Lumberjack NASA. OCT 2021

What lumberjacks have to do with human-automation interaction

The higher the tree, the harder it falls

Chris Wickens Colorado State University

The Lumberjack talk: A tribute to Raja Parasuraman: My friend, co-author and colleague: complacency, degree of automation and Air Traffic Control.

- Background: imperfect automation and situation awareness, what **is** the lumberjack
- Two NASA related Experiments (S and A)
- Experiment 1: the Robotics Arm
- Experiment 2: the imperfect Air Traffic Control conflict Resolution aid and the benefits of **automation transparency.**

- **A computational model of automation complacency**

The Lumberjack Analogy (Wickens & Sebok, 2016): Automation helps when it is correct (the tree stands) and hurts when it fails (the tree falls): the higher it stands, the harder it falls.

Five layers of argument.

1. Automation helps
2. Automation can fail
3. When automation fails, human may show a poor response in failure management: related to loss of situation awareness.
4. The higher **degree of automation**, the greater benefits when automation **performs correctly** and the greater problems when **automation fails**.
5. These greater problems can be mitigated by **transparency** which supports **situation awareness**.

1. **Automation provides a benefit**, reducing workload and increasing performance, safety and productivity. (If it didn't we wouldn't use it)

2. **Imperfect automation**. Automation CAN "fail":

- The software or hardware failure (the software bug; the calculator battery dies)
- The Tesla Crash

- Fails to act in ways expected by the human who has a different "mental model" of the process/logic: the 737 Max automation-related crashes. Aircraft FMS accidents.

3. WHEN automation fails, the human sometimes cannot deal with the failure; cannot cope. Disaster in safety critical systems that may offset the benefits when automation works correctly.

- Three Mile Island nuclear power plant
- The 2008 financial crash and "complex derivatives", triggered automation trading decisions
- **The 2013 Asiana airlines crash** approaching San Francisco Airport. The aircraft flight management system.
- Air France crash over the Atlantic
- The Tesla crash again

Why does the human sometimes fail to cope with automation failures?: The Psychology of the operator.

OOTLUF: Out of the loop unfamiliarity:

The human poor response when automation unexpectedly fails. Breakdowns of:

- **Noticing that** something is wrong (change blindness). **Loss of Endsley's stage 1 situation Awareness** Complacency: The failure to **monitor** the otherwise reliable system. The China Airlines barrel roll. Failure to monitor engine thrust.

- **Understanding what** is wrong. A loss of stage 2: **situation awareness** of the state of the system (what is the current aircraft attitude) → delayed or incorrect response. This results from the fact that the human is no longer making active responses when automation is doing so: the **generation effect**)

and because much automation lacks "transparency": what is it doing and why?

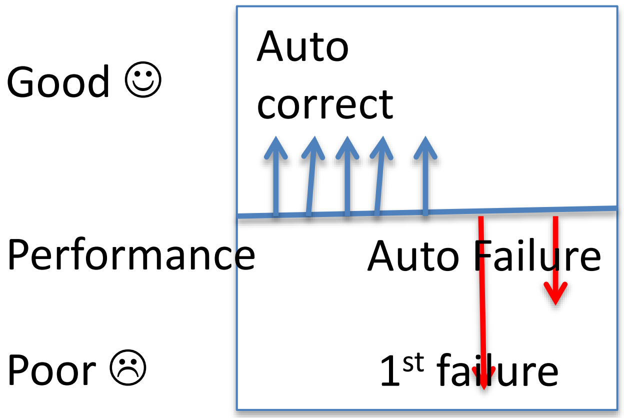
- **De-skilling**. Used automation for so long that the manual skills it replaced have degraded: Hand flying the airplane. Forgetting long division when the calculator fails.

OOTLUF: Out of the loop unfamiliarity:

The human poor response when automation unexpectedly fails. Breakdowns of:

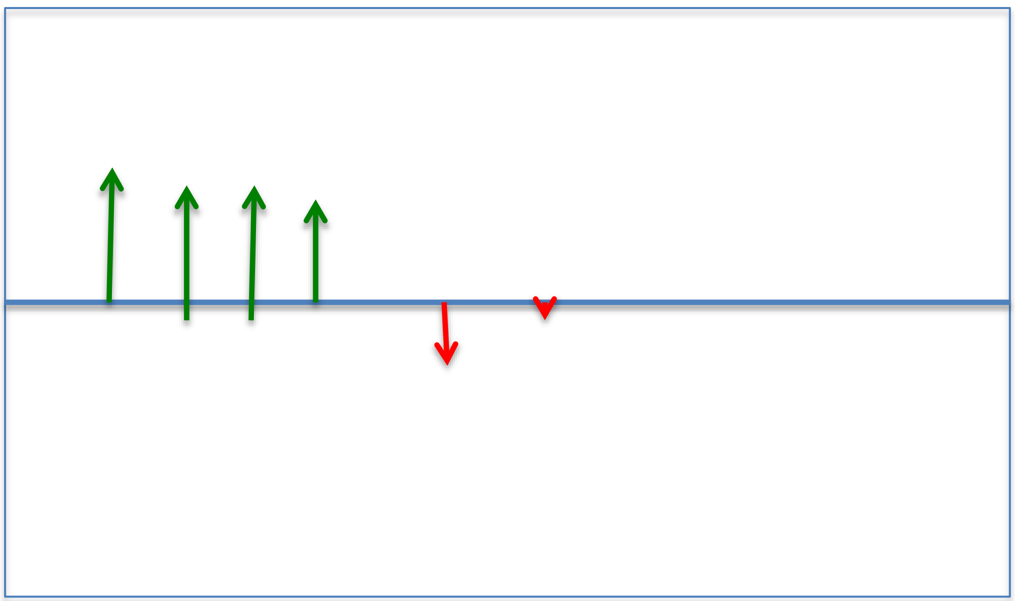
- **Noticing that** something is wrong (change blindness). Complacency. The failure to monitor the otherwise reliable system. The China Airlines barrel roll. Failure to monitor engine thrust. **Loss of stage 1 situation Awareness.**
- **Understanding what** is wrong. A loss of stage 2: **situation awareness** of the state of the system → delayed or incorrect response. This results from the fact that the human is no longer making active responses (generation effect), and because much automation lacks "transparency": what is it doing and why?
- **De-skilling**. Used automation for so long that the manual skills it replaced have degraded: Hand flying the airplane. Forgetting long division when the calculator fails.

Thus:



This is a poor OOTLUF human response to the automation failure (the "failure response") is particularly poor the **first time** automation fails: the **"first failure",** *because this failure is totally unexpected. (Sometimes a black swan).*

The challenge: Can we fix the automation **failure response**, without reducing the automation benefits? We think so: transparency

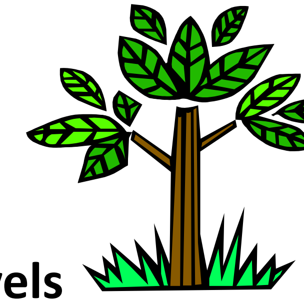


a “free lunch”?

Five layers of argument.

1. Automation helps
2. Automation can fail
3. When automation fails, human may show a poor response in failure management: related to loss of SA.
4. The higher **degree of automation**, the greater benefits when automation **performs correctly** and the greater problems when **automation fails**.
5. These greater problems can be mitigated by **transparency** which supports **situation awareness**.

4. Degree of Automation: The Higher Tree



Approaches to automation have characterized it by **Stages & Levels**
(Parasuraman Sheridan & Wickens, 2000; Onnasch Wickens Li & Manzey, 2015). NASA Johnson.

Stages: Automation assists human performance at 2 stages of information processing. {actually 4 stages; but for simplicity here, 2}

Situation Assessment
“what’s going on”

→

Decision and Action
“what do we do?”

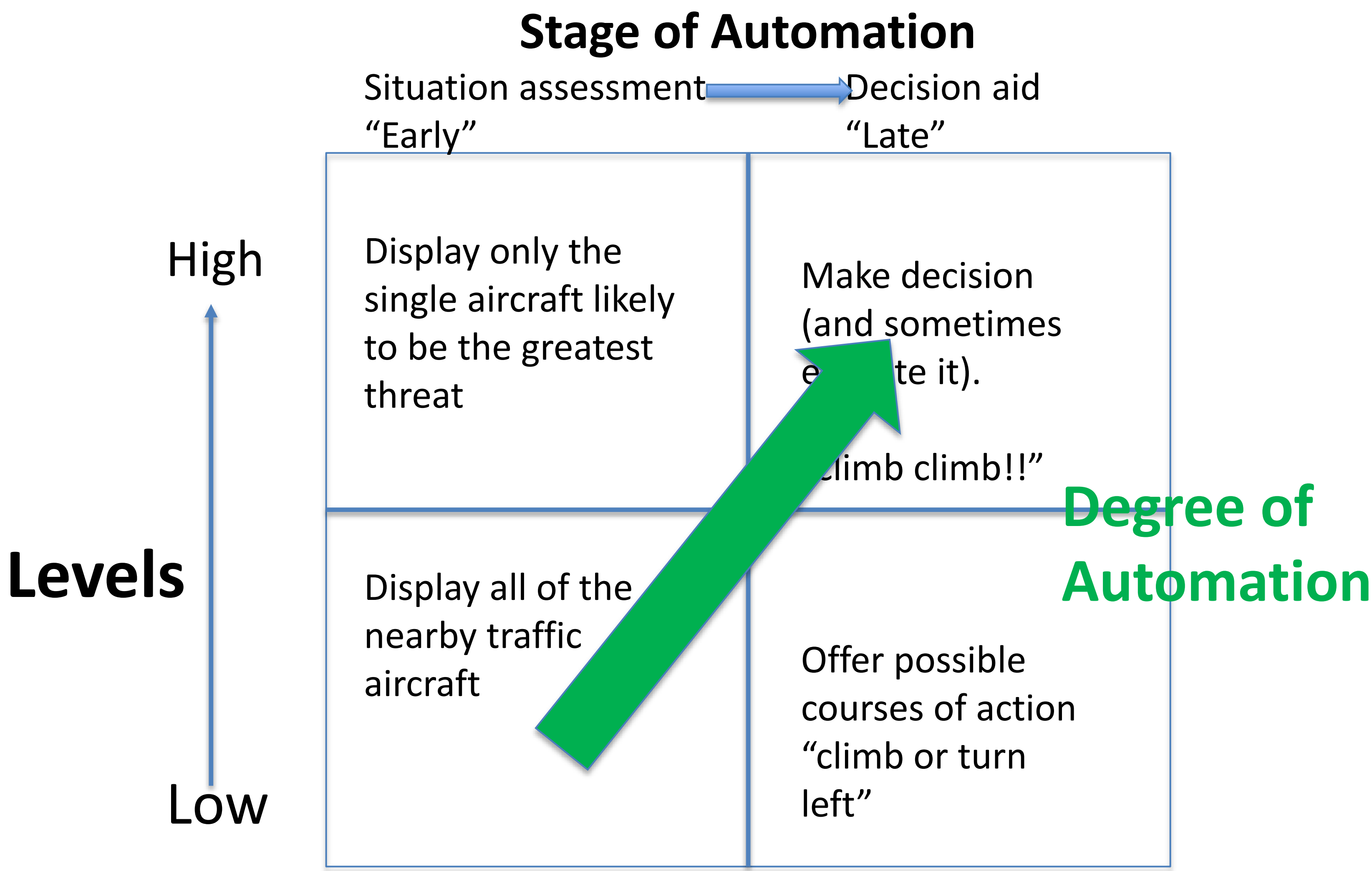
Levels: At each stage, designers can vary The **level of automation**: How “aggressively”/ autonomously does automation serve at each stage: how much human cognitive and motor “work” does automation replace.

Stage of Automation

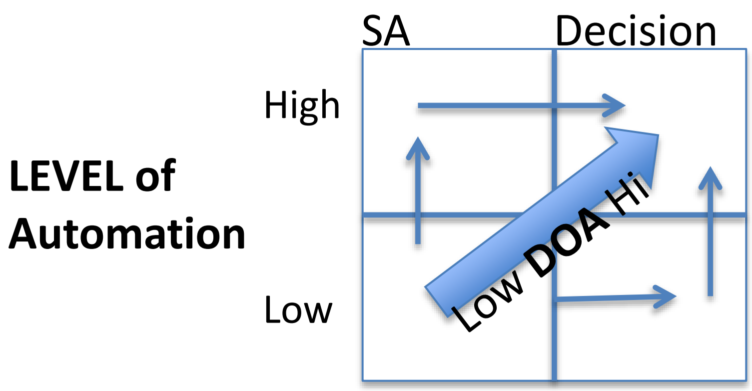
Situation assessment → Decision aid

Levels	“Early”	“Late”
High	Display only the single aircraft likely to be the greatest threat	Make decision (and sometimes execute it). “climb climb!!”
Low	Display all of the nearby traffic aircraft	Offer possible courses of action “climb or turn left”

An aircraft collision avoidance aid.
TCAS or a smart CDTI


Stage of Automation
Situation assessment
“Early”
Decision aid
“Late”
High
Levels
Low
Display only the single aircraft likely to be the greatest threat
Make decision (and sometimes e te it).
limb climb!!”
Display all of the nearby traffic aircraft
Offer possible courses of action “climb or turn left”
Degree of Automation

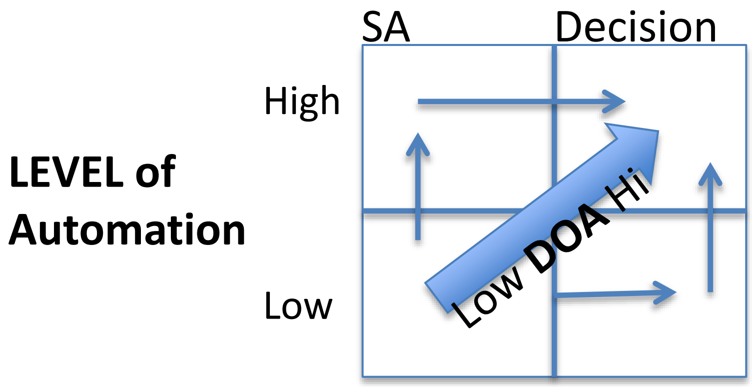
4. Degree of automation: Now the LUMBERJACK

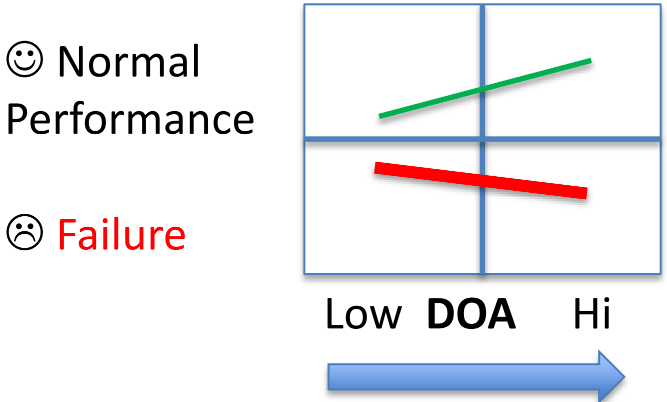


The Lumberjack analogy:
The **higher the DOA** (later stage&/or higher level), the higher the tree (better help when automation correct) and the

harder the fall (greater hurt when it fails, **particularly on the 1st failure**).

4. Degree of automation: Now the LUMBERJACK





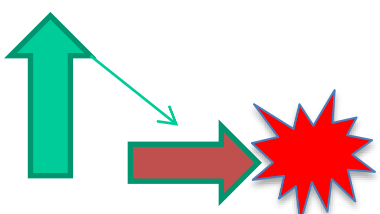
The Lumberjack analogy:
The **higher the DOA** (later stage&/or higher level), the higher the tree (better help when automation correct) and the:

harder the fall (greater hurt when it fails, **particularly on the 1st failure**).
Takeover response time in self-driving cars.

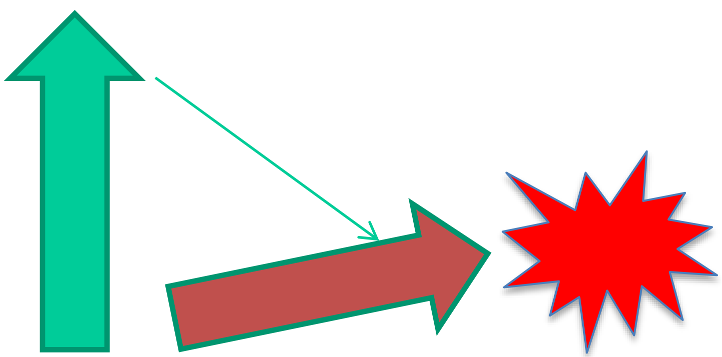
The lumberjack analogy of automation dependence: The higher the tree, the harder it falls.

•.

- The Higher the Degree of Automation (DOA), the greater benefit when it **works**: the greater cost **when** it **“fails**”.



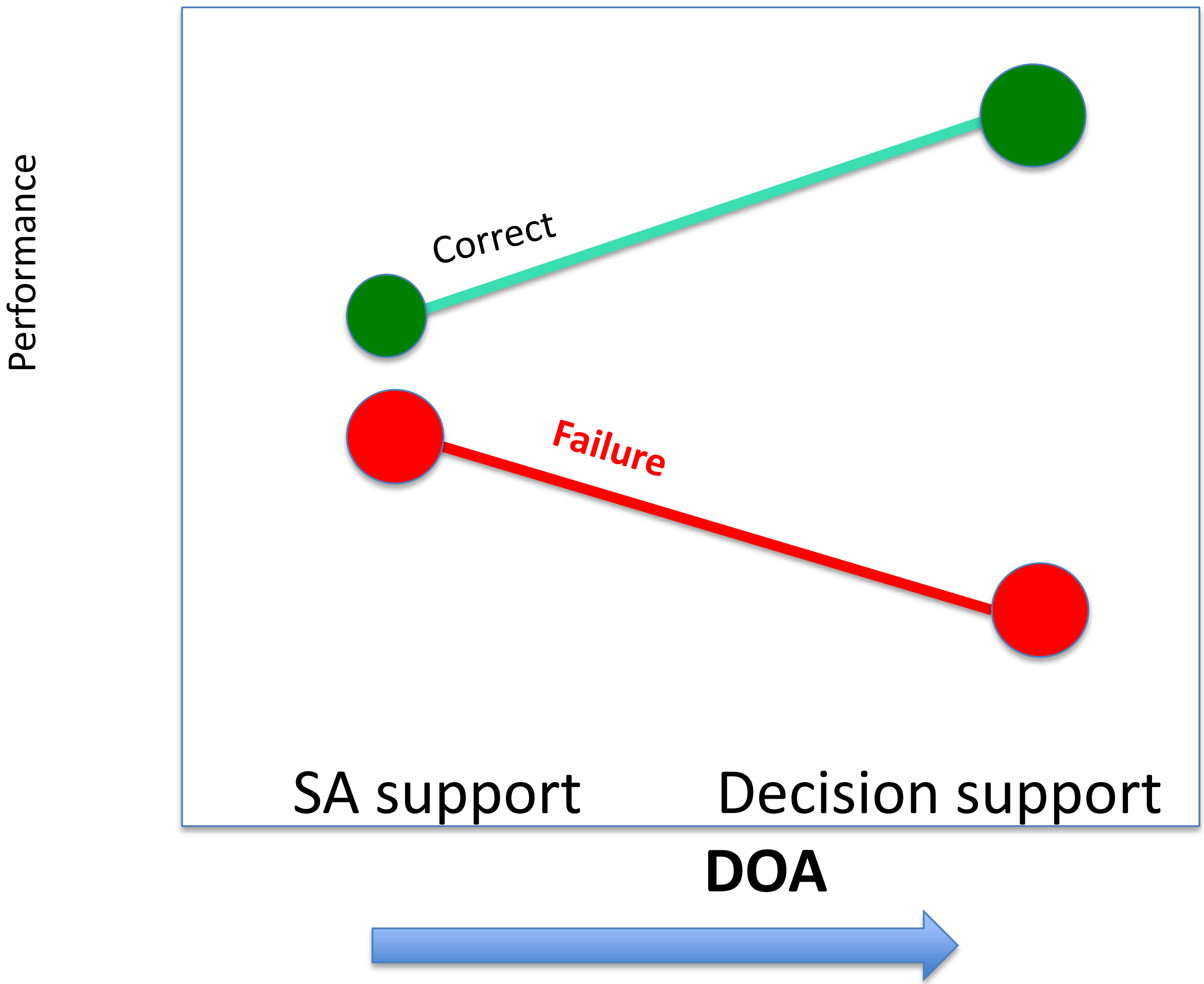
Low DOA SA support



High DOA: Decision/action support

The Generic Lumberjack: 4 data points

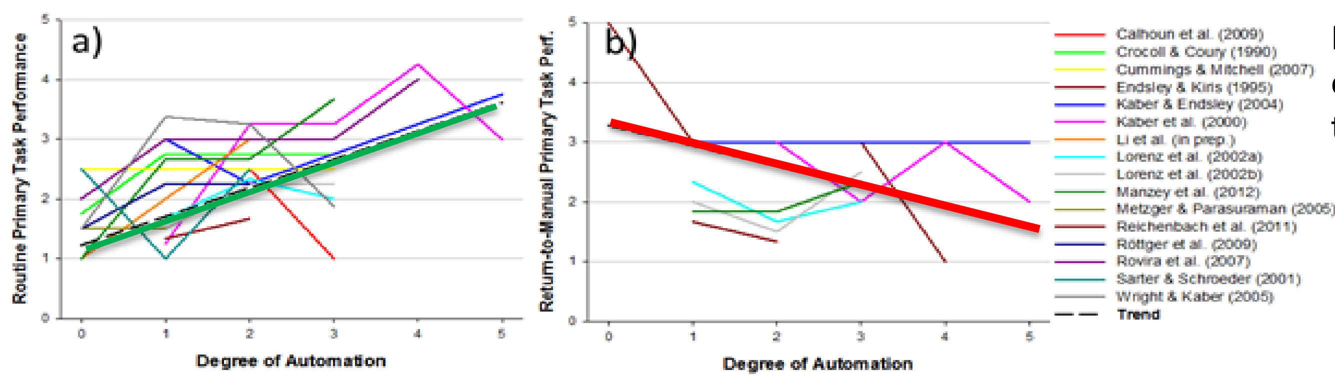
Based on a meta-analysis for supporting data



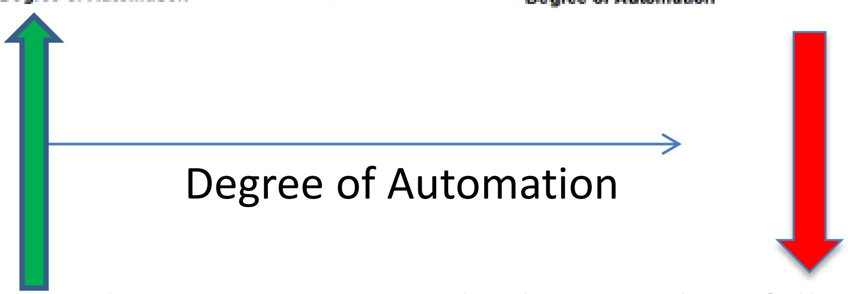
What are the Lumberjack Data?

Onnasch, Wickens, Manzey & Li, (2014)

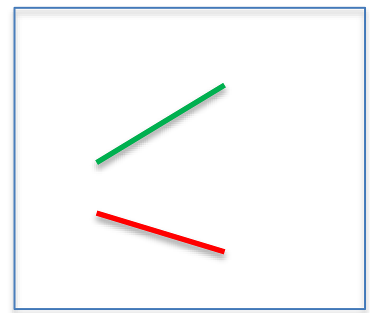
Nominal Performance | Failure Performance



Results of a meta-analysis of studies that have varied the degree of automation



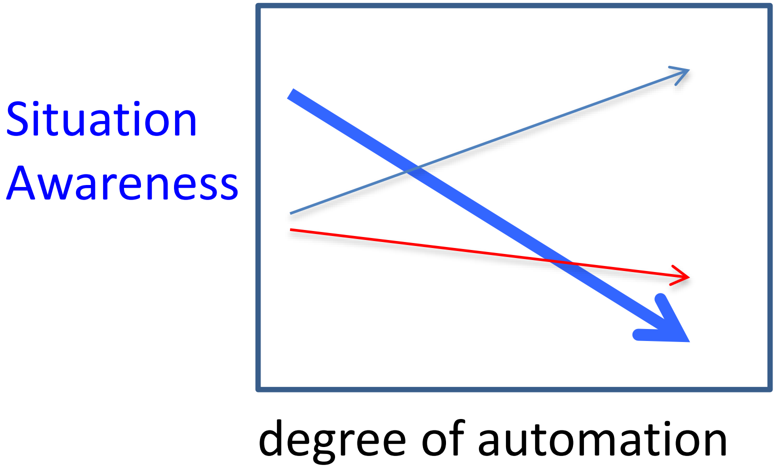
The higher they stand | The harder they fall



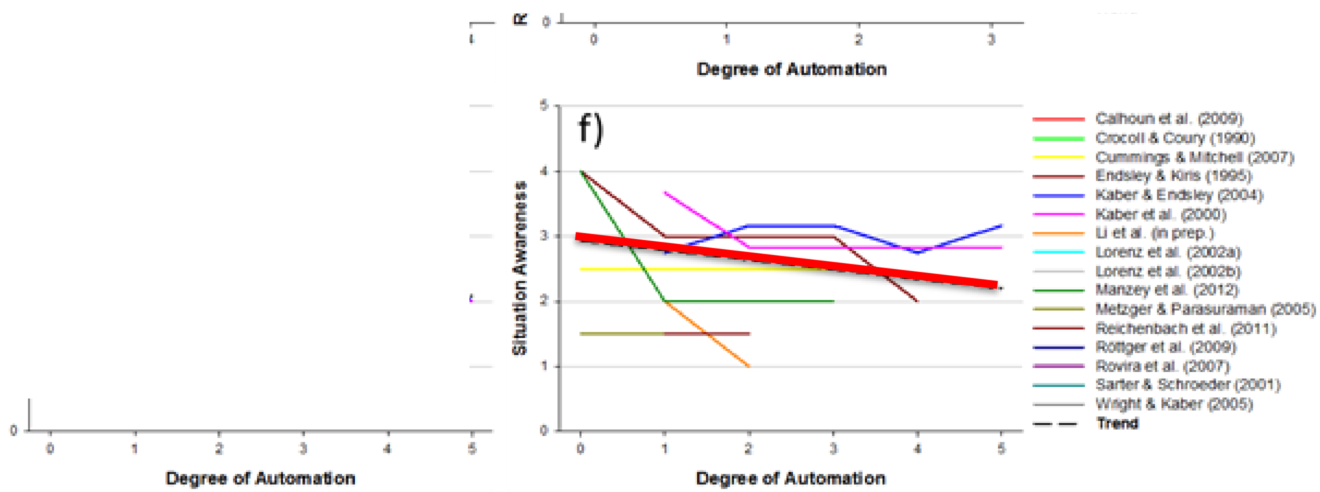
What underlies this increased cost, at higher degree of automation, when automation "fails"?

The Loss of Situation Awareness

Loss of Situation Awareness (LSA) at higher degree of automation



But the LSA effect was weak (although significant), and the correlation between the 2 variables does not necessarily imply causality.



Situation awareness is lost at:

Endsley's Level 1: At higher degree of automation, people **look less** at the process that automation is supporting.

Evaluated in our first experiment.

Endsley's level 2: At higher DOA, people fail to **understand** what automation is doing.

5. Transparency supports Situation Awareness. And can therefore buffer the first failure response.

Transparency is **a graphic or text-based explanation of what automation is doing and why.**

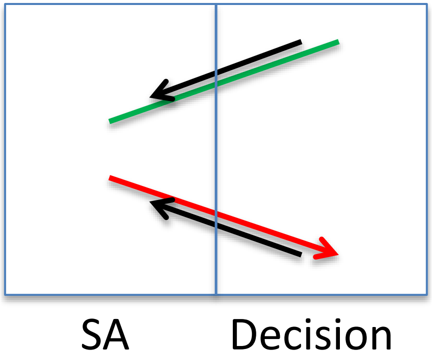
Also **a clear, intuitive depiction of the input to automation:** that is, the raw data that automation is processing.

Transparency can be offered either as SA support while automation is engaged, or as training, prior to automation use.

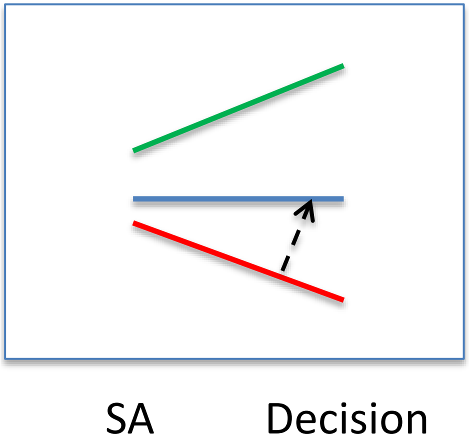
We will evaluate in the second experiment.

5. Addressing the tradeoff through SA support: Is there a "free lunch"? Fixing the costs for failure without reducing the **benefits of success.** The higher tree does **not** fall harder. Recall the two causes of loss of Level 2 SA: loss of the generation effect: **"if you don't act, you can't remember"** Lack of transparency: "**can't see what automation is doing"**

Solution 1. If lack of generation was responsible, then simply lower the DOA so human generates active responses. Less of a cost but also less of a benefit. No free lunch.



Solution 2: **Transparency** to support SA. If LSA is responsible, then a good "transparent display" should restore failure performance without reducing performance when automation is correct. A "free lunch" ☺

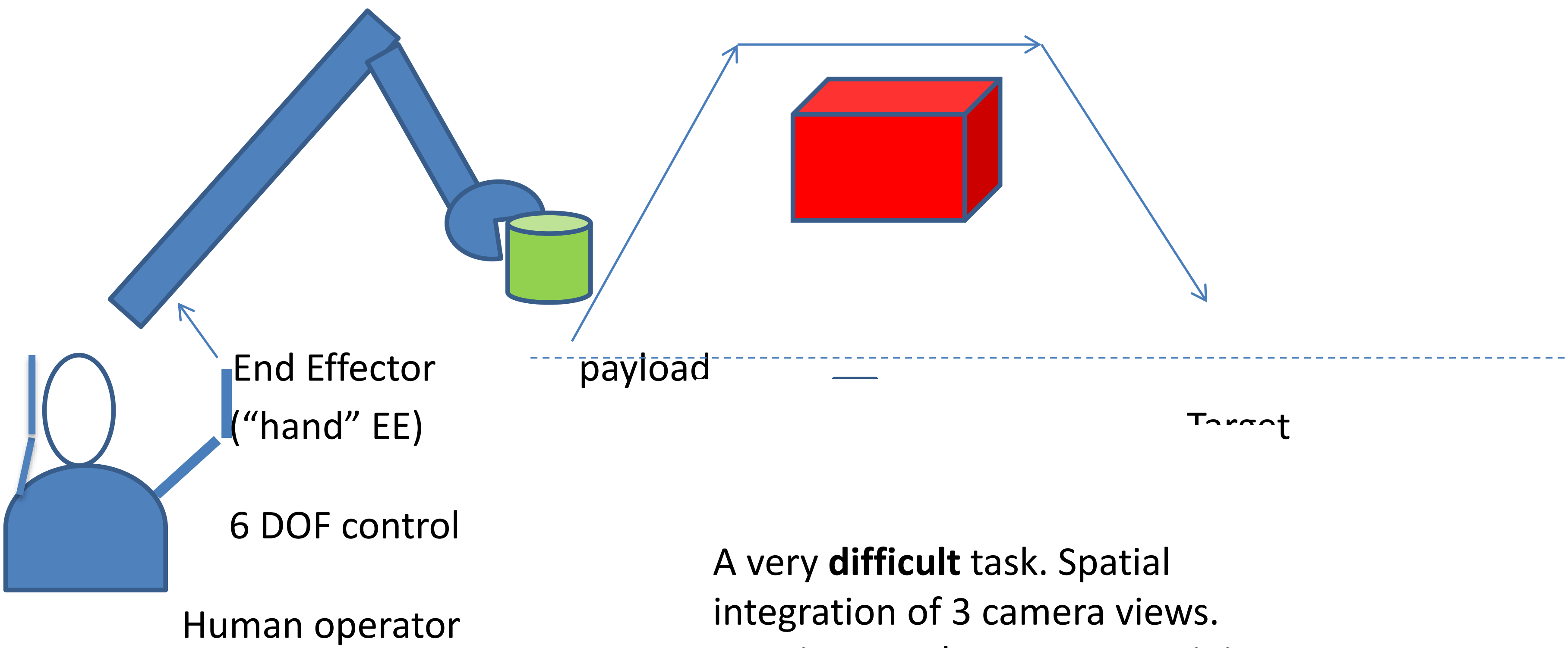


Experiment 1: showing the tradeoff: Robotics at 2 DOAs

Robotic Arm controller Task (NASA Johnson Space)

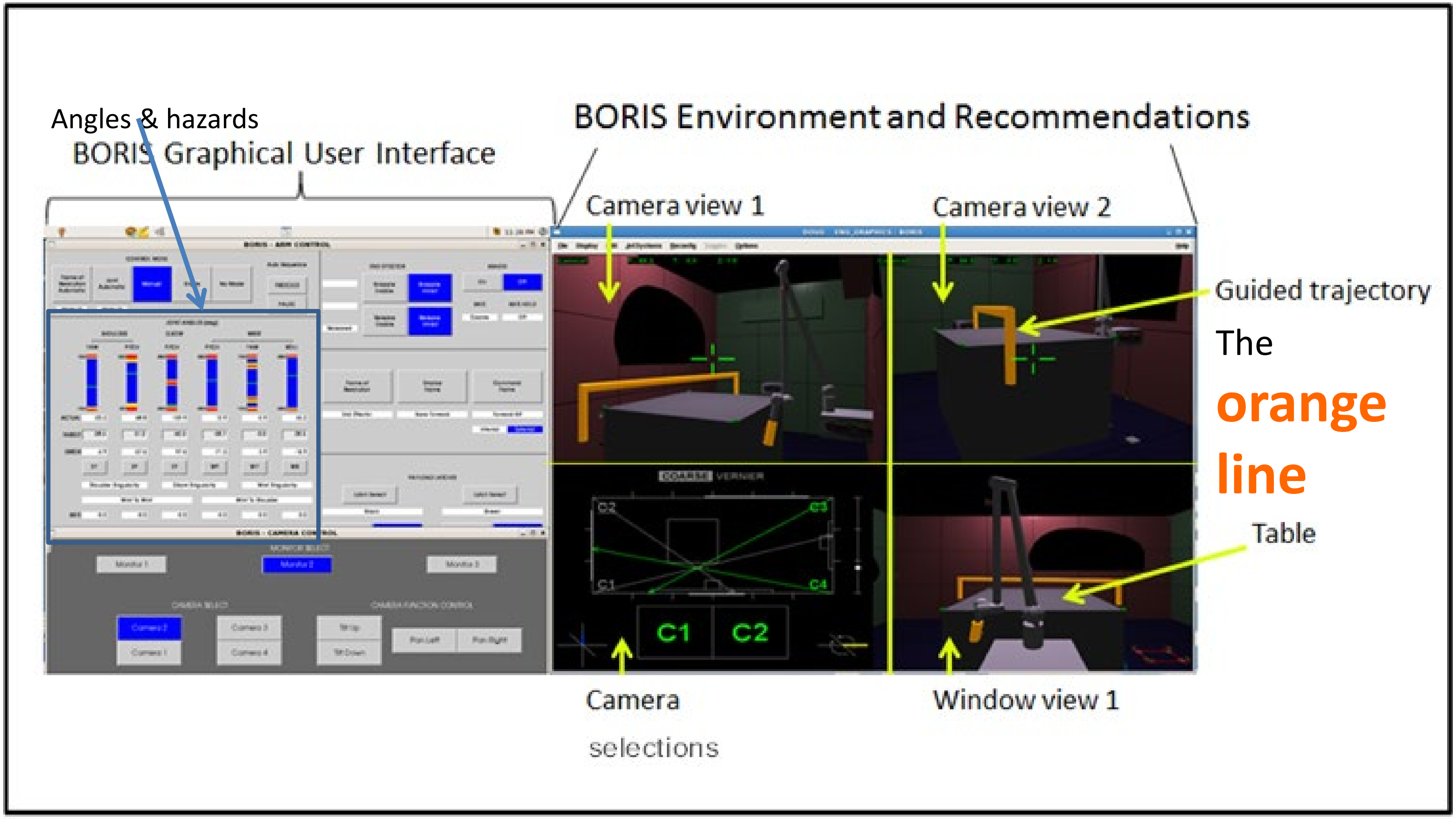
Li Wickens Sarter & Sebok,HFJ 2014; Wickens Sebok et al, HFJ 2015

Testing the Lumberjack Analogy



A very **difficult** task. Spatial integration of 3 camera views. Requires much astronaut training. Our simplified (but realistic) Michigan simulation.

The NASA BORIS Robotic Arm Simulator



Experiment 1: Operating the robotic arm manually and with automation at 2 DOAs

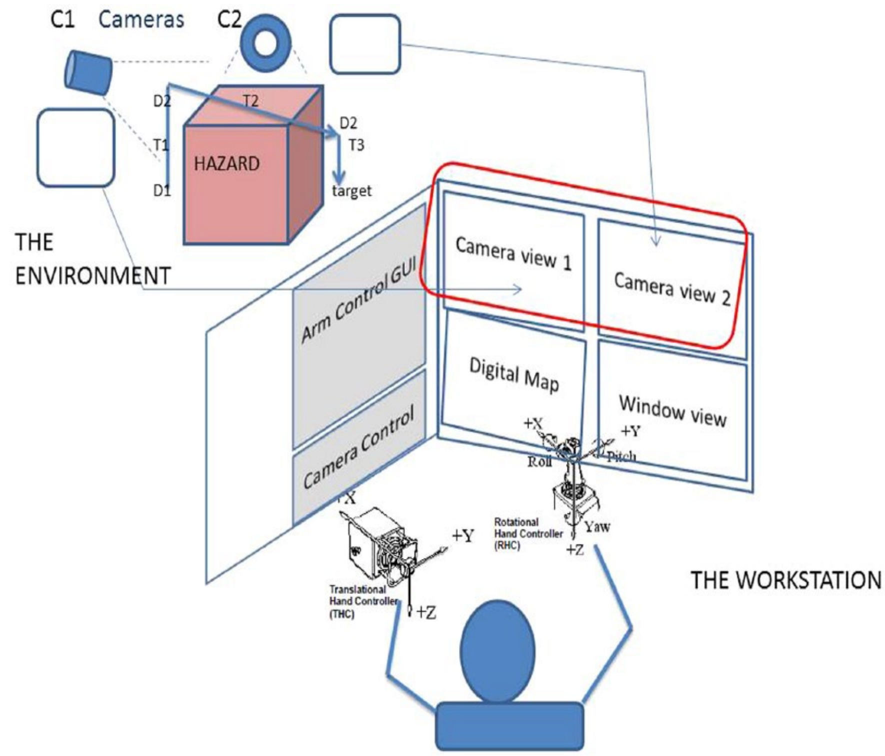
- Manual Control

- Low DOA: Automation **SA support**:

The "**orange line**" showing
target deviation.

- High DOA: Automation

decision and action support:
A full **autopilot** following the **orange** line.



The Prediction:

Introduction of automation and higher DOA supports progressively better routine performance:
(Reduced tracking error and mission time)

Inducing and measuring automation-induced "Complacency" in the Michigan Experiment

The Michigan subjects

Several hours practice with perfect automation & correct XYZ coordinates.

7 experimental Trials

Automation helps: ☺

↓ workload, ↑ precision

Experiment 1: Operating the robotic arm manually and with automation at 2 DOAs

- Manual Control

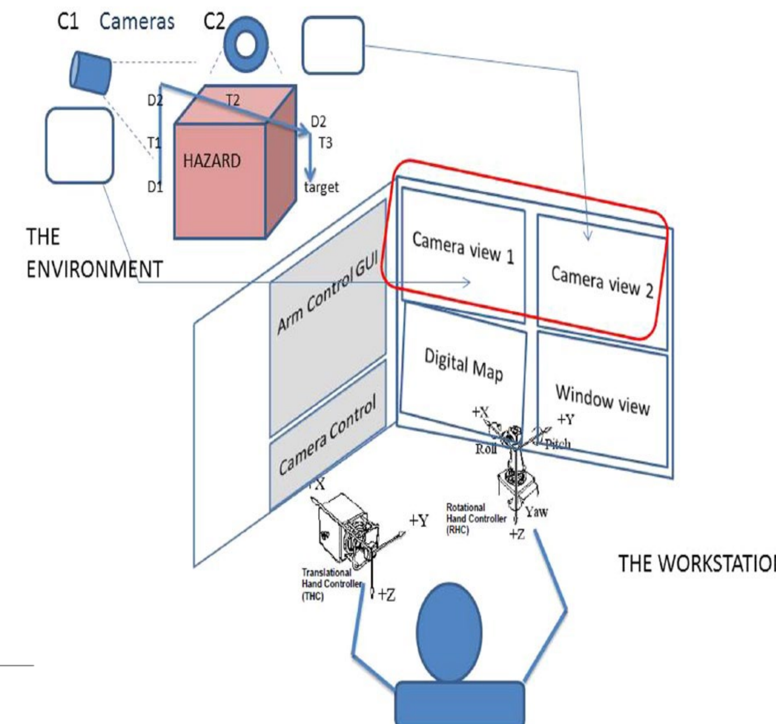
- DOA1: Automation **SA support**:

The “**orange line**” showing target deviation.

- DOA2: Automation **decision and action support**:

A full **autopilot** following the **orange** line.

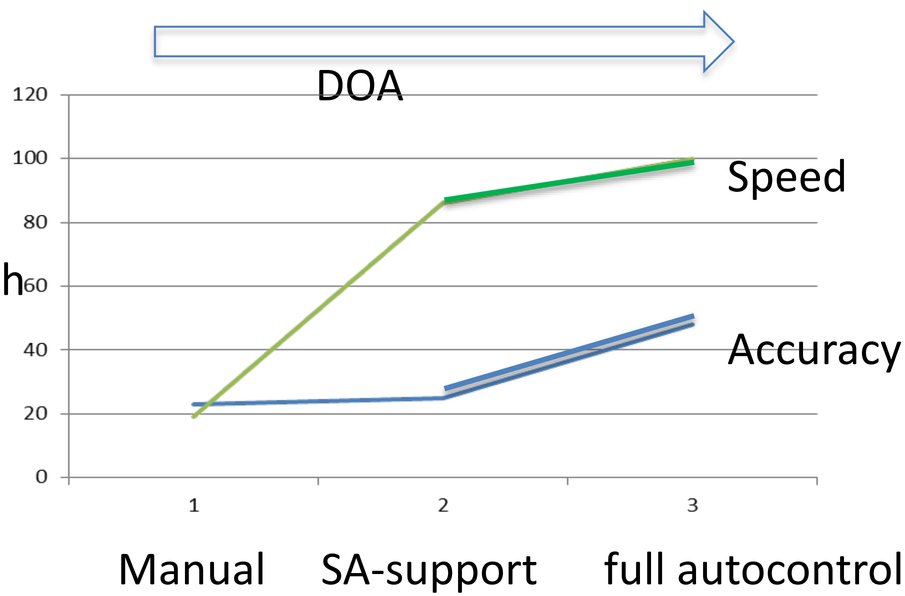
Automation and higher DOA
Supported progressively better routine performance
(tracking error, Mission time)

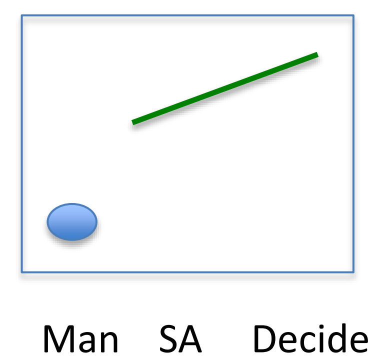


☺

Performance with
Normal automation

☹





Inducing and measuring automation-induced "Complacency" in the Michigan Experiment

The Michigan subjects

Several hours practice with perfect automation & correct XYZ coordinates.

7 experimental Trials

Automation helps: ☺

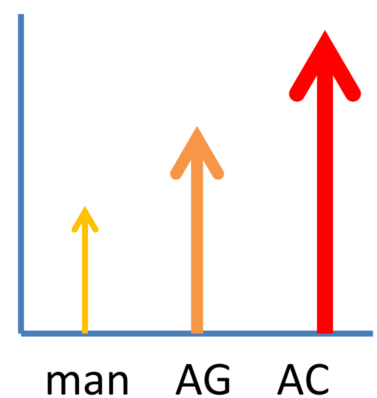
↓ workload, ↑ precision

Trial 7: **Failure**. Wrong XYZ coordinates

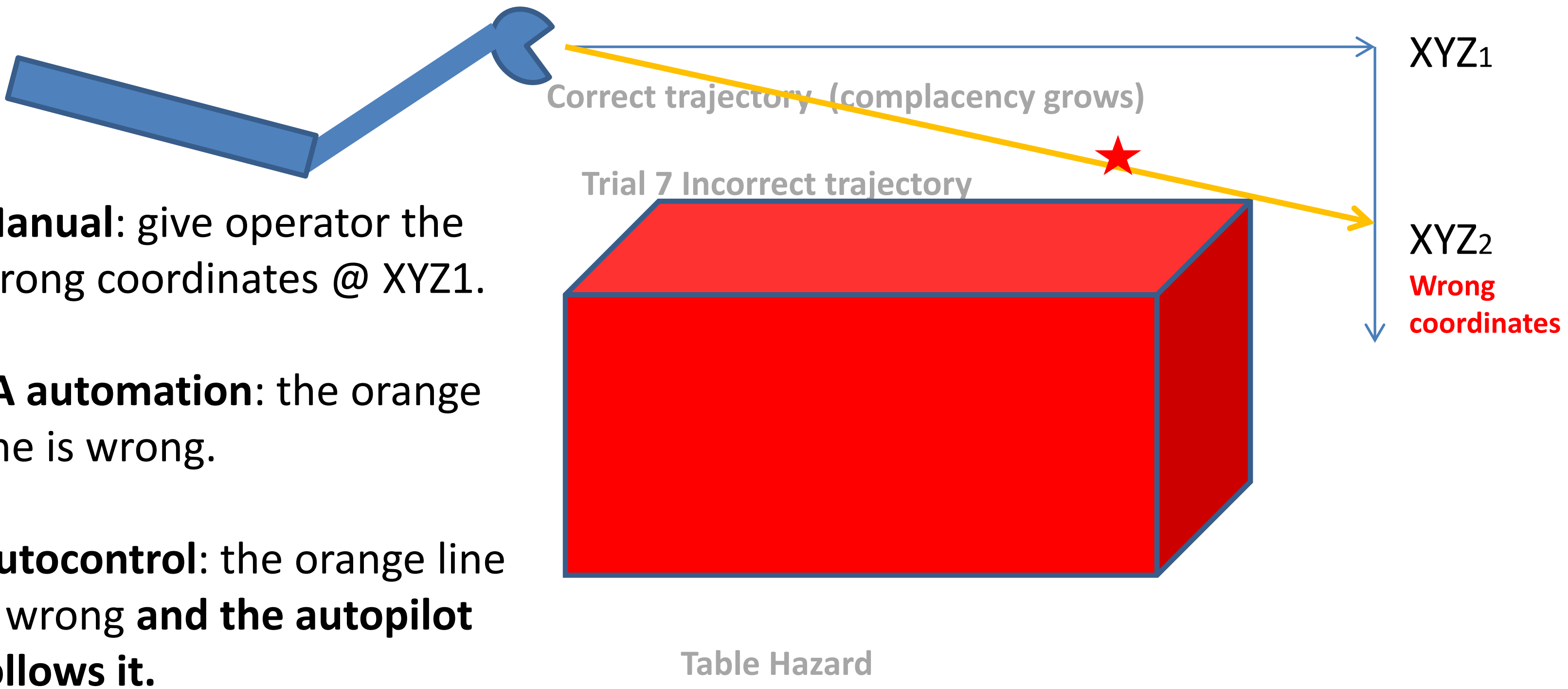
Complacency grows

Our "Lumberjack" prediction

miss rate to Failure "complacency" measured



The Trial 7 Automation Failure: After 6 scenarios and training with perfect automation, how does automation "fail"?



Manual: give operator the wrong coordinates @ XYZ1.

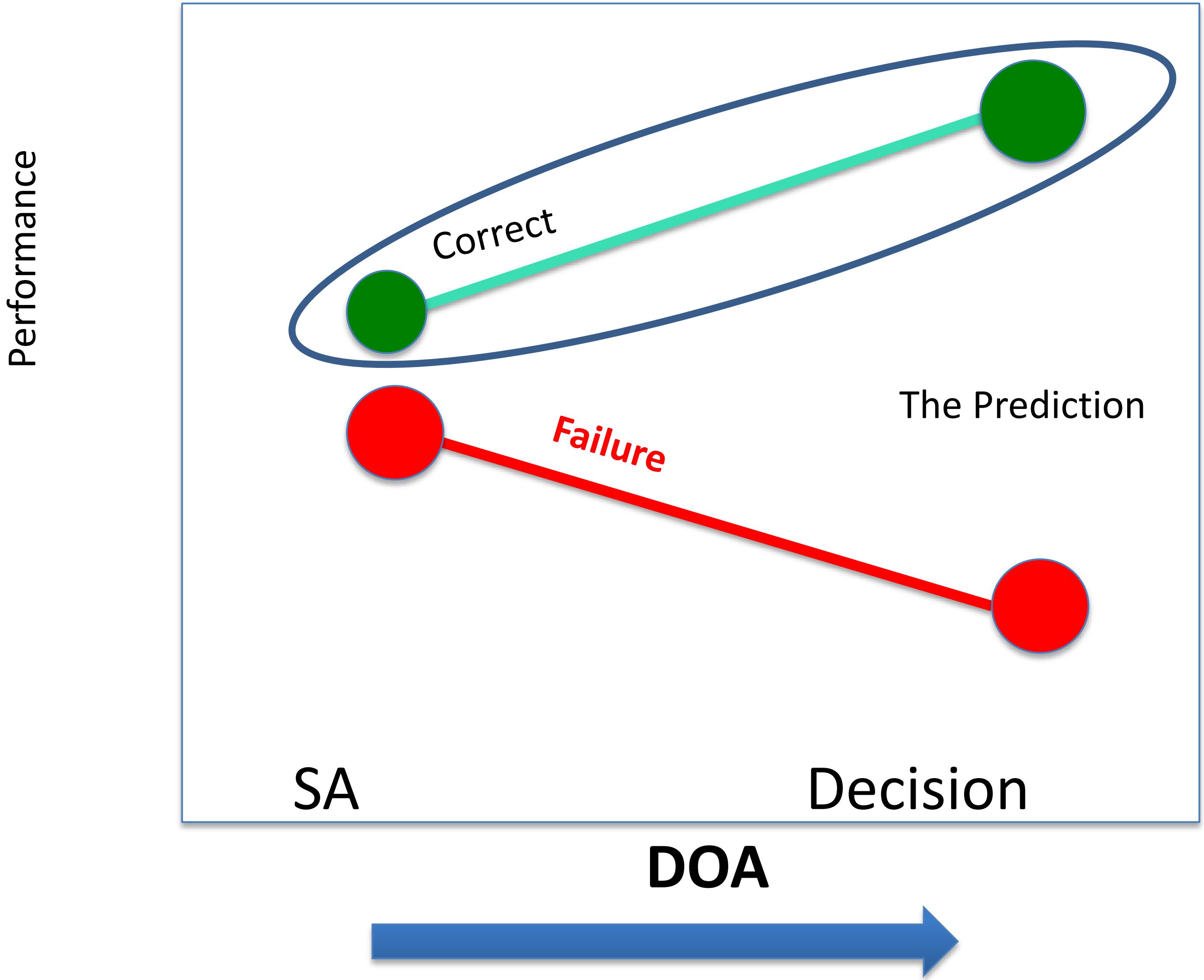
SA automation: the orange line is wrong.

Autocontrol: the orange line is wrong **and the autopilot follows it.**

NOTE that on Trial 7, the hazard proximity warning also failed

The Generic Lumberjack: 4 data points

Based on a meta-analysis

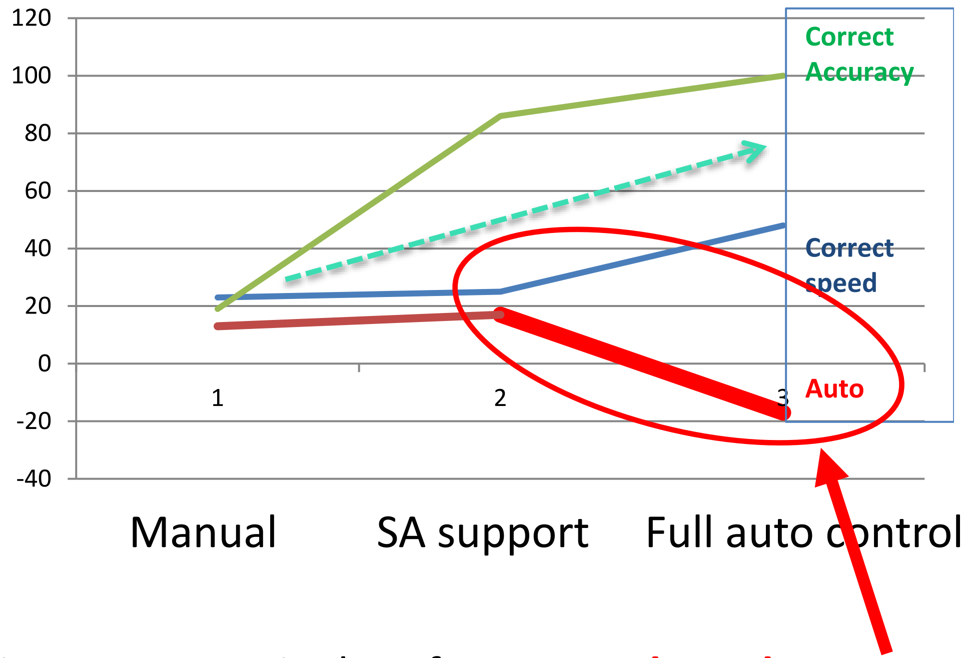


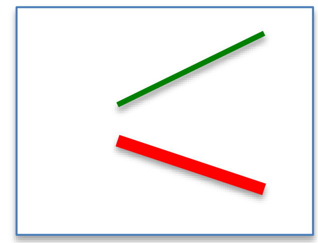
Automation Failure performance:

The guidance line was wrong (incorrect coordinate specification). It guided the trajectory into the table.

Measures: Proportion of **table hits**; hazard penetrations (no-fly zone)

Failure response performance





Conclusion:

* Higher DOA improves nominal performance, **degrades automation failure performance**.

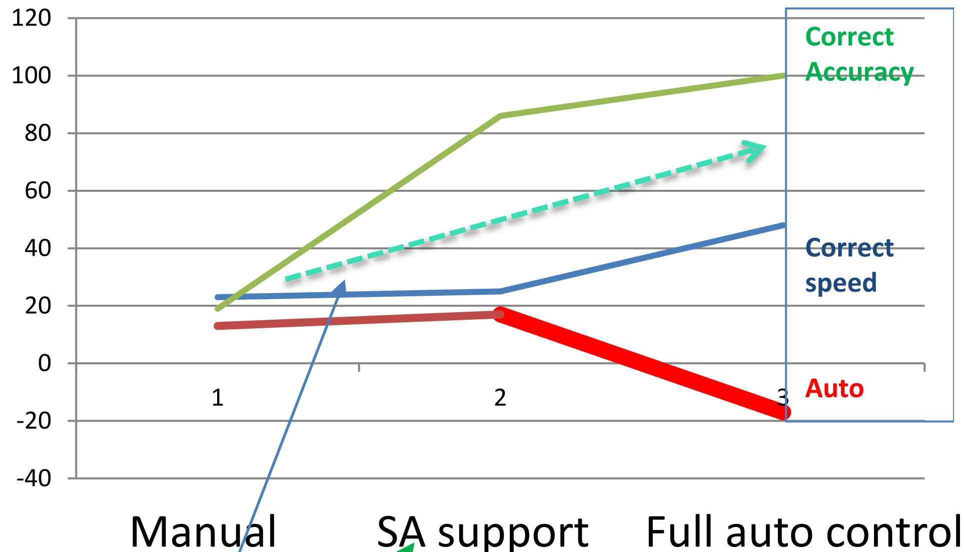
- Lower DOA offers benefits relative to manual performance with minimal “failure” cost. This supports the value of stage 2 SA automation!
- **Computational modeling** of the complacency effect provides accurate prediction of the AMOUNT of cost in failure response, attributable to OOTLUF monitoring failure.

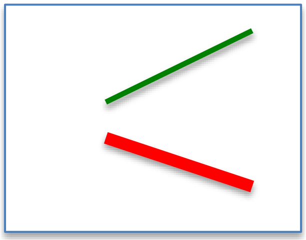
Automation Failure performance:

The guidance line was wrong (incorrect coordinate specification). It guided the trajectory into the table.

Measures: Proportion of **table hits**; hazard penetrations (no-fly zone)

Failure response performance





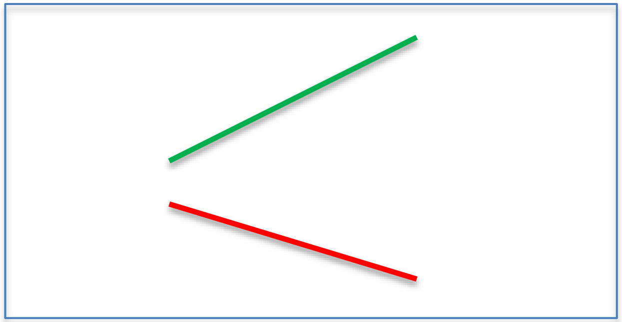
Conclusion:

* Higher DOA improves nominal performance, degrades automation failure performance.

- Lower DOA offers benefits relative to manual performance with minimal "failure" cost. This supports the value of stage 2 SA automation!
- **Computational modeling** of the complacency effect provides accurate prediction of the AMOUNT of cost in failure response, attributable to OOTLUF monitoring failure.

Conclusion: Experiment 1.

Lumberjack support. Later stage automation (autocontrol) helped performance more than earlier stage automation (SA-support) when **correct**. Hurt performance more, when **automation failed** (bad data on the trajectory).



We confirmed the Hypothesis for the 4 data points.

We modeled the complacency effect on automation failure

BUT

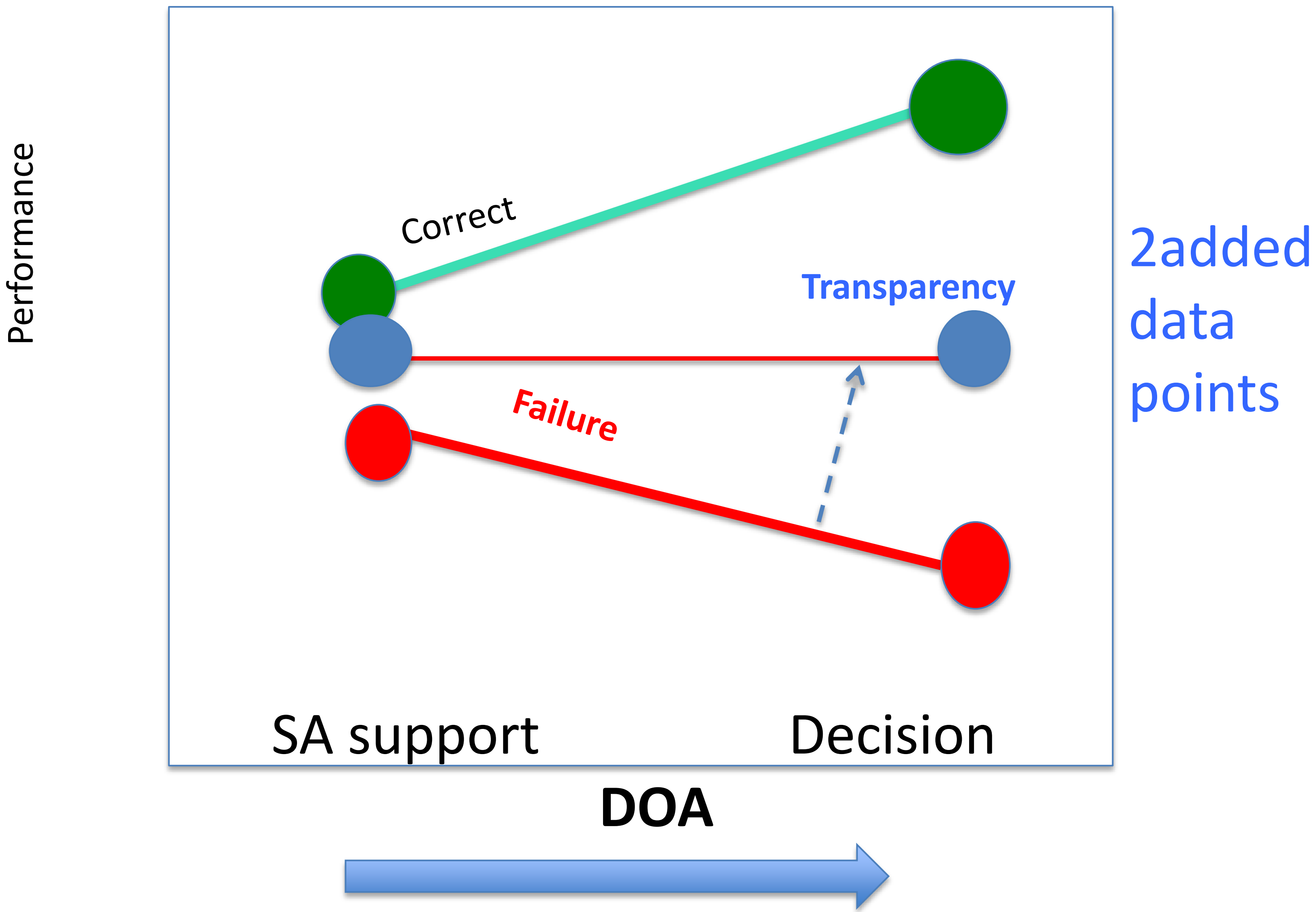
We did not explicitly provide **Transparency** to offset the late stage automation cost.

We hypothesize that compromised automation failure response results from operator **loss of situation awareness** at the higher DOA.

This loss can be mitigated by **transparency** of the display that **depicts the raw data which automation is processing.**

The goal of Experiment 2, to test these hypotheses in the context of imperfect **decision support automation** in air traffic management. Late DOA correct and failure performance

The Generic Lumberjack: 4 points



Experiment 2: Fitri Trapsilawati Dissertation,.

Nanyang Technical Institute: Singapore

* ATC: 2 responsibilities: **Productivity** (deliver planes on time) and **SAFETY** (avoid collisions.)

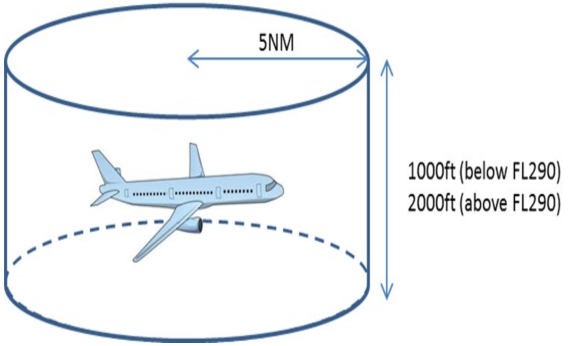
- SAFETY: Keep any plane out of another's "protected zone": a "Loss of separation" (**LOS**)
- Two stages of conflict avoidance task:

Situation Assessment →Decision

Conflict detection → Conflict Resolution.

Will there be LOS? → Maneuver the plane

Both stages are supported by automation tools



DOA

Lower: Situation Assessement **Higher: Decision Aid**

Conflict Detection Aids (CDA).

70-80% reliable.

Still assists the controller relative to manual unaided detection.

The **75% reliability** threshold, above which automation helps relative to manual.

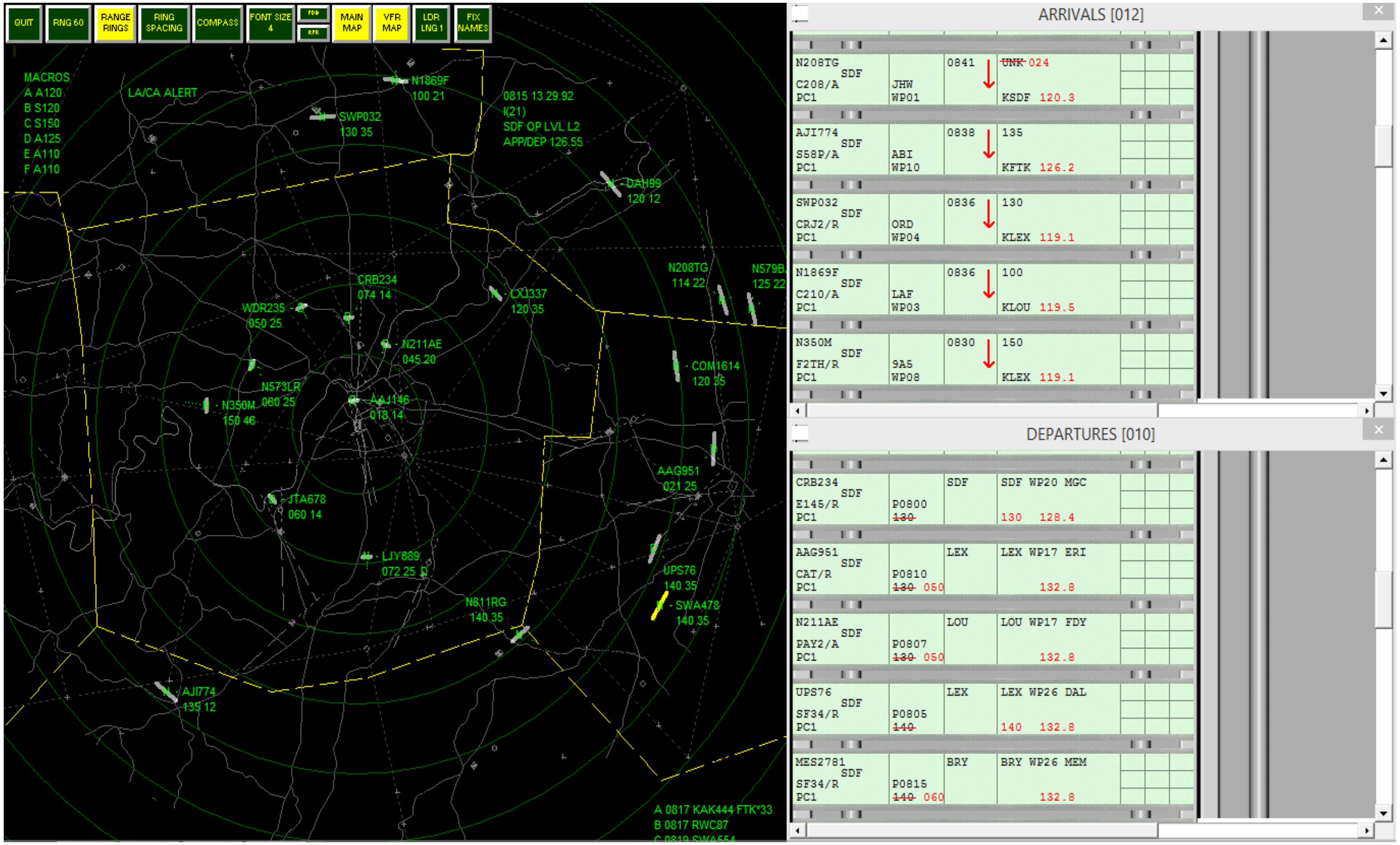
Conflict Resolution Aids (CRA)

Not yet implemented in ATC.

Perfect CRA has been investigated.

Imperfect CRA has not yet been investigated.

The 2D plan view radar Display for ATC. Low vertical SA



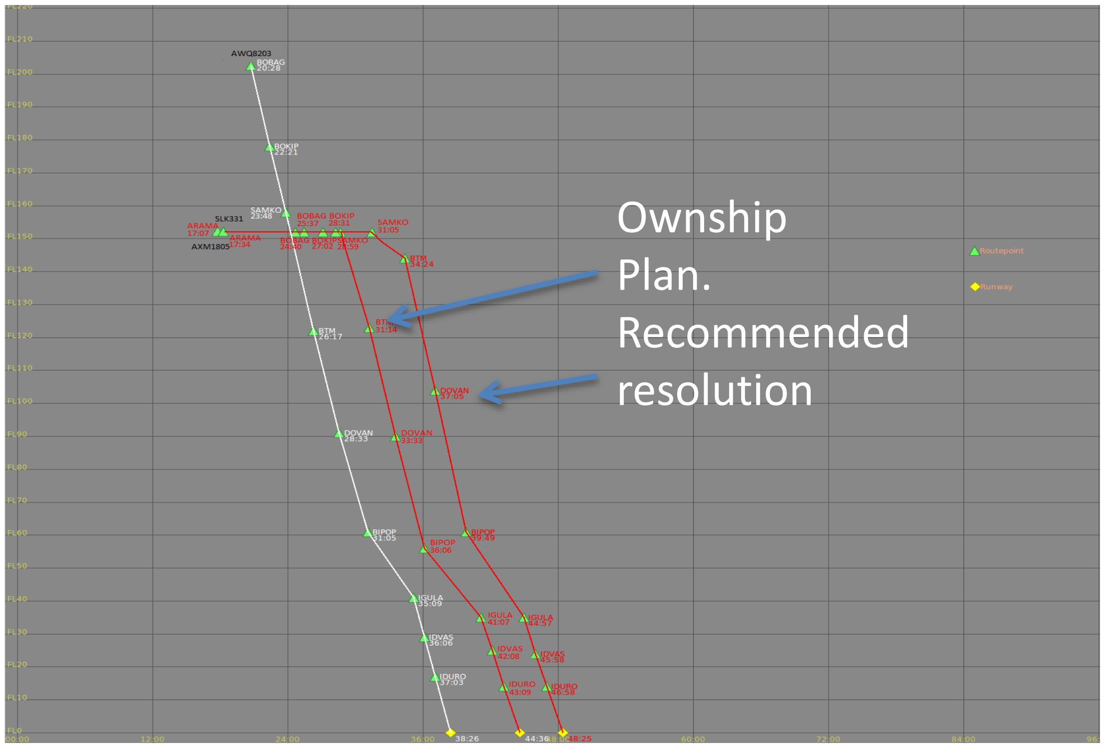
The Human Factors Challenge: Conflicts sometimes occur with vertical flight (ascents and descents). The imperfect CRA automation resolves vertical conflicts (sometimes with vertical maneuvers). **BUT THE CONTROLLER HAS NO VERTICAL DISPLAY. No Transparency: LIMITED VERTICAL SITUATION AWARENESS. ☹**

Transparency can be created by the

Vertical Situation Display

The Vertical Situation Display (VSD) designed to support controller vertical situation awareness of a conflict, and thus offset the costs of imperfection

Altitude

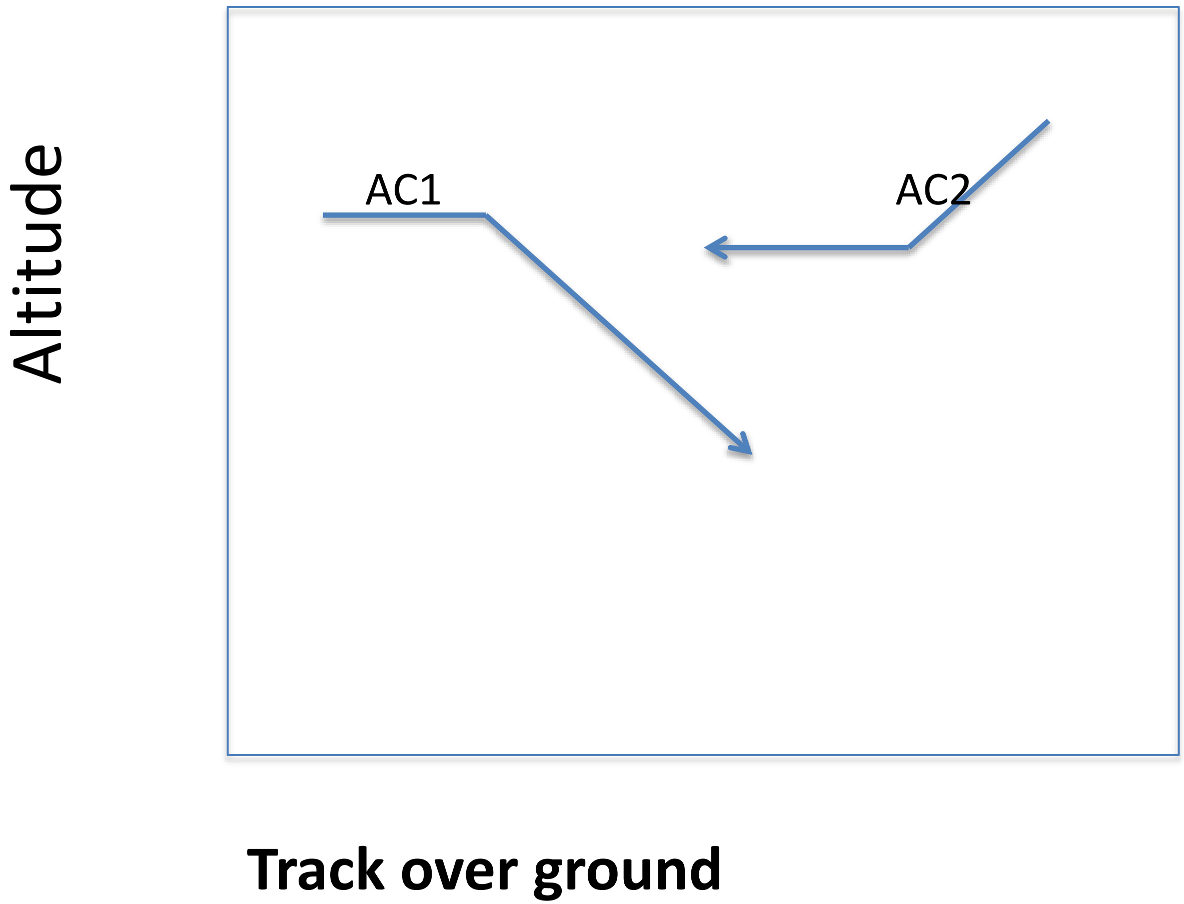


Vertical situation Displays (VSD) have rarely been examined in ATC, and **never before** in conjunction with the **Conflict Resolution Aid**.

These 2 forms of ATC automation: VSD and Decision aid examined in parallel

Simplified: Transparency created by the Vertical Situation Display

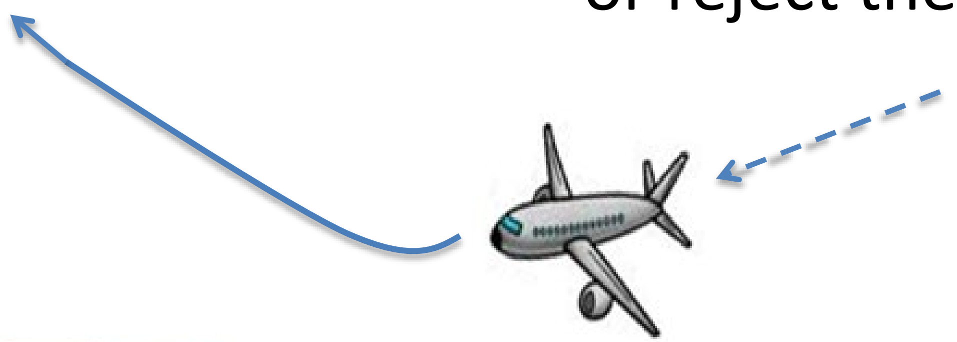
The Vertical Situation Display (VSD) designed to support controller vertical situation awareness of a conflict and thus offset the costs of imperfection

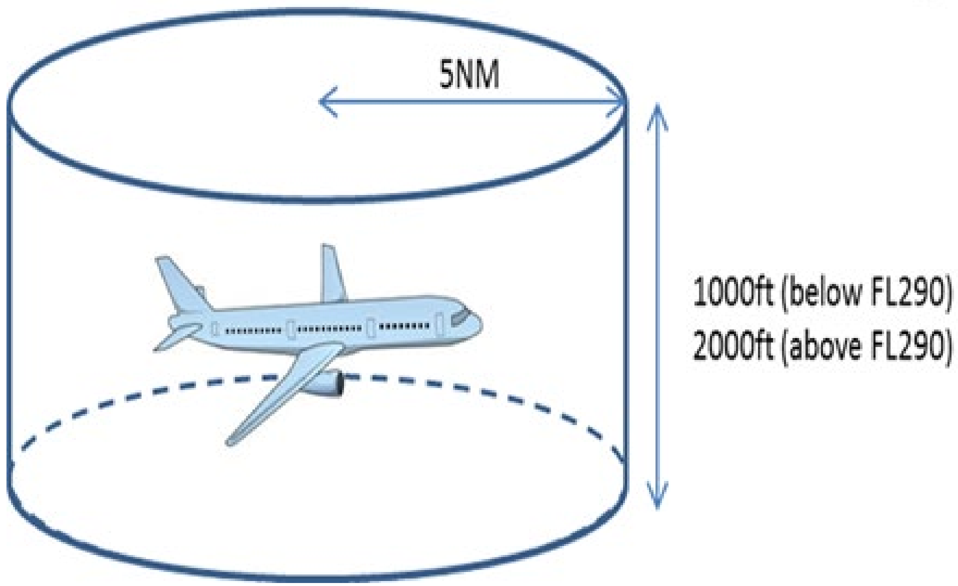


Profile View

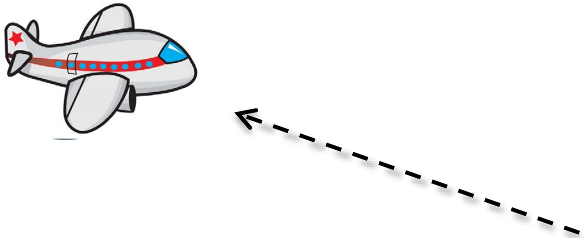
Such transparency could offset the costs of automation error when automation(CRA) is imperfect.

The Automation stage 2 Decision Aid **CRA**: "Advise climb". Controller can accept or reject the advice

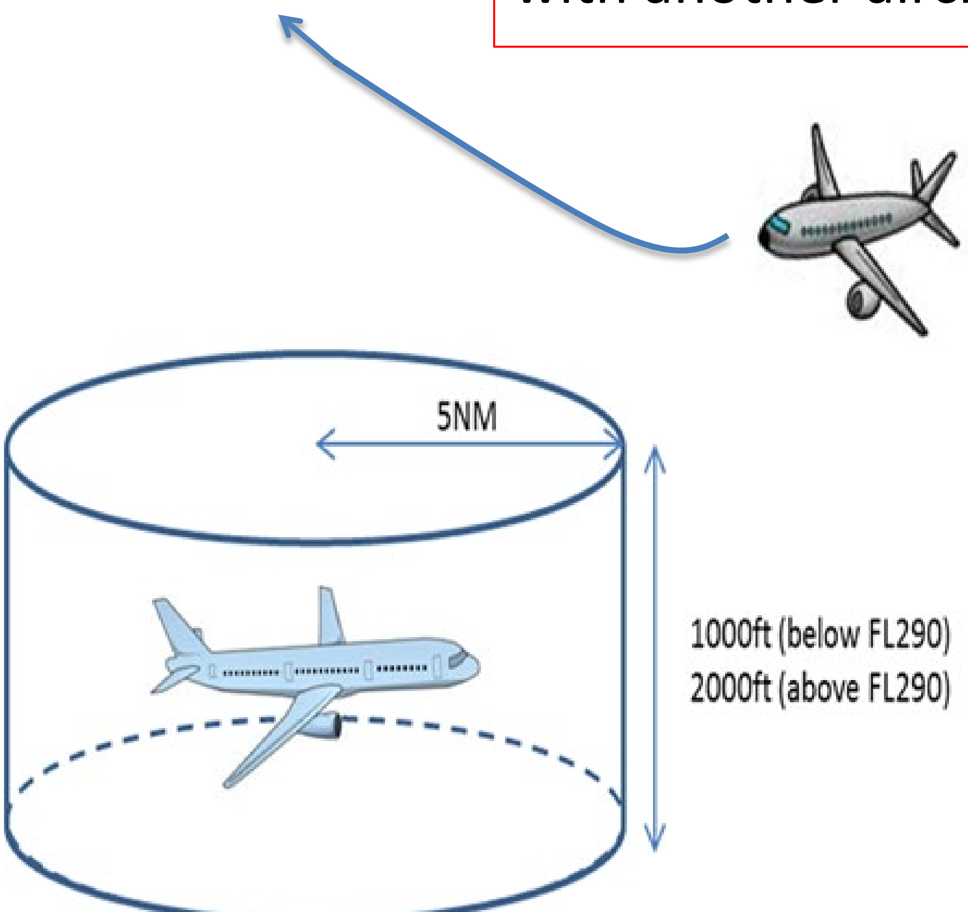




But the automation is imperfectly reliable. It **CAN fail. How?**



Imperfect CRA automation. The **automation failure:** the secondary conflict. “Advise climb”. This will produce LOS with another aircraft.



1 trial of 6:
83% Reliable

Experiment 2: combining results of 2 experiments by Fitri.

Trapsilyawati, F., Chun-Hsien, Wickens, C.D., & Qu, X (2021). The Integration of Conflict Resolution Automation and Vertical Situation Display for On-ground Air Traffic Control Operations. *Journal of Navigation*.

48 participants; All with prior ATC experience:

Performed several blocks of traffic management with perfectly reliable conflict avoidance aid. Then the aid **fails** (recommendation directs aircraft into secondary traffic). We measured response on the **FIRST FAILURE.**

Half participants had vertical situation display (VSD: transparency supporting SA)

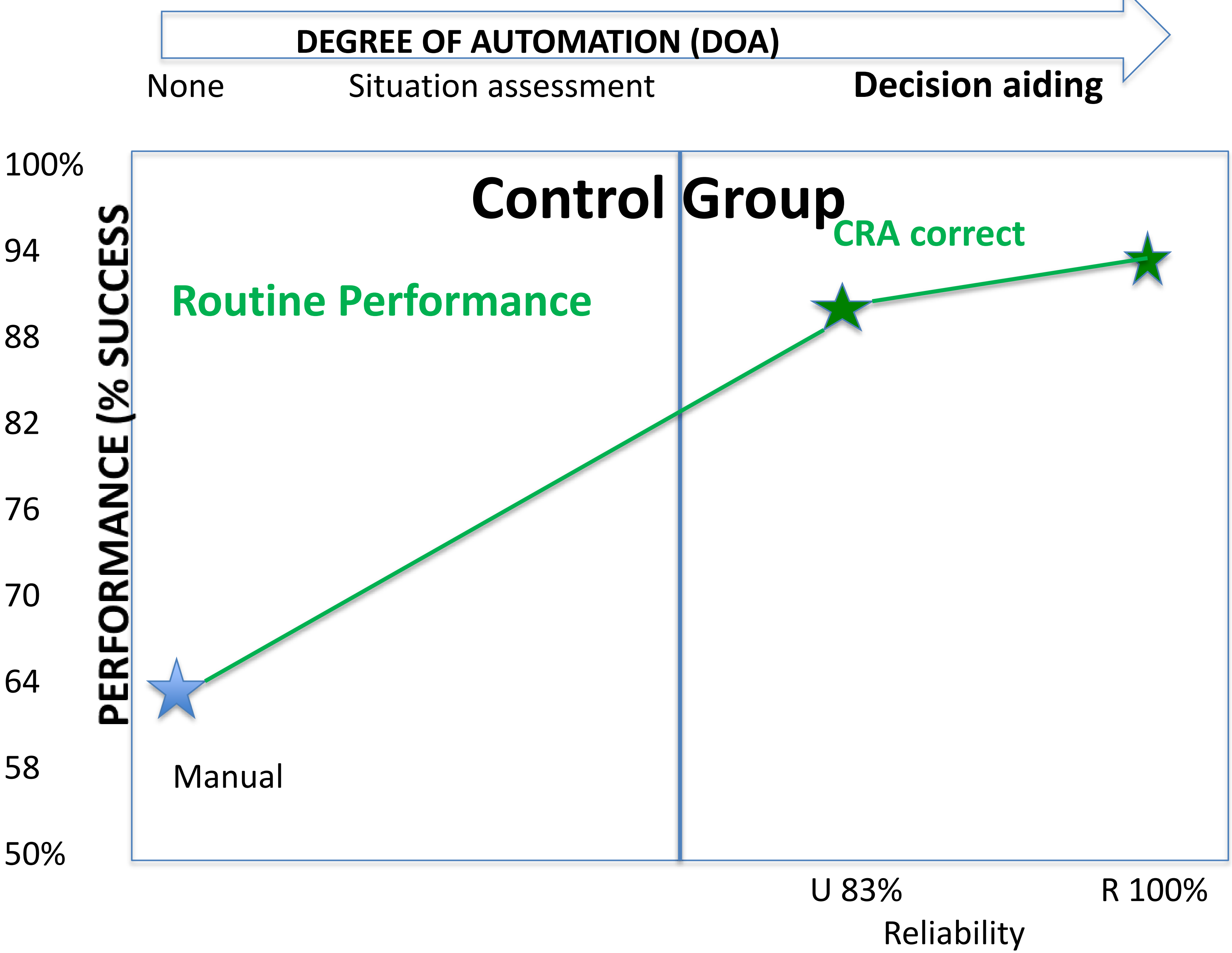
Half did not.

Primary Performance Measure:

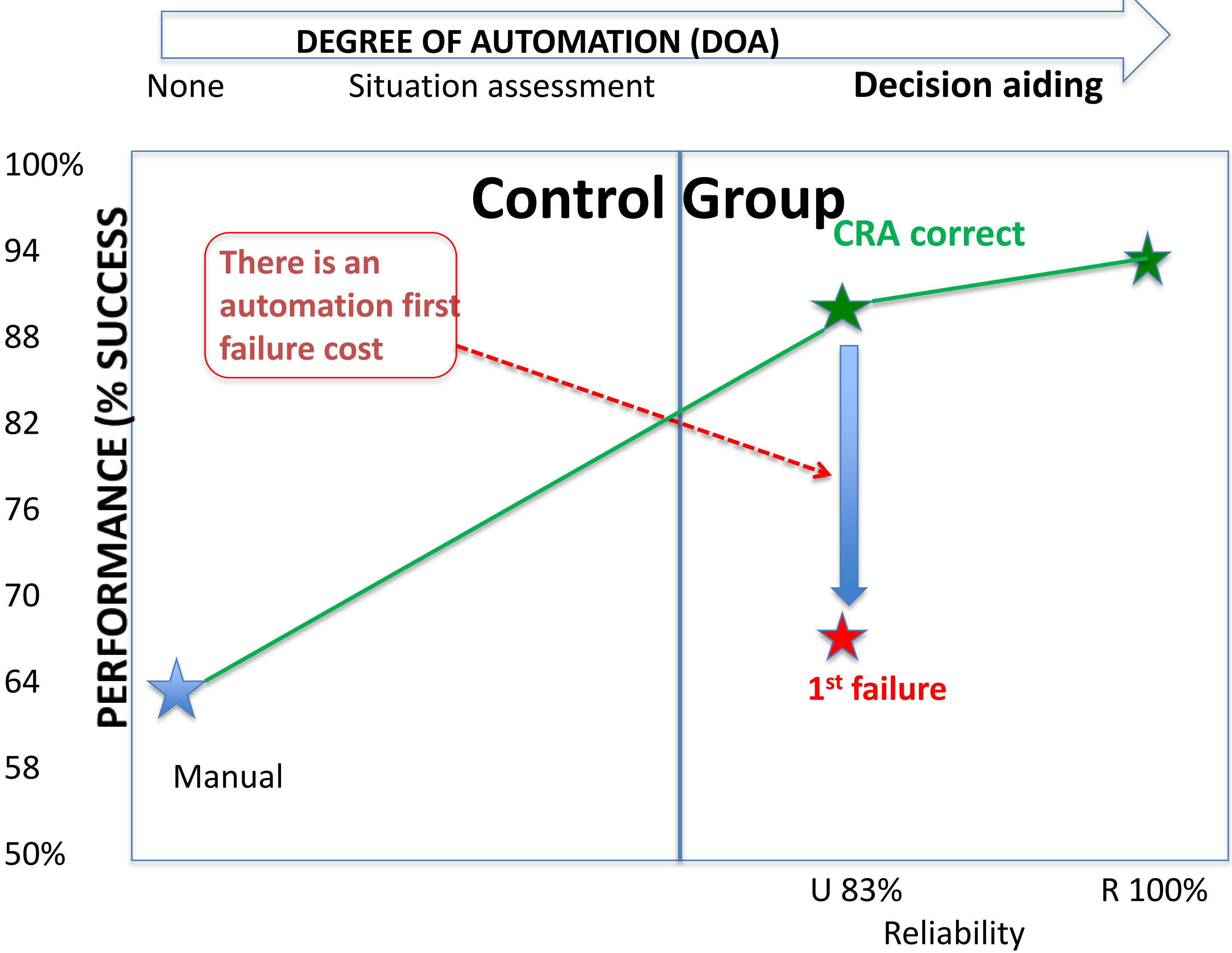
% successfully resolved conflicts (no Loss of Separation: LOS)

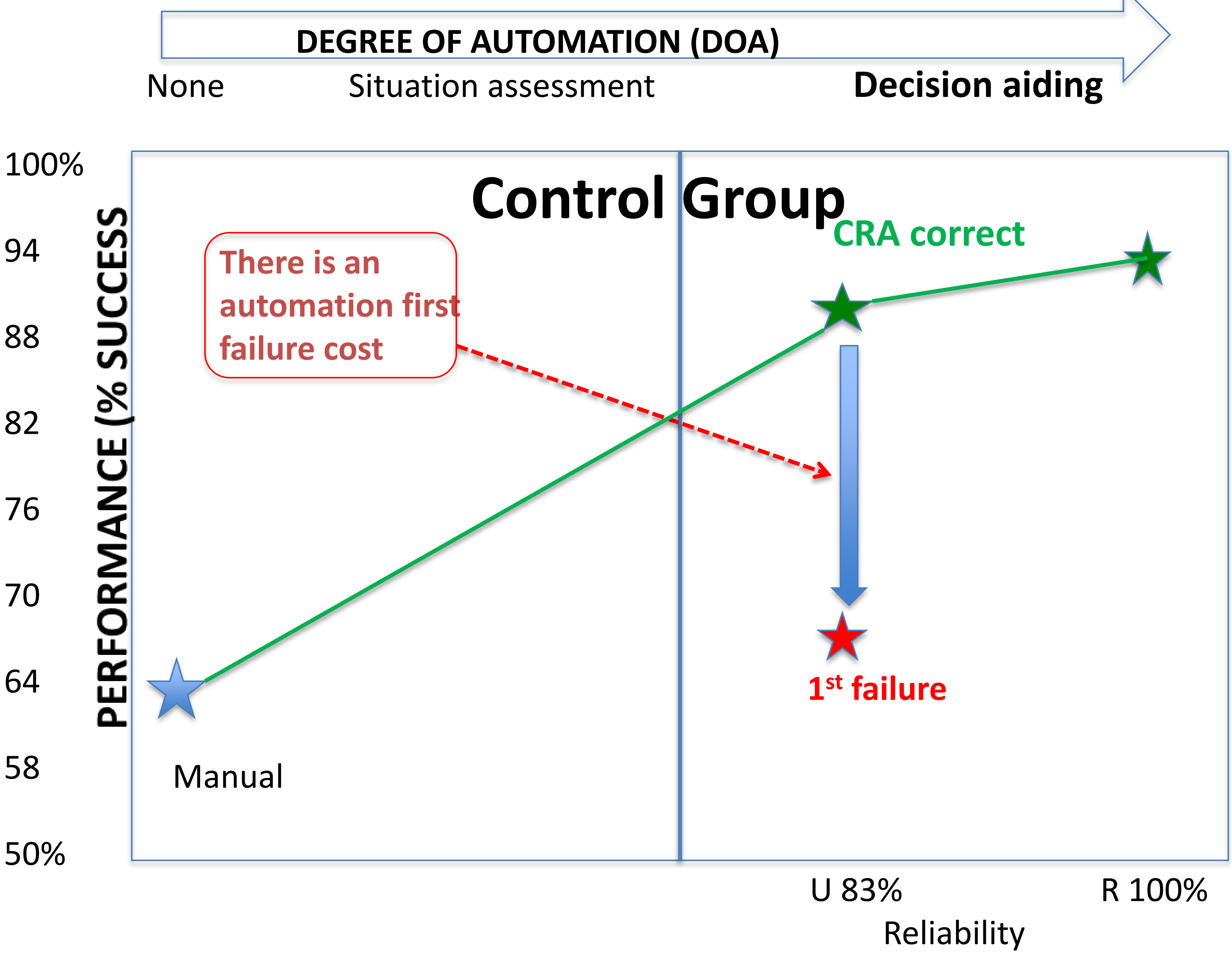
Secondary measure: **Situation Awareness**: answering probe questions about the status of the airspace. Endsley's SAGAT (e.g., "what is the relative altitude of SQ 459?")

The Results (abbreviated)

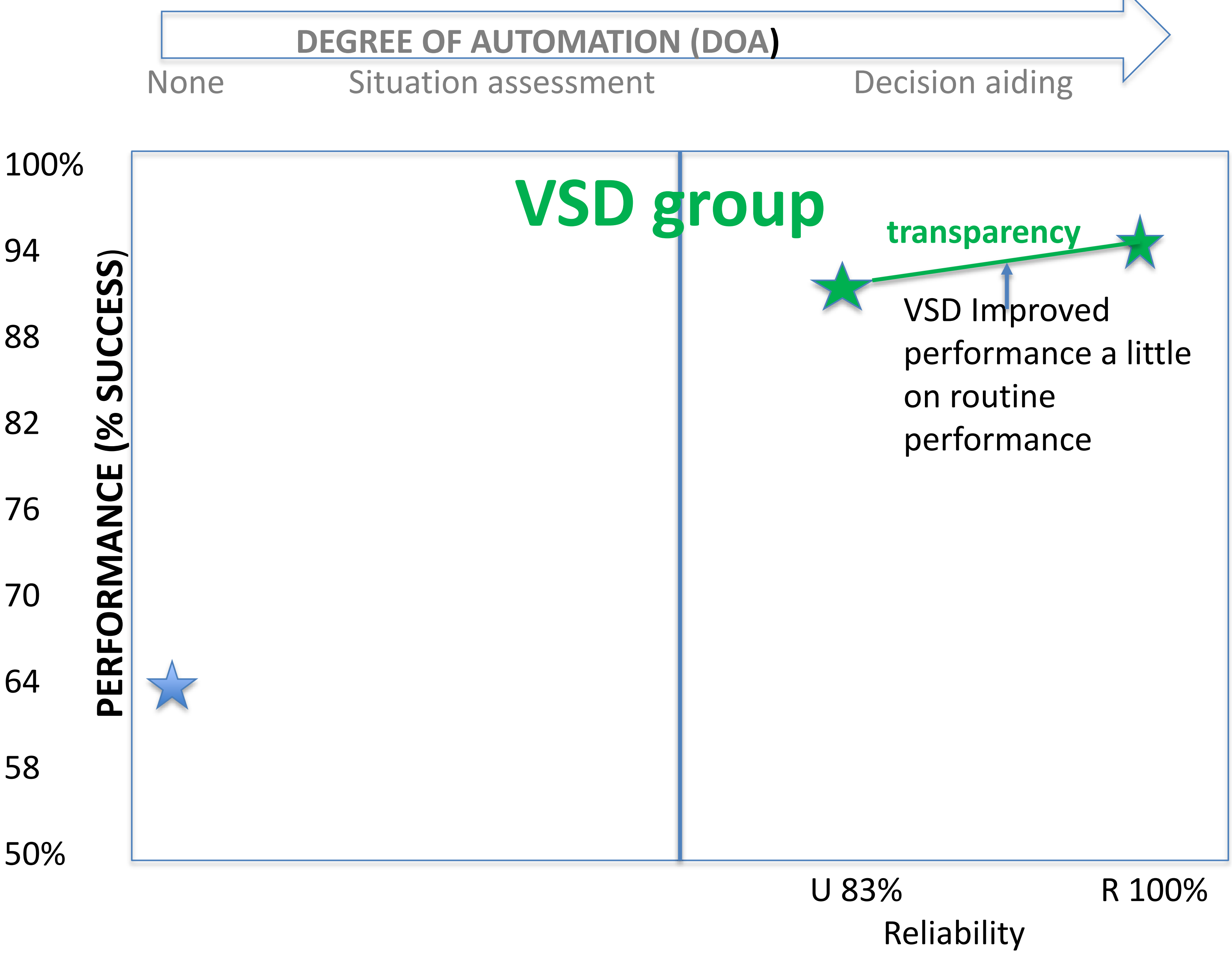


DEGREE OF AUTOMATION (DOA)
None
Situation assessment
Decision aiding
Control Group
CRA correct
Routine Performance
PERFORMANCE (% SUCCESS
100%
94
88
82
76
70
64
58
50%
Manual
U 83%
R 100%
Reliability

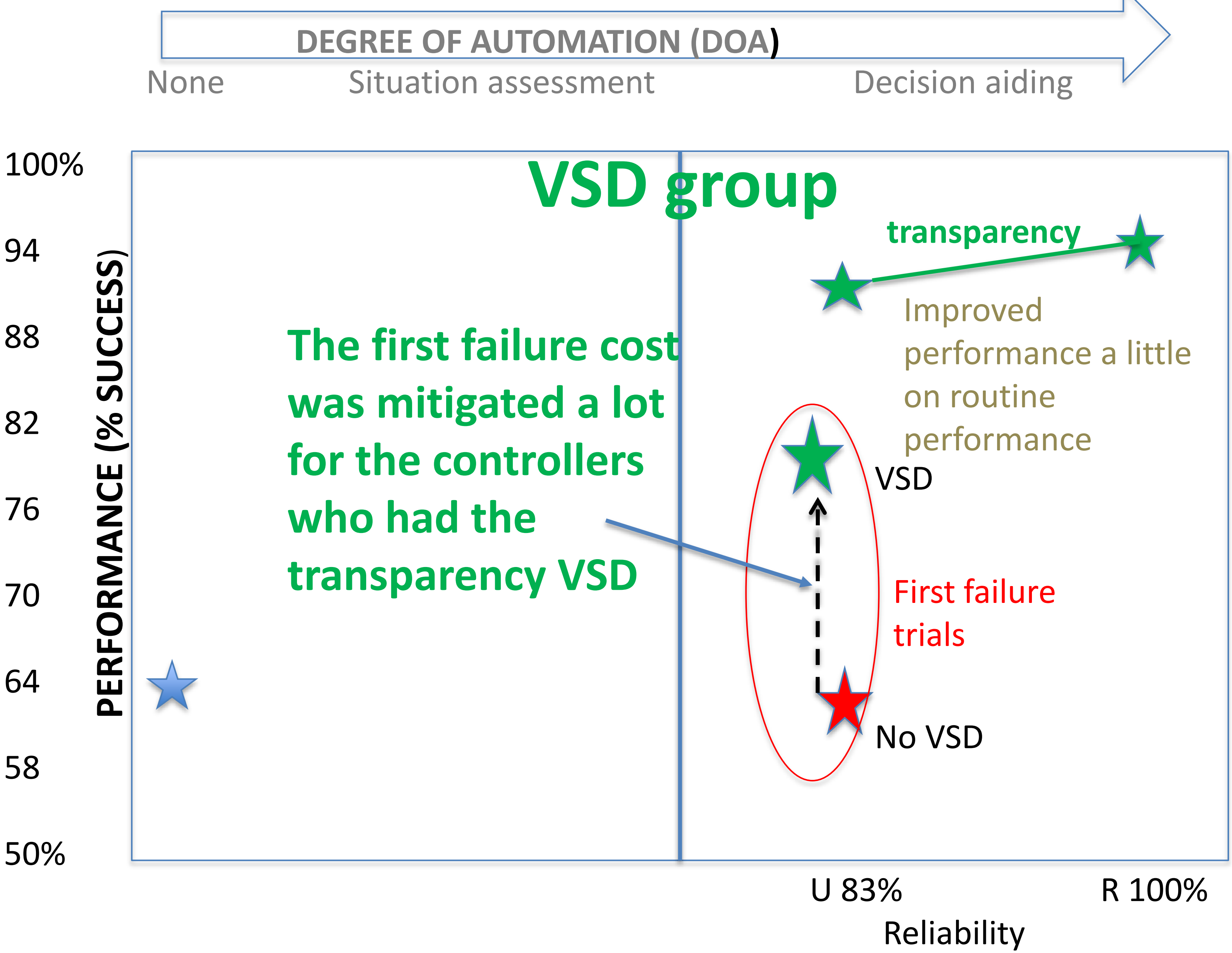

DEGREE OF AUTOMATION (DOA)
None
Situation assessment
Decision aiding
100%
94
88
82
76
70
64
58
50%
PERFORMANCE (% SUCCESS)
Control Group
CRA correct
There is an automation first failure cost
1st failure
Manual
U 83%
R 100%
Reliability



How did the VSD transparency help?

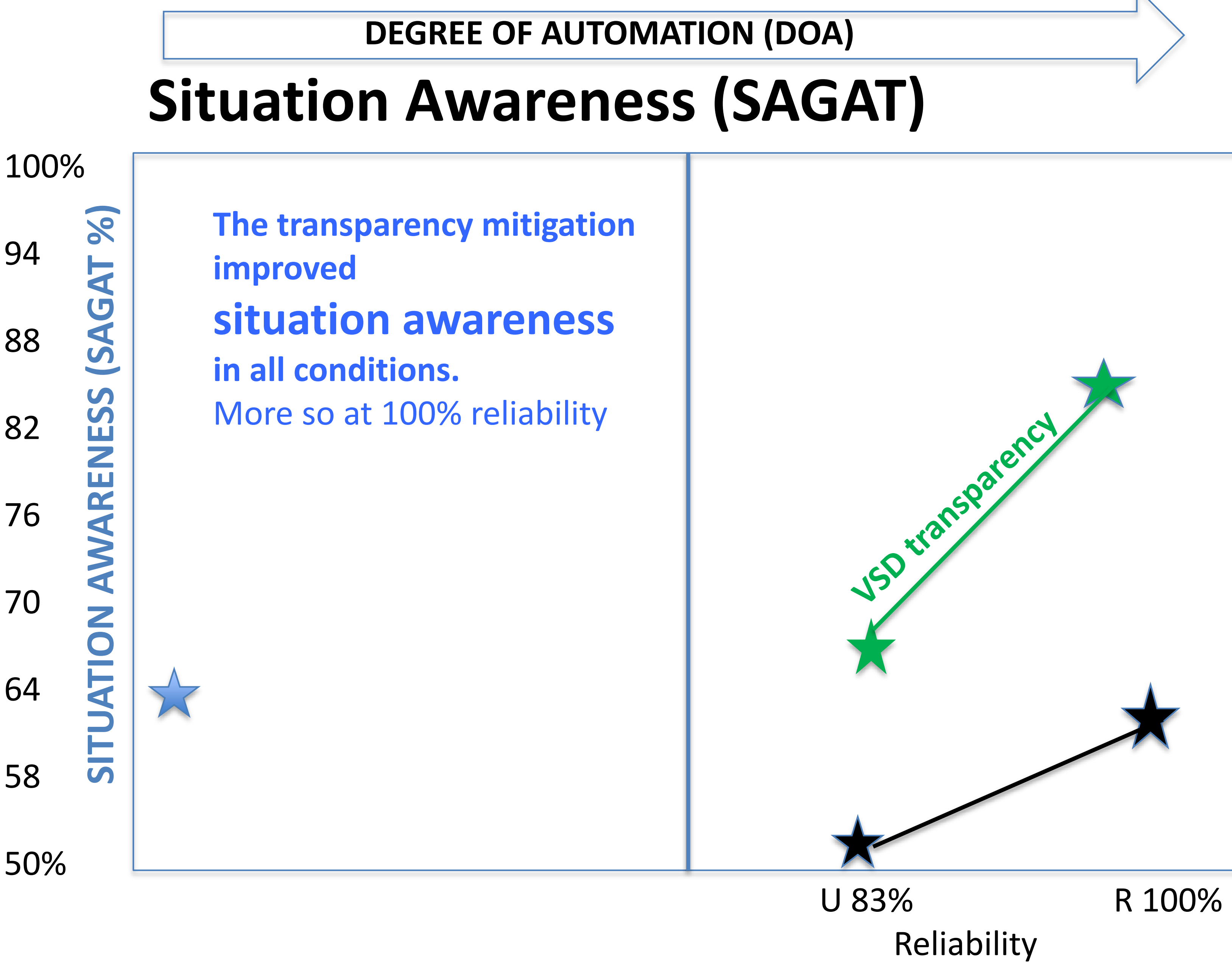


Did the VSD transparency buffer the fall of the tree

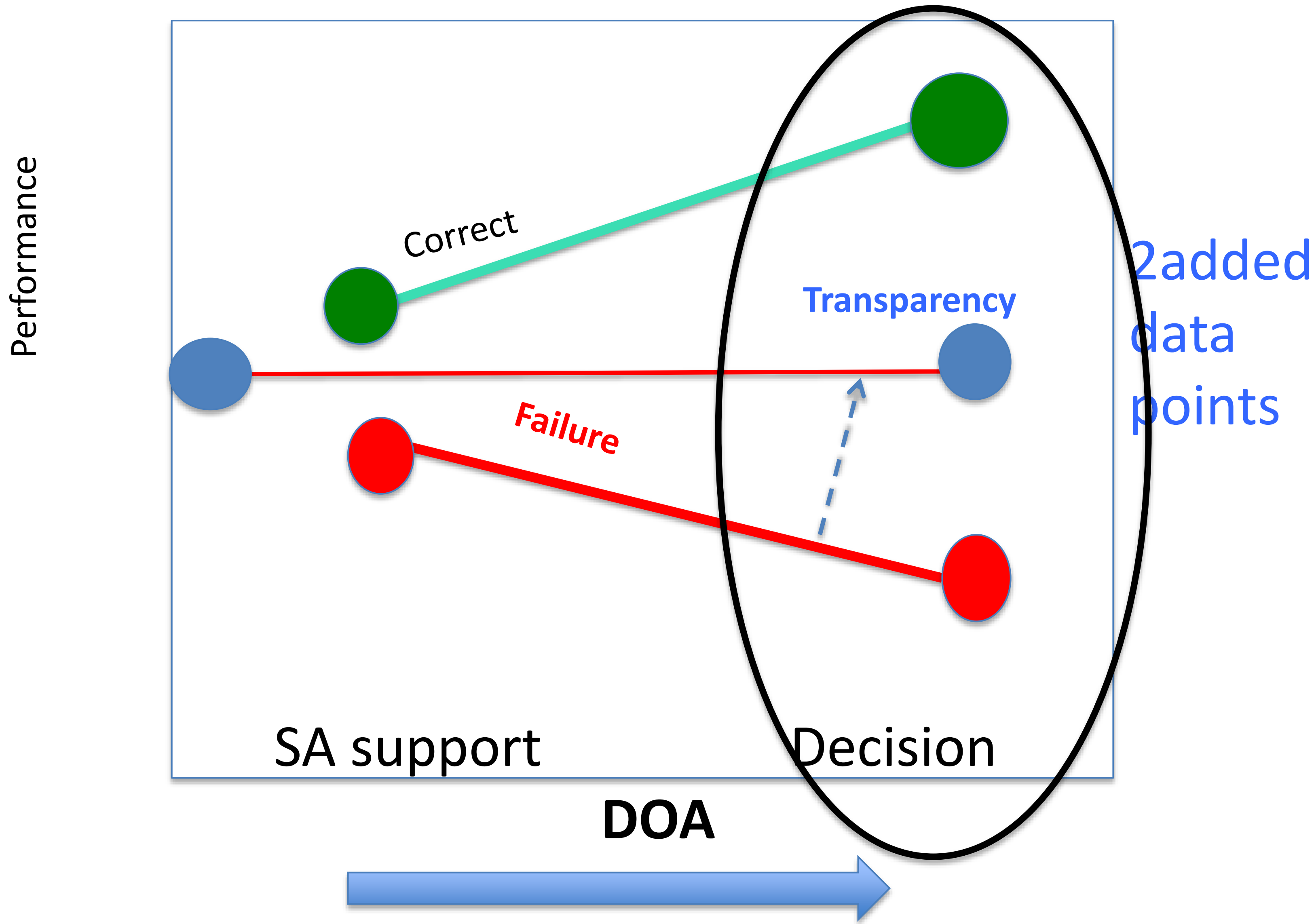


Why did the VSD help? By Improving situation awareness

Situation Awareness (SAGAT)



The Generic Lumberjack: 4 points

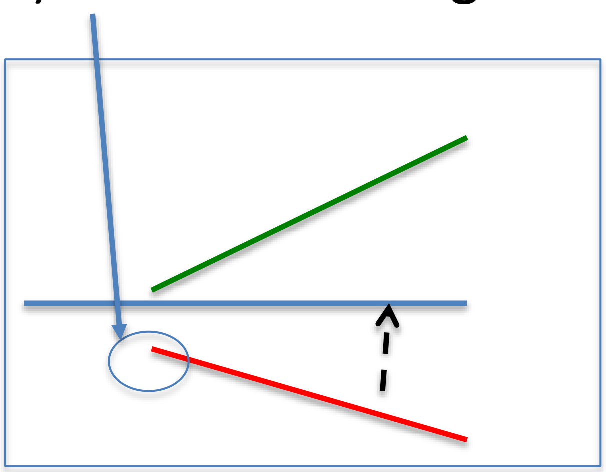


GENERAL CONCLUSIONS: TWO EXPERIMENTS

- Experiment 1: the lumberjack analogy holds from low to high DOA. But we did not examine transparency.
- We modeled automation-induced complacency.
- Experiment 2: At higher DOA (decision support):

- Large benefits benefits of automation over manual: Auto helps
- Large first failure costs
- Large reduction of these costs with VSD display **transparency**
- The VSD also increases **situation awareness**.

But we did not examine failure of SA support automation (low DOA) within the design.

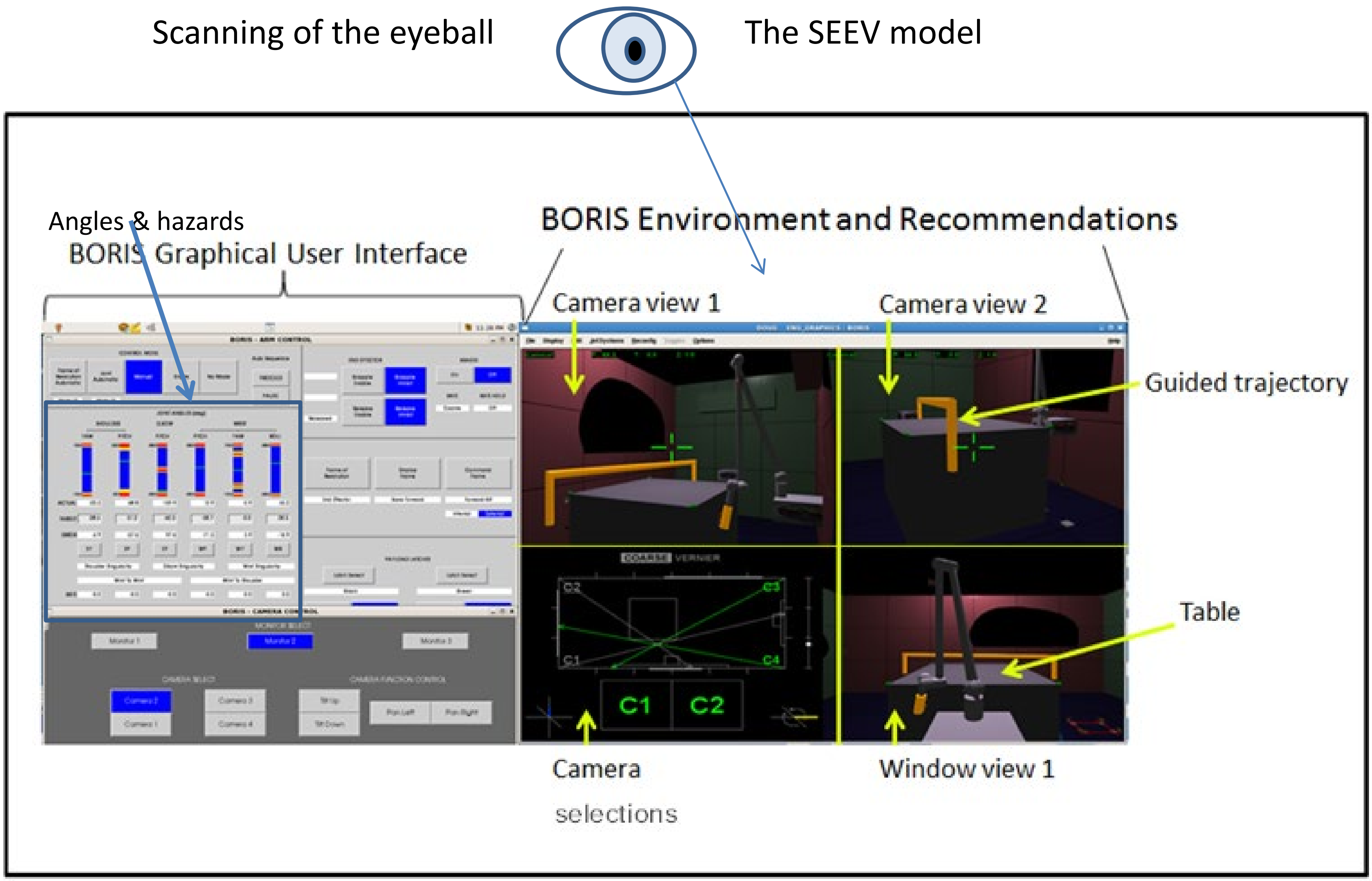


The full experiment remains to be done. But we believe we can soften the fall of the bigger tree with transparency.

Any questions?

Pandawickens94@aol.com

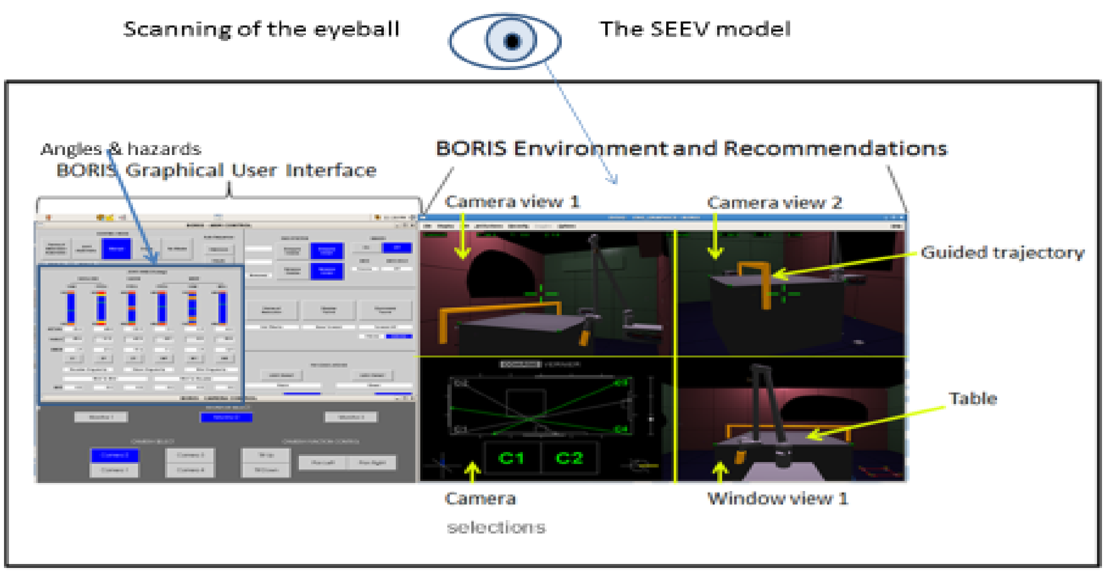
Backup slides



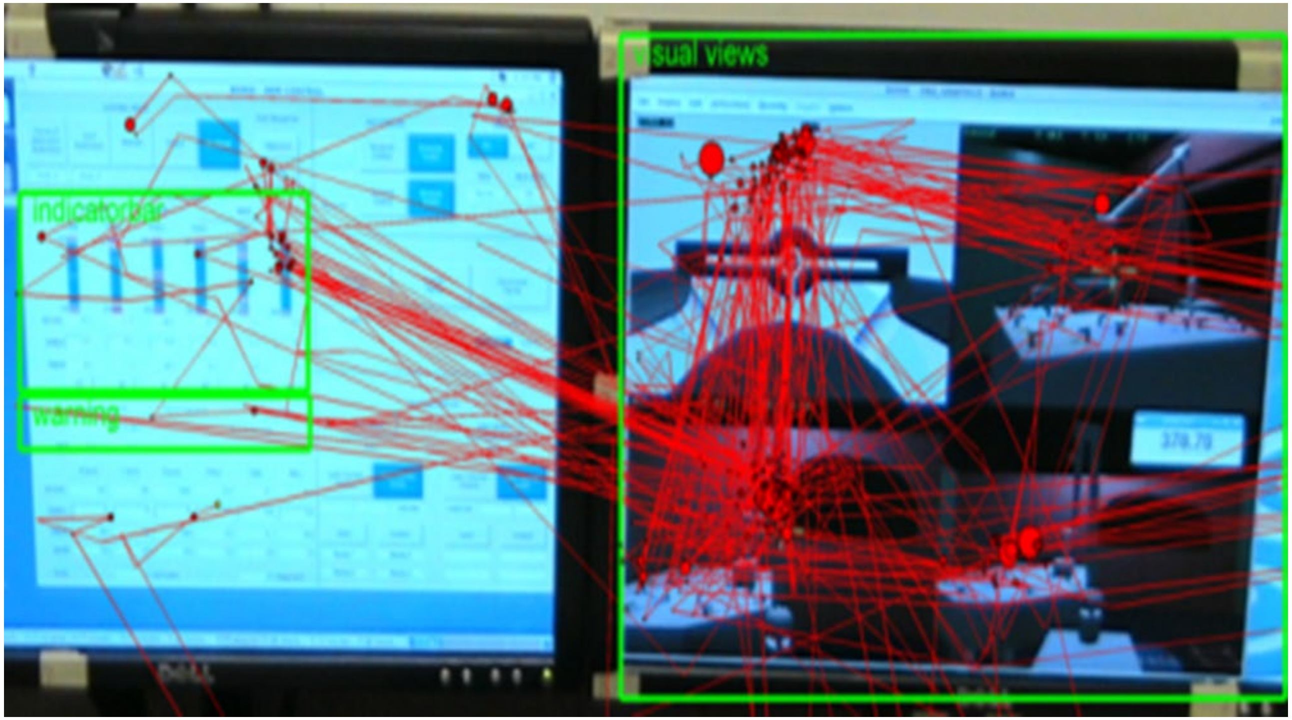
Scanning of the eyeball
The SEEV model
Angles & hazards
BORIS Graphical User Interface
BORIS Environment and Recommendations
Camera view 1
Camera view 2
Guided trajectory
Table
Camera
selections
Window view 1
C1
C2

The computational Human Performance Model of complacency in Automation, as affected by Degree of Automation.

1. Predictive model of how the operator scans displays in the workspace, given the bandwidth of each display, its value to the operator’s task, and the layout of displays. (low bandwidth peripheral displays are rarely sampled even if they are valued).



IN SEEV, we assign coefficients for how fast things change in each display (**expectancy**) and how **valuable** is the information in each display, for the TASK of arm control, within each of the 3 automation conditions.



indicatorbar
warning
visual views
378.70

SEEV. Predicts scanning

The visual workplace is divided by AREAS OF INTEREST (AOI)

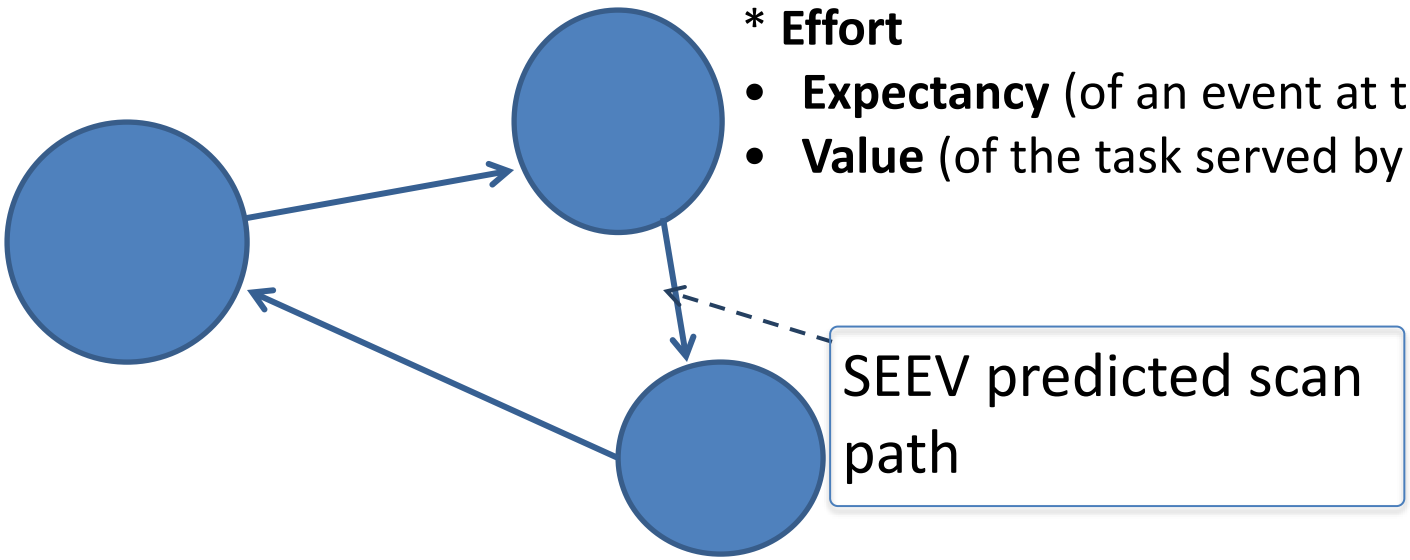
Scanning to an AOI is driven by:

- **Salience** (perceptual models)

Scanning is **inhibited** by moving between more distant AOIs because this takes more

* **Effort**

- **Expectancy** (of an event at the AOI
- **Value** (of the task served by the AOI

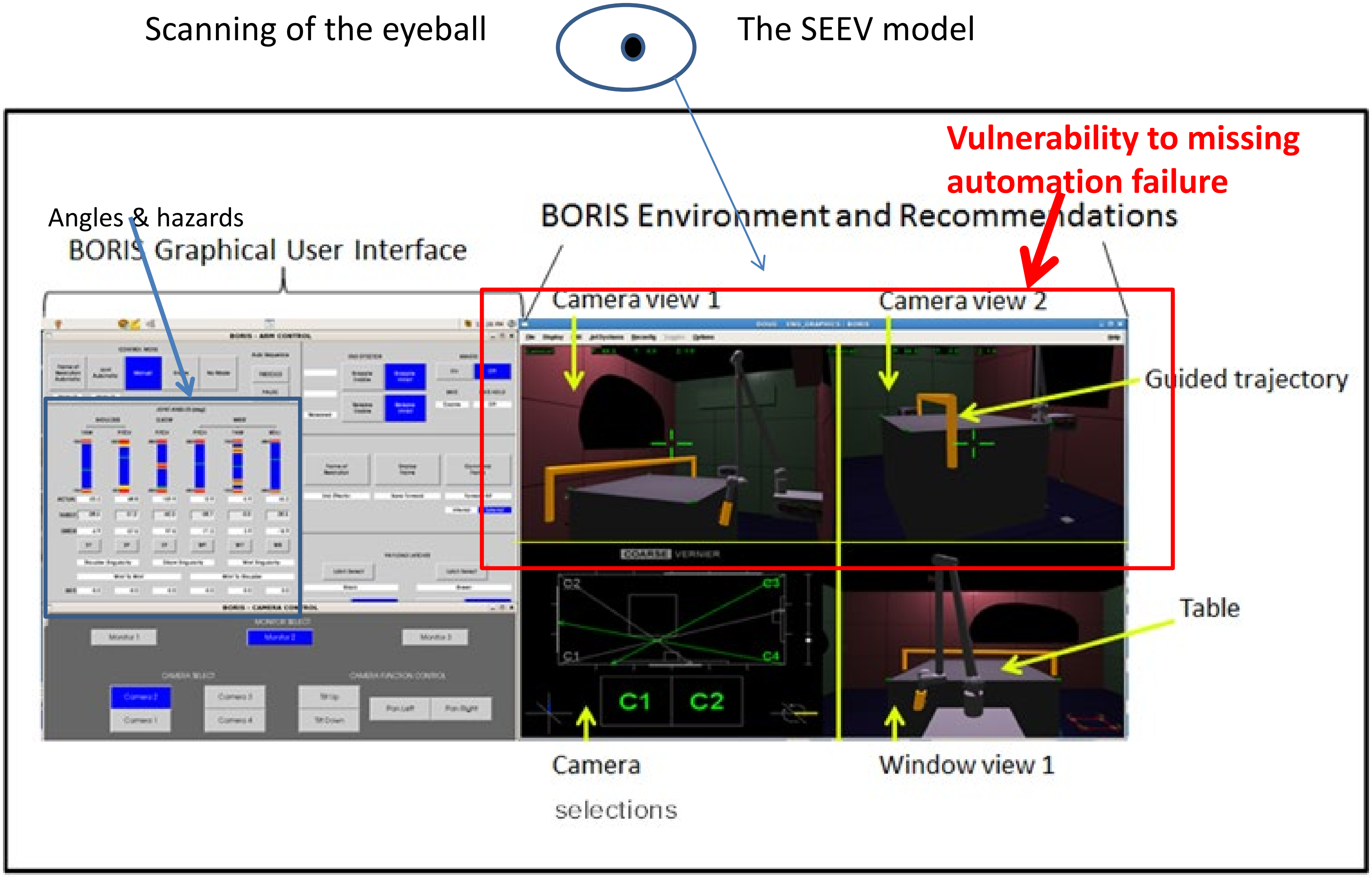


Each parameter (Salience Effort Expectancy Value) can be quantified.

Thus P(scan) to an AOI = [S – E + E + V]

The computational Human Performance Model of complacency in Automation, as affected by Degree of Automation.
(SEEV: Salience Effort Expectancy Value)

1. Predictive model of how the operator scans displays in the workspace, given the bandwidth of each display, its value to the operator's task, and the layout of displays. (low bandwidth peripheral displays are rarely sampled even if they are valued).
2. **This model predicts the delay in scanning the display where automation error is evident (depicting the deviation between programmed versus actual trajectory.**

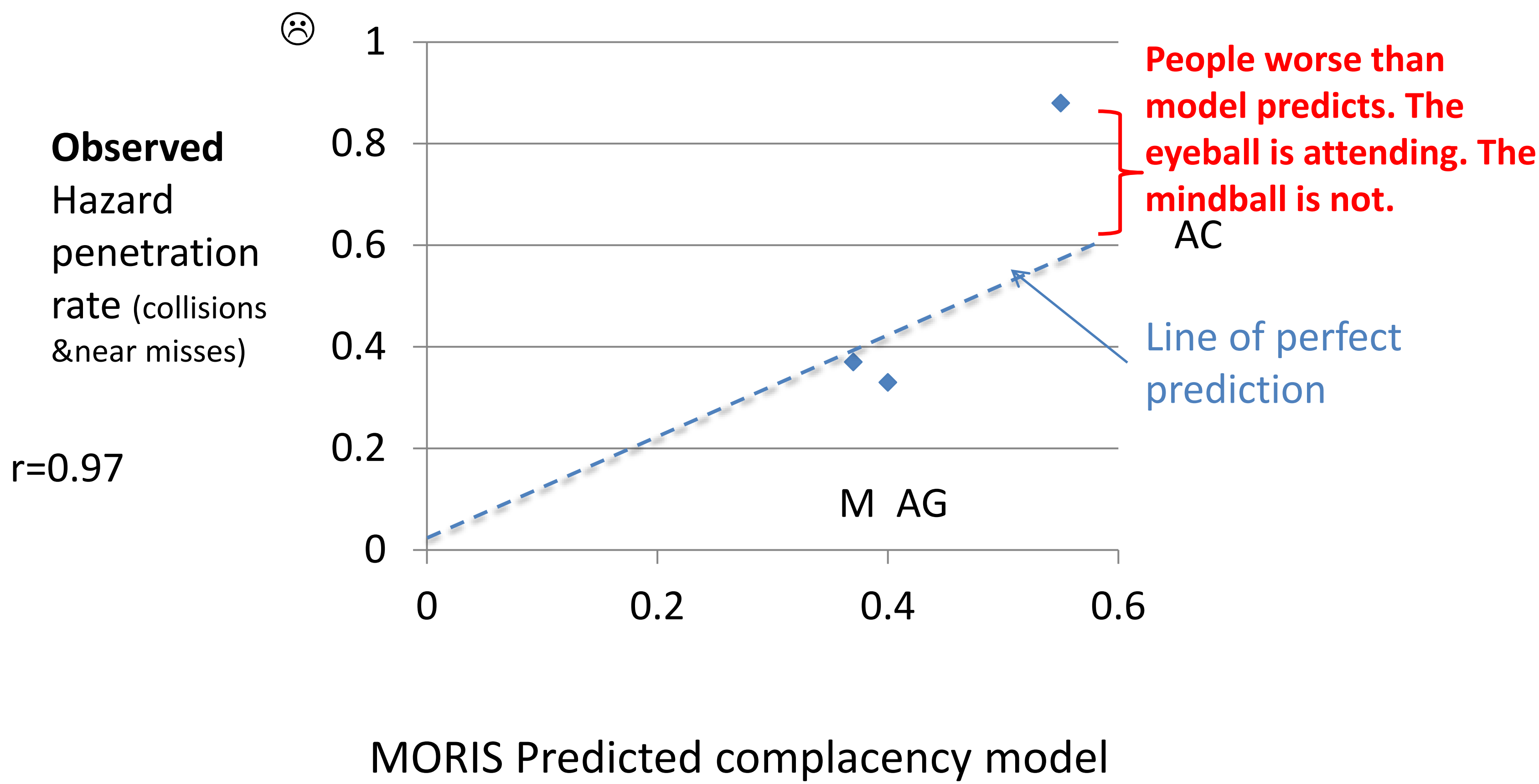

Scanning of the eyeball
The SEEV model
Vulnerability to missing automation failure
Angles & hazards
BORIS Graphical User Interface
BORIS Environment and Recommendations
Camera view 1
Camera view 2
Guided trajectory
Table
C1
C2
Camera selections
Window view 1

The computational Human Performance Model of complacency in Automation, as affected by Degree of Automation.

1. Predictive model of how the operator scans displays in the workspace, given the bandwidth of each display, its value to the operator's task, and the layout of displays. (low bandwidth peripheral displays are rarely sampled even if they are valued).
2. This model predicts the delay in scanning the display where automation error is evident (depicting the deviation between programmed versus actual trajectory.
3. **Since the model predicts different bandwidth and values for different displays under different degrees of automation (none, SA-guidance (the orange line), Auto-control), it can predict different delays in noticing the automation failure, and hence incurring a collision or penetration of the no-fly hazard zone**.

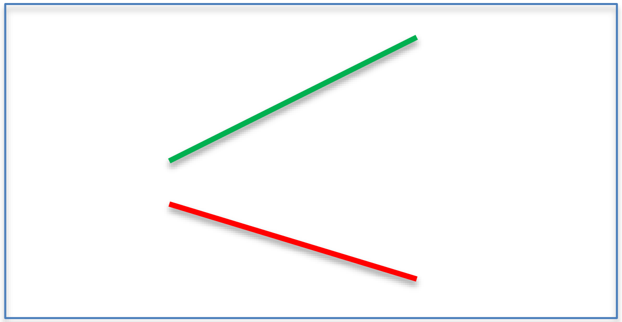
Validation Results

MORIS complacency model Predicted and Michigan obtained failure detection. Missing or delayed responding to incorrect coordinates that direct EE into the hazard.



Conclusion: Experiment 1.

Lumberjack support. Later stage automation (autocontrol) helped performance more than earlier stage automation (SA-support) when correct. Hurt performance more, when **automation failed** (bad data on the trajectory).



We confirmed the Hypothesis for the 4 data points.

We modeled the complacency effect on automation failure

BUT

We did not explicitly provide **Transparency** to offset the late stage automation cost.

Hence applications to both Aero and Space.

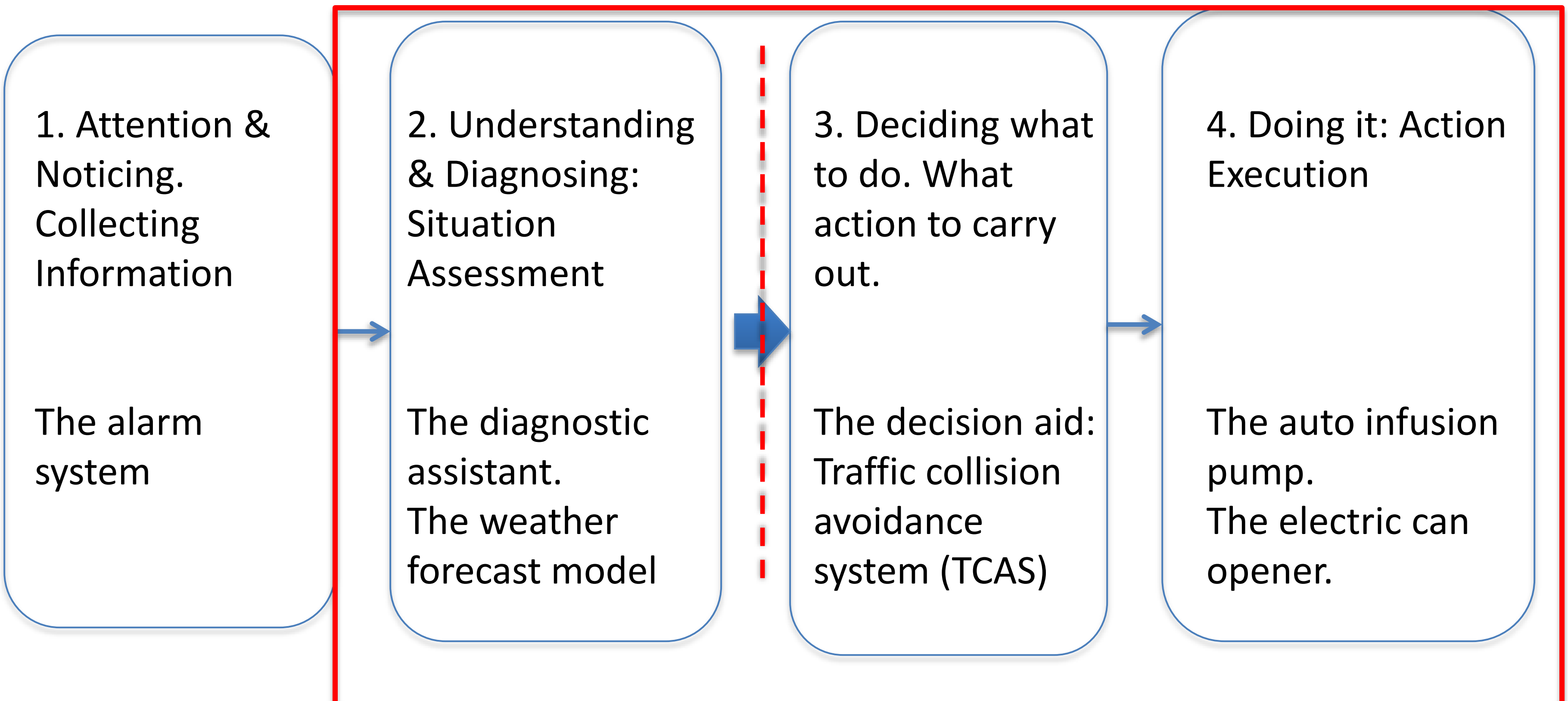
Questions?

cwickens@alionscience.com
pandawickens94@aol.com

Recent approaches to automation have characterized it by **Stages & Levels**

(Parasuraman Sheridan & Wickens, 2000).

Stages: Automation assists human performance at 4 stages of information processing



We will focus today on Stage 2: situation assessment
& Stage 3&4. Decision support: Action choice & execution.