Looking Back: Presenting User Study Results

- Keep in mind that there are various types of data
- Need to summarize the (vast amount of) collected data
  - Graphs, e.g. histogram
  - Characteristics
    » minimum, maximum, outliers
    » mean, mode, median
    » Standard deviation, variance
    » Quantile (quartile, percentile)
  - Boxplots

- **Statistical significance**
  - Low probability (e.g. < 5%) that a difference occurred by chance
  - T-test, ANOVA

- **User study report**
Looking Back: Fitts’ Law

- Predicts movement time for rapid, aimed pointing tasks
- One of the few stable observations in HCI

\[ MT = a + b \log_2 \left( 1 + \frac{D}{W} \right) \]  
Index of Difficulty: \[ \log_2 \left( 1 + \frac{D}{W} \right) \]

- How to get a and b for a specific device / interaction technique
  - vary D and W and measure MT; fit a line by linear regression

- Various implications for HCI
  - Consider button sizes
  - Use edges and corners
  - Use current location of the cursor
  - Use average location of the cursor(?)
  - **Possibility to compare different input devices**
Looking Back: Steering Law

• Models the movement time of a pointer through a 2D tunnel
• Extension of Fitts’ Law

• Tunnels with constant width: \( MT = a + b \frac{D}{W} \)  
  Index of Difficulty: D / W

• Extension for arbitrary tunnel shapes: \( MT = a + b \int_{c}^{d} \frac{ds}{W(s)} \)

• Implications for HCI
  – Nested menus
  – Navigation tasks
  – Extensions for virtual reality / 3D movements possible
3 Basic HCI Principles and Models

3.1 Predictive Models for Interaction: Fitts’ / Steering Law

3.2 Descriptive Models for Interaction: GOMS / KLM

3.3 Users and Developers

3.4 3 Usability Principles by Dix et al.

3.5 3 Usability Principles by Shneiderman

3.6 Background: The Psychology of Everyday Action
To Recap: *Predictive Models*

- **Model:**
  - Simplification of a complex situation / action, e.g. human interaction

- **Predictive:**
  - Make educated guesses about the future
    » relying on knowledge about past actions / states
    » relying on a model of interaction

- **Examples:**
  - Fitts’ Law (directed aimed movement)
  - Law of Steering (navigation through a tunnel)
  - Hick’s Law / Hick-Hyman Law (choose an item within a menu)
  - ...
Descriptive Models

• *(The categorisation is not sharp, for more insights, see [MacKenzie 2003])*

• Descriptive models
  – provide a basis for understanding, reflecting, and reasoning about certain facts and interactions
  – provide a conceptual framework that simplifies a, potentially real, system
  – are used to inspect an idea or a system and make statements about their probable characteristics
  – used to reflect on a certain subject
  – can reveal flaws in the design and style of interaction

• Examples:
  – Descriptions, statistics, performance measurements
  – Taxonomies, user categories, interaction categories

In *HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science (Book)*, 27-54
Example: Three-State Model (W. Buxton)

- Describes graphical input
- Simple, quick, expressive

- Possible extensions:
  - multi-button interaction
  - stylus input
  - direct vs. indirect input

Dragging tasks: (a) mouse (b) lift-and-tap touchpad. [MacKenzie 2003]

Buxton, W, 1990, A Three-State Model of Graphical Input
*In INTERACT’90, 449-456*
Example: Guiard’s Model of Bimanual Skill (1 / 2)

- Many tasks are asymmetric with regard to left / right hand
- Guiard’s model identifies the roles and actions of the non-preferred and preferred hands

**Non-preferred hand**
- leads the preferred hand
- sets the spatial frame of reference for the preferred hand
- performs coarse movements

**Preferred hand**
- follows the non-preferred hand
- works within established frame of reference set by the non-preferred hand
- performs fine movements
### Example: Guiard’s Model of Bimanual Skill (2 / 2)

<table>
<thead>
<tr>
<th>Task</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scrolling</td>
<td>• precedes/overlaps other tasks</td>
</tr>
<tr>
<td></td>
<td>• sets the frame of reference</td>
</tr>
<tr>
<td></td>
<td>• minimal precision needed (coarse)</td>
</tr>
<tr>
<td>Selecting, editing, reading, drawing, etc.</td>
<td>• follows/overlaps scrolling</td>
</tr>
<tr>
<td></td>
<td>• works within frame of reference set by scrolling</td>
</tr>
<tr>
<td></td>
<td>• demands precision (fine)</td>
</tr>
</tbody>
</table>

![Microsoft Office Keyboard](image-url)
The GOMS Model

• **G**: goals
  – (Verbal) description of what a user wants to accomplish
  – Various levels of complexity possible

• **O**: operators
  – Possible actions in the system
  – Various levels of abstraction possible (sub-goals / ... / keystrokes)

• **M**: methods
  – Sequences of operators that achieve a goal

• **S**: selection rules
  – Rules that define when a user employs which method

• User tasks are split into goals which are achieved by solving sub-goals in a divide-and-conquer fashion

GOMS Example: Move Word (1 / 2)

Goal: move the word starting at the cursor position to the end of the text
[select use-keyboard delete-and-write use-mouse]
verify move

Main goal with methods

Goal: use-keyboard
Goal: select word
[select use <shift> and n*<cursor right>
use <shift> and <ctrl> and <cursor right>]
verify selection

Sub-goal

Method 1

Goal: delete-and-write
...

Method 2

Goal: use-mouse
Goal: select word
[select click at beginning and drag till the end of the word
double-click on the word]
verify selection

Goal: move word
[select click on word and drag till end of text
Goal: copy-paste-with-mouse
...]

Method 3
GOMS Example: Move Word (2 / 2)

• Selection rules:
  – Rule 1: use method **use-keyboard** if no mouse attached
  – Rule 2: use method **delete-and-write** if length of word < 4
  – Rule 3: use method **use-mouse** if hand at mouse before action
  – ...

• Selection rules depend on the user (→ remember user diversity?)

• GOMS models can be derived in various levels of abstraction
  – e.g. goal: write a paper about X
  – e.g. goal: open the print dialog
**GOMS Example: Closing a Window**

**GOAL**: CLOSE-WINDOW  

[select **GOAL**: USE-MENU-METHOD  

MOVE-MOUSE-TO-FILE-MENU  
PULL-DOWN-FILE-MENU  
CLICK-OVER-CLOSE-OPTION  

**GOAL**: USE-CTRL-F4-METHOD  
PRESS-CONTROL-F4-KEYS]

For a particular user:

Rule 1: Select USE-MENU-METHOD unless another rule applies  

Rule 2: If the application is GAME, select CTRL-F4-METHOD
GOMS Example: ATM Machine

- GOMS gives an early understanding of interactions
- “How to not loose you card”

**GOAL: GET-MONEY**
- GOAL: USE-CASH-MACHINE
  - INSERT-CARD
  - ENTER-PIN
  - SELECT-GET-CASH
  - ENTER-AMOUNT
  - COLLECT-MONEY
  *(outer goal satisfied!)*
- COLLECT-CARD

**GOAL: GET-MONEY**
- GOAL: USE-CASH-MACHINE
  - INSERT-CARD
  - ENTER-PIN
  - SELECT-GET-CASH
  - ENTER-AMOUNT
  - COLLECT-CARD
  *(outer goal satisfied!)*
- COLLECT-MONEY
Some GOMS Variations

- **(CMN-)GOMS**
  - Plain GOMS
  - Pseudo-code
  - First introduced by Card, Moran and Newell
  - (This is the version we looked at)

- **KLM**
  - Keystroke-Level Model
  - Simplified version of GOMS
  - (See next slides)

- **NGOMSL**
  - Natural GOMS Language
  - Stricter version of GOMS
  - Provides more well-defined, structured natural language
  - Estimates learning time

- **CPM-GOMS**
  - Cognitive Perceptual Motor analysis of activity
  - Critical Path Method
  - Based on the parallel multi-processor stage of human information processing

GOMS – Characteristics

• Usually one high-level goal

• Measurement of performance: high depth of goal structure
  → high short term-memory requirements

• Predict task completion time (see KLM in the following)
  → compare different design alternatives
Keystroke-Level Model

- Simplified version of GOMS
  - only operators on keystroke-level
  - no sub-goals
  - no methods
  - no selection rules
- KLM predicts how much time it takes to execute a task
- Execution of a task is decomposed into primitive operators
  - Physical motor operators
    » pressing a button, pointing, drawing a line, …
  - Mental operator
    » preparing for a physical action
  - System response operator
    » user waits for the system to do something
Models: Levels of Detail

- Different levels of detail for the steps of a task performed by a user

- **Abstract:** correct wrong spelling

- **Concrete:** mark-word
delete-word
type-word

- **Keystroke-Level:** hold-shift
  n · cursor-right
  recall-word
del-key
  n · letter-key
## KLM Operators

- Each operator is assigned a duration (amount of time a user would take to perform it):

<table>
<thead>
<tr>
<th>Operator</th>
<th>Description</th>
<th>Associated Time</th>
</tr>
</thead>
</table>
| K        | Keystroke, typing one letter, number, etc. or function key like ‘CTRL’, ‘SHIFT’ | Expert typist (90 wpm): 0.12 sec  
Average skilled typist (55 wpm): 0.20 sec  
Average non-secretarial typist (40 wpm): 0.28 sec  
Worst typist (unfamiliar with keyboard): 1.2 sec |
| H        | ‘Homing’, moving the hand between mouse and keyboard                           | 0.4 sec                                                                         |
| B / BB   | Pressing / clicking a mouse button                                            | 0.1 sec / 2*0.1 sec                                                            |
| P        | Pointing with the mouse to a target                                           | 0.8 to 1.5 sec with an average of 1.1 sec  
Can also use Fitts’ Law                      |
| D(nD, lD)| Drawing nD straight line segments of length lD                                | 0.9*nD + 0.16*lD                                                               |
| M        | Subsumed time for mental acts; sometimes used as ‘look-at’                    | 1.35 sec (1.2 sec according to [Olson and Olson 1995])                          |
| R(t) or W(t) | System response (or ‘work’) time, time during which the user cannot act | Dependent on the system, to be determined on a system-by-system basis           |
Predicting the Task Execution Time

• Execution Time
  – OP: set of operators
  – $n_{op}$: number of occurrences of operator op

  \[ T_{execute} = \sum_{op \in OP} n_{op} \cdot op \]

• Example task on Keystroke-Level:

  1. hold-shift
  2. $n \cdot$ cursor-right
  3. recall-word
  4. del-key
  5. $n \cdot$ letter-key

  Sequence:
  K (Key)
  $n \cdot K$
  M (Mental Thinking)
  K
  $n \cdot K$

• Operator Time Values: K = 0.28 sec. and M = 1.35 sec

  \[ 2n \cdot K + 2 \cdot K + M = 2n \cdot 0.28 + 1.91 \text{ sec} \]

• $\rightarrow$ time it takes to replace a n=7 letter word: $T = 5.83 \text{ sec}$
Keystroke-Level Model – Example Task

**Task:** in MS Word, add a 6pt space after the current paragraph

→ Word 2003:

<table>
<thead>
<tr>
<th>Actions</th>
<th>Operator (keyboard)</th>
<th>Time allocated</th>
<th>Operator (mouse)</th>
<th>Time allocated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locate menu ‘Format’</td>
<td>M</td>
<td>1.35</td>
<td>M</td>
<td>1.35</td>
</tr>
<tr>
<td>Press ALT-o or mouse click</td>
<td>KK</td>
<td>2*0.28</td>
<td>PB</td>
<td>1.10+0.10</td>
</tr>
<tr>
<td>Locate entry ‘Paragraph’</td>
<td>M</td>
<td>1.35</td>
<td>M</td>
<td>1.35</td>
</tr>
<tr>
<td>Press ‘p’ or mouse click</td>
<td>K</td>
<td>0.28</td>
<td>PB</td>
<td>1.10+0.10</td>
</tr>
<tr>
<td>Locate item in dialogue</td>
<td>M</td>
<td>1.35</td>
<td>M</td>
<td>1.35</td>
</tr>
<tr>
<td>Point to item</td>
<td>KK</td>
<td>0.28</td>
<td>PB</td>
<td>1.10+0.10</td>
</tr>
<tr>
<td>Enter a 6 for a 6pt space</td>
<td>K</td>
<td>0.28</td>
<td>K</td>
<td>0.28</td>
</tr>
<tr>
<td>Close the dialogue (ENTER)</td>
<td>K</td>
<td>0.28</td>
<td>K</td>
<td>0.28</td>
</tr>
</tbody>
</table>


→ Word 2007:
## GOMS vs. KLM

<table>
<thead>
<tr>
<th><strong>(CMN-)GOMS</strong></th>
<th><strong>KLM</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Pseudo-code (no formal syntax)</td>
<td>• Simplified version of GOMS</td>
</tr>
</tbody>
</table>
| • Very flexible | • Only operators on keystroke-level  
  ⇒ focus on very low level tasks |
| • Goals and subgoals | • No multiple goals |
| • Methods are informal programs | • No methods |
| • Selection rules | • No selection rules  
  ⇒ strictly sequential |
|  ⇒ tree structure: use different branches for different scenarios | • Quick and easy |
| • Time consuming to create | |

### Problem with GOMS in general

- Only for well defined routine cognitive tasks
- Assumes statistical experts
- Does not consider slips or errors, fatigue, social surroundings, …
Extensions for Novel Mobile Interactions

• Current mobile interactions use
  – Keypad, hotkeys
  – Microphone, camera (marker detection)
  – Sensors like accelerometers
  – Tag readers (NFC)
  – Bluetooth

• Method
  – Large set of studies
  – Software on the phone
  – Video frame-by-frame analysis
  – Eye-tracker
  → Total number of actions measured: 2134
### KLM – Original and New Operators

<table>
<thead>
<tr>
<th>Original Operators</th>
<th>New Operators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental Act, M</td>
<td>Micro attentions Shift, $S_{Micro}$</td>
</tr>
<tr>
<td>System Response, R</td>
<td>Macro attention shift, $S_{Macro}$</td>
</tr>
<tr>
<td>Keystroke / button press, K</td>
<td>Finger movement F</td>
</tr>
<tr>
<td>Homing, H</td>
<td>Distraction X</td>
</tr>
<tr>
<td>Pointing, P</td>
<td>Gesture G</td>
</tr>
<tr>
<td></td>
<td>Initial preparation I</td>
</tr>
</tbody>
</table>

- Mental Act, M and System Response, R are unchanged.
- Micro attentions Shift, $S_{Micro}$, Macro attention shift, $S_{Macro}$, Finger movement F, Distraction X, Gesture G, and Initial preparation I are added.
Micro Attention Shift, $S_{\text{Micro}}$

Switch attention between phone parts

- display
- hot keys
- keypad
$S_{\text{Micro}}$ – Operator Time Estimation

- Measured with a standard eye tracker
- Mobile phone in front of the monitor
**S\textsubscript{Micro} – Operator Time Estimation**

**Study**
- 10 participants, 24-34 years, 6 female
- 1500 shifts detected
- Using automatic eye-tracking
- 3 pre-set tasks

<table>
<thead>
<tr>
<th>Time (sec)</th>
<th>Micro Attention Shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>display ↔ hotkeys: 0.14 sec.</td>
</tr>
<tr>
<td></td>
<td>display ↔ keypad: 0.12 sec.</td>
</tr>
<tr>
<td></td>
<td>keypad ↔ hotkeys: 0.04 sec.</td>
</tr>
</tbody>
</table>
Distraction, X

Study

- 10 participants, 24-33 years, 3 female
- Short message in 3 settings (quiet room, standing outside, walking)
- Relative slow-down (significant: $t=2.23$, $p<0.03$ and $t=3.28$, $p<0.01$)

Distraction: multiplicative

$X_{\text{slight}} = 6\%$, $X_{\text{strong}} = 21\%$
Extended KLM – Time Prediction

Total Execution Time:

\[ T_{execute} = \sum_{op \in OP} (n_{op} + d_{op} \cdot X_{slight} + D_{op} \cdot X_{strong}) \cdot op \]

Set of Available Operators:

\[ OP = \{ A, F, G, H, I, K, M, P, R, S_{Micro}, S_{Macro} \} \]
Extended KLM – Empirical Validation

• Task: buy a public transportation ticket from A to B
• Implemented 2 ways of performing the task
  – Access through mobile web browser
  – Direct interaction with NFC tags
• Created the two Keystroke-Level Models
• Study: 9 people, 23-34 years, 3 female
Extended KLM – Empirical Validation

Browser Interaction

- KLM
- Study

NFC Interaction

- KLM
- Study

Time in seconds

Predicted speed loss: 17%
Actual speed loss: 14%
## Advanced Mobile Phone KLM – Values

<table>
<thead>
<tr>
<th>Operator</th>
<th>Time</th>
<th>Qu. 1</th>
<th>Qu. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, Action</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>picture / marker</td>
<td>1.23</td>
<td>0.61</td>
<td>1.44</td>
</tr>
<tr>
<td>NFC</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>in general</td>
<td>variable, input to model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B, Mouse Button Press</td>
<td>not applicable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D, Mouse Drawing</td>
<td>not applicable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F, Finger Movement</td>
<td>0.23</td>
<td>0.20</td>
<td>0.29</td>
</tr>
<tr>
<td>G, Gestures</td>
<td>0.80</td>
<td>0.73</td>
<td>0.87</td>
</tr>
<tr>
<td>H, Homing</td>
<td>0.95</td>
<td>0.81</td>
<td>1.00</td>
</tr>
<tr>
<td>I, Initial Act</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>external trigger</td>
<td>5.32</td>
<td>3.98</td>
<td>7.51</td>
</tr>
<tr>
<td>self triggered</td>
<td>3.89</td>
<td>2.23</td>
<td>4.89</td>
</tr>
<tr>
<td>optimal setting</td>
<td>1.18</td>
<td>1.10</td>
<td>1.26</td>
</tr>
<tr>
<td>no assumptions</td>
<td>4.61</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>K, Keystroke</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>keypad average</td>
<td>0.39</td>
<td>0.37</td>
<td>0.48</td>
</tr>
<tr>
<td>keypad quick</td>
<td>0.33</td>
<td>0.32</td>
<td>0.37</td>
</tr>
<tr>
<td>hotkey</td>
<td>0.16</td>
<td>0.15</td>
<td>0.20</td>
</tr>
<tr>
<td>M, Mental Act</td>
<td>1.35</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>P, Pointing</td>
<td>1.00</td>
<td>0.84</td>
<td>1.20</td>
</tr>
<tr>
<td>R, System Response Time</td>
<td>2.58</td>
<td>2.46</td>
<td>2.80</td>
</tr>
<tr>
<td>visual marker</td>
<td>2.22</td>
<td>2.09</td>
<td>2.82</td>
</tr>
<tr>
<td>in general</td>
<td>variable, input to model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S_{Macro}, Macro Attention Shift</td>
<td>0.36</td>
<td>0.28</td>
<td>0.44</td>
</tr>
<tr>
<td>S_{Micro}, Micro Attention Shift</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>keypad ↔ display</td>
<td>0.14</td>
<td>0.14</td>
<td>0.19</td>
</tr>
<tr>
<td>hotkey ↔ display</td>
<td>0.12</td>
<td>0.02</td>
<td>0.14</td>
</tr>
<tr>
<td>keypad ↔ hotkey</td>
<td>0.04</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>in general</td>
<td>0.14</td>
<td>0.10</td>
<td>0.16</td>
</tr>
<tr>
<td>X, Distraction</td>
<td>slight</td>
<td>6%</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>strong</td>
<td>21%</td>
<td>11%</td>
</tr>
</tbody>
</table>
Using KLM

• KLM can help evaluate UI designs, interaction methods and trade-offs
• If common tasks take too long or consist of too many statements, shortcuts can be provided
• Predictions are mostly remarkably accurate: +/- 20%
Mobile Phone KLM Task: Set an Alarm

• Write down the KLM
• With the help of your neighbour, take 3 measurements
• Evaluate the accuracy

Back to KLM values
Set an alarm

M + S\textsuperscript{micro} (display – hot key)
+ K, hot key (CLOCK)

S\textsuperscript{micro} (hot key–display) + M + S\textsuperscript{micro} (hot key–display) + K, key pad (RIGHT)
Set an alarm

\[
\text{Set an alarm:}
\begin{align*}
\text{S} & \text{micro (key pad)} \\
& \quad \text{display} \\
& \quad + \\
& \quad \text{M} \\
& \quad + \\
& \quad \text{S} \text{micro (display – hot key)} \\
& \quad + \\
& \quad \text{K, hot key} \\
& \quad \text{(OPTIONS)} \\
& \quad + \\
& \quad \text{S} \text{micro (hot key – display)} \\
& \quad + \\
& \quad \text{M} \\
& \quad + \\
& \quad \text{K, hot key} \\
& \quad \text{(SELECT)}
\end{align*}
\]
Set an alarm

M + S\text{micro (display – key pad)}
+ K, key pad
avg(1)
+ K, key pad
avg(3)

\text{DONE}
Set an alarm

![Image of a mobile phone showing an alarm set for 13:15 on Monday, 01/01.](image-url)
Weaknesses of GOMS et al.

• Just spending time is not modelled
• Difficult to target specific users
• No real users
• Difficult to model novel interactions
• Various variable parameters
• Users like to have impact
Strengths of GOMS et al.

• Good treatment of learning effects
  – Measurement of learnability
  – Independence of sequences
  – Measurement of knowledge requirements

• Good results
  – Gives reasons
  – Helps in decision making
  – Identifies bottlenecks
  – Provides illustrative figures
  – Combines various views
  – Treats feasibility and cognitive load

• Less cost in money and time
  – Quick to apply
  – Quick to prepare
  – Helpful to design
  – Cheap to apply
  – Easy to repeat
  – Quick to analyse
  – Precise to interpret
  – Easy to convey
GOMS / KLM Summary Example

• Example prototype: the Combimouse

• Ergonomic models followed

• Follows Guiard’s model of bimanual control
  (for right handed people scrolling with the non-preferred hand)

• Removes KLM’s Homing operator (H ~ 1 sec.)

http://www.combimouse.com
References

GOMS


KLM


Mobile Phone KLM