

Midterm Review

CSI 60: User Interfaces

John Canny

MT this Weds, 306 Soda

- Closed book, no cheat sheets, no electronic devices – sample MTs on the wiki, solutions are up now.
- Format
 - Short answer and longer answer questions
 - Will involve some recall, correct conceptual models...
- Test-taking strategy
 - Questions will not be ordered in difficulty
 - Go through entire test, read questions, answer simple ones first
 - Read questions thoroughly
- Covers material in lectures and readings
 - Lectures mostly go over material in readings
 - Use lectures as guide to what to review in readings

Introduction/ Design Cycle and Brainstorming

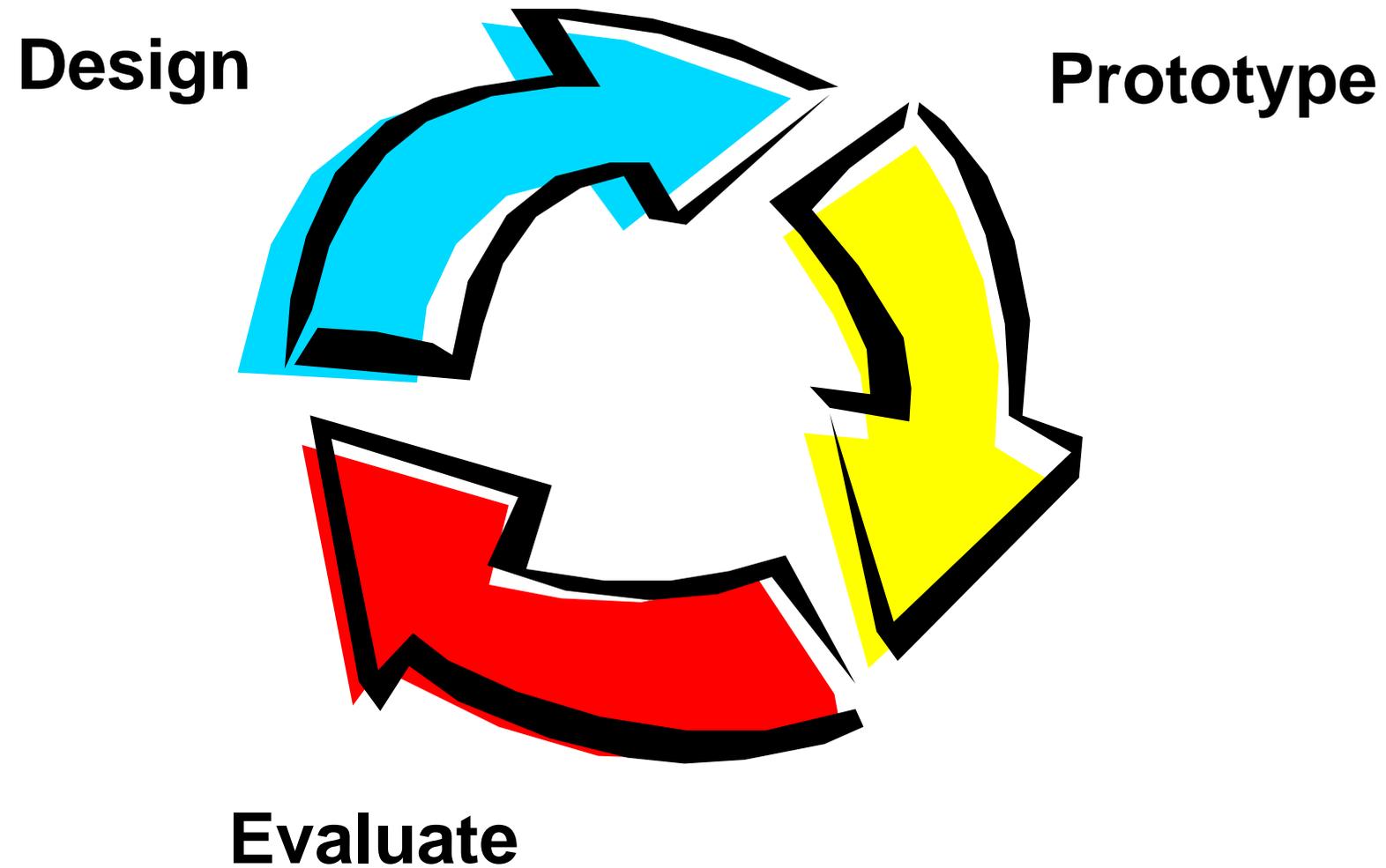
User Interface Design

We're studying the science of UI design, a Human-centered process.



The Design Cycle

(suitable for recall questions)



Lewis and Rieman's cycle

(suitable for recognition questions)

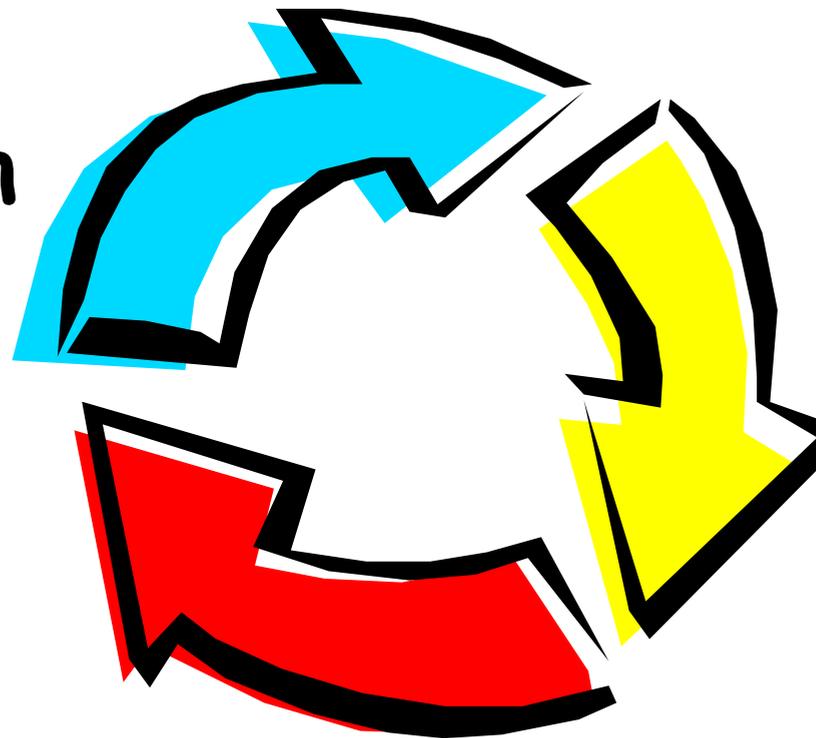
1. Choose Users

2. Select tasks

3. Plagiarize

4. Rough out a design

5. Think about it



6. Prototype

7. Evaluate

8. Iterate

9. Build the design

10. Track the design

11. Change the design

Understand Users

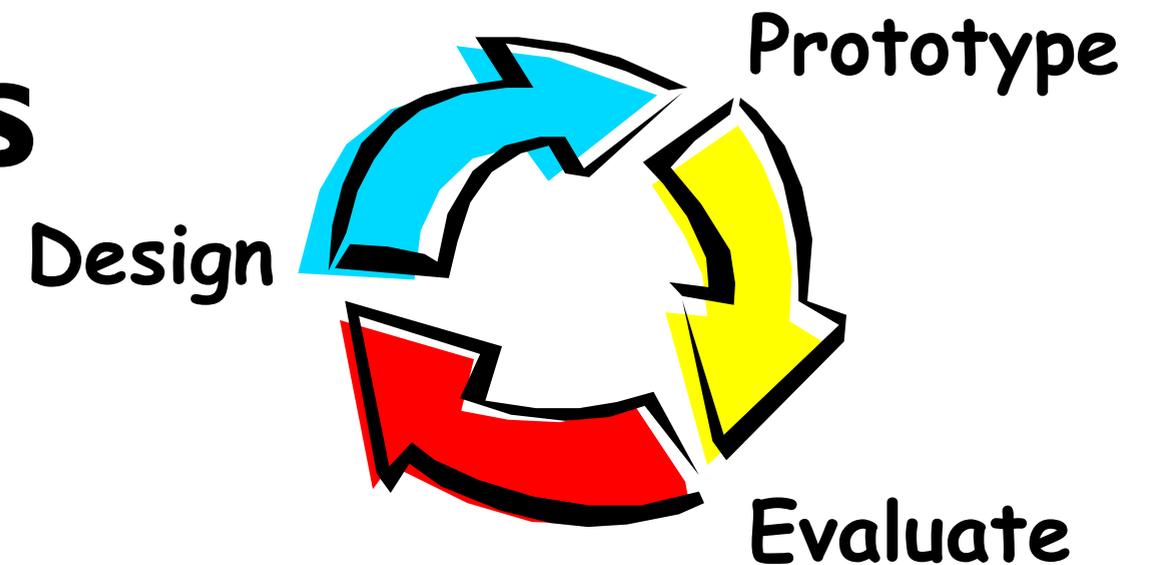
User-centered design starts
and ends with real users.

Observation, surveys, interviews

Two ways to summarize traits:

- Abstraction
- **Archetypes**

Personae



Rough it out

Sketch

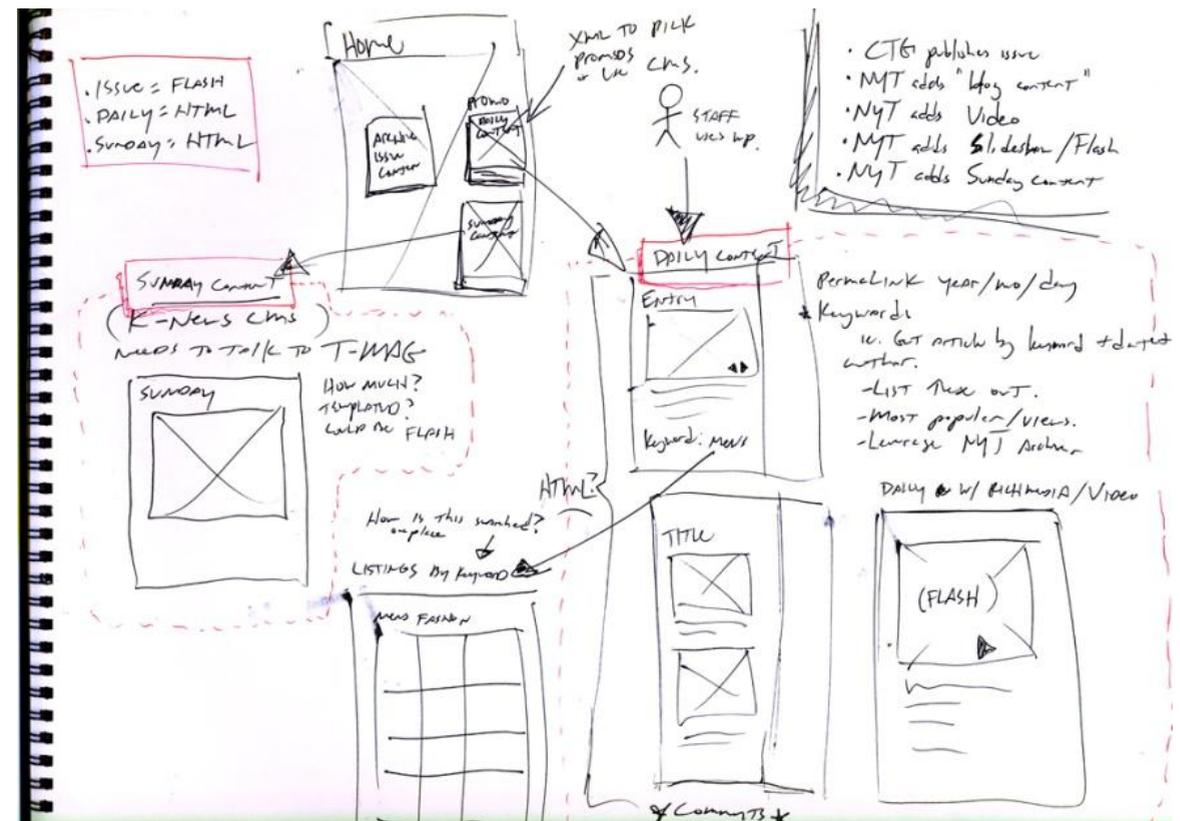
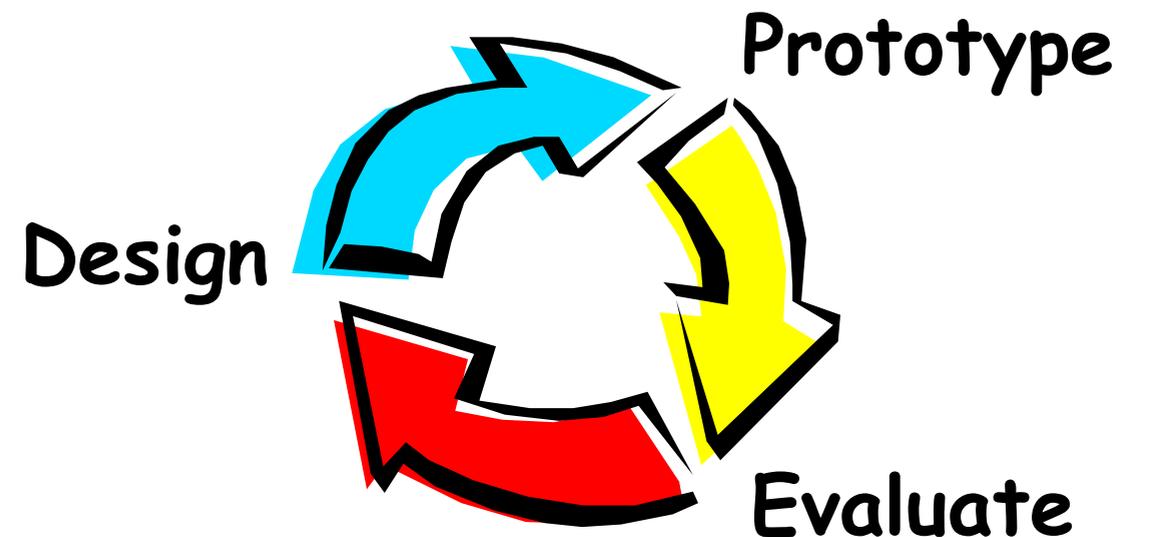
Argue

Get criticism from others

- Seeing through many eyes

Studio model

- The space is a cognitive extension



Task Analysis and Contextual Inquiry

Task analysis: in context



Why context?

Task Analysis Questions

- 1. Who is going to use system?
- 2. What tasks do they now perform?
- 3. What tasks are desired?
- 4. How are the tasks learned?
- 5. Where are the tasks performed?
- 6. What's the relationship between user & data?
- 7. What other tools does the user have?
- 8. How do users communicate with each other?
- 9. How often are the tasks performed?
- 10. What are the time constraints on the tasks?
- 11. What happens when things go wrong?
- **What is the purpose of task analysis?**

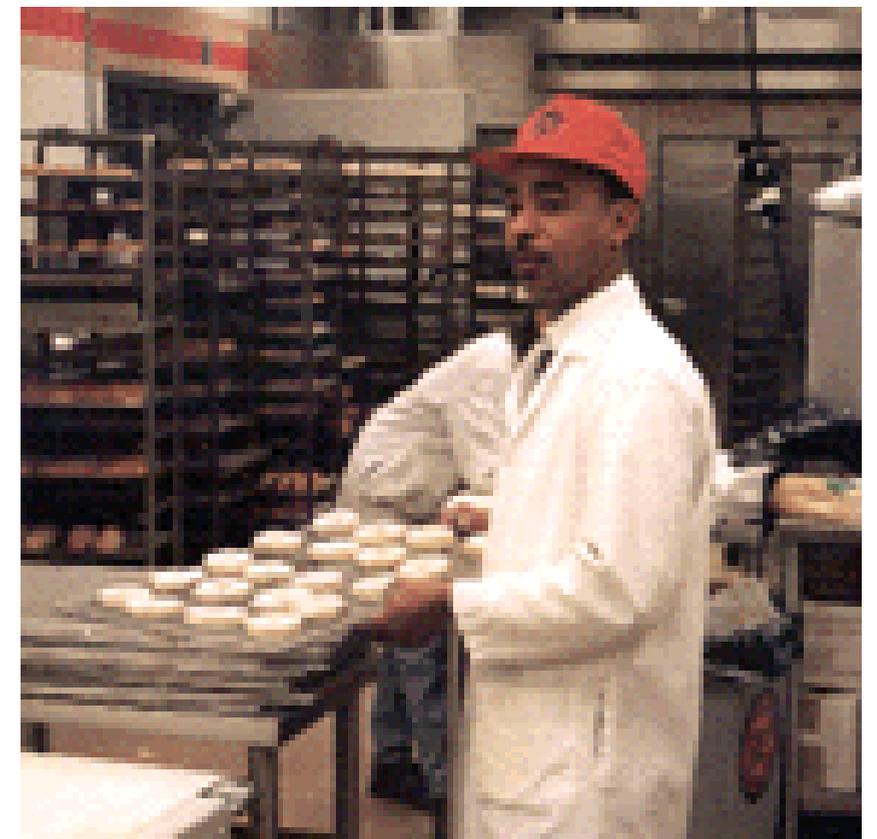
Master-Apprentice Model

Allows user to teach us what they do

- Master (user) works & talks
- We interrupt to ask questions as they go
- Each step reminds master of the next
 - Better than asking user to summarize work habits

What are other models?

How do other models compare?

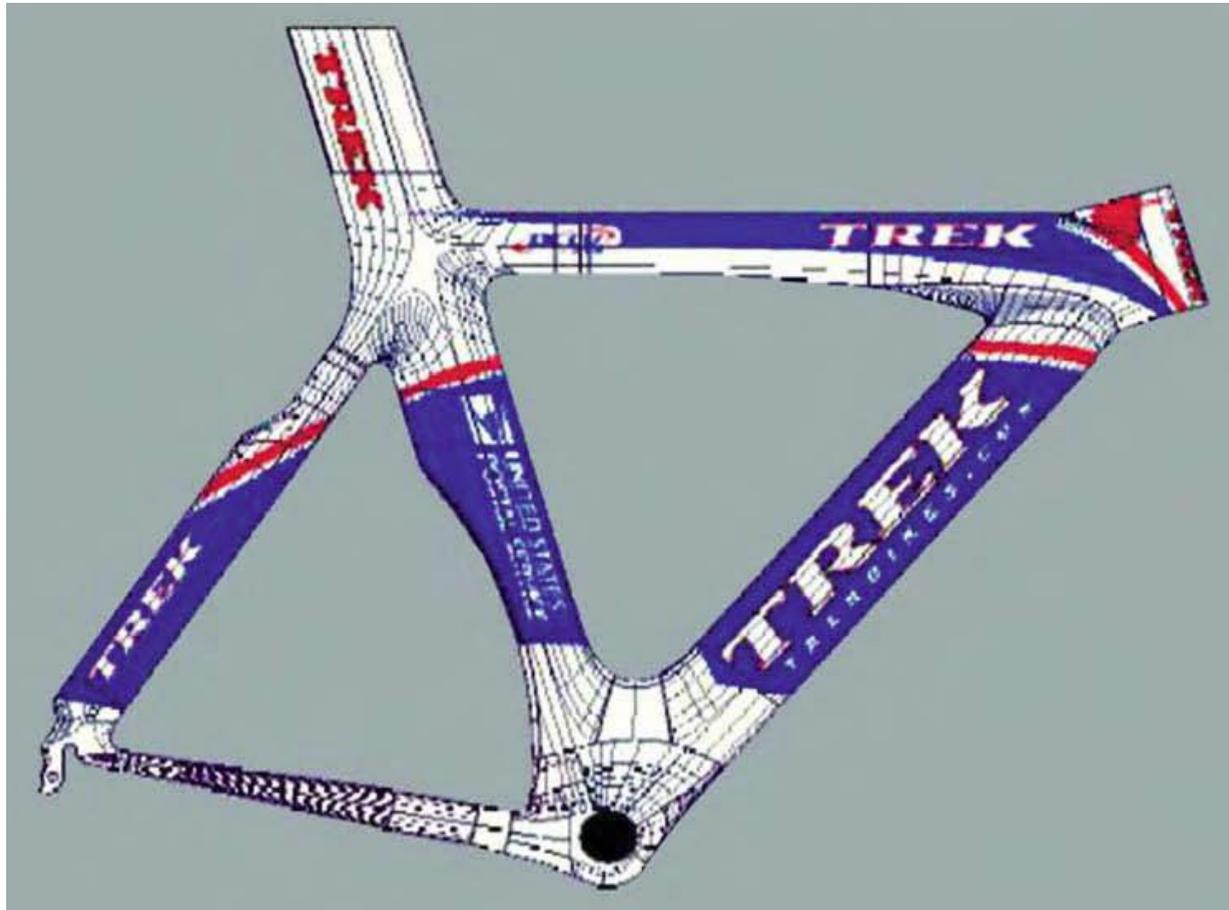


Principles of Contextual Inquiry

- 1. Context**
- 2. Partnership**
- 3. Interpretation**
- 4. Focus**

Sketching and Storyboarding

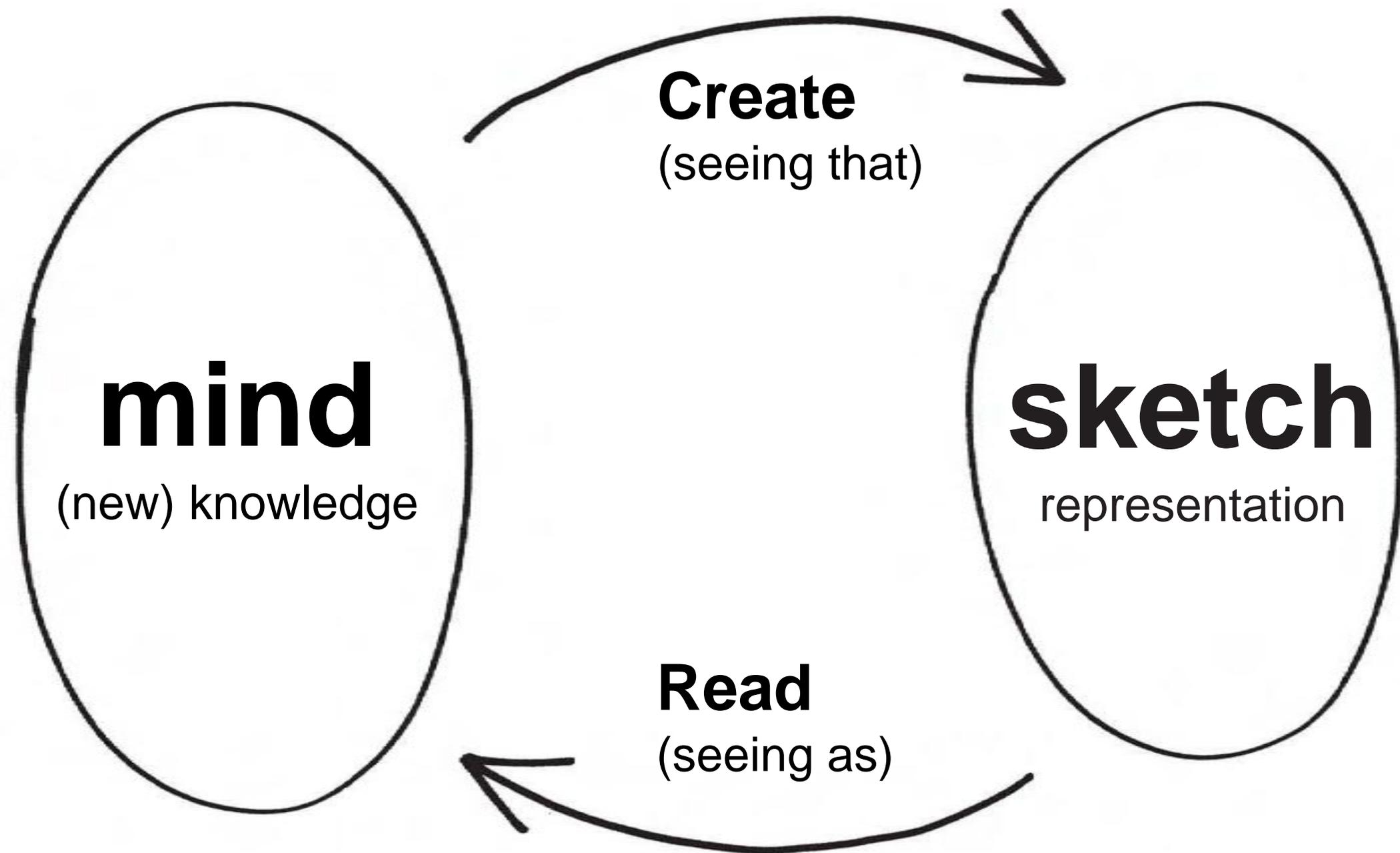
Sketchiness – why?



Sketchiness

- Communicates an unfinished design
- Users and designers are more willing to suggest changes
- Evokes emotion and ideas





The “conversation” between the sketch (right bubble) and the mind (left bubble). A sketch is created from current knowledge (top arrow). Reading, or interpreting the resulting representation (bottom arrow), creates new knowledge. The creation results from what Goldschmidt calls “seeing that” reasoning, and the extraction of new knowledge results from what she calls “seeing as.”

SKETCH

PROTOTYPE

EVOCATIVE → DIDACTIC

SUGGEST → DESCRIBE

EXPLORE → REFINE

QUESTION → ANSWER

PROPOSE → TEST

PROVOKE → RESOLVE

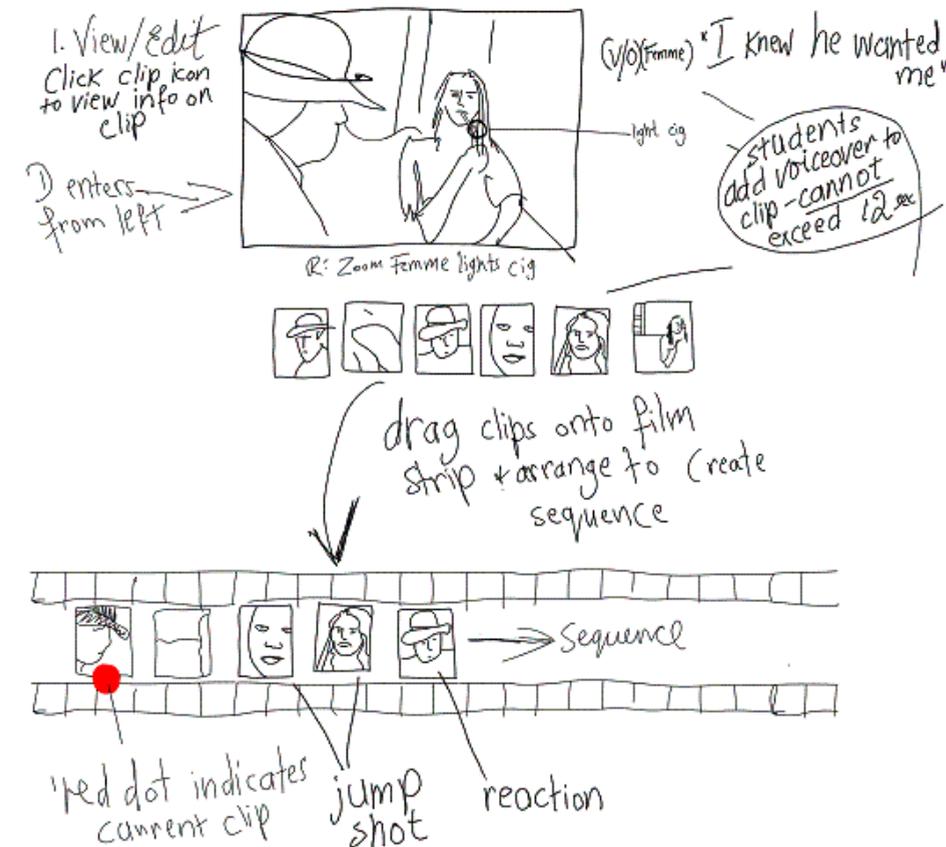
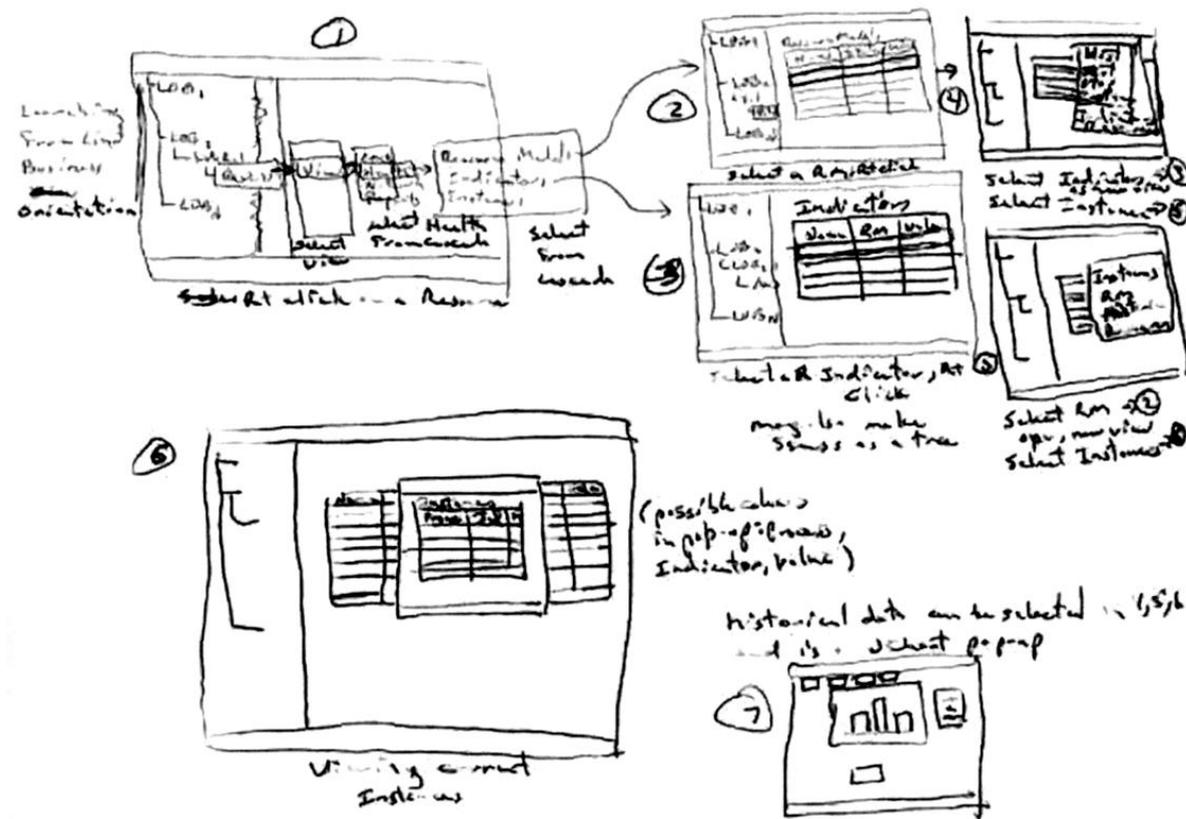
TENTATIVE → SPECIFIC

NONCOMMITTAL → DEPICTION

Storyboarding

Series of key frames depicting key steps in reaching a goal

- Describe the interaction in context
- Often useful to show user in at least 1st frame (establishing shot)



The Psychology of Creativity

Inhibitors:

Facilitators:

The Psychology of Creativity

Inhibitors:

- Peer pressure
- Conformity
- Groupthink

Facilitators:

- Authentic dissent
- Scholarship/Mastery
- Non-conformity



IDEO's Brainstorming Rules

1. Sharpen the Focus
2. Playful Rules
3. Number your Ideas
4. Build and Jump
5. The Space Remembers
6. Stretch Your Mental Muscles
7. Get Physical



Aim for quantity

Hope for quality



Build and Jump

Premature idea rejection is a serious barrier to good design.

One of the biggest differentiators between good designers and great ones is the latter's ability to successfully develop unusual ideas.

This requires a strong instinct to be able to distinguish fatal vs. minor flaws in an idea.



Conceptual Models

Affordances

What is an affordance?

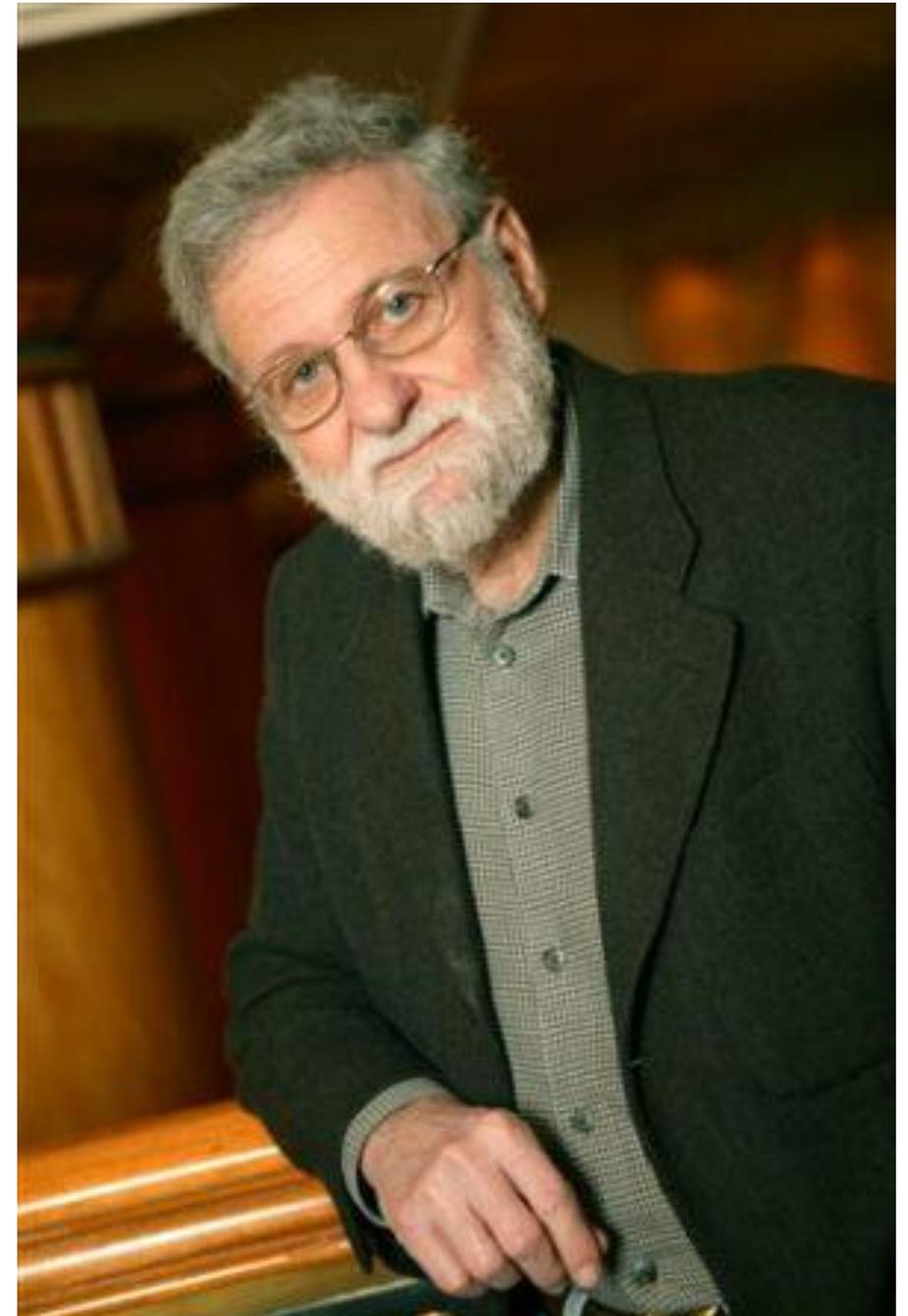
“... the term **affordance** refers to the *perceived* and *actual* properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used.

Some affordances obvious

- Knobs afford turning
- Buttons afford pushing
- Glass can be seen through

Some affordances learned

- Glass breaks easily
- Floppy disk
 - Rectangular – can't insert sideways
 - Tabs prevent backwards insertion



Norman's Design Principles

- Make controls visible
- Make sure mapping is clear
- Provide feedback

Metaphor

Definition

The transference of the relation between one set of objects to another set for the purpose of brief explanation

Examples?

When are they effective?

When are they not effective?

Direct Manipulation

Direct Manipulation

- An interface that behaves as though the interaction was with a real-world object rather than with an abstract system

Central ideas

- Visibility of the objects of interest
- Rapid, reversible, incremental actions
- Manipulation by pointing and moving
- Immediate and continuous feedback

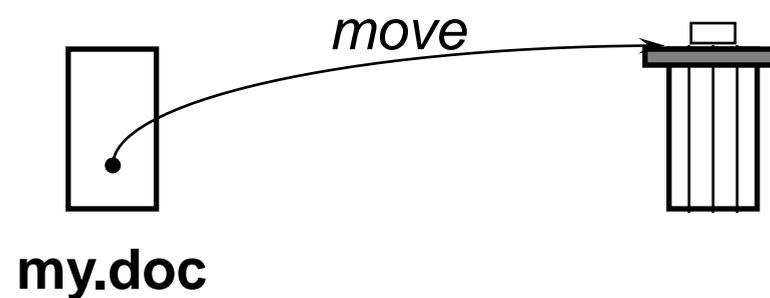
Noun-Verb VS Verb-Noun

Noun-Verb: Select object, *then* perform action

Verb-Noun: Select action, then perform it on object

What are some examples of these two approaches?

What are the pros and cons of these two approaches?



Usability Inspection Methods

Usability Heuristics

“Rules of thumb” describing features of usable systems

- Can be used as design principles
- Can be used to evaluate a design

Example: *Minimize users' memory load*

Pros and cons

- Easy and inexpensive
 - Performed by experts
 - No users required
 - Catch many design flaws
- More difficult than it seems
 - Not a simple checklist
 - Cannot assess how well the interface will address user goals

Phases of Heuristic Eval. (1-2)

1) Pre-evaluation training

- Provide the evaluator with domain knowledge if needed

2) Evaluation

- Individuals evaluate interface then aggregate results
 - Compare interface elements with heuristics
- Work in 2 passes
 - First pass: get a feel for flow and scope
 - Second pass: focus on specific elements
- Each evaluator produces list of problems
 - Explain why with reference to heuristic or other information
 - Be specific and list each problem separately

Phases of Heuristic Eval. (3-4)

3) Severity rating

- Establishes a ranking between problems
 - Cosmetic, minor, major and catastrophic
- First rate individually, then as a group

4) Debriefing

- Discuss outcome with design team
- Suggest potential solutions
- Assess how hard things are to fix

Prototyping

Fidelity in Prototyping

Fidelity refers to the level of detail

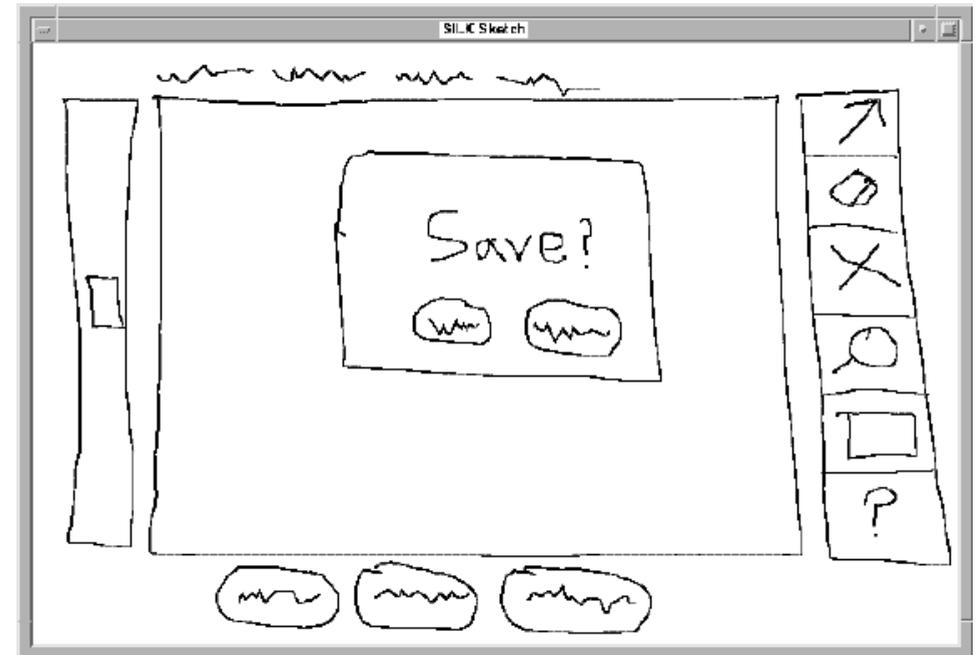
High fidelity?

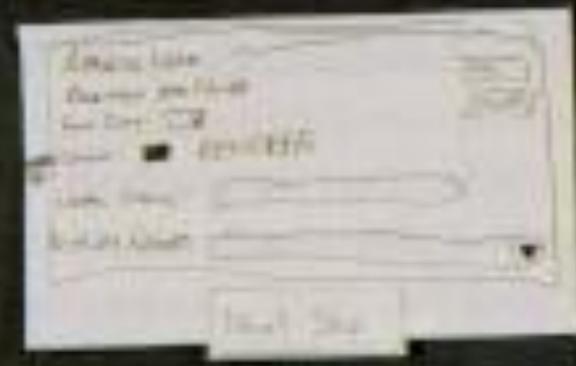
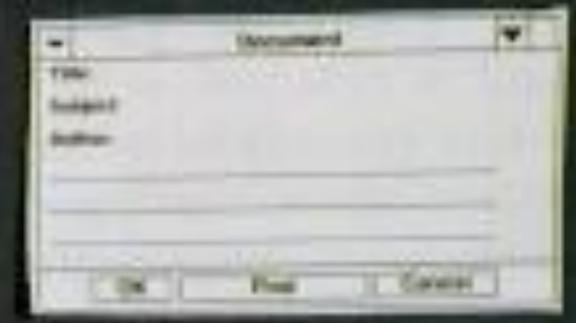
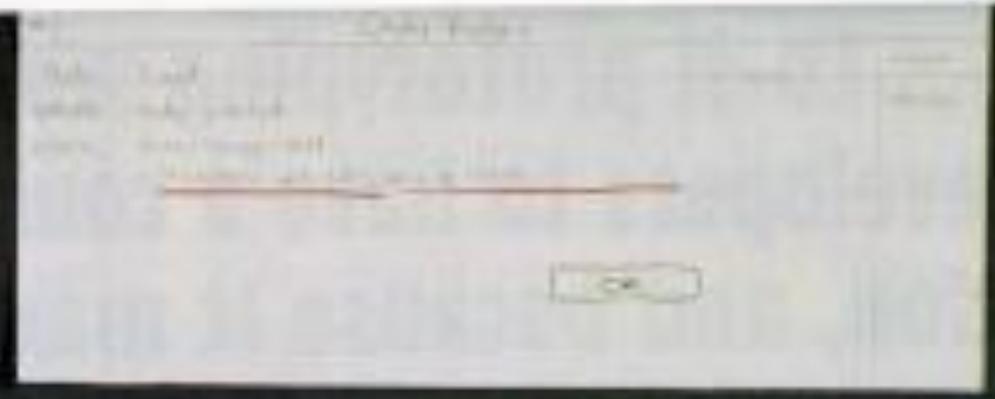
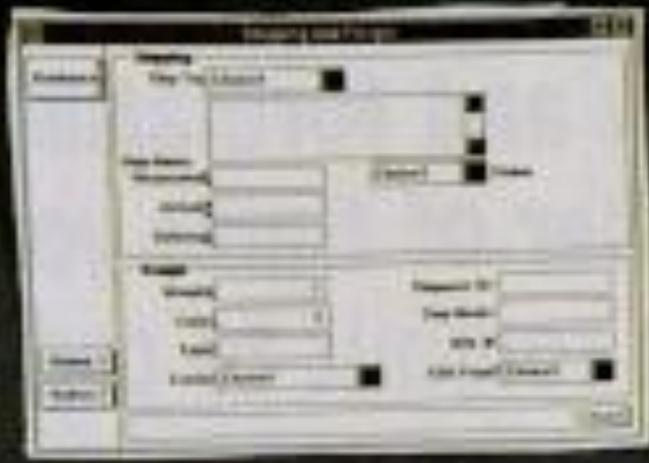
- Prototypes look like the final product

Low fidelity?

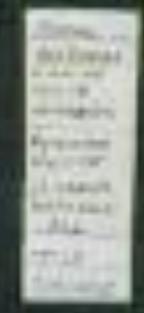
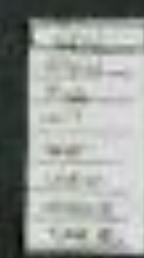
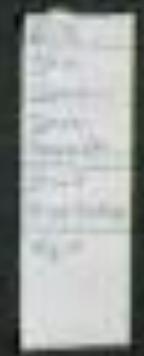
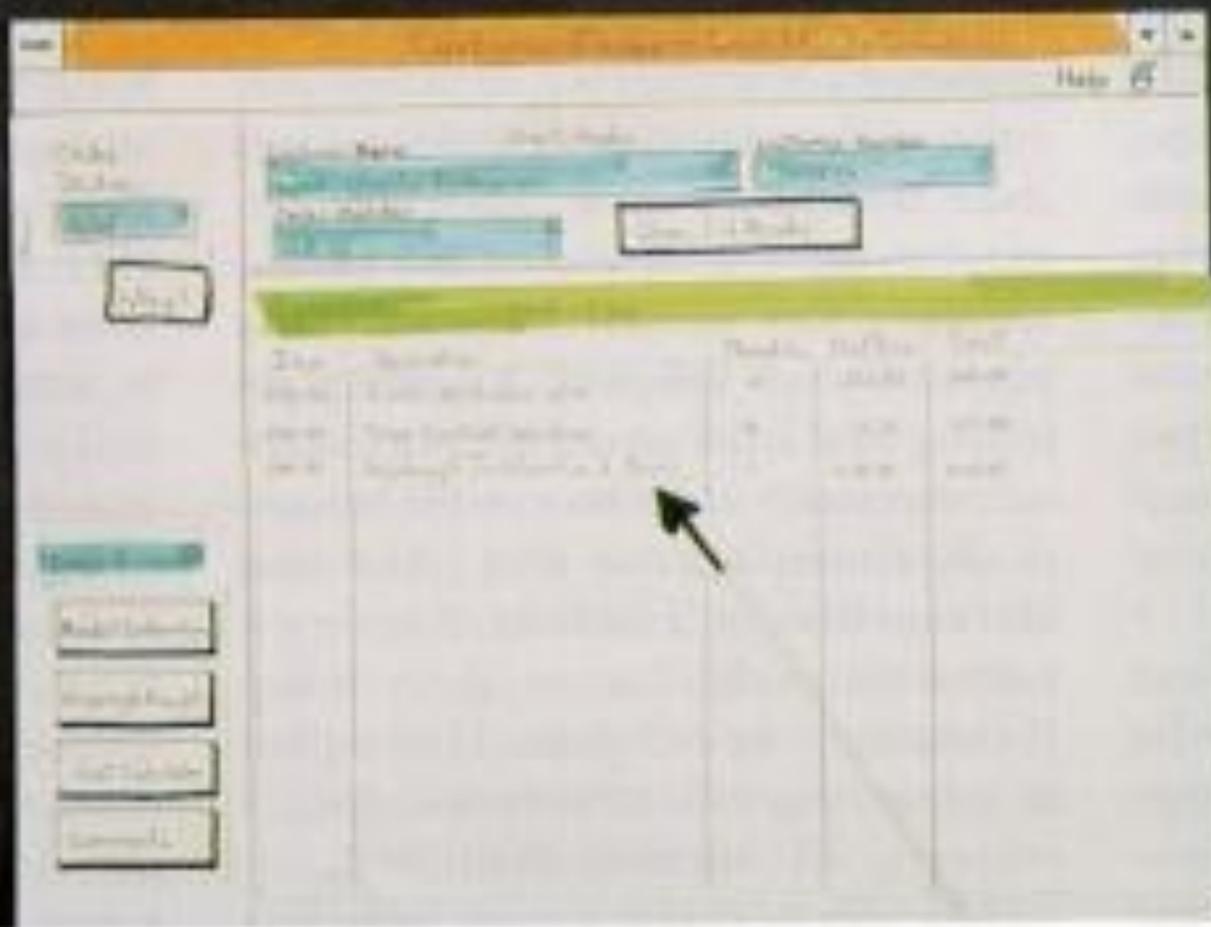
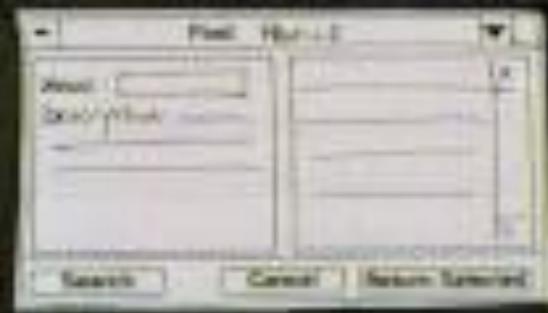
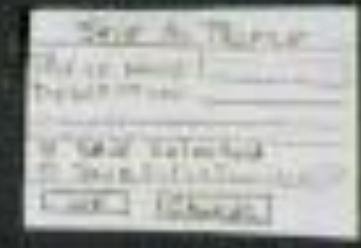
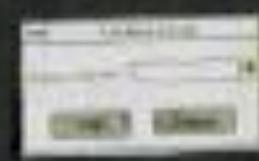
- Artists renditions with many details missing

- Why do we prototype?
- What are the disadvantages of Low- and High-fidelity prototypes?



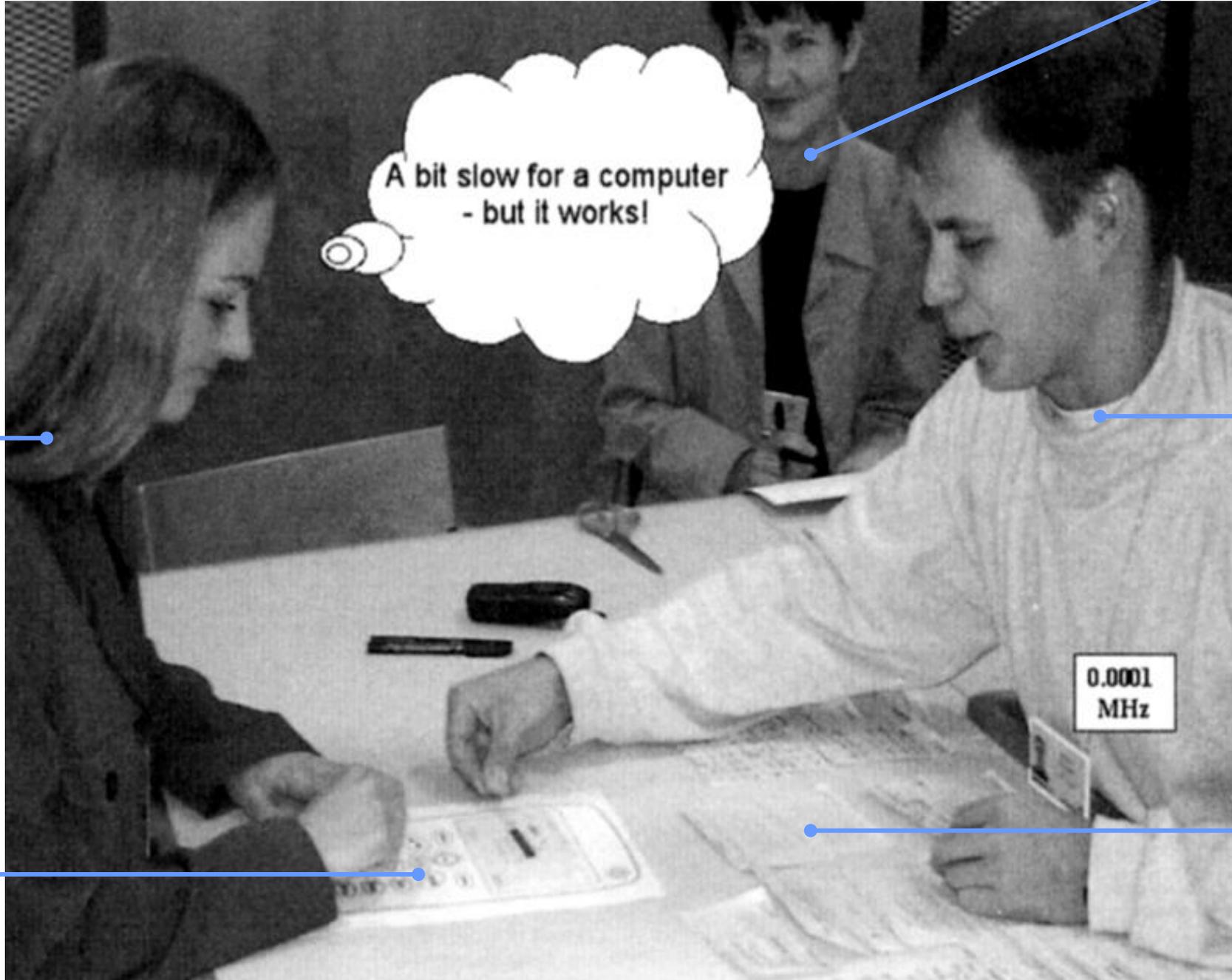


OK



Time	Location	Speed	Altitude
10:00	1000m	1000	1000
10:05	1000m	1000	1000
10:10	1000m	1000	1000
10:15	1000m	1000	1000
10:20	1000m	1000	1000
10:25	1000m	1000	1000
10:30	1000m	1000	1000
10:35	1000m	1000	1000
10:40	1000m	1000	1000
10:45	1000m	1000	1000
10:50	1000m	1000	1000
10:55	1000m	1000	1000
11:00	1000m	1000	1000
11:05	1000m	1000	1000
11:10	1000m	1000	1000
11:15	1000m	1000	1000
11:20	1000m	1000	1000
11:25	1000m	1000	1000
11:30	1000m	1000	1000
11:35	1000m	1000	1000
11:40	1000m	1000	1000
11:45	1000m	1000	1000
11:50	1000m	1000	1000
11:55	1000m	1000	1000
12:00	1000m	1000	1000

Time	Location	Speed	Altitude
10:00	1000m	1000	1000
10:05	1000m	1000	1000
10:10	1000m	1000	1000
10:15	1000m	1000	1000
10:20	1000m	1000	1000
10:25	1000m	1000	1000
10:30	1000m	1000	1000
10:35	1000m	1000	1000
10:40	1000m	1000	1000
10:45	1000m	1000	1000
10:50	1000m	1000	1000
10:55	1000m	1000	1000
11:00	1000m	1000	1000
11:05	1000m	1000	1000
11:10	1000m	1000	1000
11:15	1000m	1000	1000
11:20	1000m	1000	1000
11:25	1000m	1000	1000
11:30	1000m	1000	1000
11:35	1000m	1000	1000
11:40	1000m	1000	1000
11:45	1000m	1000	1000
11:50	1000m	1000	1000
11:55	1000m	1000	1000
12:00	1000m	1000	1000



Observer
(or video camera)

User

“Computer”

Interface

Interface elements

Conducting a Test

Three or more testers

- **Greeter** - Puts users at ease & gets data
- **Facilitator** - only team member who speaks
 - Gives instructions & encourages thoughts, opinions
- **Computer** - knows application logic & controls it
 - Always simulates the response, w/o explanation
- **Observer(s)** - Take notes & recommendations

(Greeter can serve other roles)

Typical session should be approximately 1 hour

- Preparation, the test, debriefing

Qualitative Methods

Rapid Ethnography

- A true ethnographic field study will often be beyond a companies means. Rapid ethnography may be the answer. It involves
- **Narrow the focus:** to those activities most relevant to the application to be designed. Use key informants to guide you.
- **Use multiple interactive observation techniques** to increase the likelihood of discovering *exceptional* and useful behavior.
- **Use collaborative and computerized iterative data analysis methods.**

Quantitative Methods

Steps in Designing an Experiment

1. State a lucid, testable hypothesis
2. Identify variables
(independent, dependent, control, random)
3. Design the experimental protocol
4. Choose user population
5. Apply for human subjects protocol review
6. Run pilot studies
7. Run the experiment
8. Perform statistical analysis
9. Draw conclusions

Experiment Design

- Testable hypothesis
 - Precise statement of expected outcome
- Independent variables (factors)
 - Attributes we manipulate/vary in each condition
 - Levels – values for independent variables
- Dependent variables (response variables)
 - Outcome of experiment (measurements)
 - Usually measure user performance

Experiment Design

- Control variables
 - Attributes that will be fixed throughout experiment
 - Confound – attribute that varied and was not accounted for
 - Problem: Confound rather than IV could have caused change in DVs
 - Confounds make it difficult/impossible to draw conclusions
- Random variables
 - Attributes that are randomly sampled
 - Increases generalizability

Common Metrics in HCI

- Performance metrics:
 - Task success (binary or multi-level)
 - Task completion time
 - Errors (slips, mistakes) per task
 - Efficiency (cognitive & physical effort)
 - Learnability
- Satisfaction metrics:
 - Self-report on ease of use, frustration, etc.

Between vs. Within Subjects

- Between subjects
 - Each participant uses one condition
 - +/- Participants cannot compare conditions
 - + Can collect more data for a given condition
 - - Need more participants
- Within subjects
 - All participants try all conditions
 - + Compare one person across conditions to isolate effects of individual diffs
 - + Requires fewer participants
 - - Fatigue effects
 - - Bias due to ordering/learning effects

Null Hypothesis

e.g. for the hypothesis interfaceA faster than interfaceB, the null hypothesis would be that *the times are the same*.

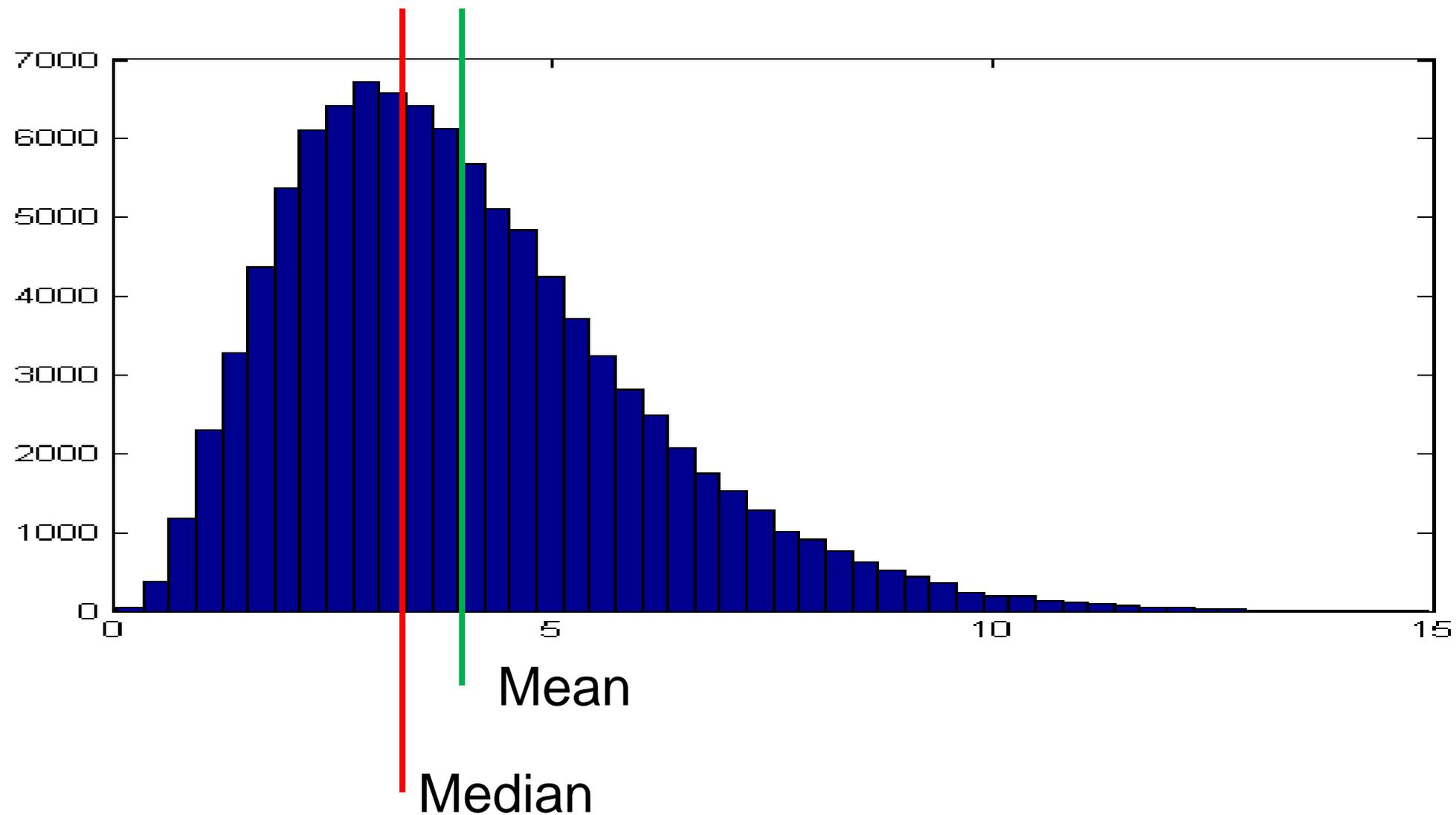
Note: refuting the null hypothesis typically does not prove the hypothesis.

Anything else, however unlikely, that could cause the measurement difference could be the real explanation. Standard tests don't consider any of these situations.

Mean and Median

The median keeps equal numbers of elements (equal curve areas) on either side. It is not influenced by magnitude.

The mean is sensitive to values, the larger the values, the larger the mean. So it will move toward the “tail” of the distribution.

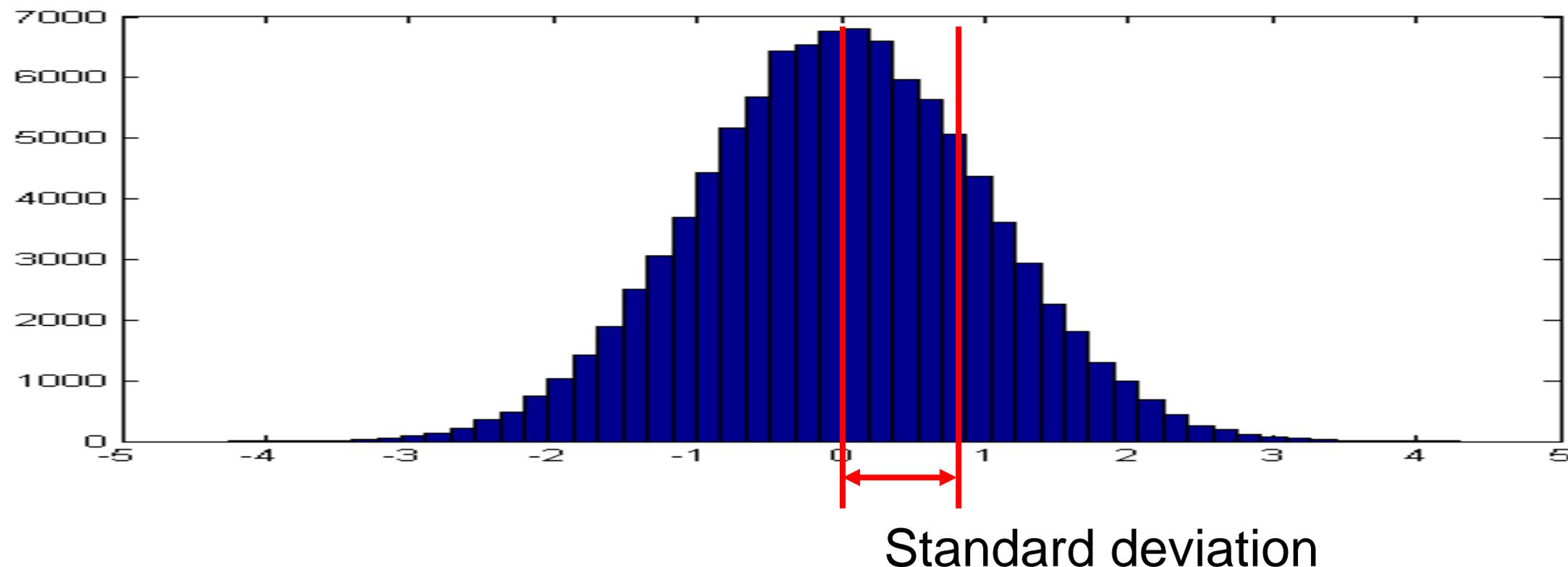


Variance

Is a measure of the width of a distribution. Specifically, it is the average squared deviation of samples from their mean:

$$\text{Var}(X) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

The related quantity called **standard deviation** is the square root of variance and can be used to measure the width of the distribution:



One-sample t-test

The t-statistic is defined as:

Where \bar{x} is the sample mean, s is the sample standard deviation, and n is the number of samples.

The distribution of this statistic depends on the number of *degrees of freedom*, which is $n-1$.

Significance

Hypothesis testing is a probabilistic process.

It will never tell you “X is true” or “X is false.”

So researchers have come to declare that certain probabilities represent “statistically significant” effects.



.05

Significance: is an a-priori determined probability σ , such as 0.05 or 0.01, such that when $\text{Pr}(\text{Observation} \mid \text{Null Hypothesis}) < \sigma$, the result can be declared to be “statistically significant”.

Core Concepts

- Variables – independent, dependent, control, random
- Data distributions: skew, mean, median, variance
- Hypothesis – Initial and then a null hypothesis
- Test statistic to measure “how unusual” the data are
- Significance – probability of type I errors
- P-values – probabilities derived from the statistic

Qualitative Methods

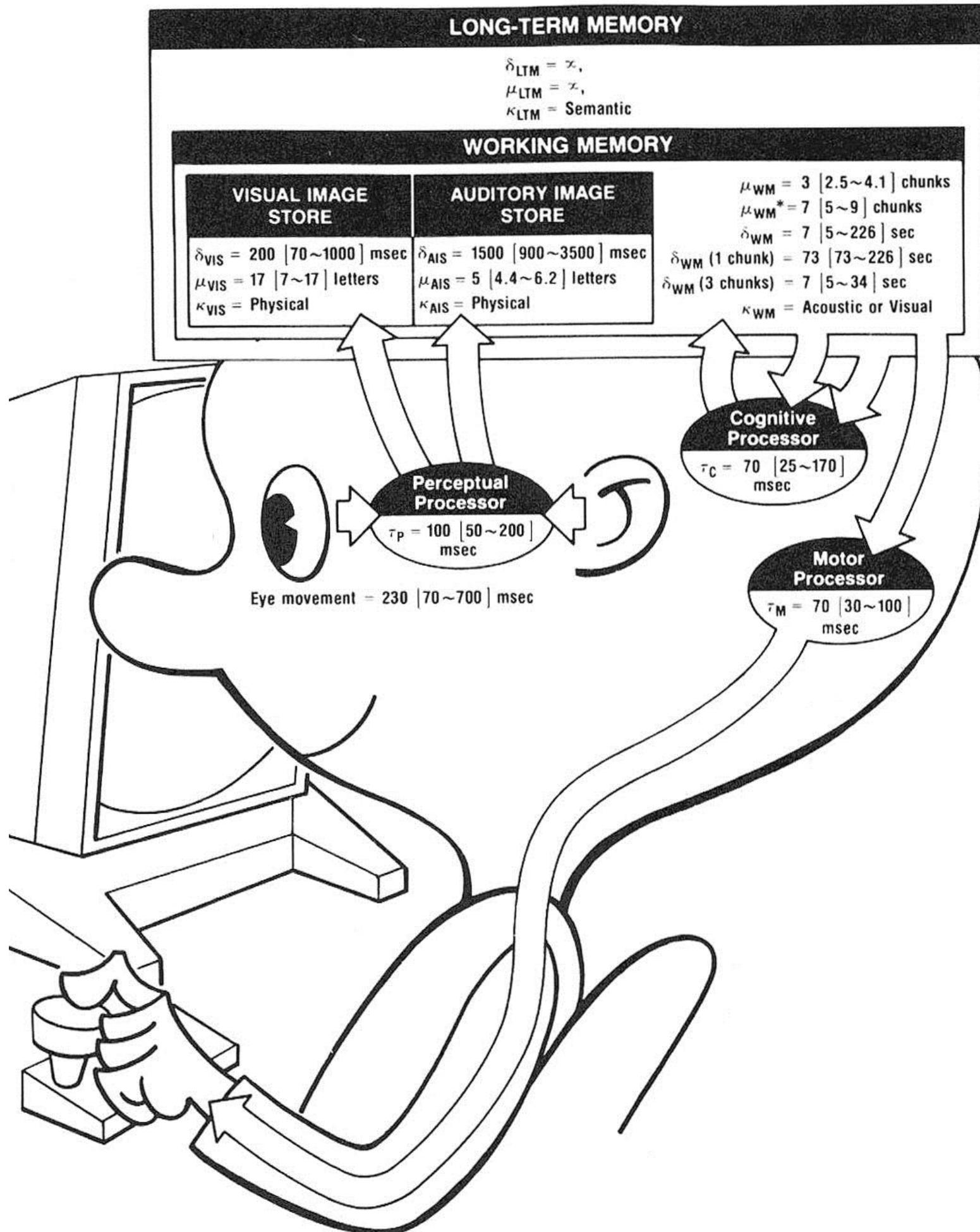
Forms of Ethnography

- **Contextual Inquiry:** is a specific type of ethnographic method which follows the master/apprentice model. Two others are:
- **Observational study:** is an approach which attempts to minimize the impact of the observation on users' behavior (c.f. Star Trek's prime directive).
- **Participant Observation:** The ethnographer participates in subjects' normal practices in order to better understand them. (what actually happens in Star Trek episodes).
- In practice "ethnography" is an open, long-term process which contrasts with a task-focused application of contextual inquiry.

Rapid Ethnography

- A true ethnographic field study will often be beyond a companies means. Rapid ethnography may be the answer. It involves
- **Narrow the focus:** to those activities most relevant to the application to be designed. Use key informants to guide you.
- **Use multiple interactive observation techniques** to increase the likelihood of discovering *exceptional* and useful behavior.
- **Use collaborative and computerized iterative data analysis methods.**

Human Information Processing



Human Info. Processor

- Perceptual
- Cognitive
- Motor (will discuss later)
- Working memory
- Long-term memory

Unified model

- Probably inaccurate
- Predicts perf. well
- Very influential

Review: Memory

Working memory is small

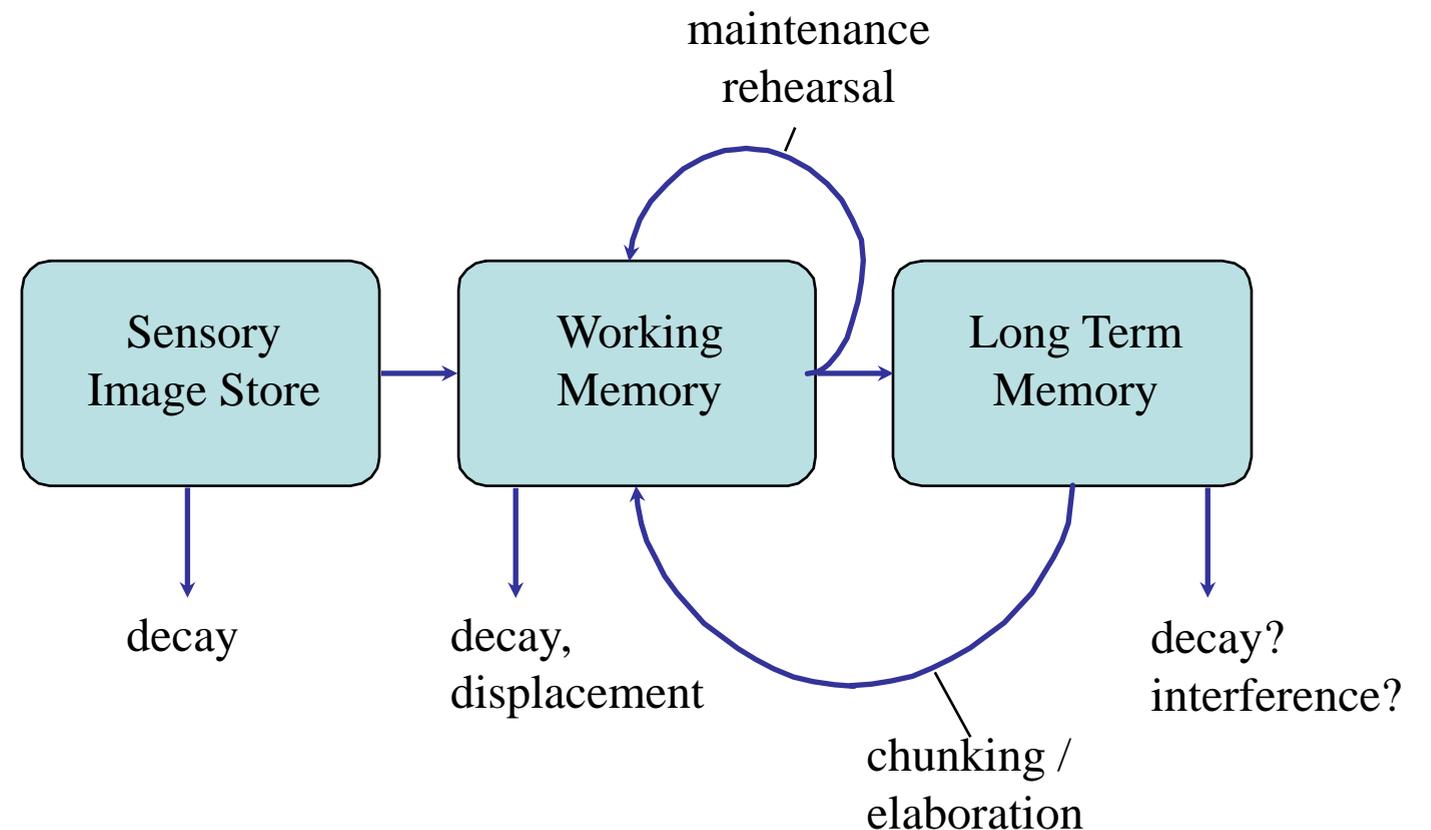
- Temporary storage
 - decay
 - displacement

Long term memory

- Rehearsal
- Relate new to learned material
- Link to existing knowledge
- Attach meaning
 - Make a story

Design interface to facilitate retrieval

- Recognition rather than recall



Recognition over Recall

Recall

- Information reproduced from memory

Recognition

- Presentation of info helps retrieve info (helps remember it was seen before)
- Easier because of cues to retrieval

Power Law of Practice

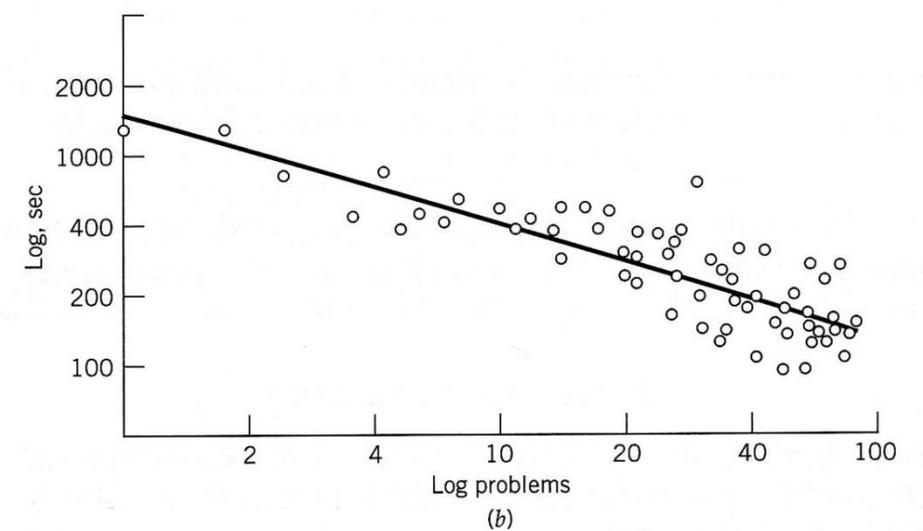
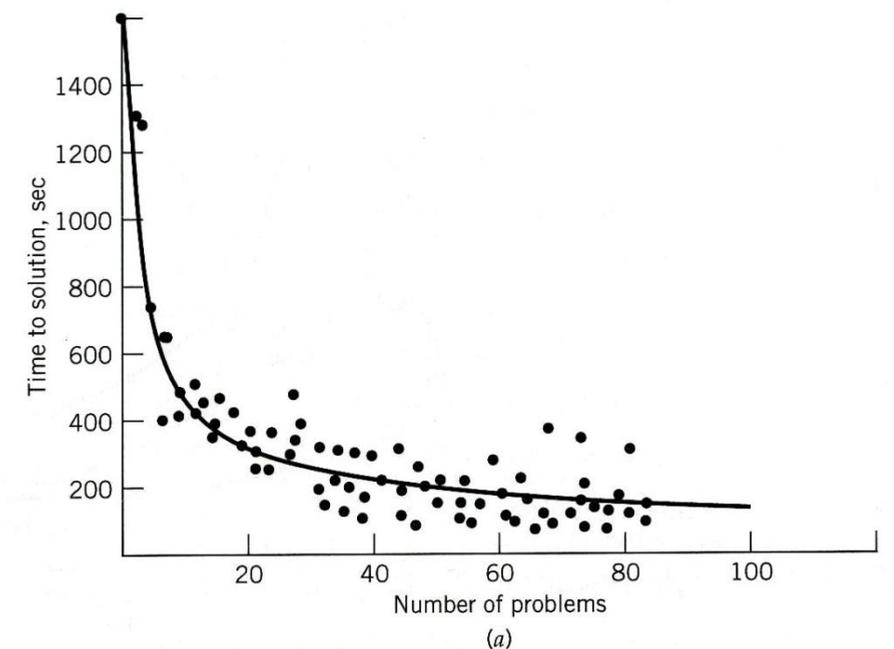
Task time on the n th trial follows a power law

$$T_n = T_1 n^{-a} + c$$

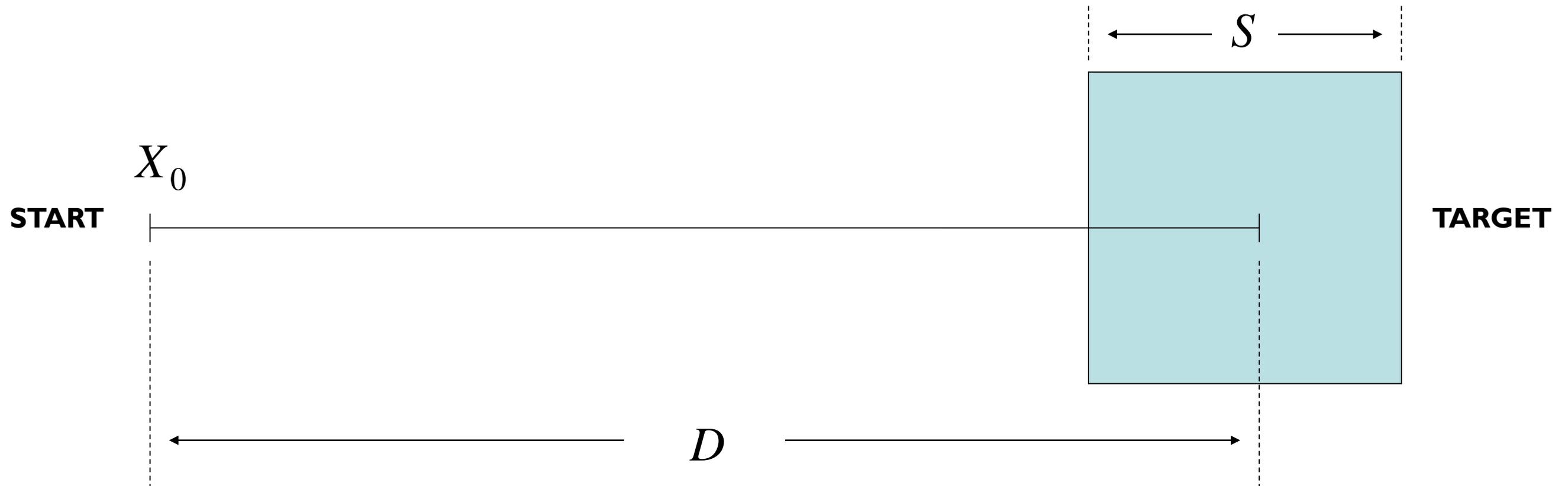
where $a = .4$, $c =$ limiting constant

Applies to skilled behavior

- Sensory
- Motor



Fitts' Law



Hand movement based on series of microcorrections

X_i = remaining distance after i th move

relative movement accuracy remains constant \rightarrow

$$\frac{X_i}{X_{i-1}} = \varepsilon$$

Then

$$T = I_m \log_2 \left(\frac{2D}{S} \right)$$

Fitts' Law

$$T = a + b \log_2(D/S + 1)$$

a, b = constants (empirically derived)

D = distance

S = size

ID is Index of Difficulty = $\log_2(D/S+1)$

- Models well-rehearsed selection task
- T increases as the **distance** to the target increases
- T decreases as the **size** of the target increases

Interactive Programming

Event Dispatch Loop



Mouse moved (t_0, x, y)

Event Queue

- Queue of input events

Event Loop (runs in dedicated thread)

- Remove next event from queue
- Determine event type
- Find proper component(s)
- Invoke callbacks on components
- Repeat, or wait until event arrives

Component

- Invoked callback method
- Update application state
- Request repaint, if needed



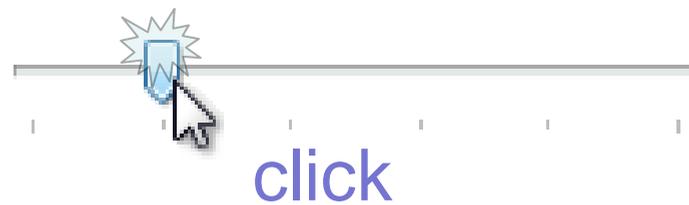
Callbacks

Window Manager/Browser

Your Code



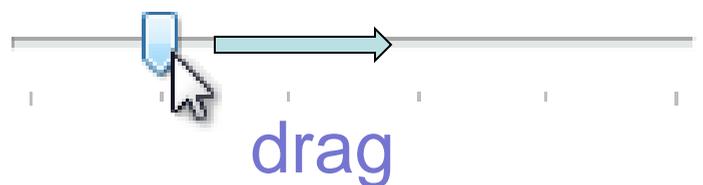
`onmouseover(Event e){...}`



`onmouseup(Event e){...}`

`onmousedown(Event e){...}`

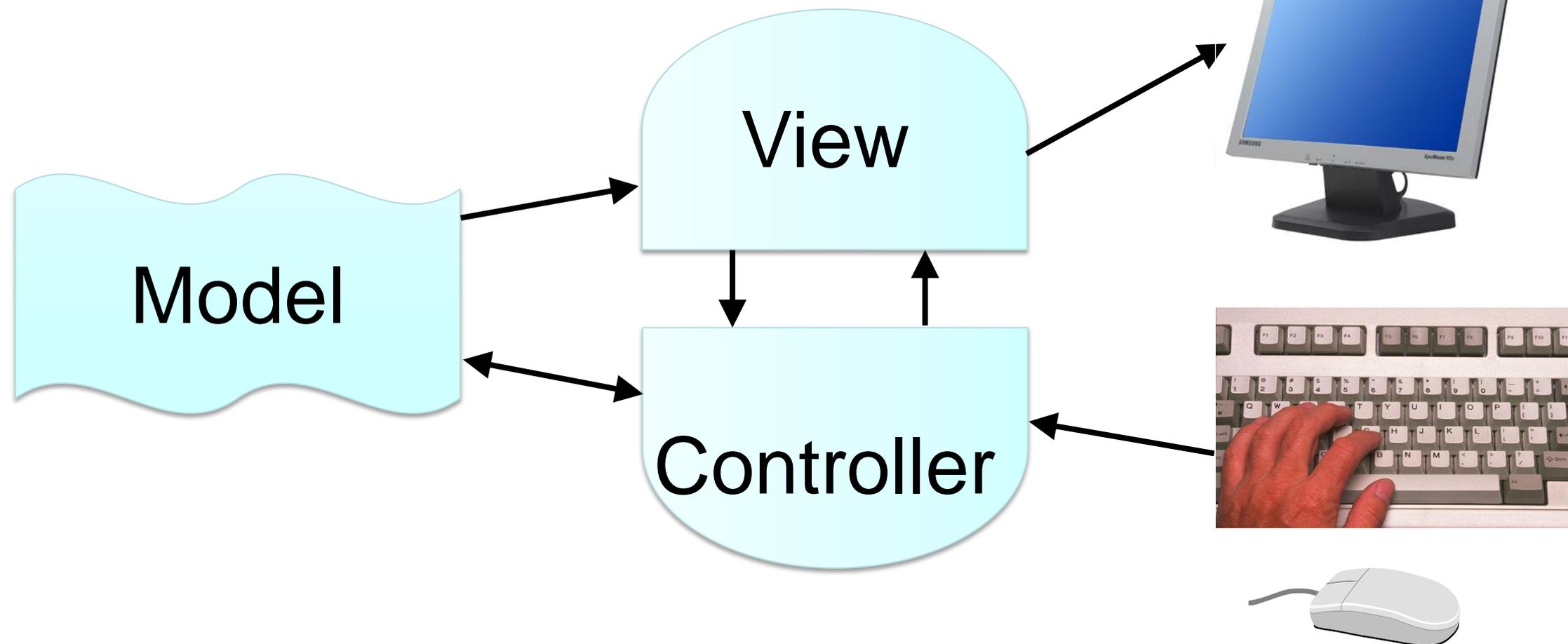
`onclick(Event e){...}`



`ondrag(Event e){...} // Maybe!`

Model-View-Controller

- OO Architecture for interactive applications
 - introduced by Smalltalk developers at PARC circa 1983



Why MVC?

- Combining MVC into one class will not scale
 - Like HTML < 4, massive code changes to improve web page appearance even if content fixed.
- Separation eases maintenance and extensibility
 - Separates design work: Content creators don't need to be designers, designers don't need to be content experts.
 - can change a view later, e.g., draw shapes in 3D
 - flexibility of changing input handling when using separate controllers

Concurrency Design Patterns

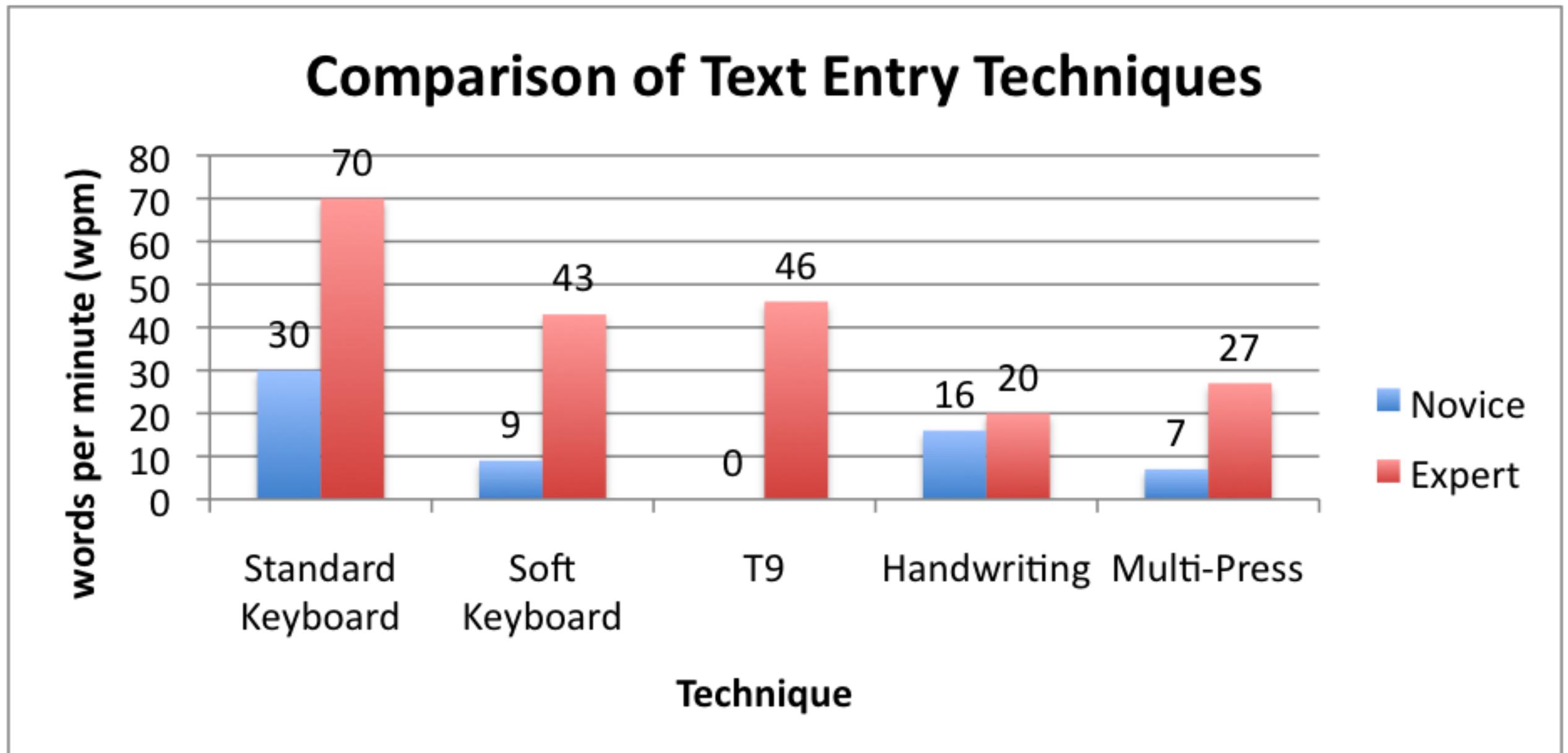
- Message queue
- GUI Thread / Worker thread pool
- Database / Model-View-Controller
- Actor

Input Devices

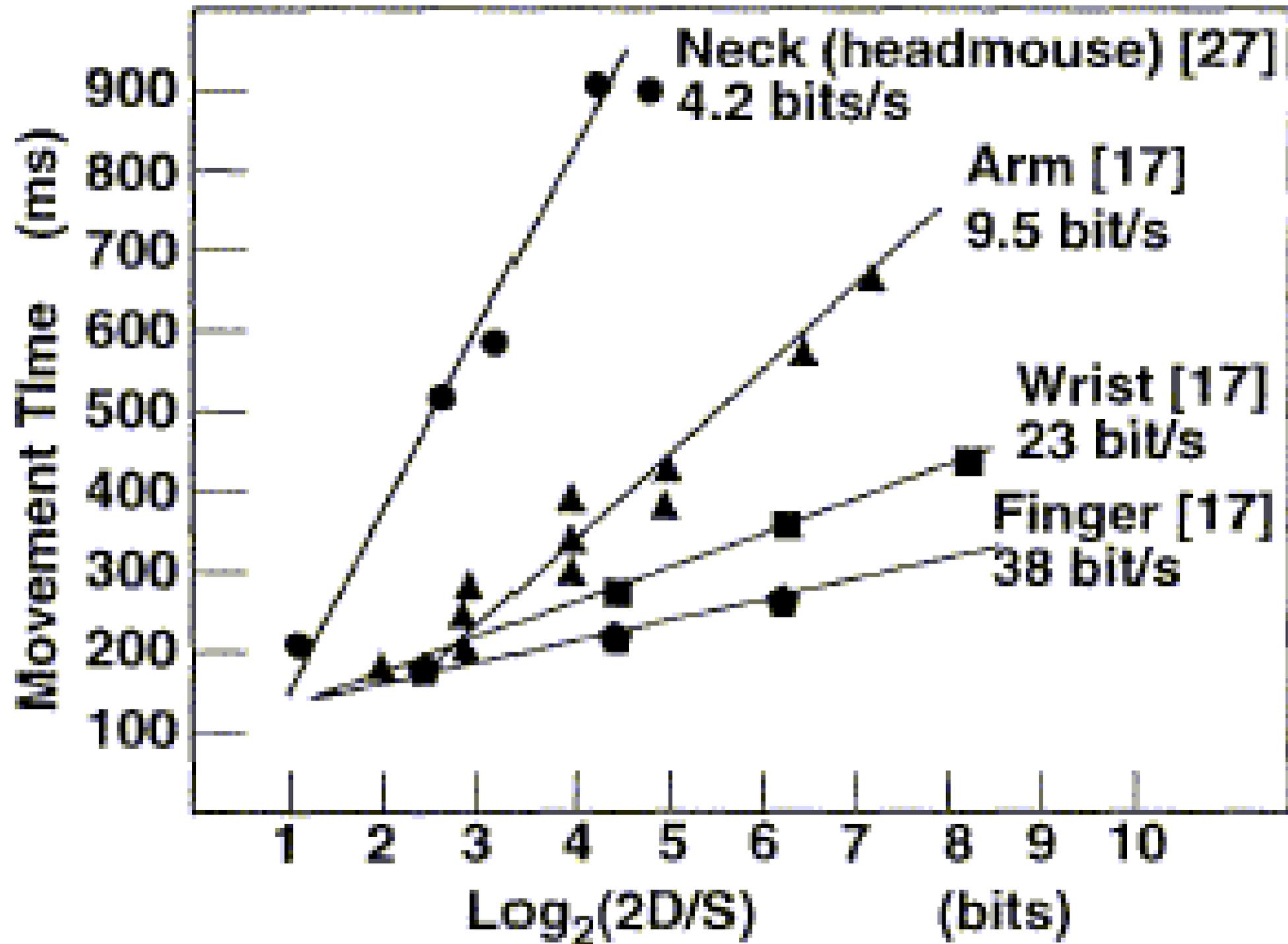
Input Devices: Important Tasks

- Text Entry
- Pointing/Marking
 - Target acquisition
 - Steering / positioning
 - Freehand drawing
 - Drawing lines
 - Tracing and digitizing
 - ...
- Voice: like text entry but limited
 - better to be “dialogic”

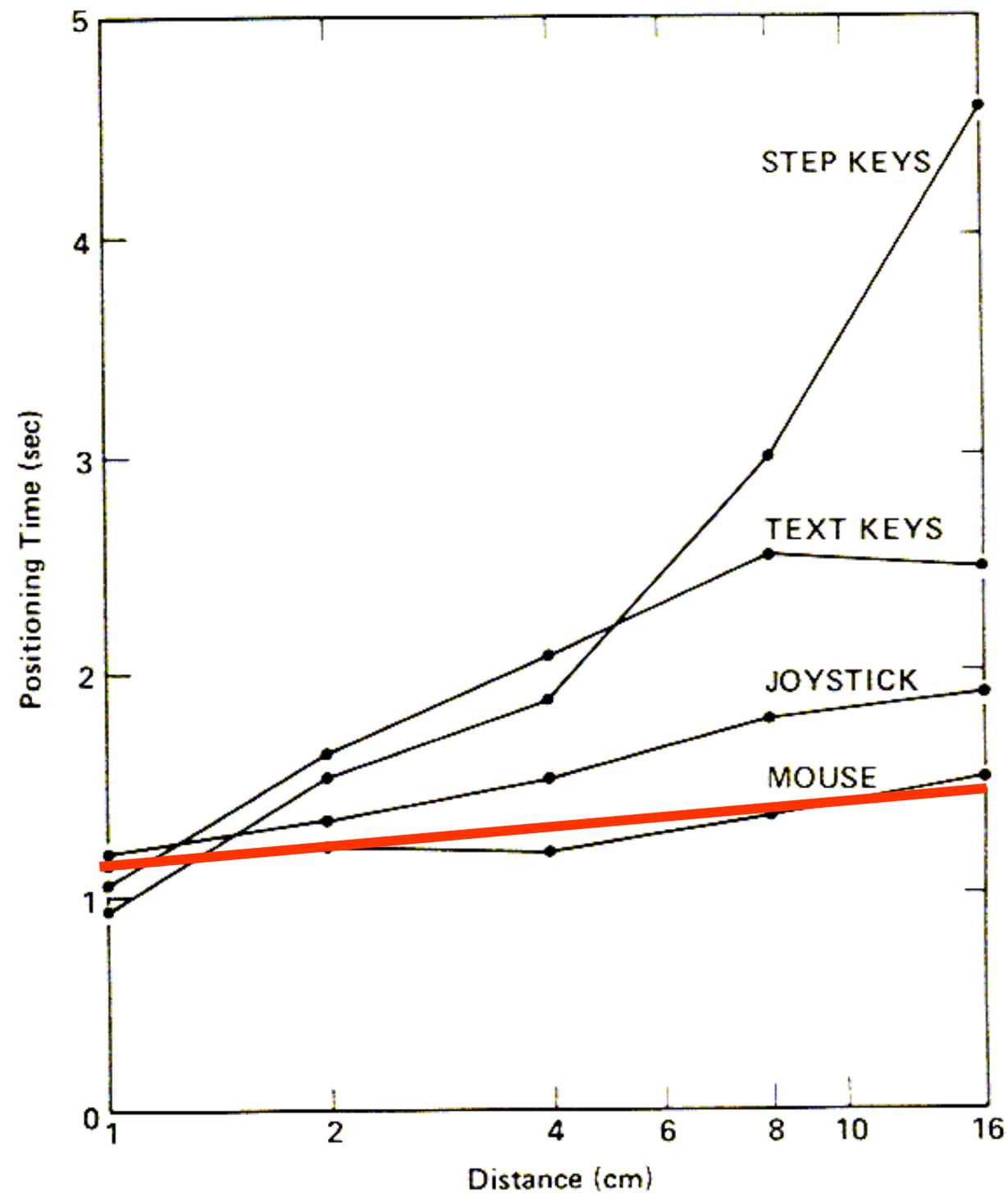
Which is fastest?



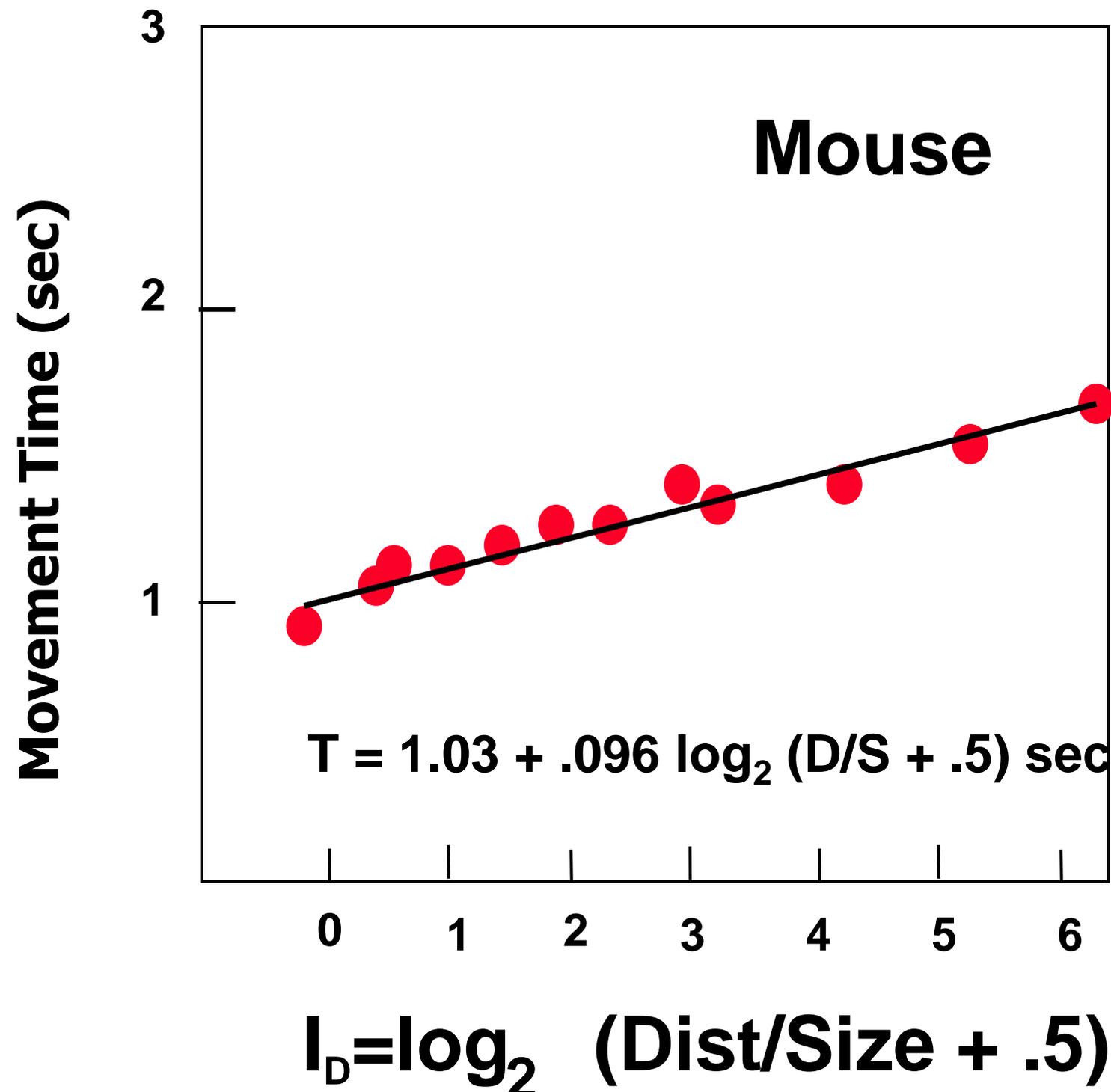
Bandwidth of Human Muscle Groups



Experiment: Mice are fastest!



Why is the mouse fastest?



Why these results?

Time to position mouse proportional to Fitts' Index of Difficulty I_D .

[i.e. how well can the muscles direct the input device]

Therefore speed limit is in the eye-hand system, not the mouse.

Therefore, mouse is a near optimal device.

What about Speech Recognition?

- Dictation is faster than typing (~100 wpm), BUT:
 - Speech is different from written language:
Speaking in well-formed, complete, print-ready sentences is cognitively challenging
 - High cost of correcting errors through speech channel alone
 - Social awkwardness?

Beyond text input...

- Even with keyboards, there is more information gathered than the text.
- Biometrics: you can partially identify a person from the timing of their keystrokes. e.g. timing of password typing has shown accuracies as high as 99%
- Speech: can detect emotion fairly well (70%) and stress very well (90%).
- Can also measure stress from mouse movement...

Next Time

Midterm Exam – good luck!