Model-based evaluations

Lecture 15

dr. Kristina Lapin

Predictive models

- Provide a way of evaluating products or designs
  - without directly involving users.
- Less expensive than user testing.
- Usefulness limited to systems with predictable tasks
  - e.g., telephone answering systems, mobiles, cell phones, etc.
- Based on expert error-free behavior.

What is GOMS?

- A family of user interface modeling techniques
- Goals, Operators, Methods, and Selection rules
  - Higher level than KLM
  - Input: detailed description of UI and task(s)
  - Output: various qualitative and quantitative measures

Slide adapted from Chris Long

GOMS

- **G**oals
  - what the user wants to achieve eg. find a website.
- **O**perators
  - the cognitive processes & physical actions needed to attain goals,
  - e.g. decide which search engine to use.
- Methods
  - the procedures to accomplish the goals,
  - eg. drag mouse over field, type in keywords, press the go button.
- Selection rules
  - decide which method to select when there is more than one.

What can GOMS model?

- Task must be goal-directed
  - Some activities are more goal-directed than others
  - Even creative activities contain goal-directed tasks
- Task must be a routine cognitive skill
- Can include serial and parallel tasks

Slide adapted from Chris Long

GOMS Output

- Functionality coverage and consistency
  - Does UI contain needed functions?
  - Are similar tasks performed similarly? (NGOMSL only)
- Operator sequence
  - In what order are individual operations done?
  - Abstraction of operations may vary among models

Slide adapted from Chris Long

GOMS Output (cont'd)

- Execution time
  - By expert user
- Error recovery
- Procedure learning time (NGOMSL only)
  - Useful for relative comparison only
  - Does not include time for learning domain knowledge

Slide adapted from Chris Long

Applications of GOMS analysis

- Comparing UI designs
- Profiling
- Building a help system
  - GOMS modelling makes user tasks and goals explicit
  - Can suggest questions users will ask and the answers

Slide adapted from Chris Long

GOMS Example 1: PDA Text Entry

- goal: enter-text-PDA
  - move-pen-to-text-start
  - goal: enter-word-PDA
  - ...repeat until no more words
    - write-letter ...repeat until no more letters
    - [select: goal: correct-misrecognized-word] ...if incorrect
- expansion of correct-misrecognized-word goal:
  - move-pen-to-incorrect-letter
  - write-letter

Slide adapted from Chris Long

GOMS Example

- Move text in a word processor
  - (example from Hochstein)

```
GOAL: EDIT-MANUSCRIPT
.     GOAL: EDIT-UNIT-TASK ... repeat until no more unit tasks
.     .     GOAL: ACQUIRE UNIT-TASK
.     .     .     GOAL: GET-NEXT-PAGE ... if at end of manuscript page
.     .     .     GOAL: GET-FROM-MANUSCRIPT
.     .     GOAL: EXECUTE-UNIT-TASK ... if a unit task was found
.     .     .     GOAL: MODIFY-TEXT
.     .     .     .   [select: GOAL: MOVE-TEXT* ...if text is to be moved
.     .     .     .         GOAL: DELETE-PHRASE ...if a phrase is to be deleted
.     .     .     .         GOAL: INSERT-WORD] ... if a word is to be inserted
.     .     .     .     VERIFY-EDIT
```

Members of GOMS Family

- Keystroke-Level Model (KLM) –
  - Card, Moran, Newell (1983)
- CMN-GOMS
  - Card, Moran, Newell GOMS
- Natural GOMS Language (NGOMSL)
  - -Kieras (1988+)
- Critical Path Method or Cognitive, Perceptual, and Motor GOMS (CPM-GOMS)
  - John (1990+)

Slide adapted from Chris Long

Real-world Applications of GOMS

- KLM
  - Mouse-based text editor
  - Mechanical CAD system
- NGOMSL
  - TV control system
  - Nuclear power plant operator's associate
- CPM-GOMS
  - Telephone operator workstation

Slide adapted from Chris Long

Advantages of GOMS

- Gives several qualitative and quantitative measures
- Model explains **why** the results are what they are
- Less work (?) than usability study
- Easy (?) to modify when interface is revised
- Research ongoing for tools to aid modeling process

Slide adapted from Chris Long

Disadvantages of GOMS

- Not as easy as heuristic analysis, guidelines, or cognitive walkthrough
- Only works for goal-directed tasks
- Assumes tasks are performed by expert users
- Evaluator must pick users' tasks/goals
- Does not address several important UI issues, such as
  - readability of text
  - memorability of icons, commands
- Does not address social or organizational impact

Slide adapted from Chris Long

GOMS Summary

- Provides info about many important UI properties
- But does not tell you most of what you want to know about a UI
- Substantial effort to do initial model,
  - but still (potentially) easier than user testing
- Changing later is much less work than initial generation

Slide adapted from Chris Long

Keystroke-level Model

- Another "discount" usability method
- Main idea:
  - Walk through the interface, counting how many operations it would take an expert user to perform
  - Look for ways to optimize
  - Look for potential sources of error
- KLM is very low-level (tiny operations)

Keystroke level model

- GOMS has also been developed to provide a quantitative model - the keystroke level model.
- The keystroke model allows predictions to be made about how long it takes an expert user to perform a task.

Keystroke-Level Model

- How to make a KLM
  - List specific actions user does to perform task
    - Keystrokes and button presses
    - Mouse movements
    - Hand movements between keyboard & mouse
    - System response time (if it makes user wait)
  - Add Mental operators
  - Assign execution times to steps
  - Add up execution times
- Only provides execution time and operator sequence

Slide adapted from Chris Long

Response times for keystroke level operators (Card et al., 1983)

Operator	Description	Time (sec)
K	Pressing a single key or button	
	Average skilled typist (55 wpm)	0.22
	Average non-skilled typist (40 wpm)	0.28
	Pressing shift or control key	0.08
	Typist unfamiliar with the keyboard	1.20
P	Pointing with a mouse or other device on a display to select an object. This value is derived from Fitts' Law which is discussed below.	0.40
P1	Clicking the mouse or similar device	0.20
H	Bring 'home' hands on the keyboard or other device	0.40
M	Mentally prepare/respond	1.35
R(t)	The response time is counted only if it causes the user to wait.	t

Rules for placing Ms

- Rule 0, initial insertion for candidate Ms
  - insert Ms in front of all Ks
  - place Ms in front of Ps that select commands, but not Ps that select arguments for the commands
- Rule 1, deletion of anticipated Ms
  - delete M between two operators if the second operator is fully anticipated in the previous one (e.g., PMK → PK)
- Rule 2, deletion of Ms within cognitive units
  - in a string of MKs that form a cognitive unit, delete all Ms except the first (e.g., "Helen of Troy", 745.8)

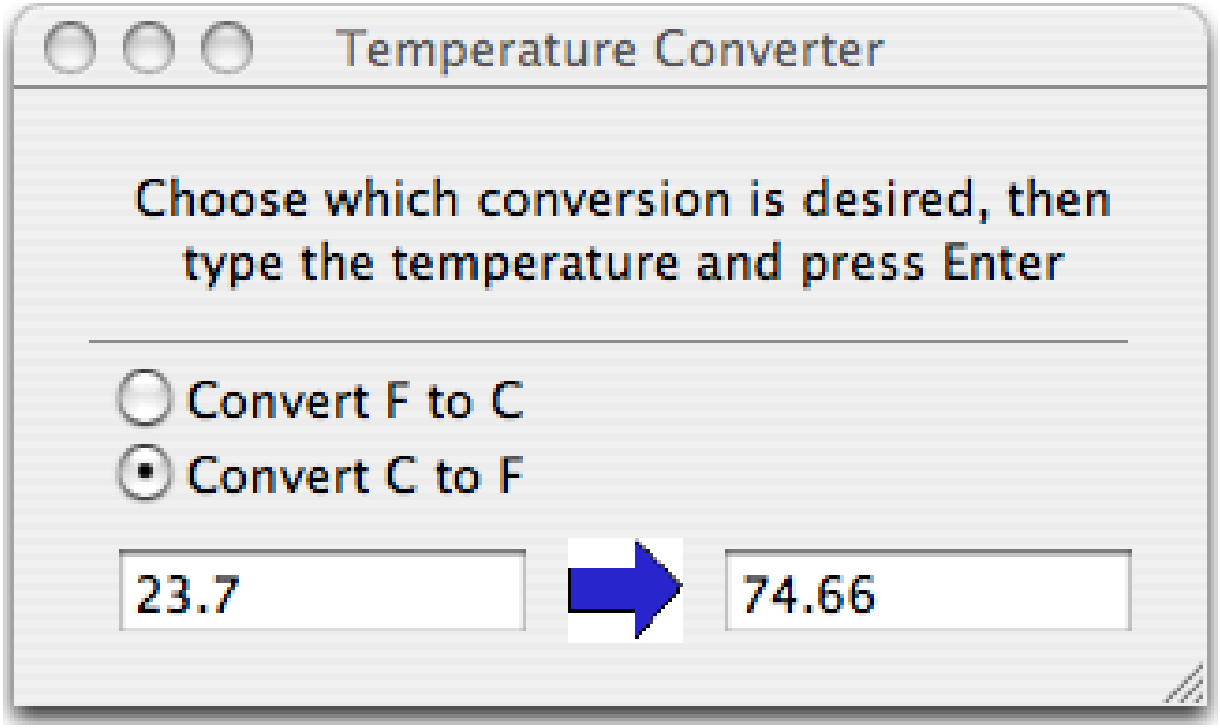
Rules for placing Ms

- Rule 3, deletion of Ms before consecutive terminators
  - if K is redundant delimiter at end of a cognitive unit, delete the M in front of it, e.g.,)’
- Rule 4, deletion of Ms that are terminators of commands
  - if K is a delimiter that follows a constant string then delete the M in front of it
    - (not for arguments or varying strings)
- Rule 5, deletion of overlapped Ms
  - do not count any M that overlaps an R
    - (e.g., user waiting for computer response)

Exercise: temperature converter

- Convert from degrees Fahrenheit (F) to Celsius (C) or v.v., requests equally distributed
- Use keyboard or GID to enter temperature
- Assume active window awaiting input, an average of four typed characters (including point and sign), and no typing errors
- Task: create and analyze your own interface!

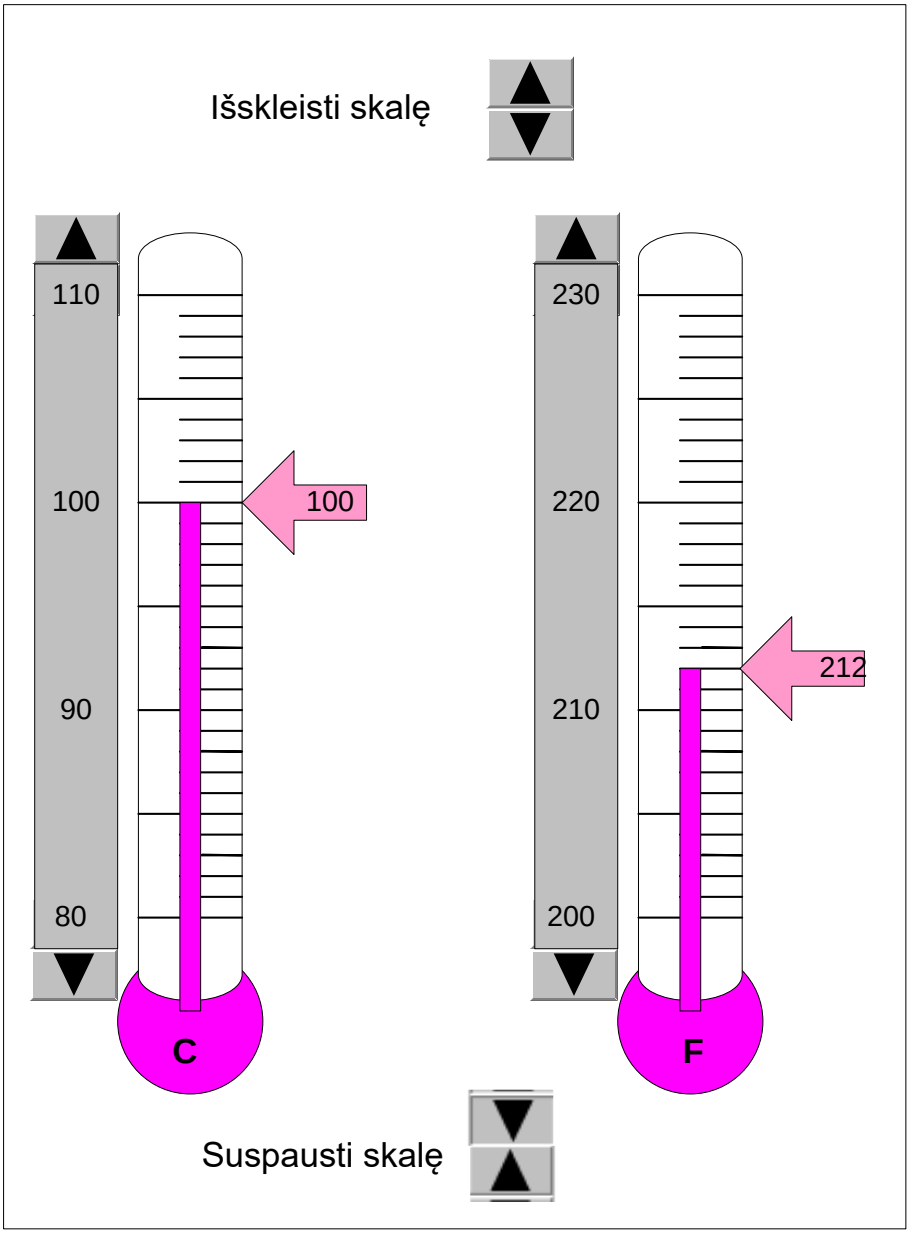
The dialog box solution with radio buttons…



Keystroke-level model

- Case 1: select conversion direction
  - move hand to mouse,
  - point to desired button,
  - click on radio button
    - (HPK)
  - move hands back to keyboard, type four characters, tap enter
    - (HPKHKKKKK)
  - Rule 0 (HMPMKHMKMKMKMKMK)
  - Rule 1, 2, 4 (HMPKHMKKKKMK)
  - Estimated time = 7.15 sec
- Case 2: correct conversion direction already selected
- MKKKKKMK = 3.7 sec
- Average time = (7.15 + 3.7) / 2 = 5.4 sec

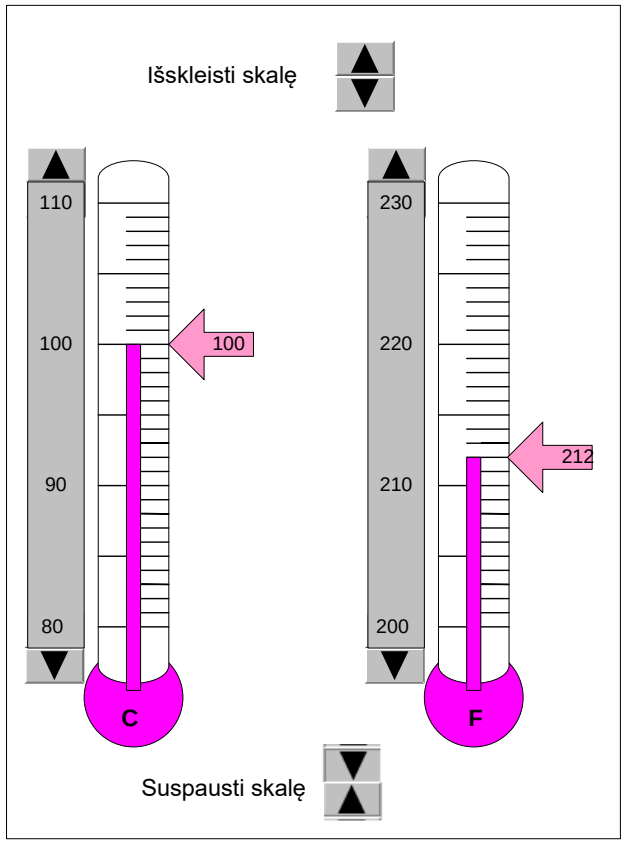
Temperature converter: alternative mockup



- Temperature scale and range is correct:
  - HPK
  - PK
  - HMPMKMPMK
  - HMPKPK=

  0.4+1.35+1.1+0.2+1.1 +0.2=4.35

Temperature converter: alternative mockup



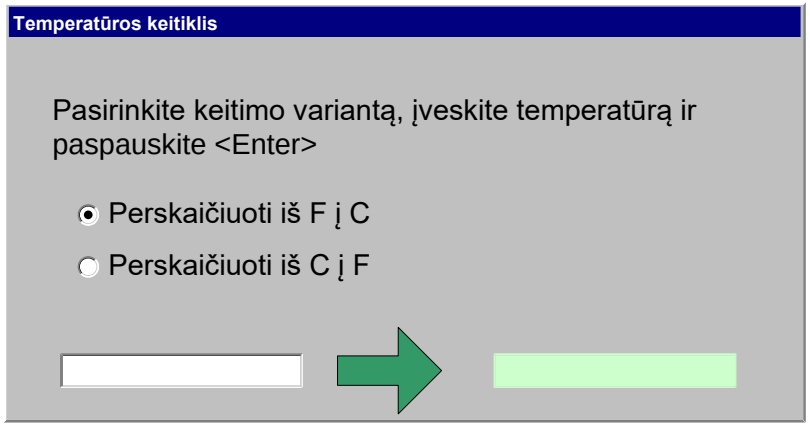
Set the scale and range.

Try to convert few temperatures: scale and range is set in 3 sec. (operacija S)

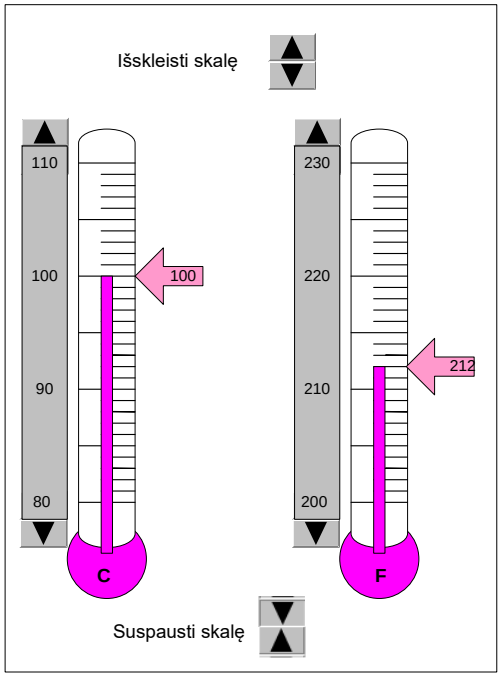
- HPKSKPKSKPKSKPKK
- H+3(M+P+K+S+K)+M+P+K+K = 0,4+3*(1,35+0,2+3,0+0,2)+1,35+0,4+0,2+0,2=
- 16,8 sec.

Comparison

Average time: 5,7 sec.



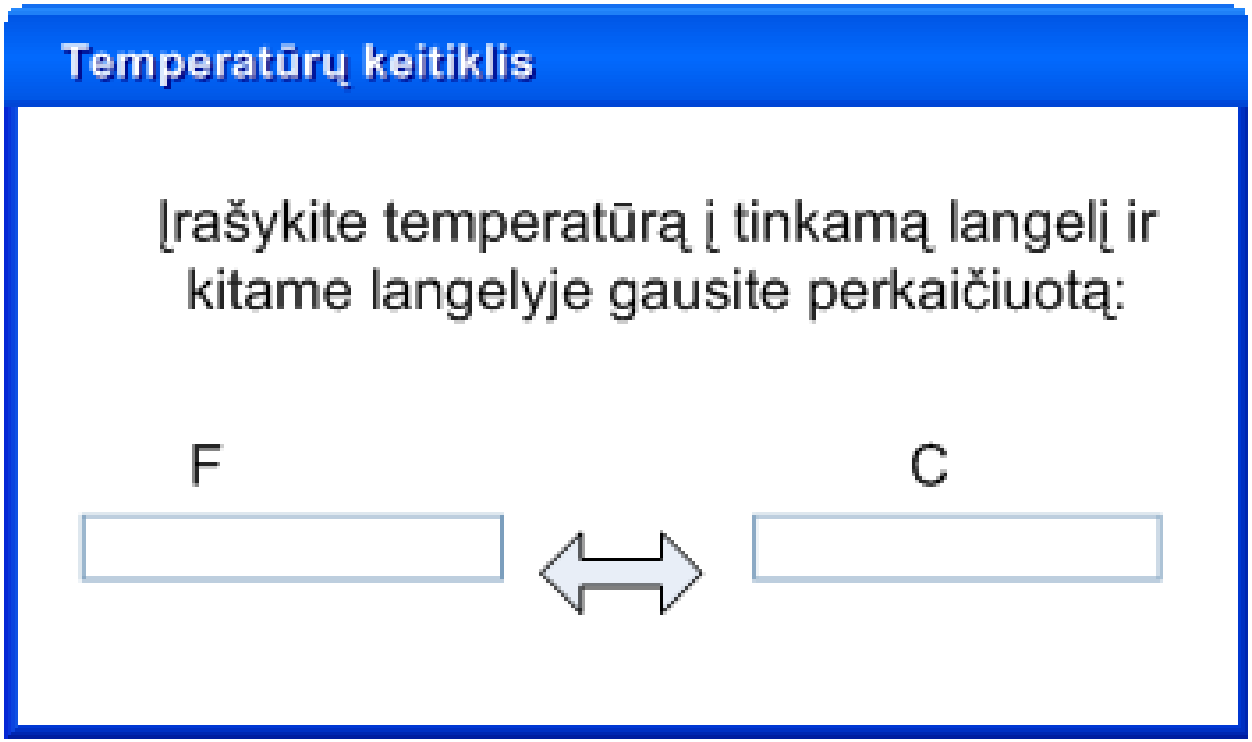
Average time: 10,58 sec.



Summary

- First mockup is more efficient
- Alternatyve mockup suggested an improvement
  - Reduce the radio buttons

Improvement



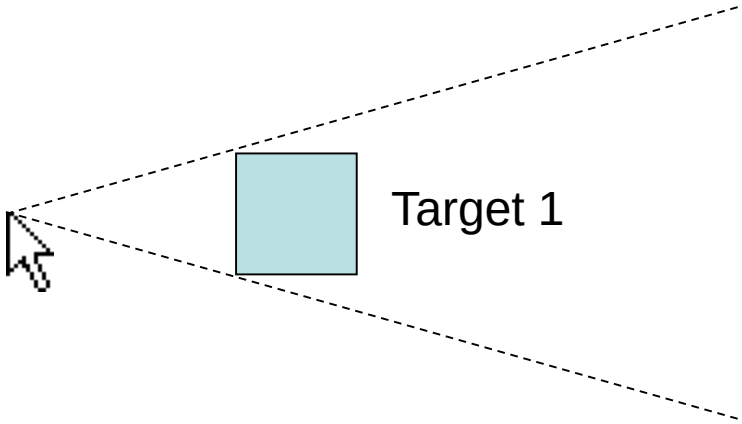
Fitts' Law (Fitts, 1954)

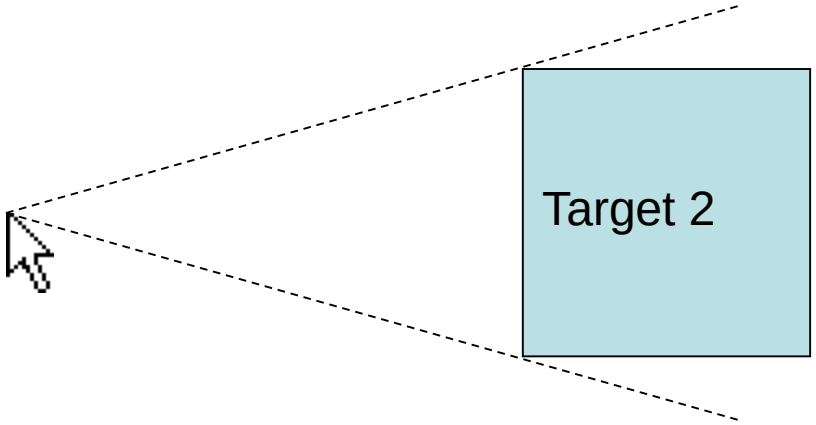
- Fitts' Law predicts that the time to point at an object using a device is a function of the distance from the target object & the object's size.
- The further away & the smaller the object, the longer the time to locate it & point to it.
- Fitts' Law is useful for evaluating systems for which the time to locate an object is important, e.g., a cell phone, a handheld devices.

Fitts' Law

Time = k $\log_2$(D/S+1)

Time = time to move the pointer to a target
D = distance between the pointer and the target
S = size of the target
k is a constant of approximately 200 ms/bit





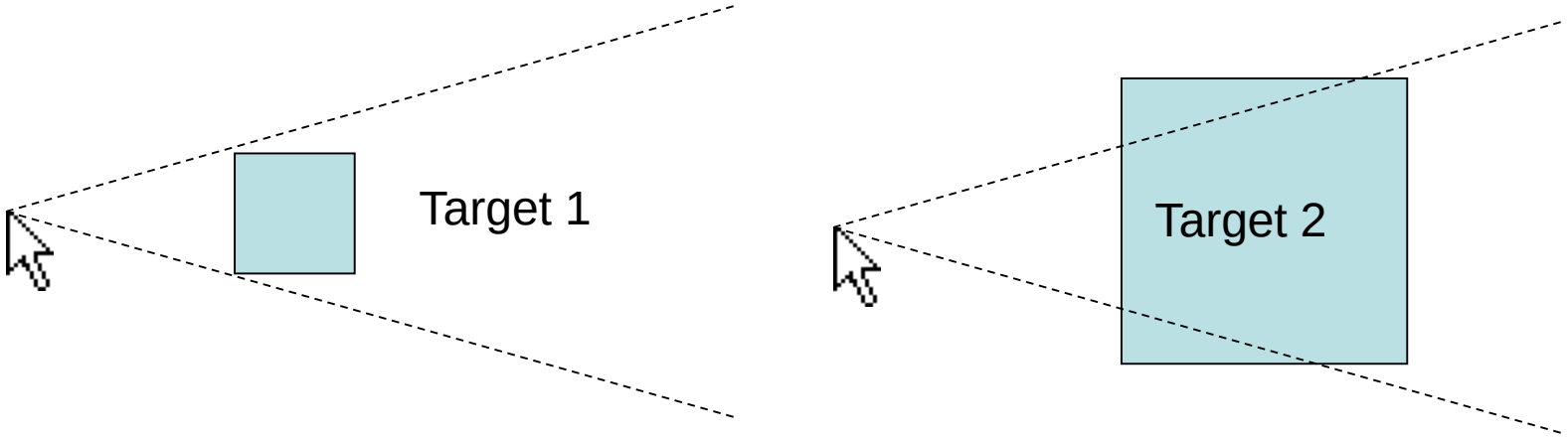
Same ID → Same Difficulty

Slide adapted from
Pourang Irani

Fitts' Law

Time = k $\log_2$(D/S+1)

Time = time to move the pointer to a target
D = distance between the pointer and the target
S = size of the target
k is a constant of approximately 200 ms/bit



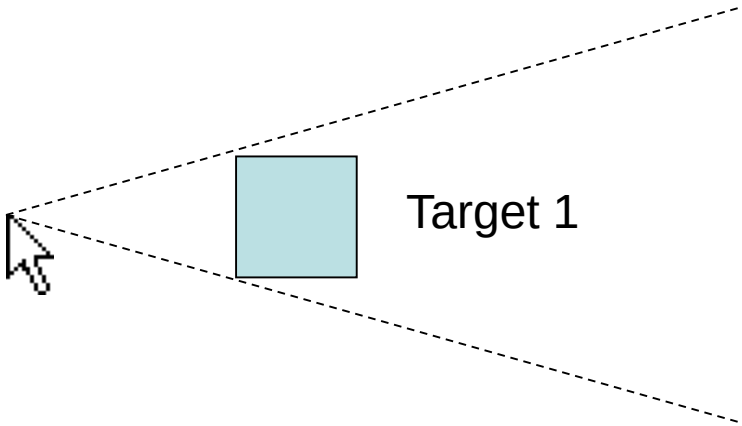
Smaller ID → Easier

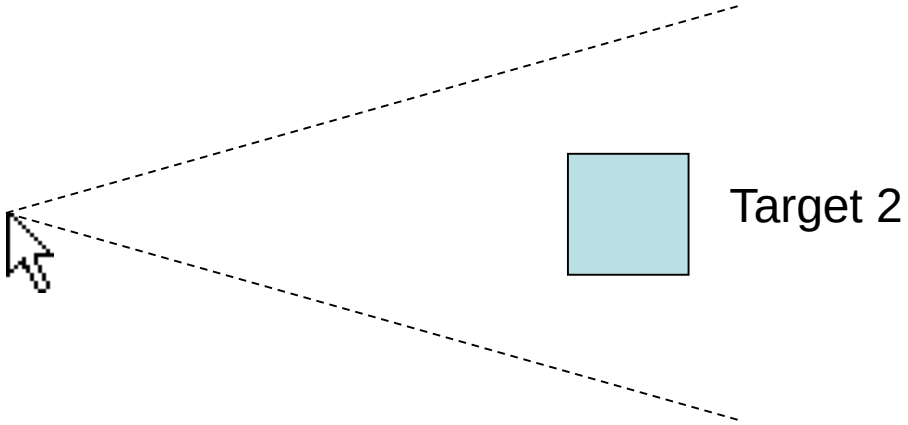
Slide adapted from
Pourang Irani

Fitts' Law

Time = k $\log_2$(D/S+1)

Time = time to move the pointer to a target
D = distance between the pointer and the target
S = size of the target
k is a constant of approximately 200 ms/bit



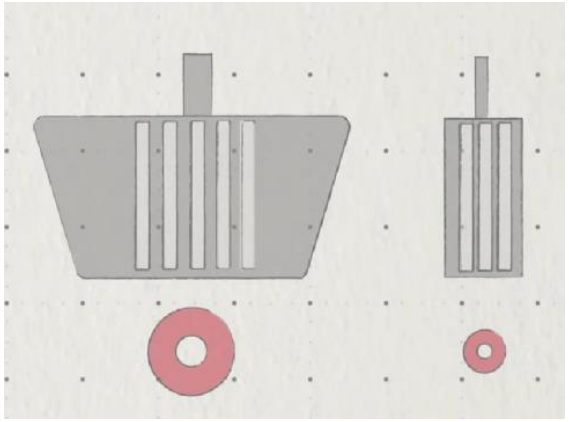


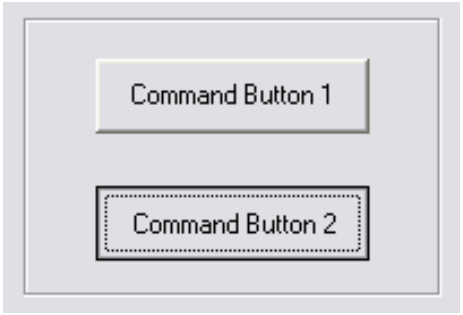
Larger ID → Harder

Slide adapted from
Pourang Irani

Fitts Law in interfaces

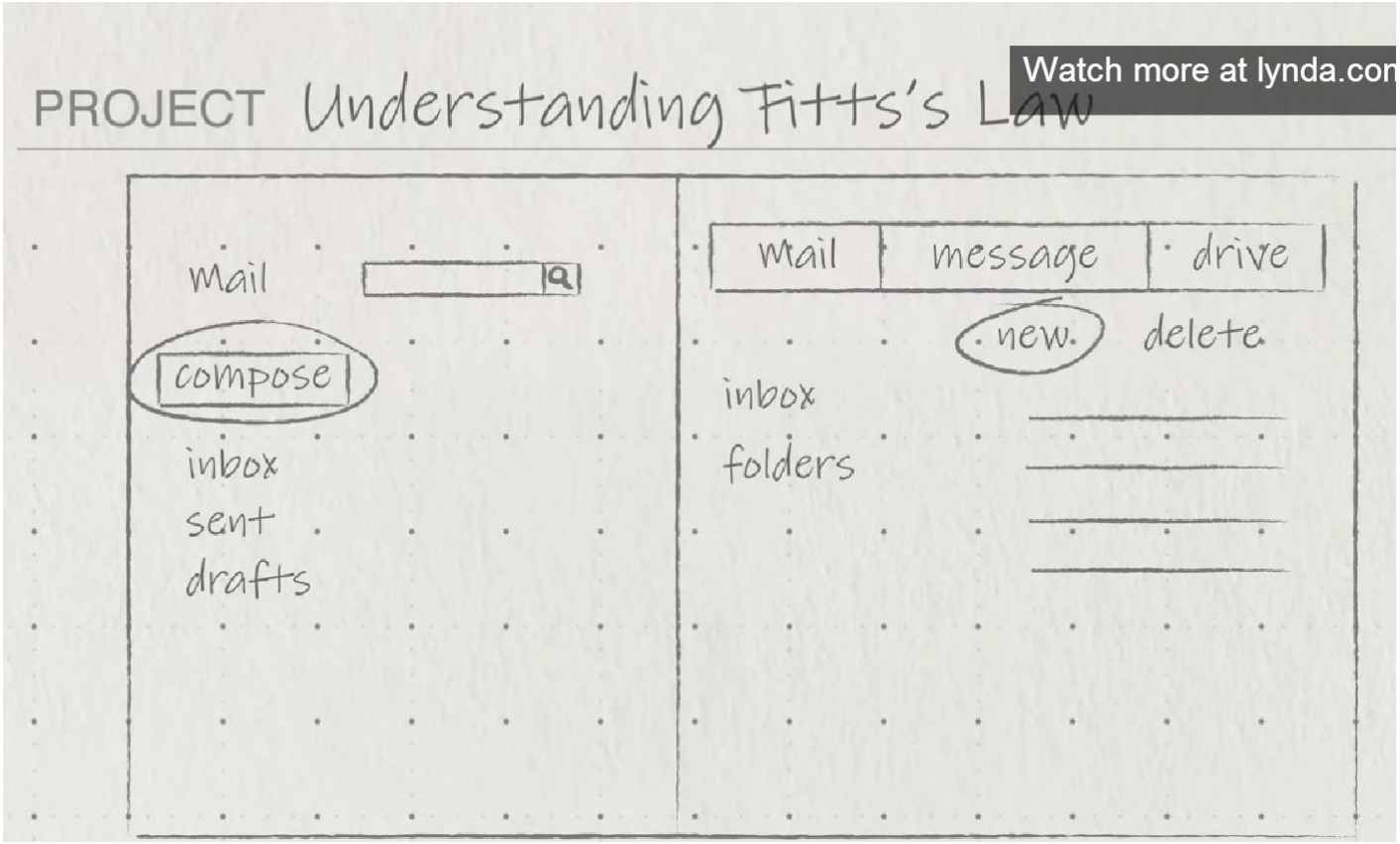
- As the distance increases, movement takes longer and
  - as the size decreases selection again takes longer





Fitts's Law: The Importance of Size and Distance in UI Design. Interaction Design Foundation

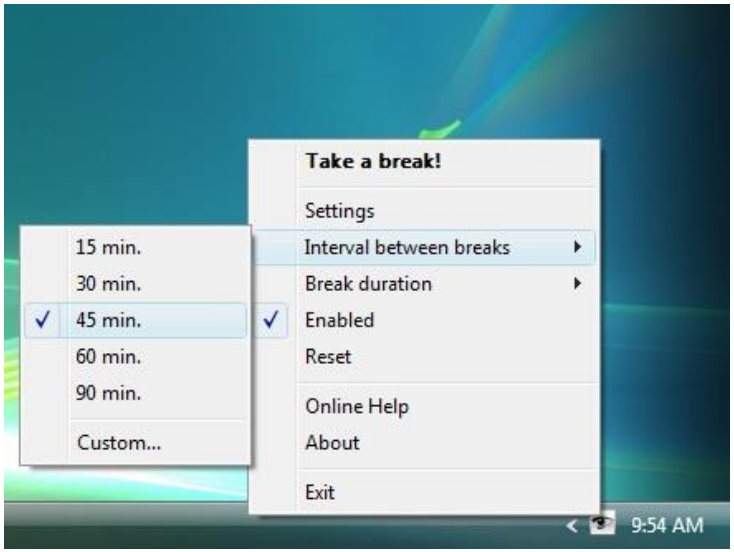
Which selection is quicker: compose or new?



User experience tutorial: Fitts's Law | lynda.com

Fitts law in interface design

- Context menu is quick,
  - Because is placed close to cursor position



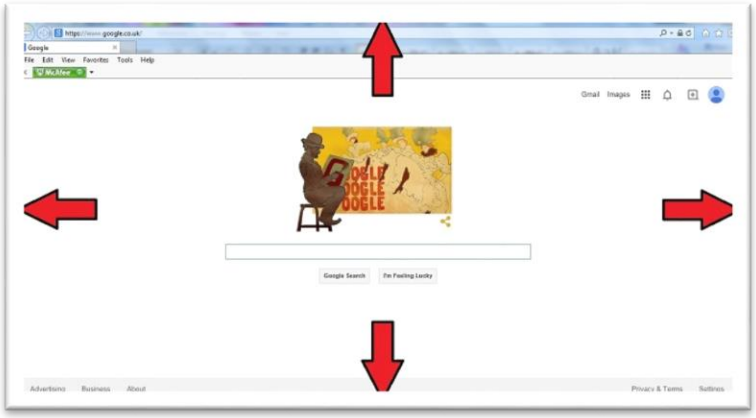
Fitts's Law: The Importance of Size and Distance in UI Design. Interaction Design Foundation

Fitts law in interface design

- Macintosh task selection is quicler than Windows
  - Windows menu height: 5 mm
    - *50+150 log_2(80/5+1) ≈ 663 msec*
  - Mac menu can be selected in 50 mm
    - *50+150 log2(80/50+1) ≈ 256 msec*

Fitts law in interface design

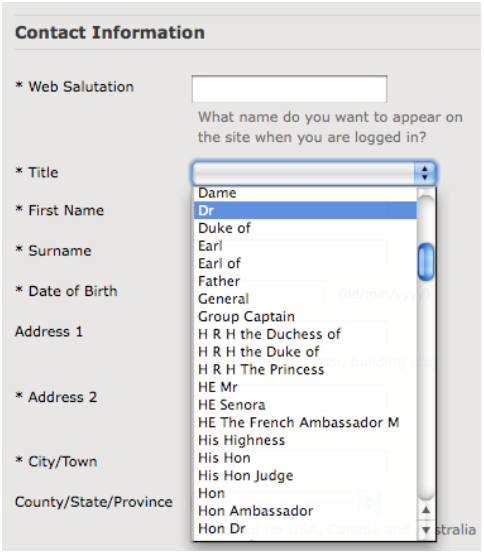
- **The outer edges and corners** of the graphical user interface can be acquired with greater speed than anywhere else in the display,
  - due to the pinning action of the screen.
    - As the user is restricted in their movements the pointing device cannot move any further when they reach the outermost points of the screen; fixing the cursor at a point on the periphery of the display.

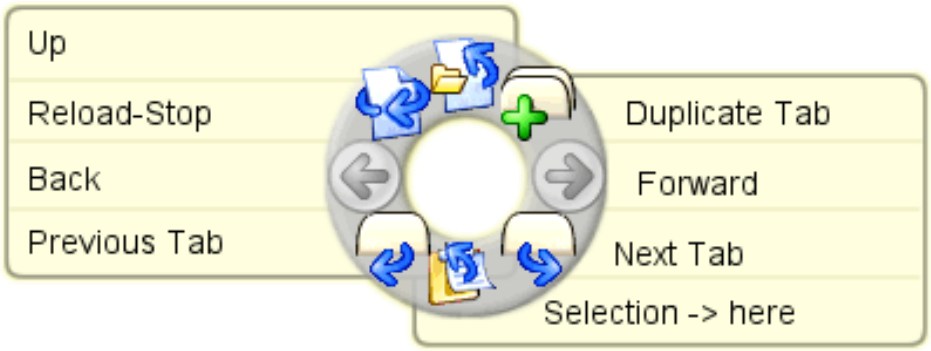


Fitts's Law: The Importance of Size and Distance in UI Design. Interaction Design Foundation

Fitts law in interface design

- **Pop-up menus** better support immediate selection of interactive elements than dropdown menus as
  - the user does not have to move the cursor from its current position.
    - Therefore, graphical designs that allow the user to interact without moving help to reduce the 'travel time'.

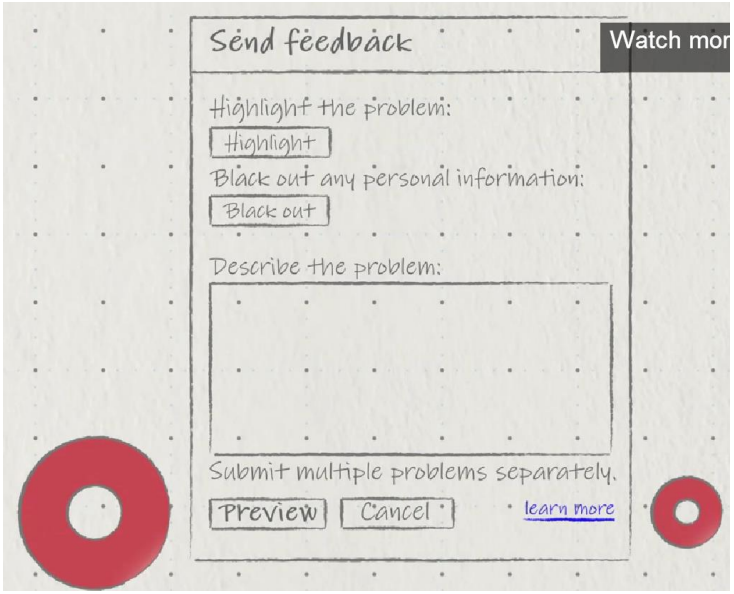




Fitts's Law: The Importance of Size and Distance in UI Design.
Interaction Design Foundation

Fitts law in interface design

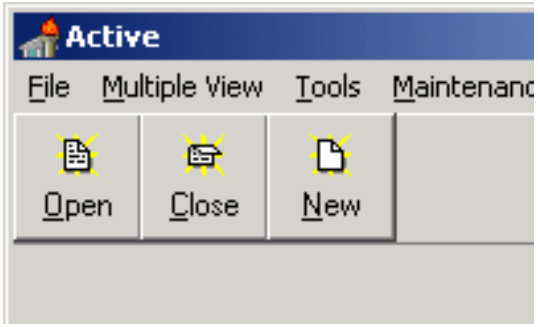
- More frequent actions could be bigger than others



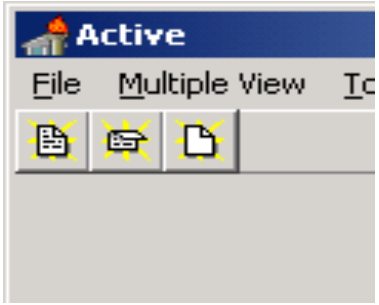
User experience tutorial: Fitts's Law | lynda.com

A Quiz Designed to Give You Fitts

http://www.asktog.com/columns/022DesignedToGiveFitts.html



Microsoft Toolbars offer the user the option of displaying a label below each tool. Name at least one reason why labeled tools can be accessed faster. (Assume, for this, that the user knows the tool.)



Slide adapted from Pourang Irani

A Quiz Designed to Give You Fitts

1. The label becomes part of the target. The target is therefore bigger. Bigger targets, all else being equal, can always be acccessed faster, by Fitt's Law.

2. When labels are not used, the tool icons crowd together.

Slide adapted from
Pourang Irani

Fitts' Law evaluations

Input Technique	Mouse	Stylus	Eye with Manual Click	Eye with Dwell Time Click
Movement time *(s.e.)*	*.66 (.03)*	*.63 (.03)*	*.57 (.04)*	*.45 (.02)*

Table 1. Mean movement times (s) and standard error per input technique.

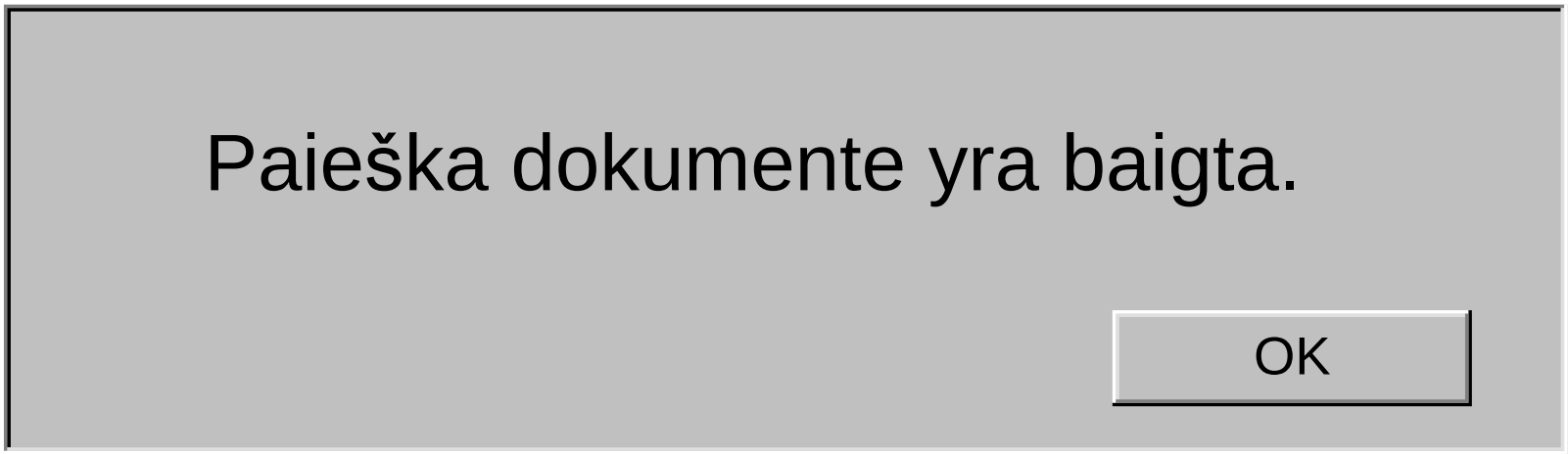
Vertegaal (2008)

Input Technique	Mouse	Stylus	Eye with Manual Click	Eye with Dwell Time Click
Error rate *(s.e.)*	*4.6% (1.3%)*	*6.2% (1.5%)*	*11.7% (3.5%)*	*42.9% (3.7%)*

Table 2. Mean error rates and standard error per input technique.

Interface efficiency

Non-efficient interface

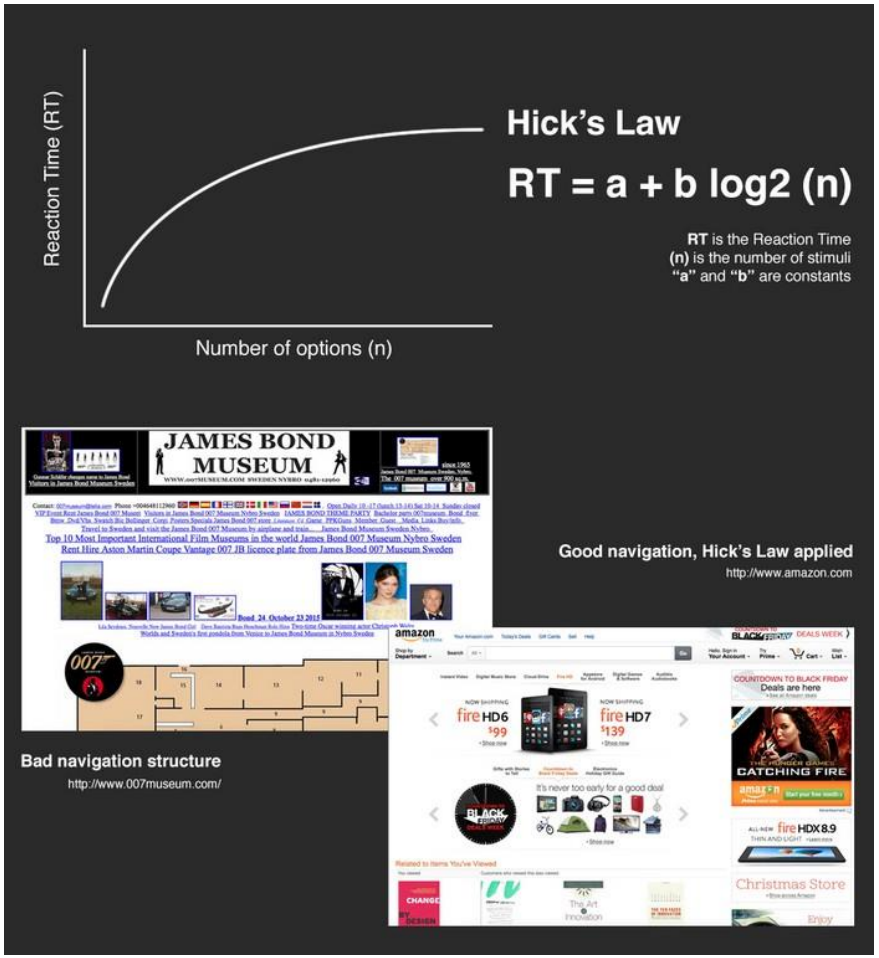


- This input is non-efficient
  - Does not provide any useful information to a system

Hick´s Law

- Hick's Law: Making the choice easier for users
  - *The more alternative options, the more time is required to select*
  - *Laikas = k $log_2(n+1)$, k ~150 msec*

Hick´s Law


Reaction Time (RT)
Number of options (n)
Hick's Law
RT = a + b log2 (n)
RT is the Reaction Time
(n) is the number of stimuli
"a" and "b" are constants
JAMES BOND MUSEUM
Top 10 Most Important International Film Museums in the world James Bond 007 Museum Nybro Sweden
Rent Hire Aston Martin Coupe Vantage 007 JB licence plate from James Bond 007 Museum Sweden
Bad navigation structure
http://www.007museum.com/
Good navigation, Hick's Law applied
http://www.amazon.com
amazon
fire HD6
$99
fire HD7
$139
It's never too early for a good deal
COUNTDOWN TO BLACK FRIDAY
Deals are here
THE HUNGER GAMES
CATCHING FIRE
fire HDX 8.9
Christmas Store

Hick´s Law

- *1 of 8 choices:*

$$Time = 150 * \log_2(8+1) = 480(m\text{sec})$$

- *2 of 4 choices:*

$$Time = 150 * \log_2(2+1) + 150 * \log_2(4+1) = 240 + 350 = 590(m\text{sec})$$

- *Conclusions?*

Summary

- We can use Cognitive Modeling to make predictions about interface usability
- Complementary to Usability Studies
- In practice:
  - GOMS And KLM not used often
  - Fitts' law *often* used for determining best case for new kinds of input methods

Key points

- Inspections can be used to evaluate requirements, mockups, functional prototypes, or systems.
- User testing & heuristic evaluation may reveal different usability problems.
- Walkthroughs are focused so are suitable for evaluating small parts of a product.
- Analytics involves collecting data about users activity on a website or product
- The GOMS and KLM models and Fitts' Law can be used to predict expert, error-free performance for certain kinds of tasks.

References

- Rogers, Sharp, Preece (2011). Interaction design: Beyond Human Computer Interaction. Wiley.
- Fitts P.M. (1954) The information capacity of the human motos system in controlling amplitude of movement, *Journal of Experimentaql psychology* **47,** 381-391.

References

- Nielsen Norman Group Reports. Usability of iPad Apps and Websites: 2 Reports With Research Findings
- K. Väänänen-Vainio-Mattila, Minna Wäljas (2009) Development of Evaluation Heuristics for Web Service User Experience. In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, ACM, pp. 3679-3684