Humans and Risk: Insights from Fossil Fuels, Nuclear, Medical, Aerospace, and AI

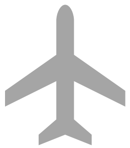
S. Camille Peres, Ph.D., C.H.F.P

Nuclear Regulatory Commission

Disclaimer

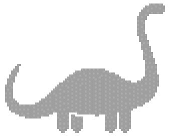
The contents of this presentation are the views of the individual presenter and do not necessarily represent those of the Nuclear Regulatory Commission.

Human Factors & Safety



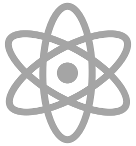
Aviation

Space



Fossil Fuels

Medical



Nuclear

Alphonse Chapanis – WWII B17s

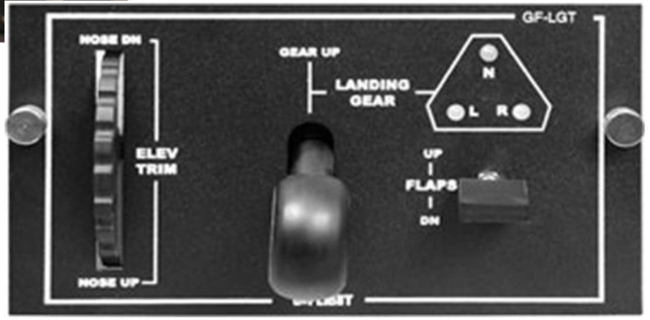
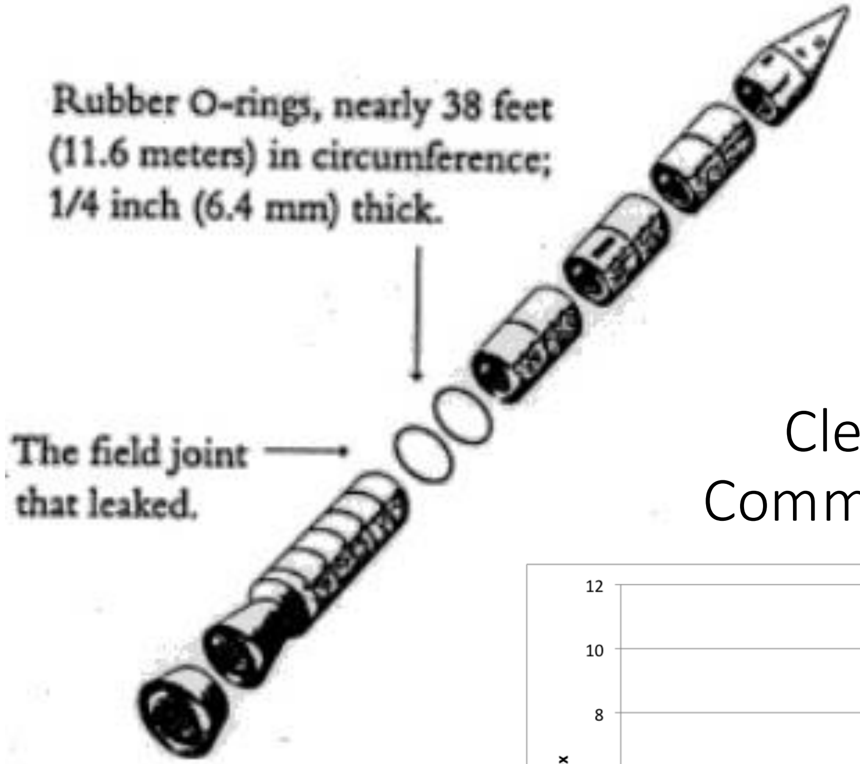


Image source: https://documentaryheaven.com/crash-of-flight-111/

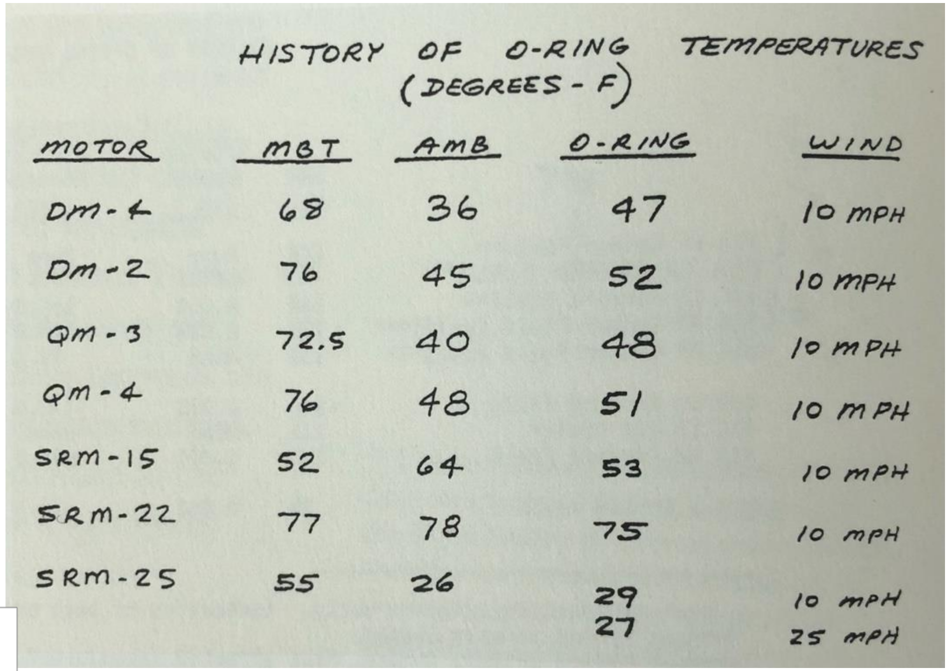
Swissair MD-11
New York to Geneva 1988

- Co-pilot suggested ignoring checklist and landing swiftly in the minutes before crash
- Pilot wanted to follow standard procedure
  Wall Street Journal reported

Challenger Disaster



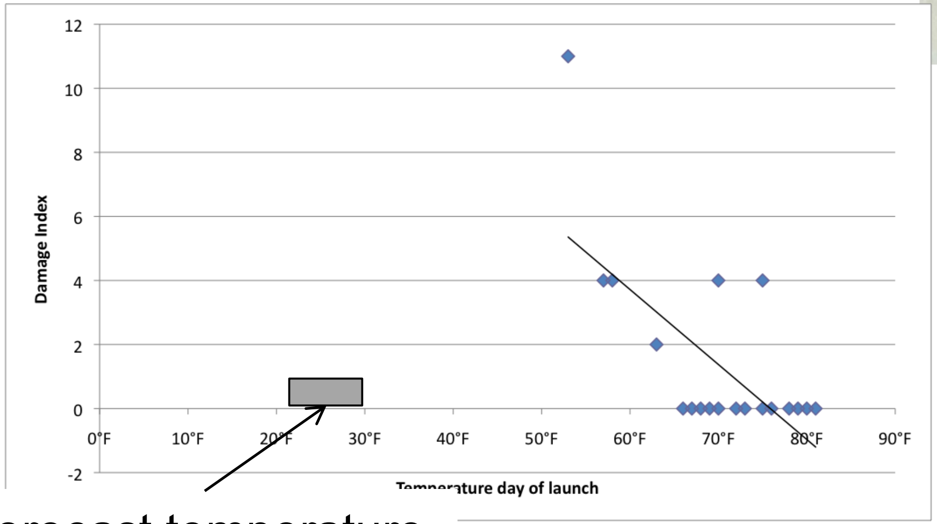
Engineers' Communication of Problem



HISTORY OF O-RING TEMPERATURES
(DEGREES - F)

MOTOR	MBT	AMB	O-RING	WIND
DM-4	68	36	47	10 MPH
DM-2	76	45	52	10 MPH
QM-3	72.5	40	48	10 MPH
QM-4	76	48	51	10 MPH
SRM-15	52	64	53	10 MPH
SRM-22	77	78	75	10 MPH
SRM-25	55	26	29 27	10 MPH 25 MPH

Clearer Method of Communicating Problem



Forecast temperature for launch

Derived from:

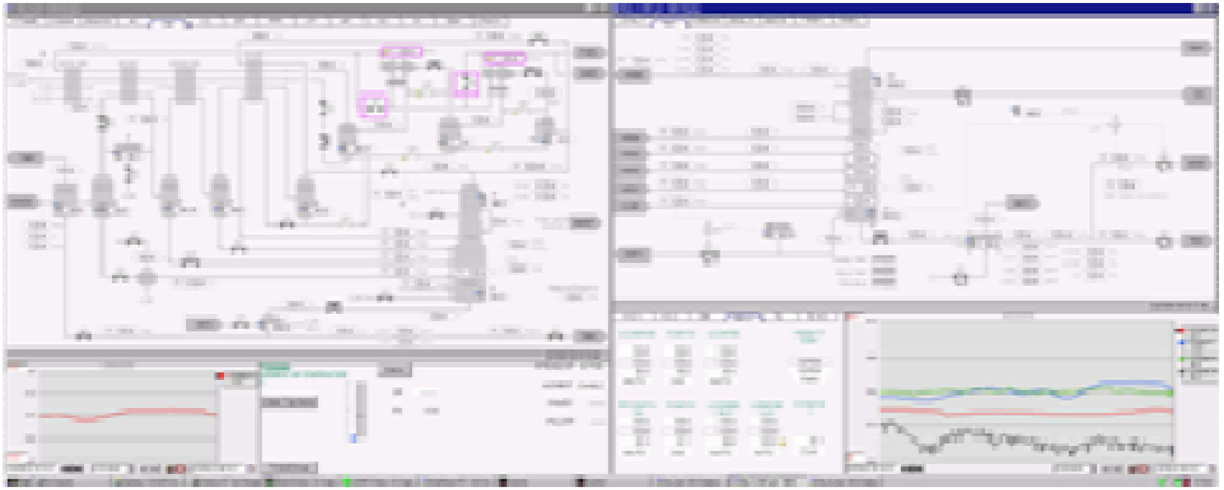
Tufte, E. (2001). *Visual Display of Quantitative Information* (Vol. 2). Cheshire, Conn: Graphics Press.

Zero Gravity Behaviors in ISS

- Best optical window ever installed in any human space vehicle
  - Astronauts want to use it
  - In zero gravity need handle to stay by window
  - There was something by window that had affordance of a handle
- It was a pressure hose
  - Crew members used it to stabilize themselves while taking pictures out of the window
  - This bent the pressure hose and was cause of air leak

https://www.nbcnews.com/id/wbna3969567

Control operations in Chemical Plant



Interface Designed Using Human Factors Methods

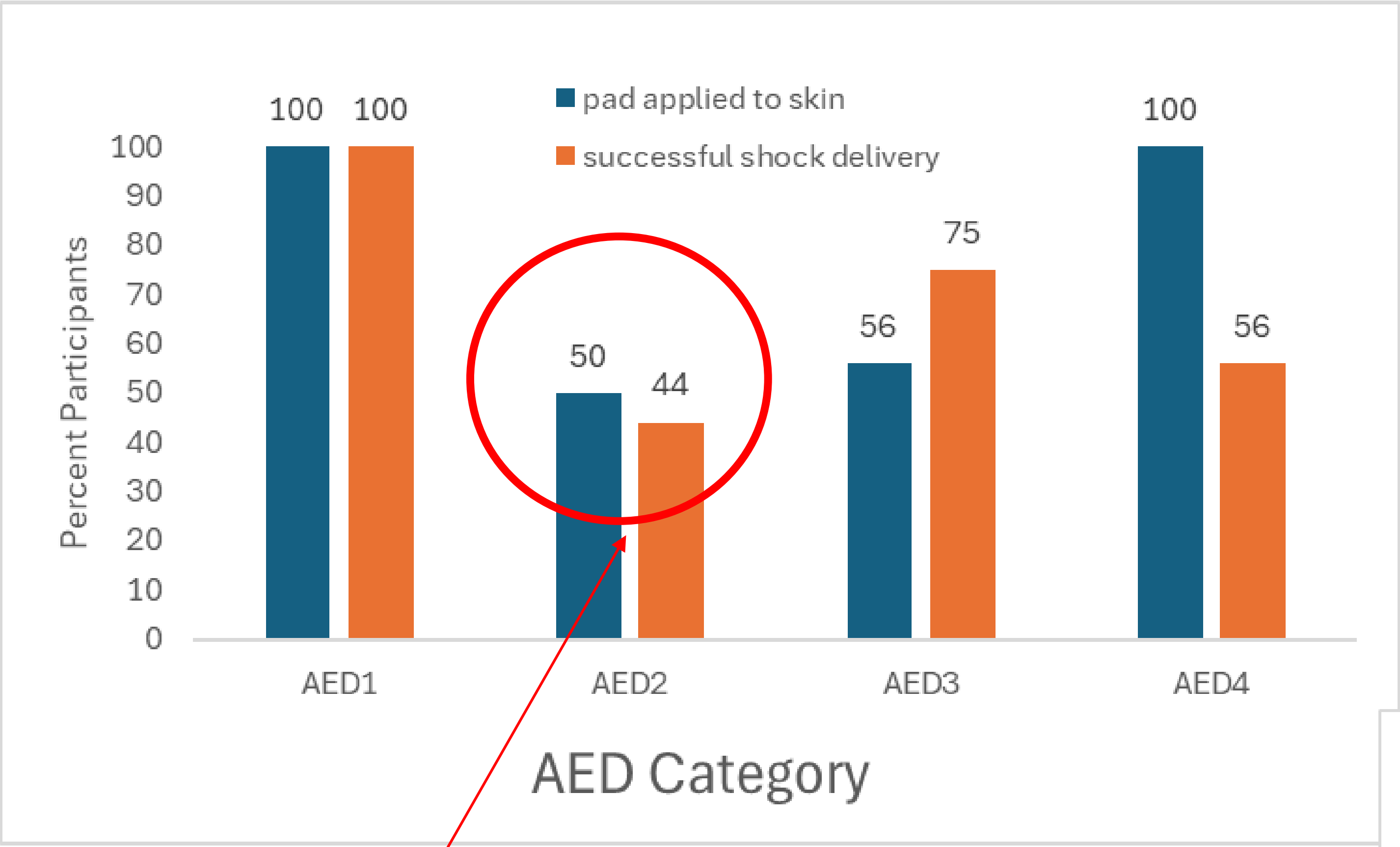
Traditional Interface

With different interface design, workers

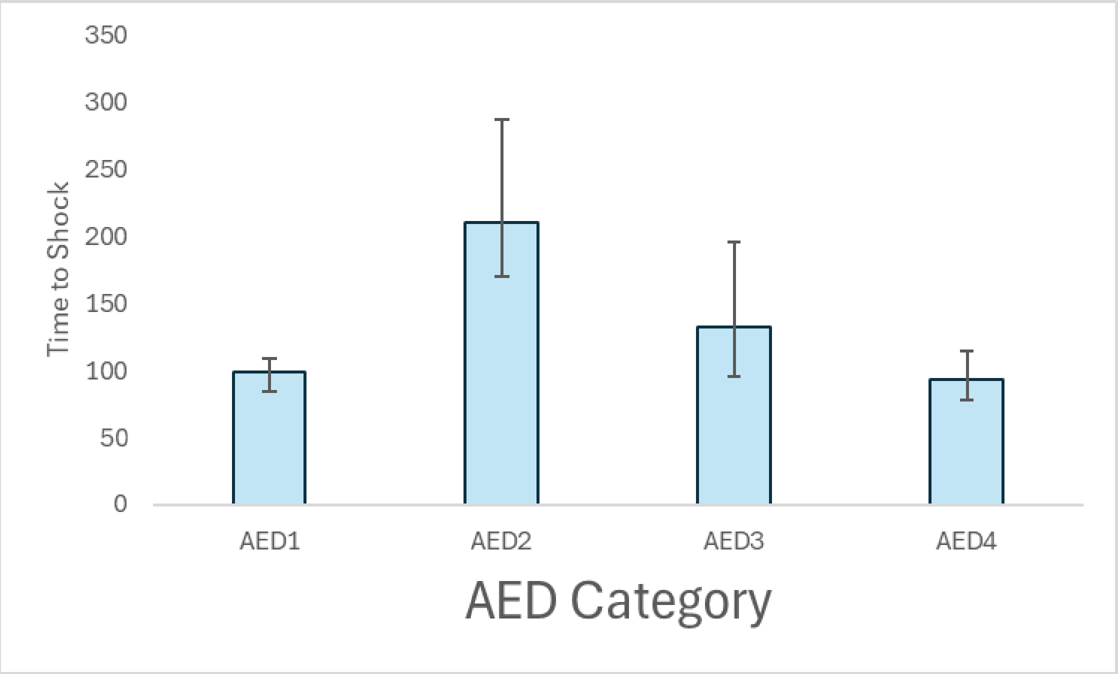
- completed scenarios 7.5 min faster (41% improvement)
- successfully dealt with failures in 96% of the cases (26% improvement)
- recognized failure before 1st process alarm in 48% of cases (38% improvement)

Economic benefit estimated at $1,090,000 CAD per year

Errington, Reising, Bullemer, DeMaere, Coppard, Doe, & Bloom, 2005



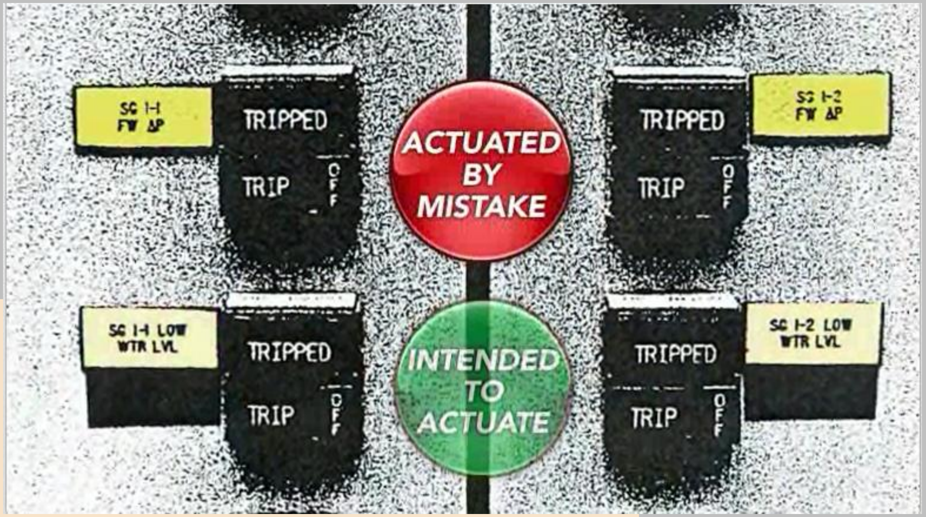
Medical



Completed these despite clearly inferior interaction design

Davis Bessie
Loss of Feedwater

June 9, 1985

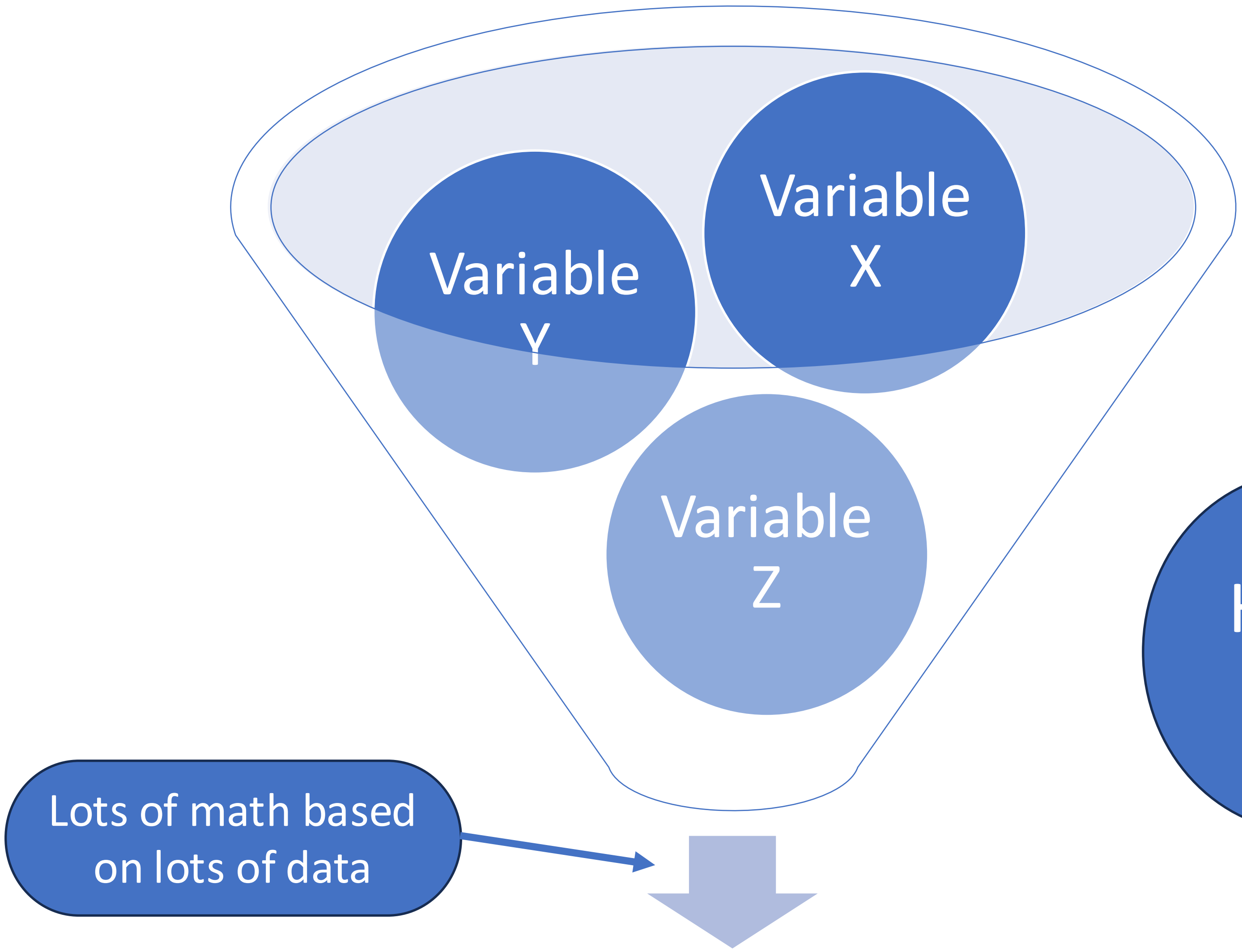


Initiating Event	Main Feed pump #1 tripped because of capacitor failure Isolation valve closed and then tripped on high pressure
In Control Room – *exacerbation of event*	Unclear layout of auxiliary feedwater displays & controls This caused isolation of the auxiliary feed water
In Secondary – *resolution of event*	Communication method for operator not close to turbine Lack of training regarding system operation No display of the turbine's RPM

Operator prevented escalation DESPITE these challenges

https://nuclepedia.usalearning.gov/index.php/Davis-Besse_Loss_of_Feedwater_Event_-_June_9,_1985

Predicting Things….
to Improve Safety



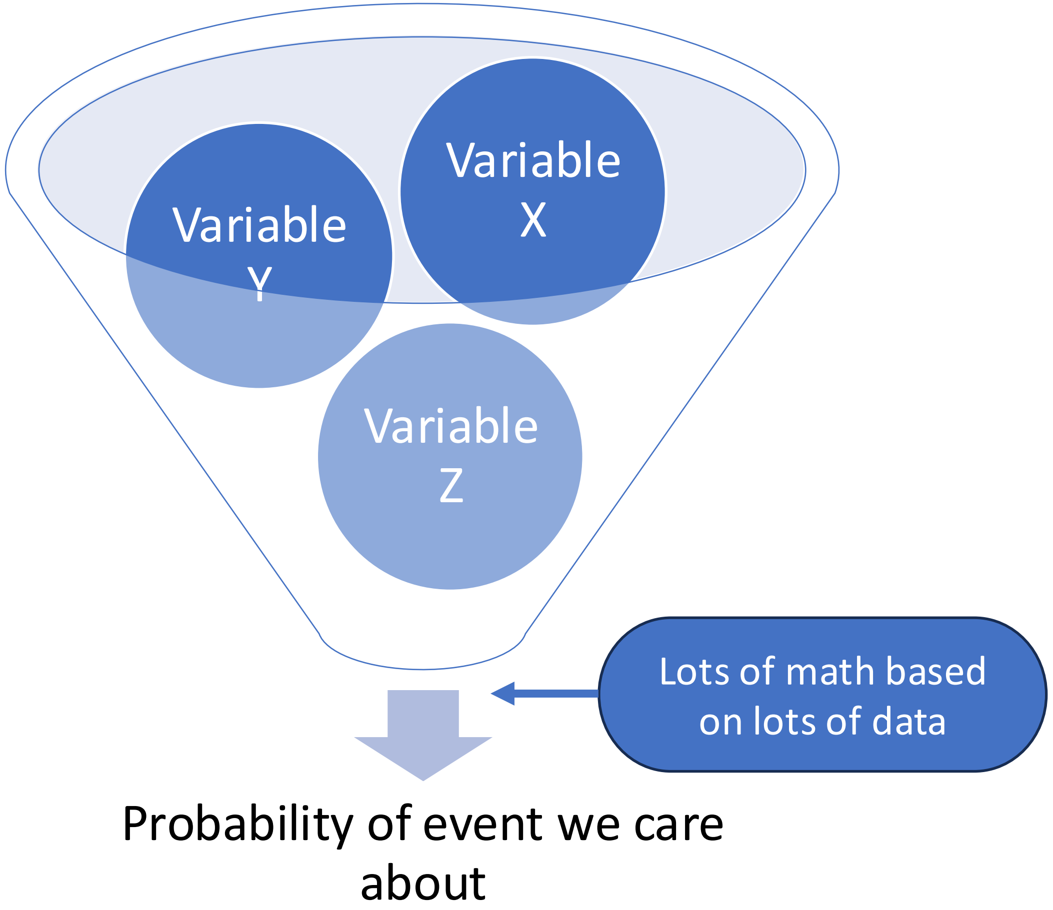
Probability of event we care about

Higher *quality and quantity* data improves predictions.

Questions I think about Regarding Humans & Safety

Are we

- including the appropriate predictors?
- trying to predict the right things?
- effectively integrating the different approaches?
- leveraging computational advances for “the math”?



Are we including the appropriate predictors?

- There are many HRA models and collectively they include performance shaping factors (PSF) from all aspects of human experience
  - Organizational

We are mostly including appropriate predictors and even interactions between them, so YES!!

  - Machine Design
  - Task
- Many investigations on interactions between the PSFs.

{still debate about accuracy of the models though}

Good Practices PSFs (NUREG-1792)
Available Staffing
Human-System Interface
Environment
Accessibility/Operability of Equipment
Need for Special Tools
Communications
Special Fitness Needs
Consideration of "Realistic" Accident Sequence Diversions and Deviations

Are we trying to predict the right things?

- Human Errors (HRA)
  - Slips, trips, lapses, etc.
- Consequential event (PRA)
  - Nuclear:
    - Significant core damage frequency (CDF);
    - Large early release of fission products from containment (LERF)
    - Conditional containment failure probability

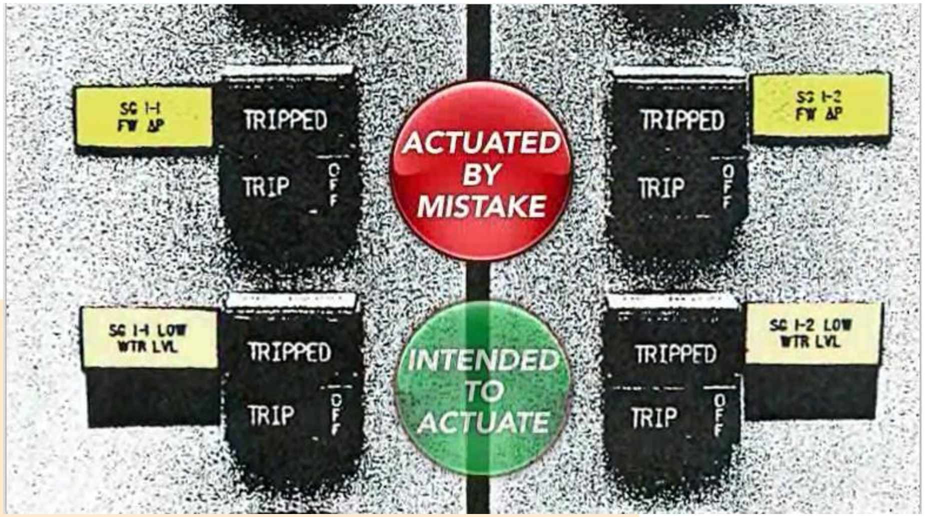
Basically, BAD THINGS

But what about predicting when humans "save the day"?

And what can increase the likelihood of this…

Davis Bessie
Loss of Feedwater

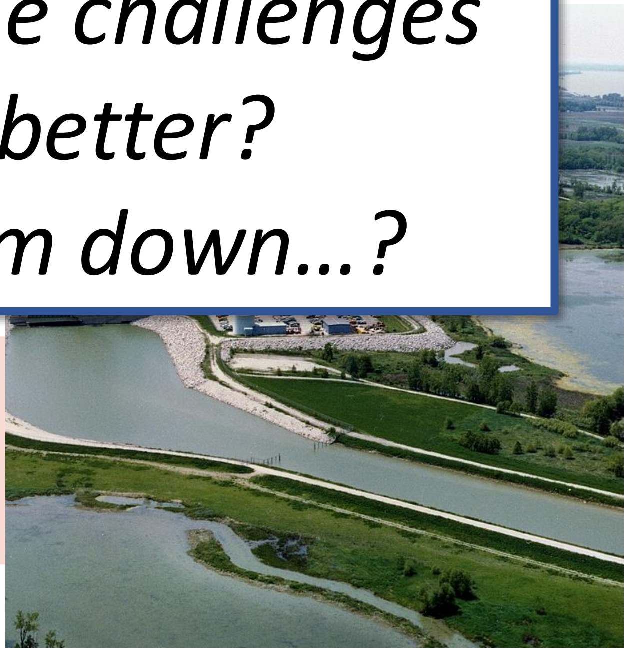
June 9, 1985



Initiating Event	Main Feed pump #1 tripped because of capacitor failure Isolation valve closed and then tripped on high pressure

In C
exac

Operator prevented escalation DESPITE the challenges
What if interface had been designed better?
Faster resolution, fewer days with system down…?



In Secondary – *resolution of event*	Communication method for operator not close to turbine Lack of training regarding system operation No display of the turbine's RPM

https://nuclepedia.usalearning.gov/index.php/Davis-Besse_Loss_of_Feedwater_Event_-_June_9,_1985

US Airways Flight 1549
New York City

Qantas Flight 72, Perth

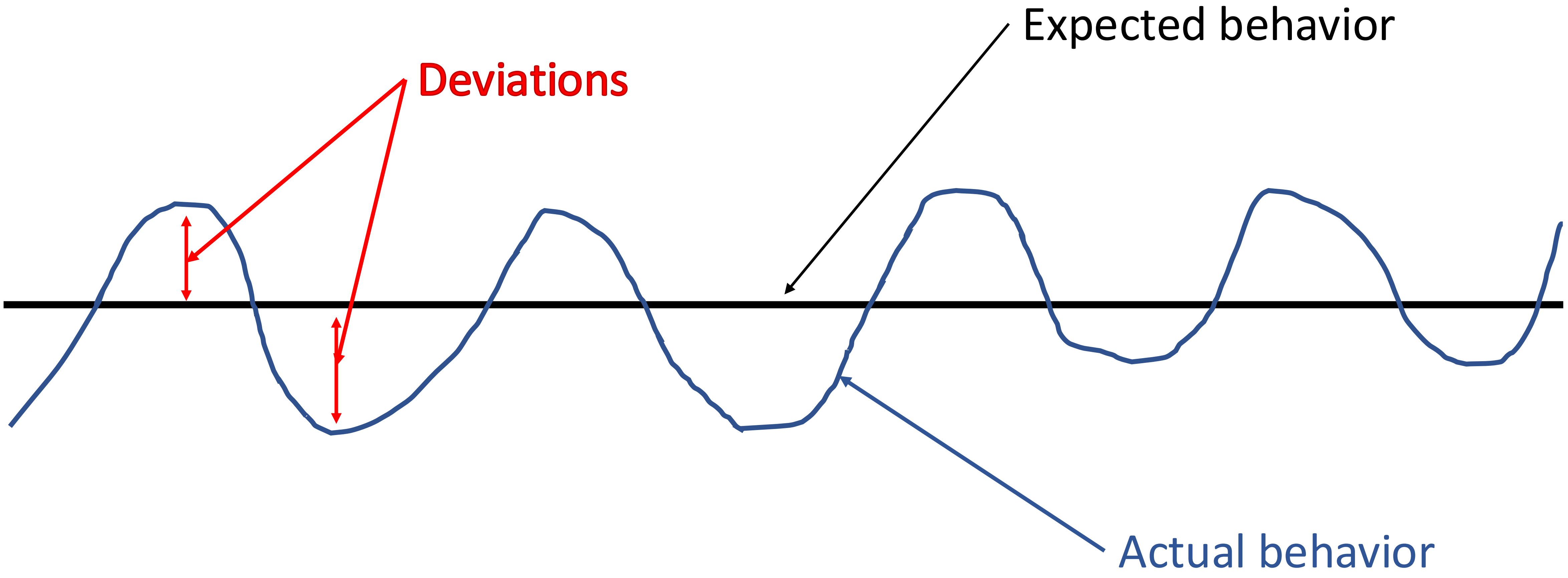
Qantas Flight 72, Perth

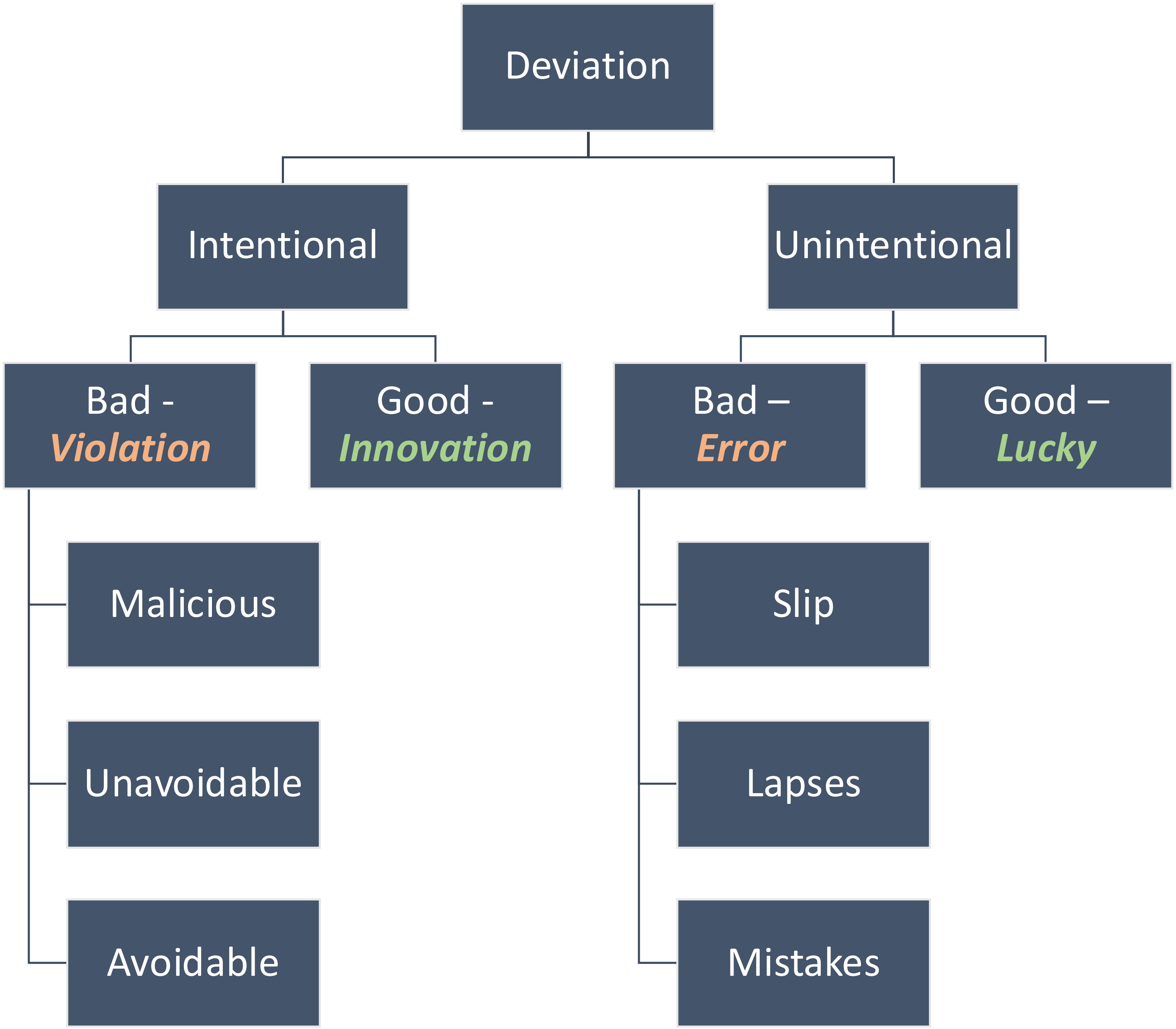
Qantas Flight 32, Changi, Singapore

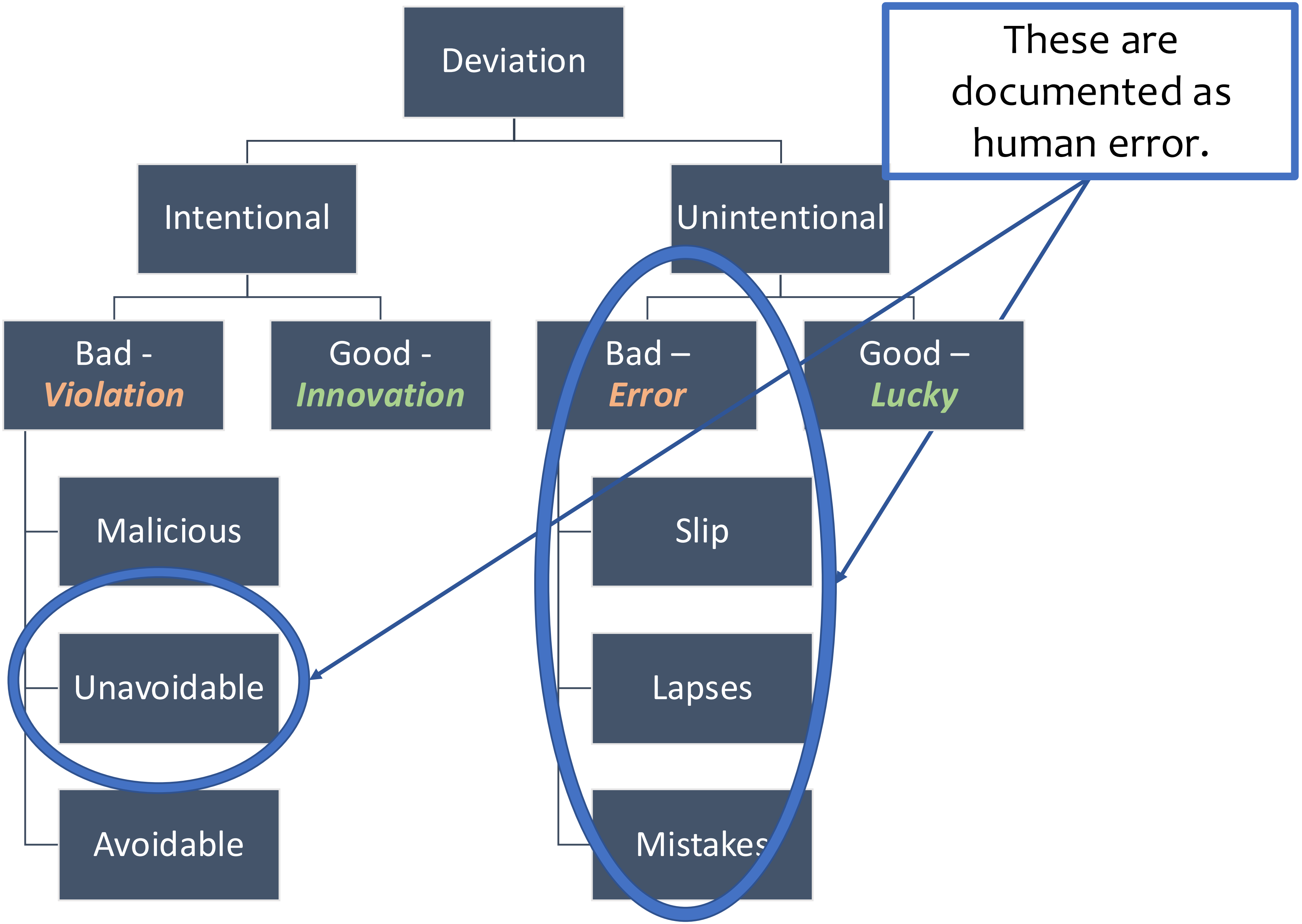
British Airways Flight 38, London

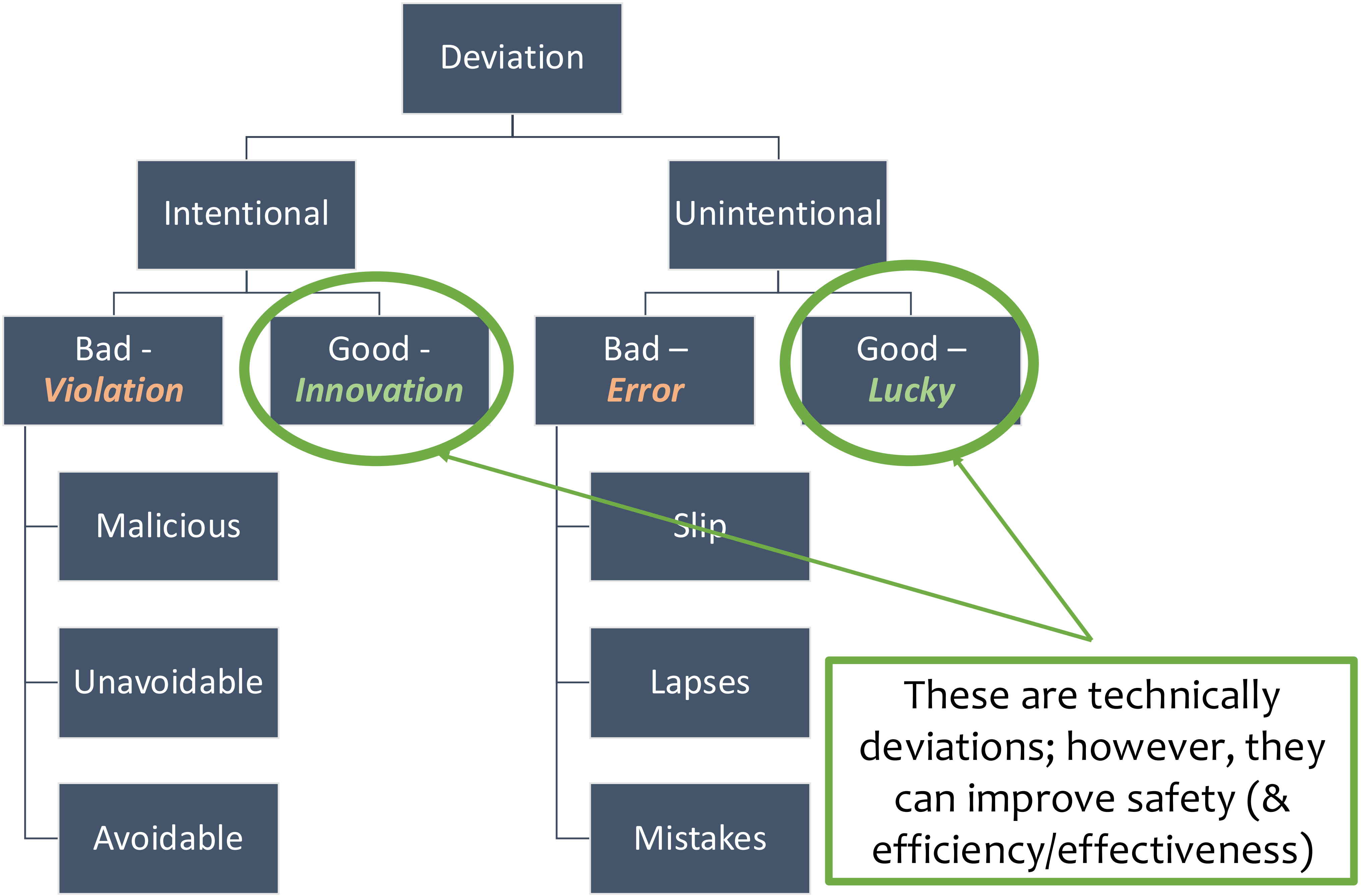
Lufthansa Flight 1829, Munich

Deviation




Deviation
Intentional
Unintentional
Bad - Violation
Good - Innovation
Bad – Error
Good – Lucky
Malicious
Unavoidable
Avoidable
Slip
Lapses
Mistakes


Deviation
Intentional
Unintentional
Bad - Violation
Good - Innovation
Bad – Error
Good – Lucky
Malicious
Unavoidable
Avoidable
Slip
Lapses
Mistakes
These are documented as human error.



Deviation
Intentional
Unintentional
Bad - Violation
Good - Innovation
Bad – Error
Good – Lucky
Malicious
Unavoidable
Avoidable
Slip
Lapses
Mistakes
These are technically deviations; however, they can improve safety (& efficiency/effectiveness)

Focus now is on behavior “below the line”

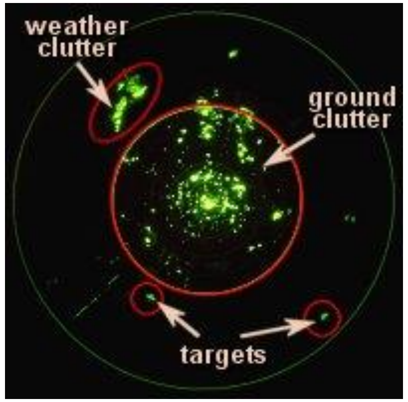
What are we losing by not integrating the innovations and needed support for the “above the line” behaviors?

I am still wondering if we are focusing on all the right things…

Sources of predictors

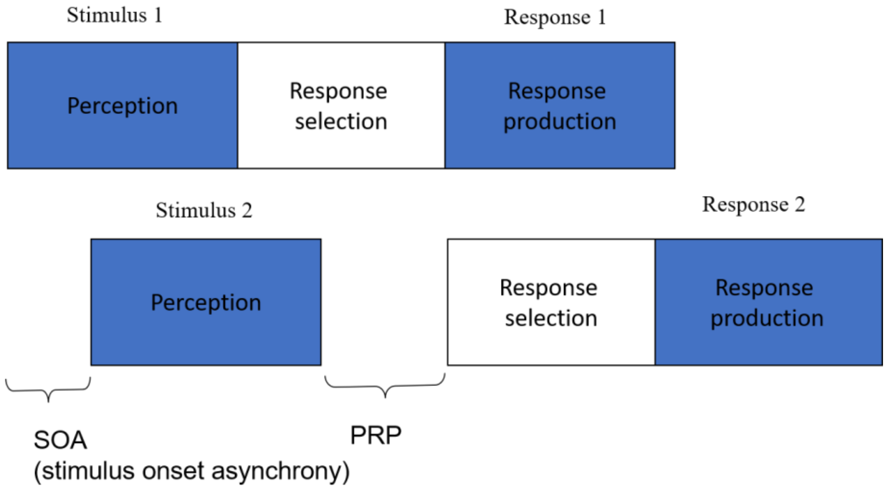
Cognitive Examples

Signal Detection

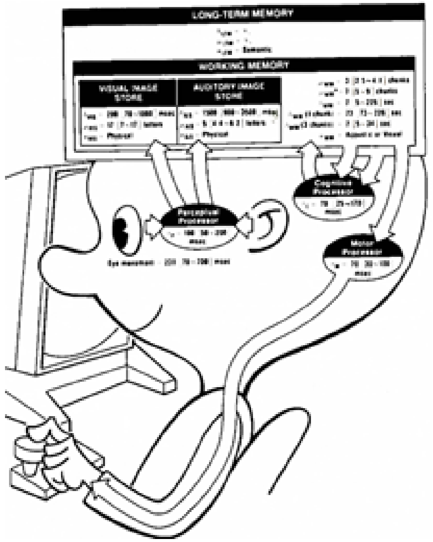


USS Vincennes

Psychological Refractory Period



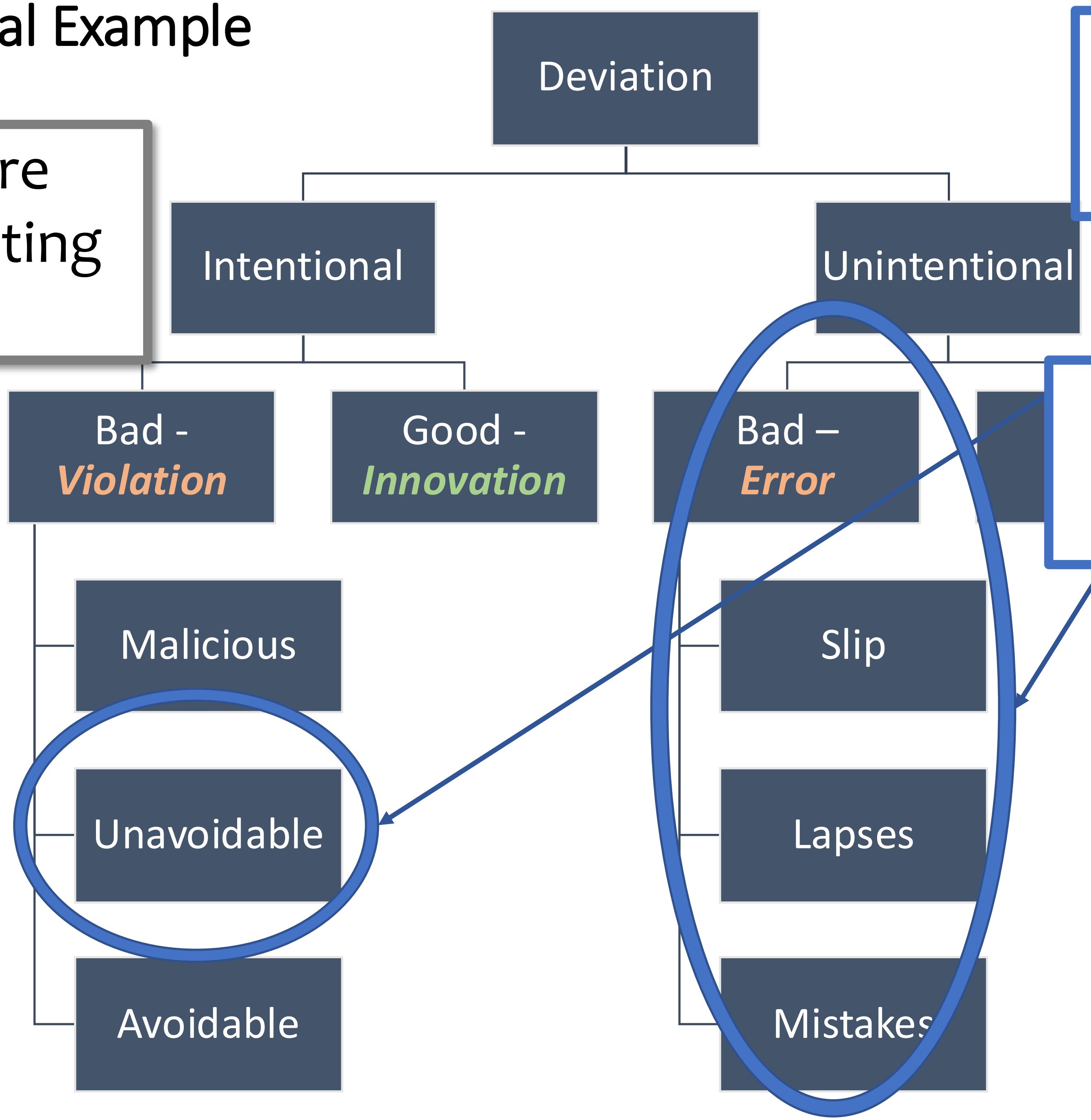
Computational Cognitive Modeling



Psychologists are really good at this…

Psychosocial Example

Psychologists are good at predicting this as well.



If a negative consequence is presented for these

Creates a climate where workers avoid reporting issues.

I’m sure there are others, but let’s start with these--
Human Reliability Analysis and
Computational Cognitive Modeling

<u>Are we integrating sources of predictors?</u>

Types of modeling systems

Human reliability analysis

- Part of probabilistic risk assessment (PRA) that includes hardware and human reliability
- “A structured approach used to **identify potential human failure events** and to systematically estimate the probability of those events using data, models, or expert judgment” (ASME RA-Sb-2013)

Computational cognitive modeling

- Computational cognitive modeling to **simulate how the human mind works and user responses to inputs.**
- Used to test psychological theories as well as **predict user actions**. [1]

[1] Trafton, J. G., Harrison, A. M., & Hiatt, L. M. (2025). *Using cognitive models to generate data for machine learning models* (Report No. AD1334810). Naval Research Laboratory. https://www.dmi-ida.org/knowledge-base-detail/Using-Cognitive-Models-to-Generate-Data

Thoughts about integration

Creating new designs

- Cognitive models can support the creation of designs that reduce distractions[1]
- This could inform the likelihood of human error

Simulations

- Cognitive model can be used to create simulations of scenarios with different types of reactors, interfaces, staffing models

Generation of data to inform human error rates

- Cognitive models can generate synthetic data of human behavior

[1]Trafton, J. G., Harrison, A. M., & Hiatt, L. M. (2025). *Using cognitive models to generate data for machine learning models* (Report No. AD1334810). Naval Research Laboratory. https://www.dmi-ida.org/knowledge-base-detail/Using-Cognitive-Models-to-Generate-Data

Are we integrating sources of predictors?

- Short answer – No, we are not.
- Questions now
  - Should we? And if so, how?
- Answer is for better minds than mine.
  - But I do have thoughts about it.
- Start with the question:
  - Are we leveraging computational advances for "the math"?

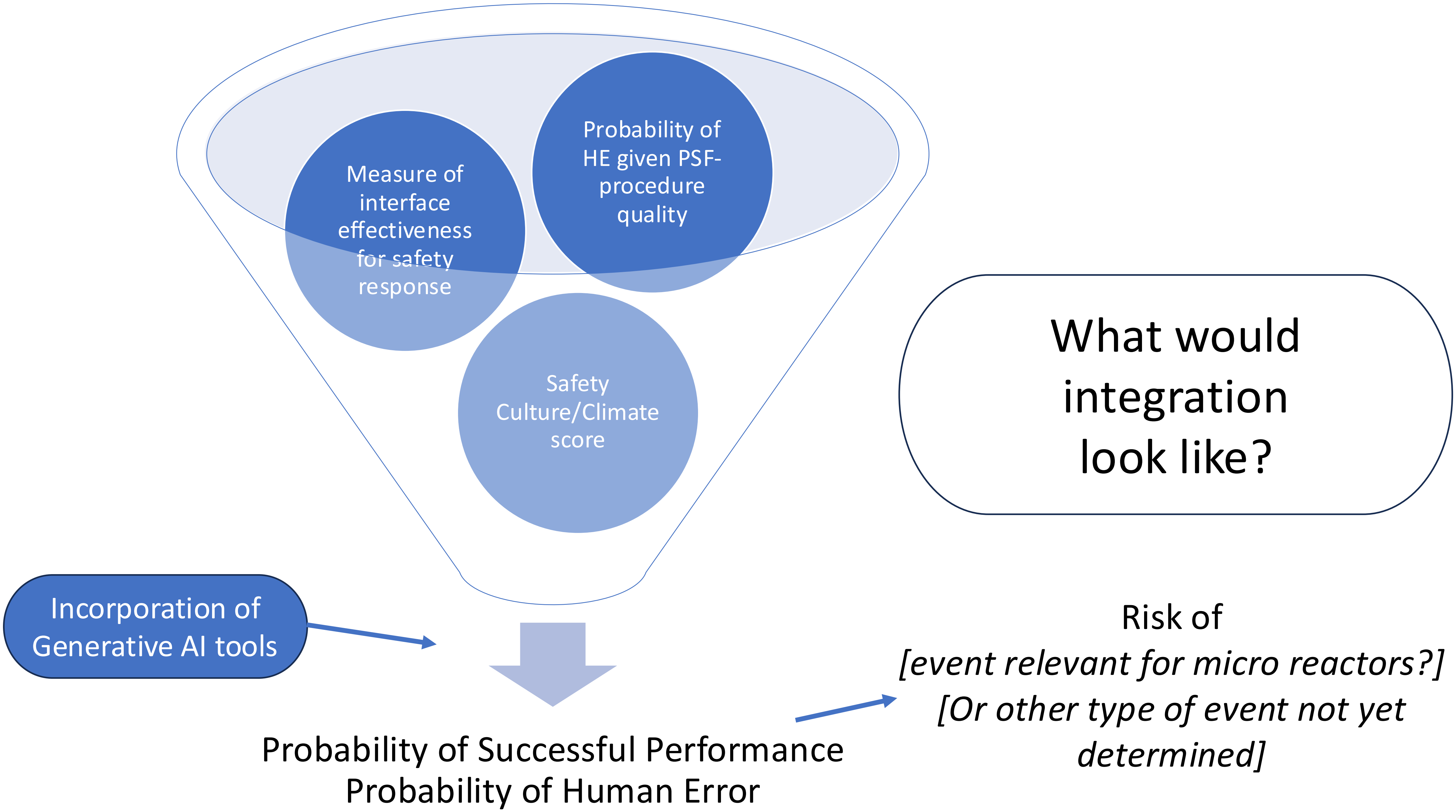
Are we leveraging computational advances for "the math"?

Types of AI

- Generative Adversarial Networks (GANS)
  - Neural network with a generator and discriminator
- Variation Autoencoders (VAES)
  - Encodes input data into compressed format, decodes it into new output with a new variation
- Large Language Models

Can we or are we leveraging these to more easily:

- Isolate important variables
- Simulate scenarios that allows for answering "what if" questions
- Simplify variable using cognitive model's methods
- Quantify the role of each variable in a model
- Integrate alternative sources of predictors into existing PRA models


Measure of interface effectiveness for safety response
Probability of HE given PSF-procedure quality
Safety Culture/Climate score
What would integration look like?
Incorporation of Generative AI tools
Probability of Successful Performance
Probability of Human Error
Risk of
[event relevant for micro reactors?]
[Or other type of event not yet determined]

We can’t wait for lessons learned

We need to learn how to effective leverage these new tools to get ahead of potential issues so, we can help develop new solutions

Going back to Davis Bessie

Bad design contributed to human error and exacerbated the event	Bad design impeded humans’ ability to mitigate the extent of the event	BUT they were able to do it anyway

It’s not JUST about preventing human error; it’s ALSO about supporting mitigation of the consequences

This changes the ultimate outcome of the PRA	Which may be different for reactors in the future

Let’s think about the future

Current limitation of HRA

- Original data with only one type of reactor
- Limited ability to run/create models with specific interfaces

Can generative AI help?

- Generative Adversarial Networks (GANS)
- Variations Autoencoders (VAES)
- Large Language Models (LLMs)

Can advanced models help predict more than risk?

- Productivity
- Efficiency
- Fatigue

THANKS!

camille.peres@nrc.gov

LinkedIn: s-camille-peres-phd