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: $MT = a + b\int_C \frac{ds}{W(s)}$
- 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)
  - ...

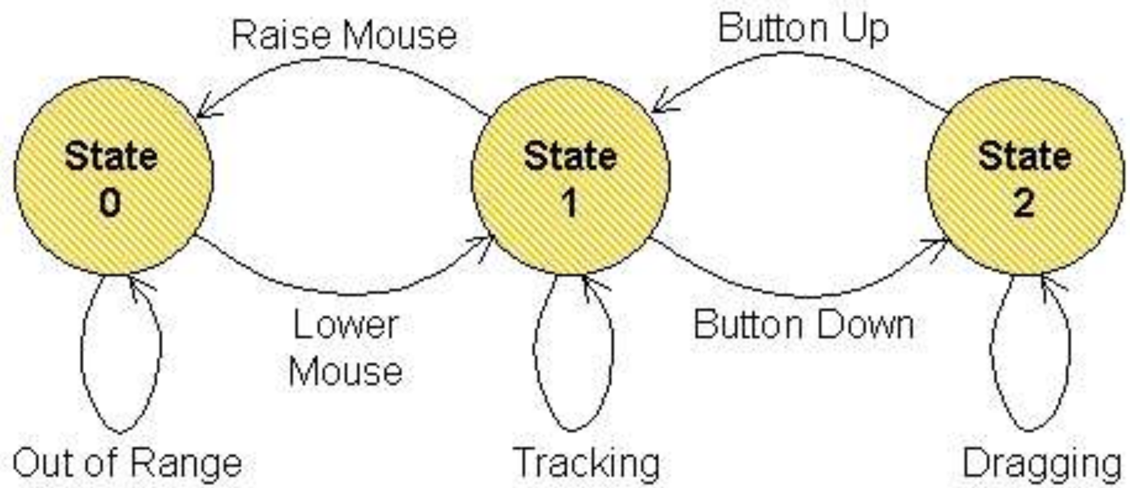
Descriptive Models

- *(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

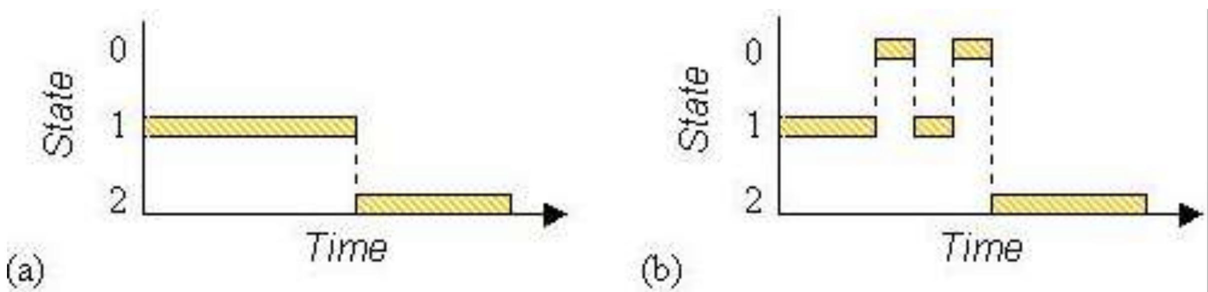
MacKenzie, I. S., 2003, Motor Behaviour Models for Human-computer Interaction
In *HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science (Book), 27-54*

Example: Three-State Model (W. Buxton)

- Describes graphical input
- Simple, quick, expressive
- Possible extensions:
  - multi-button interaction
  - stylus input
  - direct vs. indirect input



Buxton, W, 1990, A Three-State Model of Graphical Input
In INTERACT'90, 449-456

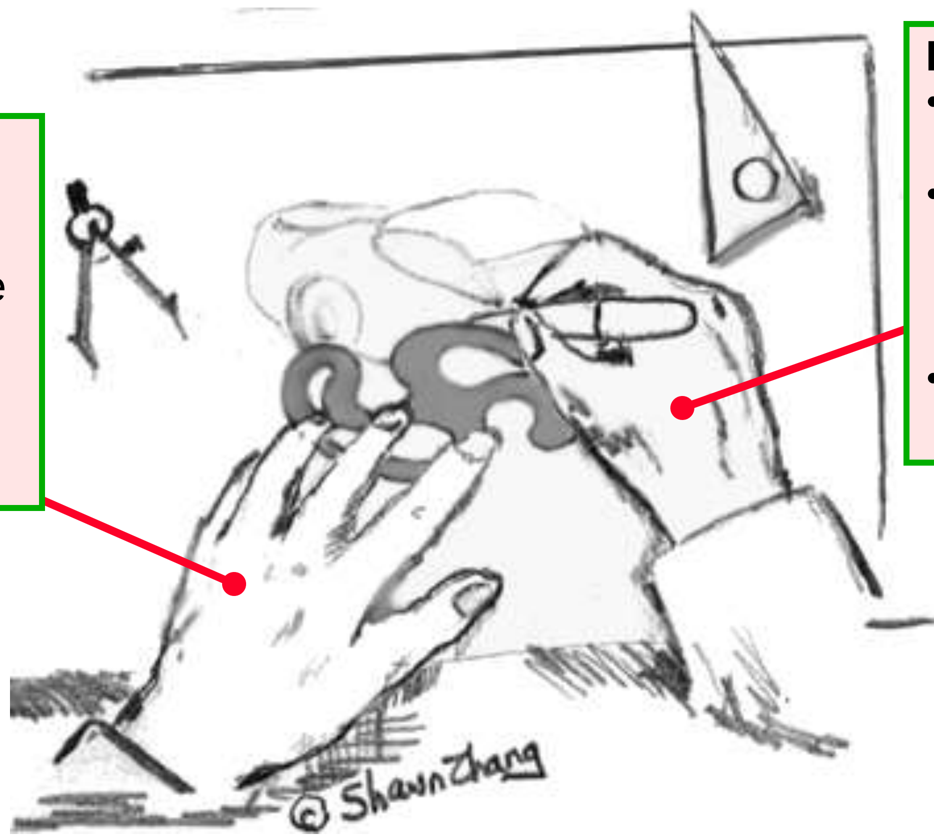


Dragging tasks: (a) mouse (b) lift-and-tap touchpad. [MacKenzie 2003]

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)

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

Microsoft Office Keyboard

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

Card, S. K.; Newell, A.; Moran, T. P., 1983, The Psychology of Human-Computer Interaction (Book)

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 ($\rightarrow$ 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
  .   COLLECT-MONEY
(outer goal satisfied!)
```

Some GOMS Variations

GOMS

(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

John, B., Kieras, D., 1996, Using GOMS for user interface design and evaluation: which technique?
ACM Transactions on Computer-Human Interaction, 3, 287-319.

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

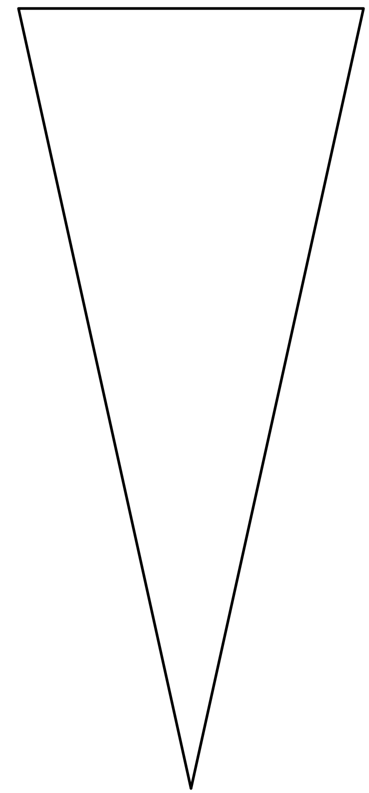
Models: Levels of Detail

- 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):

Operator	Description	Associated Time
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
$D(n_D, l_D)$	Drawing n_D straight line segments of length l_D	$0.9*n_D + 0.16*l_D$
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:

	Sequence:
1. `hold-shift`	K (Key)
2. `n·cursor-right`	n · K
3. `recall-word`	M (Mental Thinking)
4. `del-key`	K
5. `n·letter-key`	n · K

- Operator Time Values: K = 0.28 sec. and M = 1.35 sec
  $2n \cdot K + 2 \cdot K + M = 2n \cdot 0.28 + 1.91$ sec

- → time it takes to replace a n=7 letter word: T = 5.83 sec

Keystroke-Level Model – Example Task

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

→ Word 2003:

Actions	Operator (keyboard)	Time allocated	Operator (mouse)	Time allocated
Locate menu 'Format'	***M***	1.35	***M***	1.35
Press ALT-o or mouse click	***K***,***K***	2*0.28	***P***,***B***	1.10+0.10
Locate entry 'Paragraph'	***M***	1.35	***M***	1.35
Press 'p' or mouse click	***K***	0.28	***P***,***B***	1.10+0.10
Locate item in dialogue	***M***	1.35	***M***	1.35
Point to item	***K***,***K***	0.28	***P***,***B***	1.10+0.10
Enter a 6 for a 6pt space	***K***	0.28	***K***	0.28
Close the dialogue (ENTER)	***K***	0.28	***K***	0.28

→ Word 2007:

Sum (keyboard): 7.22 sec. **Sum (mouse): 7.65 sec.**

GOMS vs. KLM

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

Problem with GOMS in general

- Only for well defined routine cognitive tasks
- Aassumes 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

- Mental Act, M
- System Response, R

← unchanged

- Keystroke / button press, K
- Homing, H
- Pointing, P

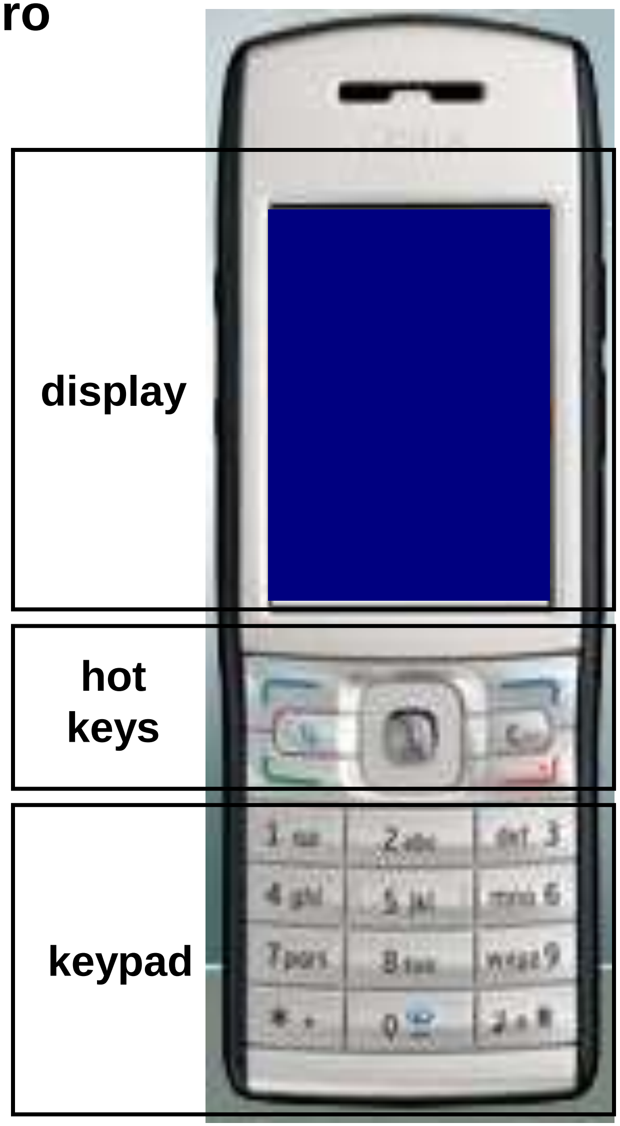
← adopted

- Micro attentions Shift, S_{Micro}
- Macro attention shift, S_{Macro}
- Finger movement F
- Distraction X
- Gesture G
- Initial preparation I

← added

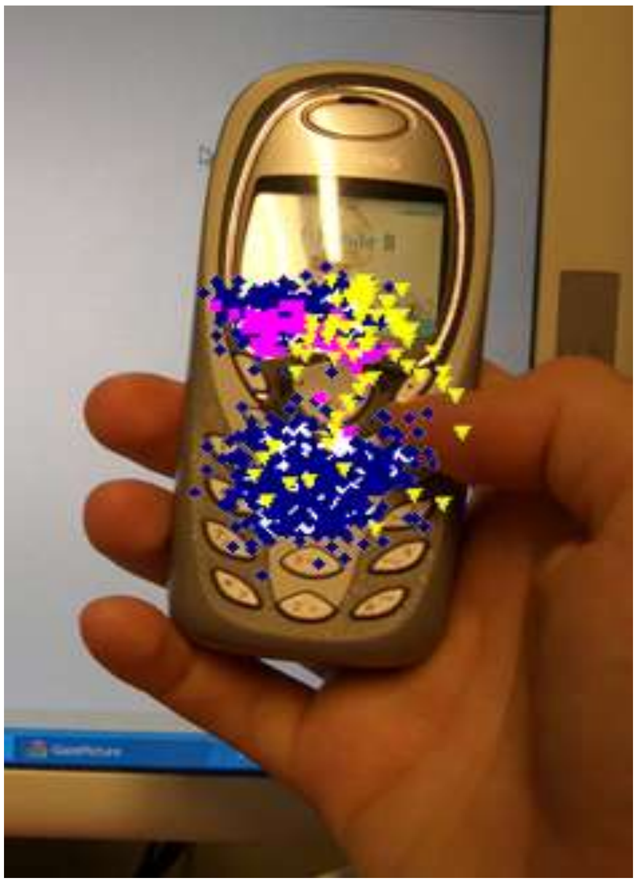
Micro Attention Shift, S_{Micro}

Switch attention between phone parts



S_{Micro} – Operator Time Estimation

- Measured with a standard eye tracker
- Mobile phone in front of the monitor

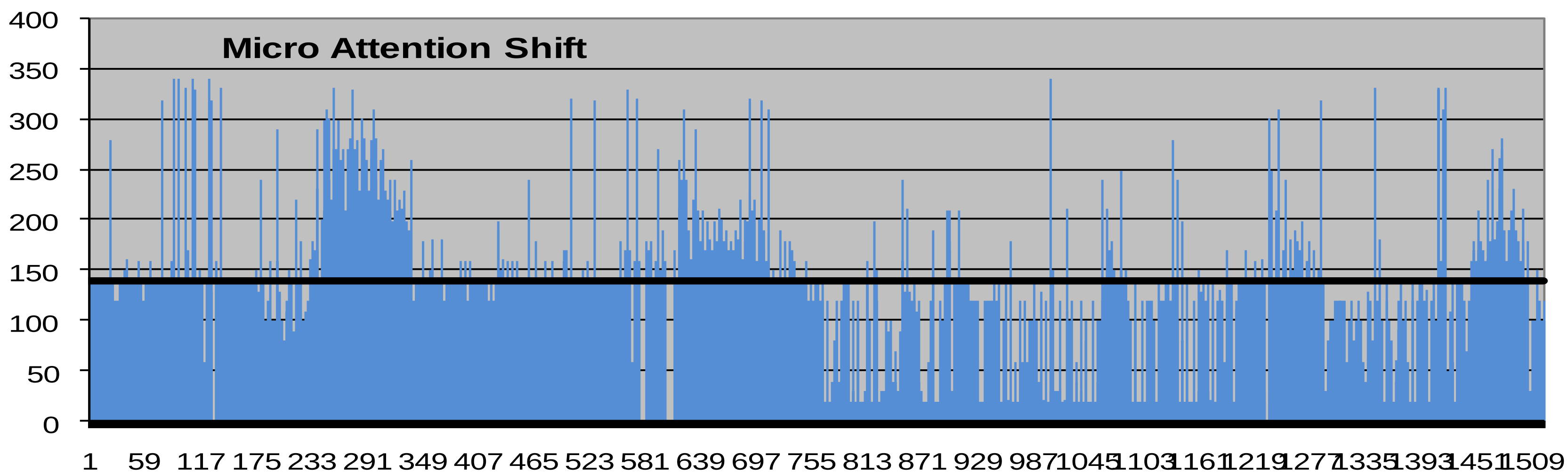


S_{Micro} – Operator Time Estimation

Study

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

display ↔ hotkeys: 0.14 sec.
display ↔ keypad: 0.12 sec.
keypad ↔ hotkeys: 0.04 sec.



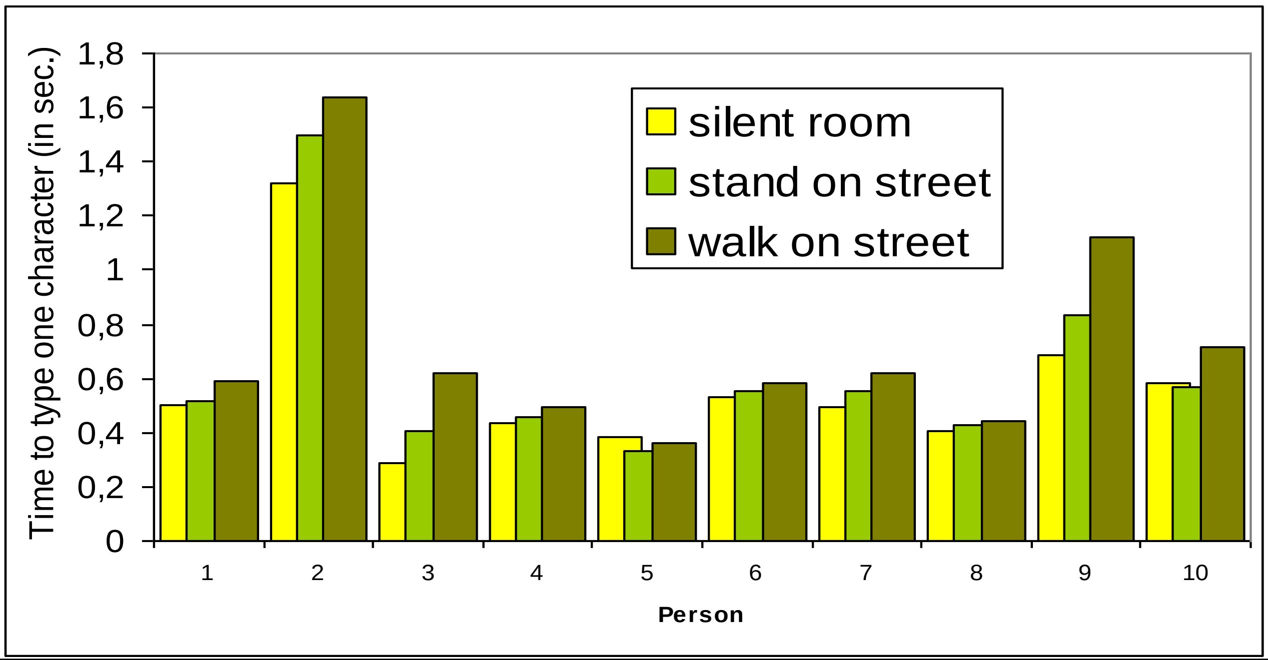
Distraction, X

Distraction: multiplicative

X_{slight} = 6%, X_{strong} = 21%

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 $t=3.28$, $p<0.01$)



Extended KLM – Time Prediction

Total Execution Time:

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

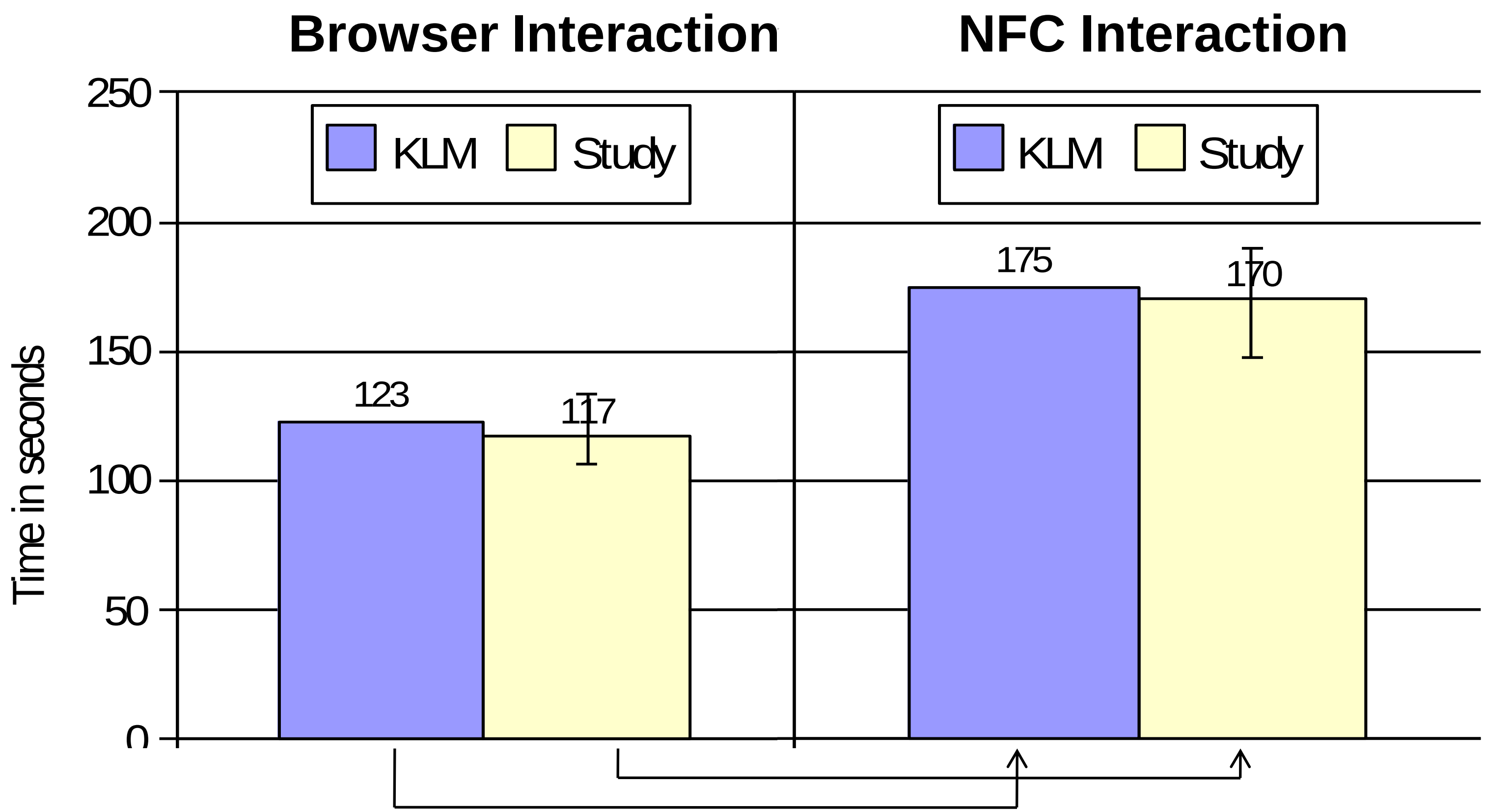
Set of Available Operators:

$OP =$ {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



Predicted speed loss: 17%

Actual speed loss: 14%

Advanced Mobile Phone KLM – Values

Operator		Time	Qu. 1	Qu. 3
***A**, Action*	picture / marker	**1.23**	0.61	1.44
	NFC	**0.00**	-	-
	in general	*variable, input to model*		
***B**, Mouse Button Press*		*not applicable*		
***D**, Mouse Drawing*		*not applicable*		
***F**, Finger Movement*		**0.23**	0.20	0.29
***G**, Gestures*		**0.80**	0.73	0.87
***H**, Homing*		**0.95**	0.81	1.00
***I**, Initial Act*	external trigger	**5.32**	3.98	7.51
	self triggered	**3.89**	2.23	4.89
	optimal setting	**1.18**	1.10	1.26
	no assumptions	**4.61**	-	-
***K**, Keystroke*	keypad average	**0.39**	0.37	0.48
	keypad quick	**0.33**	0.32	0.37
	hotkey	**0.16**	0.15	0.20

Operator		Time	Qu. 1	Qu. 3
***M**, Mental Act*		**1.35**	-	-
***P**, Pointing*		**1.00**	0.84	1.20
***R**, System Response Time*	NFC	**2.58**	2.46	2.80
	visual marker	**2.22**	2.09	2.82
	in general	*variable, input to model*		
*S_{Macro}, **Macro Attention Shift***		**0.36**	0.28	0.44
*S_{Micro}, **Micro Attention Shift***	keypad ↔ display	**0.14**	0.14	0.19
	hotkey ↔ display	**0.12**	0.02	0.14
	keypad ↔ hotkey	**0.04**	0.02	0.12
	in general	**0.14**	0.10	0.16
***X**, Distraction*	slight	**6 %**	3 %	13 %
	strong	**21 %**	11 %	25 %

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 remarkable 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
+
Smicro (display – hot key)
+
K, hot key (CLOCK)

Smicro (hot key–display)
+
M
+
Smicro (hot key–display)
+
K, key pad (RIGHT)

Set an alarm

Smicro (key pad -display)
+
M
+
Smicro (display – hot key)
+
K, hot key (OPTIONS)
+
Smicro (hot key– display)

M
+
K, hot key (SELECT)

Set an alarm

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

S_{micro} (key pad– display)
+
M
+
S_{micro} (display – hot key)
+
K, hot key (DONE)

Set an alarm

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

- Card S. K., Newell A., Moran T. P. (1983). The Psychology of Human-Computer Interaction. *Lawrence Erlbaum Associates Inc*.
- Card S. K., Moran T. P., Newell A. (1980). The Keystroke-level Model for User Performance Time with Interactive Systems. *Communication of the ACM 23(7)*. 396-410
- John, B., Kieras, D. (1996). Using GOMS for user interface design and evaluation: which technique? *ACM Transactions on Computer-Human Interaction, 3,* 287-319.

KLM

- Kieras, D. (1993, 2001). Using the Keystroke-Level Model to Estimate Execution Times. *University of Michigan. Manuscript.*

Mobile Phone KLM

- Holleis, P., Otto, F., Hussmann, H., Schmidt, A. (2007). Keystroke-Level Model for Advanced Mobile Phone Interaction, *CHI '07*