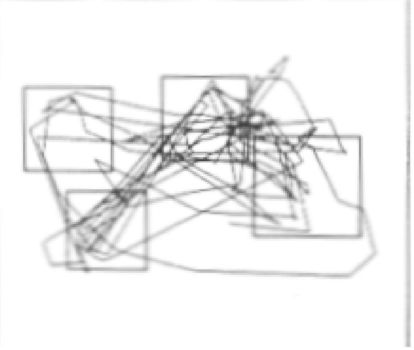




How High is your Behavioral Entropy? When & Why does it Drop?

Image

Raw Eye Scan Patterns



Fixation Location Eye Scan Patterns

Human Operator Models

Why Models?

What Models?

When Mode

Models embody our kno
operators perform a

Models off
what is

at lie beyond

ation requires a sufficiently

Opportunities: Only for the Simplest Tasks can we Predict Behavior Without Experimentation

nced models can be used to evaluate new designs without experimentation; VIRTUAL PROTOTYPING.

Model Coefficient Estimation

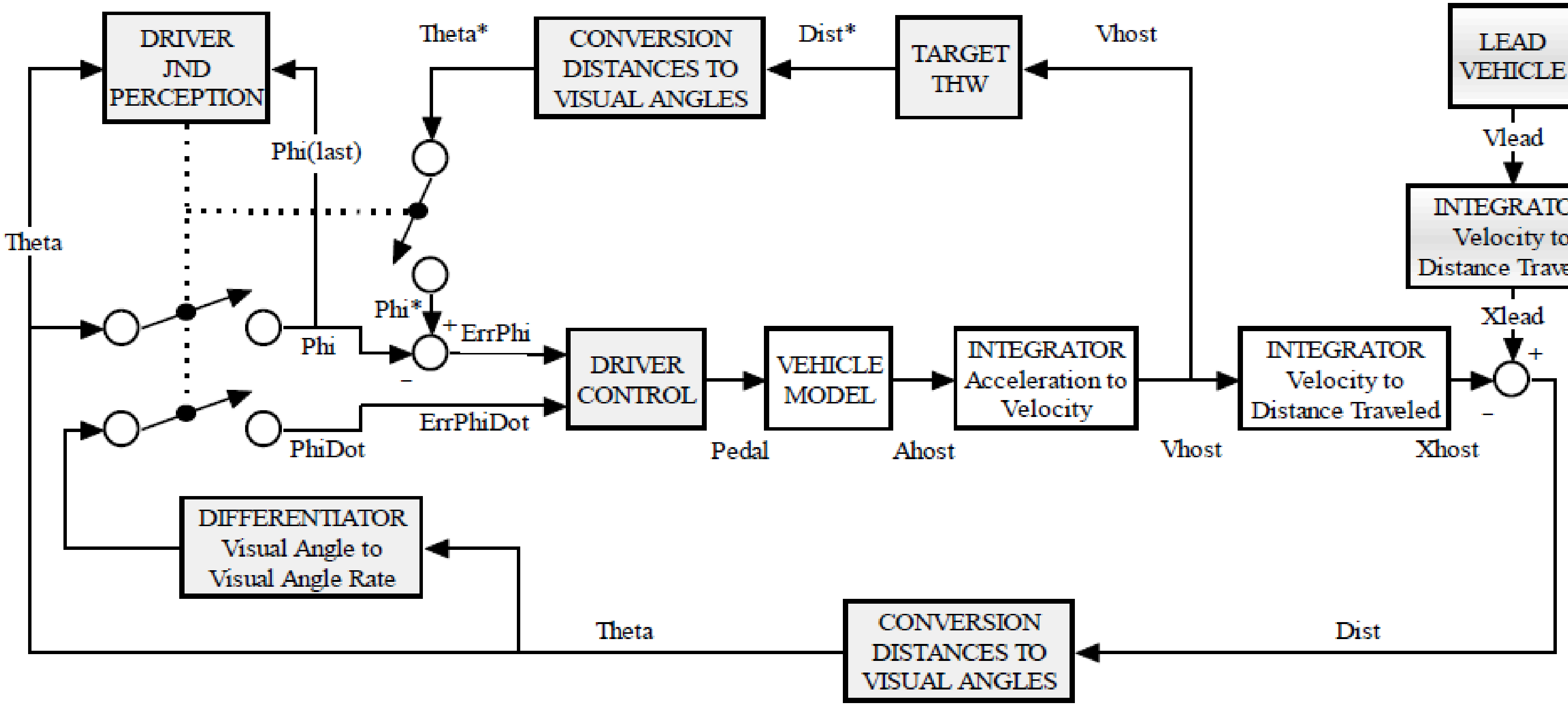
- Linear vs. nonlinear
- Static vs. dynamic
- Explicit vs. implicit
- Discrete vs. continuous
- Deterministic vs. probabilistic (stochastic)
- Deductive, inductive, or floating

- Focus next on linear, static, explicit, continuous, deterministic, inductive model.
  - Car following in Fog
  - Car following with Secondary Tasks

- Estimate model coefficients such that model predicted and observed behavior match in time, frequency or aggregate performance.

From Cost Function to Controller Gains

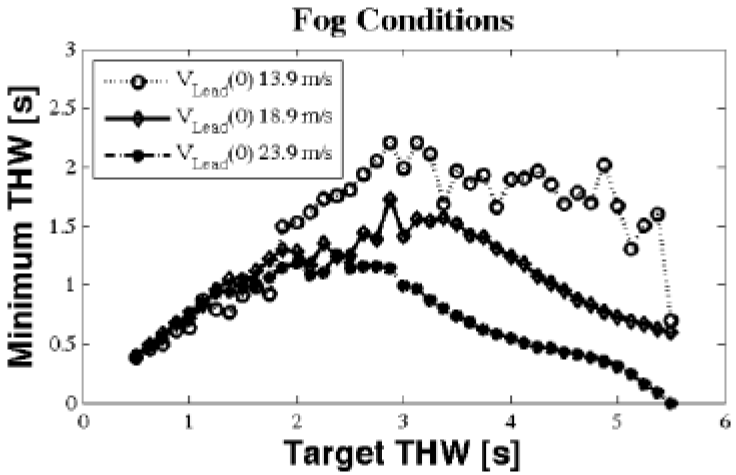
Fix Perceptual Model – Focus on Control Model

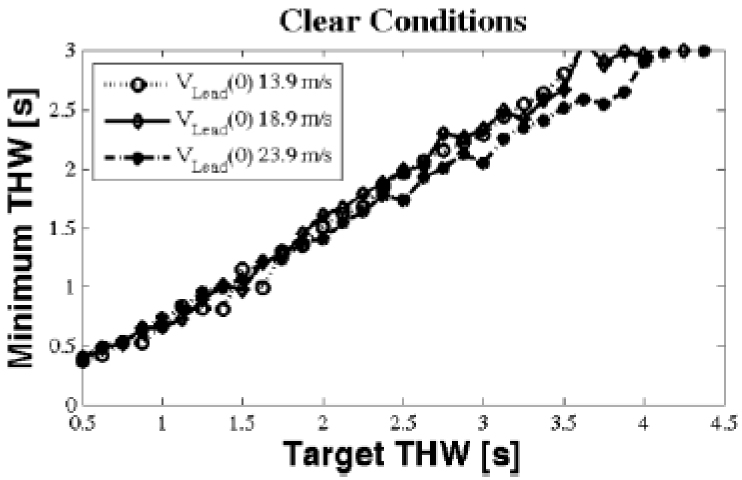


Maximize Minimum THW

Search for Optimal Model Coefficients that Minimize Cost (Maximize Safety Margin):

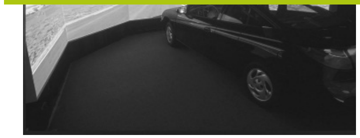
$$\{c_i\} = \min_{\{g_i\}}\left(C \mid M\left(\{g_i\}\right)\right)$$

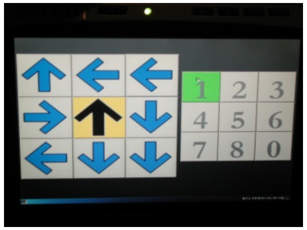


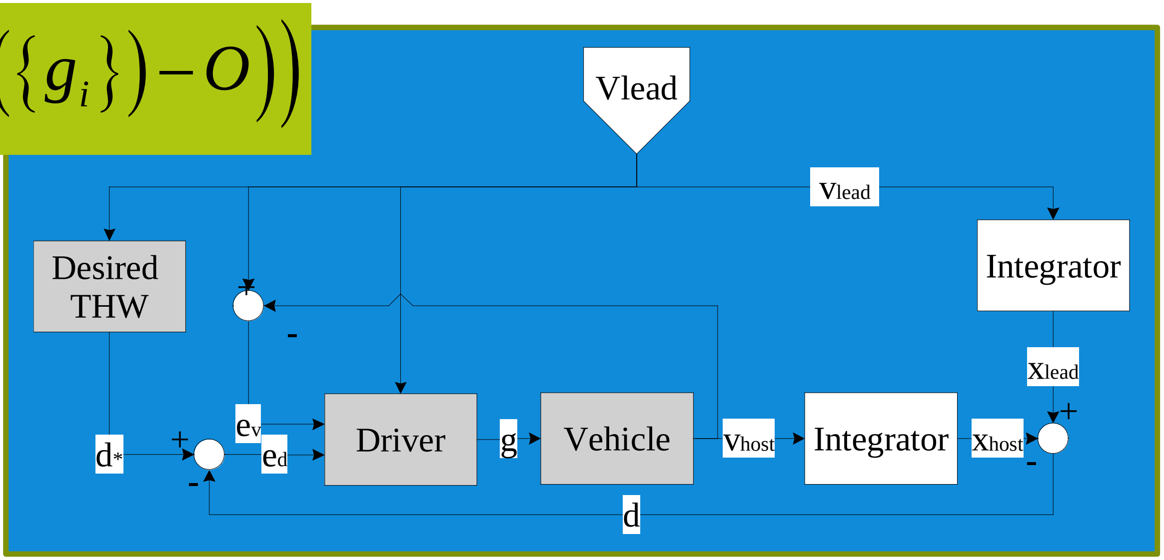


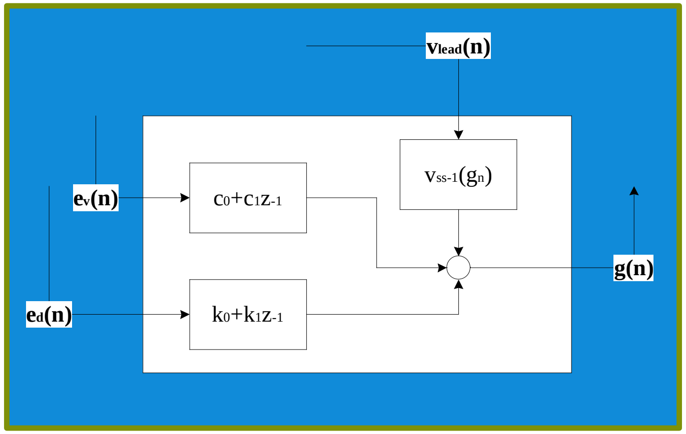
Driver Adaptation in Distracted Car Following

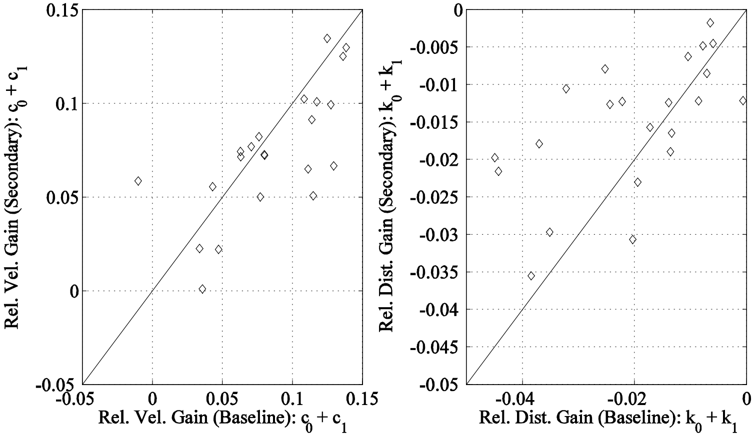
$$\{c_i\} = \min_{\{g_i\}} \left(Fit\left(M\left(\{g_i\}\right) - O \right) \right)$$





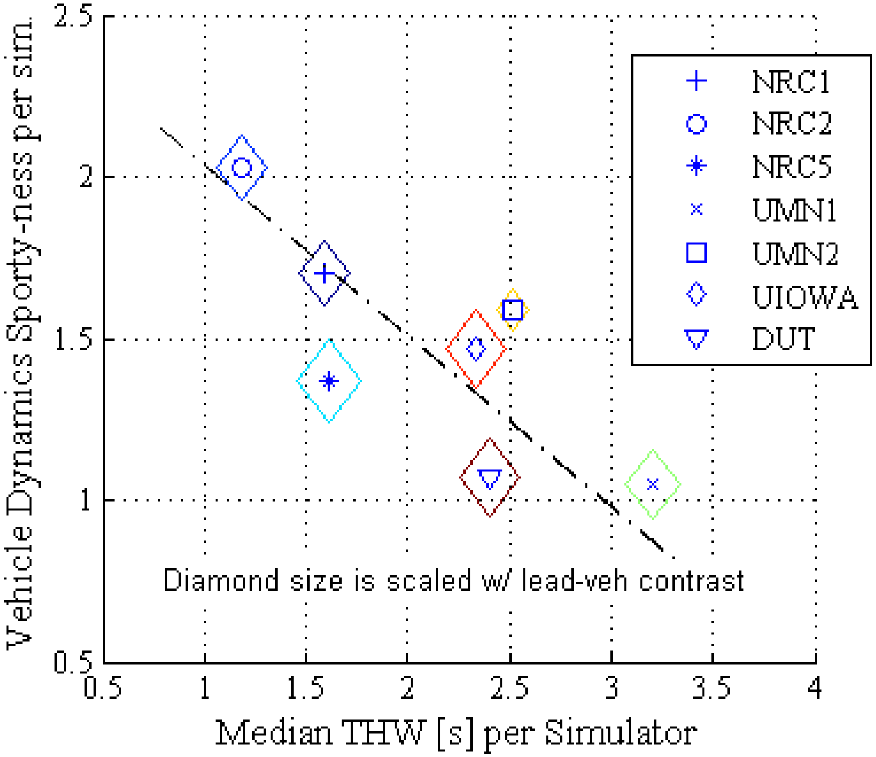


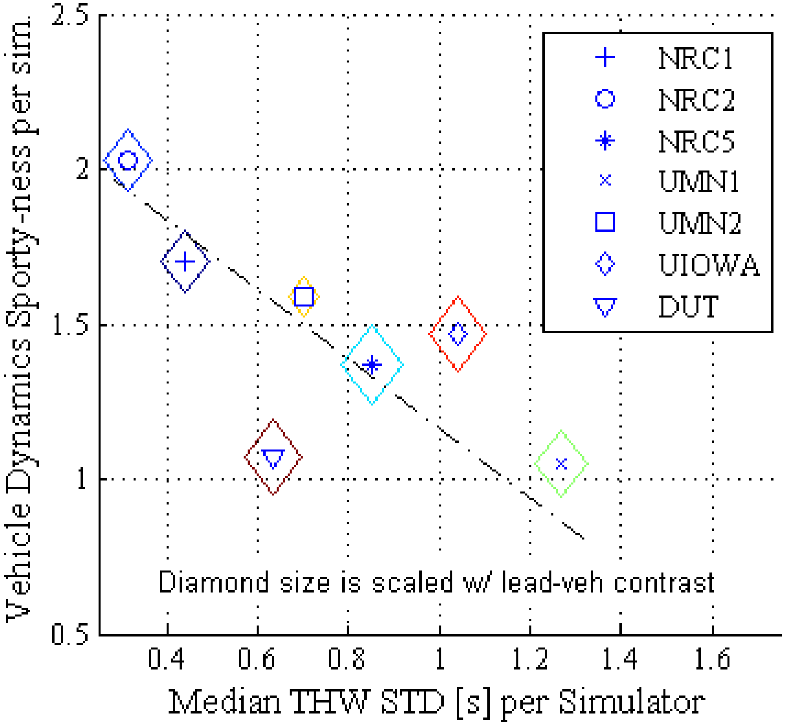




Are we Forgetting a Strong Factor that Impacts Human Machine Interacttion?

Relationship of Vehicle Dynamics and Contrast & THW Control.

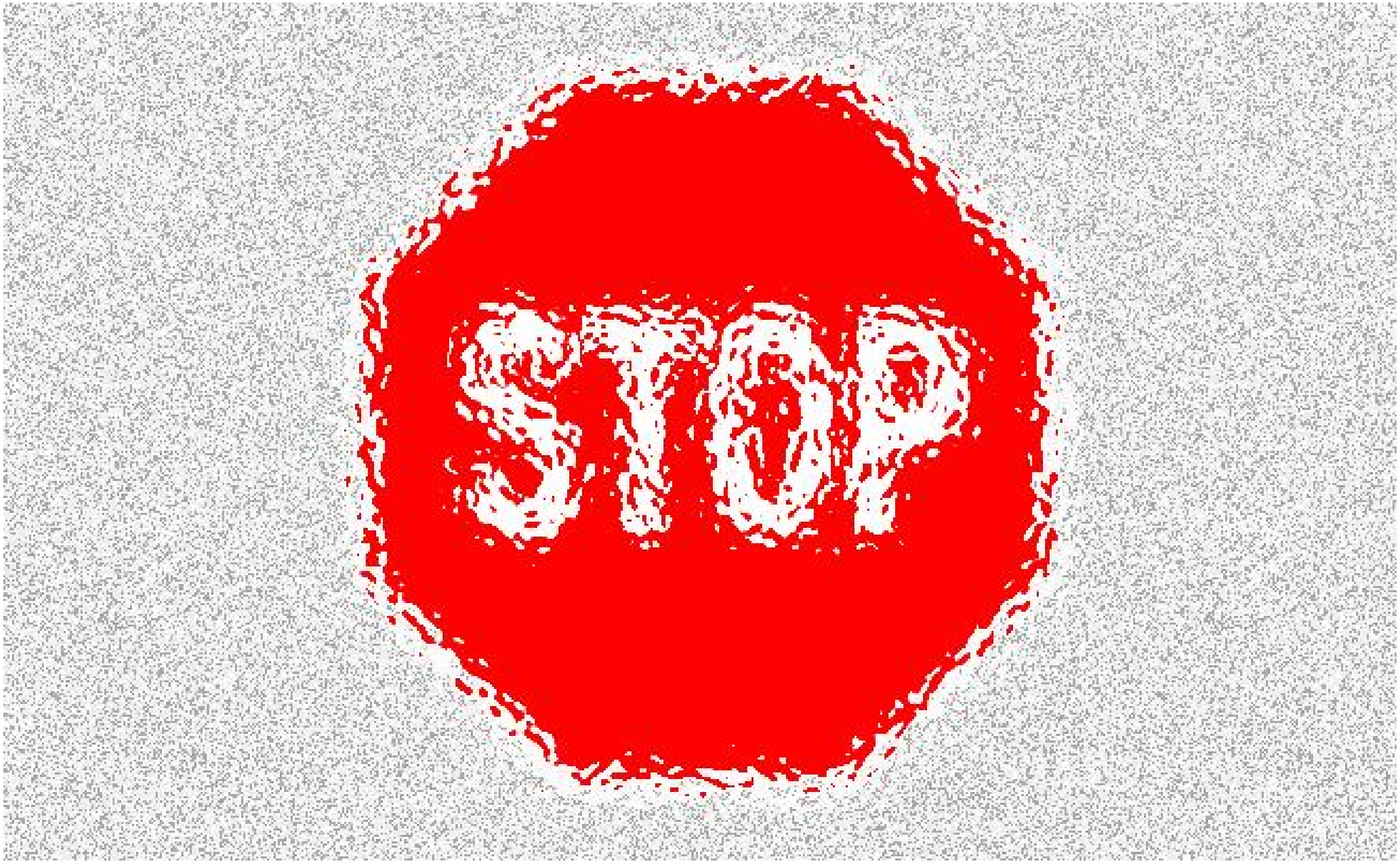




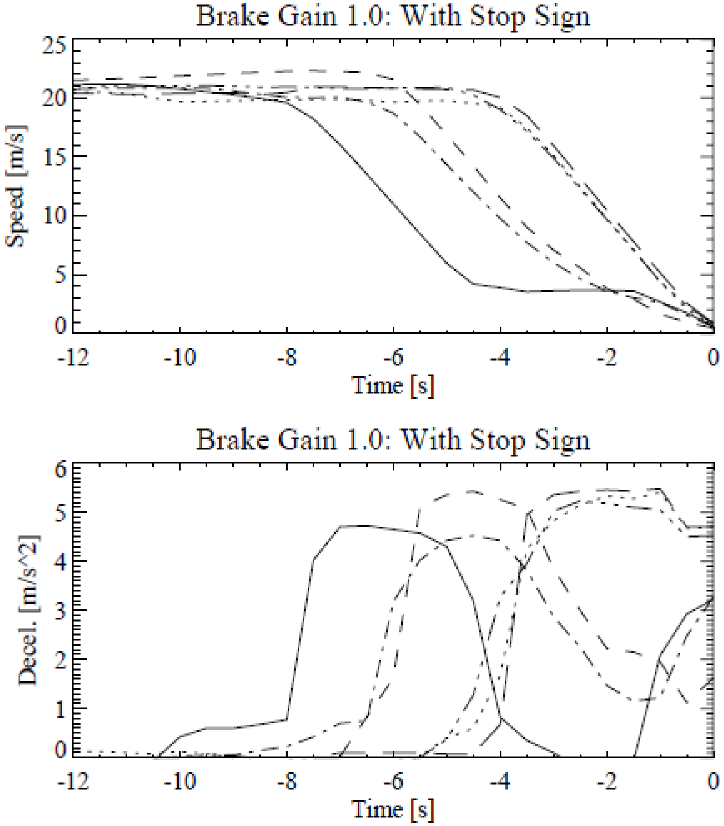
StdTHW lower for shorter THW

Follow closer when perceptibility and controllability of gap change are high.

Evaluation in Simulated Environments



Evaluation in Simulated Environments



Compared to reality:

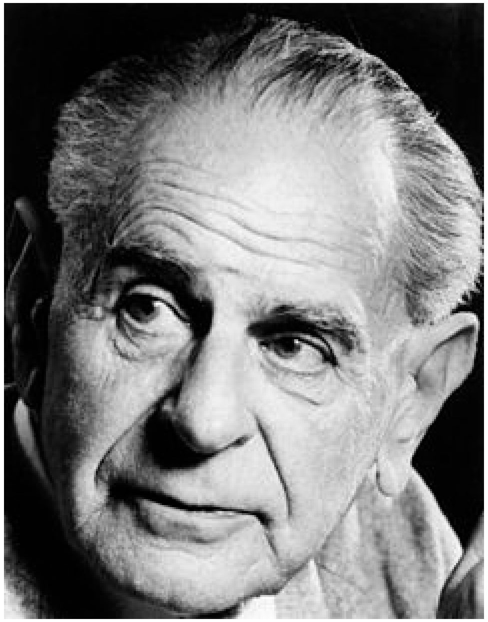
- Why do people start decelerating earlier?
- Why do people not stop with a constant deceleration rate?

Understand limitations of driving and other simulators in HMI evaluation: WORKLOAD!

Boer, E. R., Kuge, N., & Yamamura, T. (January 01, 2001). Affording realistic stopping behavior: A cardinal challenge for driving simulators. *Human-centered Transportation Simulation Conference.*

Testing Hypotheses

Popper is known for his rejection of the classical inductivist views on the scientific method, in favour of empirical falsification: A theory in the empirical sciences can never be proven, but it can be falsified, meaning that it can and should be scrutinized by decisive experiments.



Karl Popper c. 1980s

What is the probability that a theory is false? When is the probability that an experimental result correctly rejects the hypothesis and what is the probability that it falsely accept the tested hypothesis?

http://en.wikipedia.org/wiki/Karl_Popper

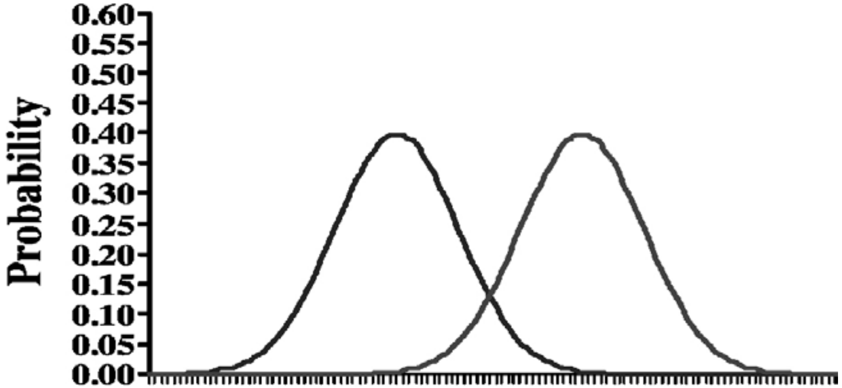
Statistical Comparison of Conditions

Possible Variables:

- cost weights
- model coefficients
- time series metrics

- subjective rating
- ...

Group A | Group B
or | or
Cond A | Cond B



Observed Behavior

SD: Standard deviation of observations
M: Mean of the observations
SE: Standard deviation of mean of observations

Distributions are of the mean across the population!

For normal distributions 95% confidence intervals for M are:
Upper 95% limit = M + 1.96 SE
Lower 95% limit = M - 1.96 SE

5% significance means that when the experiment would be repeated 100 times, only 5 times would the Null no the results be opposite.

Significance of Experimental Observations

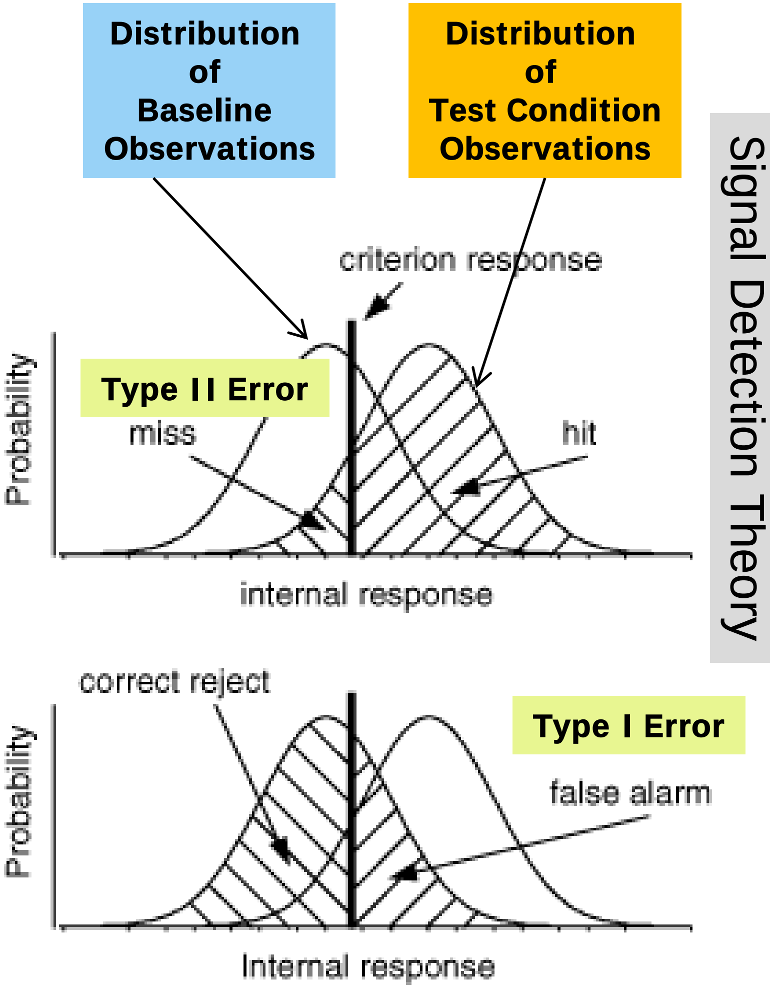
NULL hypothesis H0 assumes that the tested condition or chosen group is NOT DIFFERENT from the baseline condition or groups; H0 assumes that baseline and test are the SAME

- Type I Error: False rejection of H0; tested condition assumed to be different where as in reality it is not.
- Type II Error: False acceptance of H0; test condition assumed to be the same where as in reality it is different.

Distributions are of the mean across the population!

Sir Ronald Aylmer Fisher (1890–1962) stressed that the "null hypothesis":
… is never proved or established, but is possibly disproved, in the course of experimentation. Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis.
—1935, p.19

Signal Detection Theory



Distribution of Baseline Observations

Distribution of Test Condition Observations

Type II Error

Type I Error

Significance assessment of Null Hypothesis

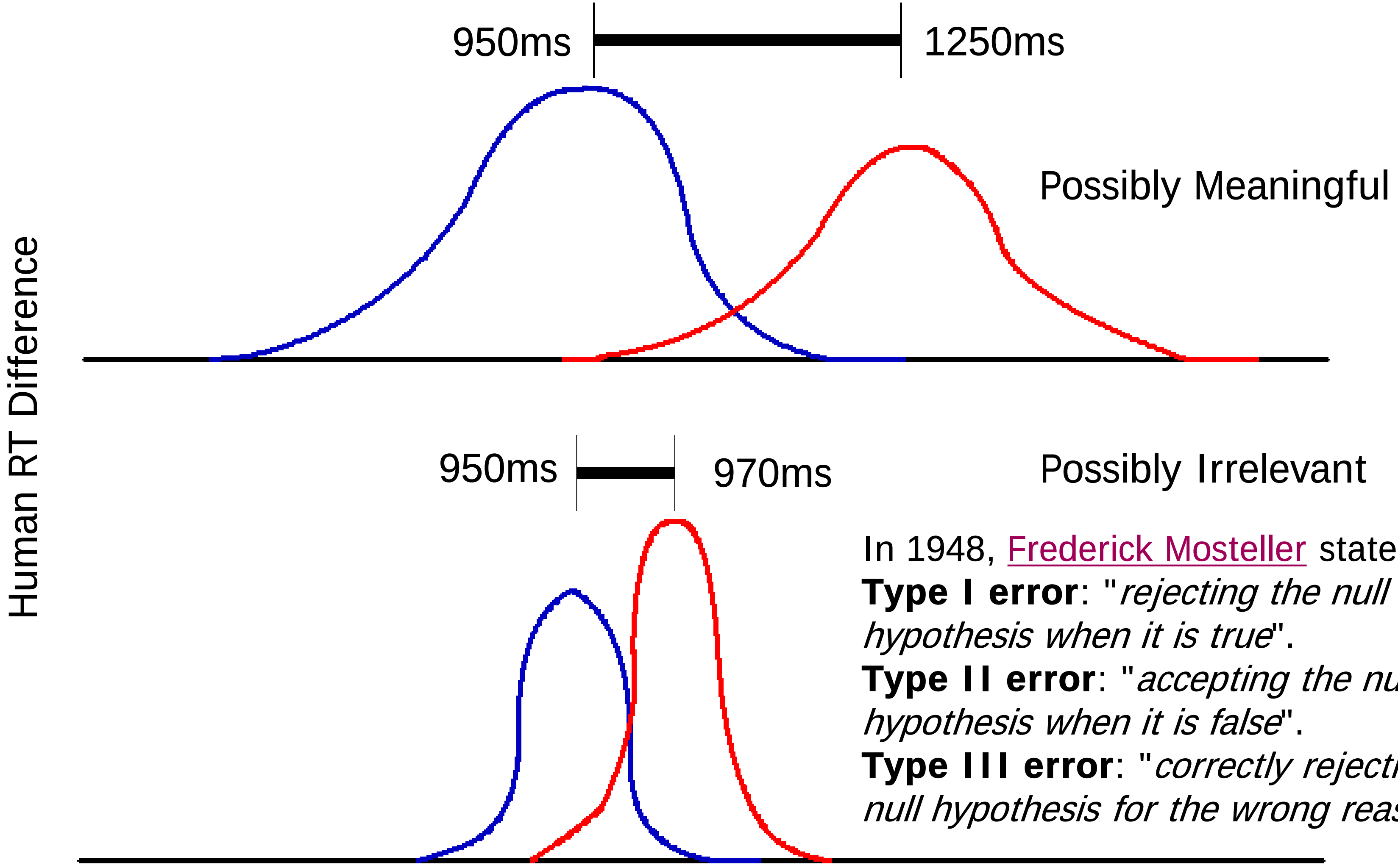
Parametric:

- Makes explicit assumptions about the type of distribution that characterizes the observations (e.g. LogNormal for human reaction times)
- Type I and Type II errors can be computed directly.

Non-Parametric:

- Does not make assumptions about the type of distribution that describes the observations.
- Type I and Type II errors require calculation of a metric of the difference between the two sets of observations; the distribution of this metric is know analytically or in tabular form and used to determine significance of similarity between baseline (control) condition and test condition.
- Example: Mann-Whitney U-Test.

Statistical Significance vs Magnitudal Relevance



In 1948, Frederick Mosteller stated:
Type I error: "*rejecting the null hypothesis when it is true*".
Type II error: "*accepting the null hypothesis when it is false*".
Type III error: "*correctly rejecting the null hypothesis for the wrong reason*".

Learning Goals Lecture 4

After this lecture, you will be able to:

1. Reproduce:
    a. Progression in our understanding of human functioning through time.
    b. Progression in design of systems and human machine interaction through time.
    c. Key types of designs that enhance human machine interaction.
    d. Classical ways to assess human machine interaction performance.
    e. Ways to characterize and measure human behavior.
    f. Statistical means to compare systems or conditions in human machine interaction
2. Apply:
    a. Methods to characterize human machine interaction behavior.
    b. Methods of evaluation to determine what system or interface is better.
3. Be critical of:
    a. The ways in which a human machine interaction task can be supported.
    b. The limitations in evaluating human machine interaction from a single narrow perspective.

Questions?