**Empirical Evaluation**

Uncontrolled Studies -> Observational

**Usability**

* The ease with which people who are using the interface can learn to use it to achieve their goals

**User Testing:**

* Can never tell how good/bad something is until we use it
* Other methods like experts know too much about the system/little about the stakeholder’s habits
* Its hard to predict what real people will do with an interface

**Observational user study**

* Field studies: Watch people as they do something in a real situation
  + Realistic activities
  + Hard to set up/expensive
  + Act of observing skews the data
  + Privacy
* Observing in the lab: get people in controlled settings and watch how they use something
  + Can be seen less realistic

**Protocols**

* Protocol are an approach to understanding a person’s intention and actions while they’re using something

**Concurrent protocol**

* The person using an interface says what they’re doing while they perform a task

**Retrospective protocol**

* The person using an interface says what they did after they’ve finished something

**Designing an experiment: Variables**

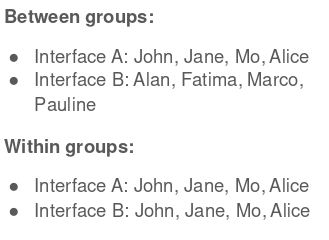
* Dependent variables are things that we can measure
* Independent variables things that we can change
* Confounding variables are things that have to be kept constant so our results are valid

**Designing an experiment: Participants**

* Must be representative of eventual users
* Incentives help

**Designing an experiment: Assigning subjects**

* Usually need a large number for statistically valid results
* Subjects can be assigned:
  + Between-groups: subjects are used only once per condition. We need lots of people but no carry-over effects
  + Within-groups: Subjects are used between conditions



**Designing an experiment: Selecting tasks**

* Tasks should be representative of real tasks
* Avoid using only tasks for testing that are best supported by your interface
* Never choose tasks that are too fragmented (where the purpose isn’t clean)

**Designing an experiment: Collecting data**

* Deciding how to collect data and what data to collect is crucial
* Quantitative:
  + How long a user stayed on a page, how long it took them to complete
* Qualitative:
  + Observations of what the users is doing; users reports

Quantitative data will give us statistical analysis and tell if us the data is significant

Qualitative data can give us analysis of what people say can be done

**P value**

* A statistical value calculated from hypothesis to show significance of the results

**Empirical Testing vs Heuristic**

* Heuristic evaluation is faster
* Heuristic evacuation doesn’t require interpreting a user’s actions
* Empirical testing is far more accurate (since it actually uses statistics)

**GOMS**

* A human processor model that describes the human cognitive structure through four components:
  + Goals
  + Operators
  + Methods
  + Selections
* GOMS is used for predicting the time it takes to use an interface (based on a task)
* Action-level task analysis

**GUI interaction requires**

* Establishing a goal
* Forming a sub-goal
* Specifying a sequence of actions
* Executing the actions
* Perceiving the state of the system
* Interpreting that system state
* Evaluating the system with respect to goals

**GOMS principle**

* Tasks have a constant property: Execution time
* We can make a prediction of how long a task will take by summing the time required for all the sub-tasks in a given GUI

**Keystroke-Level Model**

* These operators are called keystroke-level if they involve pressing keys/buttons or moving a mouse
* Goal -> Develop a simple model to describe time to do a task with a given method on an interactive system



* Task = total time to complete task
* Acquire = time to select method to complete task
* Execute = time to perform method
* Assume expert users and no errors



* Where primitive operators are:
* Keying, pointing, homing, mentally preparing and responding

**Primitive operators:**

Keying

* Time it takes to tap a key on the keyboard or click a mouse

Pointing

* Time it takes to point to something on a display using a mouse

Homing

* Time it takes for a person to move their hand to the keyboard from the mouse

Mentally Preparing

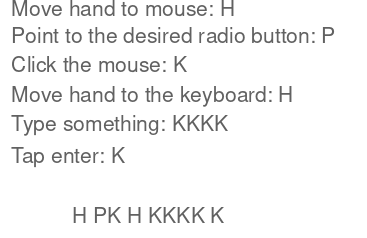
* Time it takes for a person to mentally prepare for the next step

Responding

* The time it takes for a computer to respond to a person

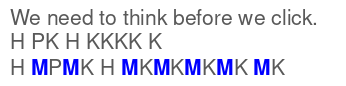
**KLM calculations**

* Step 1:
  + List the actions involved. We begin by calculating the time it takes to perform a task by listing the operations from the GOM list of gestures used
  + Example:



**Rules for placing mental operators**

* Rule 0:
  + Place before K, and any Ps that select commands



* Rule 1:
  + If an operating after M is fully anticipated in the action previous t the M, don’t add M (Pointing + clicking doesn’t need an M)



* Rule 2:
  + If a strike of MKs belong to a cognitive unit, delete Ms expect the first one
  + Like typing words



* Rule 3:
  + Delete Ms before a consecutive terminator
  + Like tasks require holding down Ctrl+C
* Rule 4:
  + Delete Ms that are the terminators of a command
  + Like a user writes a word and then presses enter



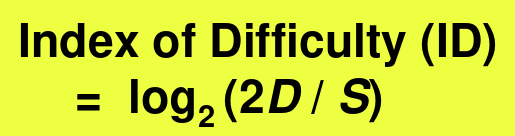
* Delete overlapping Ms

**Hick’s Law and Fitt’s Law**

* Hick’s Law
  + Estimates the time required to make a selection decision
* Fitt’s Law
  + Estimates movement time required to select something on a computer screen

**Fitt’s Law**

* Enables the prediction of human movement based on rapid, aimed movement
* Movement is affected by the distance moved and the precision required
* The precision is measured by the target’s index of difficult

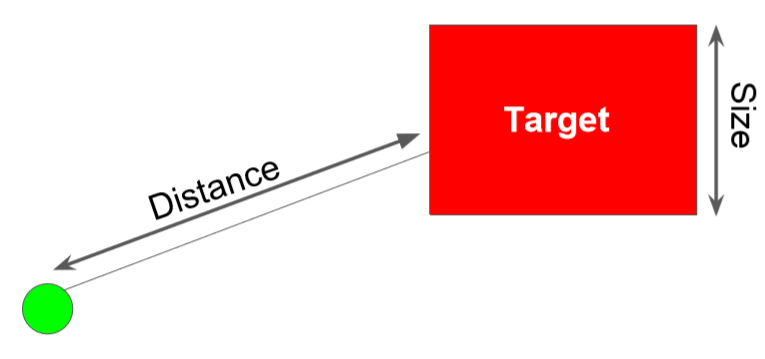


D = The distance to move

Straight line from the cursor to the target

S = The size of the target

Or the tolerance region around it

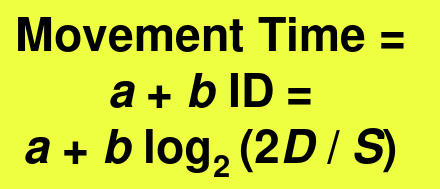


D = The distance to move

* Straight line from the cursor to the target

S = The size of the target

* Or the tolerance region around it

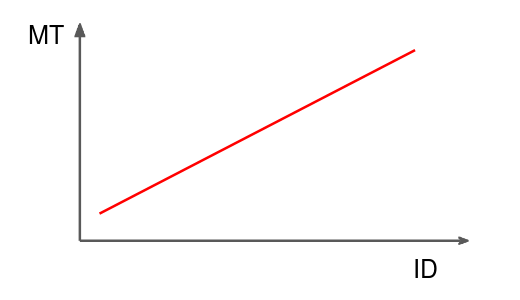


* In HCI, this formula is adapted to



**Shannon-Hartley theorem**

* Maximum rate at which information can be transmitted over a channel of specified bandwidth in the presence of noise
* Noise might include:
  + People talking
  + Bird chirping
  + Your phone vibrating



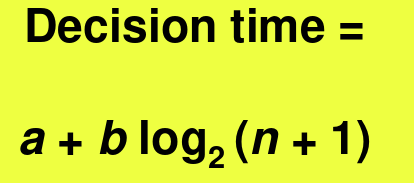
* Movement time correlates with Index difficulty

**Fitt’s Law application**

* Small
* Uninterrupted (made in one continuous motion
* We may use the law to estimate the execution time, and to pick the right size of targets

**Hicks Law**

* The more choices you have, the longer it takes you to come to a decision
* N = the number of possible choices



Where the coefficient a and b are determined experimentally, and are mainly dependent on the number of choices, layouts

**In Summary**

* Fitt’s Law:
  + Estimates the movement time required to select something on a computer screen
* Hick’s Law:

Estimates the time required to make a decision