Mathematical Cognitive Modeling

Joe Houpt

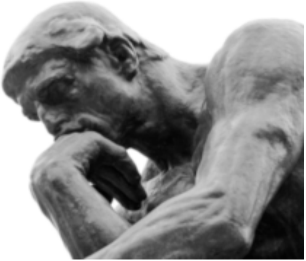
Interdisciplinary College

2019

Outline for section 1

What is mathematical cognitive modeling?

What is mathematical cognitive modeling?



Why bother?

Many of the same motivations for models in general apply.

- Reduces complexity

Why bother?

Many of the same motivations for models in general apply.

- Reduces complexity
- Allows precise predictions
  - Enables theory testing

Why bother?

Many of the same motivations for models in general apply.

- Reduces complexity
- Allows precise predictions
  - Enables theory testing
- Makes assumptions explicit
- *Facilitates communication*

Aspects of cognition

- Perception
- Decision-making
- Memory
- Learning
- Attention

Random Variables

$$X$$

Random Variables

$$X$$

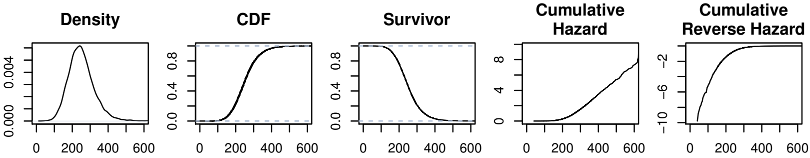
- Instead of just defining X's range, X is associated with a function defining the probability that X could be within any interval on its range.

Random Variables

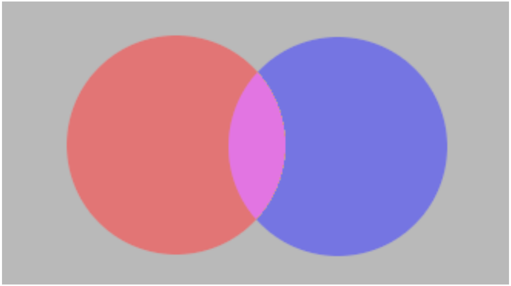
$$X$$

- Instead of just defining X's range, X is associated with a function defining the probability that X could be within any interval on its range.
- This association can be summarized by X's cumulative distribution function

CDF	PDF	Survivor	Hazard
$F(x) = P(X \leq x)$	$\frac{dF}{dx} = f(x)$	$1 - F(x)$	$\frac{f(x)}{1-F(x)}$



Joint and Conditional Probability



$$P(x \in A \text{ and } x \in B)$$
$$P(x \in A \text{ or } x \in B)$$
$$P(x \in A \mid x \in B)$$

Bayes Rule

$$
\begin{aligned}
P(A \mid B) &= \frac{P(A \text{ and } B)}{P(B)} \\
P(B \mid A) &= \frac{P(A \text{ and } B)}{P(A)}
\end{aligned}
$$

Bayes Rule

$$
\begin{aligned}
P(A \mid B) &= \frac{P(A \text{ and } B)}{P(B)} \\
P(B \mid A) &= \frac{P(A \text{ and } B)}{P(A)}
\end{aligned}
$$

$$
\begin{aligned}
P(A \text{ and } B) &= P(A \mid B)P(B) \\
P(A \text{ and } B) &= P(B \mid A)P(A)
\end{aligned}
$$

Bayes Rule

$$
\begin{aligned}
P(A \mid B) &= \frac{P(A \text{ and } B)}{P(B)} \\
P(B \mid A) &= \frac{P(A \text{ and } B)}{P(A)}
\end{aligned}
$$

$$
\begin{aligned}
P(A \text{ and } B) &= P(A \mid B)P(B) \\
P(A \text{ and } B) &= P(B \mid A)P(A)
\end{aligned}
$$

$$
P(A \mid B)P(B) = P(B \mid A)P(A)
$$

Bayes Rule

$$\begin{aligned} P(A \mid B) &= \frac{P(A \text{ and } B)}{P(B)} \\ P(B \mid A) &= \frac{P(A \text{ and } B)}{P(A)} \end{aligned}$$

$$\begin{aligned} P(A \text{ and } B) &= P(A \mid B)P(B) \\ P(A \text{ and } B) &= P(B \mid A)P(A) \end{aligned}$$

$$\begin{aligned} P(A \mid B)P(B) &= P(B \mid A)P(A) \\ P(A \mid B) &= \frac{P(B \mid A)P(A)}{P(B)} \end{aligned}$$

Likelihood

Likelihood usually refers to the probability of the data D given a model $\mathcal{M}$ and/or parameters.

$$P(D \mid \mathcal{M})$$

In the case of a model with continuous probabilities, the likelihood is hte density evaluated at the data:

$$\prod_{d \in D} f_{\mathcal{M}}(d)$$

(assuming independent data)

Likelihood

Likelihood usually refers to the probability of the data D given a model $\mathcal{M}$ and/or parameters.

$$P(D \mid \mathcal{M})$$

In the case of a model with continuous probabilities, the likelihood is hte density evaluated at the data:

$$\prod_{d \in D} f_{\mathcal{M}}(d)$$

(assuming independent data)

$$P(\mathcal{M} \mid D) = \frac{P(D \mid \mathcal{M}) P(\mathcal{M})}{P(D)}$$

A case study

How does one learn categories and use that information to reason about new items?

- Prototype theory
  - Store a representation of the average of all training examples
  - Categorize new stimuli based on minimizing distance to stored averages

A case study

How does one learn categories and use that information to reason about new items?

- Prototype theory
  - Store a representation of the average of all training examples
  - Categorize new stimuli based on minimizing distance to stored averages
- Exemplar theory
  - Store a all training examples
  - Categorize new stimuli based on minimizing distance to stored examples

Formalizing prototype theory

First, a choice rule:

$$p(r_A \mid \mathbf{S}) = \frac{e^{br_A}}{e^{br_A} + e^{br_B}}$$

Formalizing prototype theory

First, a choice rule:

$$p(r_A \mid \mathbf{S}) = \frac{e^{br_A}}{e^{br_A} + e^{br_B}}$$

Second, a similarity metric: E.g., Euclidean distance.

$$P(A|S) = \frac{e^{br_A}}{e^{br_A} + e^{br_B}}$$

$$
\begin{aligned}
P(A|S) &= \frac{e^{br_A}}{e^{br_A} + e^{br_B}} \\
&= \frac{e^{br_A}}{e^{br_A} + e^{br_B}} \frac{e^{-br_A}}{e^{-br_A}} \\
&= \frac{1}{1 + e^{b(r_B - r_A)}}
\end{aligned}
$$

$$
\begin{aligned}
P(A|S) &= \frac{e^{br_A}}{e^{br_A} + e^{br_B}} \\
&= \frac{e^{br_A}}{e^{br_A} + e^{br_B}} \frac{e^{-br_A}}{e^{-br_A}} \\
&= \frac{1}{1 + e^{b(r_B - r_A)}}
\end{aligned}
$$

$$
r_B - r_A = \sqrt{(\mu_{x,B} - x)^2 + (\mu_{y,B} - y)^2} - \sqrt{(\mu_{x,A} - x)^2 + (\mu_{y,A} - y)^2}
$$

Code

and eventually the slides.

https://github.com/jhoupt/IK2019

https://www.psychopy.org/

Outline for section 2

What is ideal observer analyis?

An ideal observer:

- Best possible performance on a task.
- Quantification of the task relevant information available.
- Allows a relative measure of sub-ideal (e.g., human observer) task efficiency

Why do ideal observer analysis?

1. Determine what variation in performance is due to task properties and
   1.1 Allows us to focus theory building on the variation that is not due to task information content
2. Efficiency parcels out task and information specifics, so it can be compared across tasks

R	Response
r_i	ith response option
$\mathbf{S}$	Stimulus

R	Response
r_i	ith response option
$\mathbf{S}$	Stimulus

$$R = \arg\max_i \left[p(r_i|\mathbf{S}) \right]$$

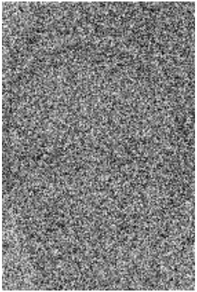
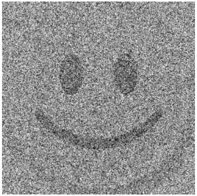
R	Response
r_i	ith response option
$\mathbf{S}$	Stimulus

$$
\begin{aligned}
R &= \arg\max_i \left[p(r_i|\mathbf{S})\right] \\
p(r_i|\mathbf{S}) &= \frac{p(\mathbf{S}|r_i)p(r_i)}{p(\mathbf{S})}
\end{aligned}
$$

$$
\begin{array}{cc}
R & \text{Response} \\
r_i & i\text{th response option} \\
\mathbf{S} & \text{Stimulus}
\end{array}
$$

$$
\begin{aligned}
R &= \arg\max_i \left[p(r_i|\mathbf{S})\right] \\
p(r_i|\mathbf{S}) &= \frac{p(\mathbf{S}|r_i)p(r_i)}{p(\mathbf{S})} \\
R &= \arg\max_i \left[p(\mathbf{S}|r_i)p(r_i)\right]
\end{aligned}
$$

An in depth example

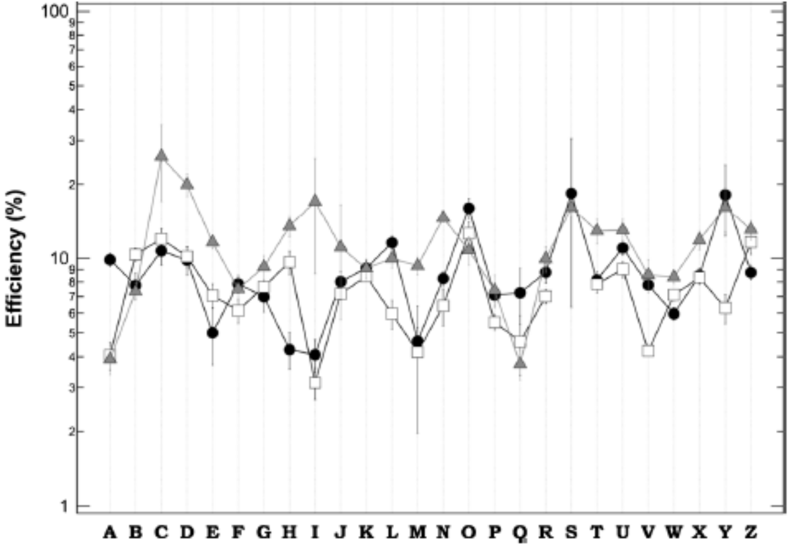
$\mathbf{S}$	Stimulus
$\mathbf{I}$	Clean Image
l	Contrast (signal) level
σ	Noise level
ϵ	$\sim \mathcal{N}(0, \sigma)$

$$S = (I/255 - 0.5) * l + \epsilon$$

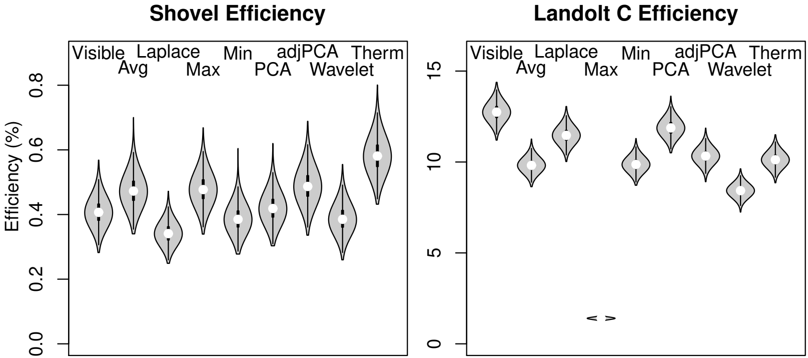
Python interlude

Survey of other applications

Eidels & Gold (2014). Measuring single-item identification efficiencies for letters and 3-D objects.



Houpt & Bittner (2018). Analyzing thresholds and efficiency with hierarchical Bayesian logistic regression.


Shovel Efficiency
Landolt C Efficiency
Efficiency (%)
Visible
Avg
Laplace
Max
Min
PCA
adjPCA
Wavelet
Therm
0.0
0.2
0.4
0.6
0.8
0
5
10
15

Classification Images

- Do people use the wrong or incomplete information, or just not use the information well enough?

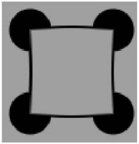
Classification Images

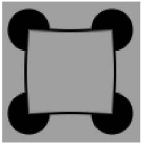
- Do people use the wrong or incomplete information, or just not use the information well enough?
- Examine noise patterns: noise when you responsd happy is more likely to be "happy"
- Assuming Gaussian white noise:

$$C = (\hat{\epsilon}_{H:H} + \hat{\epsilon}_{S:H}) - (\hat{\epsilon}_{H:S} + \hat{\epsilon}_{S:S})$$

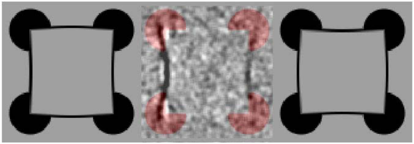
Python interlude

Sekuler, Gold, Murray & Bennett (2000). Visual completion of partly occluded objects: Insights from behavioral studies

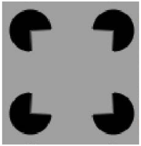


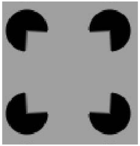


Sekuler, Gold, Murray & Bennett (2000). Visual completion of partly occluded objects: Insights from behavioral studies

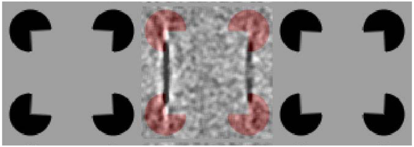


Sekuler, Gold, Murray & Bennett (2000). Visual completion of partly occluded objects: Insights from behavioral studies

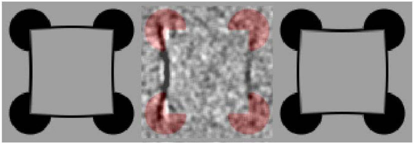


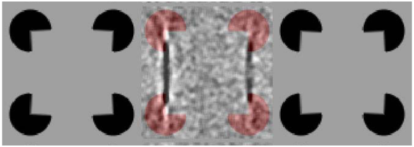


Sekuler, Gold, Murray & Bennett (2000). Visual completion of partly occluded objects: Insights from behavioral studies



Sekuler, Gold, Murray & Bennett (2000). Visual completion of partly occluded objects: Insights from behavioral studies





Outline for section 3

Modeling the cognitive tools

- Multidimensional scaling
- Categorization: The Generalized Context Model
- Decision Making: From Expected Utility to Decision Field Theory; Quantum, etc.

- Goal: A spatial representation of the relationships among stimuli or categories.

The Theory

- Exemplar theory
  - Store a all training examples
  - Categorize new stimuli based on minimizing distance to stored examples
- Distance between exemplars is the same for categorization as it is for perception.

The Model: Generalized Context Model

$$\begin{aligned} r_j &: \text{response} \\ S_i &: \text{stimulus} \\ b_j &: \text{bias} \\ \eta_{ij} = f(d_{ij}) &: \text{montonic function of psychological distance} \end{aligned}$$

$$P(r_j \mid S_i) = \frac{b_j \eta_{ij}}{\sum_{k=1}^{n} b_k \eta_{ik}}$$

The Model

$$
\begin{aligned}
C_J &: \text{collections of exemplars in category J} \\
r_j &: \text{response} \\
S_i &: \text{stimulus} \\
b_j &: \text{bias} \\
\eta_{ij} = f(d_{ij}) &: \text{montonic function of psychological distance}
\end{aligned}
$$

$$P(R_j \mid S_i) = \frac{b_J \sum_{j \in C_J} \eta_{ij}}{\sum_{K=1}^{m} b_K \sum_{k \in C_K} \eta_{ik}}$$

Theory

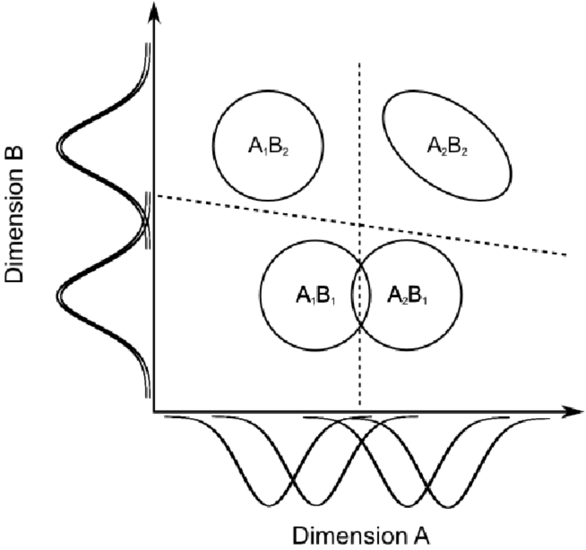
- Store a representation of the average of all training examples
- Categorize new stimuli based on minimizing distance to stored averages

General Recognition Theory

Generalize the signal detection model to multidimensional classificatons

General Recognition Theory

Generalize the signal detection model to multidimensional classificatons



Theory

People are rational.

Model I.A: Expected Value

$$\mathrm{E}V = \sum_i v_i p_i$$

Model I.A: Expected Value

$$\mathrm{E} V = \sum_i v_i p_i$$

Choice:

1. Keep your 1M€
2. 2^n€ where n is the number of coin flips before a heads.

Model I.B: Expected Utility Theory

- Completeness: For any pair of lotteries either only one is preferred to the other or they are preferred to each other: $A \succcurlyeq B$, $B \succcurlyeq A$, or both.

Model I.B: Expected Utility Theory

- Completeness: For any pair of lotteries either only one is preferred to the other or they are preferred to each other: $A \succcurlyeq B, B \succcurlyeq A$, or both.
- Transitivity: for any three lotteries, A, B, C, if $A \succcurlyeq B$ and $B \succcurlyeq C$, the $A \succcurlyeq C$.

Model I.B: Expected Utility Theory

- Completeness: For any pair of lotteries either only one is preferred to the other or they are preferred to each other: $A \succcurlyeq B, B \succcurlyeq A$, or both.
- Transitivity: for any three lotteries, A, B, C, if $A \succcurlyeq B$ and $B \succcurlyeq C$, the $A \succcurlyeq C$.
- Continuity: For all lotteries A, B, C such that $A \succcurlyeq B$ and $B \succcurlyeq C$, there is some p that mixes A and C such that $(pA, (1-p)C) \sim B$ (where indicated indifference).

Model I.B: Expected Utility Theory

- Completeness: For any pair of lotteries either only one is preferred to the other or they are preferred to each other: $A \succcurlyeq B, B \succcurlyeq A$, or both.
- Transitivity: for any three lotteries, A, B, C, if $A \succcurlyeq B$ and $B \succcurlyeq C$, the $A \succcurlyeq C$.
- Continuity: For all lotteries A, B, C such that $A \succcurlyeq B$ and $B \succcurlyeq C$, there is some p that mixes A and C such that $(pA, (1-p)C) \sim B$ (where indicated indifference).
- Independence: For all lotteries A, B, C, if $A \succcurlyeq B$ then $(pA, (1-p)C) \succcurlyeq (B, p; C, 1-p)$ for all p.

$$\mathrm{E}U = \sum_i u(v_i) p_i$$

Allais Paradox

Choice 1 :

1. Win 1M€
2. Win 5M€ with probaiblity 0.1, 1M€ with probability 0.89, or 0 with probability 0.01.

Allais Paradox

Choice 2 :

1. Win 1M€ with probability 0.11 or 0 with probability 0.89
2. Win 5M€ with probability 0.1 or 0 with probability 0.9.

1. $(y, p; c, 1-p)$
2. $(A, p; c, 1-p)$, where $A = (x, q; 0, 1-q)$ $q \in (0, 1)$

1. $(y, p; c, 1-p)$
2. $(A, p; c, 1-p)$, where $A = (x, q; 0, 1-q)$ $q \in (0, 1)$

$x = 5\text{M€}, y = 1\text{M€}, p = 0.11, q = {}^{10}\!/_{11}$
Choice 1 $c = 1\text{M}$:

1. $(1\text{M€}, 1)$
   $= (y = 1\text{M}, p = 0.11; c = 1\text{M}, 1-p = 0.89)$
2. $(5\text{M€},\ 0.1;\ 1\text{M€}, 0.89; 0, 0.01)$
   $= ([x = 5\mathrm{M}, p = {}^{10}\!/_{11}; 0, 1 - {}^{10}\!/_{11}], p = 0.11; c = 1\text{M}, 1-p = 0.89)$

1. $(y, p; c, 1-p)$
2. $(A, p; c, 1-p)$, where $A = (x, q; 0, 1-q)$ $q \in (0,1)$

$x = 5M€, y = 1M€, p = 0.11, q = {}^{10}/_{11}$
Choice 1 $c = 1M$:

1. $(1M€, 1)$
   $= (y = 1M, p = 0.11; c = 1M, 1 - p = 0.89)$
2. $(5M€,\ 0.1;\ 1M€, 0.89; 0, 0.01)$
   $= ([x = 5\mathrm{M}, p = {}^{10}/_{11}; 0, 1 - {}^{10}/_{11}], p = 0.11; c = 1M, 1 - p = 0.89)$

Choice 2 $c = 0$:

1. $(1M€, 0.11; 0, 0.89)$
   $= (y = 1M, p = 0.11; c = 0, 1 - p = 0.89)$
2. $(5M€, 0.1; 0, 0.9)$
   $= ([x = 5\mathrm{M}, p = {}^{10}/_{11}; 0, 1 - {}^{10}/_{11}], p = 0.11; c = 0, 1 - p = 0.89)$.

Imagine that the E.U. is preparing for the outbreak of an unusual disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

- If program A is adopted, 200 people will be saved.
- If program B is adopted, there is a 1/3probability that 600 people will be saved, and a 2/3 probability that no people will be saved.

- If program C is adopted, 400 people will die
- If program D is adopted, there is a $1/3$ probability that nobody will die, and a $2/3$ probability that 600 people will die

Imagine that the E.U. is preparing for the outbreak of an unusual disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

- If program A is adopted, 200 people will be saved.
- If program B is adopted, there is a 1/3probability that 600 people will be saved, and a 2/3 probability that no people will be saved.

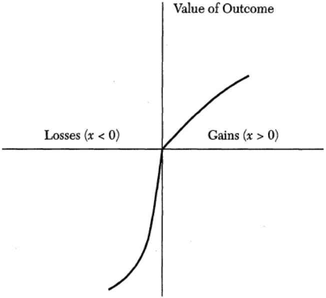
- If program C is adopted, 400 people will die
- If program D is adopted, there is a 1/3 probability that nobody will die, and a 2/3 probability that 600 people will die

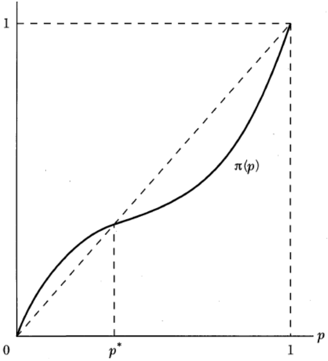
Theory II

People are not computers, but still do something reasonable.

Model II: Cumulative Prospect Theory

$$V = \sum_i v_i \pi_i$$





Theory

- Completeness: For any pair of lotteries either only one is preferred to the other or they are preferred to each other: $A \succcurlyeq B, B \succcurlyeq A$, or both.
- ~~Transitivity: for any three lotteries,~~ $\cancel{A, B, C}$~~, if~~ $\cancel{A \succcurlyeq B}$ ~~and~~ $\cancel{B \succcurlyeq C}$~~, the~~ $\cancel{A \succcurlyeq C}$~~.~~ Tversky, A. (1969). Intransitivity of preferences. Psychological review, 76(1), 31.
- Continuity: For all lotteries A, B, C such that $A \succcurlyeq B$ and $B \succcurlyeq C$, there is some p that mixes A and C such that $(pA, (1-p)C) \sim B$ (where indicated indifference).
- ~~Independence: For all lotteries~~ $\cancel{A, B, C}$~~, if~~ $\cancel{A \succcurlyeq B}$ ~~then~~ $\cancel{(pA, (1-p)C) \succcurlyeq (B, p; C, 1-p)}$ ~~for all~~ $\cancel{p}$~~.~~

Model III: Decision Field Theory

- Subjective expected utility theory:

$$
\begin{aligned}
V &= \sum_i v_i \pi_i \\
d &= V_A - V_B
\end{aligned}
$$

Model III: Decision Field Theory

- Subjective expected utility theory:

$$\begin{aligned} V &= \textstyle\sum_i v_i \pi_i \\ d &= V_A - V_B \end{aligned}$$

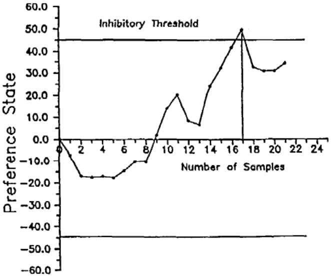
- Random SEU Theory (W depends on attention)

$$\begin{aligned} V &= \textstyle\sum_i W_i v_i \\ d &= \mathrm{E}\left[V_A - V_B\right] \\ P &= d + \epsilon \end{aligned}$$

Model III: Decision Field Theory

- Sequential SEU

$$
\begin{aligned}
V(t) &= \textstyle\sum_i W_i(t) v_i \\
d(t) &= \mathrm{E}\left[V_A(t) - V_B(t)\right] \\
P(t) &= P(t-1) + d(t) + \epsilon
\end{aligned}
$$



Model III: Decision Field Theory

- Sequential SEU with bias and decay

$$
\begin{aligned}
V(t) &= \textstyle\sum_i W_i(t) v_i \\
P(t) &= (1-s)P(t-1) + d(t) + \epsilon_t \\
P(t) &= \textstyle\sum_k (1-s)^{t-k} d(k) + \epsilon_k
\end{aligned}
$$

Model III: Decision Field Theory

- Sequential SEU with bias and decay and goal gradients

$$
\begin{aligned}
V(t) &= \sum_i f_{\text{sign } v_i} v(\theta - P(t)) W_i(t) v_i \\
P(t) &= (1-s)P(t-1) + d(t) + \epsilon_t \\
P(t) &= \sum_k (1-s)^{t-k} d(k) + \epsilon_k
\end{aligned}
$$

Theory

People's behavior is not well reporesented by classical probability theory:

Classical

- For all A, $P(A) \in \mathbb{R}, P(A) > 0$.
- $P(\Omega) = 1$
- If E_i are disjoint, $P\left(\cup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$
- Outcomes are elements of subsets of an overall sample space.
- Events are subsets of the overall sample space.
- Probabilities are functions defined on subsets.

Theory

People's behavior is not well reporesented by classical probability theory:

Classical

- For all A, $P(A) \in \mathbb{R}, P(A) > 0$.
- $P(\Omega) = 1$
- If E_i are disjoint, $P\left(\cup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$
- Outcomes are elements of subsets of an overall sample space.
- Events are subsets of the overall sample space.
- Probabilities are functions defined on subsets.

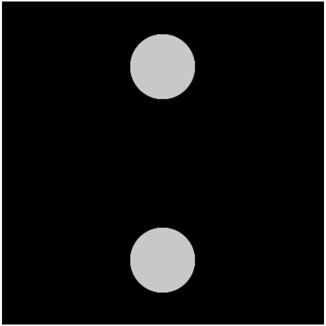
Quantum

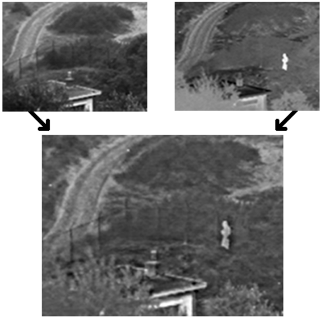
- Outcomes are vectors from an orthonormal set that spans a vector space.
- Events are subspaces of the overall vector space.
- Probabilities are the magnitude of the projection of the state onto an event.

Model IV: Quantum Decision Making

https://osf.io/rqj3b/

Outline for section 4





Theory Driven Methodology

- Systems Factorial Technology (SFT) is a framework for addressing the general question: how do different sources of information combine in mental processing?

Theory Driven Methodology

- Systems Factorial Technology (SFT) is a framework for addressing the general question: how do different sources of information combine in mental processing?
    - Are both sources used concurrently, or do we use one at a time?

Theory Driven Methodology

- Systems Factorial Technology (SFT) is a framework for addressing the general question: how do different sources of information combine in mental processing?
  - Are both sources used concurrently, or do we use one at a time?
  - How many sources are enough to respond?

Theory Driven Methodology

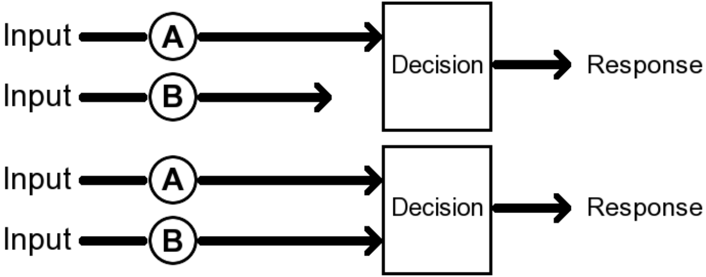
- Systems Factorial Technology (SFT) is a framework for addressing the general question: how do different sources of information combine in mental processing?
  - Are both sources used concurrently, or do we use one at a time?
  - How many sources are enough to respond?
  - Does knowledge of one source affect how we process another?

Theory Driven Methodology

- Systems Factorial Technology (SFT) is a framework for addressing the general question: how do different sources of information combine in mental processing?
    - Are both sources used concurrently, or do we use one at a time?
    - How many sources are enough to respond?
    - Does knowledge of one source affect how we process another?
    - Can we dedicate the same amount of resources to processing each source when there are more sources?

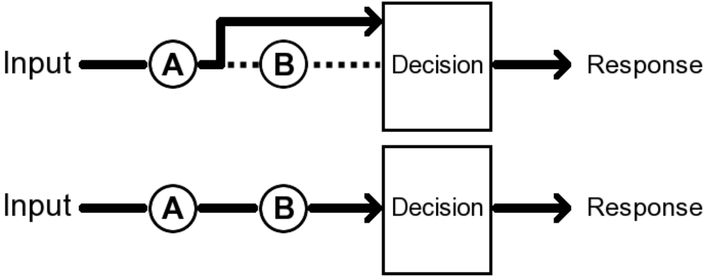
Architecture

- Are both sources used concurrently, or do we use one at a time?
  - Using sources concurrently: Parallel information processing.



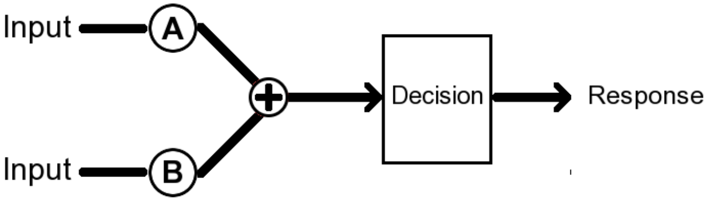
Architecture

- Are both sources used concurrently, or do we use one at a time?
  - Using sources one at a time: Serial information processing.



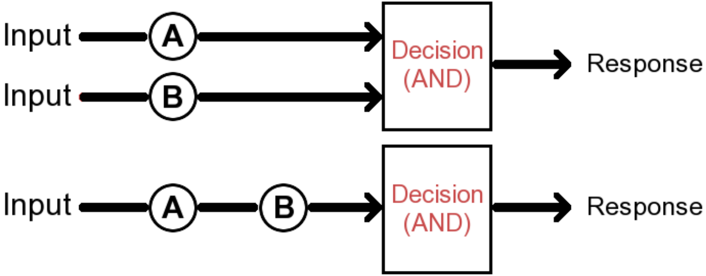
Architecture

- Are both sources used concurrently, or do we use one at a time?
  - Pooled information for a single detector: Coactive processing



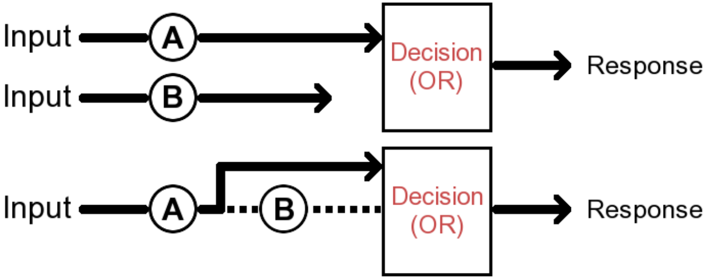
Stopping Rule

- How many sources are enough to respond?
  - All of them: Exhaustive processing (AND)



Stopping Rule

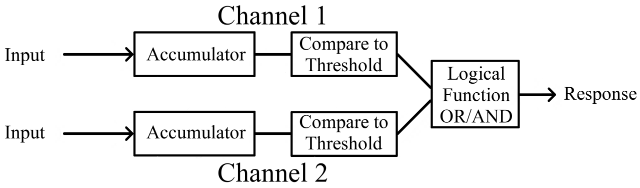
- How many sources are enough to respond?
  - Any of them: First-terminating processing (OR)
  - When there are more than two sources and not all sources are be required, but possibly more than one: Self-terminating.



Stochastic Dependence

- Does knowledge of one source affect how we process another?
  - Stochastic independence of the decision times.

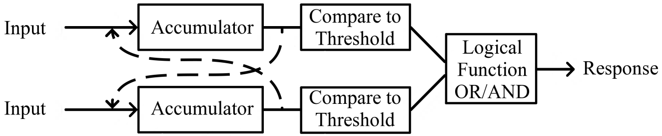
$$\Pr\{T_A \leq t_1 \text{ AND } T_B \leq t_2\} = \Pr\{T_A \leq t_1\}\Pr\{T_B \leq t_2\}$$



Stochastic Dependence

- Does knowledge of one source affect how we process another?
  - Stochastic dependence of the decision times.

$$\Pr\{T_A \leq t_1 \text{ AND } T_B \leq t_2\} \neq \Pr\{T_A \leq t_1\}\Pr\{T_B \leq t_2\}$$

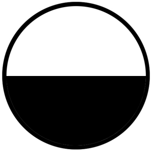


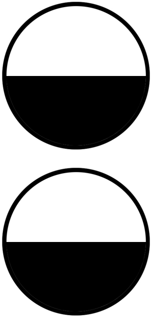
Workload Capacity

- Can we dedicate the same amount of resources to processing each source when there are more sources?
  - Fewer resources available for each process as the number of sources increases: Limited capacity.

Workload Capacity

- Can we dedicate the same amount of resources to processing each source when there are more sources?
  - Unchanged amount of resources available for each process as the number of sources increases: Unlimited capacity.

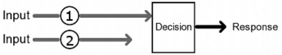




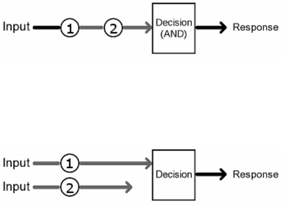
Workload Capacity

- Can we dedicate the same amount of resources to processing each source when there are more sources?
  - More resources available for each process as the number of sources increases: Super capacity.

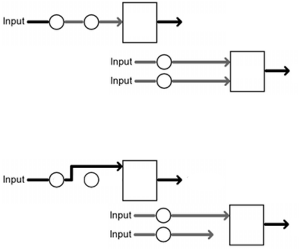
Architecture



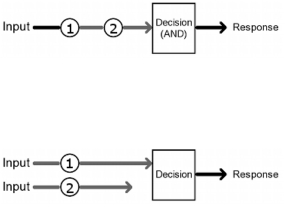
Architecture



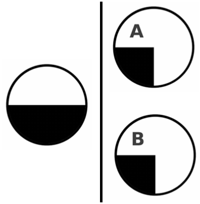
Stopping Rule



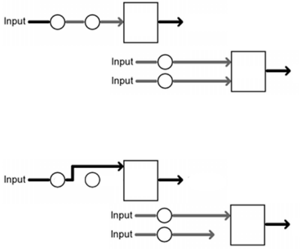
Architecture



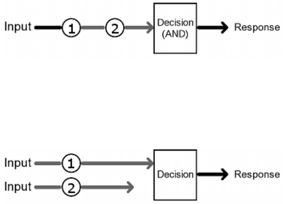
Workload Capacity



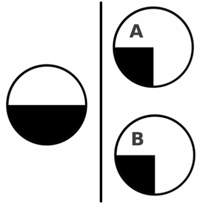
Stopping Rule



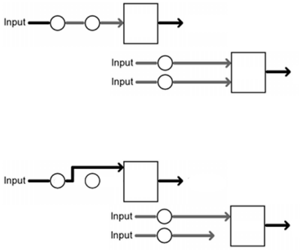
Architecture



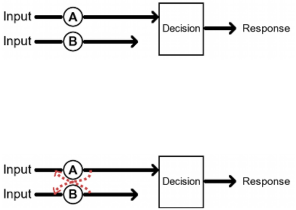
Workload Capacity



Stopping Rule



Dependence



Workload Capacity Coefficient

- Indicates changes in processing the parts due to increased workload.
- Fix a baseline for performance: UCIP model.

Workload Capacity Coefficient

- Indicates changes in processing the parts due to increased workload.
- Fix a baseline for performance: UCIP model.
  - Unlimited Capacity: No change due to increased workload

Workload Capacity Coefficient

- Indicates changes in processing the parts due to increased workload.
- Fix a baseline for performance: UCIP model.
  - Unlimited Capacity: No change due to increased workload
  - Independent: Stochastic independence

Workload Capacity Coefficient

- Indicates changes in processing the parts due to increased workload.
- Fix a baseline for performance: UCIP model.
  - Unlimited Capacity: No change due to increased workload
  - Independent: Stochastic independence
  - Parallel: Simultaneous processing

Workload Capacity Coefficient

- Unlimited Capacity: No change due to increased workload
- Independent: Stochastic independence
- Parallel: Simultaneous processing

First-Terminating (OR)

Workload Capacity Coefficient

- Unlimited Capacity: No change due to increased workload
- Independent: Stochastic independence
- Parallel: Simultaneous processing

First-Terminating (OR)

$$P\left(T_{AB} > t\right) = P\left(T_{A(B)} > t \mathrm{andT_{B(A)} > t}\right)$$

Workload Capacity Coefficient

- Unlimited Capacity: No change due to increased workload
- Independent: Stochastic independence
- Parallel: Simultaneous processing

First-Terminating (OR)

$$
\begin{aligned}
P\left(T_{AB} > t\right) &= P\left(T_{A(B)} > t \mathrm{and} \mathrm{T}_{\mathrm{B(A)}} > \mathrm{t}\right) \\
&= P\left(T_{A} > t \mathrm{and} \mathrm{T}_{\mathrm{B}} > \mathrm{t}\right)
\end{aligned}
$$

Workload Capacity Coefficient

- Unlimited Capacity: No change due to increased workload
- Independent: Stochastic independence
- Parallel: Simultaneous processing

First-Terminating (OR)

$$\begin{aligned} P\left(T_{AB} > t\right) &= P\left(T_{A(B)} > t \mathrm{and} \mathrm{T}_{\mathrm{B(A)}} > \mathrm{t}\right) \\ &= P\left(T_{A} > t \mathrm{and} \mathrm{T}_{\mathrm{B}} > \mathrm{t}\right) \\ &= P\left(T_{A} > t\right) P\left(T_{B} > t\right) \end{aligned}$$

Workload Capacity Coefficient

- Unlimited Capacity: No change due to increased workload
- Independent: Stochastic independence
- Parallel: Simultaneous processing

First-Terminating (OR)

$$
\begin{aligned}
P\left(T_{AB} > t\right) &= P\left(T_{A(B)} > t \mathrm{and T_{B(A)} > t}\right) \\
&= P\left(T_A > t \mathrm{and T_B > t}\right) \\
&= P\left(T_A > t\right) P\left(T_B > t\right) \\
S_{AB}(t) &= S_A(t) S_B(t)
\end{aligned}
$$

Workload Capacity Coefficient

- Unlimited Capacity: No change due to increased workload
- Independent: Stochastic independence
- Parallel: Simultaneous processing

First-Terminating (OR)

$$\begin{aligned}
P\left(T_{AB} > t\right) &= P\left(T_{A(B)} > t \mathrm{and T_{B(A)} > t}\right) \\
&= P\left(T_A > t \mathrm{and T_B > t}\right) \\
&= P\left(T_A > t\right) P\left(T_B > t\right) \\
S_{AB}(t) &= S_A(t) S_B(t) \\
\log\left[S_{AB}(t)\right] &= \log\left[S_A(t)\right] + \log\left[S_B(t)\right]
\end{aligned}$$

Workload Capacity Coefficient

- Unlimited Capacity: No change due to increased workload
- Independent: Stochastic independence
- Parallel: Simultaneous processing

First-Terminating (OR)

$$
\begin{aligned}
P\left(T_{AB} > t\right) &= P\left(T_{A(B)} > t \mathrm{and T_{B(A)}} > \mathrm{t}\right) \\
&= P\left(T_{A} > t \mathrm{and T_{B}} > \mathrm{t}\right) \\
&= P\left(T_{A} > t\right) P\left(T_{B} > t\right) \\
S_{AB}(t) &= S_{A}(t) S_{B}(t) \\
\log\left[S_{AB}(t)\right] &= \log\left[S_{A}(t)\right] + \log\left[S_{B}(t)\right] \\
H_{AB}(t) &= H_{A}(t) + H_{B}(t)
\end{aligned}
$$

Exhaustive (AND)

$$
\begin{aligned}
P\left(T_{AB} \leq t\right) &= P\left(T_{A(B)} \leq t \mathrm{andT_{B(A)}} \leq \mathrm{t}\right) \\
&= P\left(T_{A} \leq t \mathrm{andT_{B}} \leq \mathrm{t}\right) \\
&= P\left(T_{A} \leq t\right) P\left(T_{B} \leq t\right) \\
F_{AB}(t) &= F_{A}(t)F_{B}(t) \\
\log\left[F_{AB}(t)\right] &= \log\left[F_{A}(t)F_{B}(t)\right] \\
K_{AB}(t) &= K_{A}(t) + K_{B}(t)
\end{aligned}
$$

First-Terminating (OR)

$$\mathrm{C}_{\mathrm{or}}(t) = H_{AB}(t) - [H_A(t) + H_B(t)]$$

Exhaustive (AND)

$$\mathrm{C}_{\mathrm{and}}(t) = [K_A(t) + K_B(t)] - K_{AB}(t)$$

Test Statistic

OR Task

$$\hat{H}_X(t) = \sum_{t_j \in \{\mathrm{RT}(X) < t\}} \frac{1}{Y_{AB}(t_j)}$$

$$\hat{\mathrm{Var}}[H_X](t) = \sum_{t_j \in \{\mathrm{RT}(X) < t\}} \frac{1}{Y_{AB}{}^2(t_j)}$$

$$\hat{C}_{\mathrm{or}}(t) = \hat{H}_{AB}(t) - \hat{H}_A(t) - \hat{H}_B(t)$$

$$\hat{\mathrm{Var}}[C_{\mathrm{or}}](t) = \hat{\mathrm{Var}}[H_{AB}](t) + \hat{\mathrm{Var}}[H_A](t) + \hat{\mathrm{Var}}[H_B](t)$$

Test Statistic

OR Task

$$\hat{H}_X(t) = \sum_{t_j \in \{\mathrm{RT}(X) < t\}} \frac{1}{Y_{AB}\left(t_j\right)}$$

$$\hat{\mathrm{Var}}\left[H_X\right](t) = \sum_{t_j \in \{\mathrm{RT}(X) < t\}} \frac{1}{{Y_{AB}}^2(t_j)}$$

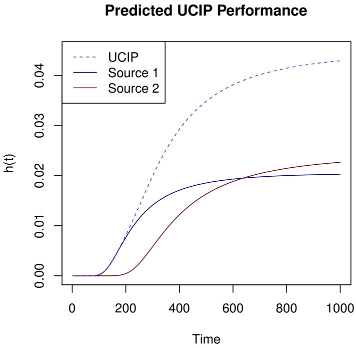
$$\hat{C}_{\mathrm{or}}(t) = \hat{H}_{AB}(t) - \hat{H}_A(t) - \hat{H}_B(t)$$

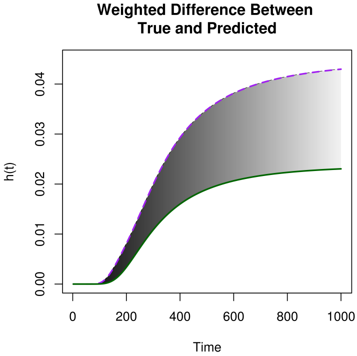
$$\hat{\mathrm{Var}}\left[C_{\mathrm{or}}\right](t) = \hat{\mathrm{Var}}\left[H_{AB}\right](t) + \hat{\mathrm{Var}}\left[H_A\right](t) + \hat{\mathrm{Var}}\left[H_B\right](t)$$

$$U_{\mathrm{or}} = \frac{Z_{\mathrm{or}}(t)}{\mathrm{Var}\left[Z_{\mathrm{or}}(t)\right]} \sim \mathcal{N}(0,1)$$

Test Statistic

OR Task

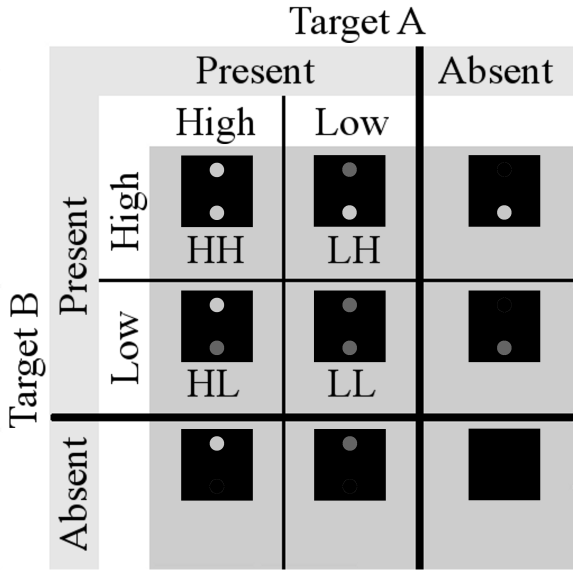




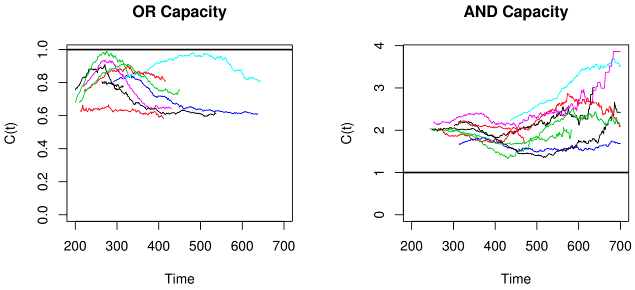
Why use $C(t)$?

- Provides a yardstick to measure changes in performance due to workload across types of stimuli.
- Response time based measure.
  - More sensitive to many model characteristics than accuracy.
  - Dynamic measure: The value of $C(t)$ can change across time.

Task


Target A
Present
Absent
High
Low
Target B
Present
Absent
High
Low
HH
LH
HL
LL

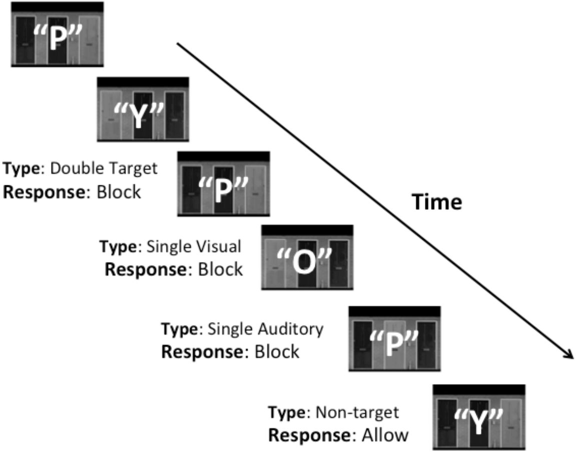
Results



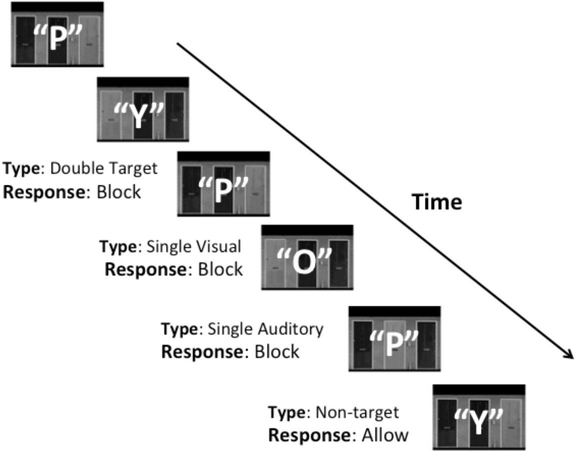
- All participants are limited on OR and super on AND

Supertaskers

The Gatekeeper Task



The Gatekeeper Task

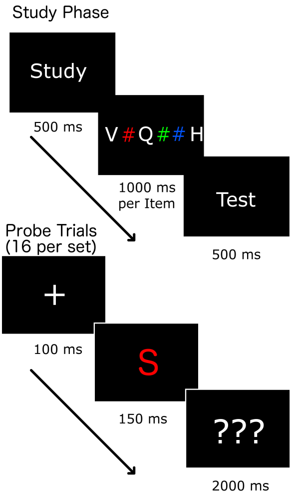


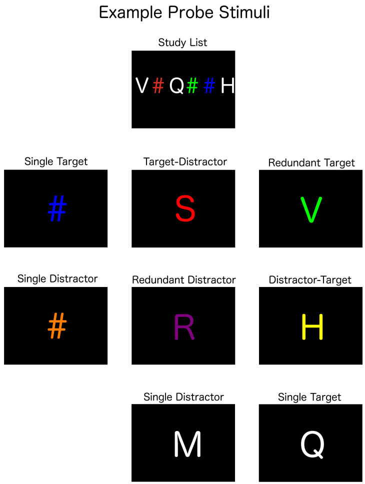
- 20 Participants of 272 were not significantly limited-capacity (potential supertaskers)

Externalizing Psychopathology

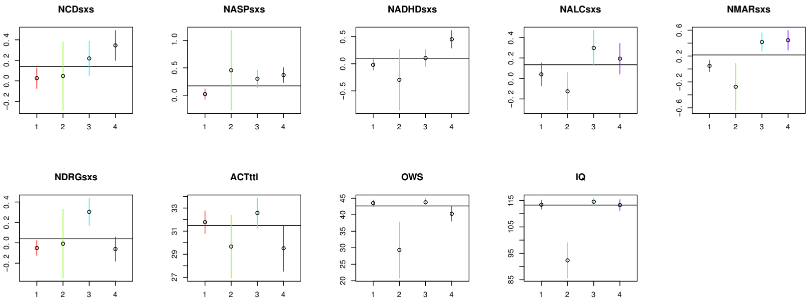
Imagens Evangelicas @ Flikr

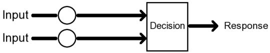
The Redundant Memory Probe Task

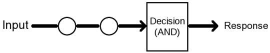


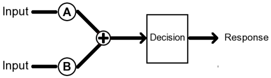


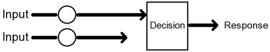
The Redundant Memory Probe Task

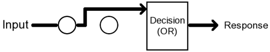



Input
Input
Decision
Response

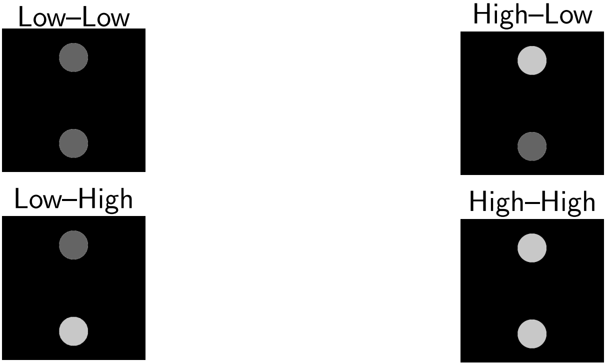

Input
Decision (AND)
Response


Input
A
Input
B
Decision
Response

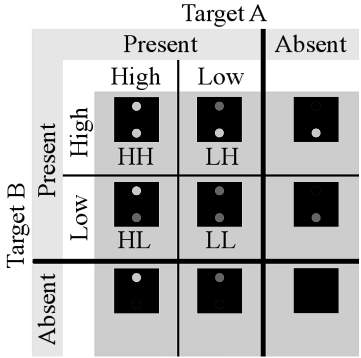

Input
Input
Decision
Response


Input
Decision (OR)
Response

- When there are multiple sources, what is the overall effect of speeding up and slowing down the processing of some subset of those sources?

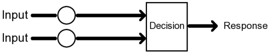


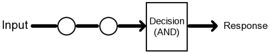
The Mean Interaction Contrast

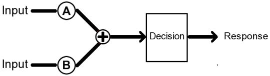


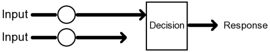
$$\mathrm{MIC}(t) = [\mathrm{M_{LL}}(t) - \mathrm{M_{LH}}(t)] - [\mathrm{M_{HL}}(t) - \mathrm{M_{HH}}(t)]$$

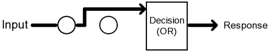
Here, the subscripts indicate the salience of each source of information.


Input
Input
Decision
Response


Input
Decision
(AND)
Response


Input
A
Input
B
Decision
Response


Input
Input
Decision
Response


Input
Decision
(OR)
Response

$\mathrm{MIC} < 0$ $\mathrm{MIC} = 0$

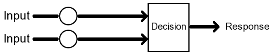
$\mathrm{MIC} > 0$

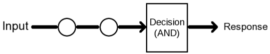
$\mathrm{MIC} > 0$ $\mathrm{MIC} = 0$

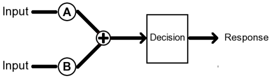
The Survivor Interaction Contrast

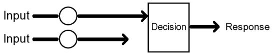
- The SIC is a measure of interaction between the salience manipulations.
  - Instead of just using the mean time, we use the survivor function: $S(t) = \Pr\{T > t\} = 1 - F(t)$.

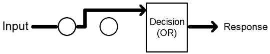
$$\mathrm{SIC}(t) = [\mathrm{S}_{\mathrm{LL}}(t) - \mathrm{S}_{\mathrm{LH}}(t)] - [\mathrm{S}_{\mathrm{HL}}(t) - \mathrm{S}_{\mathrm{HH}}(t)]$$


Input
Input
Decision
Response

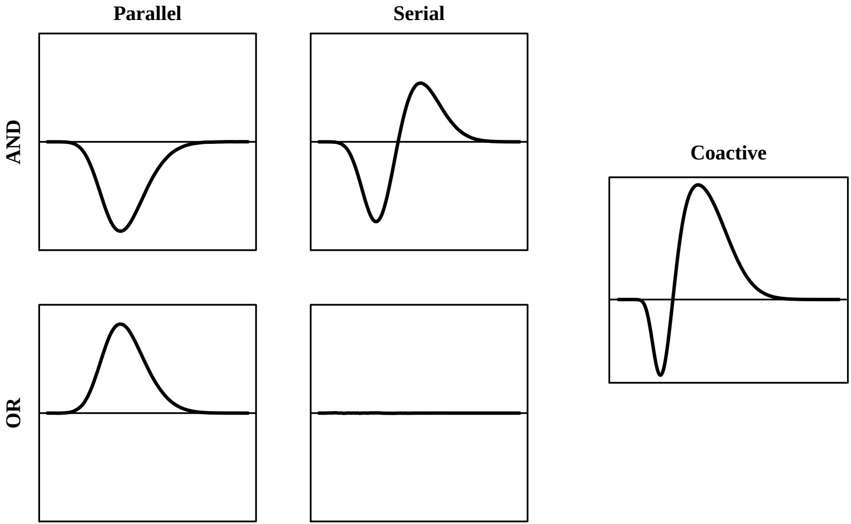

Input
Decision
(AND)
Response

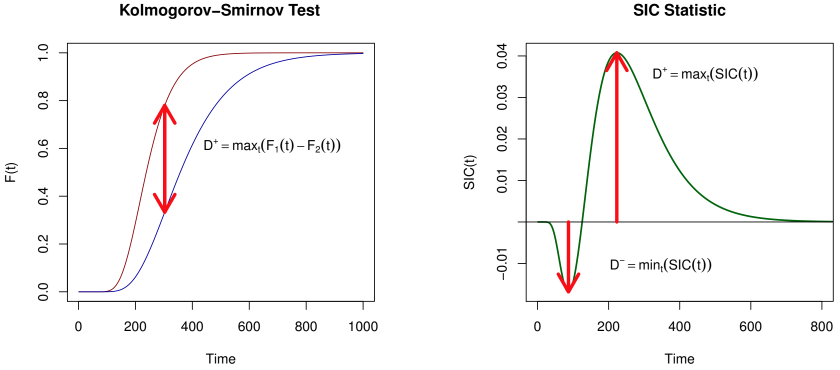

Input
A
Input
B
Decision
Response


Input
Input
Decision
Response


Input
Decision
(OR)
Response

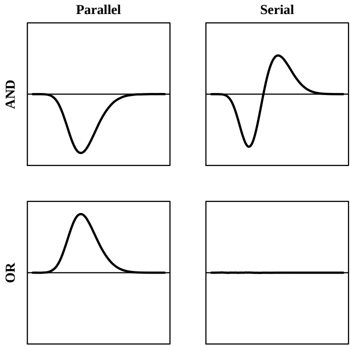
The Survivor Interaction Contrast





$$\lim_{N\to\infty} \Pr\{\sqrt{N}D^+ \geq x\} = \Pr\{\sqrt{N}D^- \geq x\} = e^{-2x^2}$$

$$N_{\mathrm{KS}} = \frac{1}{1/m + 1/n} \qquad N_{\mathrm{SIC}} = \frac{1}{1/k + 1/l + 1/m + 1/n}$$

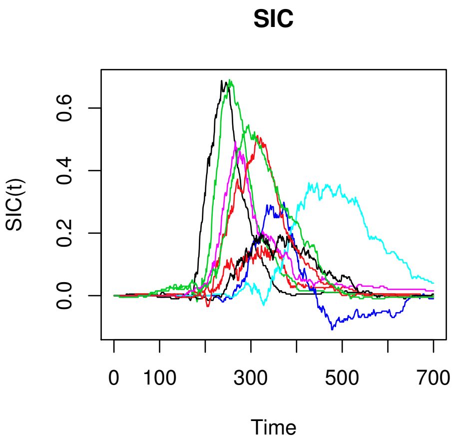


Model	$\hat{D}^{+}$	$\hat{D}^{-}$	MIC
Serial-OR	$\emptyset$	$\emptyset$	$\emptyset$
Serial-AND	$\checkmark$	$\checkmark$	$\emptyset$
Parallel-OR	$\checkmark$	$\emptyset$	$\checkmark$
Parallel-AND	$\emptyset$	$\checkmark$	$\checkmark$
Coactive	$\checkmark$	$\checkmark$	$\checkmark$

$\checkmark$: Reject null hypothesis

$\emptyset$: Fail to reject null hypothesis

Results



- D^+ significant for 9 of 10 participants
- D^- not significant for any participant

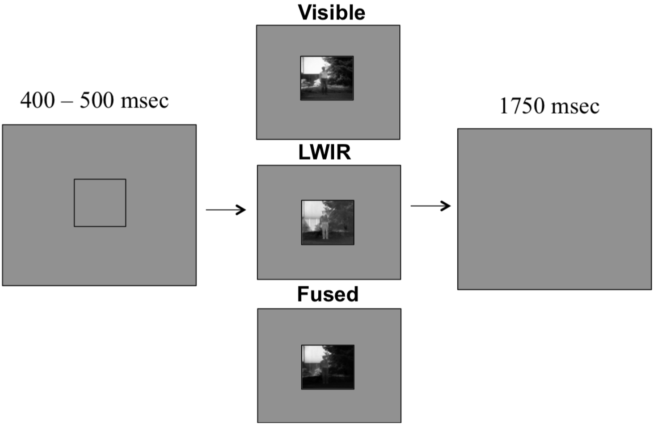
Multisensor Image Fusion


Visible

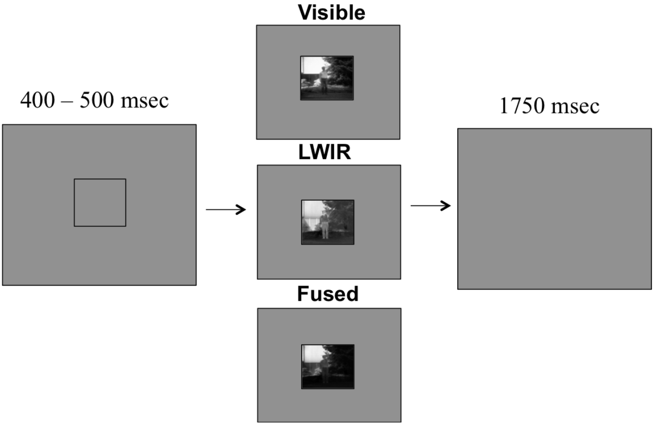

LWIR


Fused

Multisensor Image Fusion

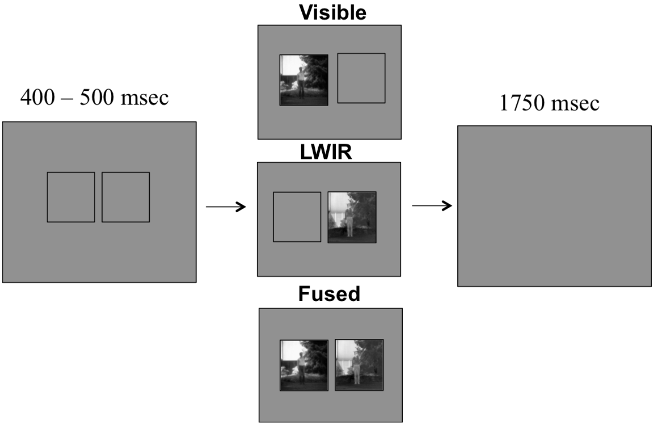


Multisensor Image Fusion

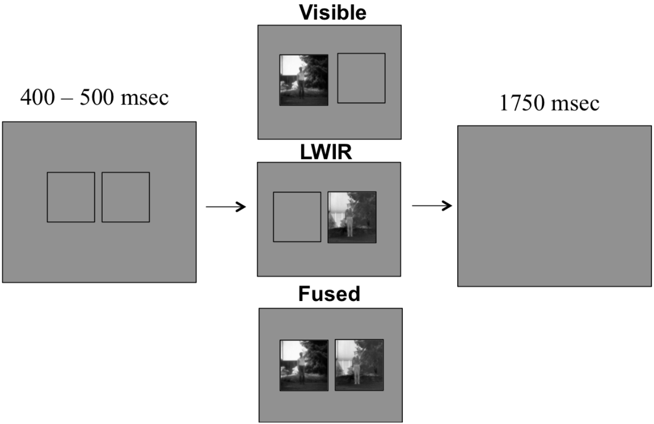


- 10/10 Participants Limited Capacity

Multisensor Image Fusion

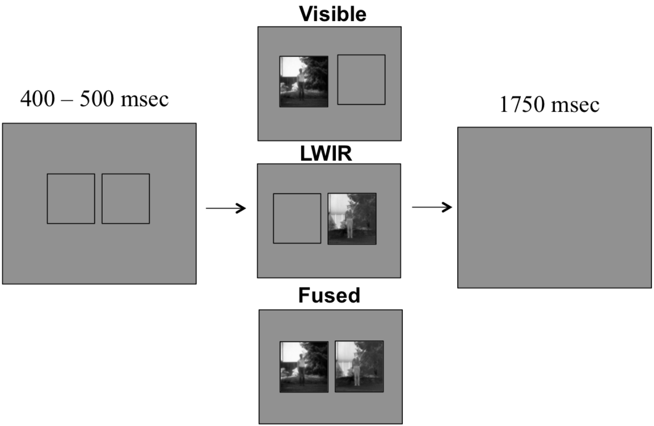


Multisensor Image Fusion



- 5/9 Limited, 4/9 Unlimited
- 9/9 Had higher capacity with side by side information

Multisensor Image Fusion



- 3 Used Parallel-OR, 2 used Parallel-AND, 2 Serial-OR

Conclusions

- SFT is a framework for addressing the general question: how do different sources of information combine in mental processing?
  - Capacity Coefficient for measuring changes across workload.
  - Survivor Interaction Contrast for measuring architecture and stopping rule.