Mensch-Computer-Interaktion 2

Modeling Interaction

Human-Computer Interaction Group

Prof. Dr. Michael Rohs
michael.rohs@hci.uni-hannover.de

Lectures

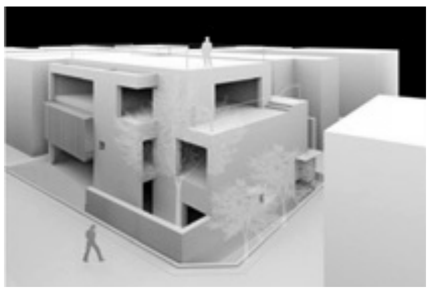
Session	Date	Topic	
1	6.4.	Introduction	
2	13.4.	Interaction elements	GUI toolkits, interaction techniques
3	20.4.	Event handling	
4	27.4.	Scene graphs	
5	4.5.	Interaction techniques	
	11.5.	no class (CHI)	
	18.5.	no class (spring break)	
6	25.5.	Experiments	design and analysis of experiments
7	1.6.	Data Analysis	
8	8.6.	Data Analysis	
9	15.6.	Modeling interaction	
10	22.6.	Visualization	current topics beyond-desktop UIs
11	29.6.	Visualization	
12	6.7.	Computer vision for interaction	
13	13.7.	Computer vision for interaction	

Klausur:
28.7.2016
8-11 Uhr
HG E214

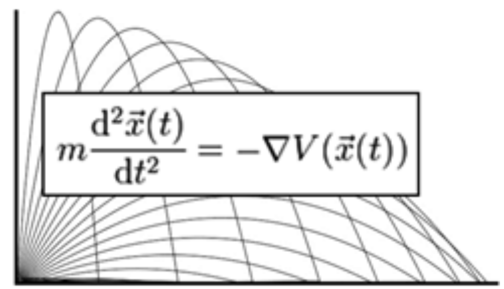
Introduction

- A model is a simplification of reality

Architect's scale model of a building



Physicist's model for the trajectory of a tossed ball



- Both are simplifications of complex phenomena
- The architect's model is a description
  - → Provides insight into space usage, movement of people, light, shade, etc.
- The physicist's model is a prediction
  - → Gives the ball's position as a function of time

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

DESCRIPTIVE MODELS

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Descriptive Models

- Descriptive models are everywhere
- Descriptive modeling is at times so simple, the process barely seems like modeling
- Any reduction or partitioning of a problem space qualifies as a descriptive model, for example:

Models	
Descriptive Models	Predictive Models

- Other names
  - Design space, framework, taxonomy, classification
- Partitioning helps us to think differently about the problem

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

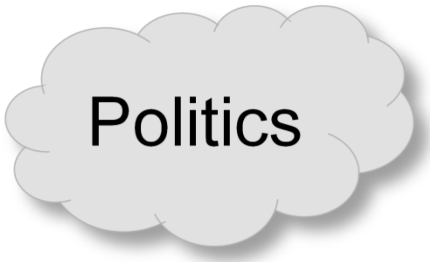
Descriptive Model Examples

- Politics
- Groupware
- Keyboards
- Two-handed input
- Emotion

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Big Fuzzy Cloud Model of Politics

- Below is a big fuzzy cloud model for politics:



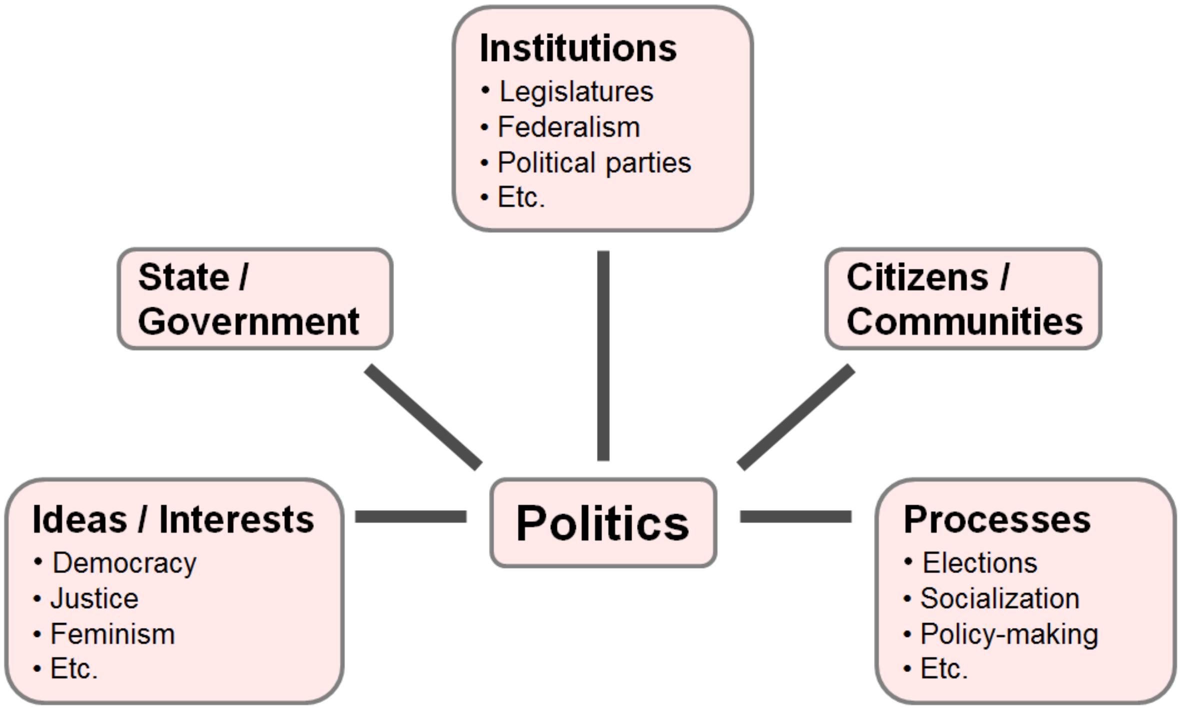
- Of course, there is no model – it's just the thing itself
- How do we go about making a descriptive model of politics?
- Break it down; partition the topic into parts
- What are the things that make up politics?
- How can they be labeled, presented, and organized?

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Descriptive Model of Politics[1]

[1] Johnston, L. (2007). Politics: An introduction to the modern democratic state (3rd ed.). Peterborough, Ontario: Broadview Press.

- With Johnston's model of politics, we are empowered to think differently about politics



- Is the model correct?
- Is there a different structure that might work better?
- Are the five items sufficient?
- Is federalism an institution?
- Is the model useful?

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Descriptive Model Examples

- Politics
- Groupware
- Keyboards
- Two-handed input
- Emotion

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

CSCW (Groupware)

- CSCW (computer supported cooperative work) is a research topic within HCI
- Concerned with people working collaboratively using computing technology
- Same challenge:
  - How do we go about making a descriptive model of groupware?
  - Break it down; partition the topic into parts
  - What are the things that make up groupware?
  - How can they be labeled, presented, and organized?

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Quadrant Model of Groupware[1]

[1] Johansen, R. (1991). Groupware: Future directions and wild cards. Journal of Organizational Computing and Electronic Commerce, 1(2), 219-227.

- A descriptive model
- Groupware partitioned into a 2 × 2 space
  - Location → same place | different places
  - Time → same time | different times

	Same Time	Different Times
Same Place	Copy boards PC Projectors Facilitation Services Group Decision Room Polling Systems	Shared Files Shift Work Kiosks Team Rooms Group Displays
Different Places	Conference Calls Graphics and Audio Screen Sharing Video Teleconferencing Spontaneous Meetings	Group Writing Computer Conferencing Conversational Structuring Forms Management Group Voice Mail

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Critiquing the Model

- The quadrant model of groupware was introduced in 1991
- The same questions apply:
  - Is the model correct? Is there a different structure that might work better? Is the model useful? Etc.
- Many of today's methods of collaborating didn't exist in 1991
- Contemporary groupware activities include
  - Sharing photos using camera phones, web cams, Skype, social media, blogging, tweeting
- Can these be positioned in the quadrant model of groupware?

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Descriptive Model Examples

- Politics
- Groupware
- Keyboards
- Two-handed input
- Emotion

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Keyboards

1874

2012

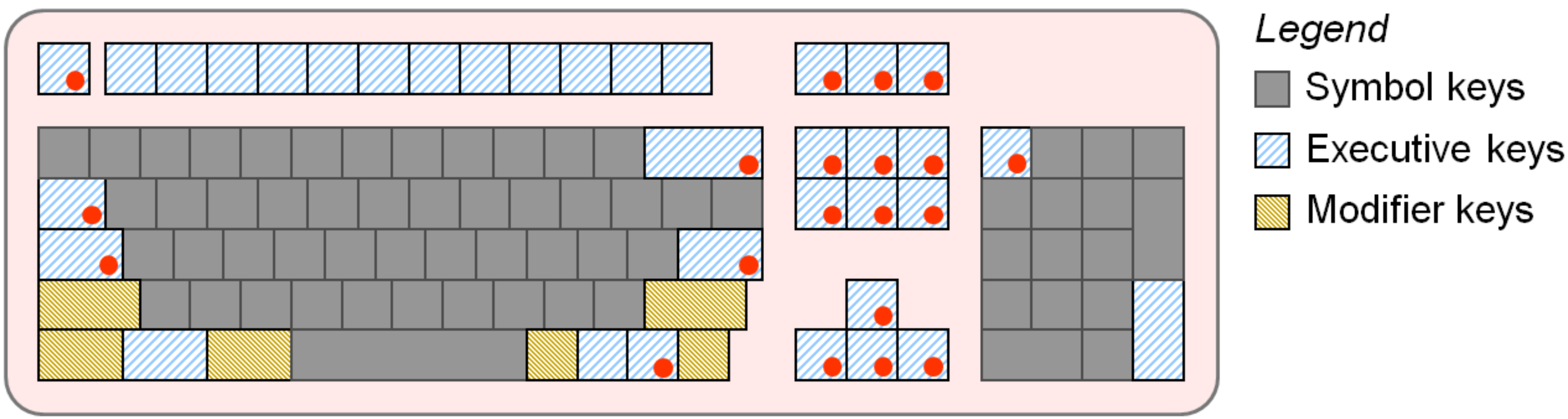
- Keyboards date at least to the 1870s with the introduction of the typewriter keyboard
- Today's keyboards retain the same core arrangement (qwerty), but have extra keys
  - 100+ keys can produce a wide variety of letters, symbols, commands, etc.
- Same challenge:
  - How do we go about making a descriptive model of keyboards?
  - Break it down; partition the topic into parts
  - What are the things that make up keyboards?
  - How can they be labeled, presented, and organized?

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Key-Action Model (KAM)[1]

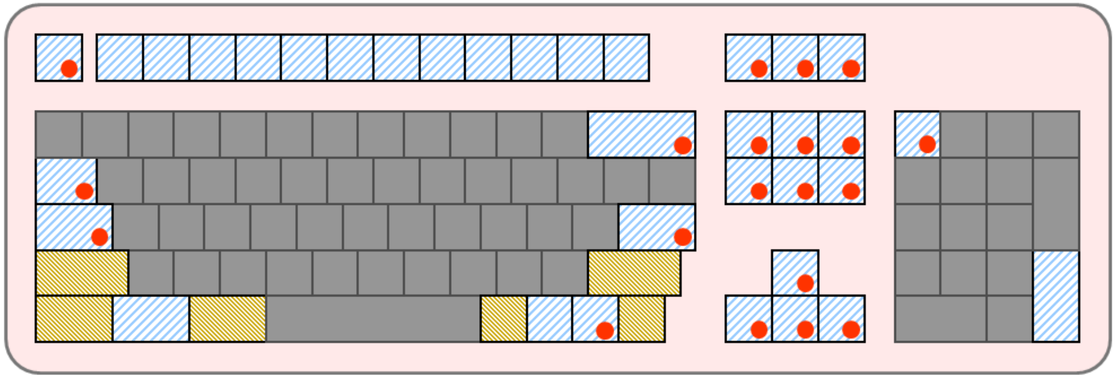
[1] MacKenzie, I. S. (2003). Motor behaviour models for human computer interaction. In J. M. Carroll (Ed.), HCI models, theories, and frameworks: Toward a multidisciplinary science (pp. 27-54).

- A descriptive model
- Three categories of keys
  - Symbol keys → produce graphic symbols (e.g., A)
  - Executive keys → invoke actions (e.g., Esc)
  - Modifier keys → modify effect of other keys (e.g., Ctrl)



MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Critiquing the Model



- Key-Action Model reveals the organization of the keyboard in terms key types
- Questions
  - Is the model correct? Do all keys fit the model? Are there additional categories to improve the model? Do some keys fit more than one category? Can the model be applied to other keyboards, such as mobile phone keyboards or soft keyboards? Is the model useful? Etc.
- Note: Red dots on previous slide identify executive keys that are not mirrored (excluding function keys)
  - Hmm... There seems to be a 3:17 right-side bias!

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Descriptive Model Examples

- Politics
- Groupware
- Keyboards
- Two-handed input
- Emotion

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Two-Handed Input

- Humans not only have two hands, they use their hands differently
- Most people have a hand preference
  - (Which hand do you use to write?)
- Study of hand usage is called laterality or bimanual control

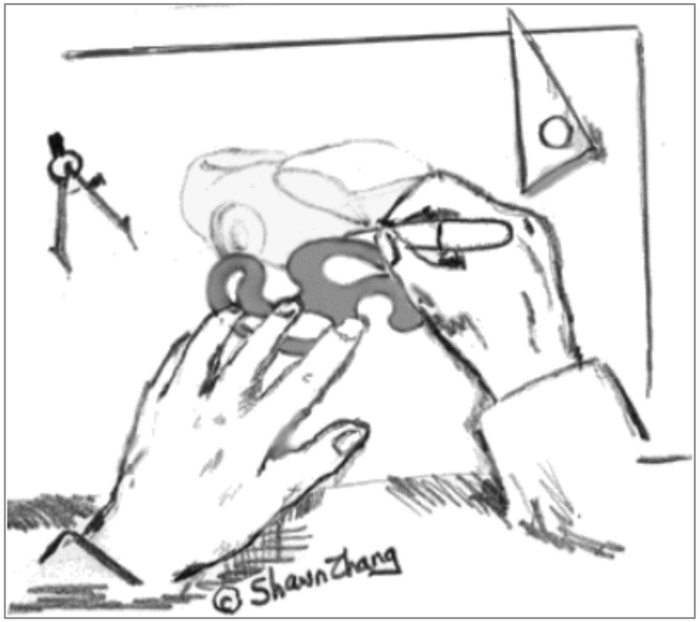
- Guiard[1] undertook a study of hand usage, examining the roles of the preferred and non-preferred hands in common tasks
- The result is a descriptive model

[1] Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. Journal of Motor Behavior, 19, 486-517.

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Guiard's Model of Bimanual Control[1]

Hand	Role and Action
Non-preferred	▪ Leads the preferred hand ▪ Sets the spatial frame of reference for the preferred hand ▪ Performs coarse movements
Preferred	▪ Follows the non-preferred hand ▪ Works within the established frame of reference set by the non-preferred hand ▪ Performs fine movements



Consider the bulleted points above in terms of the sketch on the right

[1] Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. Journal of Motor Behavior, 19, 486-517.

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Scrolling

- Scrolling in relation to other input tasks

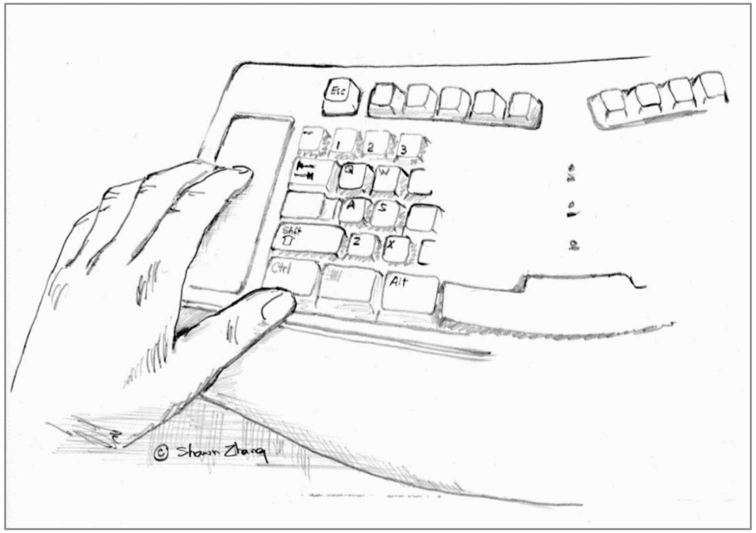
Task	Characteristics
Scrolling	▪ Precedes/overlaps other tasks ▪ Sets the frame of reference ▪ Minimal precision needed (coarse)
Selecting, editing, reading, drawing, etc.	▪ Follows/overlaps scrolling ▪ Works within frame of reference set by scrolling ▪ Demands precision (fine)

- Similarity with Guiard's model of bimanual control?

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Non-Preferred Hand Scrolling

- Guiard's model of bimanual control suggests that scrolling is a task well suited to the non-preferred hand:



- But what about mouse wheels?
- How to get scrolling into the non-preferred hand?

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Re-Engineered IntelliMouse

- Scrolling via the non-preferred hand (MacKenzie)
- Demonstration of non-preferred-hand scrolling given at Microsoft in Feb 1998 using a re-engineered IntelliMouse:

Two-years later

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Microsoft Office Keyboard

- Left-side wheel (roller) for scrolling via the non-preferred hand (for right-handed users):

- The left-side scroll wheel was eventually discontinued

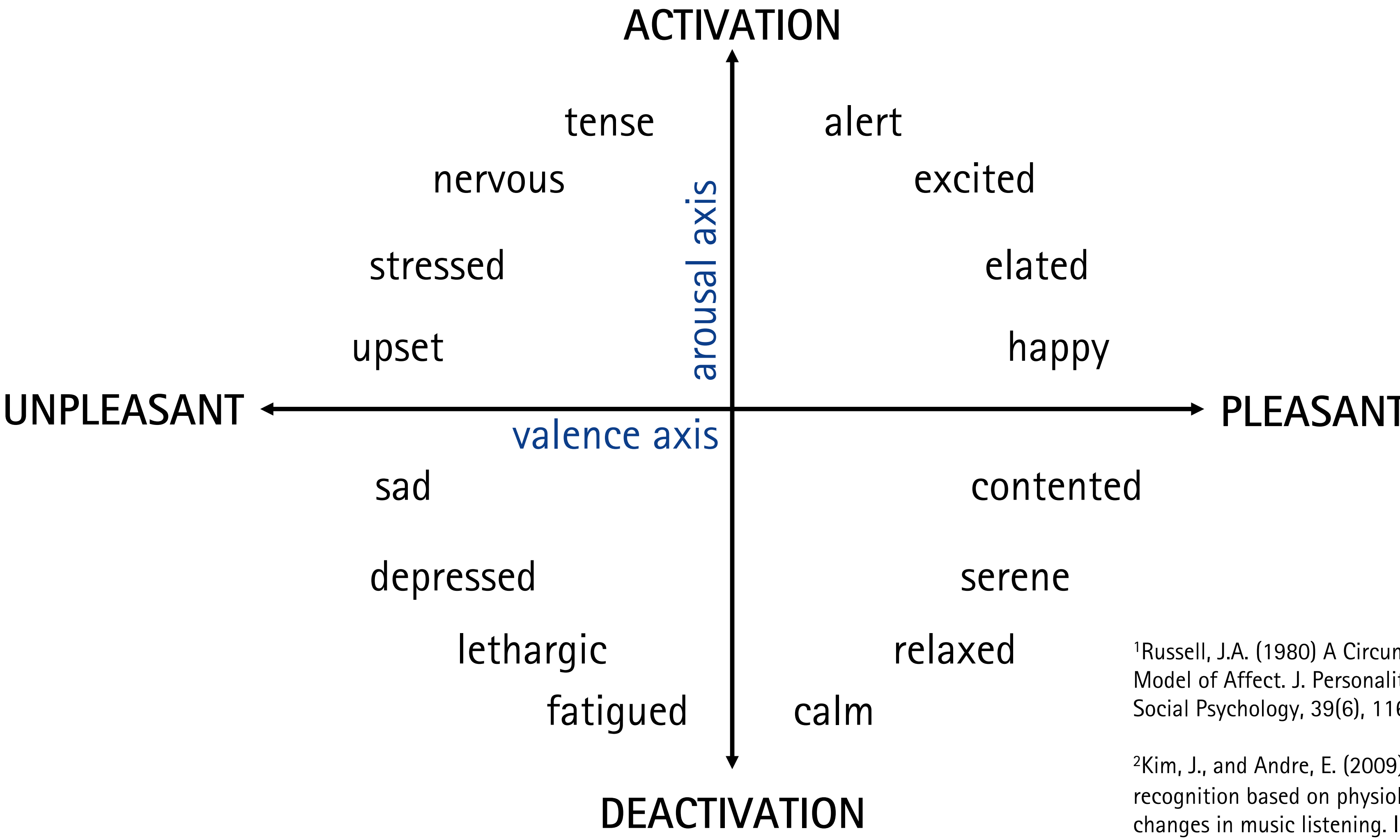
MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Descriptive Model Examples

- Politics
- Groupware
- Keyboards
- Two-handed input
- Emotion

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Circumplex Model of Emotion[1, 2]



MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

[1]Russell, J.A. (1980) A Circumplex Model of Affect. J. Personality and Social Psychology, 39(6), 1161-1178.

[2]Kim, J., and Andre, E. (2009) Emotion recognition based on physiological changes in music listening. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30, 2067-2083.

Circumplex Model of Emotion[1]

- How to come up with such a model?
- Russell[2]: "… there is other evidence that rather than being independent, these affective dimensions are interrelated in a highly systematic fashion. The evidence suggests that these interrelationships can be represented by a spatial model in which affective concepts fall in a circle in the following order: pleasure (0°), excitement (45°), arousal (90°), distress (135°), displeasure (180°), depression (225°), sleepiness (270°), relaxation (315°)."
- Evidence
  - How laymen conceptualize affective states
  - Multivariate analyses of self-reported affective states

[1]Russell, J.A. (1980) A Circumplex Model of Affect. J. Personality and Social Psychology, 39(6), 1161-1178.

Circumplex Model of Emotion[1]

- Category-sort task
  - Place each word into a category
  - 8 categories: pleasure, excitement, arousal, distress, misery, depression, sleepiness, contentment
  - 28 words: happy, delighted, excited, alarmed, afraid, annoyed, frustrated, sad, bored, tired, sleepy, calm, relaxed glad, pleased, …
- Circular ordering task
  - Place categories around the edge of a circle, such that
    (1) opposite words describe opposite feelings and
    (2) closer words describe more similar feelings

[1]Russell, J.A. (1980) A Circumplex Model of Affect. J. Personality and Social Psychology, 39(6), 1161-1178.

Result of Category-Sort Task[1]

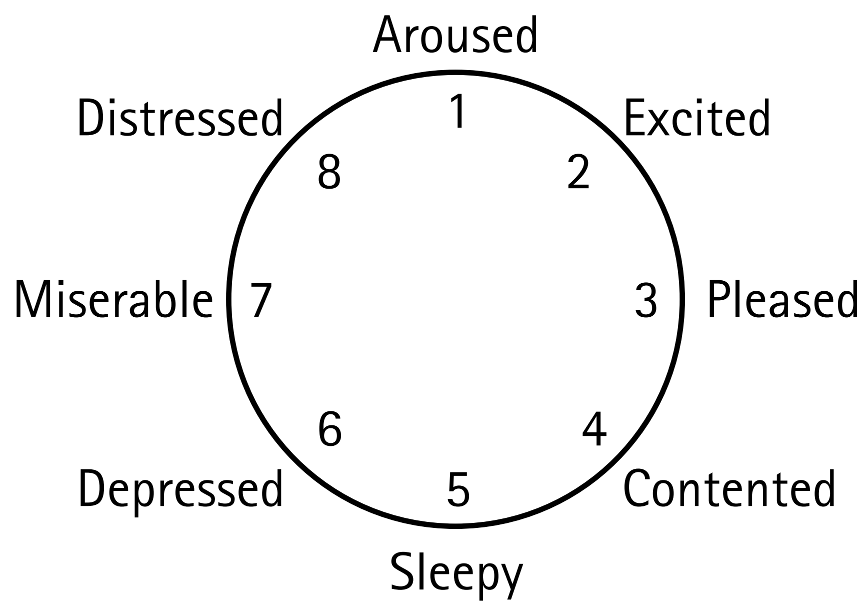
[1]Russell, J.A. (1980) A Circumplex Model of Affect. J. Personality and Social Psychology, 39(6), 1161-1178.

	Category							
Term	Pleasure	Excitement	Arousal	Distress	Misery	Depression	Sleepiness	Contentment
Happy	21	8	2					5
Delighted	15	16	3					2
Excited	2	29	5					
Astonished		17	18	1				
Aroused		14	21	1				
Tense		8	18	9		1		
Alarmed		6	19	11				
Angry		5	21	5	3	2		
Afraid		2	11	22		1		
Annoyed		1	12	14	4	4		1
Distressed			4	25	5	2		
Frustrated		2	5	19	4	6		
Miserable				3	23	10		
Sad				10	6	19		1
Gloomy				2	11	22	1	
Depressed				4	7	24		1
Bored				3	2	14	17	
Droopy				1	1	8	26	
Tired					1	1	34	
Sleepy					1		32	3
Calm	4						3	29
Relaxed	6						4	26
Satisfied	3	1						32
At ease	7						3	26
Content	6	1						29
Serene	8	2						26
Glad	20	4						12
Pleased	22	2	2					10

fuzziness of affect terms

Result of Circular Category Ordering Task[1]

Term	Position on circle							
	1	2	3	4	5	6	7	8
Aroused	*36*							
Excited		*24*	3	1				8
Pleased		9	*20*	7				
Contented		2	13	*16*	3			2
Sleepy				9	*23*	3		1
Depressed		1			5	*19*	10	1
Miserable				1	1	11	*18*	5
Distressed				2	4	3	8	*19*



[1]Russell, J.A. (1980) A Circumplex Model of Affect. J. Personality and Social Psychology, 39(6), 1161-1178.

PREDICTIVE MODELS

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Predictive Models

- A predictive model is an equation
- Predicts the value of a criterion variable (dependent variable) based on the value of one or more predictor variables (independent variables)

- Predictive models, like descriptive models, allow exploring a problem space
- Predictive models deal with numbers, not concepts

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Why Use Predictive Models?

- Experiments only test a single point in a larger space
- Without a model it is hard to predict other points in that space
- Better
  - Understand the problem space in theory
  - Develop a predictive model: Predicts how the dependent variable varies with changes of the independent variables
- A predictive model may reduce the need for experiments
- Requires understanding of problem space
- Experiments can produce evidence in favor of or against the theory / model

Predictive Model Examples

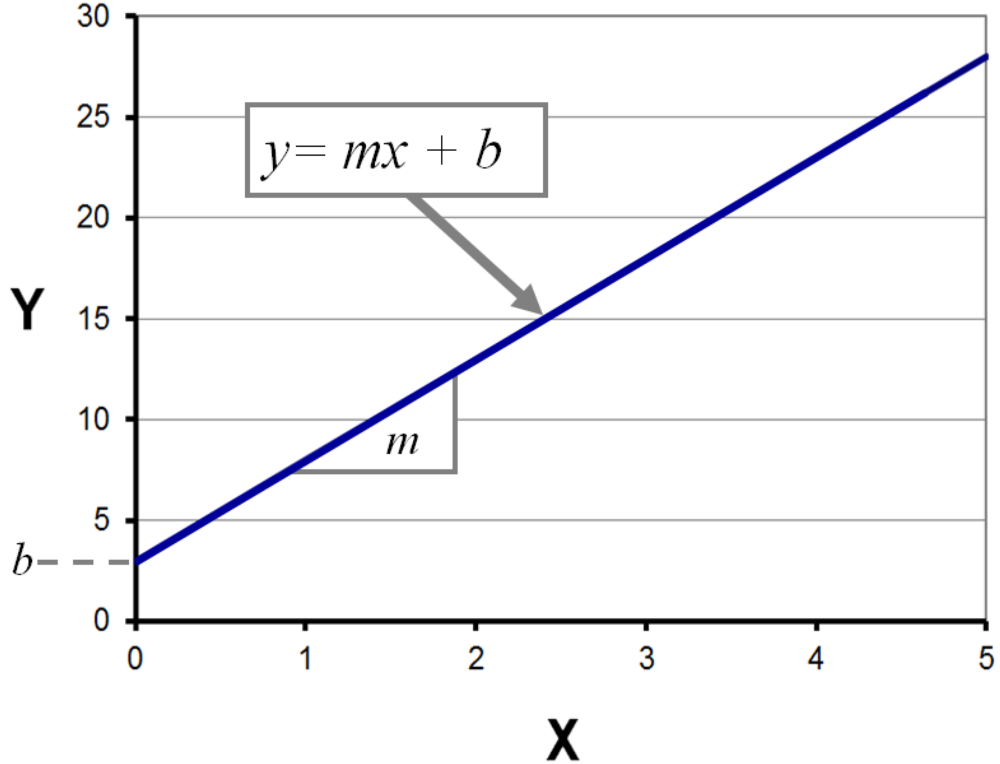
- Linear prediction equation
- Fitts' law
- Choice reaction time
- Keystroke-level model (KLM)
- Multiple linear regression

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Leibniz Universität Hannover

Linear Prediction Equation

- The basic prediction equation expresses a linear relationship between a predictor variable (x) and a criterion variable (y):



MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Linear Regression

- A linear prediction equation is found using linear regression
- Goal
  - Given a set of (x,y) sample points, find the coefficients m and b for the line that minimizes the squared distances of the points from the line
- Result
  - Prediction equation that yields the best estimate of y given x
- Assumption
  - Relationship is linear

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

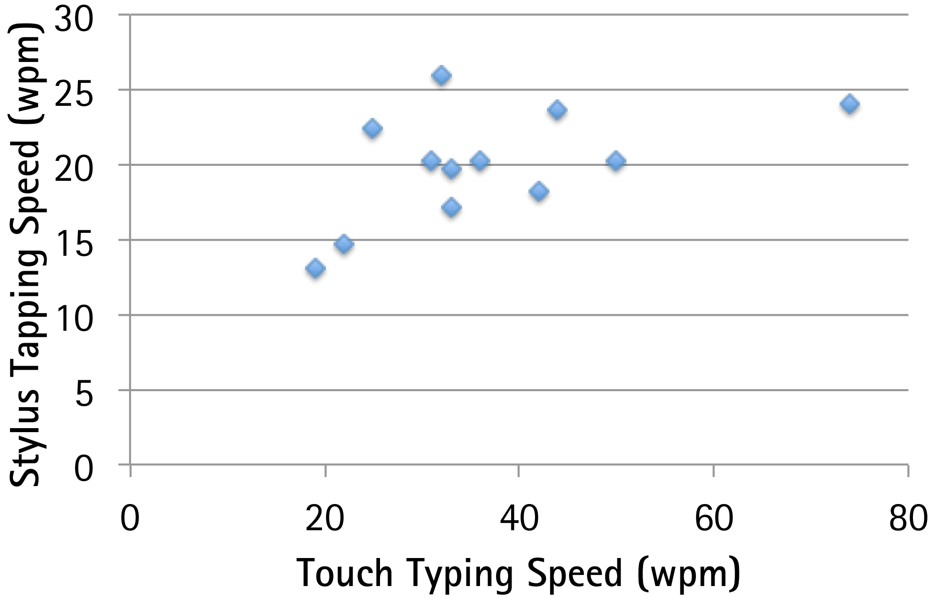
Example

- A research project investigated text entry on soft keyboards[1]
- Research question
  - Can stylus text entry speed be predicted from touch typing entry speed?
- Touch typing speed is the predictor variable
  - (x, measured in a pre-test)
- Stylus typing speed is the criterion variable
  - (y, measured experimentally)
- Data and scatter plot

[1] MacKenzie, I. S., & Zhang, S. X. (2001). An empirical investigation of the novice experience with soft keyboards. Behaviour & Information Technology, 20, 411-418.

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Data and Scatter Plot

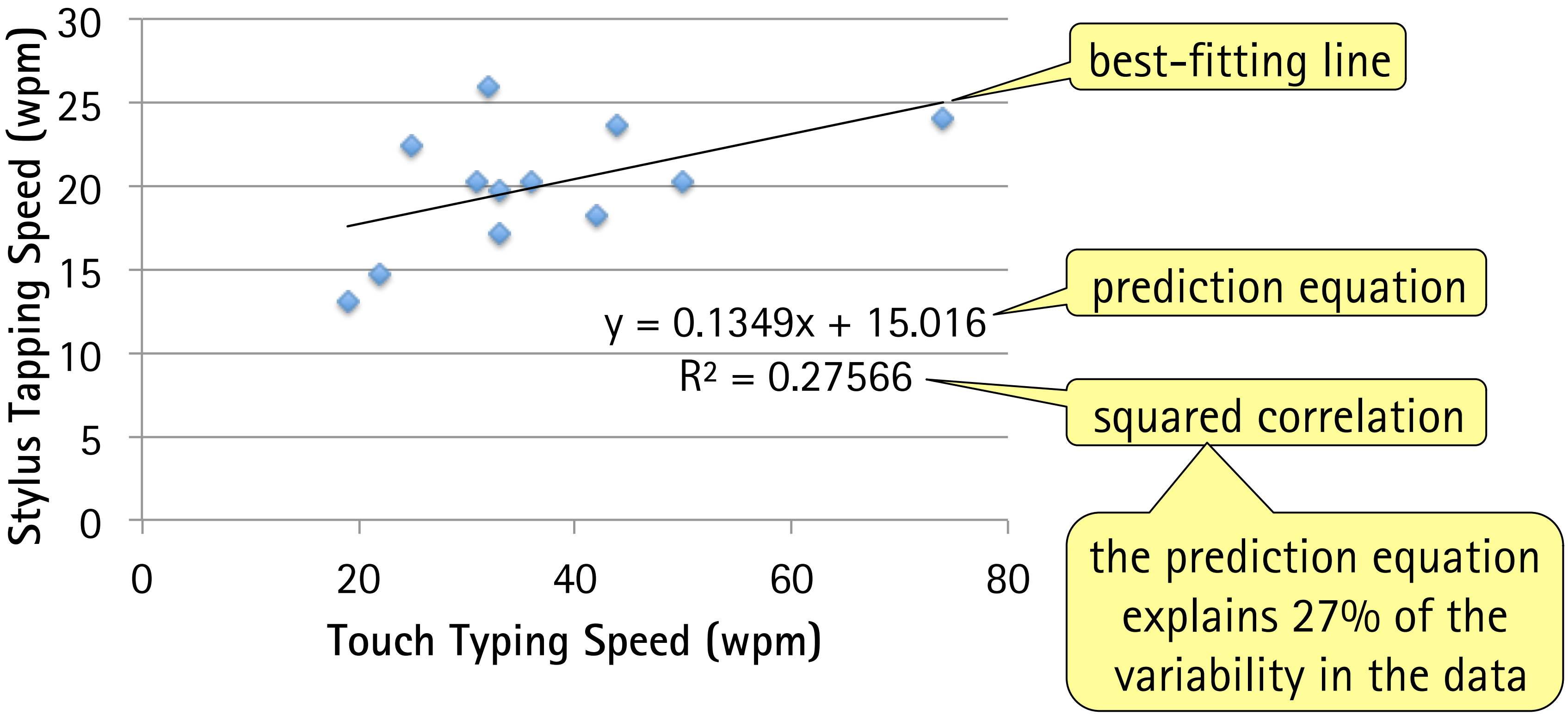


Participant	Stylus Tapping Speed (wpm)	Touch Typing Speed (wpm)
P1	18.2	42
P2	23.6	44
P3	26.0	32
P4	20.3	50
P5	20.3	36
P6	17.1	33
P7	24.0	74
P8	14.7	22
P9	20.3	31
P10	19.7	33
P11	22.4	25
P12	13.1	19

- There seems to be a relationship
  - Faster touch typists seem to be faster at stylus tapping
- Questions
  - What is the prediction equation?
  - How strong is the relationship?

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Prediction Equation



95% confidence interval for slope (via bootstrapping): [-0.048, 0.866]
mean slope (via bootstrapping): 0.511

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Predictive Model Examples

- Linear prediction equation
- Fitts' law
- Choice reaction time
- Keystroke-level model (KLM)
- Multiple linear regression

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Fitts' Law

- One of the most widely used models in HCI
  - Origins: Papers in experimental psychology (from 1954[1] and 1964[2])
- Model for rapid aimed movements
  - E.g., moving a cursor toward a target and selecting the target
- Applications of the model
  - Analyze and compare design alternatives
  - Determine if a device or technique "conforms to Fitts' law"

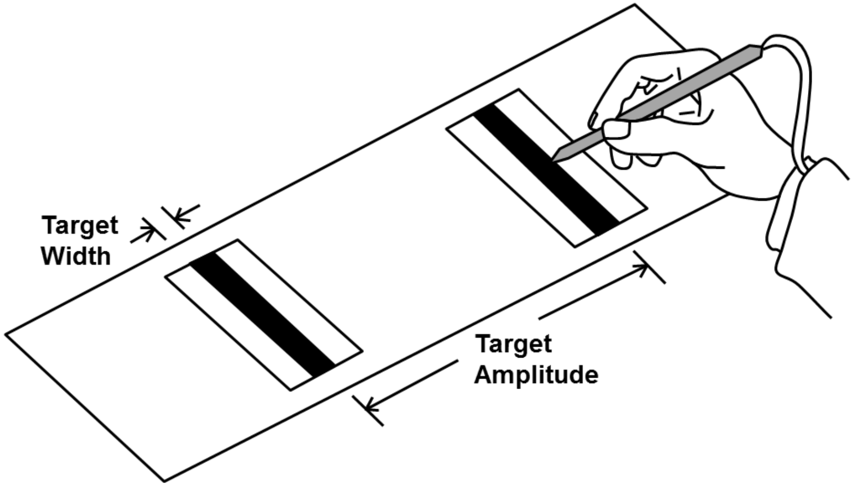
[1] Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47, 381-391.
[2] Fitts, P. M., & Peterson, J. R. (1964). Information capacity of discrete motor responses. Journal of Experimental Psychology, 67, 103-112.

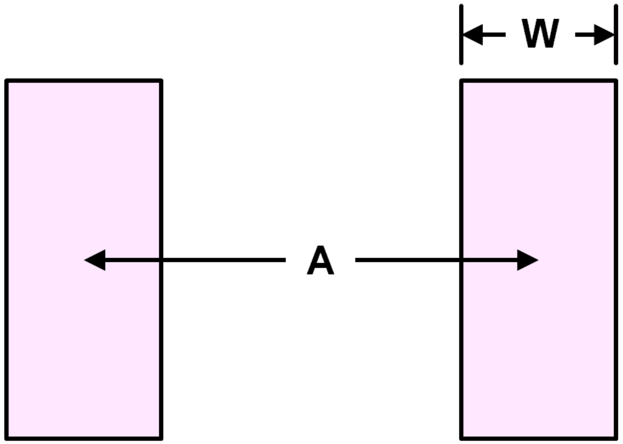
MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Fitts' Index of Difficulty (ID)

- Fitts' index of difficulty (ID) is a measure of the difficulty of a target selection task:

$$ID = \log_2\left(\frac{A}{W}+1\right)$$





- Fitts hypothesized that the relationship between movement time (MT) and ID is linear

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Fitts' Law Models for Pointing Devices

- Comparison of pointing devices[1]
- Twelve participants performed serial target selection tasks using these devices:

Interlink RemotePoint

Microsoft Mouse 2.0

[1] MacKenzie, I. S., & Jusoh, S. (2001). An evaluation of two input devices for remote pointing. Proceedings - EHCI 2000, 235-249, Heidelberg, Germany: Springer-Verlag.

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Experiment Conditions and Observations

Conditions			Mouse Observations				RemotePoint Observations			
			Mouse				RemotePoint			
A (pixels)	*W* (pixels)	*ID* (bits)	W_e (pixels)	ID_e (bits)	*MT* (ms)	*TP* (bits/s)	W_e (pixels)	ID_e (bits)	*MT* (ms)	*TP* (bits/s)
40	10	2.32	11.23	2.19	665	3.29	13.59	1.98	1587	1.25
40	20	1.58	19.46	1.61	501	3.21	21.66	1.51	1293	1.17
40	40	1.00	40.20	1.00	361	2.76	37.92	1.04	1001	1.04
80	10	3.17	10.28	3.13	762	4.11	10.08	3.16	1874	1.69
80	20	2.32	18.72	2.40	604	3.97	25.21	2.06	1442	1.43
80	40	1.58	35.67	1.70	481	3.53	37.75	1.64	1175	1.40
160	10	4.09	10.71	3.99	979	4.08	10.33	4.04	2353	1.72
160	20	3.17	21.04	3.11	823	3.77	19.09	3.23	1788	1.81
160	40	2.32	41.96	2.27	615	3.69	35.97	2.45	1480	1.65
		Mean	23.25	2.38	644	3.60	23.51	2.35	1555	1.46

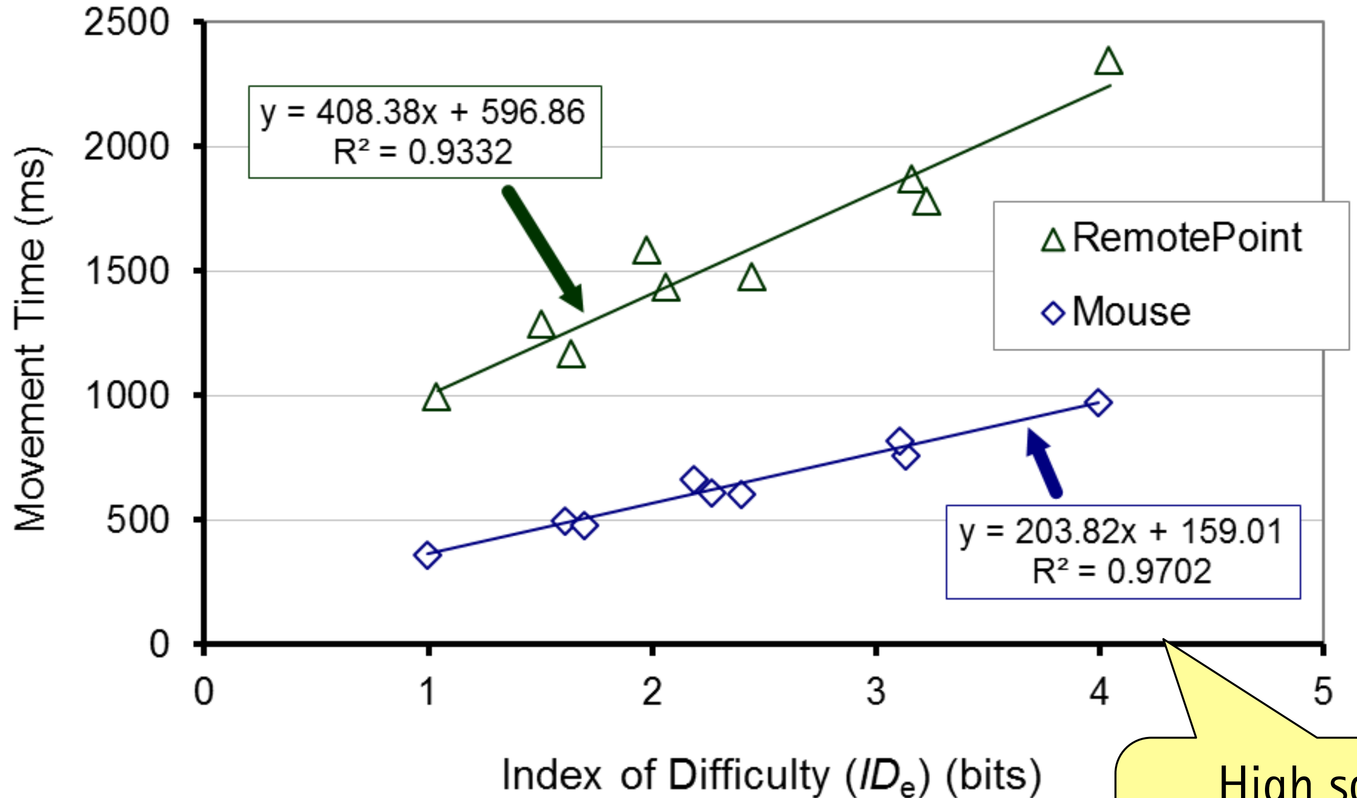
For model building...

x sample points
y sample points

x sample points
y sample points

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Fitts' Law Prediction Equations



High squared correlations; linear MT-ID relationship

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Predictive Model Examples

- Linear prediction equation
- Fitts' law
- Choice reaction time
- Keystroke-level model (KLM)
- Multiple linear regression

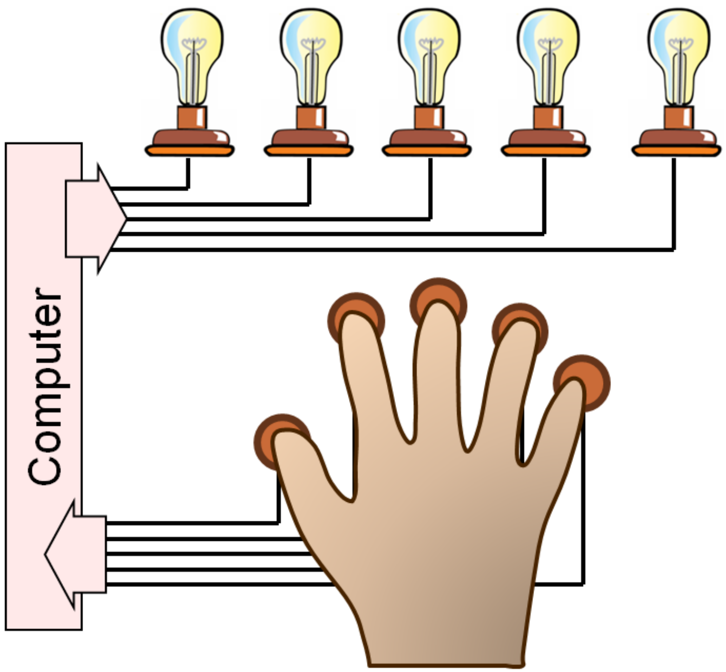
MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Reaction Times

- Simple reaction time: 1 → 1
  - One stimulus, one response
- Recognition reaction time: n → 1
  - n stimuli, only respond to one of them, ignore others
- Choice reaction time: n → n
  - n stimuli, n responses, answer with the right response

Choice Reaction Time

- Given n stimuli, associated with n responses, the time to react to a stimulus is the choice reaction time
- Modeled by the Hick-Hyman law:[1,2]
  - $RT = a + b \log_2(n + 1)$
- Coefficients:
  - $a \approx 200$ ms
  - $b \approx 150$ ms/bit
- An Information processing model
  - (like Fitts' law)



[1] Hick, W. E. (1952). On the rate of gain of information. Quarterly J Exp Psychol, 4, 11-36.
[2] Hyman, R. (1953). Stimulus information as a determinant of reaction time. J Exp Psychol, 45, 188-196.

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

HCI Applications of Choice Reaction Time

- Applications of choice reaction time
  - Telephone operator selects among ten buttons when a light behind a button turns on[1]
  - Time to select items in a hierarchical menu (visual search eliminated by practicing participants to expert levels)[2]
  - Activation time for mode switching with non-dominant hand in a tablet interface[3]
- Difficult to apply because additional behaviors are often present, such as visual search or movement

[1] Card, S. K., Moran, T. P., & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: Erlbaum.
[2] Landauer, T. K., & Nachbar, D. W. (1985). Selection from alphabetic and numeric menu trees using a touch screen: Breadth, depth, and width. Proc CHI '85, 73-77, ACM.
[3] Ruiz, J., Bunt, A., & Lank, E. (2008). A model of non-preferred hand mode switching. Proceedings of Graphics Interface 2008, 49-56, Toronto: Canadian Information Processing Society.

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Predictive Model Examples

- Linear prediction equation
- Fitts' law
- Choice reaction time
- Keystroke-level model (KLM)
- Multiple linear regression

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Keystroke-Level Model (KLM)[1,2]

- Developed for predicting human performance with interactive computing systems
- Predicts expert error-free task completion times
- Elements of a KLM prediction
  - Task (or a series of tasks)
  - Method used
  - Command language of the system
  - Motor skill parameters of the user
  - Response time parameters of the system

[1] Card, S. K., Moran, T. P., & Newell, A. (1980, July). The keystroke-level model for user performance time with interactive systems. Communications of the ACM, 23, 396-410.
[2] Card, S. K., Moran, T. P., & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: Erlbaum.

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Why Use the KLM?

- Consider a task such as "delete a file"
- Perhaps there are two ways to do the task
  - Mouse + menu selection
  - Keyboard + command entry
- The KLM can predict the time for each method
- If used at the design stage, design alternatives
  may be considered and compared without experiments

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

A KLM Prediction

- A task is broken into a series of subtasks
- Total predicted time is the sum of the operator times:
  $t_{EXECUTE} = n_K\, t_K + n_P\, t_P + n_H\, t_H + n_D\, t_D + n_M\, t_M + n_R\, t_R$
- Operators
  - K → keystroking P → pointing H → homing
  - D → drawing M → mental preparation R → system response
- Operator times are multiplied by number of occurrences in the task

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

KLM Operator Times

Operator	Description	Time (s)
K	PRESS A KEY OR BUTTON Pressing a modifier key (e.g., shift) counts as a separate operation, Time varies with typing skill: Best typist (135 wpm) Good typist (90 wpm) Average skilled typist (55 wpm) Average non-secretary typist (40 wpm) Typing random letters Typing complex codes Worst typist (unfamiliar with keyboard)	 0.08 0.12 0.20 0.28 0.50 0.75 1.20
P	POINT WITH A MOUSE Empirical value based on Fitts' law. Range from 0.8 to 1.5 seconds. Operator does *not* include the button click at the end of a pointing operation	1.10
H	HOME HAND(S) ON KEYBOARD OR OTHER DEVICE	0.40
$D(n_D.l_D)$	DRAW n_D STRAIGHT-LINE SEGMENTS OF TOTAL LENGTH l_D. Drawing with the mouse constrained to a grid.	$.9\ n_D + .16\ l_D$
M	MENTALLY PREPARE	1.35
R(t)	RESPONSE BY SYSTEM Different commands require different response times. Counted only if the user must wait.	t

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Heuristic Rules for Placing the M Operator

Begin with a method encoding that includes all physical operations and response operations. Use Rule 0 to place candidate Ms, and then cycle through Rules 1 to 4 for each M to see whether it should be deleted.

Rule 0. Insert Ms in front of all Ks that are not part of argument strings proper (e.g., text strings or numbers). Place Ms in front of all Ps that select commands (not arguments).

Rule 1. If an operator following an M is *fully anticipated* in the operator just previous to M, then delete the M (e.g., PMK → PK).

Rule 2. If a string of MKs *belong to a cognitive unit* (e.g., the name of a command), then delete all Ms but the first.

Rule 3. If a K is a *redundant terminator* (e.g., the terminator of a command immediately following the terminator of its argument), then delete the M in front of the K.

Rule 4. If a K *terminates a constant string* (e.g., a command name), then delete the M in front of the K; but if the K terminates a variable string (e.g., an argument string), then keep the M.

Card, Moran, Newell: The Keystroke-Level Model for User Performance Time with Interactive Systems. CACM July, 1980.

Original KLM Experiment

- The KLM was validated in an experiment with fourteen tasks performed using various methods and systems
- Example: Task T1 → Replace a 5-letter word with another word (one line from previous task)
- Using one system, POET, the task was broken down as follows:

Jump to next line	M K[LINEFEED]
Issue Substitute command	M K[S]
Type new word	K[word]
Terminate new word	M K[RETURN]
Type old word	K[word]
Terminate old word	M K[RETURN]
Terminate command	K[RETURN]

- 4 mental operations + 15 keystroking operations, hence →

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

KLM Prediction (Example)

$t_{EXECUTE} = 4\ t_M + 15\ t_K$

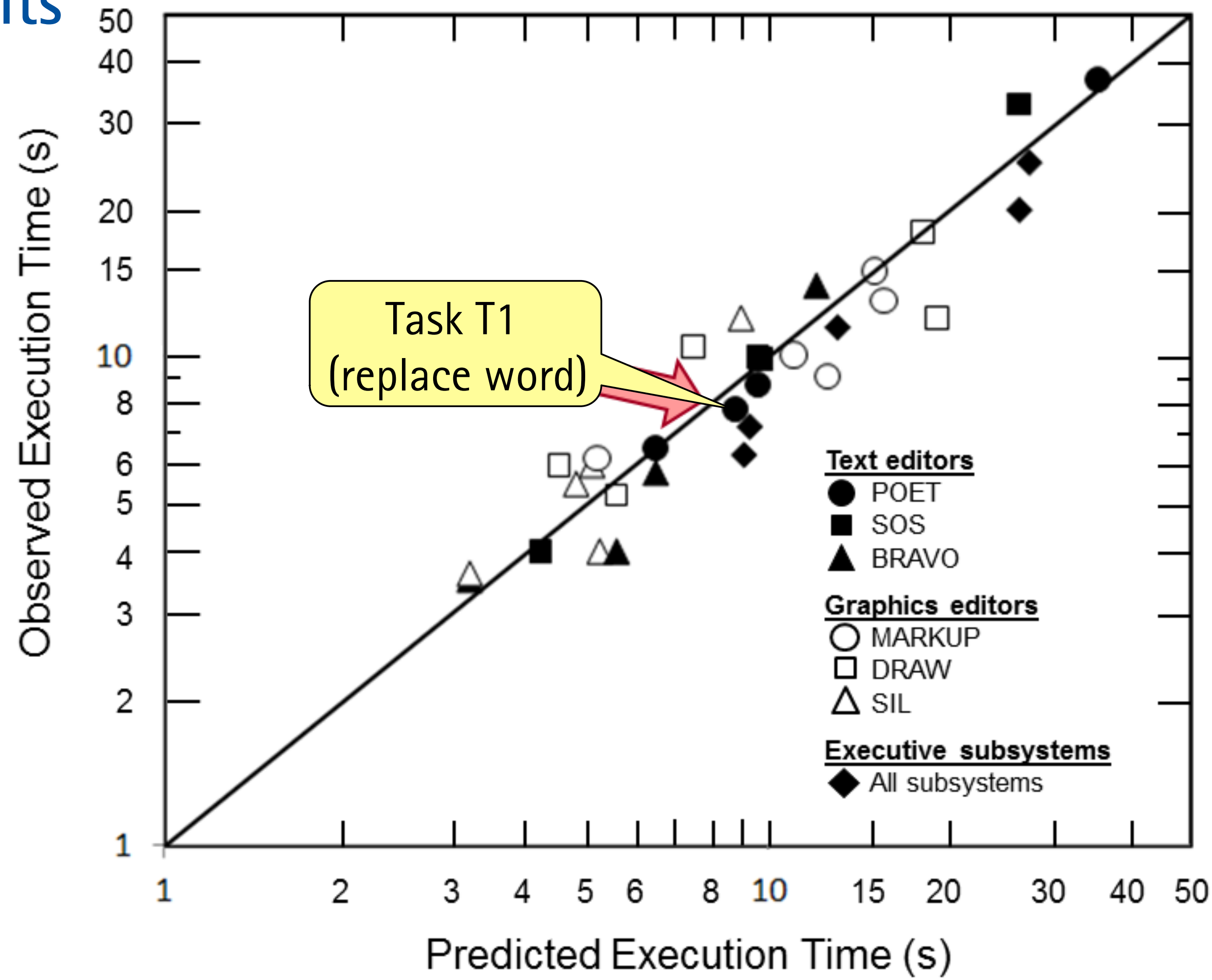
Jump to next line	M K[LINEFEED]
Issue Substitute command	M K[S]
Type new word	K[word]
Terminate new word	M K[RETURN]
Type old word	K[word]
Terminate old word	M K[RETURN]
Terminate command	K[RETURN]

- M set to 1.35 seconds (KLM operator times)
- K set to 0.23 seconds, based on a pre-test
- So: $t_{EXECUTE} = 4 * 1.35s + 15 * 0.23s = 8.85s$

- This is the prediction
- What about the observation?

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Results



MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Modern Applications

- Mouse interaction was just emerging when the KLM was introduced
- Update: Replace pointing constant (t_P) with a Fitts' law equation
  - Choose constants as appropriate for the device (e.g., mouse vs. touchpad)
  - and task (e.g., point-select vs. drag-select)
- For example, using the Fitts' law equation given earlier for the mouse

$$t_{\mathrm{P}} = 0.159 + 0.204 \times \log_2\left(\frac{A}{W} + 1\right)$$

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Pointing Operator – Update

- Example: Mouse point-select operation over 3.2 cm to click a 1.2 cm wide toolbar button:

$$t_{\mathrm{P}} = 0.159 + 0.204 \times \log_2\left(\frac{3.2}{1.2} + 1\right) = 0.45 \text{ seconds}$$

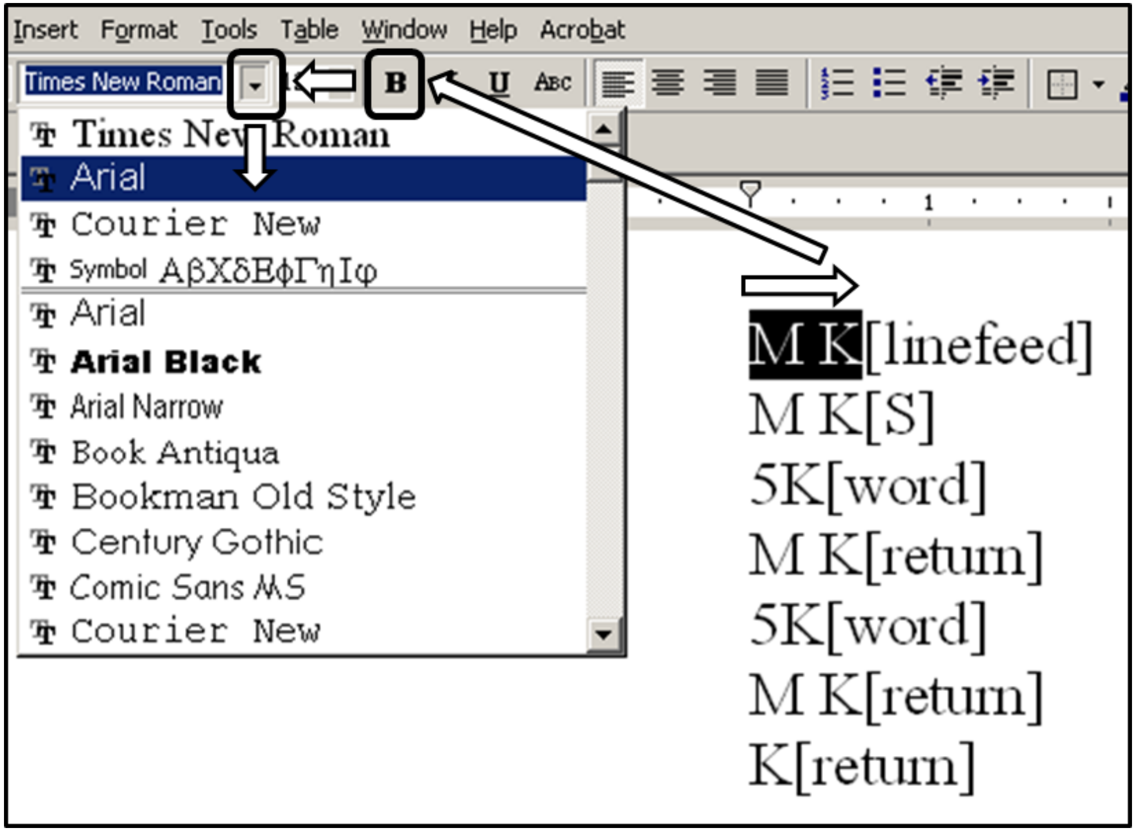
- Same task but moving the pointer 44.6 cm:

$$t_{\mathrm{P}} = 0.159 + 0.204 \times \log_2\left(\frac{44.6}{1.2} + 1\right) = 1.22 \text{ seconds}$$

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Pointing Operator – Example

- Develop KLM mouse and keyboard predictions for this GUI screen
- Task: Change the style and font for "M K" to bold, Arial
  - move left of "M K"
  - drag "M K"
  - move to B button
  - click B button
  - move to down button
  - click down button
  - move to Arial entry
  - click Arial entry



MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Mouse Analysis

- Operations

Mouse Subtasks	KLM Operators	t_P (s)
Drag across text to select "M K"	**M P**[2.5, 0.5]	0.686
Move pointer to Bold button and click	**M P**[13, 1]	0.936
Move pointer to Font drop-down button and click	**M P**[3.3, 1]	0.588
Move pointer down list to Arial and click	**M P**[2.2, 1]	0.501
	$\sum t_P =$	2.71

P[distance, width] here in characters and with K included

- Prediction

$$t_{\text{EXECUTE}} = 4 \times t_{\text{M}} + \sum t_{\text{P}} = 4 \times 1.35 + 2.71 = 8.11 \text{ seconds}$$

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Keyboard Analysis

- Operations

Keyboard Subtasks	KLM Operators
Select text	**M K**[shift] **3K**[→]
Convert to boldface	**M K**[ctrl] **K**[b]
Activate Format menu and enter Font sub-menu	**M K**[alt] **K**[o] **K**[f]
Type *a* ("Arial" appears at top of list)	**M K**[a]
Select "Arial"	**K**[↓] **K**[enter]

- Prediction

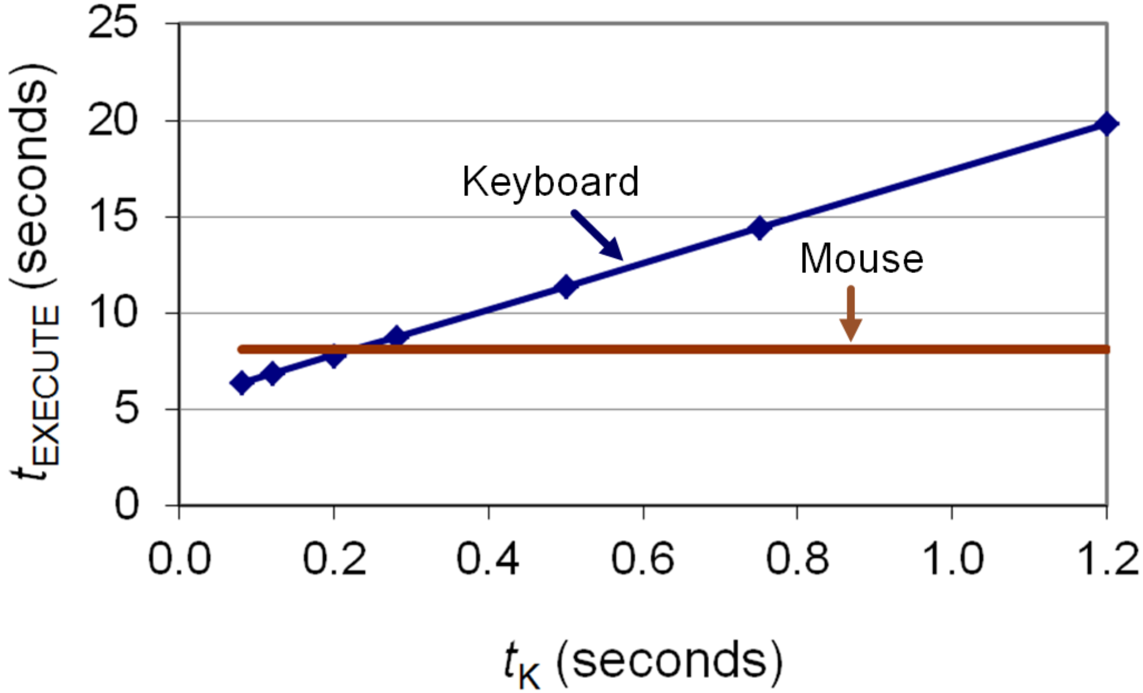
$$t_{\mathrm{EXECUTE}} = 4 \times t_{\mathrm{M}} + 12 \times t_{\mathrm{K}} = 4 \times 1.35 + 12 \times 0.75 = 14.40 \text{ seconds}$$

Use "typing complex codes"
(t_K = 0.75 s, KLM operator times)

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Sensitivity Analysis

- The keyboard prediction is sensitive to the parameter t_K, the keystroking time
- If t_K is allowed to vary, what is the effect on the predictions?



Implication: The mouse is faster than the keyboard, except for $t_k \leq 0.2$ s (which is unlikely)

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Contemporary Uses of the KLM

- The KLM continues to be widely used in HCI
  - Attention shifts with mobile phones
  - Stylus-based circling gestures
  - Managing folders and messages in email applications
  - Predictive text entry on mobile phones
  - Task switching in multi-monitor systems
  - Mode switching on tablet PCs
  - Distractions in in-vehicle information systems

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Validating KLMs

- When researchers build a KLM, an experiment may be conducted to validate the model
- What is the implication if observations ≠ predictions?
- Hinckley et al. note:[1]

"[The discrepancy] shows where the techniques deviate from the model, indicating the presence of hidden costs. These costs might include increased reaction time resulting from planning what to do next, mental pauses, or delays while the user attends to visual feedback after performing an action. Our methodology cannot attribute these costs to a specific cause. It just lets us deduce that a hidden cost must exist in a specific portion of the task. This is sufficient to generate many insights as to where the bottlenecks to performance lie and what parts of a technique might be improved."

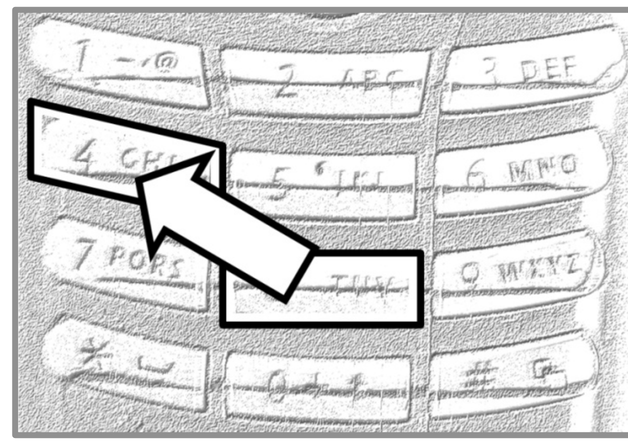
[1] Hinckley, Guimbretière, Baudisch, Sarin, Agrawala, Cutrell. The SpringBoard: Multiple modes in one spring-loaded control. CHI 2006.

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

KLM and Mobile Phone Keypad

For single finger or thumb input, replace keystroking operator K with pointing operator P[1]

$$t_{\mathrm{P(index\ finger)}} = 0.165 + 0.052 \log_2\left(\frac{A}{W} + 1\right)$$



$$t_{\mathrm{P(thumb)}} = 0.176 + 0.064 \log_2\left(\frac{A}{W} + 1\right)$$

[1] Silfverberg, M., MacKenzie, I. S., & Korhonen, P. (2000). Predicting text entry speed on mobile phones. Proc CHI 2000, 9-16, New York: ACM.

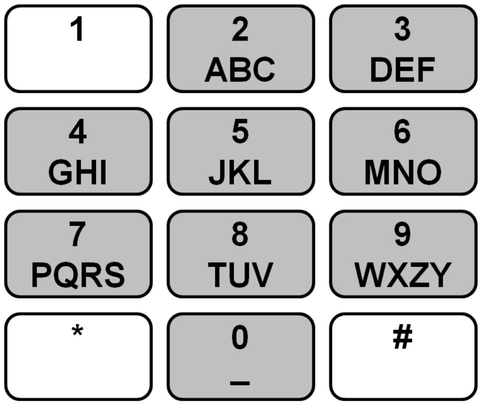
MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

KLM and Predictive Text Entry

- Interesting because of the combination of keystroking operations and mental operations
- Consider entering "beep" on a mobile phone keypad using predictive text entry (T9):

time

Keys	Display
2	a
23	be
233	bed
2337	beer
2337n	beds
2337nn	adds
2337nnn	bees
2337nnnn	beep
2337nnnn0	beep_



Mobile phone keypad

KLM model?

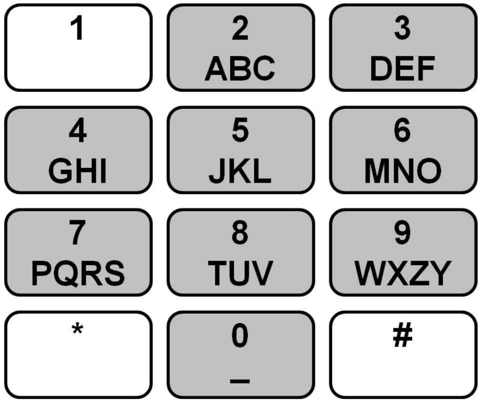
MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

KLM Operators for "beep"

- Perhaps the following…

2 3 3 7 M_P n M_P n M_P n M_P n M_P 0

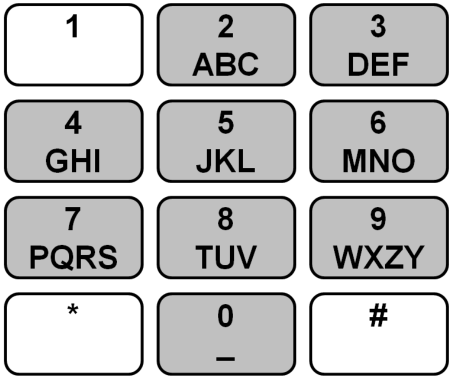
Mental operator M_P for "physical matching" (in this case, between the word on the display and the word in the user's mind)



- Where should mental operators be placed?
- What value should the mental operator assume?

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Expert Behavior



- Experts know the T9 key sequences for common words (i.e., no need for M_P at end of word)
  - the → 8430
  - of → 630
  - and → 2630
- But, how far down a word frequency list does this extend?
- What about ambiguous words?
  - if → 43n0
  - no → 66n0
  - beep → 2337nnnn0

"beep" is at rank 20767[1]. Even experts will likely require the M_P operator at the end of the word

[1] The word-frequency list used here is the 64,000 word list described by Silfverberg et al. (2000).

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Heuristics for M_P Operator

- It is not possible to precisely know when a user will hesitate to perform an M_P operation – a physical match between the word on the display and the word in the user's mind
- Two approaches for KLM modeling with M_P:
  - All-in → include M_P at every reasonable juncture
  - All-out → exclude all M_P operations
- The two approaches will produce upper bound (all-in) and lower bound (all-out) predictions

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

KLM and Visual Search

- Interfaces supporting word prediction or word completion typically provide the user with a list of choices ("candidates") as entry proceeds
- Instead of performing a "physical match", the user performs a "visual search"
- Visual search time is proportional to the number n of items to search
- How can visual search be included as a KLM mental operator?

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Mental Operator for Visual Search (M_V)

- Two scenarios using “vegetables” as an example

8 3 4 M_V W

Visual search after 3rd keystroke

8 M_V 3 M_V 4 M_V W

Visual search after each keystroke

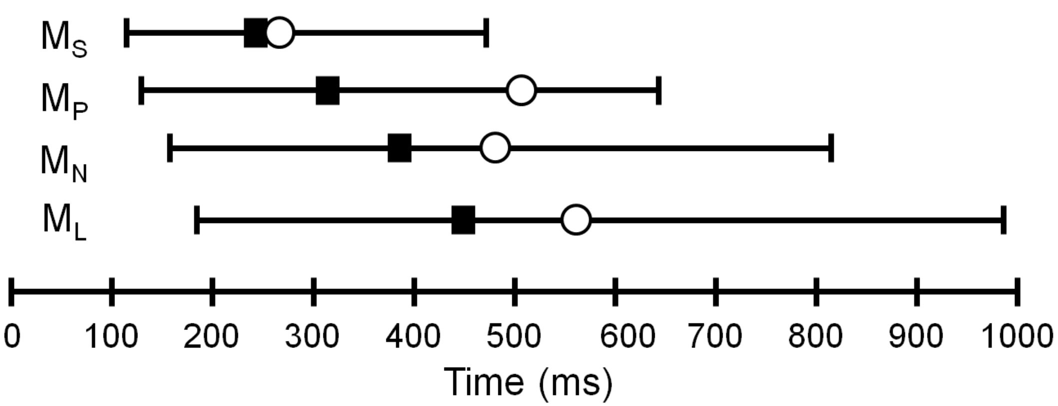
- Some considerations
  - A long candidate list bears a cost since it takes longer to visually scan but brings benefit since the desired word appears sooner
  - Viewing the candidate list after each keystroke adds time (cost) but allows entry of the intended word at the earliest opportunity (benefit)
  - There are different methods of selecting a candidate (e.g., direct selection, keystroke navigation)

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Updating the KLM's Mental Operator

Based on operators from Card et al. and on reaction time experiments

Proposed Mnemonic	Task	Execution Time (ms)	
		Card et al.	Figure 2-28 & Figure 2-30
M_S	Simple Reaction	240 [105 – 470]	277 [±44]
M_P	Physical Matching	310 [130 – 640]	510 [±59]
M_N	Name Matching	380 [155 – 810]	485 [±52]
M_L	Class Matching	450 [180 – 980]	566 [±96]
M_C	Choice Reaction	$200 + 150 \log_2(N + 1)$	
M_V	Visual Search		$498 + 41\,N$

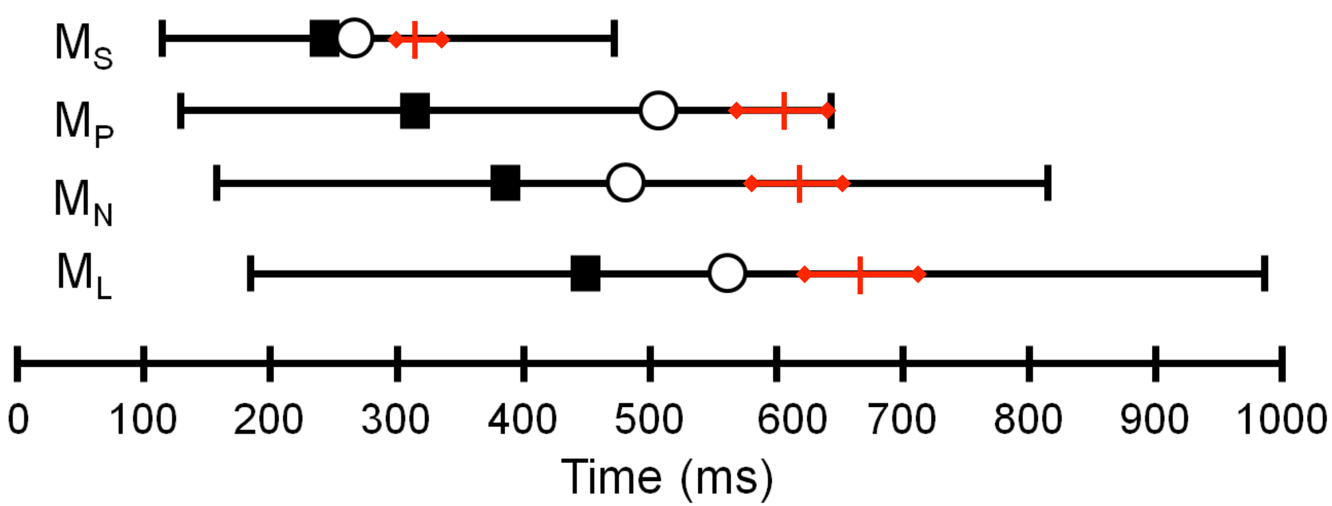


- data from Card et al.
- circles are mean times from experiments by MacKenzie

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Reaction Times from Class Experiment

- Means in sec, 95% confidence intervals from bootstrapping
  - Simple reaction: 312 [297, 330] s
  - Physical matching: 602 [569, 638] s
  - Name matching: 614 [579, 651] s
  - Class matching: 664 [622, 709] s



- Squares and black ranges from Card et al.
- Circles from experiments by MacKenzie
- Red diamonds and ranges from class experiment

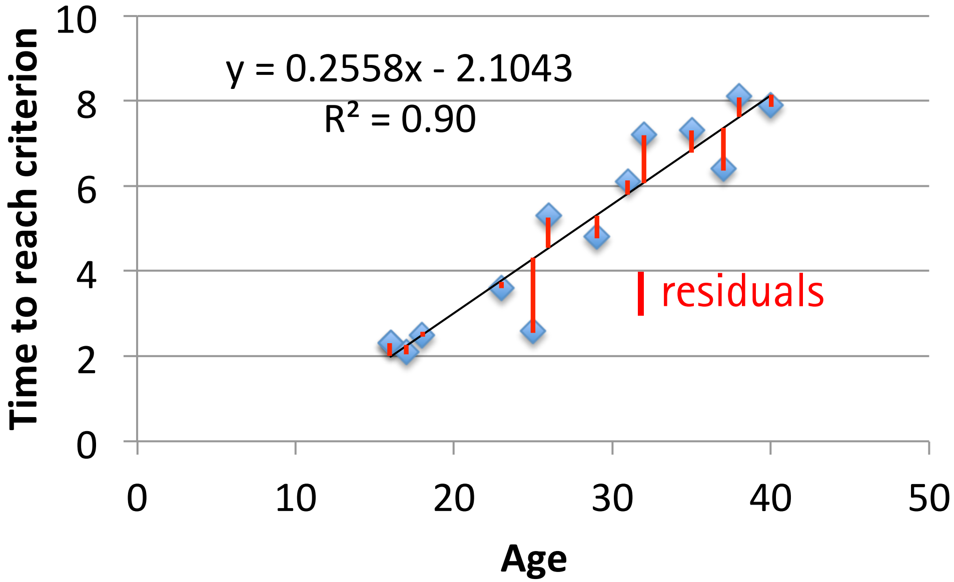
Predictive Model Examples

- Linear prediction equation
- Fitts' law
- Choice reaction time
- Keystroke-level model (KLM)
- Multiple linear regression

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Simple Linear Regression

- One independent variable, one dependent variable



- Residuals: Errors between observed and predicted values

$$\varepsilon_i = y_i - \hat{y}_i$$

- Regression equation minimizes sum of squared residuals

More Than One Predictor

- A prediction equation may have more than one predictor:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \ldots$$

- This technique is called multiple linear regression

- Conditions for linear regression
  - Linear relationship between the predictor variables and dependent variable
  - (Approximately) normal residuals
  - Constant variability of residuals
  - Independence of residuals

Hypothetical Example

- Research question
  - Is there a relationship between the time to learn a computer game and the age and computing habits of players?
- Dependent variable
  - y → the time to reach a criterion score
- Predictor variables
  - x1 → age of player
  - x2 → daily computer use in hours
- The experiment involved 14 participants
- Prediction equation
  - $y = b_0 + b_1x_1 + b_2x_2$ → how to find b_0, b_1, and b_2?

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Data

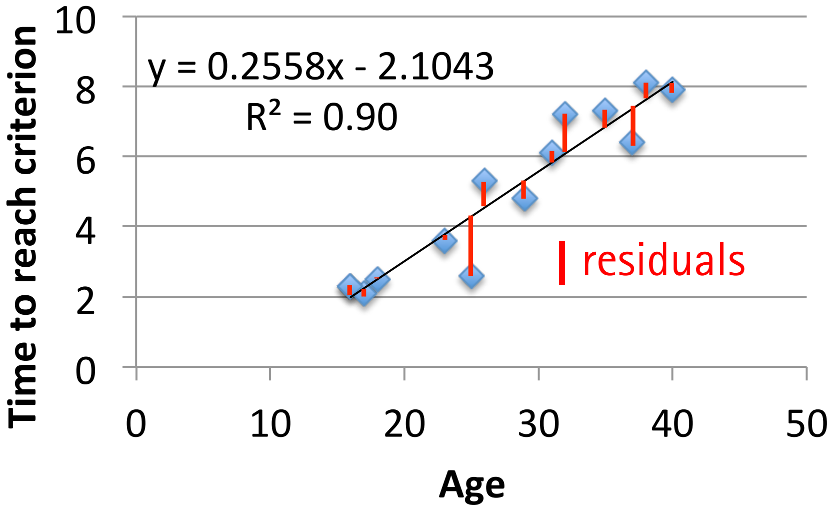
Participant	Time To Reach Criterion (hours)	Age (years)	Daily Computer Usage (hours)
P1	2.3	16	8
P2	2.1	17	7
P3	2.5	18	8
P4	3.6	23	6
P5	2.6	25	7
P6	5.3	26	5
P7	4.8	29	6
P8	6.1	31	4
P9	7.2	32	5
P10	7.3	35	3
P11	6.4	37	4
P12	8.1	38	2
P13	7.9	40	4
P14	2.3	16	8

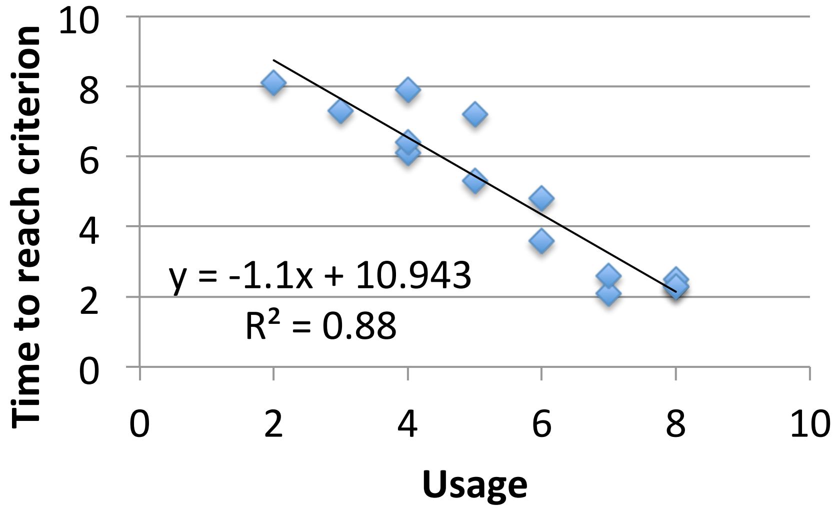
$$\text{time} = 3.3853 + 0.1533\ \text{age} - 0.4884\ \text{usage}, \quad R^2 = 0.93$$

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Simple Linear Regressions

- Based on each independent variable separately





- Residuals: Errors between observed and predicted values
- Regression equation minimizes sum of squared residuals

Prediction Equation

$$y = 3.3853 + 0.1533\ x_1 - 0.4884\ x_2 \qquad R^2 = 0.930$$

Example:

A 35 year-old player who uses a computer
5 hours per day is predicted to take

$$3.3853 + 0.1533 * 35 - 0.4884 * 5 = 6.3 \text{ hours}$$

to reach the criterion score.

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Multiple Linear Regression

- Equation of regression model

$$\begin{bmatrix} y_1 \\ \dots \\ y_{14} \end{bmatrix} = \begin{bmatrix} 1 & x_{1,1} & x_{1,2} \\ \dots & \dots & \dots \\ 1 & x_{14,1} & x_{14,2} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \dots \\ \varepsilon_{14} \end{bmatrix}$$

- Minimize sum of squared residuals

$$e(b_0, b_1, b_2) = \sum_{i=1}^{14} \varepsilon_i^2$$

- Compute partial derivatives w.r.t b_0, b_1, b_2
- Set derivatives to zero to find extrema
- Solve system of 3 linear equations

	y_i	$x_{i,1}$	$x_{i,2}$
Participant	Time To Reach Criterion (hours)	Age (years)	Daily Computer Usage (hours)
P1	2.3	16	8
P2	2.1	17	7
P3	2.5	18	8
P4	3.6	23	6
P5	2.6	25	7
P6	5.3	26	5
P7	4.8	29	6
P8	6.1	31	4
P9	7.2	32	5
P10	7.3	35	3
P11	6.4	37	4
P12	8.1	38	2
P13	7.9	40	4
P14	2.3	16	8

Excel Function Linear Estimate (LINEST)

- LINEST(ys, xs, TRUE, TRUE)
  - ys: Dependent variable (in example: time to reach criterion)
  - xs: Independent variables, columns (in example: age and daily usage)

	A	B	C	D
1	Participant	Time	Age	DailyUsage
2	1	2.3	16	8
3	2	2.1	17	7
4	3	2.5	18	8
5	4	3.6	23	6
6	5	2.6	25	7
7	6	5.3	26	5
8	7	4.8	29	6
9	8	6.1	31	4
10	9	7.2	32	5
11	10	7.3	35	3
12	11	6.4	37	4
13	12	8.1	38	2
14	13	7.9	40	4
15	14	2.3	16	8

=LINEST(B2:B15,C2:D15, TRUE, TRUE)

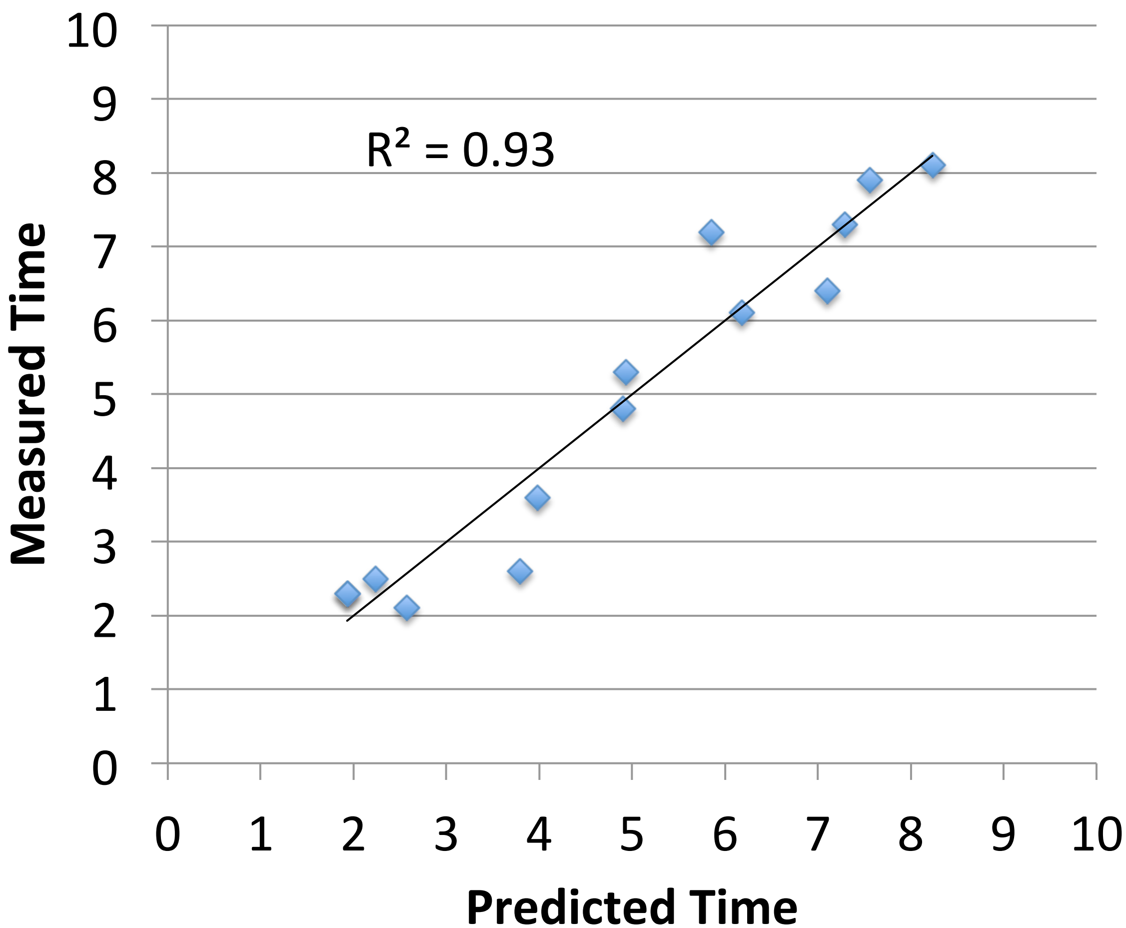
Result of LINEST:

-0.4884	0.1533	3.3853	b_2, b_1, b_0
0.2316	0.0531	2.6743	se(b_2), se(b_1), se(b_0)
0.9303	0.6574		R^2, se(y)
73.4570	11		F, df
63.4952	4.7541		$ss_{regression}$, $ss_{residual}$

se: standard error

df: deg. of freedom

Predicted vs. Measured Time



Participant	Measured	Predicted
1	2.3	1.93
2	2.1	2.57
3	2.5	2.24
4	3.6	3.98
5	2.6	3.80
6	5.3	4.93
7	4.8	4.90
8	6.1	6.18
9	7.2	5.85
10	7.3	7.29
11	6.4	7.10
12	8.1	8.23
13	7.9	7.56
14	2.3	1.93

$$\text{time} = 3.3853 + 0.1533\ \text{age} - 0.4884\ \text{usage}$$

Multiple Regression in NumPy

- Import modules

  ```
  import pandas as pd
  from statsmodels.formula.api import ols
  ```

  ols = ordinary least squares

- Load data (tab-separated values)

  ```
  data = pd.read_csv('data.txt', delimiter='\t')
  ```

- Predict Time from Age and Usage:

  ```
  result = ols('Time ~ Age + Usage', data).fit()
  ```

- Output b_0, b_1, b_2 in $y = b_0 + b_1x_1 + b_2x_2$

  ```
  b0 = result.params.Intercept
  b1 = result.params.Age
  b2 = result.params.Usage
  ```

data.txt:

Time	Age	Usage
2.3	16	8
2.1	17	7
2.5	18	8
3.6	23	6
2.6	25	7
5.3	26	5
4.8	29	6
6.1	31	4
7.2	32	5
7.3	35	3
6.4	37	4
8.1	38	2
7.9	40	4
2.3	16	8

Multiple Regression in NumPy

print result.params

```
Intercept     3.385330
Age           0.153292
Usage        -0.488381
```

Multiple Regression in NumPy

print result.summary()

```
 OLS Regression Results
==============================================================================
Dep. Variable:                   Time   R-squared:                       0.930
Model:                            OLS   Adj. R-squared:                  0.918
Method:                 Least Squares   F-statistic:                     73.46
Date:                                   Prob (F-statistic):           4.33e-07
Time:                                   Log-Likelihood:                -12.305
No. Observations:                  14   AIC:                             30.61
Df Residuals:                      11   BIC:                             32.53
Df Model:                           2
==============================================================================
                 coef    std err          t      P>|t|      [95.0% Conf. Int.]
------------------------------------------------------------------------------
Intercept      3.3853      2.674      1.266      0.232        -2.501     9.271
Age            0.1533      0.053      2.886      0.015         0.036     0.270
Usage         -0.4884      0.232     -2.109      0.059        -0.998     0.021
==============================================================================
```

Inference for Multiple Linear Regression

- Regarding the obtained data as a sample from a larger population
- Drawing conclusions given the sample w.r.t. the population

- Checking whether a predictor is significant

- Determining the confidence interval for a predictor

Interpretation of Confidence Intervals for the Example

- Age: $CI(b_2) = 0.1533 \pm 2.20 * 0.0531 = [0.036, 0.270]$
- We are 95% confident that for each additional year of age the time to reach the criterion is expected to increase by 0.036 to 0.270 hours (2.2 to 16.2 minutes).

- Daily usage $CI(b_2) = -0.4884 \pm 2.20 * 0.2316 = [-0.998, 0.021]$
- We are 95% confident that for each additional hour of average daily computer use the time to reach the criterion is expected to be -0.998 hours lower to 0.021 hours higher (59.9 minutes lower to 1.3 minutes higher).

Bootstrapping Prediction Equation

- Regard the obtained sample (size n) as a small but representative subset of the population
- Repeatedly (m times) take n random items from this original sample with replacement
  - A resample may contain an item more than once, or may omit an item
  - Each element of the original sample must have equal probability of being resampled
- Result is m resamples, each of size n
  - m is on the order of 1000 to 10000
- On each resample, compute the parameters of a linear model
- Store the parameters, sort them, take 2.5% and 97.5% percentiles as confidence intervals

Bootstrapping Prediction Equation, Example Script

```
N = len(data) # 14 data items in sample
M = 10000 # number of resamples
# indices from uniform distribution, N x M indices (N rows, M columns)
resampleIndices = np.random.randint(low=0, high=N, size=(N,M))

# M x 3 array for parameter values, initialized to 0 (M rows, 3 columns)
params = np.zeros((M,3)) # Intercept, Age, Usage

# compute least squares regressions M times
for i in xrange(M):
    indices = resampleIndices[:,i] # select i-th column (of length N)
    resample = data.loc[indices] # resample based on index vector
    lm = ols('Time ~ Age + Usage', resample).fit() # fit linear model
    params[i] = lm.params # store the result (parameters for intercept, age, usage)
```

Bootstrapping Prediction Equation, Example Script

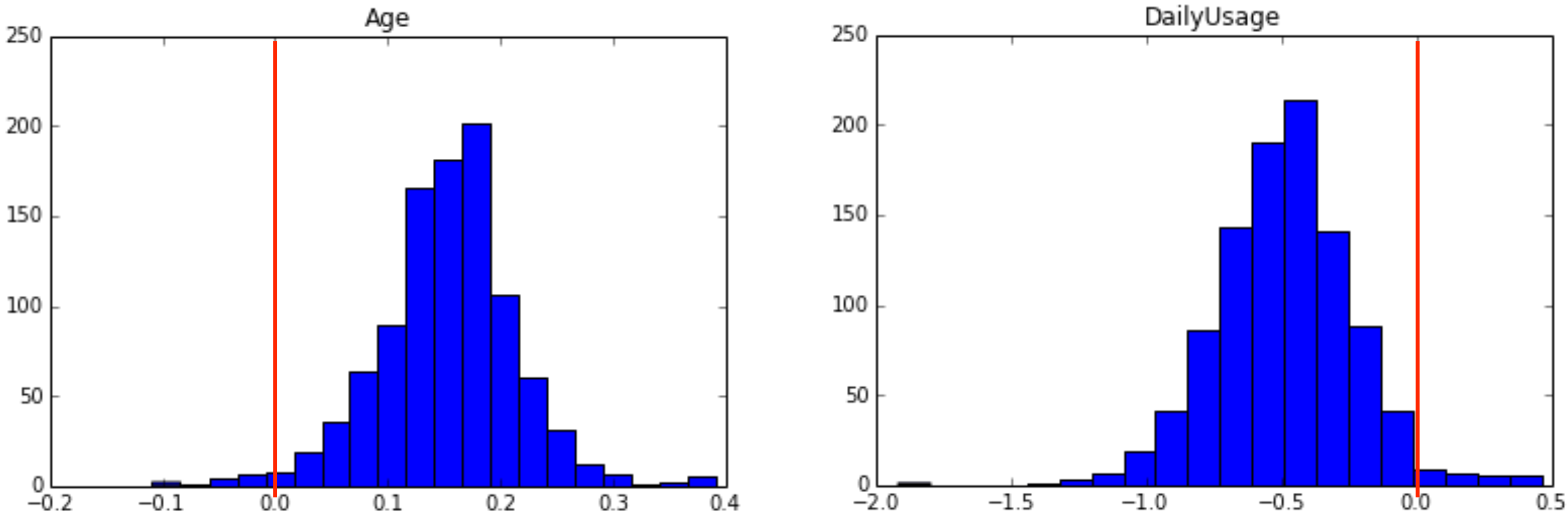
```
# sort parameter vectors, separately for age and usage
age = params[:,1] # column 2 (M elements)
usage = params[:, 2] # column 3 (M elements)
age.sort() # sort the parameter estimates
usage.sort() # sort the parameter estimates
# look up CI values at 2.5% and 97.5% percentiles
lowerIndex, upperIndex = 0.025 * M, 0.975 * M
print "95% CI for age:", age[lowerIndex], age[upperIndex]
print "95% CI for usage:", usage[lowerIndex], usage[upperIndex]
# draw histograms for age and usage
pp.figure()
pp.hist(age, bins = 20)
pp.title("Age")
```

```
pp.figure()
pp.hist(Usage, bins = 20)
pp.title("Usage")
```

Bootstrapping Prediction Equation, Confidence Intervals

95% CI from Bootstrap:
for age: [0.030, 0.268]
usage: [-0.969, -0.014]

cf. Statsmodels:
age: [0.036, 0.270]
usage: [-0.998, 0.021]



means for Intercept, Age, Usage: 3.453, 0.152, -0.495
medians for Intercept, Age, Usage: 3.384, 0.155, -0.496

cf. Statsmodels:
3.385, 0.153, -0.488

Example (Lag)[1]

Model for MT (ms)[a]	Fit[b]	Variance Explained
$MT = 435 + 190\ ID_e$	$r = .560$	31.30%
$MT = 894 + 46\ LAG$	$r = .630$	39.80%
$MT = -42 + 246\ ID_e + 3.4\ LAG$	$R = .948$	89.80%
$MT = 230 + (169 + 1.03\ LAG)\ ID_e$	$R = .987$	93.50%
[a] LAG in ms, IDe in bits		
[b] $n = 48$, $p < .0001$ for all models		

Models built using multiple regression

[1] MacKenzie, I. S., & Ware, C. (1993). Lag as a determinant of human performance in interactive systems. Proc INTERCHI '93, 488-493, New York: ACM.

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Social Science Example

- Often interested in observing and explaining human behaviour (rather than in measuring and predicting human performance)
- Example
  - A research project sought to determine the "probability of replying to an email message" (dependent variable)[1]
  - 12 predictor variables were tested using multiple linear regression
  - Some variables had a positive contribution
    - E.g., email with only one recipient
  - Some variables had a negative contribution
    - E.g., emails from close colleagues
  - Even with 12 variables, the correlation was modest: $R^2 = 0.37$

[1] Dabbish, L. A., Kraut, R. E., Fussell, S., & Kiesler, S. (2005). Understanding email use: Predicting action on a message. *Proc CHI 2005*, 691-700, New York: ACM.

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Stepwise Linear Regression

- Stepwise linear regression involves using multiple, and often many, predictor variables
- The variables are added or removed one at a time to determine which is the best model
- Goal: Model should explain as much of the variability as possible with as few predictors as possible
  - Model should be as simple as possible

MacKenzie: Human-Computer Interaction - An Empirical Research Perspective.

Adjusted R^2 for Stepwise Model Selection

- Use as few predictors as possible to get a simple model
- Adding more predictors automatically increases R^2, even if the predictors are not useful
- Adjusted R^2 penalizes the higher numbers of predictors

$$R^2_{adjusted} = 1 - \frac{SS_{residual}}{SS_{total}} \frac{n-1}{n-k-1} = 1 - \left(1 - R^2\right) \frac{n-1}{n-k-1}$$

  - n: number of observations
  - k: number of predictors (not counting the intercept)
- Adjusted R^2 does not increase if added predictor does not give new information

Stepwise Model Selection

- Criteria for good model
  - High adjusted R^2, low p-value, other criteria
- Backwards elimination
  - Initially, model contains all predictors
  - Weak predictors are dropped, one at a time (if resulting model improves)
  - Adj. R^2 criterion: Select resulting model with highest adj. R^2
  - p-value criterion: Drop insignificant predictors
- Forward selection
  - Initially, model contains no predictors
  - Add strongest predictors, one at a time (if resulting model improves)
  - Adj. R^2 criterion: Select resulting model with highest adj. R^2
  - p-value criterion: Add predictor with lowest significant p-value

Stepwise Model Selection in the Example

- Alternative 1: Backward elimination, adjusted R^2 criterion
  - Full model: Time ~ Age + Usage, Adj. $R^2 = 0.918$
  - Eliminate Usage: Time ~ Age, Adj. $R^2 = 0.894$
  - Eliminate Age: Time ~ Usage, Adj. $R^2 = 0.867$
  - No improvement through elimination, keep full model

- Alternative 2: Forward selection, adjusted R^2 criterion
  - Time ~ Age, Adj. $R^2 = 0.894$
  - Time ~ Usage, Adj. $R^2 = 0.867$
  - Age model has highest Adj. R^2, pick Age model, add Usage, result is
  - Time ~ Age + Usage, Adj. $R^2 = 0.918$
  - Adj. R^2 is better than Age alone, keep model, contains all predictors, stop