Black Swans, Complacency and Automation Bias

Chris Wickens

Colorado State University

Dept of Psychology

Over 50 years of research in human-automation interaction:
Enduring truths emerged.

*Origins in Aviation accidents.

- The Eastern Airlines Everglades Crash (over-reliance on the autopilot)
- Problems with the Flight Management System (FMS)
- Summary of Aviation automation accidents by Dornheim 1996
- Weiner, Billings, Sarter & Woods: “What is the FMS doing and why is it doing it?”: opacity of automation

- 1990s: FAA’s failure to field their ATC Advanced Automation System (AAS)

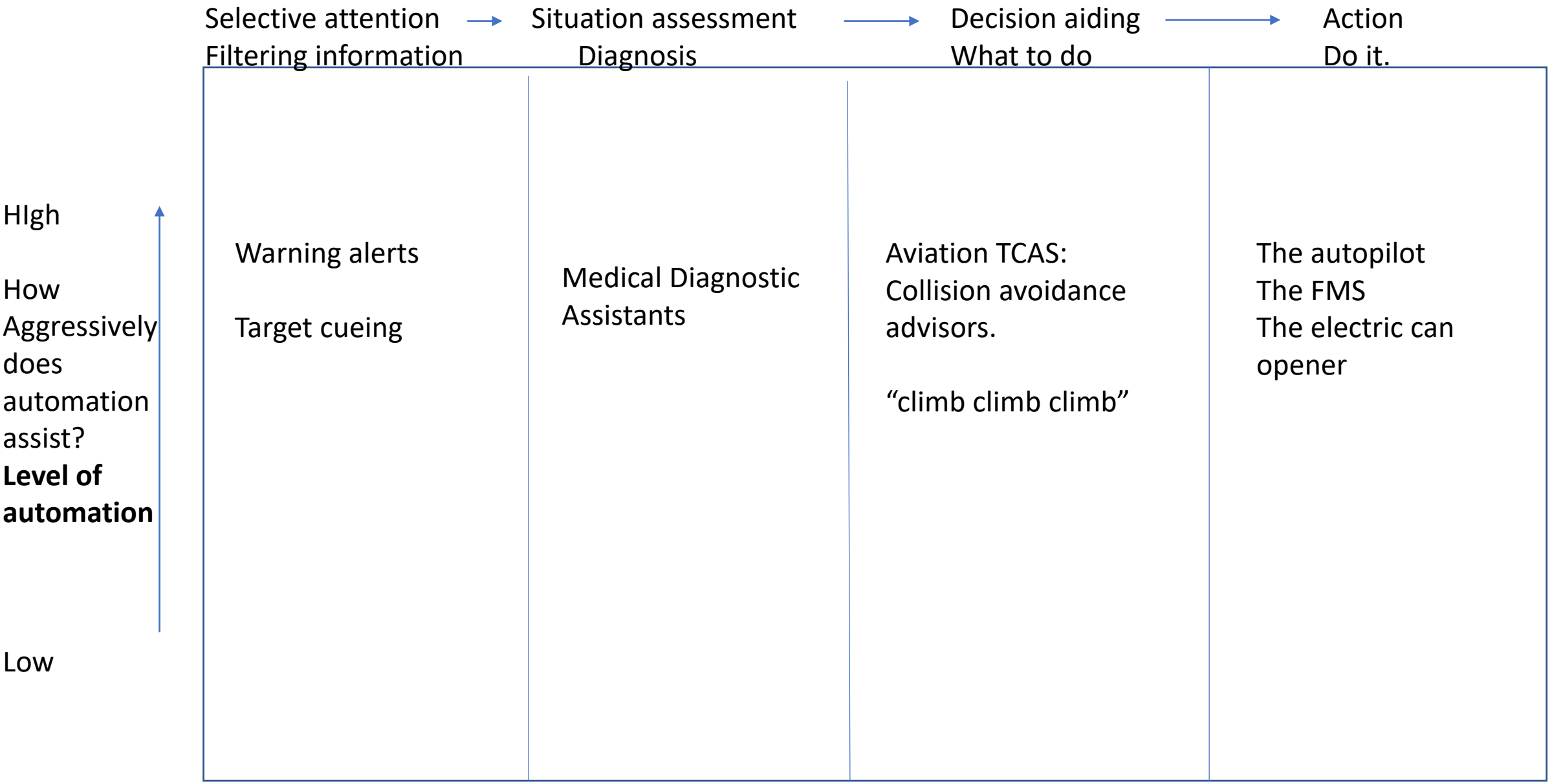
National Research Council panel on human factors of ATC automation Two NRC books: *Flight to the Future, The Future of Air Traffic Control.*

- Could ATC Automation system design learn from Aircraft Automation?

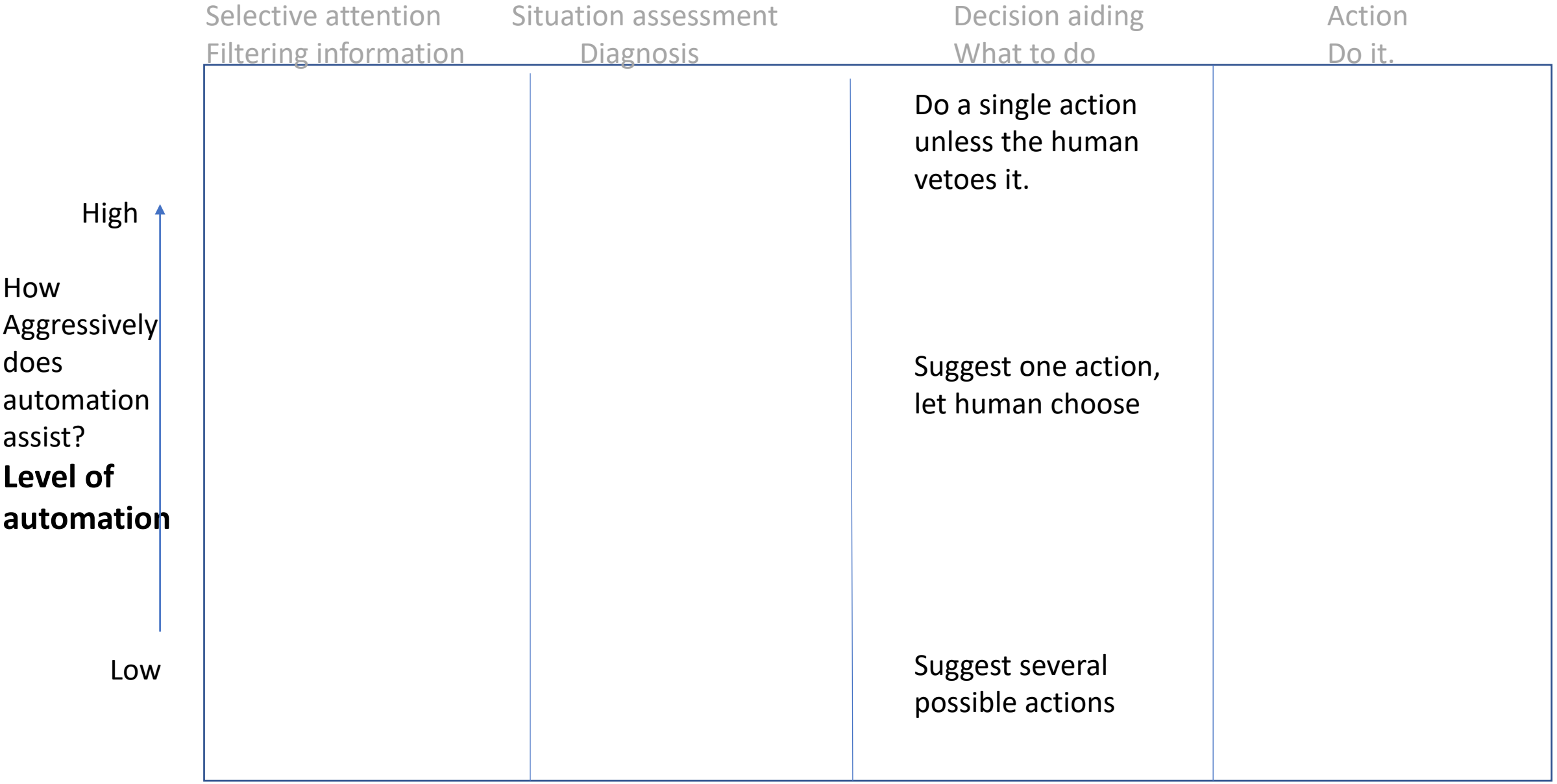
- The Parasuraman Sheridan & Wickens Degree of automation taxonomy

The PSW Taxonomy

How Can Automation Assist the Human? Stage of automation



How Can Automation Assist the Human? Stage of automation



How Can Automation Assist the Human? **Stage of automation**

Selective attention
Filtering information

Situation assessment
Diagnosis

Decision aiding
What to do

Action
Do it.

Complete autonomy

HIgh

How Aggressively does automation assist?

Level of automation

Low

Degree of automation: (DOA) More later stages and/or higher level.

Implications of the DOA Taxonomy for human automation Interaction:

* The higher the degree of automation:
    - The more it helps performance (and reduces workload) when it works
    - The more **catastrophic** the human response when automation **fails**.

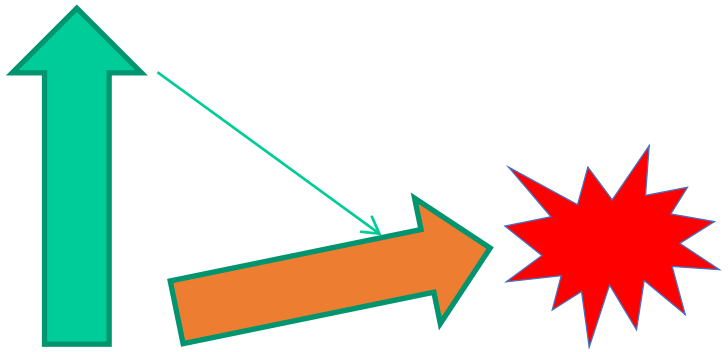
The Lumberjack Analogy: “The Higher the tree, the harder it falls.”

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: aggressive Decision/action support

Implications of the DOA Taxonomy for human automation Interaction:

* The higher the degree of automation:
    The more it helps performance (and reduces workload) when it works
    The more **catastrophic** the human response when automation **fails**.

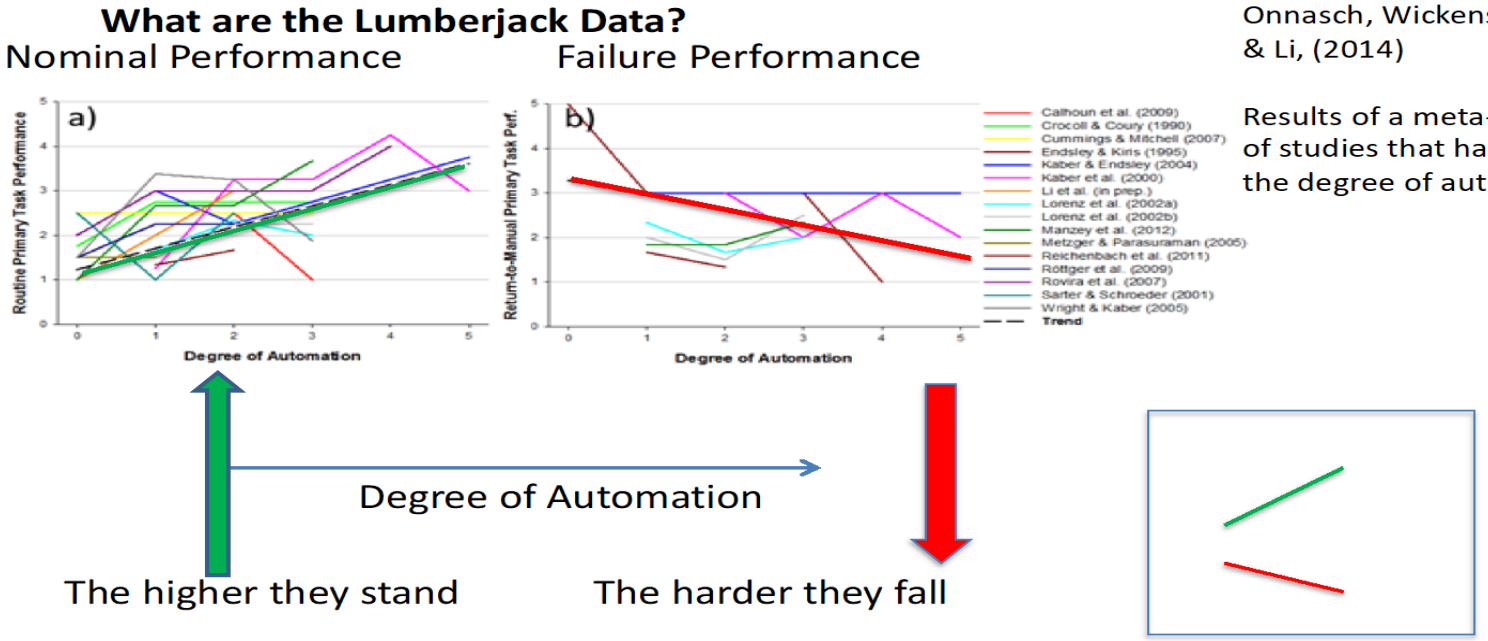
The Lumberjack Analogy: “The Higher the tree, the harder it falls.”

What is an automation failure?

* Hardware failure (the power source disconnects, the hydraulic valve stuck)
* Software Failure: Millions of lines of Code in the Aircraft FMS
* Fails to perform **as the human expects.** Human has the wrong “mental model” of what automation is doing: The Aviation FMS accidents.

What Are the Data?

- Accident/Incident reports (Mumaw)
- The Onnasch et al. Meta-analysis

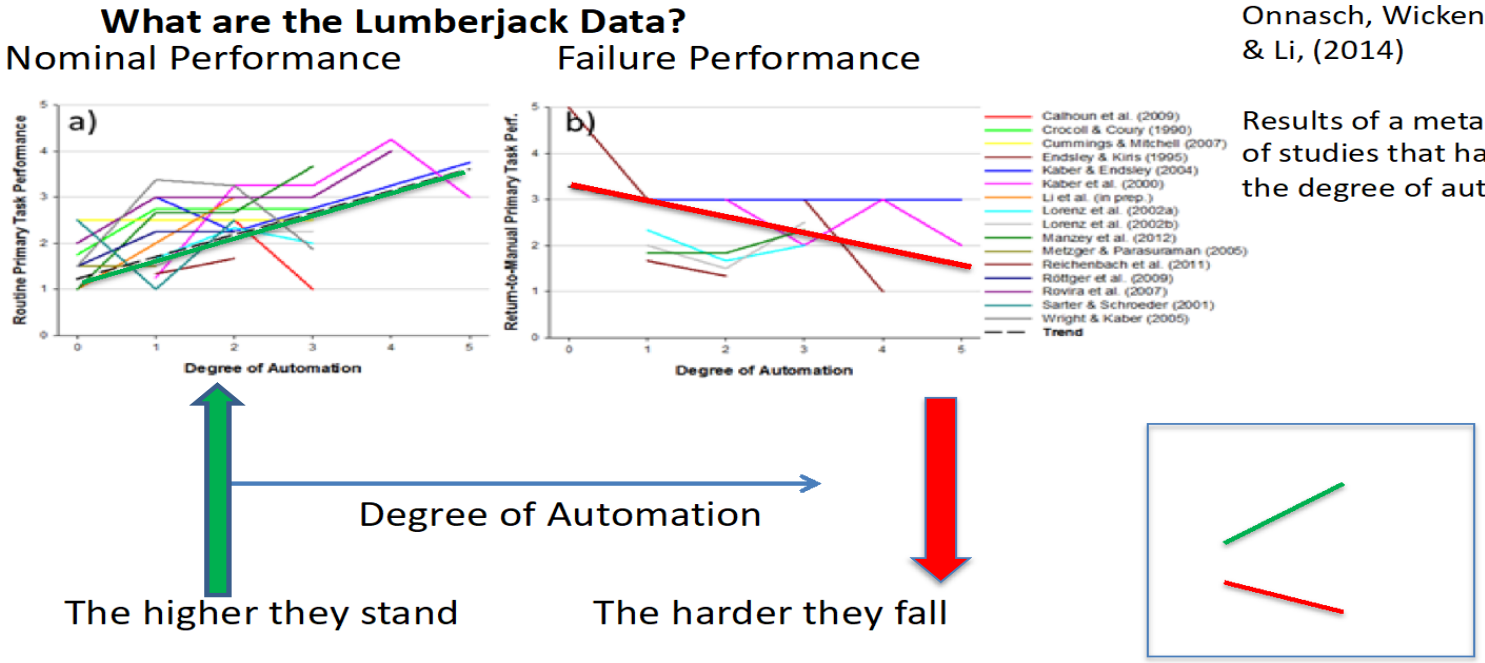


Onnasch, Wickens, Manzey & Li, (2014)

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

What Are the Data?

- Accident/Incident reports
- The Onnasch et al. Meta-analysis



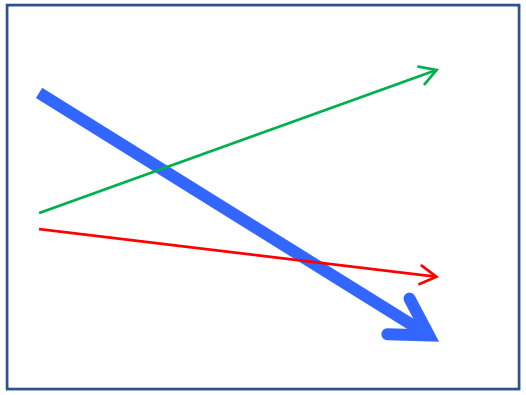
Onnasch, Wickens, Manzey & Li, (2014)

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

What underlies this increased cost, at higher degree of automation, when automation "fails"?
The Loss of Situation Awareness

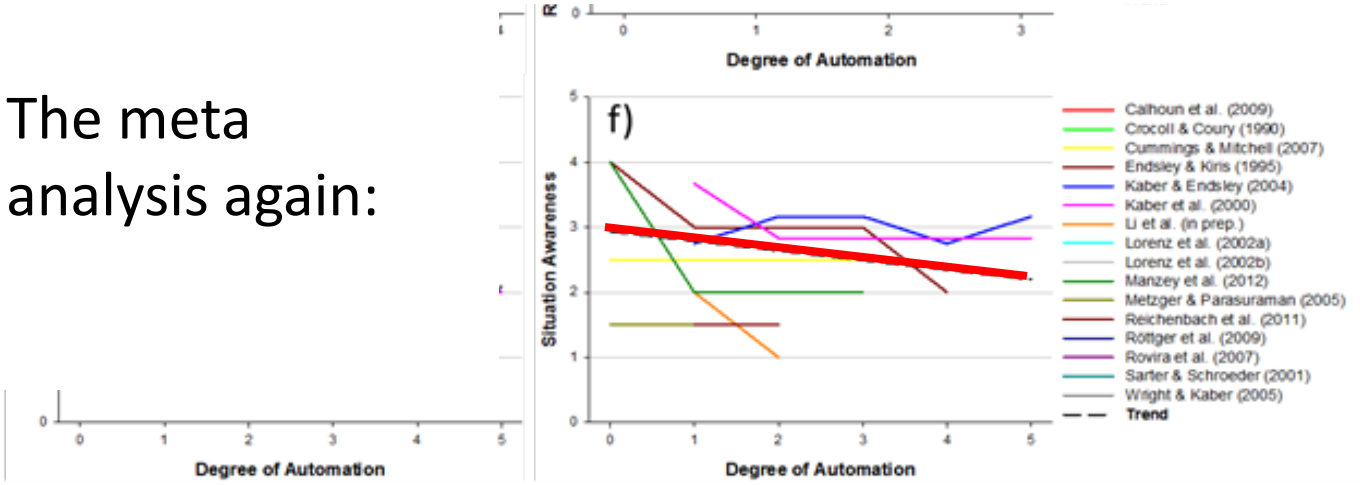
Loss of Situation Awareness (LSA) at higher degree of automation when automation is working well.

Situation Awareness



degree of automation

The meta analysis again:



Endsley's Three level model of Situation Awareness.

Perception-Understanding-Prediction. Situation awareness is lost at:

Endsley's Level 1: At higher degree of automation, people **look less** at the process that automation is supporting. They stop monitoring it. They don't notice the failure (The Tesla Crash): **Complacency.**
They do what automation says without **looking at** the data supporting the automation's decision: **The Automation Bias**.

Endsley's level 2: At higher DOA, people fail to **understand** what automation is doing at the time of the failure. They intervene incorrectly, too late, or not at all. (The 2013 Air France Crash over the Atlantic).

Collectively: loss of SA levels 1 and 2 create the **out-of-the-loop unfamiliarity. Syndrome: OOTLUF**

What about Black Swans? (Taleb)

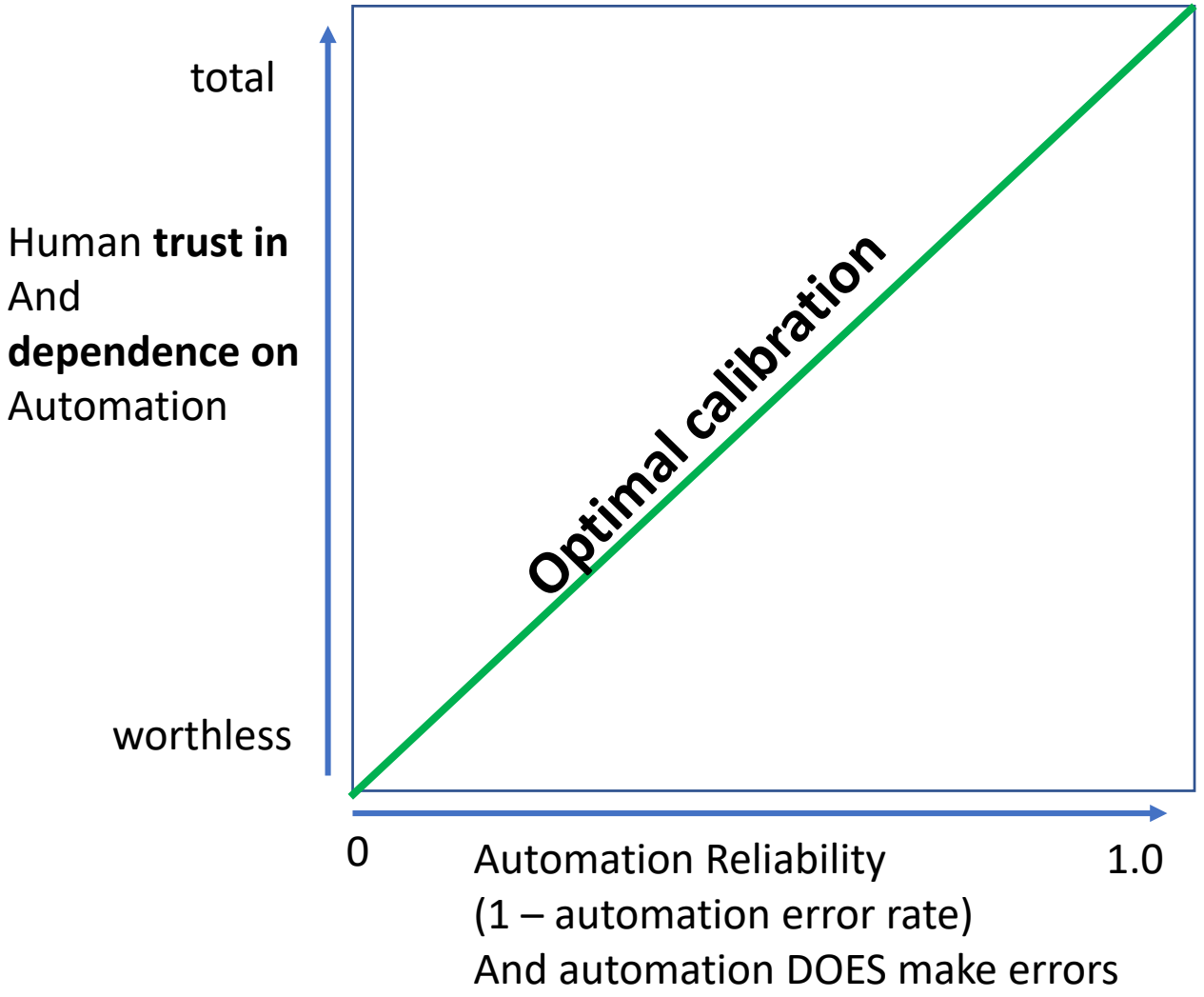
Very rare events that the user of automation does not consider will happen (at least to them).

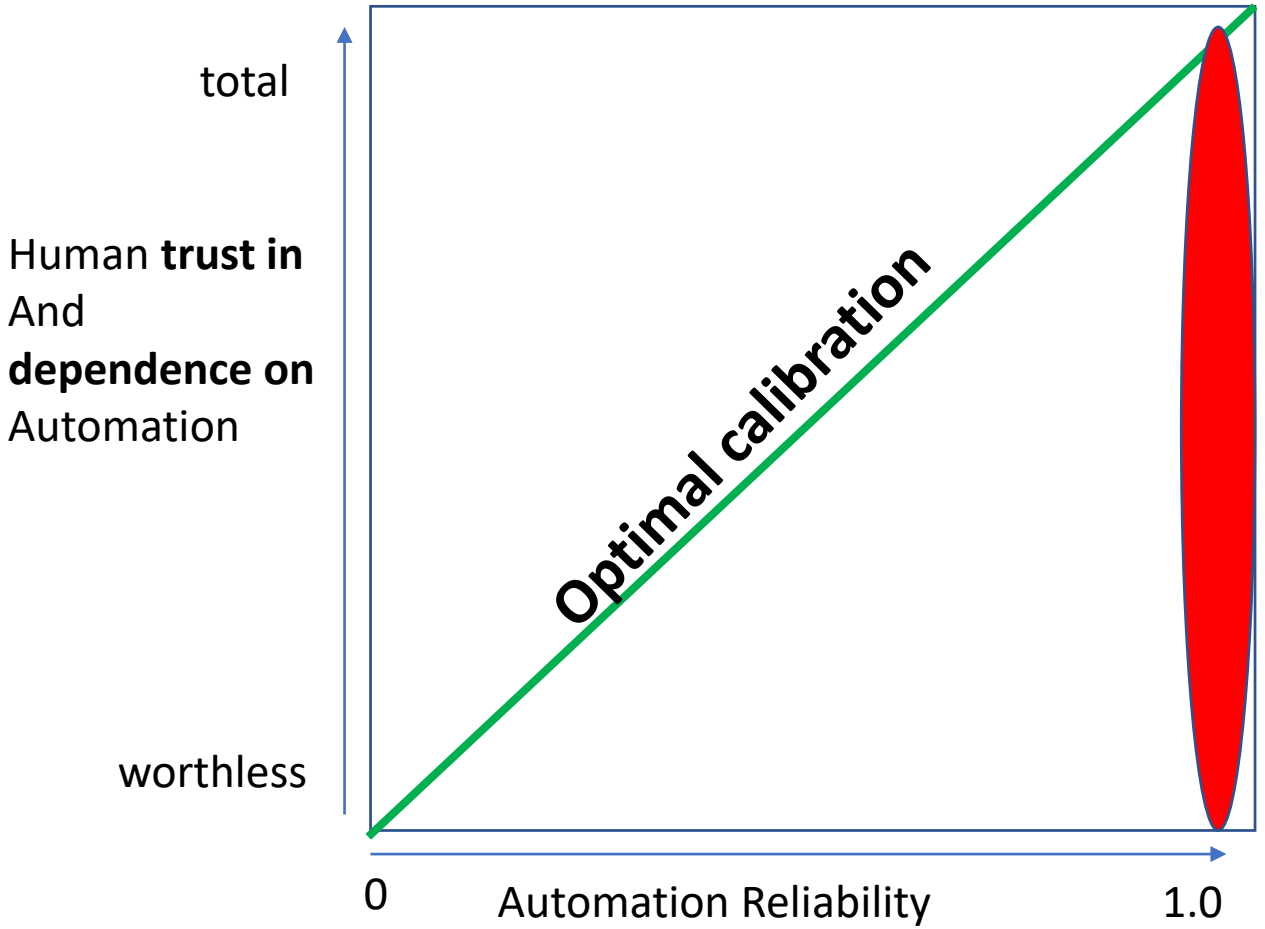
The very rare black swan event: **the automation failure** for one who has never experienced it: the Tesla Crash.

The automation failure rate determines **automation reliability**. (r= 1-p(f))

Black swans in the context of the The Dependency Calibration Space

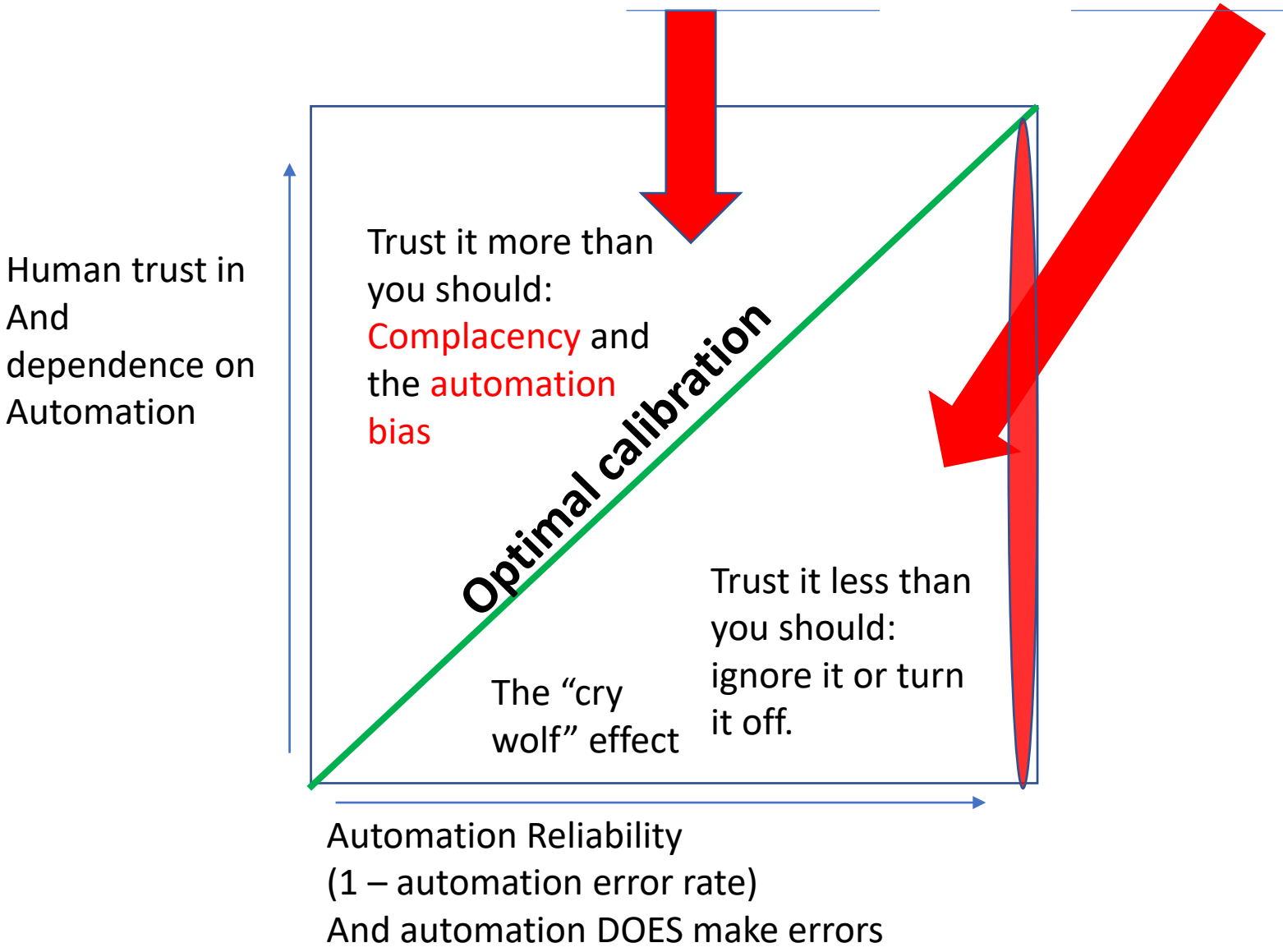
The Reliability-trust calibration space



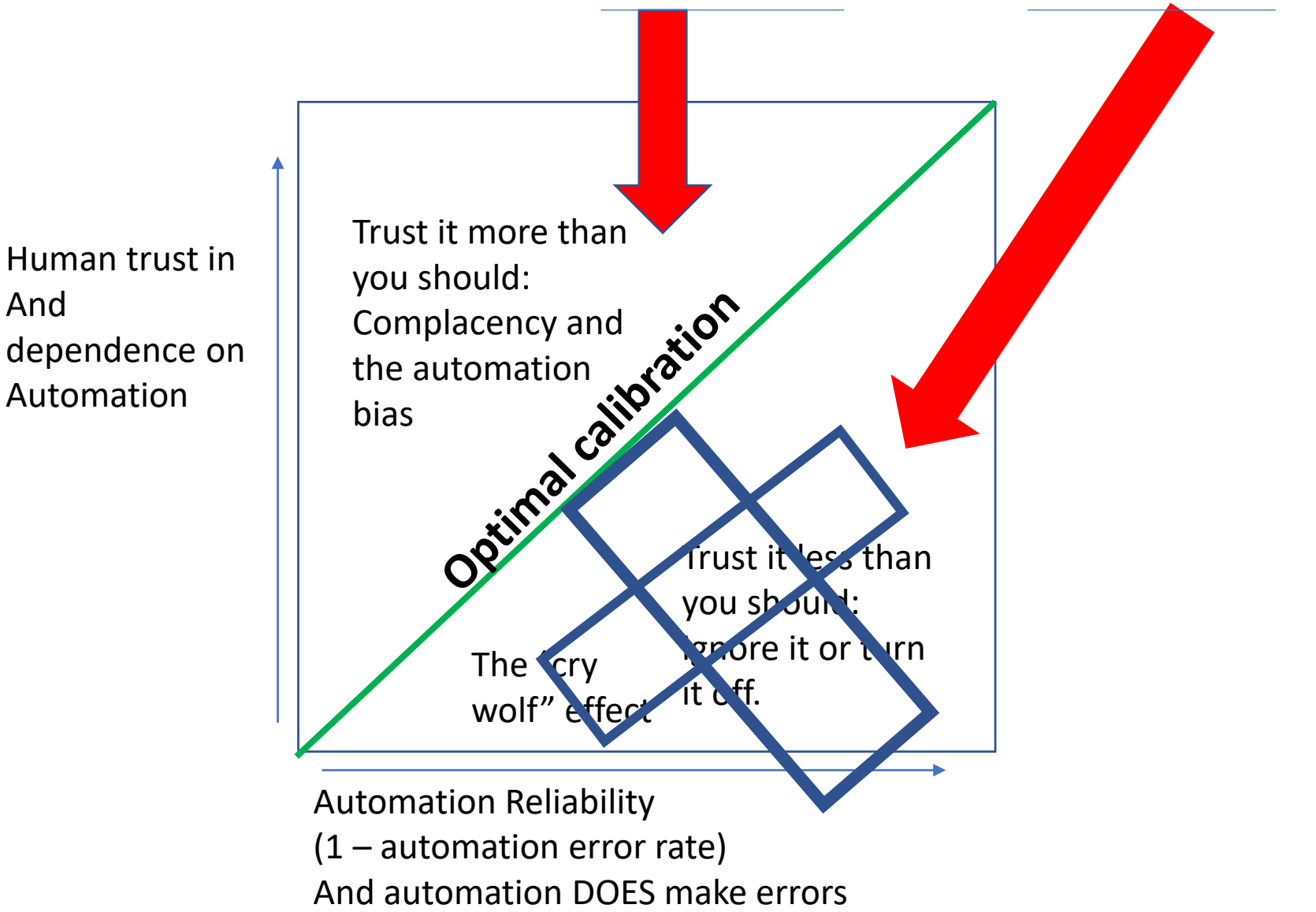


The Black swan: very rare automation failure.
0.1% in the population.
Never, for the typical individual user.

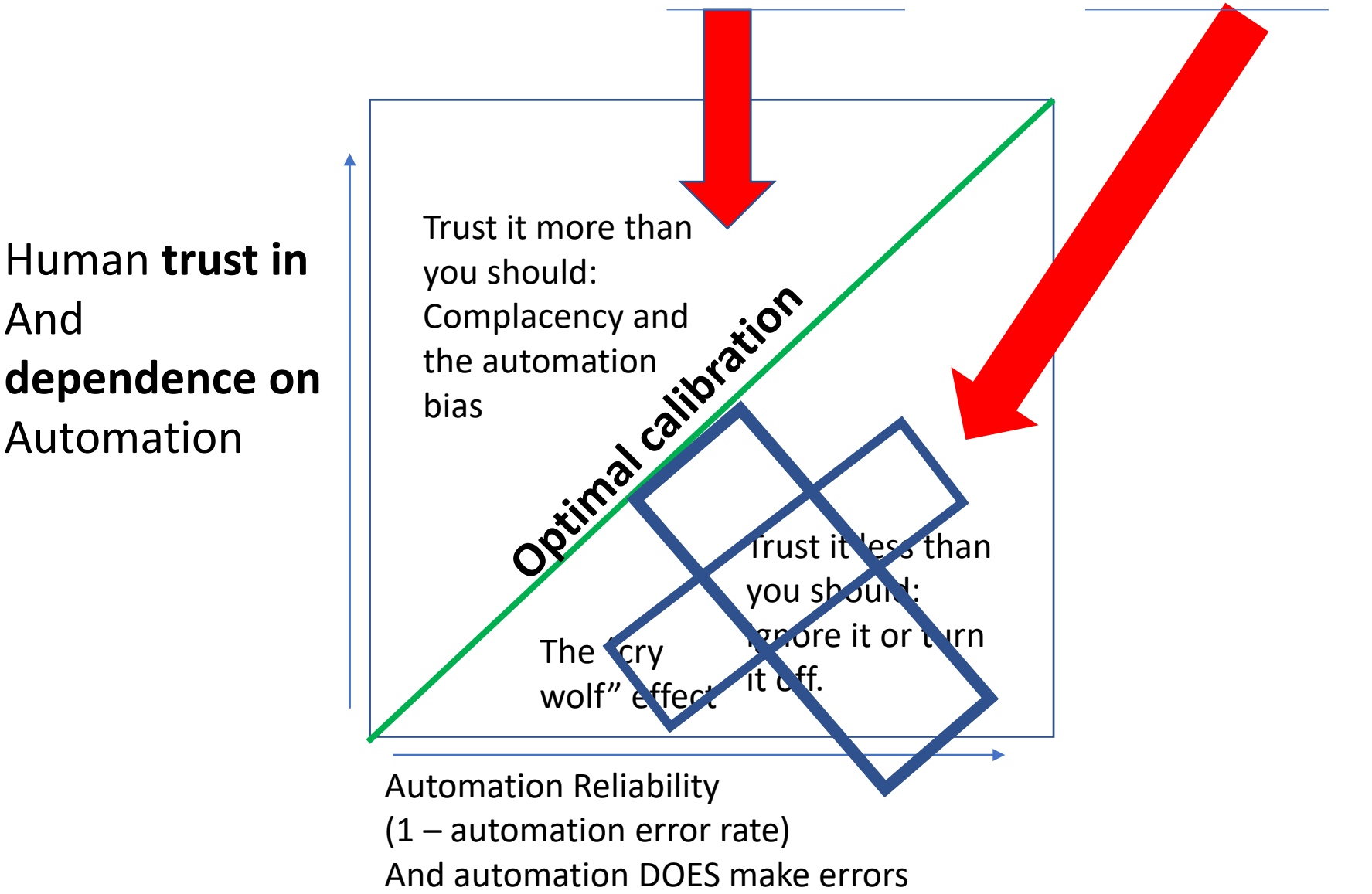
Poor Trust Calibration: Over trust and under trust



Poor Trust Calibration: Over trust and under trust



Poor Trust Calibration: Over trust and under trust



Trust/dependency: The Y axis. Often co-vary, but not the same:

* Trust: a subjective **attitude**.
* Dependence: an objective **behavior.**
* Often the correlation between the two, while positive, is low. (Pharmer, Wickens, Clegg, Smith, 2021).

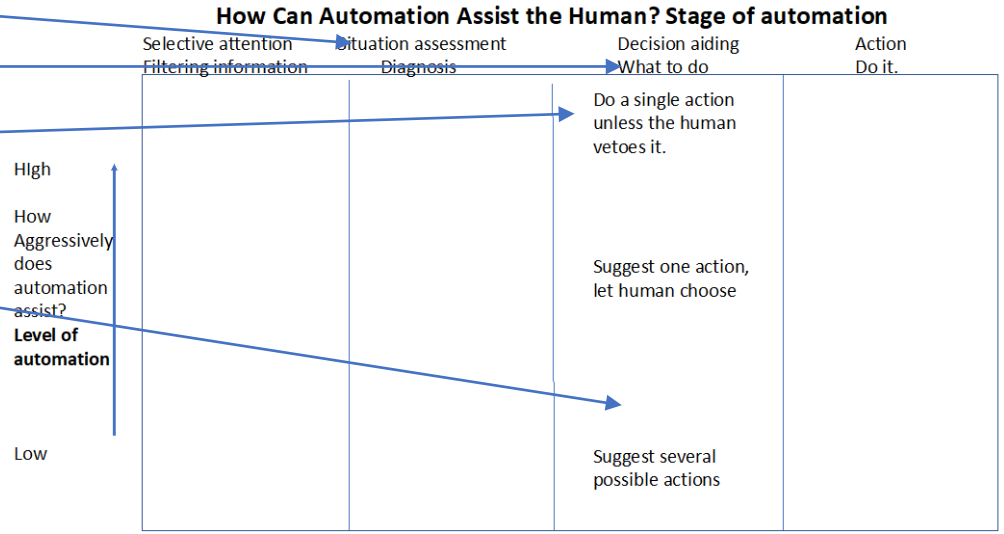
- Implications: trust is easy to measure, but we should not use it as a proxy for dependence (and automation use).

- People will ignore trusted automation and do the task themselves if they enjoy it
- People will use automation they don't trust if they have to: high workload, supervisor or company mandates that they do.

Three Human factors Solutions to address OOTLUF in human-machine teaming.

1. Lower the Degree of Automation:.

Support Diagnosis and situation assessment
Rather than decision making
Implement a lower rather than the highest level
Reduces automation failure cost,
but also reduces benefits. (no Free Lunch)



2. Implement **automation transparency**,

A solution based on findings that OOTLUF is Loss of SA.

Automation Transparency (ATP) One of the great success stories of HAI research. A free lunch?
15 studies of ATP reviewed. (Wickens Helton et al, 2022)
12 found ATP mitigated (and sometimes eliminated) costs of automation failure.
None found that it amplified those costs.
What **is** ATP?

ATP offered off line (before automation is used): training:

- How automation works
- How it could fail (Beller, 2012)
- Conditions of failure. (The self-driving car: sensor can't see low contrast in poor visibility)

ATP Offered on line (while automation is in use).

- Intuitive Display (Seppelt & Lee 2007, Mumaw, 2021)
- Automation expression of its own reliability. (Chen et al, 2018: decision aid: "I am 70% confident in this recommendation").
- Explanation of its reasoning for a particular decision (Mercado, 2016).
- Display (or easy availability) of the raw data processed by automation (Trapsilawati et al., 2021). The human can check automation's work.

For on line ATP: Keep it simple. Don't increase the workload and visual distraction of the user.

Mumaw

Vertical Behavior

The vertical/altitude targets are shown on the bottom half of the ADW display.

There are always 2 altitude targets, one from the MCP and one from the FMC (in some cases, they will be the same value).

In this example, the FMC altitude target is an altitude restriction at MOLEN; it is a “cross at” for 3500.
The MCP altitude has been dialed up to FL250, which is the current cleared altitude.

The FMC altitude is what is determining autoflight behavior at this time and the current target is 3500, which is green and boxed.
The airplane will climb to FL250 next.

Note that Top of Climb (T/C) and Top of Descent (T/D) points will be shown as a circle at the transition point between flight phases

Vertical situation display

R.J. Mumaw

3. Training:

- Off line transparency (discussed above) offered before automation use.
- First failure exposure during training. “OOTLUF inoculation” (Bahner Manzey et al).

Then the failure, if it occurs during use, won’t be the first failure.

Learning from experience is more robust than learning from description (Hartwig & Erev)

Conclusion.

- Automation, Artificial Intelligence, Machine Learning will only increase.
- Machine Learning is particularly “opaque” (non transparent)
- Automation will fail, in high risk circumstances
- Make the HAI team resilient via ATP
- Resist the temptation to implement very high DOA in safety-critical systems.

Thanks for your attention. Pandawickens94@aol.com