SYDE 543 COURSE NOTES
COGNITIVE ERGONOMICS

Paolo Torres

University of Waterloo
Winter 2021

Table of Contents

1 WHY COGNITIVE ERGONOMICS?

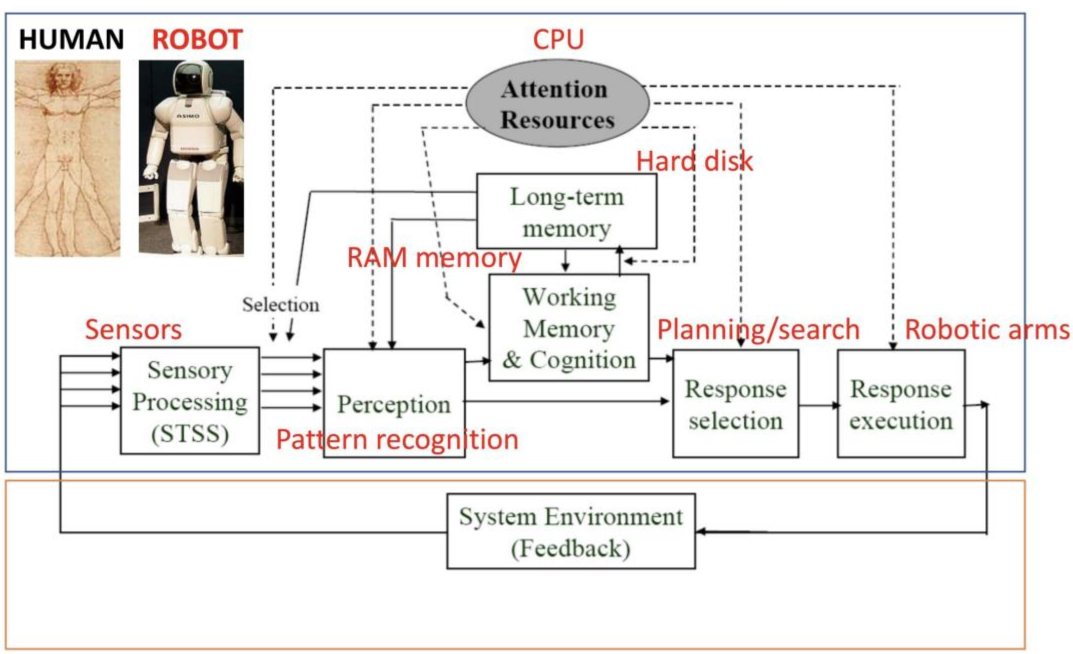
1.1 What is Cognitive Ergonomics?

- Cognitive ergonomics is the field of study that focuses on how well the use of a product matches the cognitive capabilities of users
- Mainly focuses on work activities which have an emphasized cognitive component, are in safety-critical environments, and are in a complex, changeable environment
- Domain: Environment where the system operates, presents constraints and opportunities
- Operates with two underlying theories: a theory about domain and about human cognition

1.2 The Descriptive Model of Human Information Processing

- Short Term Sensory Store (STSS): Events first processed by sight, sound, touch, etc.
- Perception: Determining meaning of events, long term memory of events
- Response Selection: A decision made based on either perception or working memory

A Robot/Computer/Automation/AI Analogy



1.3 Kind vs. Wicked Learning Environment

- Kind learning environments have next steps and goals that are clear, have rules that are clear and never change, get feedback that is quick and accurate (golf, chess, etc.)
- Wicked learning environments have next steps and goals that may not be clear, have rules that may change, may or may not get feedback
- The work world is a wicked environment, where hyper specialization can backfire
- In a wicked world, we need people who generalize first then specialize later on
- We need both frogs and birds, frogs to see the details up close, and birds to integrate the knowledge together, to succeed in a wicked world

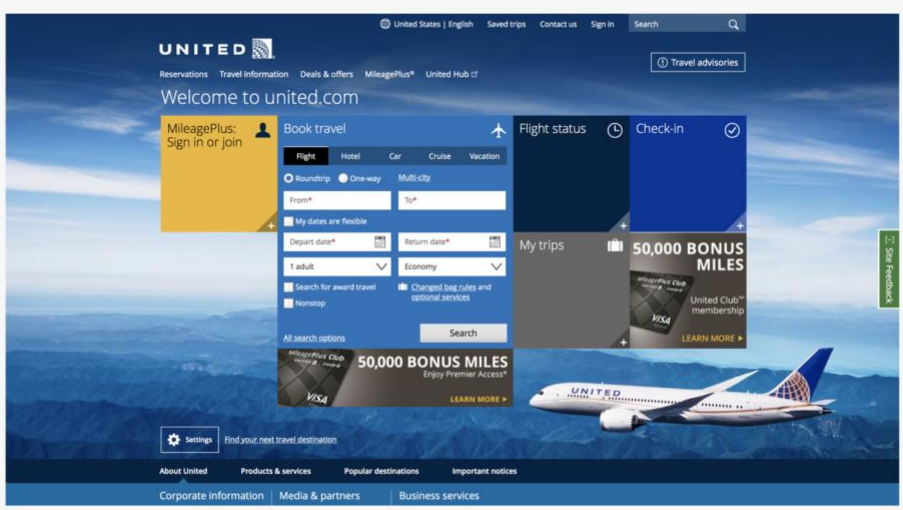
2 SIGNAL DETECTION THEORY AND UI/UX (PART 1)

2.1 Signal-to-Noise Ratio

2.1.1 Definitions

- Signal: Information that is relevant and useful to us
- Noise: Information that is irrelevant to our current need
- Signal-to-Noise Ratio: Ratio of relevant to irrelevant information in an interface

2.1.2 Example



United Airlines' homepage has several elements that might be signal in some situations, but noise in others.

- If booking a flight, “Book Travel” is a signal, but everything else is noise
- UI elements may serve functions other than simple communication or task efficiency
- Aim for a reasonable signal-to-noise ratio rather than excluding all “irrelevant” parts

2.1.3 Increasing Signal-to-Noise Ratio

- Pay attention to your content and have a strong visual hierarchy
- Start with a clear content strategy to help prioritize the information to convey
- Examples: Ensure every piece of text has some importance, avoid redundancy, separate paragraphs, bold keywords, use bullet points, etc.

2.1.4 Visual Hierarchy

- Reflects the relative importance of different elements on the interface (highly relevant, high visual weight)
- Examples: Making font large and bold, changing colour on action, adding an icon, etc.

2.1.5 Dynamic Noise

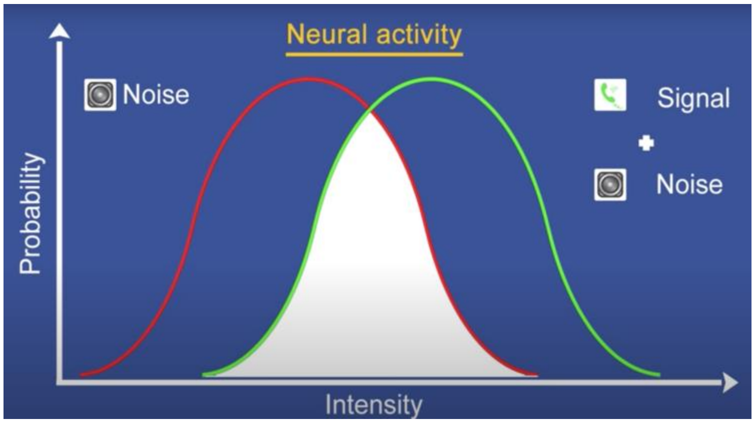
- What counts as noise can change from moment to moment, as the user’s task changes
- Example: Navigation on a website, where the navigation UI is noise while the user is focused on the page content, but becomes the signal once the user is done

2.1.6 Heuristics

- Aesthetic and minimalist design (remove unnecessary elements from the user interface)

2.2 Signal Detection Theory (SDT)

2.2.1 Neural Activity



2.2.2 Hit, Miss, False Alarm, Correct Rejection

- Hit: Positive response when there is a signal
- Miss: Negative response when there is a signal
- False Alarm: Positive response when there is no signal
- Correct Rejection: Negative response when there is no signal

	Signal + noise	Noise
Thinks phone ringing	Hit	False alarm
Thinks phone not ringing	Miss	Correct rejection

2.2.3 Perceptual Sensitivity (d')

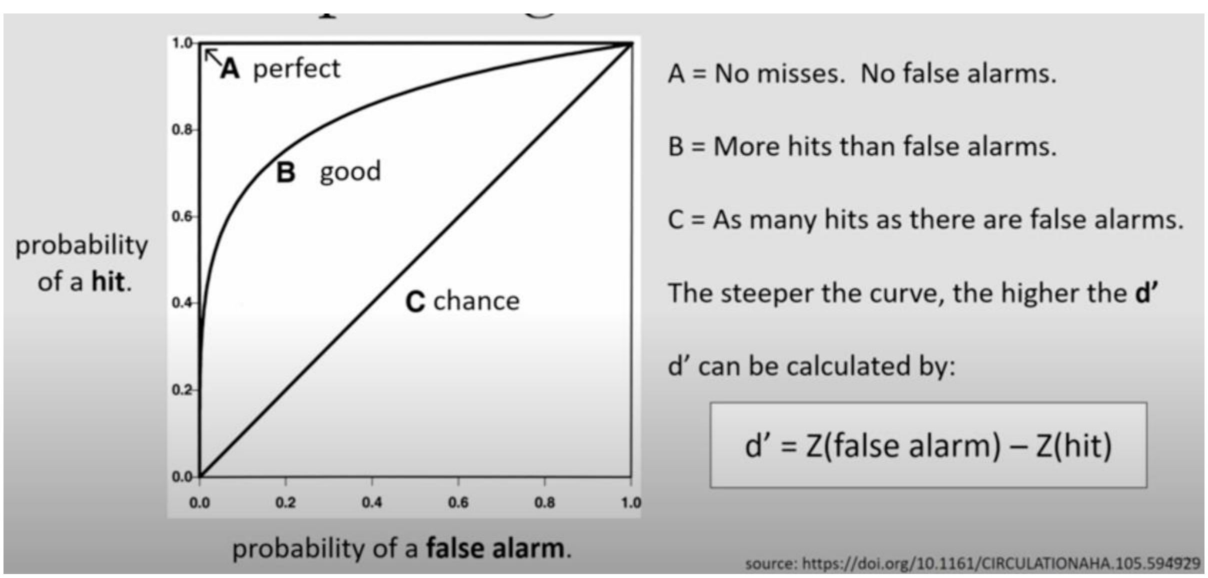
- How different the signal is from the noise
- Larger d': Signal more distinguishable from noise, more hits and correct rejections
- Smaller d': Signal less distinguishable from noise, more misses and false alarms

2.2.4 Decision Criteria (β)

- The degree at which the perceiver is biased to detect or not detect
- Conservative (large) β: Minimal detection, more misses and correct rejections
- Liberal (small) β: Maximal detection, more hits and false alarms

2.3 Receiver Operating Characteristic (ROC) Curve

2.3.1 ROC Curve



2.3.2 Relationship Between ROC Curve and d'

- The steeper the curve, the higher the d'
- $d' = Z(false\ alarm) - Z(hit)$

2.3.3 ROC Curve Axes

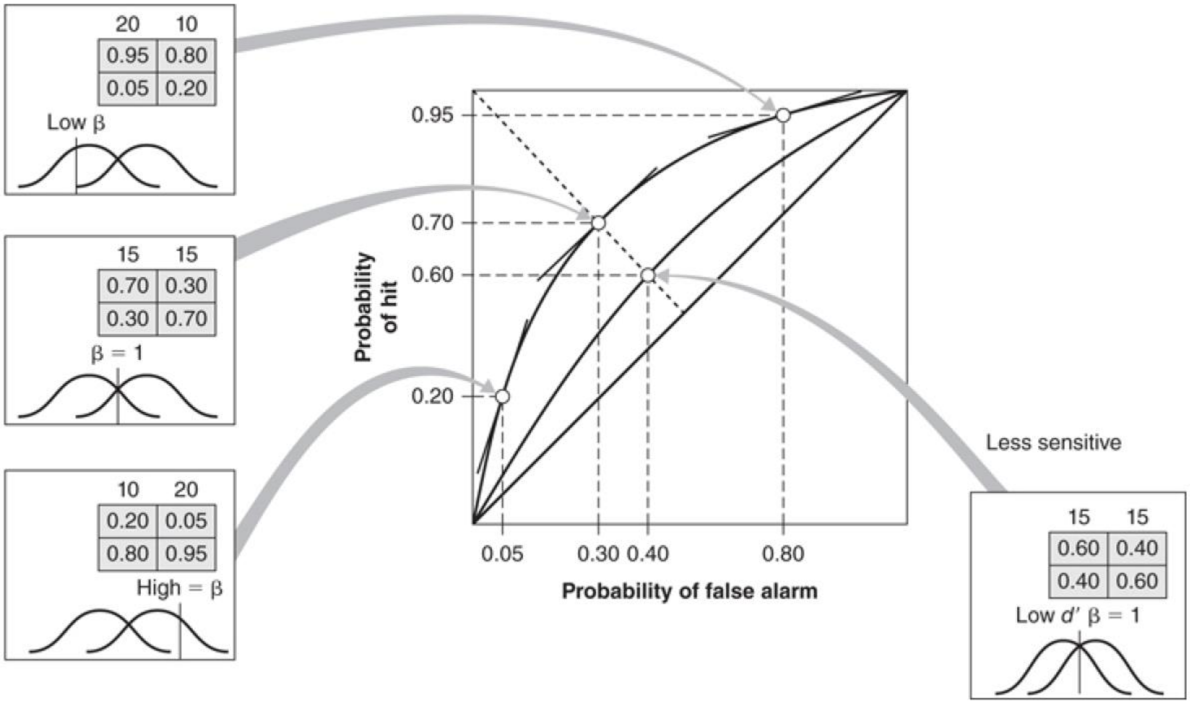
- $x - axis$: $probability\ of\ a\ false\ alarm$
- $y - axis$: probability of a hit

3 SIGNAL DETECTION THEORY AND UI/UX (PART 2)

3.1 Engineering Psychology and Human Performance

3.1.1 The ROC Curve

- Of the four values in SDT, only two are critical, $P(H)$ and $P(FA)$, since $P(M)$ and $P(CR)$ can be specified as $1 - P(H)$ and $1 - P(FA)$, respectively
- The ROC curve plots $P(H)$ against $P(FA)$ for different response criterions
- Each signal detection condition (each matrix) generates one point on the ROC
- Points falling on the curve have the *same* sensitivity
- Points in the lower left represent conservative responding, upper right risky responding

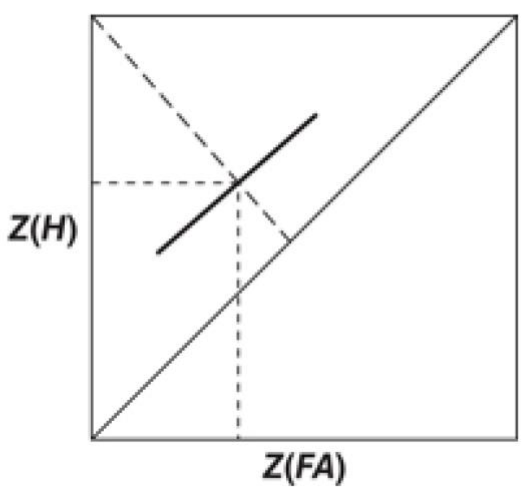


- More efficient method, **confidence levels**: $1 = confident\ no\ signal\ present, 2 = uncertain, 3 = confident\ signal\ present$
- For example, if levels 1 and 2 are "no" and level 3 "yes" classify as a conservative beta setting, but if level 1 is "no" and levels 2 and 3 are "yes", classify as a risky beta setting

TABLE 2.1 Analysis of confidence ratings in signal detection tasks

Subject's Response	Stimulus Presented		How Responses Are Judged	
	Noise	Signal		
"1" = "No Signal"	4	2	**No**	**No**
"2" = "Uncertain"	3	2	**No**	**Yes**
"3" = "Signal"	1	4	**Yes**	**Yes**
Total No. Of Trials	8	8	↓	↓
			Conservative Criterion	Risky Criterion
			P(FA) = 4/8	P(FA) = 4/8
			P(HIT) = 4/8	P(HIT)= 6/8

- The value of $beta$ at any given point along the ROC curve is equal to the slope of a tangent drawn to the curve at any point
- Slope is equal to 1 at points that fall along the **negative diagonal**
- $P(H) = 1 - P(FA)$
- Points on the **positive diagonal** represent chance performance: no matter how the criterion is set, $P(H)$ equals $P(FA)$, so the signal can't be distinguished from the noise
- Alternative way of plotting is to use z-scores, the bowed lines now become straight lines parallel to the chance diagonal



- A': A measure of the area under the ROC curve that provides an alternative sensitivity
- Represents the triangular area formed by connecting the lower left and upper right corners of the ROC space to the measured data point

$$A' = 0.5 + \frac{\big(P(H) - P(FA)\big)\big(1 + P(H) - P(FA)\big)}{4P(H)\big(1 - P(FA)\big)}$$

3.1.2 Fuzzy Signal Detection Theory

- Such "crisp" definitions of signal and noise are possible in everyday or work environments, yet more often than not, whether it is a signal or not is fuzzy
- Example: In air traffic control, a signal is when the flight paths of two aircraft come within $5\ nautical\ miles$ horizontally and $1{,}000\ feet$ vertically of each other
- However, the controller will consider a signal requiring action when these distances are exceeding these minimum values, depends on other factors like complexity and time

3.2 Wicked (Open-Ended) Problem

3.2.1 Vocal Biomarkers and COVID

- A group of researchers have discovered that they can determine if a person is infected with the coronavirus by analyzing signals hidden in their speech
- **Problem 1: Silent Spread:** One can infect others even if asymptomatic, an estimated 40% are asymptomatic, the virus has a highly variable incubation period
- **Problem 2: Delayed Results:** May take upwards of a week to get results, making results virtually meaningless for contact tracing
- **The Solution:** Have something that people can take in their own homes, with results being available within moments
- **Speech Signals:** Neurological diseases affect the brain's ability to process speech, and these changes can serve as vocal biomarkers
- **Inflammation and COVID:** Coupling between lung inflammation and speech could serve as a biomarker, would not prove COVID but could indicate presence of it

3.2.2 App to Detect COVID by Speech Analysis

- Strong evidence that COVID symptoms could be detected from human speech
- Speech contains inherent info about the physical, physiological, etc. status of a speaker
- This app would detect COVID symptoms at a much earlier stage
- **Speech Variations:** COVID will cause subtle variations to speech characteristics
- **Data Collection:** Recordings, along with body params, are measured and trained
- **Signal Processing and AI:** Techniques like filtering, voice activity detection, etc.
- **Challenges:** Minimizing false alarms (alert, no COVID) and misses (no alert, COVID)

3.2.3 New Method of Detecting Illnesses

- Examining individual molecules (biomarkers) to detect presence of disease in blood
- In theory could speed up coronavirus testing (minutes) and provide accurate results
- Involves using DNA origami, used to capture biomarkers, which are the indicators
- By modifying DNA origami to capture COVID molecules, can detect the proteins that the coronavirus uses to invade human cells

4 DESIGN FOR DECISION MAKING WITH A TWIST (PART 1)

4.1 Are We in Control of Our Own Decisions?

4.1.1 Intuition and Illusion

- Visual illusion as a metaphor for rationality
- Our intuition fools us in repeatable, predictable, consistent ways, and there is almost nothing we can do about it, aside from taking a ruler and starting to measure it
- Examples: horizontal vs. vertical table but same length, shadow on Rubik's cube but same colour

4.1.2 Cognitive Illusion

- In cognitive illusion, it is much harder to demonstrate the mistakes to people
- Example: organ donation, opt-in vs. opt-out, countries with opt-in by default have way more organ donation vs. countries with opt-out by default
- Example: Subscription options
  - $\$59$ web, $\$125$ print, $\$125$ web and print
  - Results: $16\%\ web, 0\%\ print, 84\%\ web\ and\ print$
  - Since the print only option was 0%, safe to remove right? No
  - When removed, results were: $68\%\ web, 32\%\ web\ and\ print$
  - Idea: Even though it was 0%, it had an indirect effect on the choice of the web and print option, perhaps due to it being strictly inferior to the print only option

4.1.3 Can Experts Overcome This Issue?

- Example: Medication and hip replacement

    - Send patient to hip replacement, but forgot to try one medication, most would call them back to try it
    - Send patient to hip replacement, but forgot to try two medications, most would let them go to hip replacement
- With more decisions comes more complexity, hence experts could still face problems

4.1.4 Gap Between Decisions and Actions

- Many decisions are not residing within is, but rather in the person who is designing what is being used/interacted with
- We have such a feeling that we're in the driver's seat, such a feeling that we're in control and we are making the decision, that it's very hard to even accept the idea that we actually have an illusion of making a decision, rather than an actual decision

4.2 Heuristics and Biases (HB)

4.2.1 Definitions

- Heuristics: People develop mental shortcuts to make decisions quickly
- Biases: Heuristics can be good old rules of thumbs, but they can also lead to cognitive biases (two sides of the same coin)

4.2.2 Origins of the Heuristics and Biases Approach

- Favours a skeptical attitude toward expertise and expert judgement
- Statistical predictions were more accurate than human predictions in almost every case
- Inferiority of clinical judgement was due in part to systematic errors
- Clinicians' uncritical reliance on their intuition and their failure to apply elementary statistical reasoning
- Inconsistency is a major weakness of informal judgement: when presented with the same case information on separate occasions, human judges often reach different conclusions
- Human judgements are noisy to an extent that substantially impairs their validity
- **Illusion of Validity:** Unjustified sense of confidence that often comes with clinical judgement

- Sophisticated scientists reached incorrect conclusions and made inferior choices when they followed their intuitions, failing to apply rules with which they were familiar

4.3 Bayes' Theorem

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} = \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid A')P(A')}$$

4.3.1 Example

- The probability that a woman has breast cancer (event A) is 1% ("prevalence") $P(A) = 0.01$, so $P(A') = 0.99$
- If a woman has breast cancer (event A), the probability that she tests positive (event B) is 90% ("hit/true positive of the machine"), $P(B \mid A) = 0.9$
- If a woman does not have breast cancer, the probability that she nevertheless tests positive is 9% ("false alarm rate of the machine"), $P(B \mid A') = 0.09$
- Given a woman has a positive test result, what is the probability that this woman actually has breast cancer? $P(A \mid B) = ?$

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid A')P(A')} = \frac{0.9 \times 0.01}{0.9 \times 0.01 + 0.09 \times 0.99} = 0.09174$$

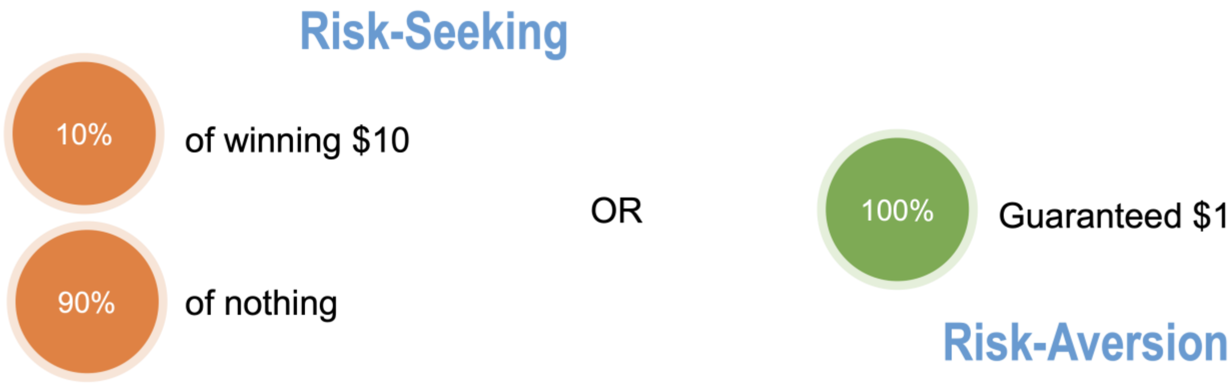
4.4 Loss Aversion

The well-known "Loss Aversion"

Probability	Outcome	
	Profit	Loss
Small	Risk seeking	Risk aversion
Large	Risk aversion	Risk seeking

- Loss Aversion: People react to losses more strongly than gains and they try to prevent losses more than they try to make gains
- But people often forget about the conditions (e.g., it is actually just small probability loss aversion)

4.4.1 Example



5 DESIGN FOR DECISION MAKING WITH A TWIST (PART 2)

5.1 Naturalistic Decision Making (NDM)

5.1.1 Origins of the NDM Approach

- Focuses on the successes of expert intuition, decision making under complex conditions
- Examples: Chess grandmasters vs. mediocre players, fireground commanders
- Draw on the repertoire of patterns that had compiled during decades of experience
- In the face of uncertainty, time pressure, high stakes (i.e., a wicked environment)
- A central goal of NDM is to demystify intuition by identifying the cues that experts use
- Use cognitive task analysis (CTA) methods to investigate the cues and strategies

5.1.2 Contrasts Between NDM and HB

Three important contrasts:

1. The stance taken by the NDM and HB researchers toward expert judgement
2. The use of field versus laboratory settings for decision making research
3. The application of different standards of performance, which leads to different conclusions about expertise.

Stance Regarding Expertise and Decision Algorithms

- NDM: Tend to stress the marvels of successful expert performance, explore the thinking of experts, little faith in formal approaches, skeptical about universal structures and rules
- HB: Tend to focus on flaws in human cognitive performance, skepticism, compare with performance by formal models or rules, replace with algorithms when possible

Field Versus Laboratory

- NDM: Operate in "real world" organizations, natural sympathy for ecological approach
- HB: Mostly based in academic departments, favor well-controlled experiments in the lab

The Definition of Expertise

- NDM: Criteria for judging expertise are based on a history of successful outcomes, rely on peer judgements, consensus and evidence, best practitioners define the standard
- HB: Evaluated by comparing the accuracy of decisions with accuracy of optimal linear combinations, compare the judgement of professionals with the outcome of a model

5.2 Ethnography

- **Ethnography:** The study of human behaviour within a culture; a research method central to knowing the world from the standpoint of its social relations
- What people say, what people do, and what people say they do are different things
- Hidden obvious: Insights that are obvious only after someone points them out
- Example: Parking signs in the city are sometimes difficult to read

5.3 "Skill, Rule, and Knowledge" Model of Decision Making (SRK)

5.3.1 Where Do Rasmussen Models Fit?

- Best known models: The abstraction hierarchy and the skill-rule-knowledge taxonomy
- Have verbal elements within categories and an order relationship between the elements
- More widely applicable (suggestive, in the sense of analogy) than other models

5.3.2 Evaluating the SRK Model Attributes

1. Applicability to Observables: Observes anything, there is no context limits built in
2. Dimensionality: Open to considering single input-output variables or complex combos
3. Metricity: Poses inherent ordering based on the hierarchy assumptions
4. Robustness: Applies to any behavior of any human, as well as to "intelligent automation"
5. Social Penetration: Penetrated as well or better than other models
6. Conciseness: Model is simply stated, not overly wordy, has good graphic representations

5.3.3 Relation of SRK to Automation

- Reasonable qualitative breakdown of hierarchical modes of behavior of living things

- Skills are inherited but are also conditioned through experience; automated routines requiring little conscious attention
- Rules are conditioned and codified by experience in the brain or in external references; pre-packaged units of behavior released (IF, THEN, ELSE)
- Knowledge is that still mysterious process inherent in creative decisions and acts, improvisation in unfamiliar environments, no routines or rules available
- Example: Is a model of how modern automatic control technology functions
- Modern automation is hierarchical in the same sense as the SRK model
- Classical feedback control relates squarely to Rasmussen's skill category
- There is adaptive control which modifies fixed parameters according to rules
- Analogous to the knowledge block are interactions with large flexible databases

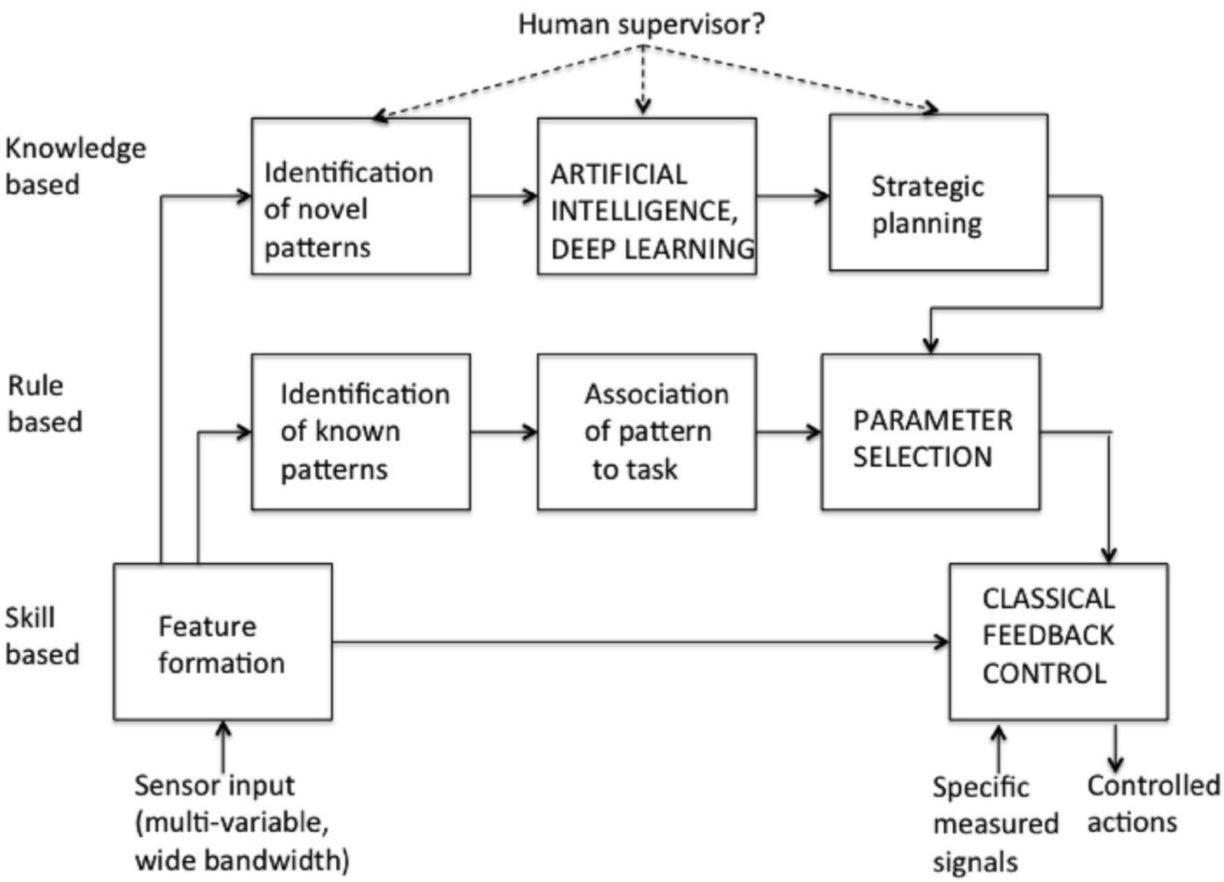


Figure 1. SRK representation of three-level automation.

6 DESIGN FOR THE STORM RATHER THAN THE CALM (PART 1)

6.1 Minimize Cognitive Load to Maximize Usability

The total cognitive load, or amount of mental processing power needed to use your site, affects how easily users find content and complete tasks.

6.1.1 Intrinsic vs. Extraneous Cognitive Load

- **Intrinsic Cognitive Load:** The effort of absorbing new information and keeping track of their own goals.
- **Extraneous Cognitive Load:** Processing that takes up mental resources but doesn't actually help users understand the content (for example, different font styles that don't convey any unique meaning).

6.1.2 Design Solutions to Minimize Cognitive Load

1. **Avoid Visual Clutter:** Redundant links, irrelevant images, and meaningless typography flourishes slow users down (note that meaningful links, images, and typography are valuable design elements; it is only when overused that these backfire and actually impair usability).
2. **Build on Existing Mental Models:** People already have mental models about how websites work, based on their past experiences visiting other sites. When you use labels and layouts that they've encountered on other websites, you reduce the amount of learning they need to do on your site.
3. **Offload Tasks:** Look for anything in your design that requires users to read text, remember information, or make a decision. Then look for alternatives: can you show a picture, re-display previously entered information, or set a smart default? You won't be able to shift all tasks away from users, but every task you eliminate leaves more mental resources for the decisions that truly are essential.

6.2 Mental Workload: Assessment, Prediction and Consequences

- Physical workload:

$$Physical\ workload = \frac{Physical\ energy\ requirement}{Physical\ energy\ expenditure\ capacity}$$

- Mental workload:

$$Mental\ workload = \frac{Mental\ energy\ requirement}{Human\ information - processing\ capacity}$$

- How to measure mental workload: secondary task performance, subjective ratings, physiological measures, and computational modeling

6.2.1 High Mental Workload, Human Decision-Making, and the "4 Ss"

- When workload becomes excessive, three things can happen:
    1. Over the long run, high workload can exert a toll on health and well-being
    2. When work us driven over the "red line", performance can start to fail
    3. People are typically effort conserving, wishing to avoid the stress of high MWL, they often make decisions with negative consequences to performance
- Decisions that people make to choose between one of two alternatives: one of **higher value**, and the other of **lower effort**
- The "4 Ss" are the choice: to **switch** attention between tasks, to **seek** information, to **study** material, and to **safely** behave

6.2.2 STOM

- **Strategic Task Overload Management:** A multi-attribute decision model that predicts, based on four task attributes, which task a person will choose to switch attention to, amongst a multi-task ensemble in overload situations.
- Examples: Management of an oil rig disaster, handling an in-flight aircraft emergency
- The four task attributes are priority, interest, salience, and task difficulty
- Two aspects of STOM that cry out for valid predictive models:
    - Within the four attributes, difficulty or effort demand has been found to produce a fairly robust influence on task switching
    - Humans have an inherent "switch resistance," or bias to keep doing what they are doing

6.2.3 SEEV

- **Salience, Effort, Expectancy, Value:** Describes the movement of *visual attention* around the *visual space* in order to acquire information necessary to accomplish a task.
- Involves the role of effort in the decision of "where" and "whether" to look

- Example: a driver's failure to decide to check the blind spot, "where" could be looking behind, and "whether" could be the decision not to scan

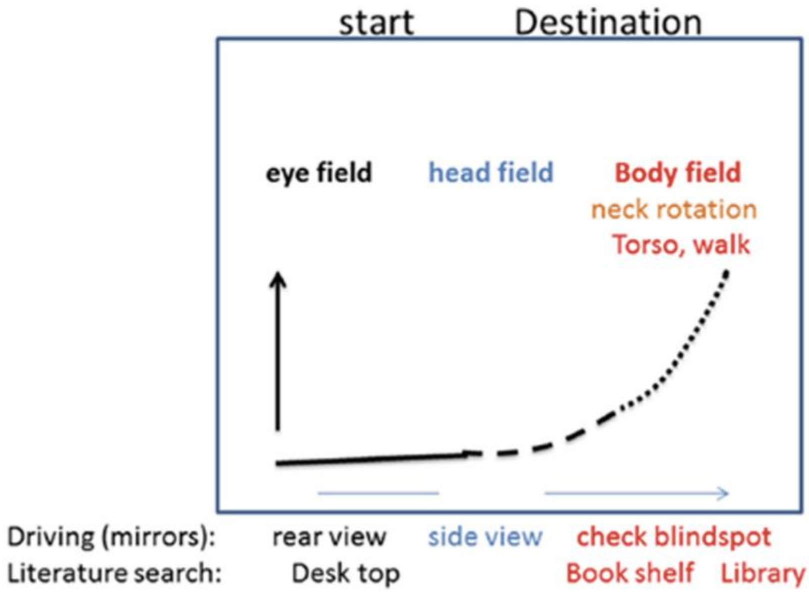


Fig. 2. Information access effort as a function of distance from the forward view: the legends on the X-axis represent the access of two different kinds of information: a view behind a vehicle (top row) and a reference citation (bottom row). From [22].

6.2.4 Trade-Offs Between Effort and the Value of Studying

- Many of the choices that students make in study strategies result, in part, from the false belief that easier (less effortful) study and learning strategies signal better retention or transfer of learning; that is, higher value
- Trade-off of effort and value in study strategies:

Lower Effort and Value	Higher Effort and Value
Massed practice	Spaced practice (contextual interference)
Listening	Note taking (active response generation)
Note taking	Self-quizzing (active retrieval practice)
Part task training	Whole task training

6.2.5 Role of Perceived Effort and Discomfort in Decision-Making

- Example: the "discomfort" of wearing a safety helmet is not truly a source of effort demands, but imposes the same sort of negative or unpleasant valence as expending effort, particularly if the driver is experiencing other sources of stress

- Example: reading long safety instructions on equipment or drug labels is clearly effortful, particularly if they are not well worded and excessive in number
- More examples: driving on a straight motorway late at night with fatigue diminishing the capacity to mobilize effort, higher degrees of automation facilitate the ability to engage in effort-light pleasure like mind wandering

6.2.6 Differences Between Intrinsic, Extraneous, and Germane Load

- Three sources of effort demand (MWL) imposed in the learning environment:
    1. More complex tasks intrinsically demand more effort to be performed (intrinsic load)
    2. More effort can be invested into the more physically and cognitively challenging but effective strategies, it is productive or "germane" for learning, known as germane load
    3. Effort can be demanded in the learning space that has little effectiveness, it is "extraneous load"

6.2.7 Bias Associated with the Relationship Between Effort and Value

- The bias to overestimate the benefits (value) of some training strategies
- The overconfidence bias to underestimate the expected costs or risks of unsafe behavior

7 DESIGN FOR THE STORM RATHER THAN THE CALM (PART 2)

Stress is related to mental workload:

- Yerkes-Dodson Law: Too much or too little stress is bad, the optimal stress level depends on the difficulty of task.
- Mental Workload vs. Stress: Mental workload is relationship between task demands and mental processing capability, stress is the response to disruption, challenge, and threats.

7.1 Expertise and Situation Awareness

7.1.1 Situation Awareness (SA) Definition and Levels

- **Situation Awareness:** The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future.

- Level 1 SA – Perception: The perception of relevant information from the environment.
- Level 2 SA – Comprehension: Situation awareness involves more than simple perception of information – it also demands that people understand the meaning and significance of what they have perceived.
- Level 3 SA – Projection: At the highest level of SA, the ability to forecast future situation events and dynamics marks individuals who have the highest level of understanding of the situation.

7.1.2 Relationship Between SA and Human Information-Processing

Key features of the environment affect how well people are able to obtain and maintain SA:

1. The capability of the system for providing the needed information
2. The design of the system interface, determining which information is available
3. System complexity, including number of components, inter-relatedness of components
4. The level of automaton present in the system
5. Stress and workload that occur as a function of the environment, interface, and domain

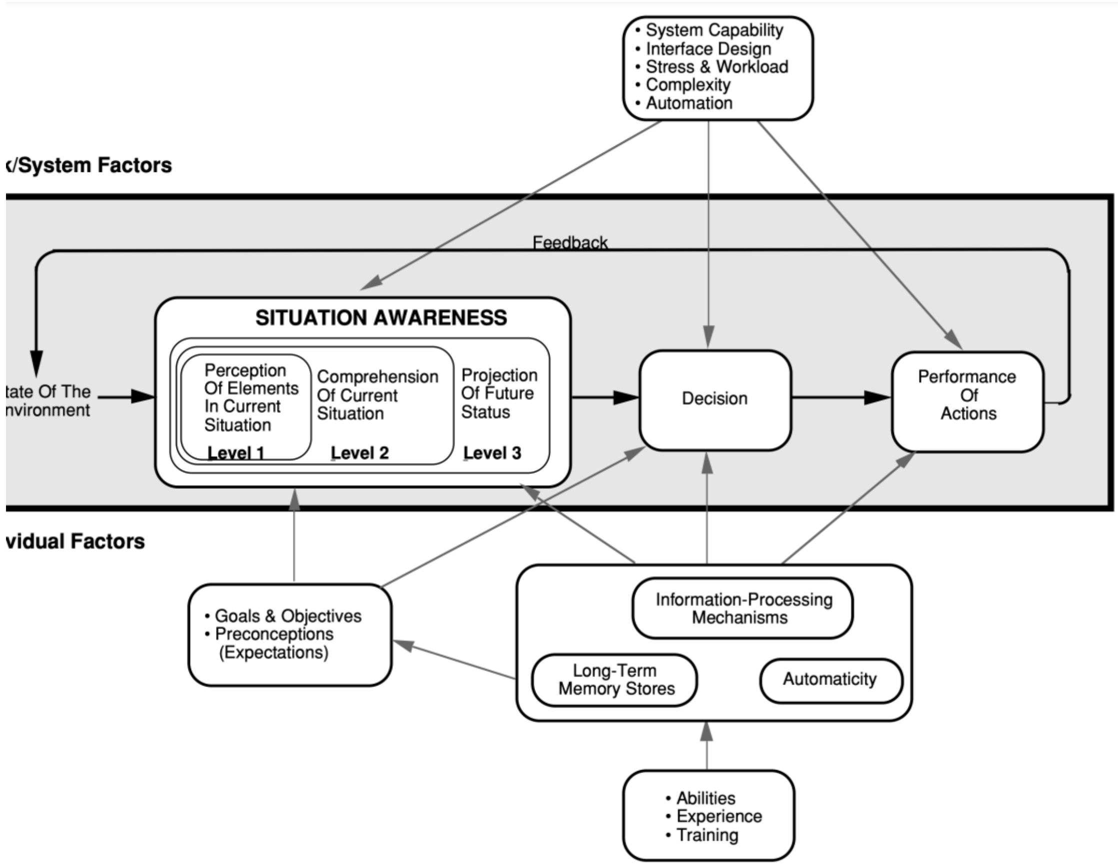


Figure 36.1. Model of Situation Awareness in Dynamic Decision Making (Endsley, 1995)

7.1.3 How to Measure SA

1. **Situation Awareness Global Assessment Technique (SAGAT)**
   - Randomly freeze a task – just like getting a phone call when you are watching an exciting Netflix show
   - Ask the operator a series of questions to assess their knowledge of the situation
   - Use correct rate (%) as a measure of SA
2. **Situational Awareness Rating Technique (SART)**
   - Similar to SAGAT, except SART is a standardized set of questions, and is asked after the completion of the task rather than in-between

7.1.4 Role of Expertise in Situation Awareness

Expertise in a particular domain has a significant role in allowing people to develop and maintain SA in the face of high volumes of information transfer and system complexity.

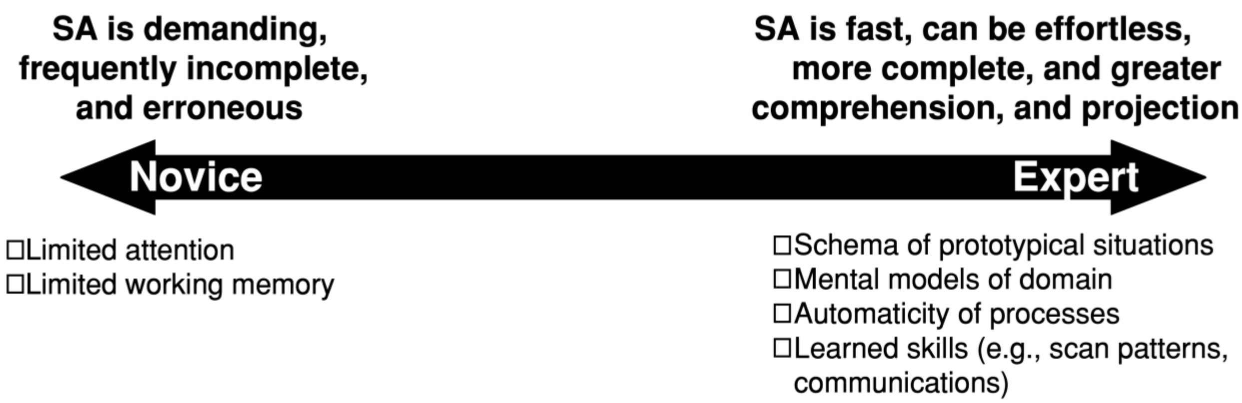


Figure 36.2. Factors effecting SA in Novices and Experts in a Domain

7.2 Context-Aware Design

7.2.1 Innovating with Awareness

- Fanvision, while a nifty idea, only is aware that the game is going on, while the changes in information are only a broadcast model.
- But in the train-catching scenario, we see an awareness to the user's position, their travel goals, and their food needs; provides an advanced view of awareness, better experience.

7.2.2 Already Here: The Apple Store App

- Apple has taken this idea of awareness and built it into a wonderful store application for their iOS devices.
- Customers can, from anywhere in the store, summon a salesman, schedule a Genius Bar appointment, or even purchase a product they are holding by scanning it.

7.2.3 Infusing Context into the Design Process

- We've never really designed for context before; things like detecting if user was logged in or already established an account aren't really user context, but rather system states.
- Real context aware applications are combining several features together, based on what we're trying to do and how far along we've gotten.
- Existing design processes don't work for this; knowing how the visual design needs to adapt to the changing needs of the user is not something we've thought about before.
- This process starts with mapping out the current experience, then seeing where the context breakpoints occur, and where information needs to shift in the application.
- Designing for situational awareness will become the norm; right now, it's still a burgeoning frontier and experimenting and seeing what works.