# Information foraging, interface design, and human factors

Hyeonsu Kang and Enhao Cui

# *Information Foraging*

Peter Pirolli and Stuart Card

# Learning goals

1. Be able to explain the context and implementation of Information Foraging Theory
    a. Information fragmentation, poverty of attention, cognitive architectures, spreading activation
    b. Information patch foraging, Charnov’s Marginal Value Theorem
    c. Information Scent
2. Understand what approach did the author take to evaluate the theory
3. Think about Information Foraging Theory’s application in the context of user interaction and design


**Max [R = energy/time]**

**Max [R = useful info/time]**

Maybe the way we seek information is an example of *exaptation** from food foraging. Can we model it using Optimal Foraging Theory?

* a term used in evolutionary biology to describe a trait that has been co-opted for a use other than the one for which natural selection has built it.

*Source: Peter Pirolli*

# Information overload and poverty of attention

"What information consumes is rather obvious: it consumes the attention of its recipients. Hence **a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently** among the overabundance of information sources that might consume it" -- Herbert Simon

The important question lies not only in how to generate good information but also in **how to design a good information architecture** that enables people an easy access to useful information.

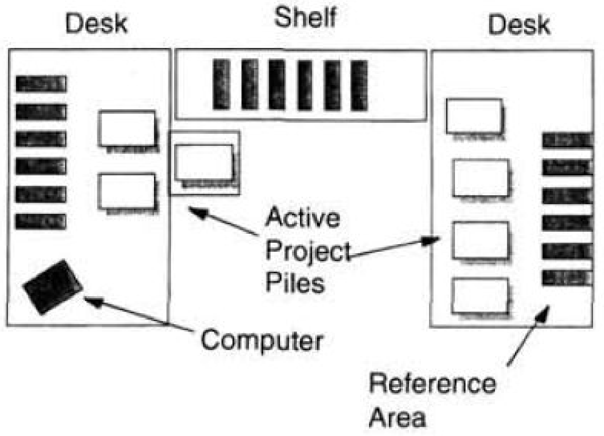



*Figure 2.* Schematic layout of the business intelligence office.

Because the environment is malleable for humans in information foraging (unlike animals in the wilderness foraging for food), a **predictive model** can help answer two design questions:

- **How can we better-shape or adapt ourselves to info. environment?**
- **How can the info. environment be designed better to match human skills/strategies/needs?**

Q. As a group of 2 - 3, think about a hypothetical situation in which an individual's planning a trip to New York. Then, as a group, discuss about how you would carry out this planning. Specifically ground your discussion around *information foragers*, the *information*, and the *strategy*:

Information foragers: what do you want to know about?

Information: what kind of information are you looking for? Where is that available, how can you obtain it?

Strategy: how would you strategize your planning? For example, would you collect as much info as possible first then filter out?

The model should be able to describe

- **“Information patch” foraging**
- **Relevance assessment by “information scent”**

# Background: The possibility of unifying cognitive **theories**?

| TIME (sec) | TIME (common units) | ACTION | MEMORY | THEORY |
| --- | --- | --- | --- | --- |
| $10^9$ | (decades) | Technology | Culture | |
| $10^8$ | (years) | System | Development | Social and |
| $10^7$ | (months) | Design | Education | Organizational |
| $10^6$ | (weeks) | Task | Education | |
| $10^5$ | (days) | Task | Skill | |
| $10^4$ | (hours) | Task | Skill | Bounded |
| $10^3$ | (ten mins) | Task | LTM | Rationality |
| $10^2$ | (minutes) | Task | LTM | |
| 10 | (ten secs) | Unit task | LTM | |
| 1 | (secs) | Operator | STM | Psychological |
| $10^{-1}$ | (tenths) | Cycle time | Buffers | |
| $10^{-2}$ | (centisecs) | Signal | Integration | Neural and |
| $10^{-3}$ | (millisecs) | Pulse | Summation | Biochemical |

***The Prospects for Psychological Science in Human-Computer Interaction***

Allen Newell and Stuart Card

# Background: The possibility of unifying cognitive **theories**?

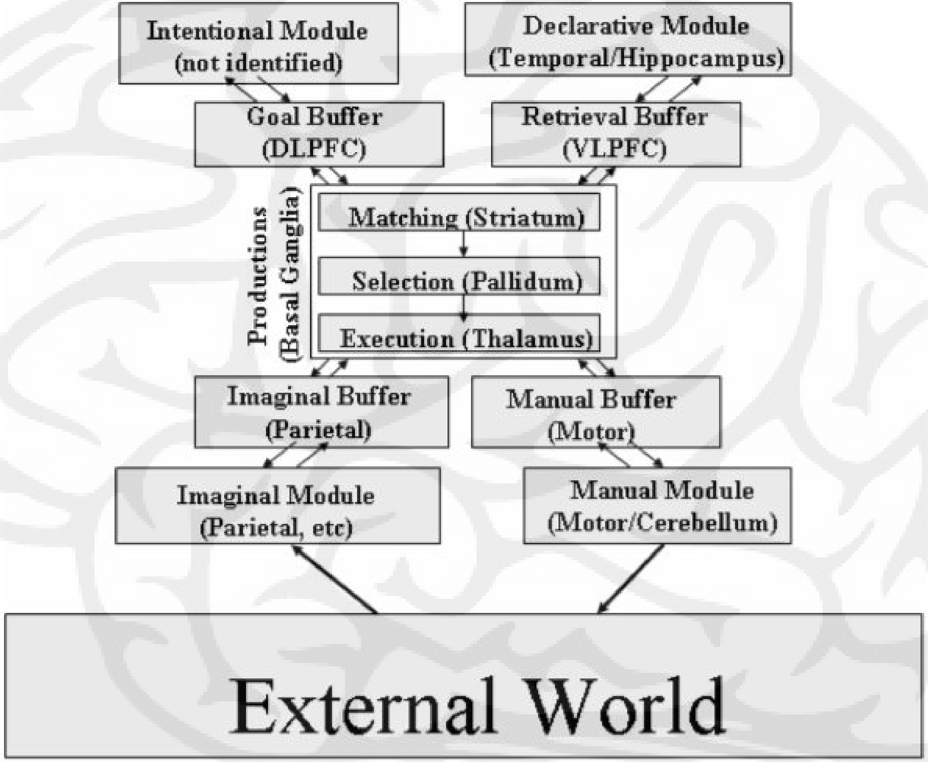
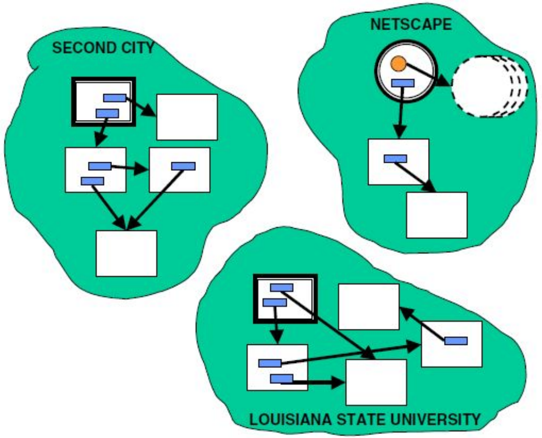



Figure 1. ACT-R Architecture

Growing interest in developing cognitive architectures:
Ex. ACT-R architecture, SOAR

***The Newell Test for a theory of cognition***
John Anderson and Christian Lebiere

# Information patch foraging




"The difficulty in finding useful information related to the balkanization* of the Web structure"
"It is difficult to solve this [Web structure] **fragmentation** problem by designing an effective and efficient classification scheme, an alternative approach is to seek regularities in **user patterns** that can then be used to develop technologies for increasing the density of relevant data for users"

***Strong Regularities in World Wide Web Surfing***
Bernardo Huberman, Peter Pirolli et al.

* originally used to describe the process of fragmentation or division of a region or state into smaller regions or states that are often hostile or uncooperative with one another.

# Information patch foraging

## Holling's disc equation

$$R = \frac{G}{T_B + T_W}$$

**Max [R = useful info / time]**

## Table of notation

| | |
|---|---|
| G | the ratio of the total net amount of valuable info. gained |
| T_B | The total amount of time spent between-patches |
| T_W | The total amount of time spent within-patches foraging |

# Information patch foraging

## Holling's disc equation

$$R = \frac{G}{T_B + T_W}$$

The total # of processed patches

$$\lambda = 1/t_B \qquad G = \lambda T_B g \qquad T_W = \lambda T_B t_W$$

## Table of notation

| | |
|---|---|
| G | the ratio of the total net amount of valuable info. gained |
| T_B | The total amount of time spent between-patches |
| T_W | The total amount of time spent within-patches foraging |
| g | The avg. gain per patch |
| t_B | The avg. time between processing patches |
| t_W | The avg. time to process patches |

# Information patch foraging

Plugging in everything yields

$$R = \frac{\lambda T_B g}{T_B + \lambda T_B t_W} = \frac{\lambda g}{1 + \lambda t_W}$$

$$\lambda = 1/t_B$$

**Prevalence**

$$\pi = g/t_W.$$

**Profitability**

## Table of notation

| | |
|---|---|
| G | the ratio of the total net amount of valuable info. gained |
| T_B | The total amount of time spent between-patches |
| T_W | The total amount of time spent within-patches foraging |
| g | The avg. gain per patch |
| t_B | The avg. time between processing patches |
| t_W | The avg. time to process patches |

# Information patch foraging

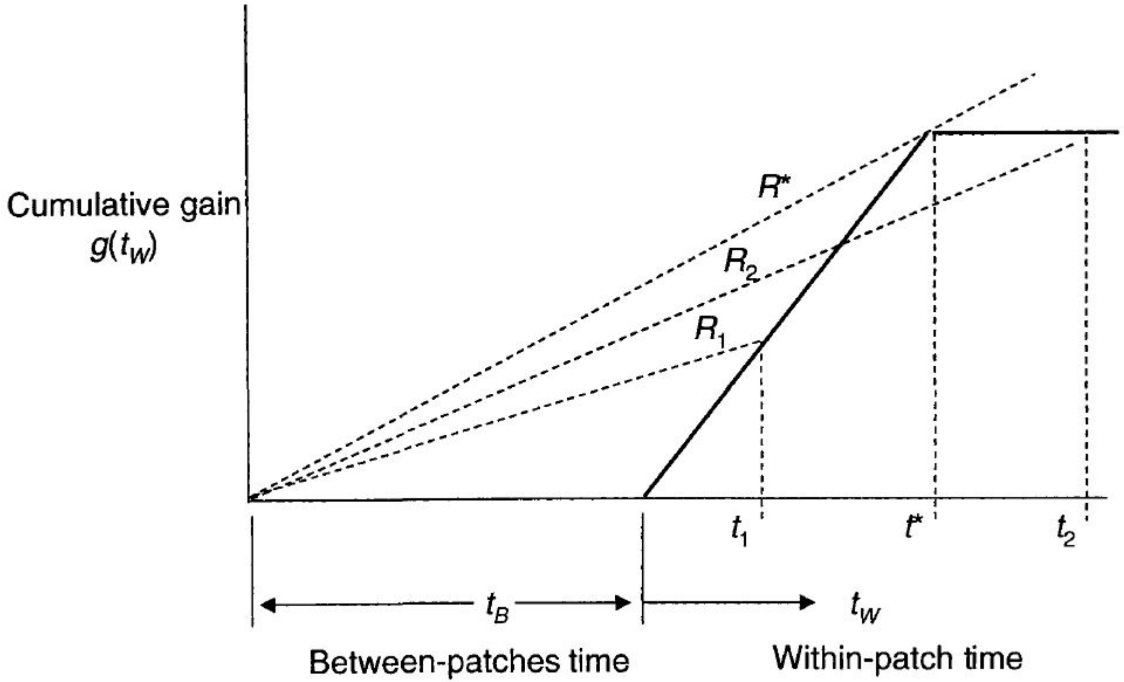
Incorporating patch types *i* ({1,2, …, P}) gives

$$R = \frac{\lambda g}{1 + \lambda t_W} = \frac{\sum_{i=1}^{P} \lambda_i g_i(t_{Wi})}{1 + \sum_{i=1}^{P} \lambda_i t_{Wi}}$$

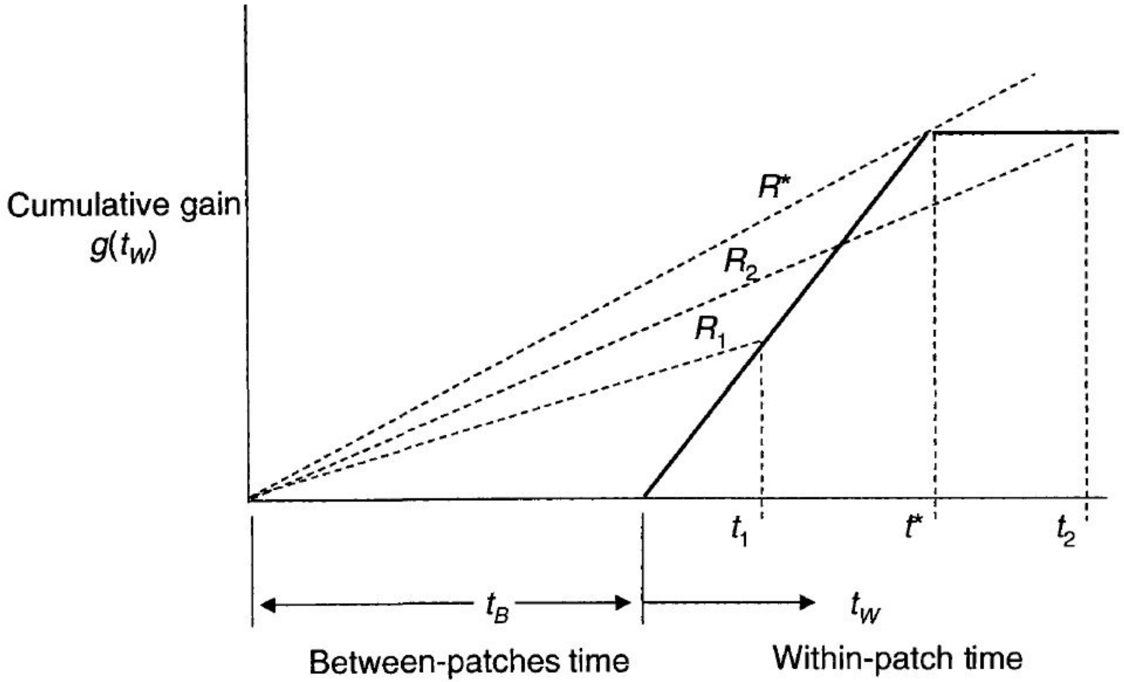
## Table of notation

| | |
|---|---|
| G | the ratio of the total net amount of valuable info. gained |
| T_B | The total amount of time spent between-patches |
| T_W | The total amount of time spent within-patches foraging |
| g | The avg. gain per patch |
| t_B | The avg. time between processing patches |
| t_W | The avg. time to process patches |

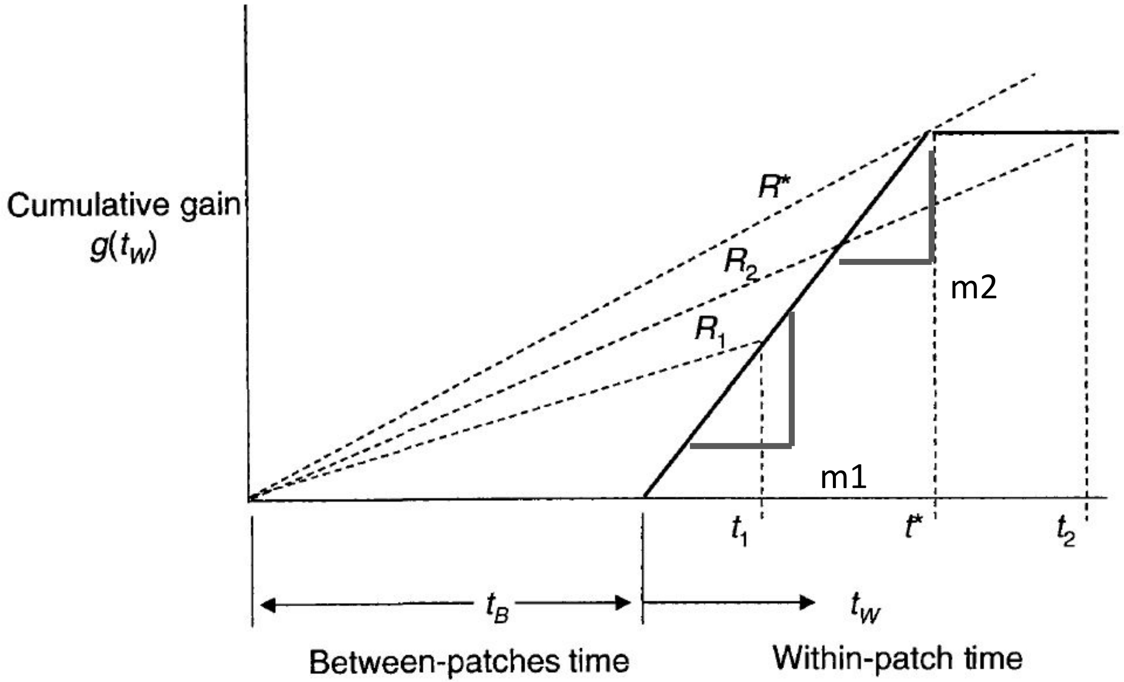
# Gain graph (for a specific type *i*)



# Q. Is this realistic enough?

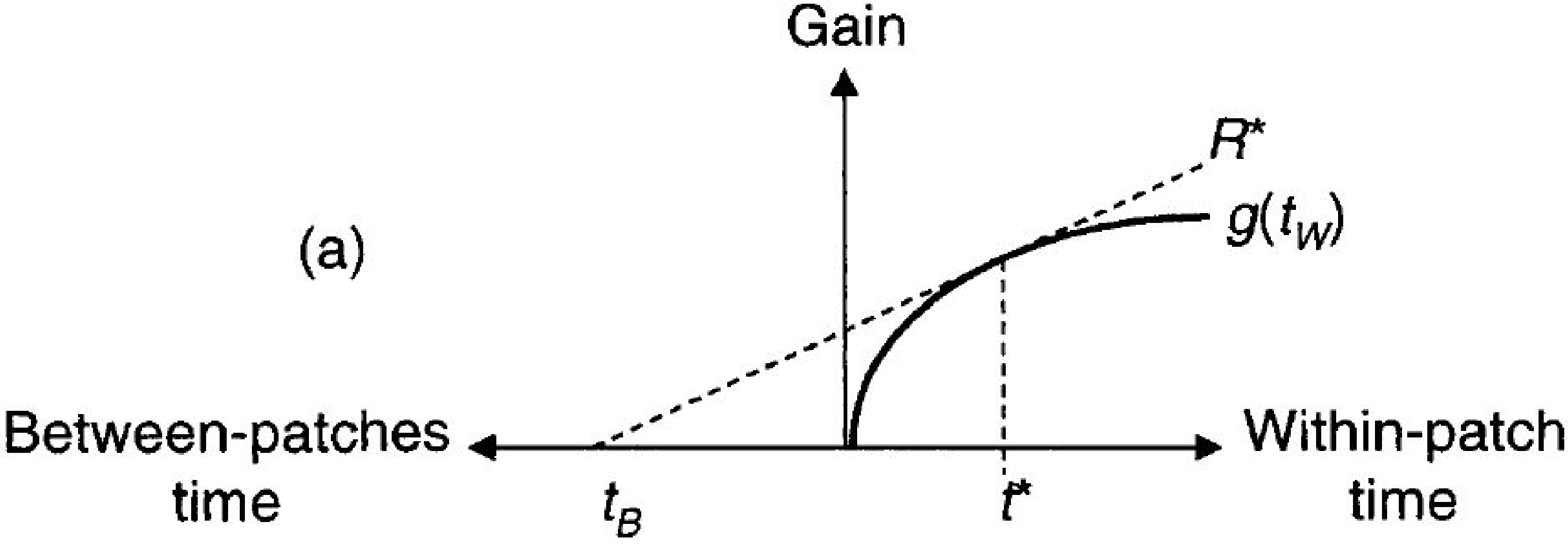


**Charnov's Marginal Value Theorem** states m1 > m2.

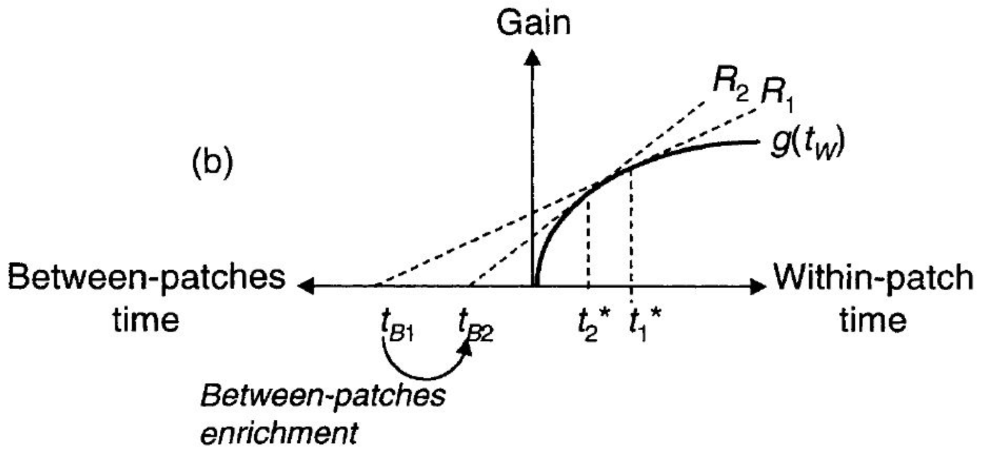


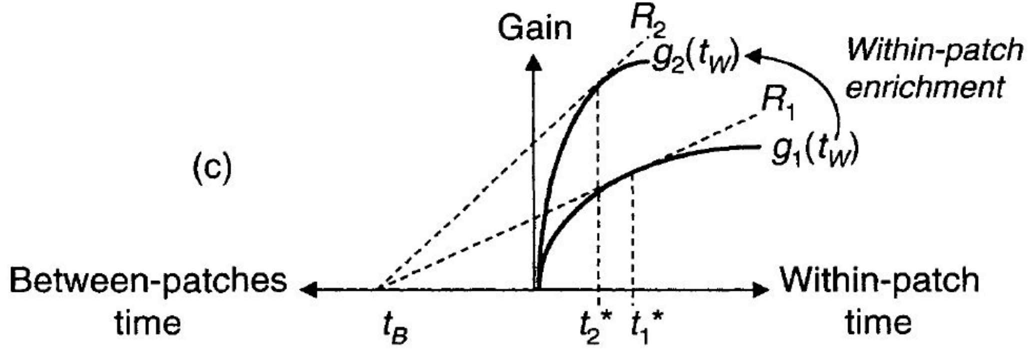
**Charnov's Marginal Value Theorem** states m1 > m2.

Q. what are other similar patterns observed in the real-world? How does marginal value theorem affect information foraging behaviors?

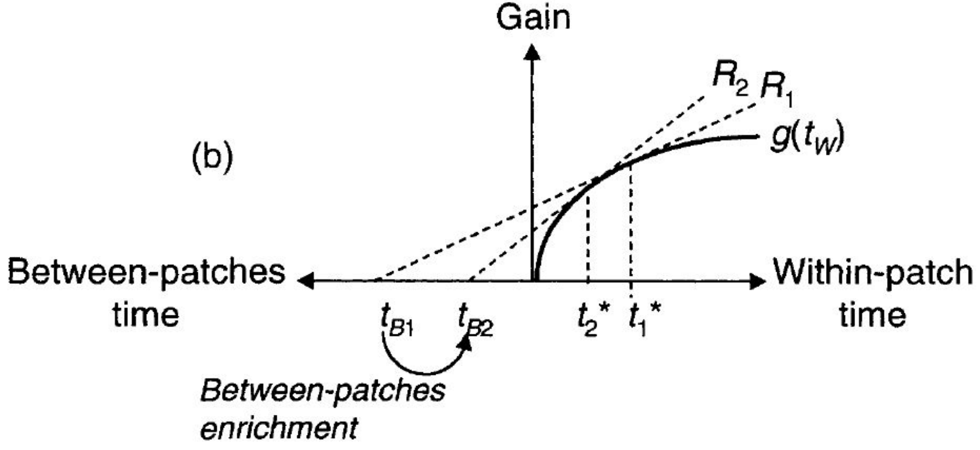


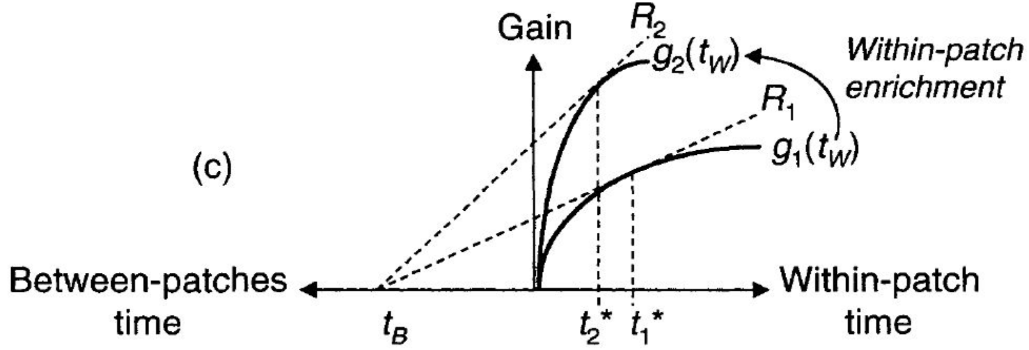
# Enrichment activities & their effect



Q. What are real-life examples of each enrichment type? And can you explain it using the changes described in the graphs?



# Information diet selection (there are not just one type of patches!)

If there are different types of information, (differing in their respective profitability), then you'll need to pursue items of the type of a specific profitability in an all-or-none manner; never have a mixed diet. **(zero-one rule)**

To create a decision model, introduce a new **probability parameter** $p_i$ -- we'll eventually solve for p_i's

$$p_i = 0 \text{ if } g_i/t_{Wi} < k_i/c_i$$

$$p_i = 1 \text{ if } g_i/t_{Wi} > k_i/c_i$$

## Table of notation

| | |
|---|---|
| G | the ratio of the total net amount of valuable info. gained |
| T_B | The total amount of time spent between-patches |
| T_W | The total amount of time spent within-patches foraging |
| g | The avg. gain per patch |
| t_B | The avg. time between processing patches |
| t_W | The avg. time to process patches |

## Q. What are the examples of how the diet selection algorithm can be used?

The optimal diet selection algorithm suggests two aspects of the information foragers' behavior:

- **Lost opportunity.** Information item types should be ignored if their profitability is less than the expected rate of gain of continuing search for other types of items
- **Decision to include a new information item type is independent of its prevalence but profitability** (but dependent of prevalence of already included item types).

**Lambda appears only on the lefthand-side**

$$R(k) = \frac{\sum_{i=1}^{k} \lambda_i g_i}{1 + \sum_{i=1}^{k} \lambda_i t_{Wi}} > \frac{g_{k+1}}{t_{Wk+1}} = \pi_{k+1}$$

Then don't go for k+1

**Righthand-side contains only the profitability**

Q. What is the main limitation of the diet selection algorithm to be practical in predicting the actual selections we make?

The optimal diet selection algorithm suggests two aspects of the information foragers' behavior:

- **Lost opportunity.** Information item types should be ignored if their profitability is less than the expected rate of gain of continuing search for other types of items
- **Decision to include a new information item type is independent of its prevalence but profitability** (but dependent of prevalence of already included item types).

**Lambda appears only on the lefthand-side**

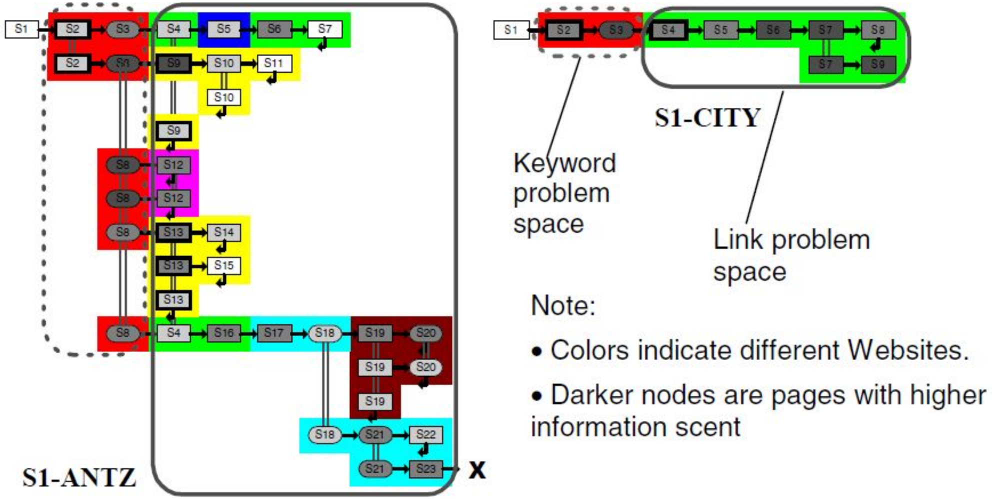
$$R(k) = \frac{\sum_{i=1}^{k} \lambda_i g_i}{1 + \sum_{i=1}^{k} \lambda_i t_{Wi}} > \frac{g_{k+1}}{t_{Wk+1}} = \pi_{k+1}$$

**Righthand-side contains only the profitability**

# Information scent

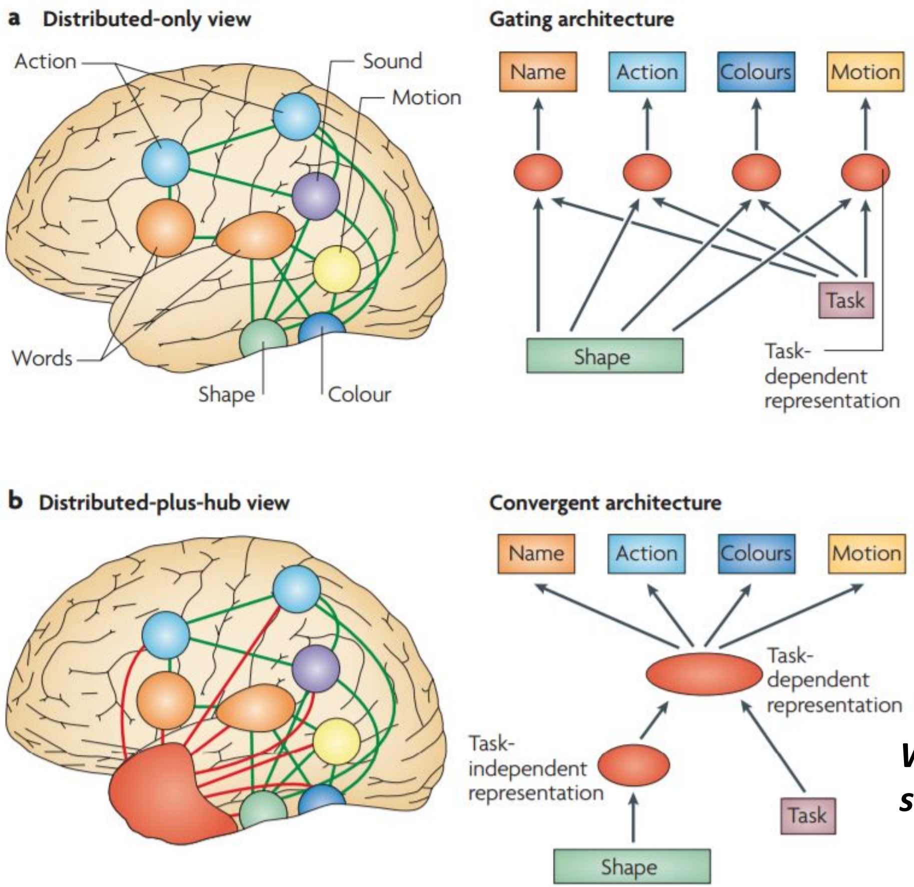
New evidences on **the representation of semantic knowledge** in the human brain (--> spreading activation)

New empirical evidence on user's web behavior




***Information Scent as a Driver of Web Behavior Graphs: Results of a Protocol Analysis Method for Web Usability***
Stuart Card et al.

# Background: The representation of semantic knowledge




How semantic knowledge is represented in our brain
→ Inspired from this discovery, the **spreading activation** mechanism became an integral part of the assessment of info. scent.

***Where do you know what you know? The representation of semantic knowledge in the human brain***

Patterson et al.

# Bayesian analysis of information scent

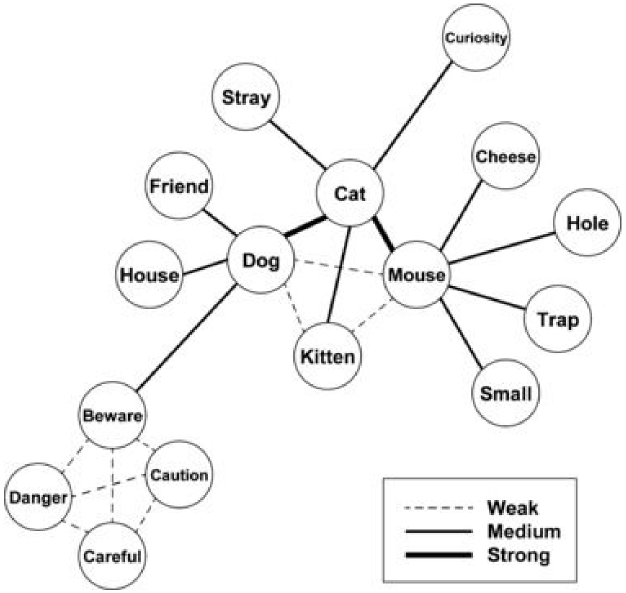
The spread of activation from one cognitive structure to another is determined by some network representation.

**Base activation for query *i***

$$A_i = \overline{B_i} + \underbrace{\sum_j W_j S_{ji}}$$

**Sum of activation from other concepts**

* Interpret A_i as Bayesian a posteriori logarithmic odds, B_i as log prior odds of *i* being relevant, and S_{ji} as the log likelihood ratios that i is relevant given that it occurs in the context of word *j*



# Bayesian analysis of information scent

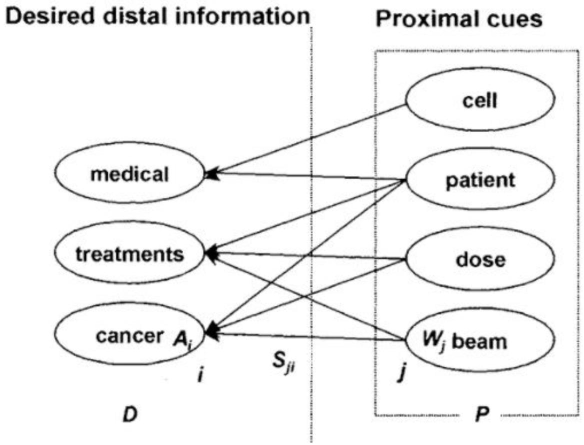
(Prior) Odds $O(i) = \frac{P(i)}{P(\neg i)}$

Posterior Odds
$$\begin{aligned} O(i|j) &= \frac{P(i|j)}{P(\neg i|j)} \\ &= \frac{P(j|i)P(i)}{P(j|\neg i)P(\neg i)} \\ &= \frac{P(i)}{P(\neg i)} \times \frac{P(j|i)}{P(j|\neg i)} \\ &= O(i) \times \frac{P(j|i)}{P(j|\neg i)} \end{aligned}$$

Making a simplifying independence assumption for each individual feature j in the set P of proximal cues yields

$$O(i|P) = O(i) \times \prod_{j \in P} \left\{ \frac{P(j|i)}{P(j|\neg i)} \right\}^{w_j}$$

Finally, taking log of both sides $$\underbrace{\log O(i|P)}_{A_i} = \underbrace{\log O(i)}_{B_i} + \sum_{j \in P} w_j \underbrace{\log \frac{P(j|i)}{P(j|\neg i)}}_{S_{ji}}$$



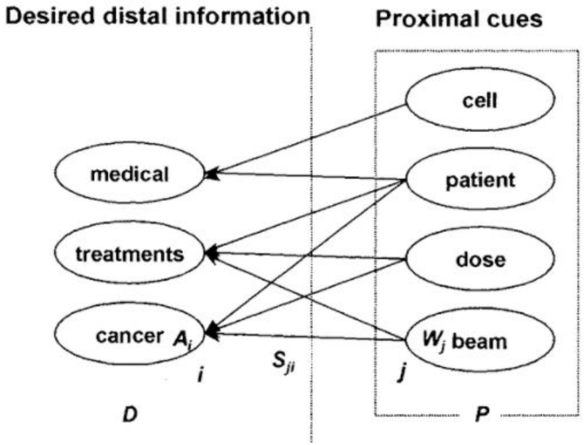

*Example by Robert Goldstone*

# Bayesian analysis of information scent

$$\underbrace{\log O(i|P)}_{A_i} = \underbrace{\log O(i)}_{B_i} + \sum_{j \in P} \underbrace{w_j \log \frac{P(j|i)}{P(j|\neg i)}}_{S_{ji}}$$

Info scent assessment model using activation spreading (adopted from Kruschke)

$$g(c, s) = \exp\left(\frac{\sum_{i \in Q} A_i}{T}\right)$$




*Example by Robert Goldstone*

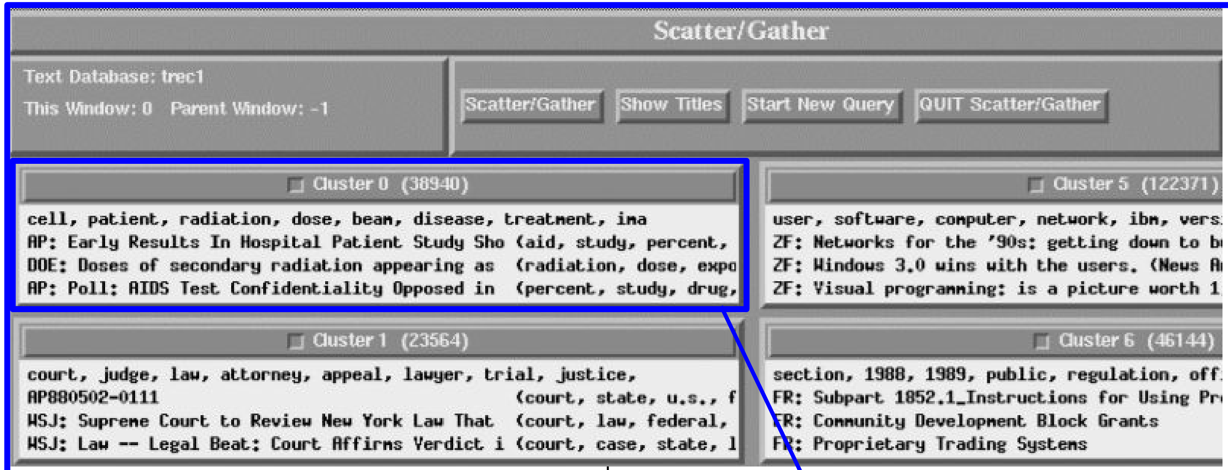
# Bayesian analysis of information scent

$$\overset{A_i}{\overline{\log O(i|P)}} = \overset{B_i}{\overline{\log O(i)}} + \sum_{j \in P} \overset{S_{ji}}{\overline{w_j \log \frac{P(j|i)}{P(j|\neg i)}}}$$

Info scent assessment model using activation spreading (adopted from Kruschke)

$$g(c, s) = \exp\left(\frac{\sum_{i \in Q} A_i}{T}\right)$$

screen




cluster

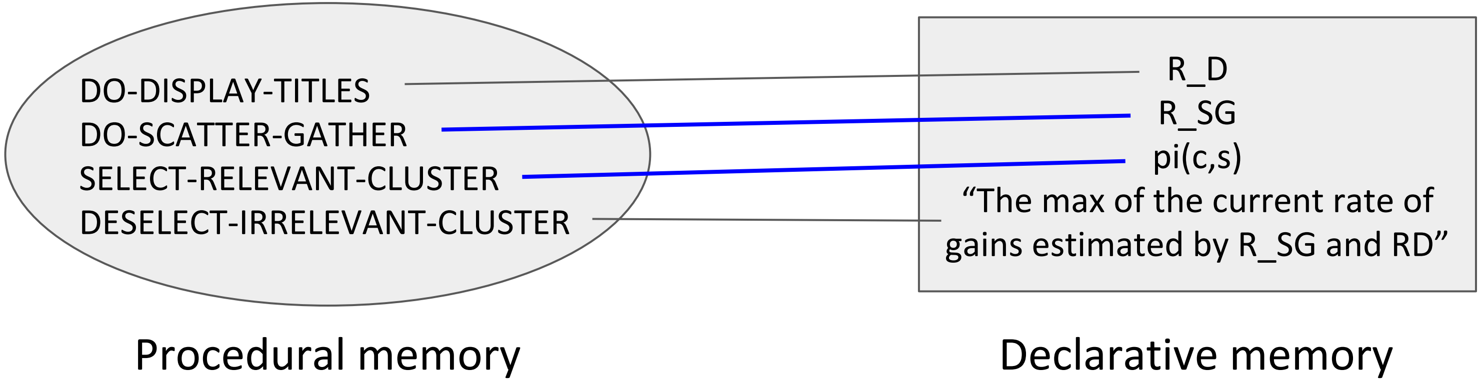
***Constant interaction-time scatter/gather browsing of very large document collections Cutting***

Cutting, Karger, and Pederson, 1993

Conceptual working of ACT-IF: human information foraging behavior can be modelled with

- Declarative (factuals and semantic relationships) memory
- Procedural (if-then rules that executes based on activation) memory
- Goal (what user wants to find out) memory
- Information Scent mechanism

In the interest of time, we examine two of the procedural memory - declarative memory mappings below (in blue):



Q. Can all human behaviors modeled in this analytic way using the notion of declarative and procedural memory?

Q. Can all human behaviors modeled in this analytic way using the notion of declarative and procedural memory?

→ Related to your discussion points:

***The Newell Test for a theory of cognition***

John Anderson and Christian Lebiere

They compared ACT-R with Connectionism using 12 criteria (distilled from Newell's original 13 criteria). Among them, criteria such as *Consciousness, Development, Evolution, and Natural language* are the ones that ACT-R is deemed to be performing worse.

SELECT-RELEVANT-CLUSTER: Clusters at state s should be selected so long as their profitability $\pi(c, s)$ is greater than the overall rate of gain for the clusters gathered at that sta $R_D(k, s, t)$.

The profitability term can be computed as

$$\pi(c, s) = \frac{g(c, s)}{t_g g(c, s) + t_N N(c, s)}$$

The numerator is the gain computed using the modeling earlier, and the denominator is time, where *tg* and *tN* are the time it takes to process a relevant document title and the title in the gathered cluster, respectively

$$\log O(i|P) = \underbrace{\log O(i)}_{A_i} + \underbrace{\sum_{j \in P} w_j \log \frac{P(j|i)}{P(j|\neg i)}}_{S_{ji}}$$

$$g(c, s) = \exp\left(\frac{\sum_{i \in Q} A_i}{T}\right)$$

SELECT-RELEVANT-CLUSTER: Clusters at state s should be selected so long as their profitability $\pi(c, s)$ is greater than the overall rate of gain for the clusters gathered at that sta $R_D(k, s, t)$.

The overall rate of gain can be computed as $R_D(k, s, t_B) = \dfrac{\sum_{i=1}^{k} g(i, s)}{t_B + t_W}$

$$R_D(k, s, t_B) = \frac{\sum_{i=1}^{k} g(i,s)}{t_B + [t_N \sum_{i=1}^{k} N(i,s) + t_g \sum_{i=1}^{k} g(i,s)]}$$

# An experiment that show the predictive power of the model (but there are more than one empirical evidence introduced here!)

**Participants:** 12 adults from Xerox PARC or Stanford
**Task:** Collect as many relevant articles as possible for a given query topic using Scatter/Gather*
**Conditions:** 12 query topics at three levels of difficulty (measured by the mean number of expert-identified relevant documents)

Hard: avg. 46 vs Medium: avg. 303 vs Easy: avg. 865

**Study design:** 4 blocks of topics were constructed, each topic-block contained 1 easy, 1 medium, and 1 hard topic (in this order). Each participant completed 2 blocks of topics using Scatter/Gather (2 other for other activities), the presentation order of blocks was counterbalanced over participants, within groups, according to a randomized Latin square.

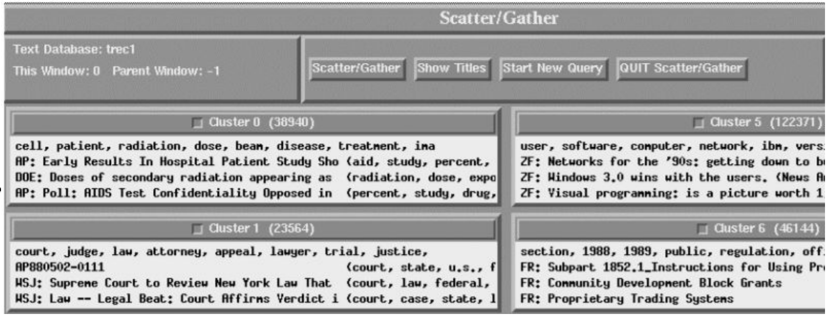
4 participants were in a timed condition, 4 were in not-timed. The latter group of participants also provided subjective ratings on what percentage of texts in a cluster seemed relevant

# Experiment - Can the information diet model predict which clusters get selected?

**Table 1.**
**Optimal information diet analysis for Scatter/Gather** (data from Pirolli et al., 1996).
**The optimal diet includes the four highest profitability clusters.**

| Task Condition (Rank of cluster within query condition) | Participants' estimate of net relevant documents ($g_i$) | Handling time in sec ($t_{Wi}$) | Estimated profitability ($\pi_I$) in relevant documents per sec. |
|---|---|---|---|
| Easy (1) | 13,670 | 957 | 14.28 |
| Medium (1) | 10,400 | 994 | 10.46 |
| Hard (1) | 11,890 | 1,261 | 9.43 |
| Easy (2) | 5,824 | 957 | 6.08 |
| Easy (3) | 2,526 | 957 | 2.64 |
| Medium (2) | 2,607 | 994 | 2.62 |
| Hard (2) | 1,991 | 1,261 | 1.58 |
| Easy (4) | 1,040 | 957 | 1.09 |
| Easy (5) | 891 | 357 | .93 |
| Hard (3) | 379 | 1,261 | .30 |

Participants chose (avg.)
1.38 clusters for Hard queries
1.63 clusters for Medium queries
2.25 clusters for Easy queries



# Experiment - Can the information diet model predict which clusters get selected?

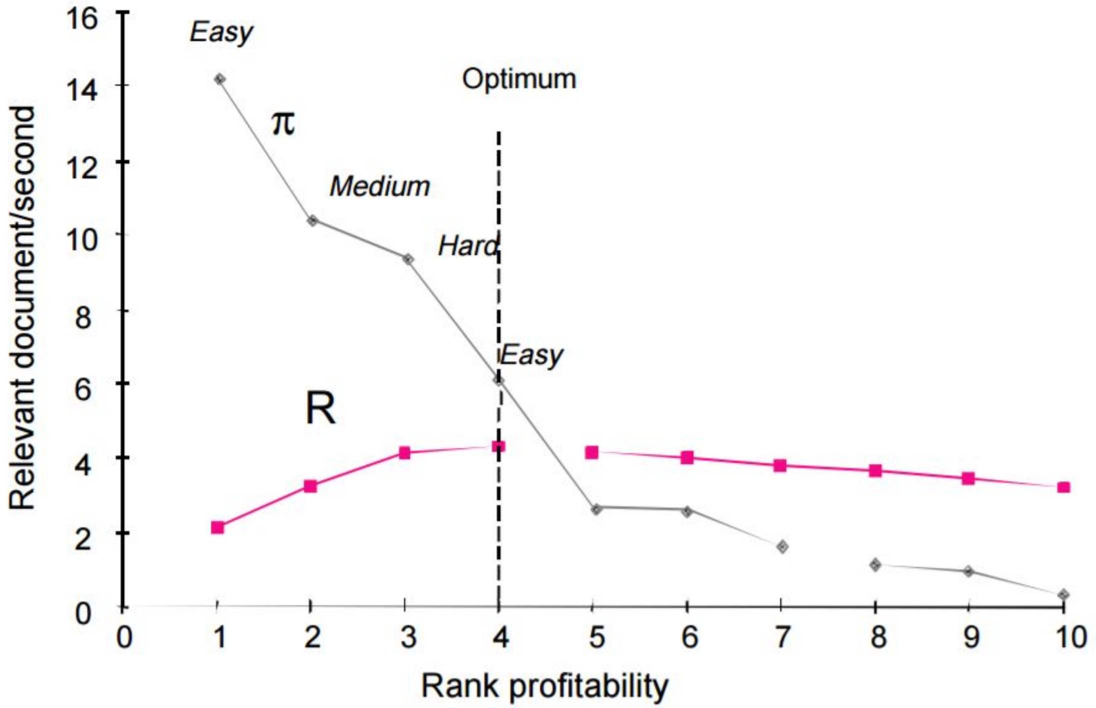



**Figure 15. Analysis of the optimal information diet. The profitability (π) of clusters is ranked and added to the diet in order of decreasing profitability until the rate of gain, R, so long as the profitability of the item is greater than R.**

Participants chose (avg.)
1.38 clusters for Hard queries
1.63 clusters for Medium queries
2.25 clusters for Easy queries

Model predicted
Top 1 cluster for Hard queries
Top 1 cluster for Medium queries
Top 2 clusters for Easy queries

# Experiment - Can the information scent model predict perceived topic relevance?

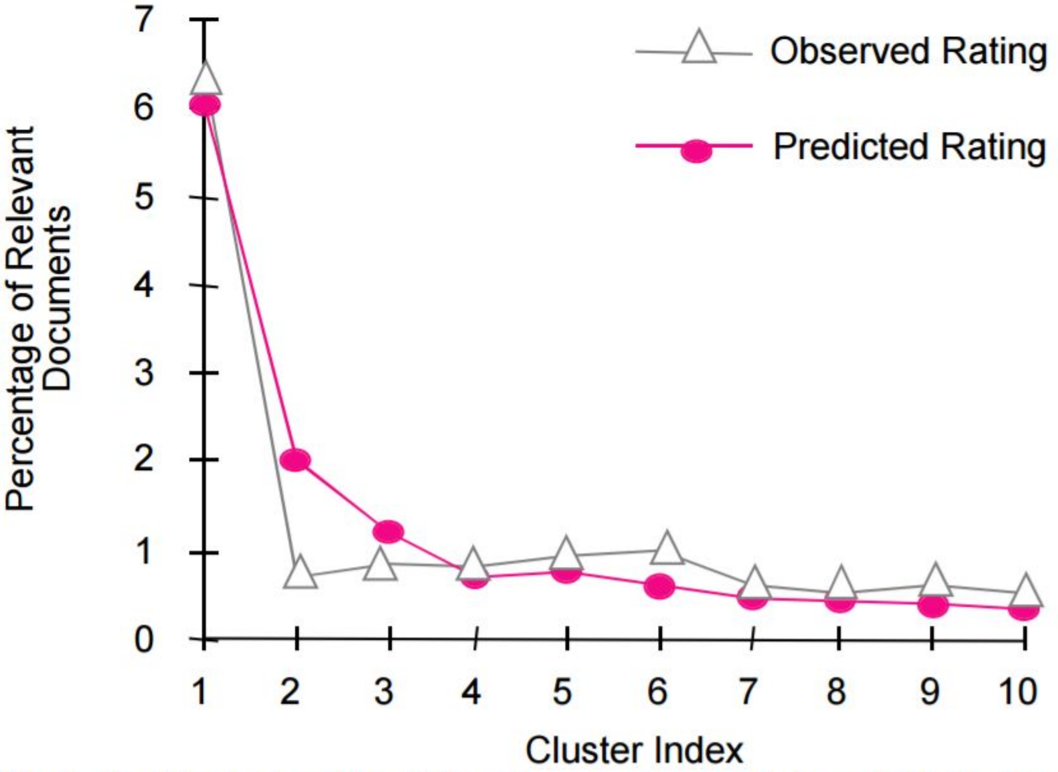



**Figure 16. Observed ratings of the percentage documents in each cluster that are relevant and the ratings predicted by activation-based assessment of information scent.**

Observed rating (triangles)
Predicted rating (circles) by g(c, s)/N
Linear fitting of

$$Rating = a + b\left(\frac{g(c,s)}{N}\right)$$

yields

$R^2 = .92; a = .32$ and $b = 232$

# Experiment - Can the IFT model predict the selection of clusters?

If we let

$$x = Cluster\ Profitability - Expected\ Rate\ of\ Gain$$
$$= \pi(c,s) - R_D(k,s,t),$$

the model states that decisions should be made by users to

(a) select a cluster when x > 0

(b) do not select a cluster when x < 0

x = 0 happens when profitability equals rate of gain.

# Experiment - Can the IFT model predict the selection of clusters?

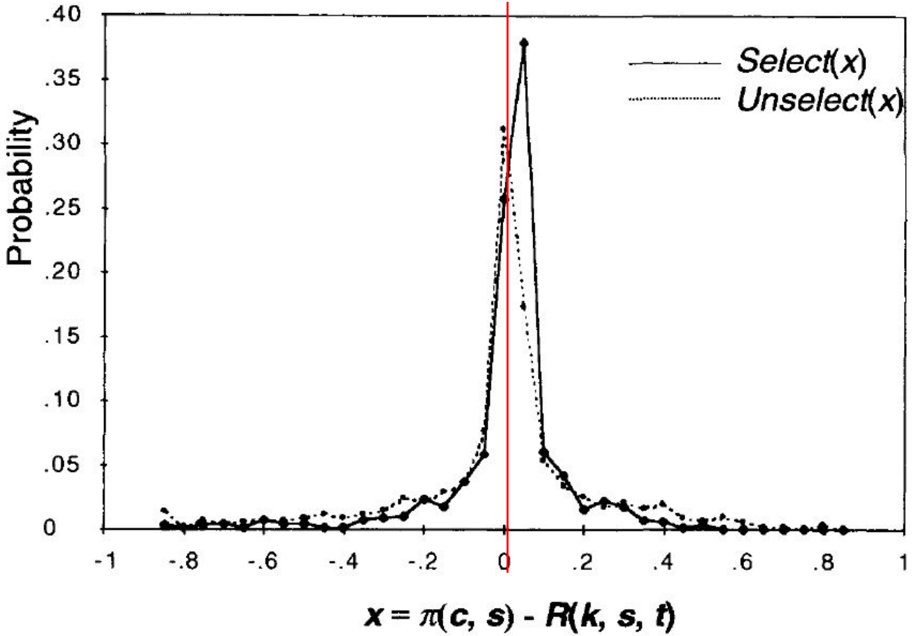



*Figure 18.* The probability density distributions for selecting clusters, *select*($x$), and not selecting clusters, *unselect*($x$), as a function of the difference between cluster profitability and current estimate of rate of gain: $x = \pi(c, s) - R(k, s, t)$.

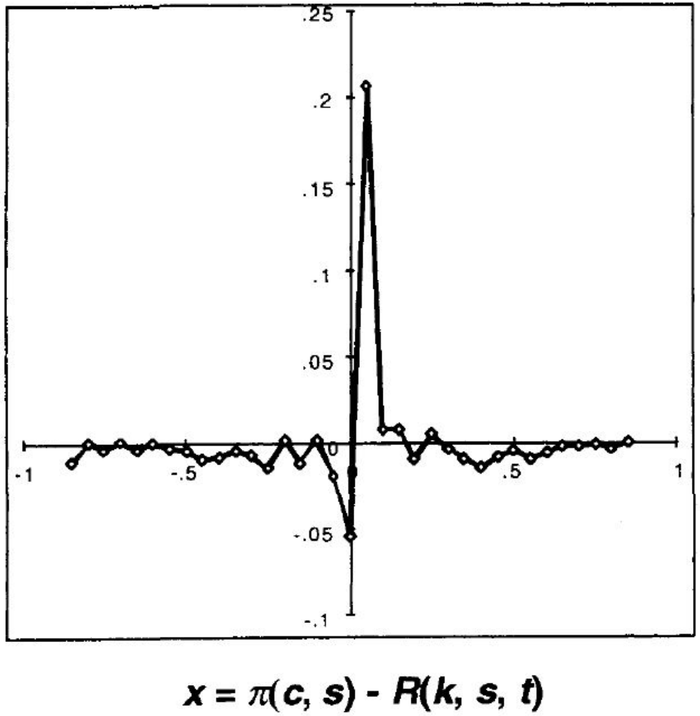



*Figure 19.* The difference in density distributions from Figure 18, *select*($x$) − *unselect*($x$) as a function of the difference between cluster profitability and rate of gain, $x = \pi(c, s) - R(k, s, t)$.

The shift in probability of selecting vs not selecting clusters across the threshold x = 0

$\chi^2(1, N = 2{,}929) = 50.65, p < .0001$

## You commented

Q. Give an example of how information foraging theory can be applied to increase the information scent of a website design.

## Discussion

Some careful rational analysis can lead to a mathematical model of spreading activation, which then can be used to predict user behaviors on the Web.

Q. What are the potential applications of this modeling?

Q. What are the limitations?

# Q. How could information foraging theory be extended to model cooperative behaviors? (e.g. Wiki, collaborative filtering)

| TIME (sec) | (common units) | ACTION | MEMORY | THEORY |
|---|---|---|---|---|
| $10^9$ | (decades) | Technology | Culture | |
| $10^8$ | (years) | System | Development | Social and |
| $10^7$ | (months) | Design | Education | Organizational |
| $10^6$ | (weeks) | Task | Education | |
| $10^5$ | (days) | Task | Skill | |
| $10^4$ | (hours) | Task | Skill | Bounded |
| $10^3$ | (ten mins) | Task | LTM | Rationality |
| $10^2$ | (minutes) | Task | LTM | |
| 10 | (ten secs) | Unit task | LTM | |
| 1 | (secs) | Operator | STM | Psychological |
| $10^{-1}$ | (tenths) | Cycle time | Buffers | |
| $10^{-2}$ | (centisecs) | Signal | Integration | Neural and |
| $10^{-3}$ | (millisecs) | Pulse | Summation | Biochemical |

Revisiting the earlier question, modify the task at hand to "a group of friends planning a trip to New York."

*The Prospects for Psychological Science in Human-Computer Interaction*
Allen Newell and Stuart Card

# Q. How could information foraging theory be extended to model cooperative behaviors? (e.g. Wiki, collaborative filtering)

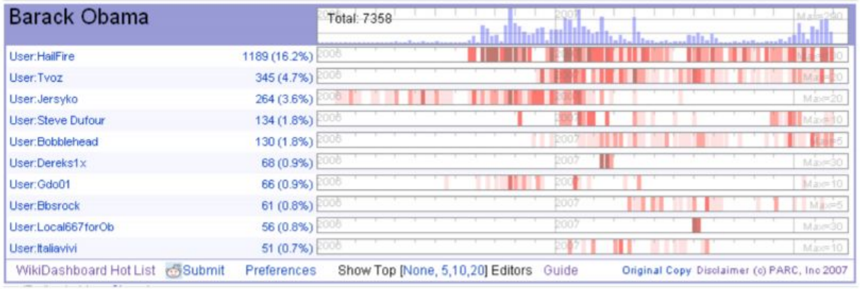



**Figure 2 Article Dashboard. The top summary graph shows the weekly edit trend of this page. Below the summary graph, there is a list of active editors and their activities on the article.**

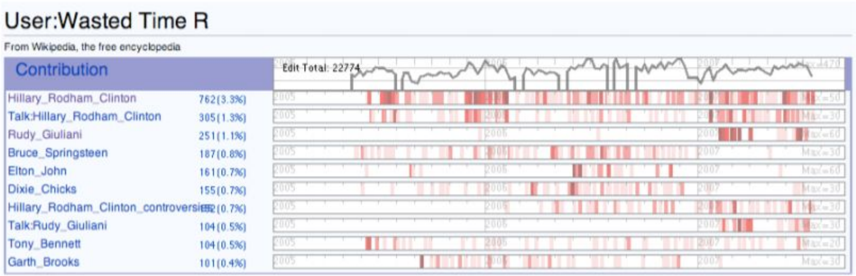



**Figure 2 User Dashboard .The dashboard displays weekly edit trend of an editor as well as the list of articles that the editor made revisions on.**

Trustworthiness is another dimension important in processing and aggregating information

***So you know you're getting the best possible information: a tool that increases Wikipedia credibility***

Peter Pirolli et al.

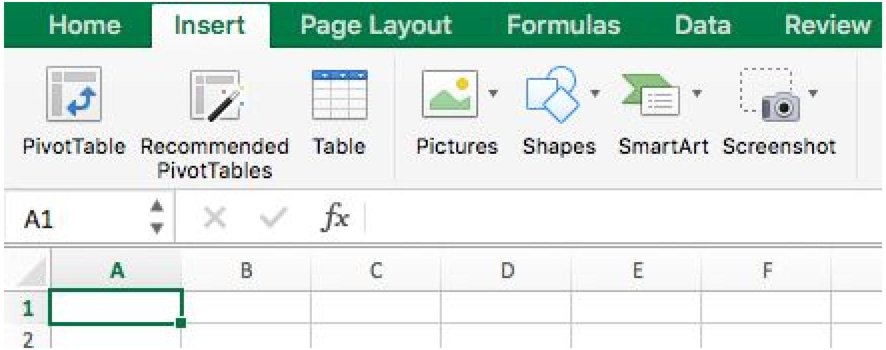
# *Beyond Performance: Feature Awareness in Personalized Interfaces*

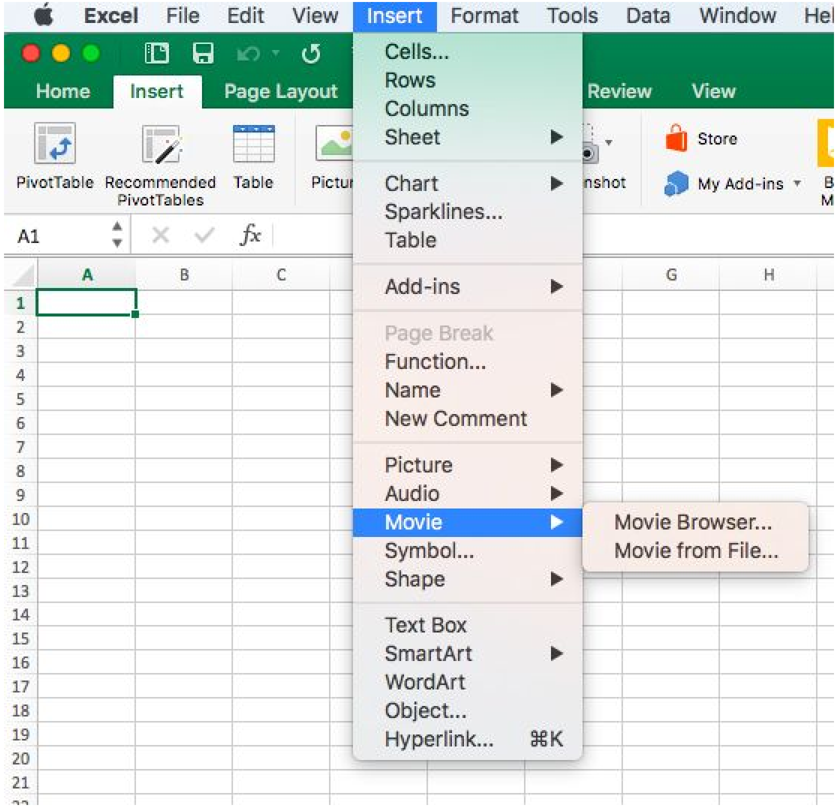
Leah Findlater and Joanna McGrenere

# Learning goals

- Be able to explain interface personalization and two related measures: performance and awareness
- Understand principles and techniques for designing experiments to maximize statistical power

# Different usage of GUI



## Components of Interface Personalization

### Performance

- Core task performance
- New task performance

### Awareness

impacts

- Awareness is about learning generally.
- Measures : Recognition rate of unused features, New task performance

## Design Space

**Control** - Adaptive (Automatically), Adaptable (Manually), or a mix of both

**Granularity** - Fine (high accuracy) & Coarse (low accuracy)

**Visibility** - Hide, mark, resize, move, replicate

**Frequency** - High/low frequency

## Design Space

Discussion

**In what cases are adaptive and adaptable personalizations desirable, respectively?**

**Design Space**

| Control | Adaptive | | Adaptable | | Mixed-initiative | |
|---|---|---|---|---|---|---|
| **Granularity** | *Coarse* | *Fine* | *Coarse* | *Fine* | *Coarse* | *Fine* |
| **Visibility of change**: *Hidden* | | MS Office 2003 adaptive menus | Layered interfaces (Clark and Matthews, 2005; Findlater and McGrenere, 2007; Gustavsson Christiernin et al., 2003; Plaisant et al., 2003; Shneiderman, 2003)<br>User role-based (Findlater et al., 2008; Greenberg, 1991) | Multiple interfaces (McGrenere et al., 2002) | | Incremental interfaces (Brusilovsky and Schwarz, 1997)<br>Adaptive bar (Debevc et al., 1995)<br>Adaptively supported multiple interfaces (Bunt et al., 2007) |
| *Moved* | | Original split menus (Sears and Shneiderman, 1994)<br>Frequency based menus (Mitchell and Shneiderman, 1989)<br>Adaptive hierarchical menus (Greenberg and Witten, 1985)<br>Adaptive split menus (Findlater and McGrenere, 2008) and toolbars (Gajos et al., 2006) | | | | |
| *Resized* | | Ability-based interfaces (Gajos et al., 2008b)<br>Morphing menus (Cockburn et al., 2007) | | | | |
| *Replicated* | | Replicated split interfaces (Findlater and McGrenere, 2008; Gajos et al., 2006; Gajos et al., 2008a) | | Facades (Stuerzlinger, et al., 2006) | | |
| *Marked* | | Colour highlighting (Tsandilas and schraefel, 2005)<br>Ephemeral adaptation (Findlater et al., 2009) | Marked layered interface (Findlater and McGrenere, 2007) | | | |

## Design Space

Concluded from the experiments are...

**Control of personalization**

- Users improve their awareness when doing the “adaptable” personalization.
- “Adaptive” personalization could trade accuracy for awareness.

**Granularity**

- “Fine” improves core task performance.
- “Coarse” could contribute to awareness if properly designed.

## Design Space

### Visibility of change

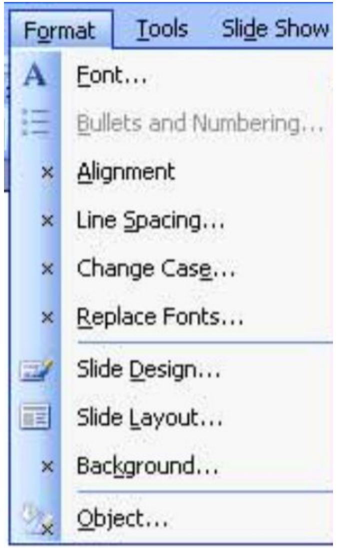
- Hiding negatively impacts awareness.
- Graphical marking may result in higher awareness than hiding.
- Direction of change could affect awareness and core task performance.

### Frequency of Change

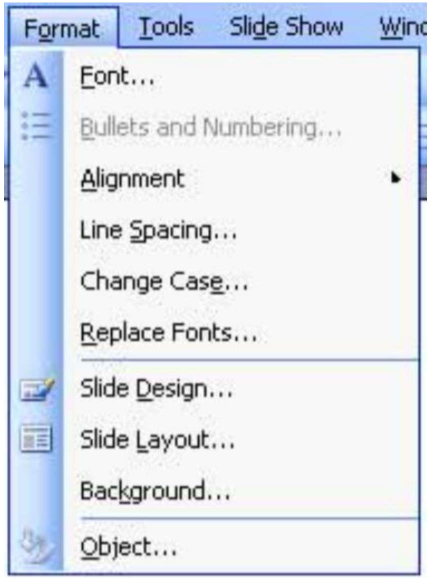
- Future work required.

## S I - Layered Interface

Minimal

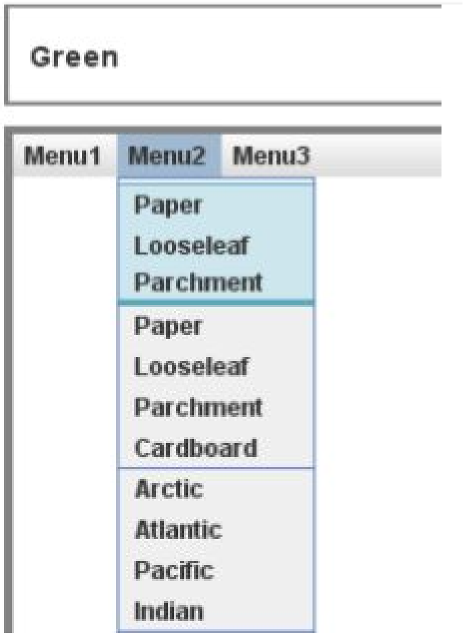




Control




Marked

## S II&III - Split Menus




High & Low Accuracy


Control

## Experiment I

**Hypothesis** - Personalization makes better core task performance but lower awareness than the control condition.

**Conditions** - Minimal, Marked, Control

**Methodology** - between-subjects design

**Result**

- Core task performance: Minimal > Control
- Awareness: Control > Minimal
- Marked shows no significant effect on performance and awareness.
- Awareness may indirectly impact new task performance.

## Experiments

**Study I**
Layered Interfaces

There is a trade off between core task performance and awareness.

**Study II**
Adaptive Split Menus

## Experiment II

**Problem** - Impact of adaptive split menus and screen size on core task performance, awareness and user satisfaction

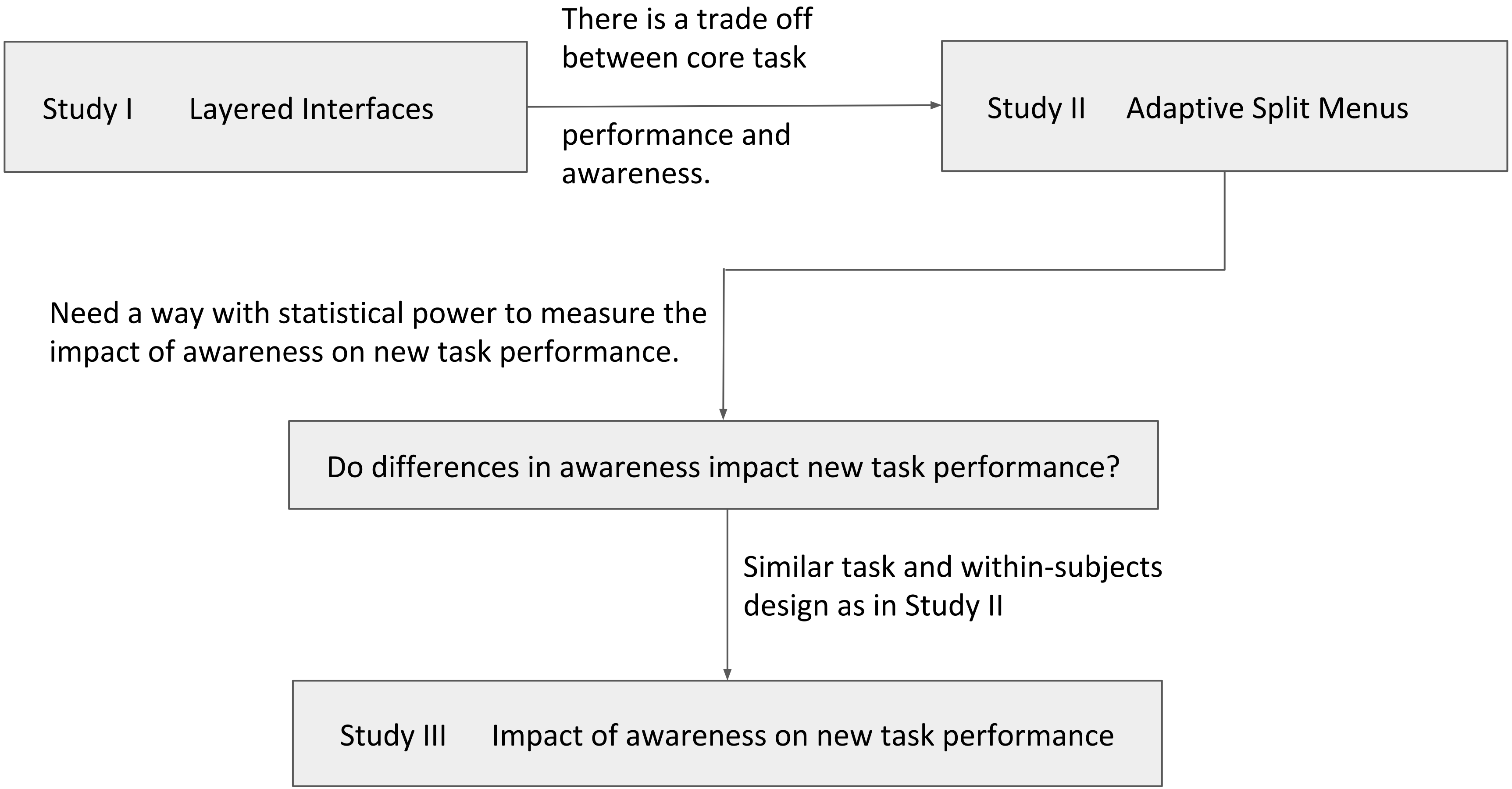
## Conditions

- Screen size (between-subjects factor) : PDA, desktop
- Menu type (within-subjects factor) : High (78%), Low (50%), Control (static)

## Result

- Tradeoff between core performance and awareness.
- Large screen leads to better performance and better awareness (more menu items).
- Awareness: control > low > high
- Performance: high > control.

# Experiments



## Experiment III

### Hypotheses

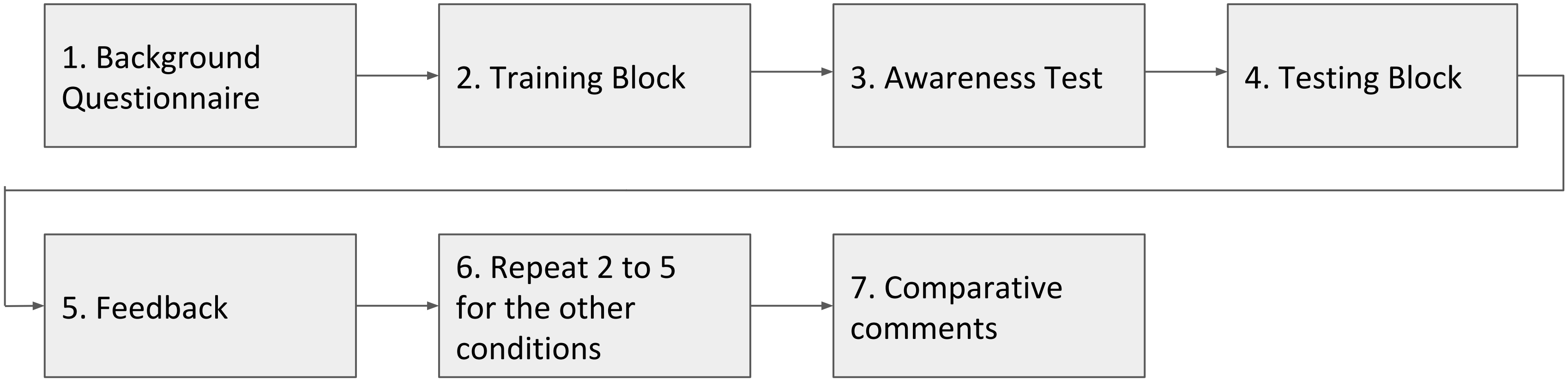
- Impact of awareness on new task performance: Control & Low > High
- Core task performance: High & Control > Low
- Perception of Awareness: Control & Low easier than High

**Conditions** - High (78% accuracy), Low (50% accuracy), Control

**Methodologies** - Within-subjects design, RM ANOVA

**Participants** - 30 (19 female)

## Experiment III



**Experiment III**

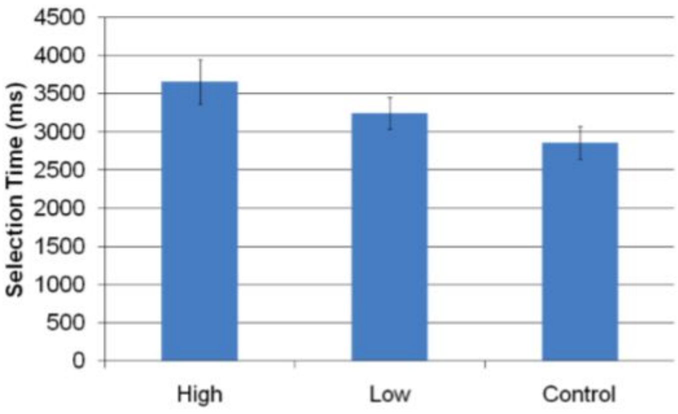
**Measures**

- **The time to select “new” items**
- **Corrected recognition rate**
- **Time to select “old” items**
- Feedback on each of the menu types

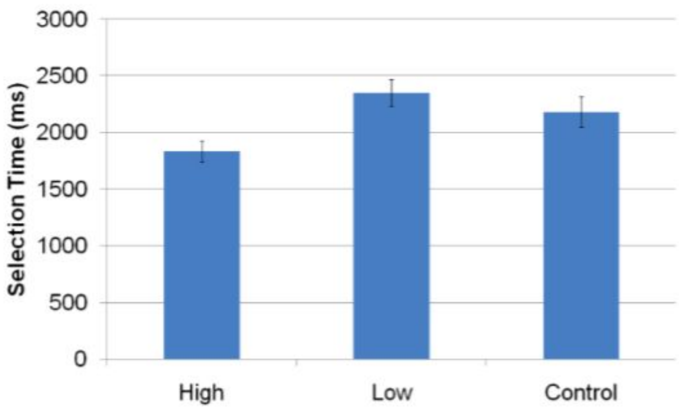
## Experiment III

## Results

- Impact of awareness on new task performance: Control > Low > High
- **Core task performance: High > Control > Low**
- Perception of Awareness: Control & Low easier than High




New task performance




Core task performance

## Experiment III

## Results

- Awareness impacts new task performance.
- Awareness and core task performance work against each other.
- Lower recognition test scores due to less exposure to the interface

**Q: The low accuracy condition does not serve as a trade-off between awareness and core task performance, why?**

## Design Implications

- Look beyond accuracy.
- Identify the **balance** between core performance and awareness.
- Match design characteristics to core performance and awareness.
- Use appropriate awareness measure in evaluations.
- Support exploratory behaviour.
- Make features easily discoverable.

## Methodologies

Within-subjects Design

Between-subjects Design

ANOVA

## Within-subjects design

- A type of experimental design in which participants are exposed to every treatment or condition
- All conditions per group

**Within-subjects design**

**Advantages**

**Relatively small applicant pool** (30 participants in experiment III)

Reduced errors due to the same participants in all conditions

- No individual difference (Everyone serves as his/her own baseline.)

**Disadvantages**

Carryover effect (Randomly generate selections.)

- Practice effects

Fatigue (Limited length procedure, short breaks)

**Between-subjects design**

- A type of experimental design in which two or more groups of subjects each is tested by a different testing factor simultaneously
- One condition per group

**Between-subjects design**

**Advantages**

No Carryover effect - Each group is assigned with one condition only.

Less fatigue - Relatively shorter compared with within-subjects design.

**Disadvantages**

Large Applicant pool

Errors due to the different participants in all conditions

- Individual difference

**ANOVA**

- Analysis of Variance
- A statistical test of whether or not the means of several groups are equal.
- **“Extended” t-test with more than two groups**

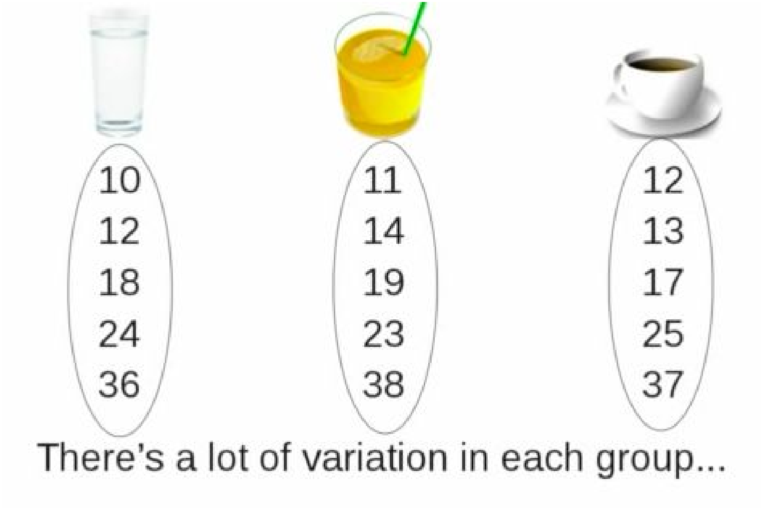
## ANOVA

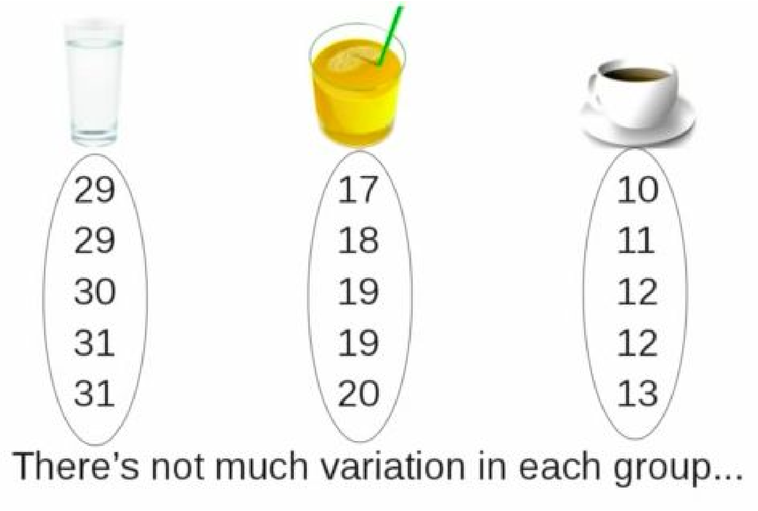
- **One-way Anova**
- Multivariate Anova
- Repeated Measures Anova

## One-Way ANOVA

Null hypothesis: The means for all three groups are the same.

# One-Way ANOVA




It's the people that make the difference, not the drink.

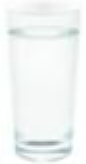
It's the drink that make the difference, not the people.

## One - Way ANOVA

$$F = \frac{between\ groups}{within\ groups}$$

- Calculate the variance between and within groups.
- The larger the ratio, the more likely that the groups have different means.

# Multi - Variable ANOVA

| |  |  |  |
|---|---|---|---|
|  | 30, 31, 31, 32, 32 | 28, 30, 27, 29, 32 | 25, 26, 25, 28, 29 |
|  | 31, 31, 33, 35, 30 | 29, 30, 28, 29, 31 | 28, 30, 27, 26, 27 |

## RM ANOVA

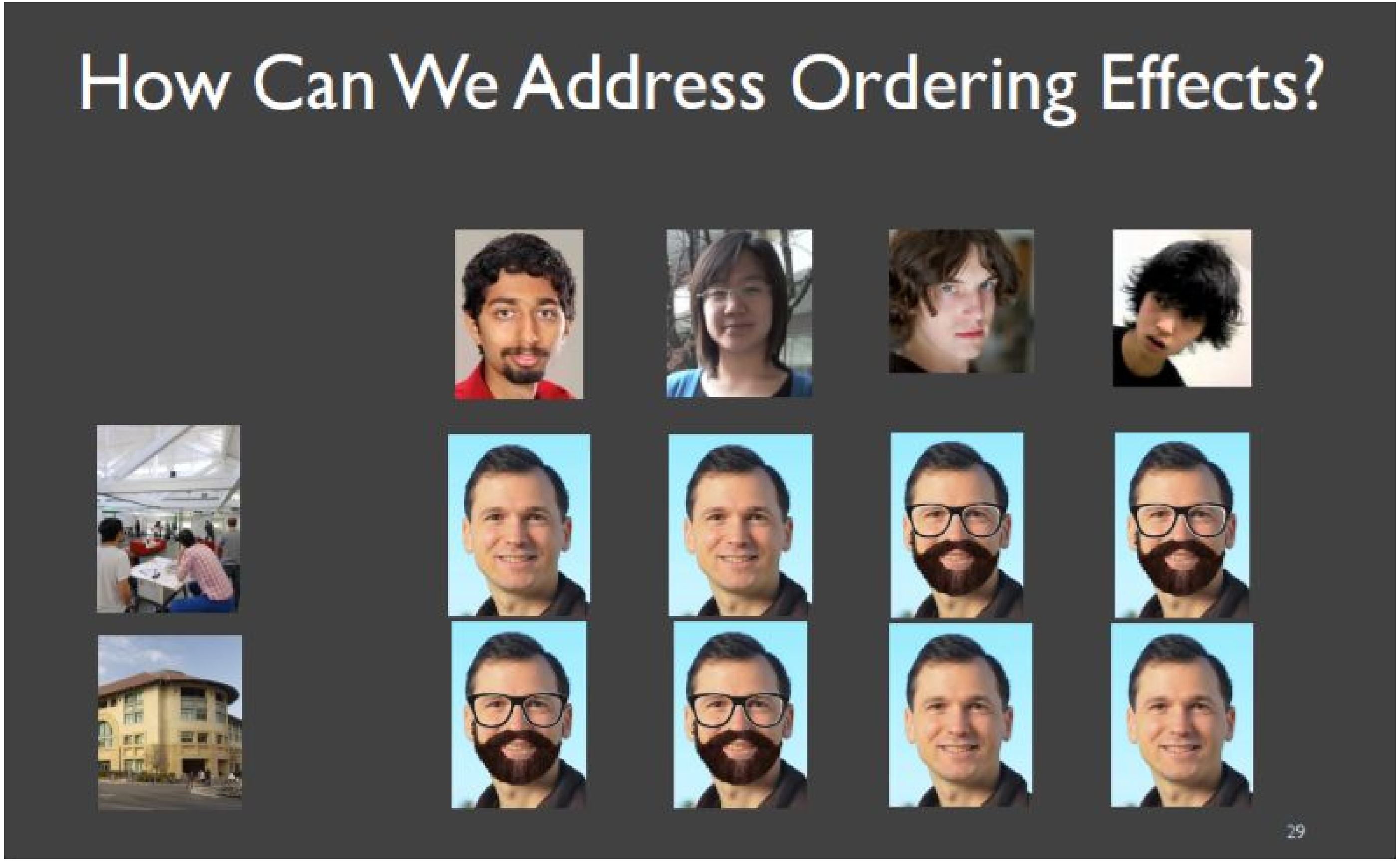
- “Analysis of dependencies”
- A test to prove an assumed cause-effect relationship between the independent variable(s) and the dependent variable(s)
- Used in within-subjects design

## Q: Will you use ANOVA in your project? Why or why not?

# Latin Square


*Adopted & modified from Scott Klemmer and Michael Bernstein*

# Latin Square




***Adopted & modified from Scott Klemmer and Michael Bernstein***

# Latin Square

Task 1 Task 2 Task 3

Order 1

Order 2

Order 3


***Adopted & modified from Scott Klemmer and Michael Bernstein***

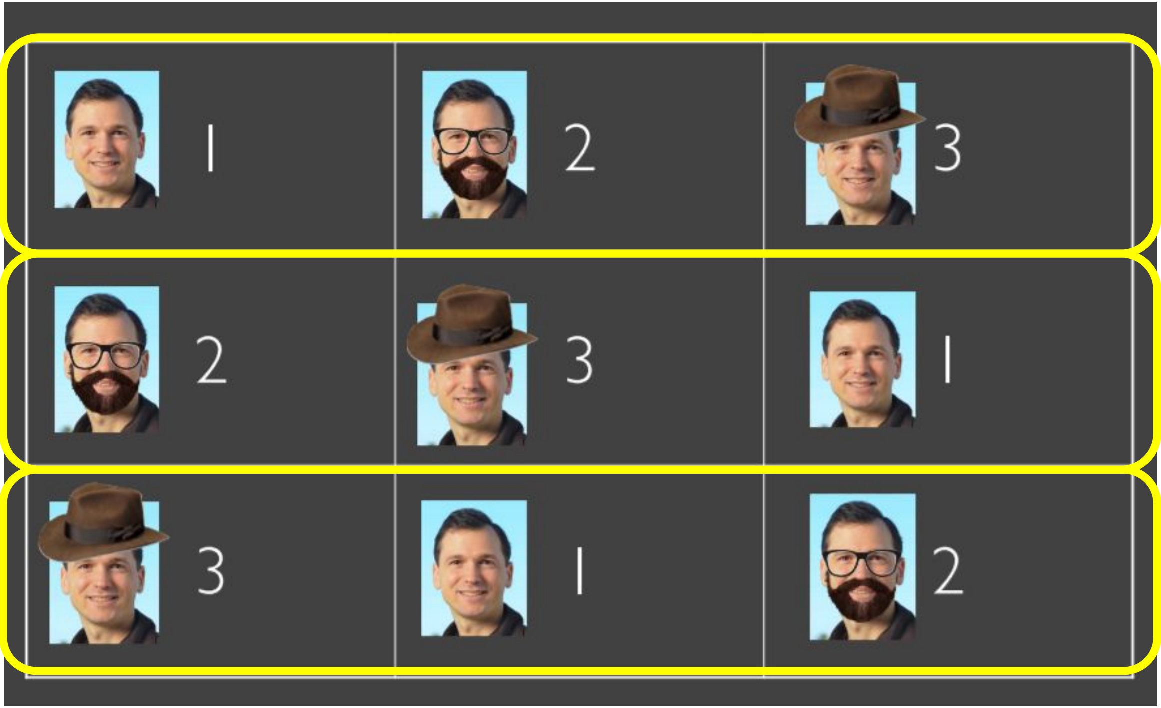
# Latin Square

Task 1 | Task 2 | Task 3

Order 1
1->2->3

Order 2
2->3->1

Order 3
3->1->2




**Detects the order effect!**

***Adopted & modified from Scott Klemmer and Michael Bernstein***

# Latin Square

Task 1 Task 2 Task 3

Order 1
1->2->3

Order 2
2->3->1

Order 3
3->1->2


**Detects the sequence effect!**

***Adopted & modified from Scott Klemmer and Michael Bernstein***

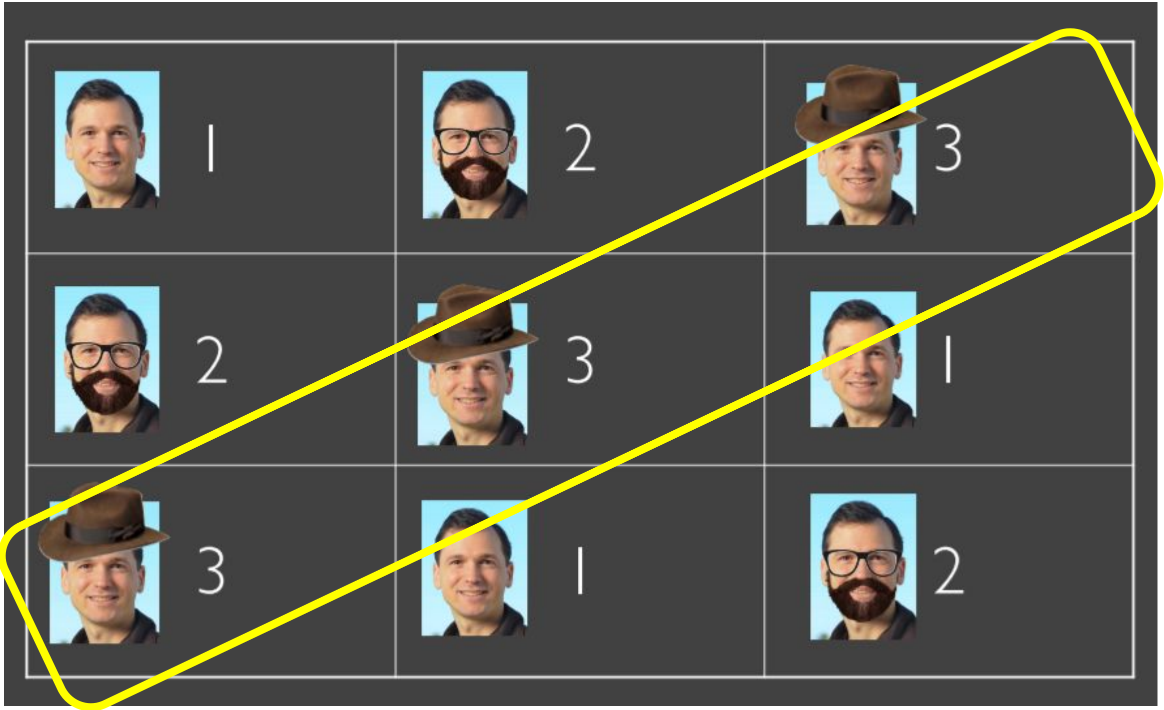
# Latin Square

Task 1 Task 2 Task 3

Order 1
1->2->3

Order 2
2->3->1

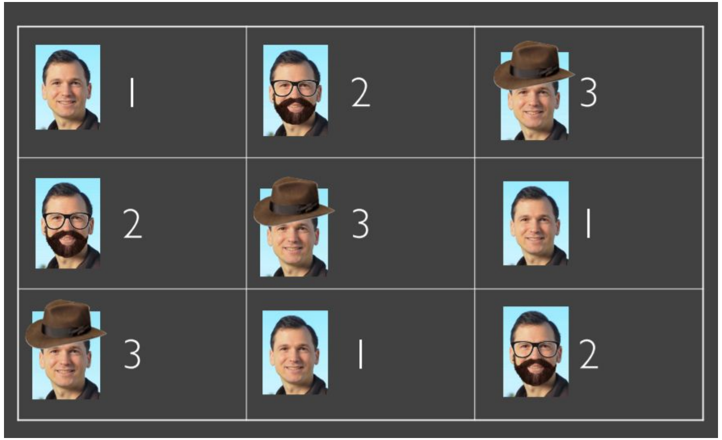
Order 3
3->1->2




**Detects the treatment effect!**

***Adopted & modified from Scott Klemmer and Michael Bernstein***

# Latin Square




$$\begin{bmatrix}1\end{bmatrix} \quad \begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix} \quad \begin{bmatrix}1 & 2 & 3\\ 2 & 3 & 1\\ 3 & 1 & 2\end{bmatrix}$$

$$\begin{bmatrix}1 & 2 & 3 & 4\\ 2 & 1 & 4 & 3\\ 3 & 4 & 1 & 2\\ 4 & 3 & 2 & 1\end{bmatrix} \quad \begin{bmatrix}1 & 2 & 3 & 4\\ 2 & 4 & 1 & 3\\ 3 & 1 & 4 & 2\\ 4 & 3 & 2 & 1\end{bmatrix}$$

$$\begin{bmatrix}1 & 2 & 3 & 4 & 5\\ 2 & 3 & 5 & 1 & 4\\ 3 & 5 & 4 & 2 & 1\\ 4 & 1 & 2 & 5 & 3\\ 5 & 4 & 1 & 3 & 2\end{bmatrix} \quad \begin{bmatrix}1 & 2 & 3 & 4 & 5\\ 2 & 4 & 1 & 5 & 3\\ 3 & 5 & 4 & 2 & 1\\ 4 & 1 & 5 & 3 & 2\\ 5 & 3 & 2 & 1 & 4\end{bmatrix}$$

***Examples of main-class Latin Squares of order 1 ~ 5***

**Compared to simple randomization, this Latin Square detects two blocking factors (sequence and order) instead of one. Simple randomization would've required 3 * 3 * 3 = 27 experiments, here, only 9. This is 18 / 27 * 100 = 66.7% reduction!**

***Adopted & modified from Scott Klemmer and Michael Bernstein***

# Latin Square

*** Careful design can further reduce the number of treatments required → *Graeco Latin Square***

# Statistical tests for subjective measures (e.g. Likert-scale questionnaire responses, etc.)

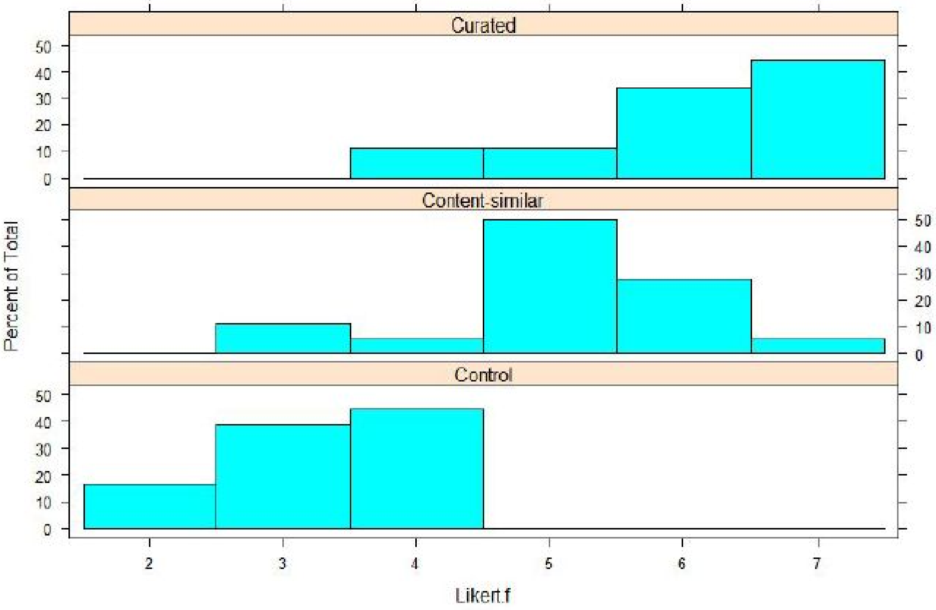
**Friedman test:** a non-parametric test for differences between groups when the dependent variable being measured is **ordinal** (or continuous). Some assumptions that have to be met:

1: Same group of subjects measured on **three or more** different occasions.
2: Group is randomly sampled from the entire population.
3: Your dependent variable should be measured at the ordinal (e.g. 7-point Likert scale) or continuous (e.g. temperature) level.
4: Samples do NOT need to be normally distributed.

→ Tells whether there were differences between groups but not exactly where they occurred.

# Statistical tests for subjective measures (e.g. Likert -scale questionnaire responses, etc.)

| Condition | Rater | Likert |
|---|---|---|
| Control | 1 | 4 |
| Content-similar | 1 | 5 |
| Curated | 1 | 6 |
| Control | 2 | 3 |
| Content similar | 2 | 5 |
| Curated | 2 | 5 |
| Control | 3 | 4 |
| Content-similar | 3 | 5 |
| Curated | 3 | 7 |
| Control | 4 | 2 |




Report the result as: "*There was a statistically significant difference in easiness of applying rubric in design critique depending on the type of critique , χ2(2) = 7.600, p = 0.022.*"

# Statistical tests for subjective measures (e.g. Likert -scale questionnaire responses, etc.)

**Wilcoxon signed-rank test:** a non-parametric post-hoc test to check for where the differences actually occurred. Assumptions
1: Your dependent variable should be measured at the ordinal (e.g. 7-point Likert scale) or continuous (e.g. temperature) level.
2: Your independent variable should consist of two categorical, related groups or matched pairs.
3: The distribution of the differences between the two related groups needs to be symmetrical in shape.

Report the result as: *"Wilcoxon signed-rank test showed that a 4 week, twice weekly acupuncture treatment course did not elicit a statistically significant change in lower back pain in individuals with existing lower back pain (Z = -1.807, p = 0.071). Indeed, median Pain Score rating was 5.0 both pre- and post-treatment."*

***source: https://statistics.laerd.com/spss-tutorials/wilcoxon-signed-rank-test-using-spss-statistics.php***

Thank you

**Appendix A.**

The original (vectorized form) of Charnov’s Marginal Value Theorem

# The formal **Charnov's Marginal Value Theorem** (in the vectorized form)

For patches {1, 2, …, P}: patch foraging times are $(t_{W1}, t_{W2}, \ldots, t_{WP})$ and the rate of gain $R = \dfrac{\lambda g_i(t_{Wi}) + k_i}{c_i + \lambda_i t_{Wi}}$

For each t_{wi}, maximization of R should satisfy (a set of  P equations)

$$\frac{\partial R}{\partial t_{Wi}} = \frac{\lambda_i g_i'(t_{Wi})[\lambda_i t_{Wi} + c_i] - \lambda_i[\lambda_i g_i(t_{Wi}) + k_i]}{(\lambda_i t_{Wi} + c_i)^2}$$

And setting the partial derivative to zero

$$g_i'(t_{Wi})[\lambda_i t_{Wi} + c_i] - [\lambda_i g_i(t_{Wi}) + k_i] = 0$$

**=>** $$g_i'(t_{Wi}) = \frac{\lambda_i g_i(t_{Wi}) + k_i}{\lambda_i t_{Wi} + c_i} = R. \longrightarrow$$

$$g_1'(\hat{t}_{W1}) = R(\hat{t}_{W1}, \hat{t}_{W2}, \ldots, \hat{t}_{WP})$$

$$g_2'(\hat{t}_{W2}) = R(\hat{t}_{W1}, \hat{t}_{W2}, \ldots, \hat{t}_{WP})$$

$$\cdots$$

$$g_p'(\hat{t}_{WP}) = R(\hat{t}_{W1}, \hat{t}_{W2}, \ldots, \hat{t}_{WP}).$$

## Appendix B.

Derivation of the optimal information diet selection algorithm

## Information diet selection (there are not just one type of patches!)

If there are different types of information, (differing in their respective profitability), then you'll need to pursue items of the type of a specific profitability in an all-or-none manner; never have a mixed diet. **(zero-one rule)**

To create a decision model, introduce a new **probability parameter** $p_i$ -- we'll eventually solve for p_i's

$$R = \frac{\sum_{i=1}^{P} p_i \lambda_i g_i}{1 + \sum_{i=1}^{P} p_i \lambda_i t_{W_i}} = \frac{p_i \lambda_i g_i + \sum_{j \in P \setminus \{i\}} p_j \lambda_j g_j}{1 + p_i \lambda_i t_{W_i} + \sum_{j \in P \setminus \{i\}} p_j \lambda_j t_{W_j} p_j \lambda_j t_{W_j}}$$

$$= \frac{p_i \lambda_i g_i + k_i}{p_i \lambda_i t_{W_i} + c_i}$$

$$k_i = \sum_{j \in P \setminus \{i\}} p_j \lambda_j g_j$$

$$c_i = 1 + \sum_{j \in P \setminus \{i\}} p_j \lambda_j t_{W_j}$$

### Table of notation

| | |
|---|---|
| G | the ratio of the total net amount of valuable info. gained |
| T_B | The total amount of time spent between-patches |
| T_W | The total amount of time spent within-patches foraging |
| g | The avg. gain per patch |
| t_B | The avg. time between processing patches |
| t_W | The avg. time to process patches |

## Information diet selection (there are not just one type of patches!)

Deriving by $p_i$ yields

$$\frac{\partial R}{\partial p_i} = \frac{\lambda_i g_i c_i - \lambda_i t_{Wi} k_i}{(c_i + p_i \lambda_i t_{Wi})^2}$$

The righthand-side of the equation is either >0 or <0, independent of $p_i$. Therefore maximization happens when

$$p_i = 0 \text{ if } g_i / t_{Wi} < k_i / c_i$$

$$p_i = 1 \text{ if } g_i / t_{Wi} > k_i / c_i$$

## Table of notation

| | |
|---|---|
| G | the ratio of the total net amount of valuable info. gained |
| T_B | The total amount of time spent between-patches |
| T_W | The total amount of time spent within-patches foraging |
| g | The avg. gain per patch |
| t_B | The avg. time between processing patches |
| t_W | The avg. time to process patches |

## Information diet selection (there are not just one type of patches!)

Deriving by $p_i$ yields

$$\frac{\partial R}{\partial p_i} = \frac{\lambda_i g_i c_i - \lambda_i t_{W_i} k_i}{(c_i + p_i \lambda_i t_{W_i})^2}$$

The righthand-side of the equation is either >0 or <0, independent of $p_i$. Therefore maximization happens when

$$p_i = 0 \text{ if } g_i/t_{W_i} < \boxed{k_i/c_i}$$

$$p_i = 1 \text{ if } g_i/t_{W_i} > \boxed{k_i/c_i}$$

**This is the profitability of "other" types acquired so far; which suggests a greedy* algorithm for diet selection (think about starting from the most profitable)**

### Table of notation

| | |
|---|---|
| G | the ratio of the total net amount of valuable info. gained |
| T_B | The total amount of time spent between-patches |
| T_W | The total amount of time spent within-patches foraging |
| g | The avg. gain per patch |
| t_B | The avg. time between processing patches |
| t_W | The avg. time to process patches |

## Information diet selection (there are not just one type of patches!)

### Algorithm for optimal diet selection

Suppose that we can sort item types in terms of their profitability

$$\pi_1 > \pi_2 > \ldots > \pi_n.$$

Add item type *k+1* from the most profitable to the least, **until** the rate of gain for a diet of *k* item types already added is greater than profitability of the *k+1*$^{st}$ type

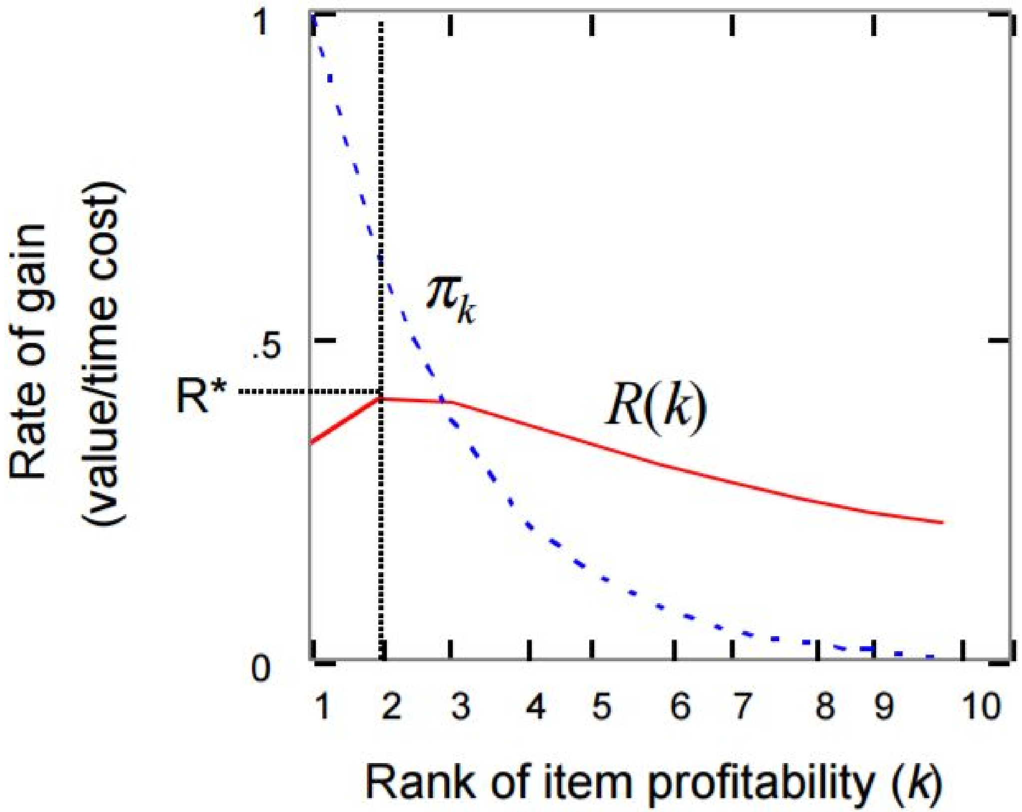
$$R(k) = \frac{\sum_{i=1}^{k} \lambda_i g_i}{1 + \sum_{i=1}^{k} \lambda_i t_{Wi}} > \frac{g_{k+1}}{t_{Wk+1}} = \pi_{k+1}$$

### Table of notation

| | |
|---|---|
| G | the ratio of the total net amount of valuable info. gained |
| T_B | The total amount of time spent between-patches |
| T_W | The total amount of time spent within-patches foraging |
| g | The avg. gain per patch |
| t_B | The avg. time between processing patches |
| t_W | The avg. time to process patches |

**Graphical representation**

Q. Which item types are chosen for optimal rate of gain?



# References https://goo.gl/VsC2IK


[1] Pirolli, Peter, and Wai-Tat Fu. "SNIF-ACT: A model of information foraging on the World Wide Web." *International Conference on User Modeling*. Springer Berlin Heidelberg, 2003.

[2] Chi, Ed H. "Information seeking can be social." *Computer* 42.3 (2009): 42-46.

[3] Cutting, Douglass R., David R. Karger, and Jan O. Pedersen. "Constant interaction-time scatter/gather browsing of very large document collections." *Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 1993.

[4] Anderson, John R. "A spreading activation theory of memory." *Journal of verbal learning and verbal behavior* 22.3 (1983): 261-295.

[5] Newell, Allen, and Stuart K. Card. "The prospects for psychological science in human-computer interaction." *Human-computer interaction* 1.3 (1985): 209-242.

[6] Anderson, John R., and Christian Lebiere. "The Newell test for a theory of cognition." *Behavioral and brain Sciences* 26.05 (2003): 587-601.

[7] Patterson, Karalyn, Peter J. Nestor, and Timothy T. Rogers. "Where do you know what you know? The representation of semantic knowledge in the human brain." Nature Reviews Neuroscience 8.12 (2007): 976-987.

[8] Huberman, Bernardo A., et al. "Strong regularities in world wide web surfing." *Science* 280.5360 (1998): 95-97.

[9] Card, Stuart K., et al. "Information scent as a driver of Web behavior graphs: results of a protocol analysis method for Web usability." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2001.

[10] Chi, Ed H., Peter Pirolli, and James Pitkow. "The scent of a site: A system for analyzing and predicting information scent, usage, and usability of a web site." *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 2000.

[11] Pirolli, Peter, Evelin Wollny, and Bongwon Suh. "So you know you're getting the best possible information: a tool that increases Wikipedia credibility." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2009.