Aggregation of Autocalculated Human Error Probabilities from Tasks to Human Failure Events in a Dynamic Human Reliability Analysis Implementation

Ronald L. Boring, PhD[1]
Martin Rasumussen, PhD[2]
Thomas A. Ulrich, PhD[1]
Nancy J. Lybeck, PhD[1]

[1]Idaho National Laboratory
[1]NTNU Social Research

www.inl.gov

Idaho National Laboratory

Best title ever, amiright!? #nailedit

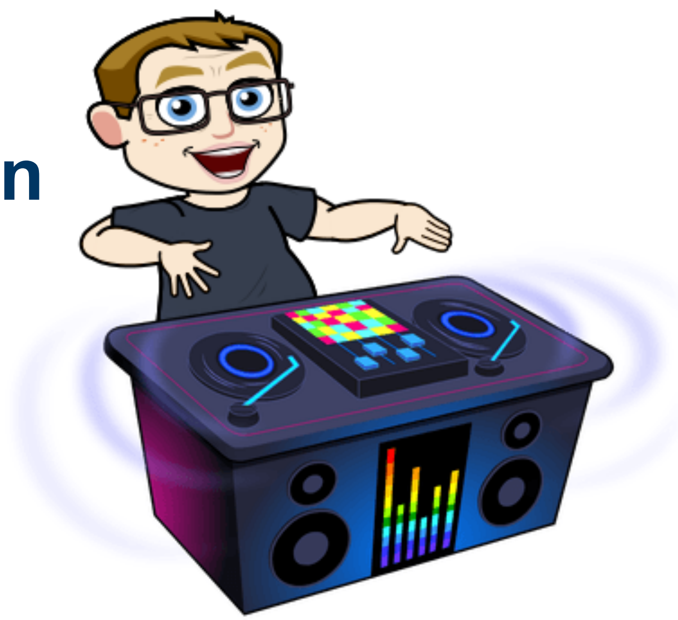
Let's break down that title…

Aggregation of
[combining]

Autocalculated Human Error Probabilities
[automatically generated HEPs]

from Tasks to Human Failure Events
[from subtasks to HFEs]

in a Dynamic HRA Implementation
[in HUNTER]



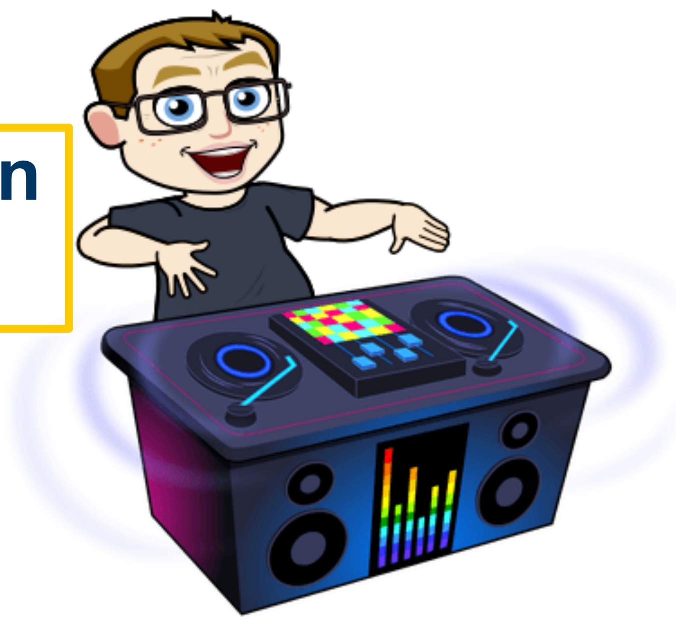
Let's break down that title…

Aggregation of
[combining]

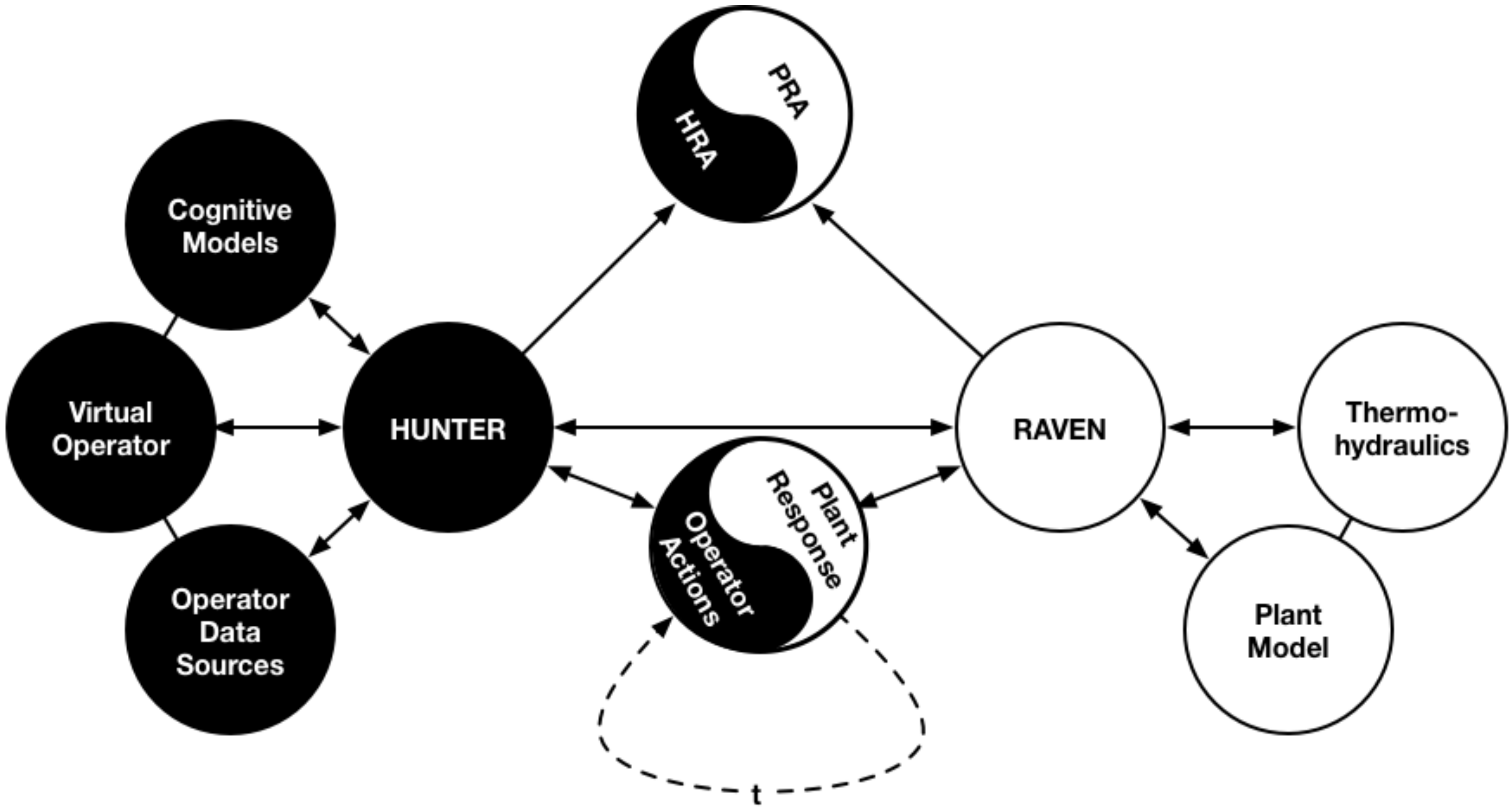
Autocalculated Human Error Probabilities
[automatically generated HEPs]

from Tasks to Human Failure Events
[from subtasks to HFEs]

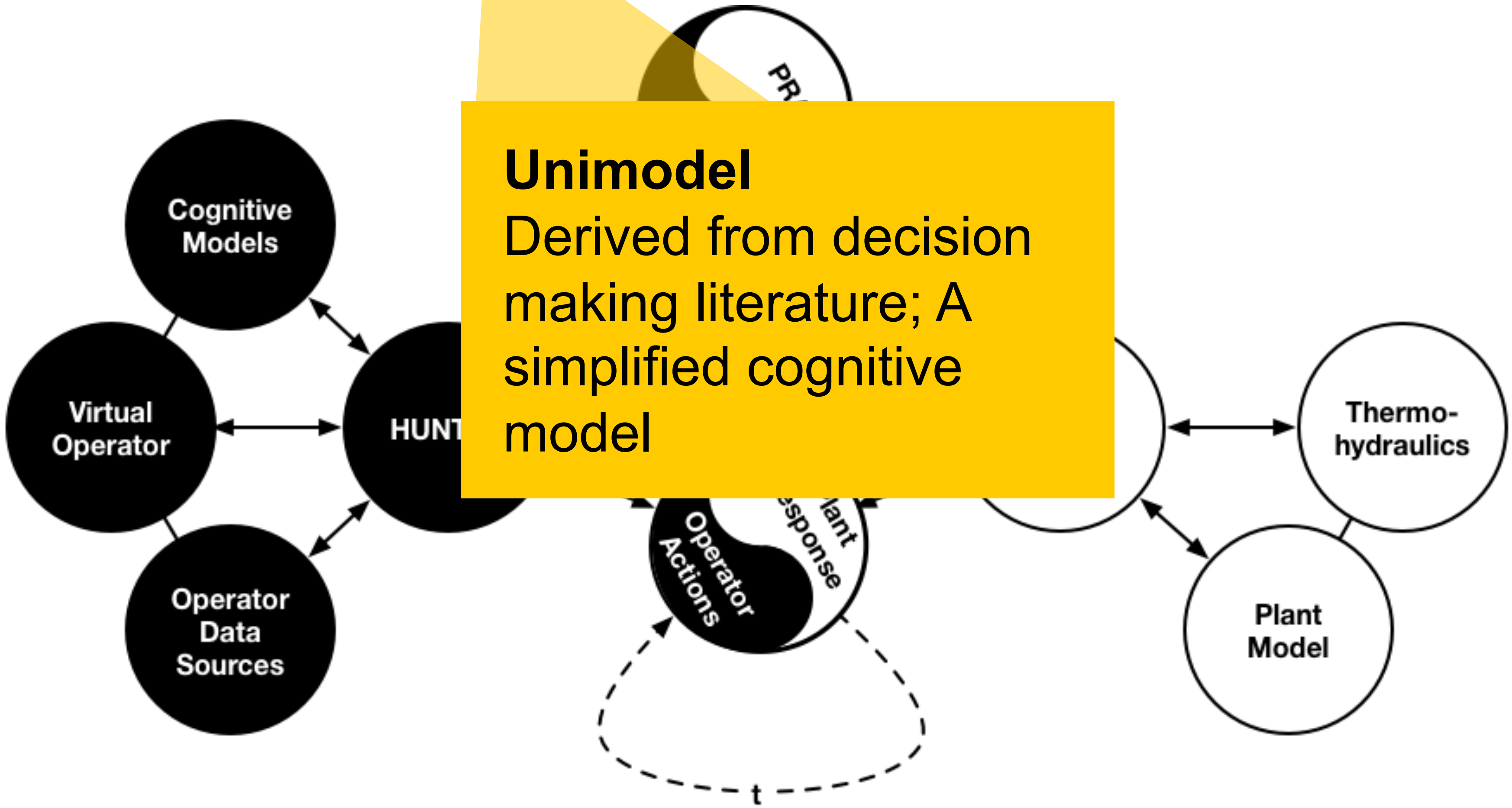
in a Dynamic HRA Implementation
[in HUNTER]



Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER)

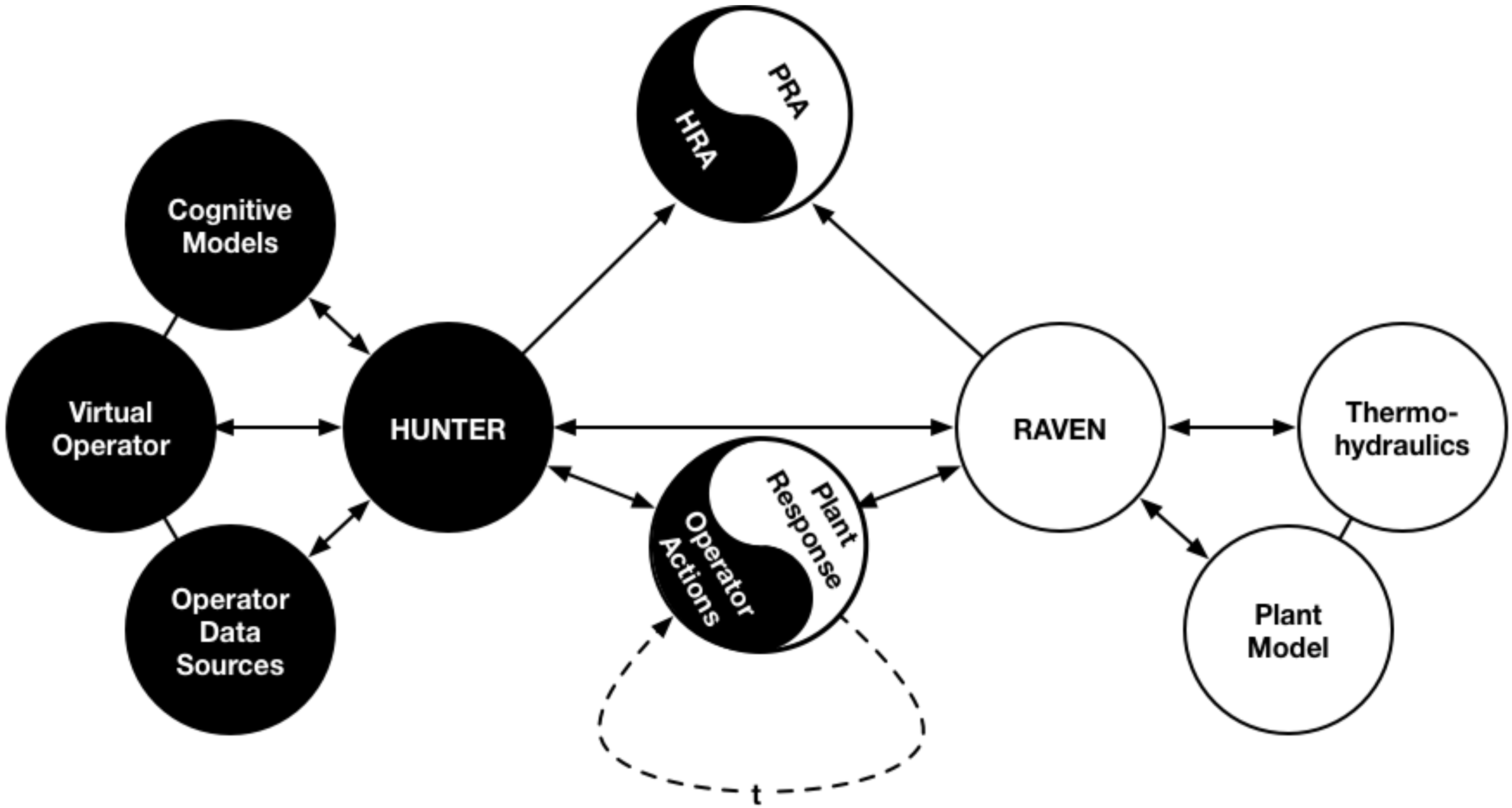


Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER)



Unimodel
Derived from decision making literature; A simplified cognitive model

Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER)



What HUNTER is Currently

Plant model

- RAVEN interface with thermo-hydraulics software (RELAP)

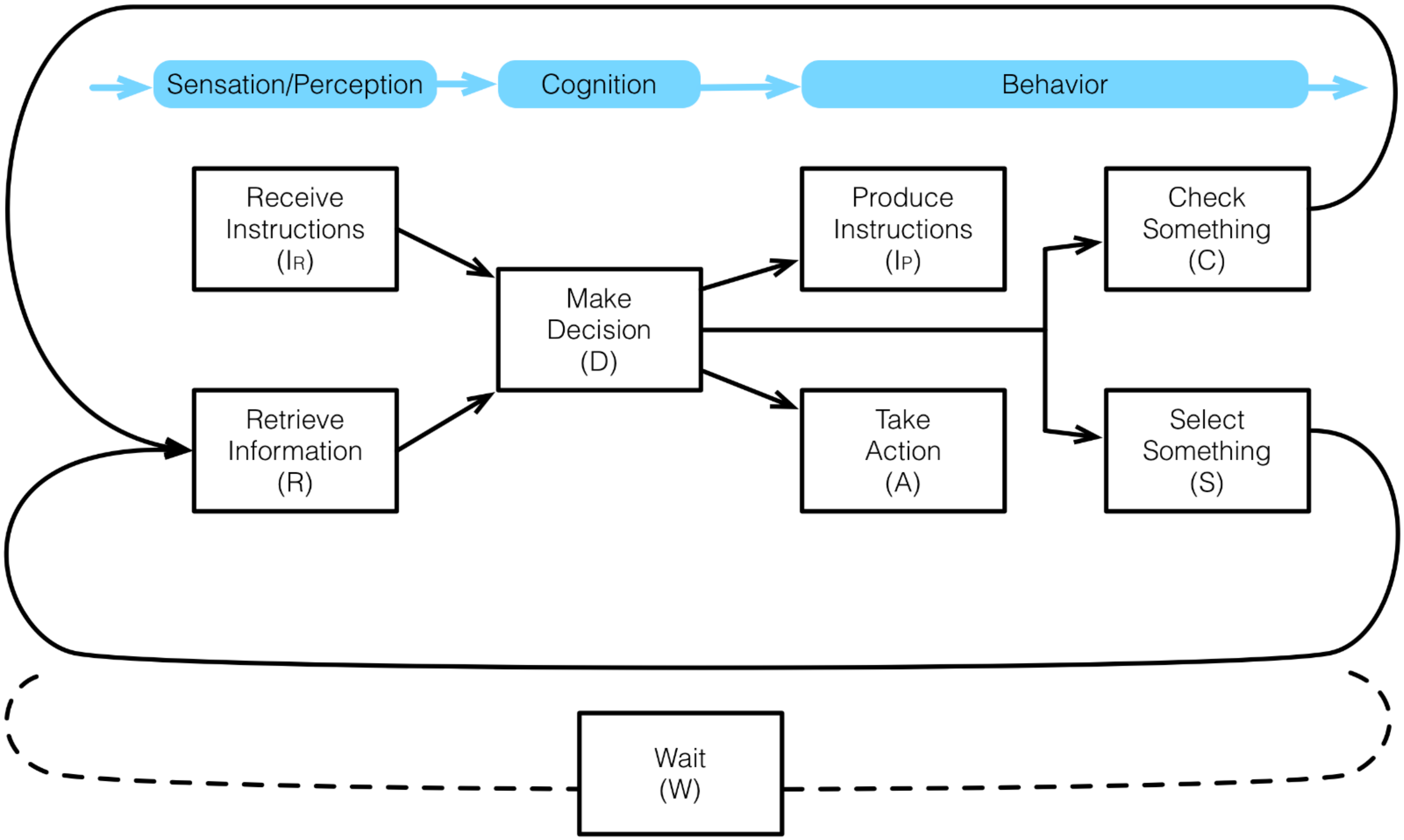
Subtask modeling system

- New approach called GOMS-HRA
- Provides mapping between procedures, error taxonomies, and task primitives needed for HRA
- Provides timing and nominal human error probability (HEP)

Auto-calculating performance shaping factor (PSF)

- SPAR-H based PSF system
- Currently auto-calculating Complexity PSF based on plant parameters
- PSF serves as multiplier to refine HEP

GOMS-HRA Cognitive Framework



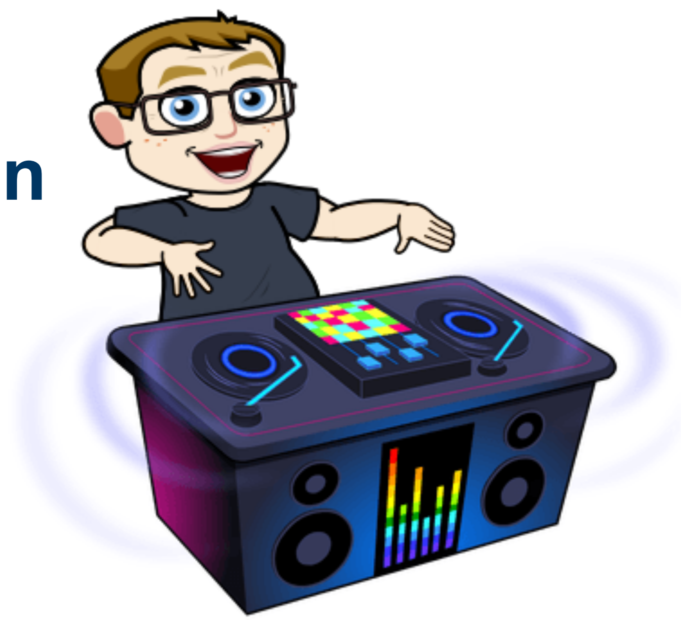
Let's break down that title…

Aggregation of
[combining]

Autocalculated Human Error Probabilities
[automatically generated HEPs]

from Tasks to Human Failure Events
[from subtasks to HFEs]

in a Dynamic HRA Implementation
[in HUNTER]



SPAR-H

- SPAR-H determines HEP based on expert estimation using calculation worksheets
- Estimation of PSFs carried out using predefined multipliers and a nominal failure probability
- $HEP = NHEP * PSF$
  - HEP is the overall human error probability
  - NHEP is the nominal human error probability (0.01 or 0.001)
  - PSF is substituted with the respective PSF level's multiplier

PSFs	PSF Levels	Multiplier for Diagnosis
Available Time	Inadequate time	P(failure) = 1.0 ☐
	Barely adequate time (≈2/3 x nominal)	10 ☐
	Nominal time	1 ☐
	Extra time (between 1 and 2 x nominal and > than 30 min)	0.1 ☐
	Expansive time (> 2 x nominal and > 30 min)	0.01 ☐
	Insufficient information	1 ☐
Stress/ Stressors	Extreme	5 ☐
	High	2 ☐
	Nominal	1 ☐
	Insufficient Information	1 ☐
Complexity	Highly complex	5 ☐
	Moderately complex	2 ☐
	Nominal	1 ☐
	Obvious diagnosis	0.1 ☐
	Insufficient Information	1 ☐
Experience/ Training	Low	10 ☐
	Nominal	1 ☐
	High	0.5 ☐

Dynamic Complexity Modeling

- **Complexity as a proof-of-concept PSF**
  - Well documented from static HRA
  - Determined to be one of the main drivers on operator performance across a number of studies (e.g., NUREG-2127)
  - Recent modifications to SPAR-H for the Norwegian Petro-HRA project give us insights on how to model and operationalize Complexity as a PSF

- In SPAR-H, the analyst subjectively assigns a level for the Complexity PSF (a multiplier on the nominal HEP)

- **In dynamic SPAR-H, the PSF multiplier is auto-calculated based on plant parameters**

- In dynamic HRA it is not possible to use a subjective evaluation for each simulation

Dynamic Complexity Modeling

- Examples of auto-populated aspects that could be used
  - Total size of the task or scenario (size complexity)
  - Number of success criteria (goal complexity)
  - Number of alternative paths to the goal(s) (goal complexity)
  - Number of steps conducted (step complexity)
  - Number of tasks per time (temporal complexity)
  - Time spent on task (temporal complexity)
  - Time in scenario (temporal complexity)

- Examples of categorization aspects that could be included
  - Amount of information the operator uses in this task (size complexity)
  - Is the task influenced by factors outside of the operators control (dynamic complexity)
  - Is the task connected to other tasks (connection complexity)
  - Number of procedures used by the operator (procedure complexity)
  - Number of operators involved (interaction complexity)

Modeling Station Blackout

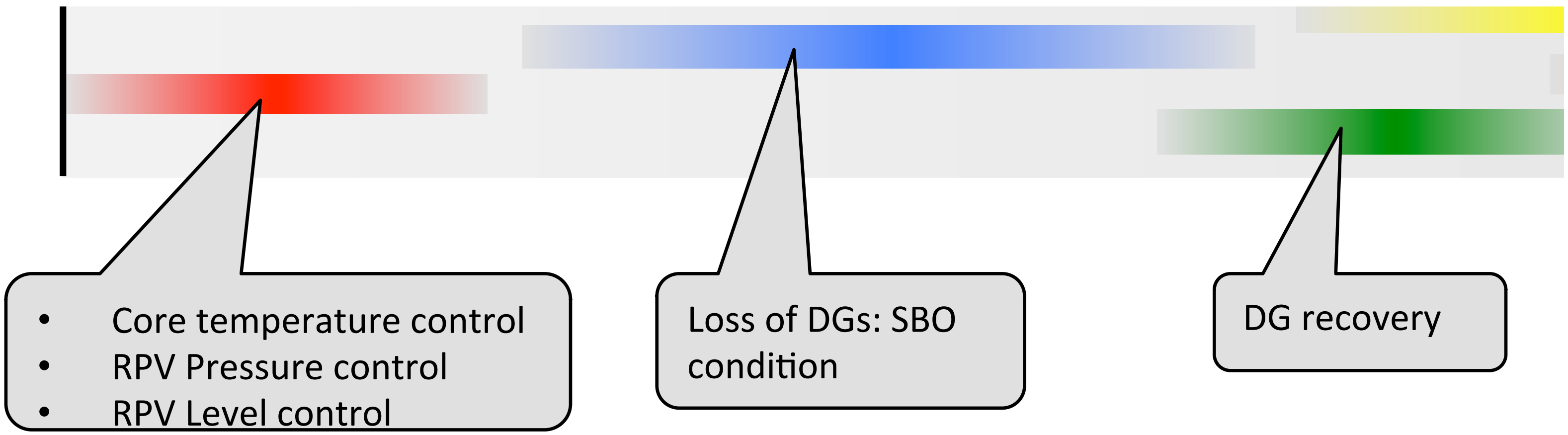


Figure 18. Sequence of events for the SBO scenario considered

Example Normalized Complexity Mapping

7	1	0	0	539.49	2.79	3
8	1	0	0	561.59	2.38	3
9	1	0	0	538.57	2.48	3
10	1	0	0	538.55	2.63	3
11	1	0	0	538.55	2.63	3
12	1	0	0	538.55	2.63	3
13	1	1	0	575.73	1.36	4
14	1	1	0	624.89	1.29	4
15	1	1	1	1775.04	0.75	5
16	1	1	1	2092.49	0.66	5
17	1	1	1	2257.35	0.60	5
18	1	1	1	2374.40	0.54	5
19	1	1	1	2407.60	0.00	5
20	1	1	1	2400.87	0.51	5

This relationship does not always hold true, because

Temporal Evolution of Complexity Multiplier

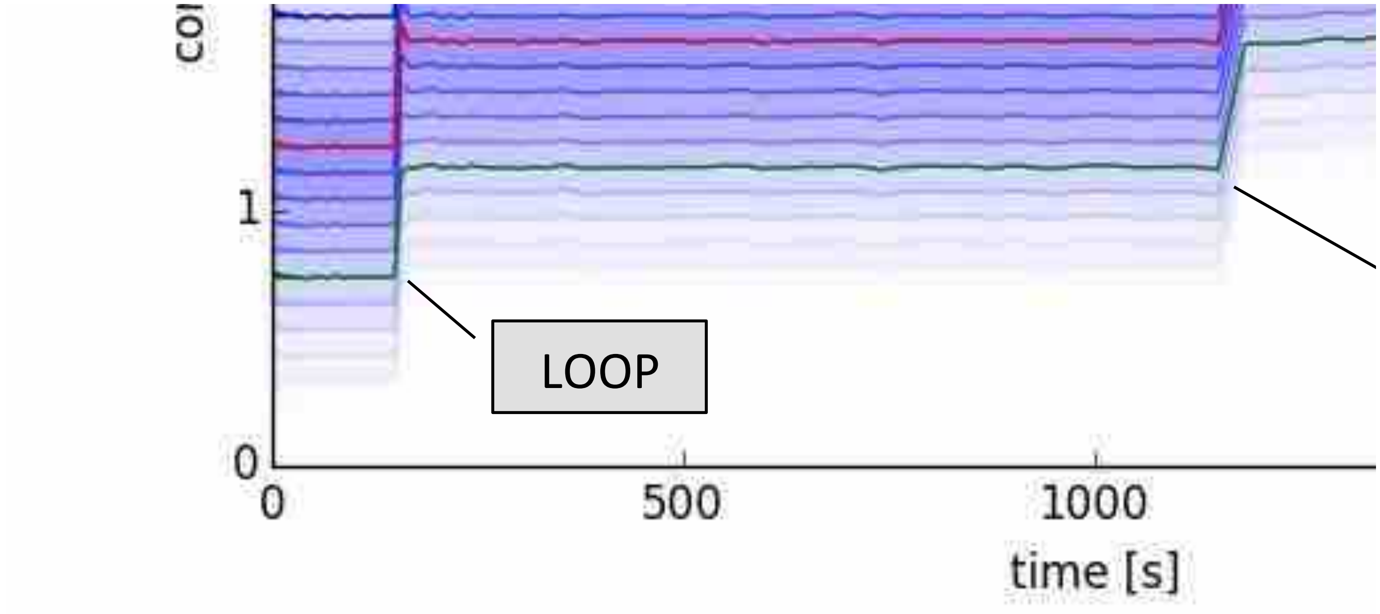


Figure 27. Temporal evolution of the complexity multi

7.8 Quantifying Operator Per

Operator performance was quantified as a final HEP value usin

Lessons Learned from Automatic PSF Calculation

Internal vs. External PSFs Matter!

- External PSFs are relatively easy to auto-calculate
  - The mapping between context and plant factors and operator performance is traceable
- Internal PSFs are not so easy to auto-calculate
  - These psychological factors must be input manually into the model

Time is a funny PSF!

- GOMS-HRA produces time estimates, which can be treated as time-reliability method rather than PSF

Aggregation of subtask HEPs must still be solved

- We have thousands of HEPs generated, but how do we combine or average them?

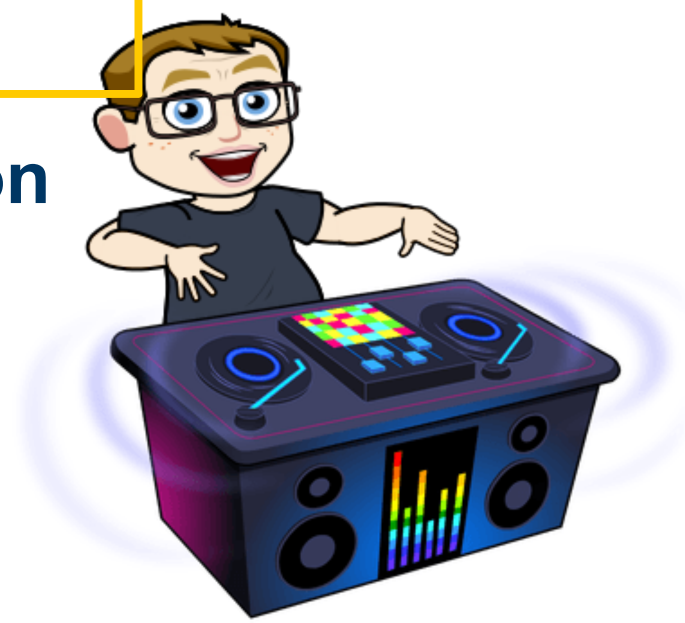
Let’s break down that title…

Aggregation of
[combining]

Autocalculated Human Error Probabilities
[automatically generated HEPs]

from Tasks to Human Failure Events
[from subtasks to HFEs]

in a Dynamic HRA Implementation
[in HUNTER]



What is an HFE?

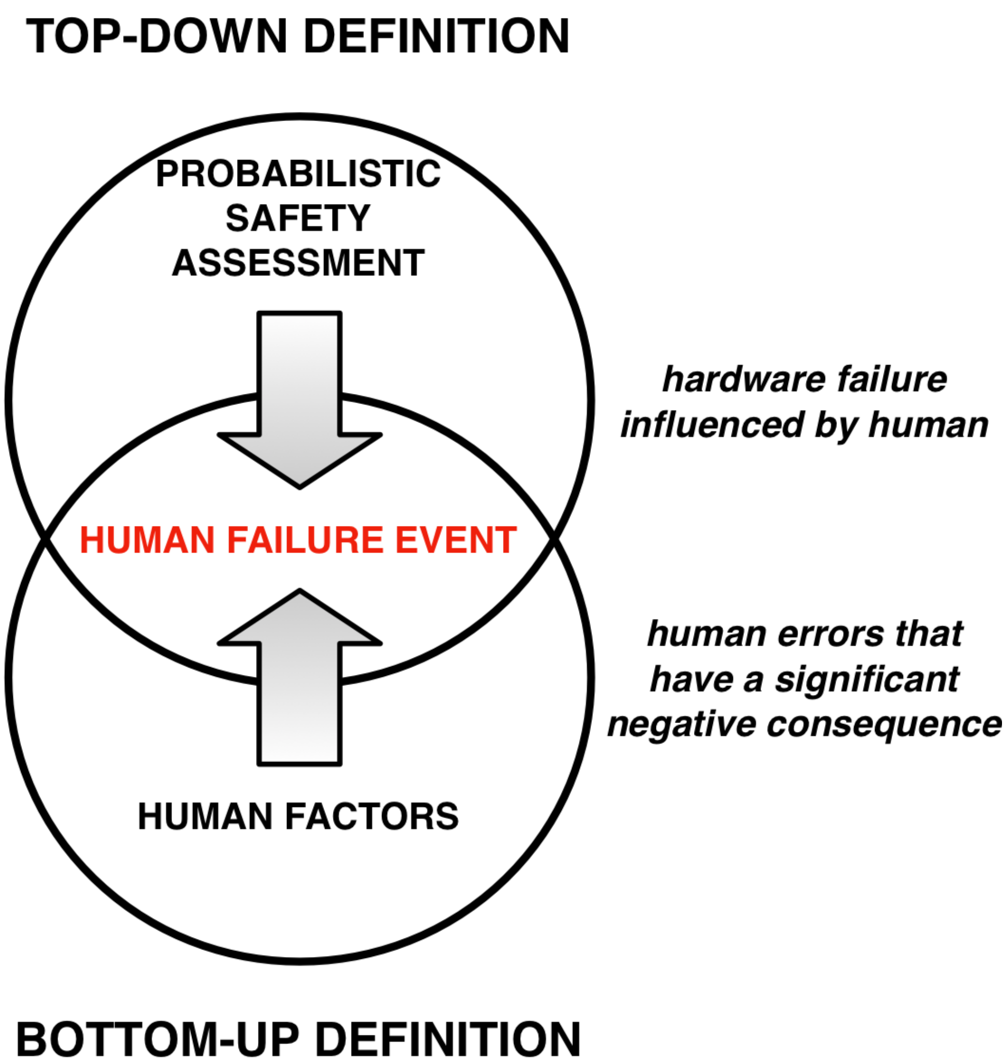
ASME defines a Human Failure Event (HFE):

- Basic event that represents a failure or unavailability of a component, system, or function that is caused by human inaction, or an inappropriate action

While this definition is helpful, it doesn't articulate:

- What tasks constitute a typical HFE
- What is the boundary between HFEs
  - Is it based on the system affected?
  - Is it based on the goals or tasks being performed by operators?
  - Are these the same?

Two Approaches to Defining HFEs/Task Decomposition



- **Ideally, these two approaches meet in the middle and come up with the same set of HFEs.**
- **But not always!**

Why Does Task Decomposition Matter?

Task decomposition shapes the analysis

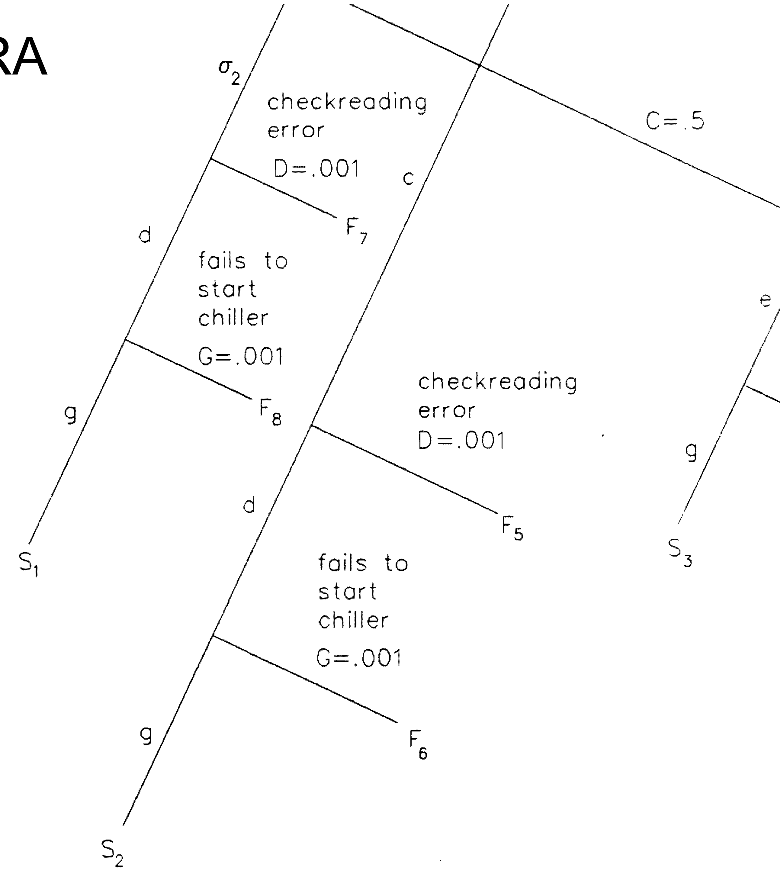
- Unit of analysis changes
  - e.g., In SPAR-H, a high-level HFE will almost always have Diagnosis and Action tasks, but a finer grained analysis will treat these separately
- Different HEPs result depending on unit of analysis
  - e.g., Ispra European HRA Benchmark in the late 1980s demonstrated that the greatest source of variability between analyses was due to different units of analysis
- Dependency analysis hinges on the basic units that are interdependent
  - e.g., original THERP notion of dependency was limited to dependency between subtasks **within** an HFE; yet most dependency treatment in HRA now uses THERP dependency **between** HFEs

Our HRAs are only as good as our definition of what we are analyzing

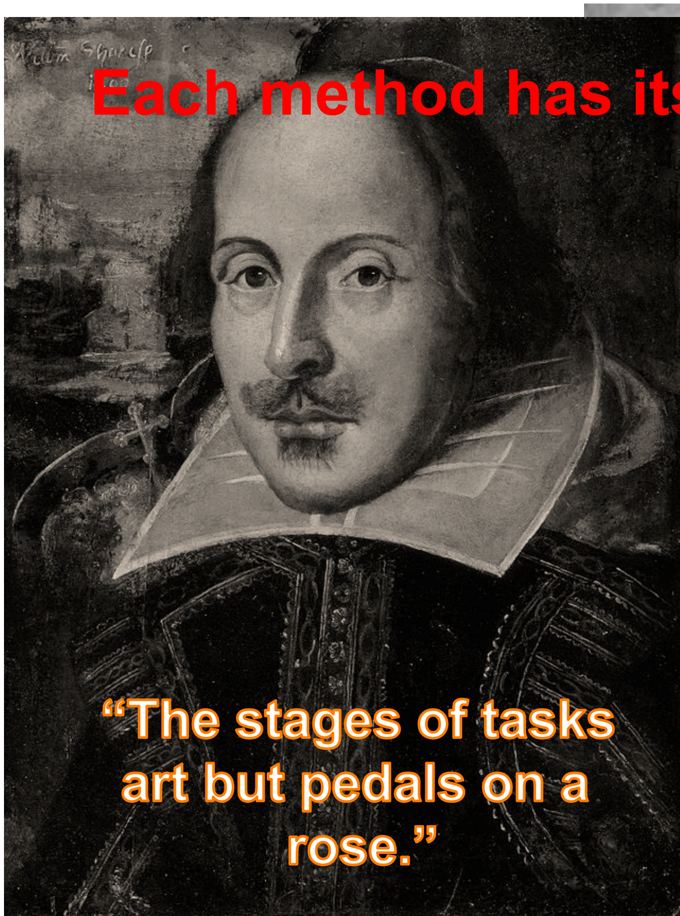
Lack of Consistent Task Decomposition

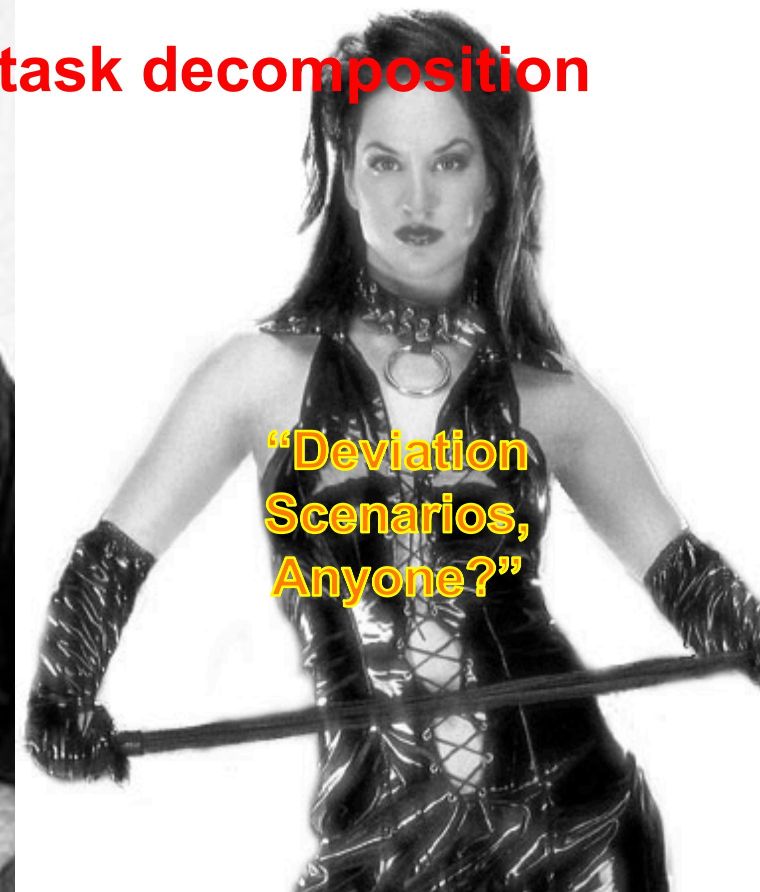
While all HRA treats HFEs, they do not treat it of define it the same

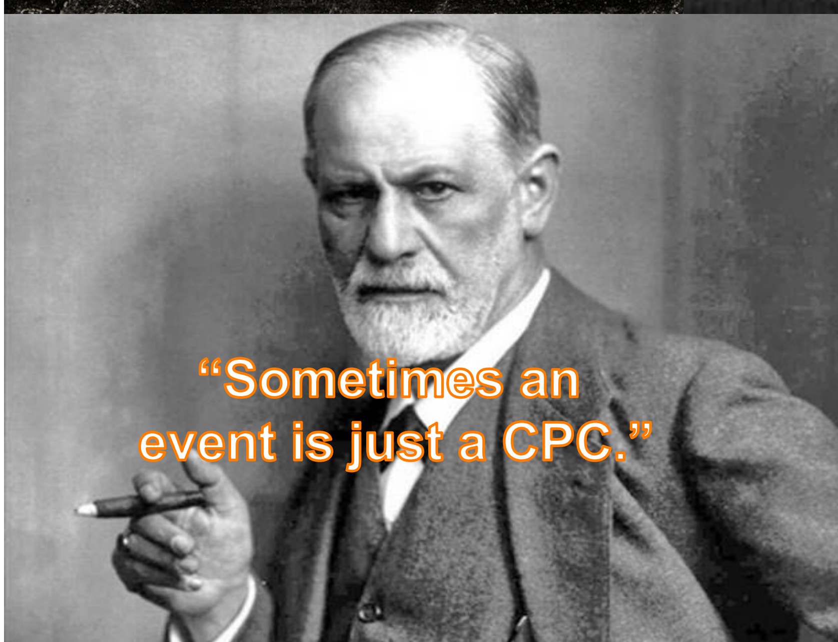
- *THERP:* Detailed subtask analysis using HRA event trees
- *SPAR-H:* Hey, whatever you want to call an HFE is cool with me
- *ATHEANA:* Deviation paths
- *EPRI HRA Calculator:* The HFE is defined in the PRA
- *CREAM:* It's all cognitive, Baby!
- *MERMOS:* The wide world of CICAs
- Etc…

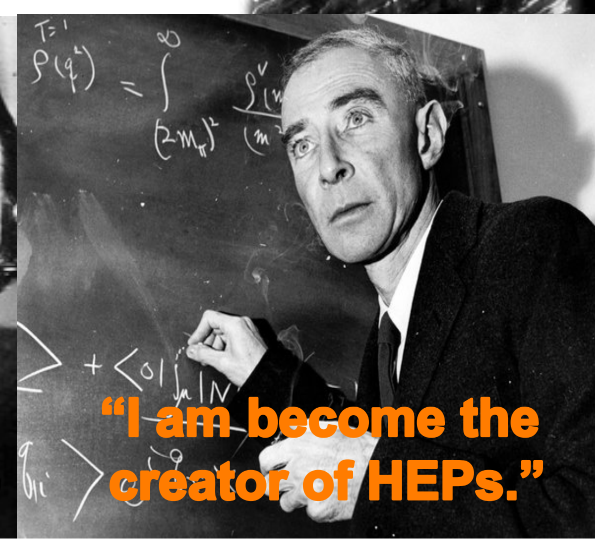


Each method has its own approach to task decomposition









The Human Error Probability (HEP) Changes Over Time

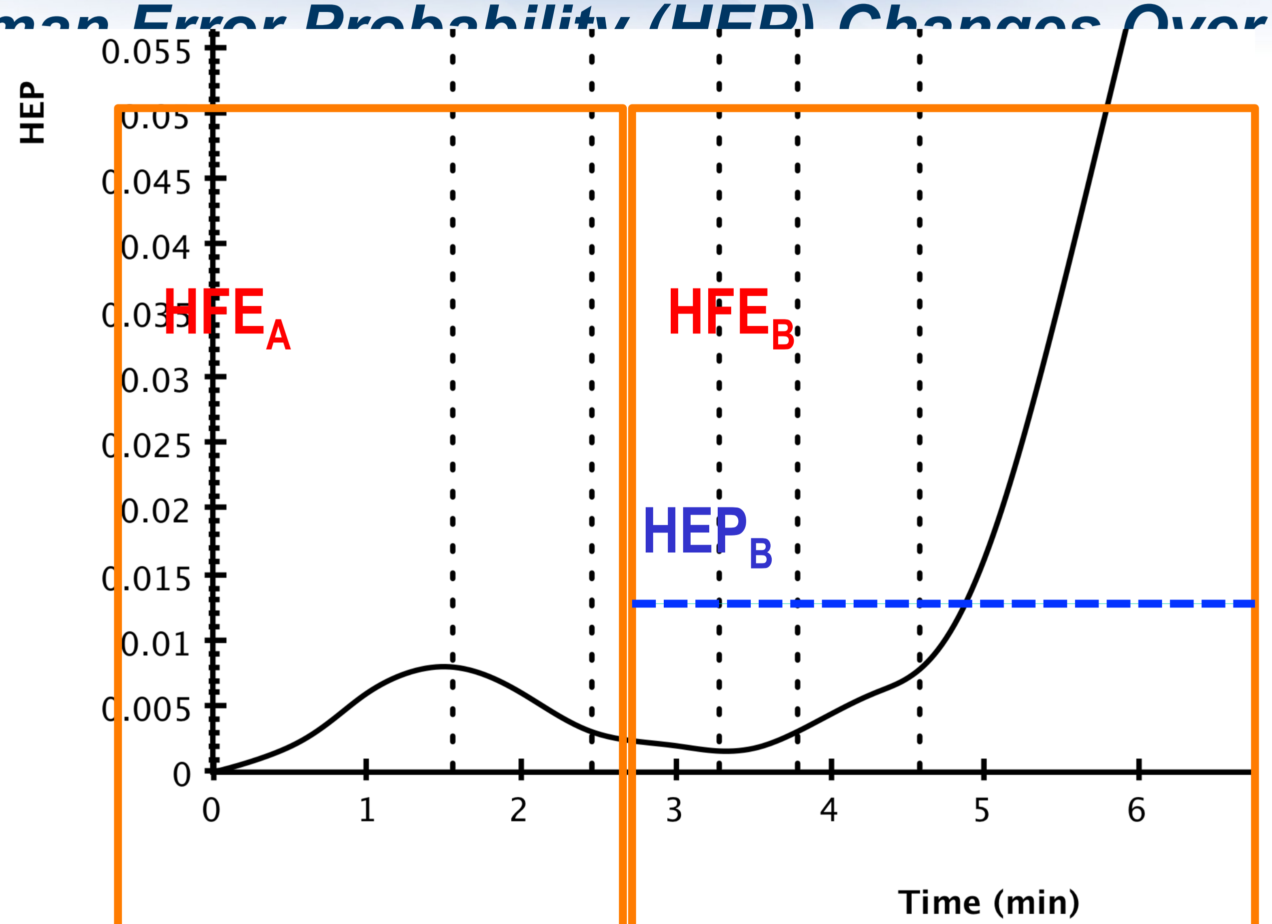
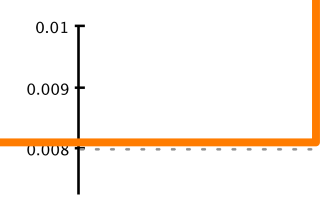
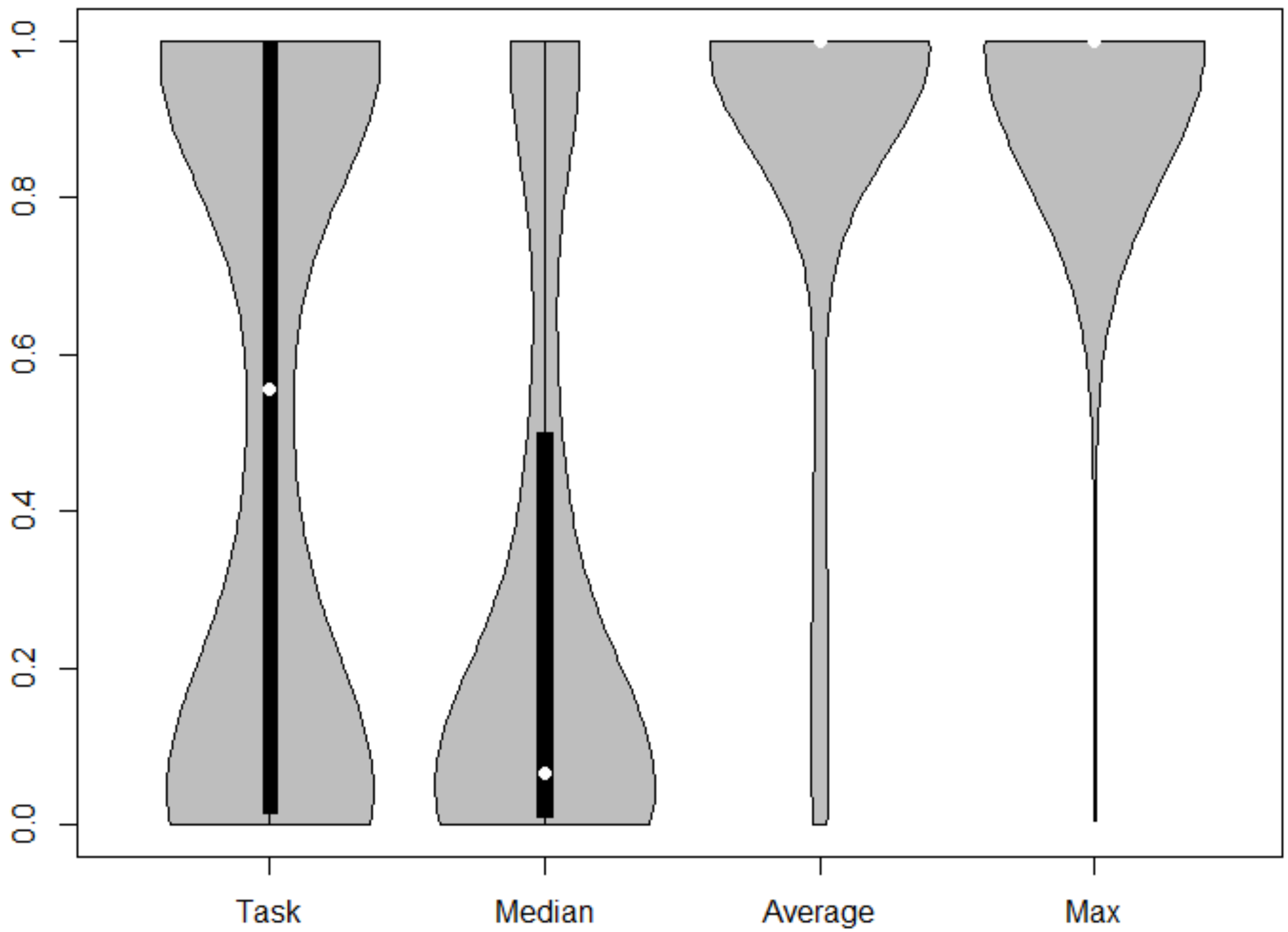


Figure 3. Hypothetical subtask HEP calculation for a dynamic event progression.

subtask will do so differently. As such, it is often convenient to consider the subtasks in terms of windows of time. Hypothetical Tasks A – I are parsed across the timeline in Figure 3. Within each subtask

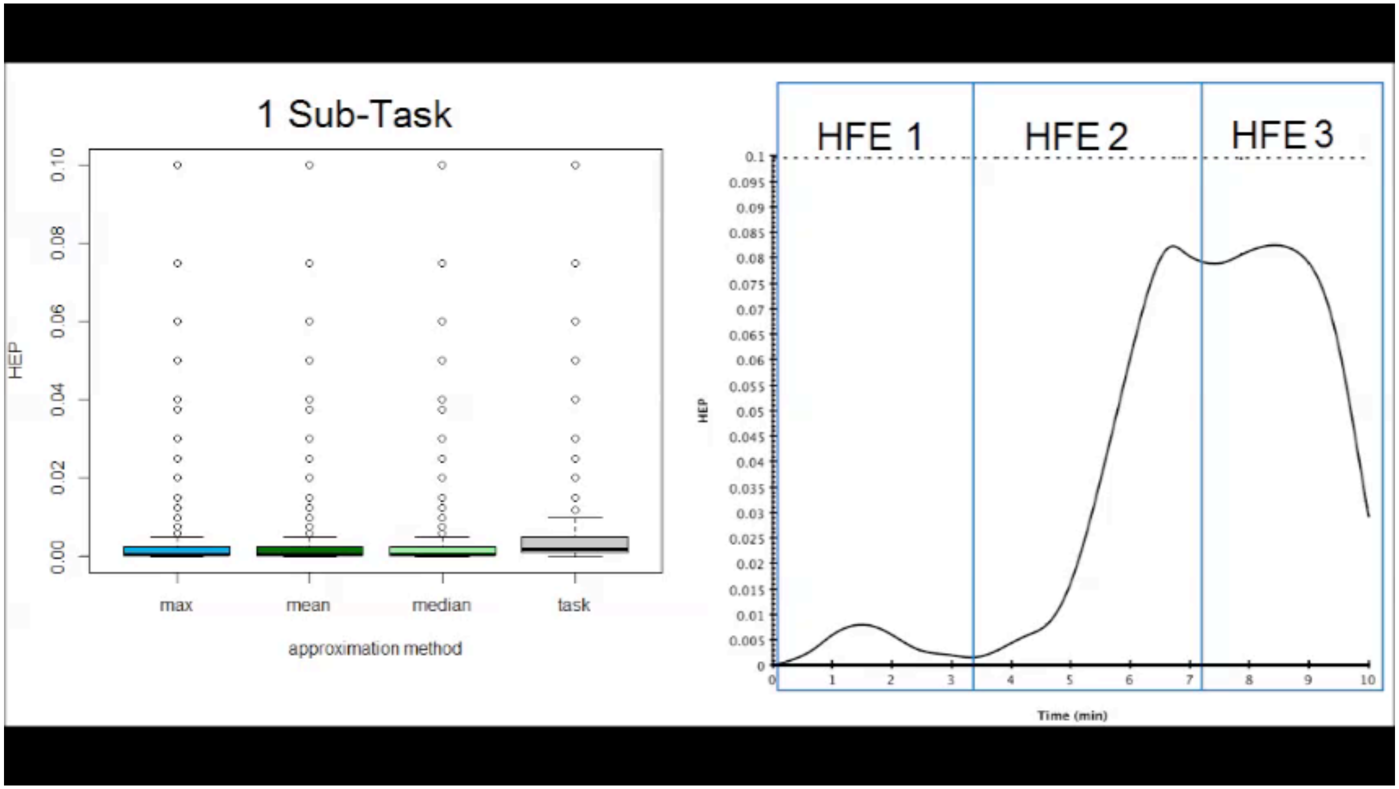


HRA Methods Like SPAR-H Produce Different HEPs Depending on How You Decompose Tasks



Sub-Task Modeling

- HEP increases as more sub-tasks are added



This Matters Because Dynamic HRA Uses Subtasks

Subtask Coupling
Virtual operator actions coupled to step-by-step plant model

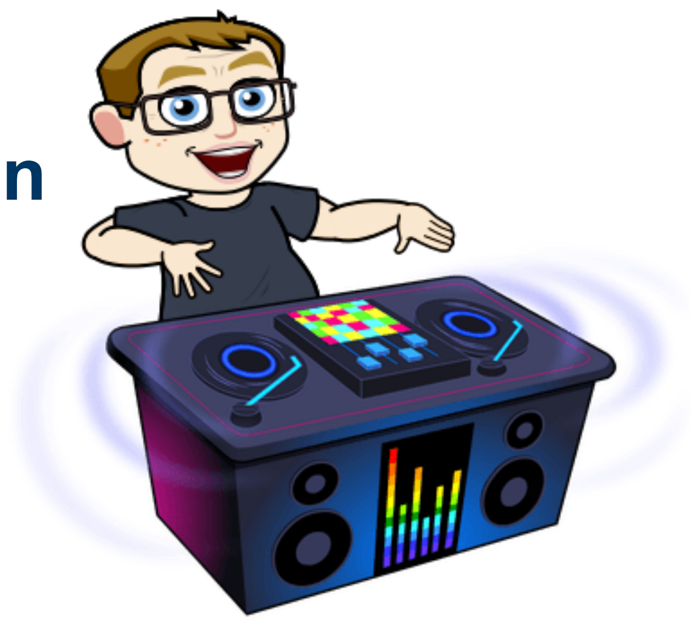
Let's break down that title…

Aggregation of
[combining]

Autocalculated Human Error Probabilities
[automatically generated HEPs]

from Tasks to Human Failure Events
[from subtasks to HFEs]

in a Dynamic HRA Implementation
[in HUNTER]



Different Ways to Average the HEP within HFEs

gle-point sub-
as a function
P within each
ormation such
also accompa-

calculated be-
Even though
defined event
x outcomes to
generation of
HEPs until all

This?
or
That?

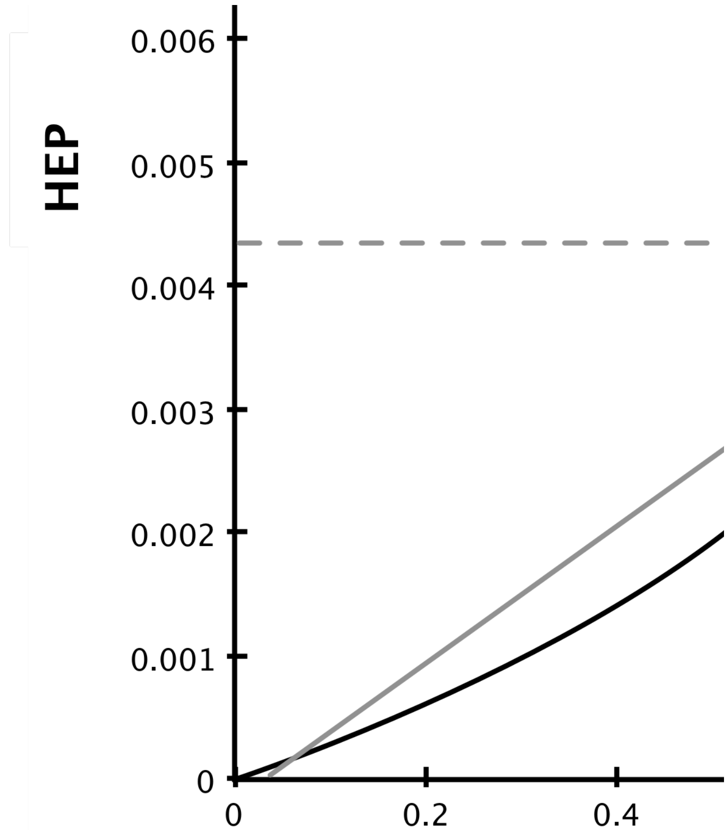
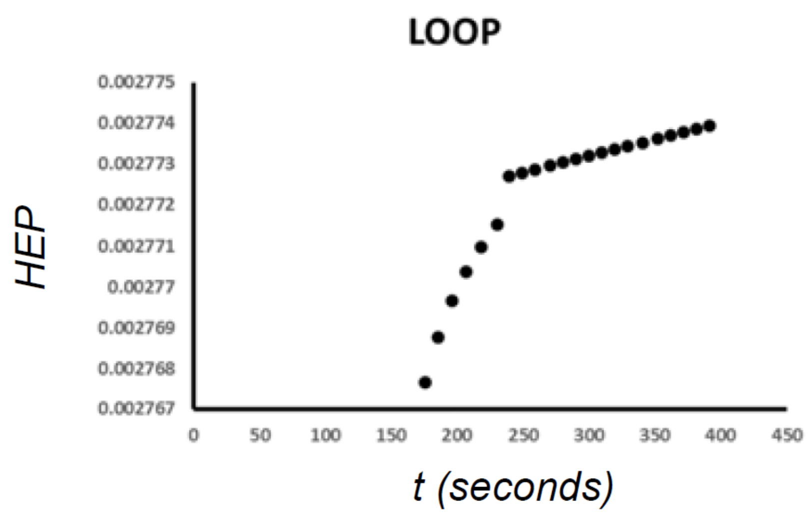


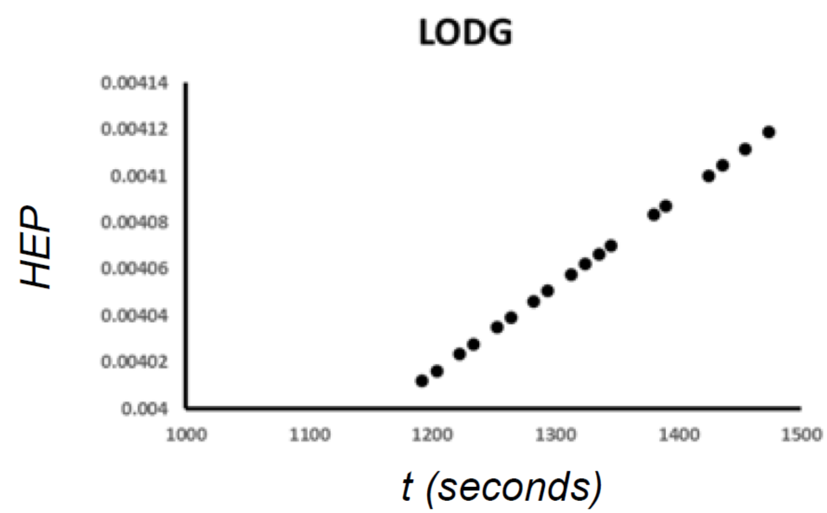
Figure 4. Four types of sub

Paper Explores Ways of Aggregating HEPs

Dynamic Generated Data Set for HEPs

- Loss of Offsite Power (LOOP)
- Loss of Diesel Generator (LODG)





- These represent two HFEs
  - Failure to prepare plant for shutdown
  - Failure to initiate backup power

Continuous data: What's the HEP for these HFEs?

Two Approaches for Aggregation Considered

Maximum HEP During the Interval

- Inherently conservative—worst case

$$HEP_{conservative} = \max_{HEP \in P} f(HEP)$$

- HEP_{LOOP} = 2.773E-3
- HEP_{LODG} = 4.118E-3
- Both HFEs have narrow ranges between minimum and maximum HEPs
  - $HFE_{LOOP\ [min,max]}$ = 6.288E-6
  - $HFE_{LODG[min,max]}$ = 1.070E-4
- **Conclusion 1:** For small ranges, maximum HEP in the range is reasonable single point estimate
- **Conclusion 2:** For larger ranges, maximum HEP in the range may present a conservative single point estimate

Two Approaches for Aggregation Considered

Central Tendency HEP Across the Interval: Median

- By definition, media provides single point estimate that falls in the middle of the data set
  - Good measure for capturing midpoint even with skews
- HEP_{LOOP} = 2.773E-3
  - Identical to maximum HEP due to narrow range of data
- HEP_{LODG} = 4.066E-3
  - Lower than maximum HEP due to broader range of data
- **Conclusion 3:** For small ranges, maximum HEP and median HEP are similar
- **Conclusion 4:** For larger ranges, median HEP will be lower than maximum HEP

Two Approaches for Aggregation Considered

Central Tendency HEP Across the Interval: Average

- Mean is average value of HEP function over the interval

$$E(HEP) = \int_{t_0}^{t_n} \frac{HEP(t)dt}{t_n - t_0}$$

  – Susceptible to outliers and skews
- HEP_{LOOP} = 2.772E-3
- HEP_{LODG} = 4.065E-3
  – Slightly lower than median values
- **Conclusion 5:** Average illustrates mathematically tractable way to summarize HEP across data range

Conclusions of Aggregation

Seemingly Trivial Solutions to a Hard Problem

- **Conclusion 1:** For small ranges, maximum HEP in the range is reasonable single point estimate
- **Conclusion 2:** For larger ranges, maximum HEP in the range may present a conservative single point estimate
- **Conclusion 3:** For small ranges, maximum HEP and median HEP are similar
- **Conclusion 4:** For larger ranges, median HEP will be lower than maximum HEP
- **Conclusion 5:** Average illustrates mathematically tractable way to summarize HEP across data range

Additional Data Sets Needed to Scope Best Aggregation Technique

Let's break down that title…

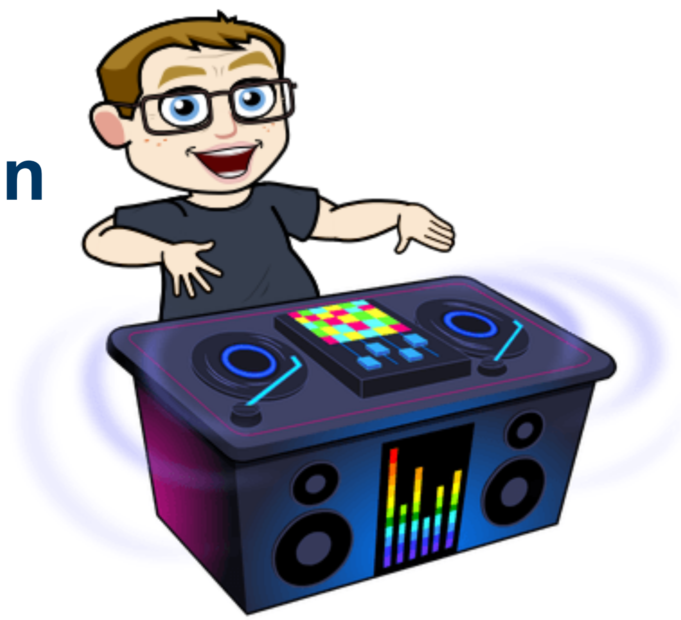
Aggregation of
[combining]

Autocalculated Human Error Probabilities
[automatically generated HEPs]

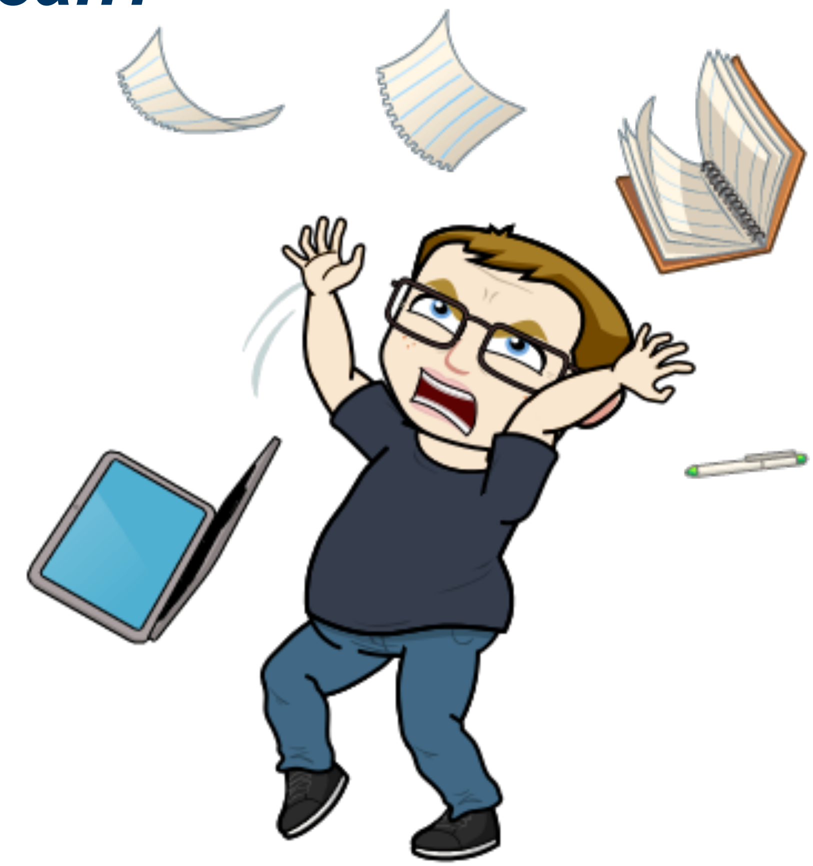
from Tasks to Human Failure Events
[from subtasks to HFEs]

in a Dynamic HRA Implementation
[in HUNTER]

Does this solve the HEP upward creep of subtask modeling?
Yes, as long as we don't consider dependency!

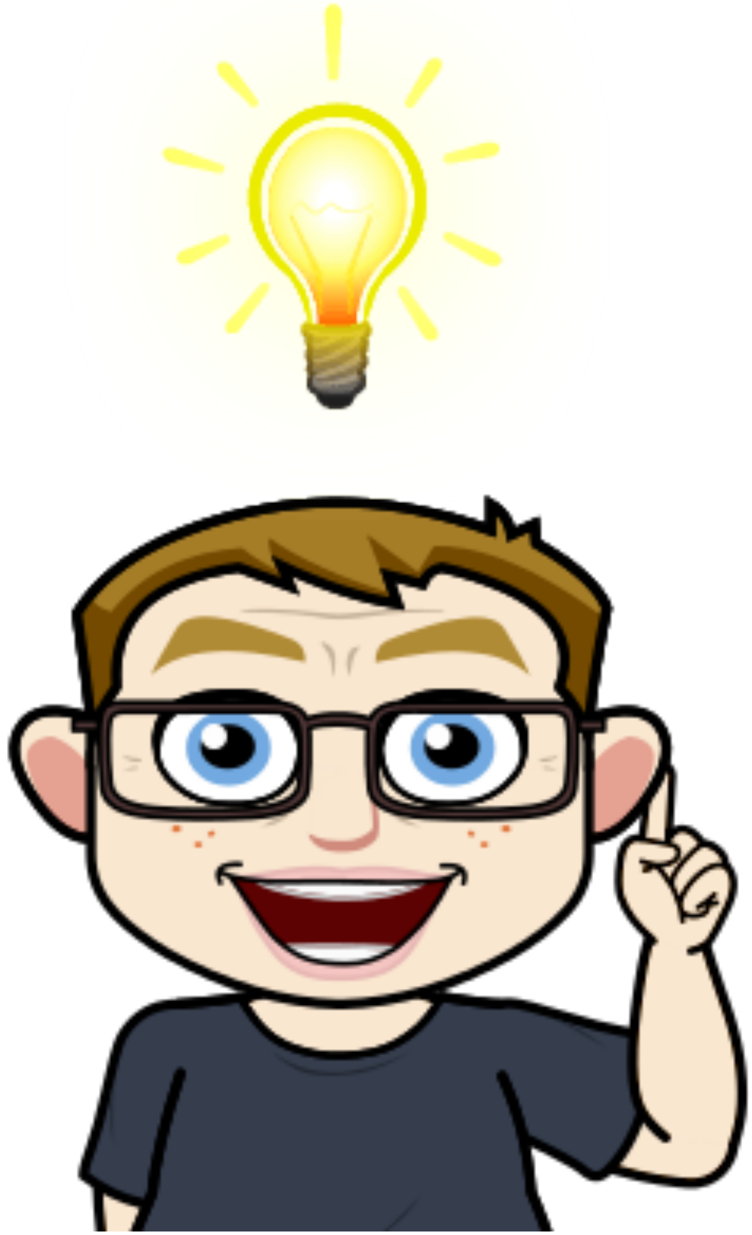


Wait…did I just spend 20 minutes telling you to take the mean?



What did I miss?

What are your ideas for subtask aggregation?




INL
Idaho National Laboratory
ronald.boring@inl.gov