

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF TORONTO

CSC 428F/2514F

HUMAN-COMPUTER INTERACTION

Lecture 21

MODELS OF ROUTINE COGNITIVE SKILLS IN
INTERACTIVE COMPUTER USAGE (e.g., TEXT EDITING)

21.1 The GOMS Model(s) of Manuscript Editing.....	2
21.2 Nature of the GOMS Model.....	5
21.3 Applications of the GOMS Model.....	9
21.4 The Keystroke Level Model.....	10
21.5 Applications of the Keystroke Level Model.....	13
21.6 Extensions and Limitations (of Both Theories).....	17

Ronald Baecker
Professor of Computer Science,
Electrical and Computer Engineering, and Management
University of Toronto

Copyright © 1987-97, Ronald Baecker.
All rights reserved.

21.1 The GOMS Model(s) of Manuscript Editing

Goal is to describe user's cognitive structure in order to predict:
Methods, sequences of operators that are used
Length of time to carry out task

Limitations

Doesn't deal well with errors
Doesn't deal with problem solving behaviour

G: Goals to be achieved

O: Operators, skilled actions

(elementary perceptual/cognitive/motor acts)

e.g., GET NEXT TASK
USE-LF-METHOD
USE-S-COMMAND
VERIFY-EDIT

Each operator has specific output and duration

Operators define grains of analysis (can be at various levels, e.g., "type a command," individual keystroke, etc.)

M: Methods, procedures for accomplishing goals

Conditional sequence of goals and operators, with conditional tests on contents of users' intermediate memory and state of task environment

e.g., GOAL: ACQUIRE-UNIT-TASK
GET-NEXT-PAGE if at end of page
GET-NEXT-TASK

Methods relatively certain of success in routine cognitive skills, relative to problem solving, although errors occur
Methods are learned procedures

S: Selection rules (control structures) for choosing among available competing methods for achieving goals (Fig. 21.3)

Example: manuscript editing w. line-oriented editor (Fig. 21.1)

*Figure 21.1: Sample GOMS Model for Manuscript Editing Task
(CMN, 1983, p. 142; BB, pp. 221-222)*

Step-by-step behaviour of a GOMS model (Figure 21.2)

*Figure 21.2: Trace Model of M4B during Performance of a Unit Task
(CMN, 1983, p. 143; BB, p. 222)*

Step	Contents of Goal Stack	Operator Executed	External User Action
1	ED-MS		
2	ED-MS, ED-UT		
3	ED-MS, ED-UT, ACQ-UT		
4	ED-MS, ED-UT, ACQ-UT	GET-NEXT-TASK	Looks at manuscript
5	ED-MS, ED-UT		
6	ED-MS, ED-UT, EX-UT		
7	ED-MS, ED-UT, EX-UT, LOC-LINE		
8	ED-MS, ED-UT, EX-UT, LOC-LINE	USE-LF-METHOD	Types LINEFEED
9	ED-MS, ED-UT, EX-UT		
10	ED-MS, ED-UT, EX-UT, MOD-TEXT		
11	ED-MS, ED-UT, EX-UT, MOD-TEXT	USE-S-COMMAND	Types sldi RETURN idi RETURN RETURN
12	ED-MS, ED-UT, EX-UT, MOD-TEXT	VERIFY-EDIT	Types /
13	ED-MS, ED-UT, EX-UT		
14	ED-MS, ED-UT		
15	ED-MS		

Figure 21.3: Selection Rules for LOCATE-LINE Goal, Experiment 5A
(CMN, 1983, p. 153; BB, p. 226)

User	Rule	This Rule		Cumulative		
		Gain	Loss	Hits	Misses	%Hits
S1 (M)	Rule 1: Use the OS-METHOD unless another rule applies.	44	0	44	24	65%
	Rule 2: If $d < 3$, use the LF-METHOD.	11	2	53	15	78%
	Rule 3: If the target line is the last line of the page, use the AN-METHOD (with \$).	5	0	58	10	85%
	Rule 4: If the current method is to use paragraph numbers for search strings and the target line is near a paragraph number, use the AN-METHOD.	2	0	60	8	88%
S4 (M)	Rule 1: Use the +N-METHOD unless another rule applies.	51	0	51	15	77%
	Rule 2: If $d < 3$, use the LF-METHOD.	12	1	62	4	94%
S4 (SP)	Rule 1: Use the LF-METHOD unless another rule applies.	45	0	45	21	68%
	Rule 2: If $d > 9$, use the +N-METHOD.	16	12	49	17	74%
	Rule 3: If the target line is on the next page of the manuscript, use the LF-METHOD.	56	10	56	10	85%
S22 (SP)	Rule 1: Use the OS-METHOD unless another rule applies.	40	0	40	25	62%
	Rule 2: If $d < 5$, use the LF-METHOD.	22	2	60	5	92%
Average Final %Hits = 90%						

21.2 Nature of the GOMS Model

Can derive a hierarchically nested family of models
(Figs. 21.4, 21.5, 21.6) at various levels

Unit-task level
Functional level
Argument level
Keystroke level

Figure 21.4: Graph of the Family Tree of GOMS Models Investigated for the POET text editor (CMN, 1983, p. 163)

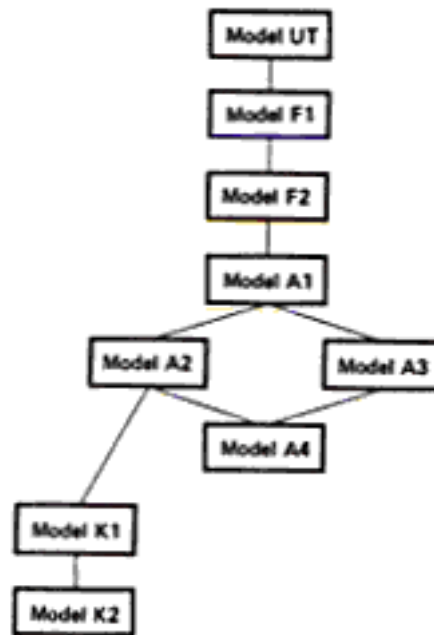


Figure 21.5: Description of Family of GOMS Models Tested
(CMN, 1983, p. 162; BB, p. 230)

UNIT-TASK LEVEL:

Model UT Constant time per unit task. Only one operator: EDIT-UNIT-TASK. (This model is like the Constant Time per Modification model of Chapter 4, except for the substitution of unit tasks for modifications.)

FUNCTIONAL LEVEL:

Model F1 Single operator for each functional step in unit task sequence: GET-NEXT-TASK, LOCATE-LINE, MODIFY-TEXT, VERIFY-EDIT.

Model F2 Like Model F1, but with operators LOCATE-LINE and MODIFY-TEXT broken into separate cases based on the methods used to accomplish them.

ARGUMENT LEVEL:

Model A1 Like Model F2, but with operators at the level of typing a system command (SPECIFY-COMMAND) or typing an argument to a command (SPECIFY-ARG).

Model A2 Like Model A1, but with SPECIFY-COMMAND and SPECIFY-ARG broken into separate cases according to whether they involve an implicit need to get information from manuscript (suffix = /g) or not (suffix = /ng).

Model A3 Like Model A1, but with SPECIFY-COMMAND and SPECIFY-ARG broken into separate cases according to four method contexts: quoted string method (suffix = /q), first argument to Substitute command (suffix = /s1), second argument to Substitute command (suffix = /s2), or Modify command (suffix = /w).

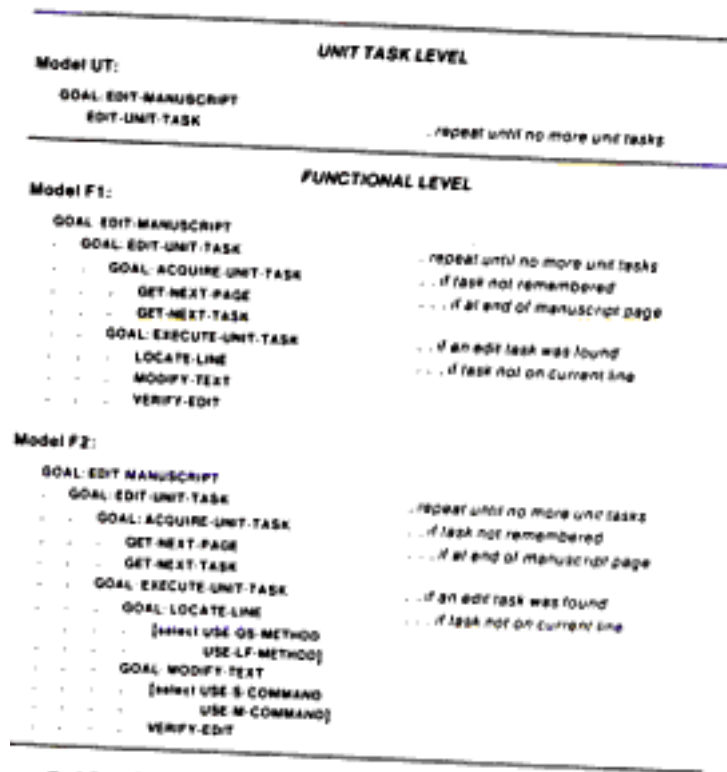
Model A4 Like Model A1, but with all the distinctions in both Model A2 and Model A3 combined multiplicatively.

KEYSTROKE LEVEL:

Model K1 Like Model A2, but with operators at the level of basic perceptual, cognitive, and motor actions: LOOK-AT, HOME, TURN-PAGE, TYPE, and MOVE-HAND. All mental actions not overlapped with motor operations are represented as the MENTAL operator.

Model K2 Like Model K1, but with MENTAL broken down into SEARCH-FOR, COMPARE, CHOOSE-COMMAND, and CHOOSE-ARG.

Figure 21.6ab: GOMS Models of POET
 CMN, 1983, pp. 164-166; BB, pp. 231-232)



	ARGUMENT LEVEL	
Model A1:		
GOAL: EDIT MANUSCRIPT		
GOAL: EDIT UNIT TASK		... repeat until no more unit tasks
GOAL: ACQUIRE UNIT TASK		... if task not remembered
GET NEXT PAGE		... if at end of manuscript page
GET NEXT TASK		
GOAL: EXECUTE UNIT TASK		... if an edit task was found
GOAL: LOCATE LINE		... if task not on current line
[select GOAL: USE QS METHOD		
SPECIFY COMMAND		
SPECIFY ARG		
GOAL: USE LF METHOD		
SPECIFY COMMAND]		... repeat until at line
VERIFY LOC		
GOAL: MODIFY TEXT		
[select GOAL: USE S COMMAND		
SPECIFY COMMAND		
SPECIFY ARG		
SPECIFY ARG		
GOAL: USE M COMMAND		
SPECIFY COMMAND		... repeat until at next
SPECIFY ARG		
SPECIFY ARG]		
VERIFY EDIT		

Model A2: as in Model A1 but substitute

SPECIFY COMMAND/G or SPECIFY COMMAND/NG for SPECIFY COMMAND

SPECIFY ARG/G or SPECIFY ARG/NG for SPECIFY ARG

Model A3: as in Model A1 but substitute

SPECIFY ARG/G or SPECIFY ARG/W or

SPECIFY ARG/S1 or SPECIFY ARG/S2 for SPECIFY ARG

Model A4: as in Model A1 but substitute

SPECIFY COMMAND/G or SPECIFY COMMAND/NG for SPECIFY COMMAND

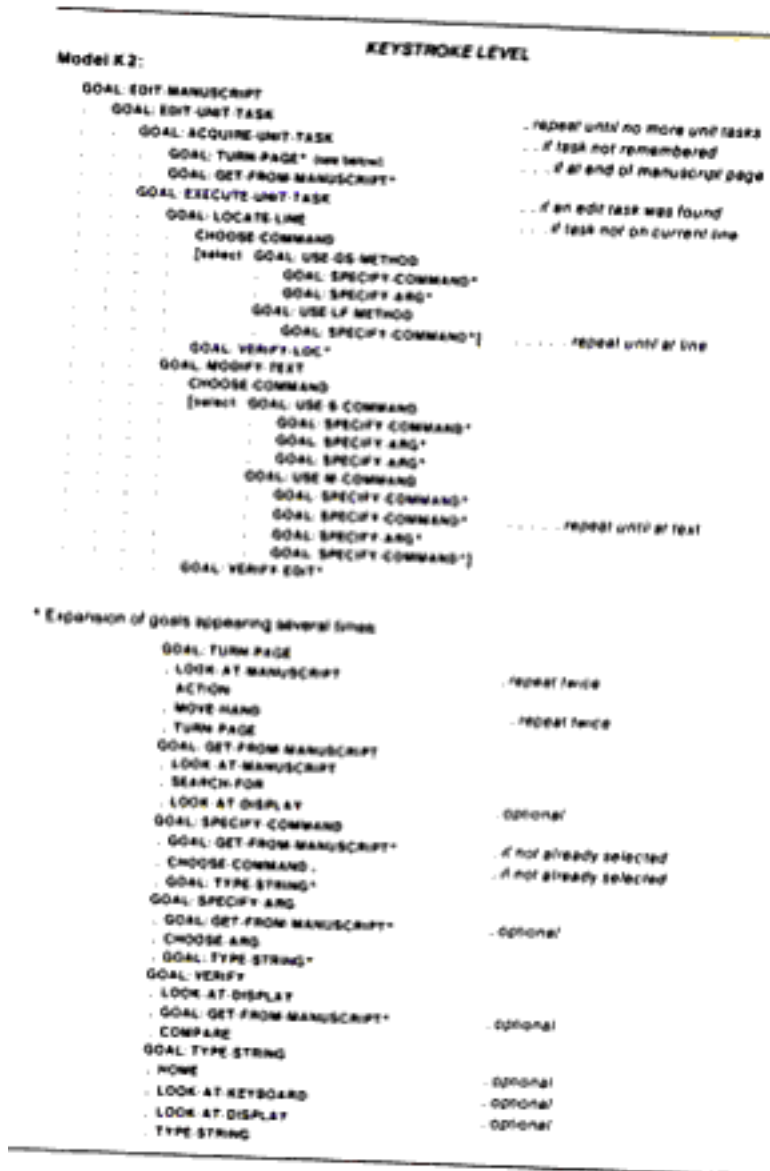
SPECIFY ARG/G/G or SPECIFY ARG/G/NG or

SPECIFY ARG/W/G or SPECIFY ARG/W/NG or

SPECIFY ARG/S1/G or SPECIFY ARG/S1/NG or

SPECIFY ARG/S2/G or SPECIFY ARG/S2/NG for SPECIFY ARG

Figure 21.6c: GOMS Models of POET continued
(CMN, 1983, pp. 164-166; BB, pp. 231-232)



21.3 Applications of the GOMS Model

Predicting task times (Fig. 21.7)

(RMS error = 33% of mean observed time)

Predicting sequence in which operators occurred (Fig. 21.8)

(Average 88%, varied from 79% to 98%)

Note grain of analysis, how do predictions depend on this?

Unit-task level

Functional level

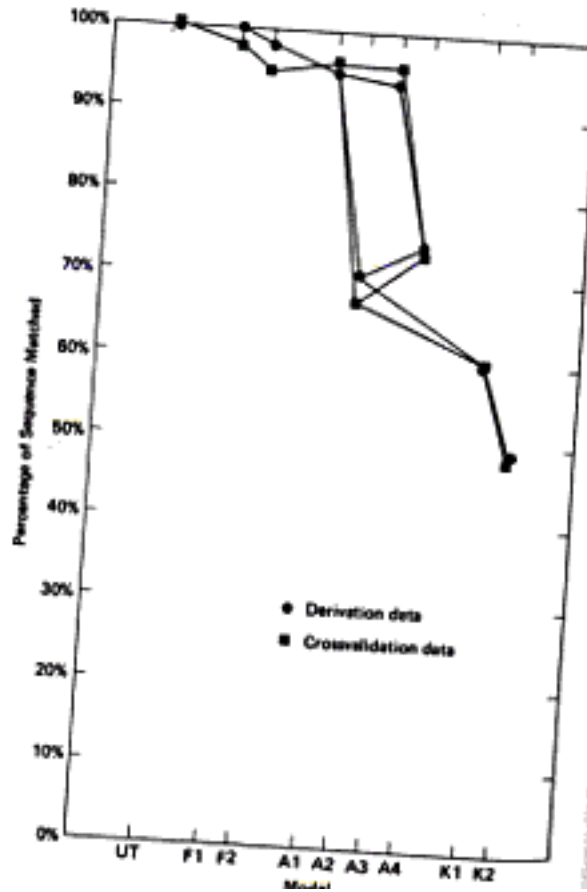
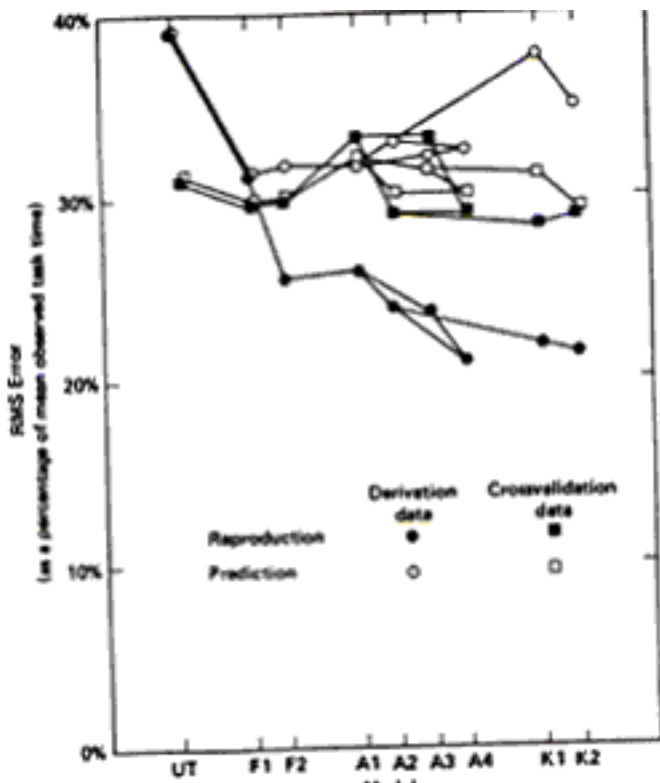
Argument level

Keystroke level

The ability to work at various grains of analysis — very powerful result!!!!

Figure 21.7: Accuracy of Time/Task Predictions, Experiment 5C (left below)
(CMN, 1983, p. 176; BB, p. 235)

Figure 21.8: Accuracy of Operator Sequence Predictions, Expt 5C (right below)
(CMN, 1983, p. 172; BB, p. 233)



21.4 The Keystroke Level Model

Card, Moran, & Newell, 1980; BB, pp. 192-206

Keystroke level GOMS model operators

MENTAL	SEARCH-FOR
TYPE	SearchFor + ChooseMethod
LOOK-AT	CHOOSE-COMMAND
HOME	CHOOSE-ARG
TURN-PAGE	COMPARE
MOVE-HAND	Compare + ChooseCommand
ACTION	EXPRESSION

A rather strange set

Recast the model, staying at this level, assuming method is known \Rightarrow *The Keystroke-Level Model*

Given:

- A task (possible involving subtasks)
- The command language of the system
- The motor skill parameters of the user
- The response time parameters of the system
- The method used for the task

Predict:

The *time* an *expert user* will take to execute the *task* using the system, provided he uses the *method* without error

$$T_{execute} = T_K + T_P + T_H + T_D + T_M + T_R$$

K = Keystroke, P = Pointing, H = Homing, D = Drawing,
M = Mental, R = Response

(See Figs. 21.9, 21.10)

Note: Much cleaner formulation than Keystroke Level GOMS

Figure 21.9: *The Operators of the Keystroke Level Model*
 CMN, 1983, p. 264; BB, p. 195; BGBG, p. 592)

Operator	Description and Remarks	Time (sec)
K	PRESS KEY OR BUTTON.	
	Pressing the SHIFT or CONTROL key counts as a separate K operation. Time varies with the typing skill of the user; the following shows the range of typical values:	
	Best typist (135 wpm)	.08
	Good typist (90 wpm)	.12
	Average skilled typist (55 wpm)	.20
	Average non-secretary typist (40 wpm)	.28
	Typing random letters	.50
Typing complex codes	.75	
Worst typist (unfamiliar with keyboard)	1.20	
P	POINT WITH MOUSE TO TARGET ON A DISPLAY. The time to point varies with distance and target size according to Fitts's Law, ranging from .8 to 1.5 sec, with 1.1 being an average. This operator does not include the (.2 sec) button press that often follows. Mouse pointing time is also a good estimate for other efficient analogue pointing devices, such as joysticks (see Chapter 7).	1.10
H	HOME HAND(S) ON KEYBOARD OR OTHER DEVICE.	.40
$D(a_D, l_D)$	DRAW a_D STRAIGHT-LINE SEGMENTS OF TOTAL LENGTH l_D CM. This is a very restricted operator; it assumes that drawing is done with the mouse on a system that constrains all lines to fall on a square .56 cm grid. Users vary in their drawing skill; the time given is an average value.	$.9a_D + .16l_D$
M	MENTALLY PREPARE.	1.35
$R(i)$	RESPONSE BY SYSTEM. Different commands require different response times. The response time is counted only if it causes the user to wait.	i

Figure 21.10: Sample Keystroke Level Model Calculations, for the text-editing task of replacing a 5-letter word with another 5-letter word, where this replacement takes place one line below the previous modification (CMN, 1983, p. 266; BB, p. 196)

Method for Task T1-POET:	
Jump to next line	MK[LINEFEED]
Issue Substitute command	MK[S]
Type new 5-letter word	5K[word]
Terminate new word	MK[RETURN]
Type old 5-letter word	5K[word]
Terminate old word	MK[RETURN]
Terminate command	K[RETURN].

Using the operator times from Figure 8.1, and assuming the user is an average skilled typist ($t_K = .2$ sec), we could predict the time it will take to execute this method:

$$T_{execute} = 4t_M + 15t_K = 8.4 \text{ sec.}$$

This method could be compared to the method for executing task T1 on the display-based system BRAVO:

Method for Task T1-BRAVO:	
Reach for mouse	H[mouse]
Point to word	P[word]
Select word	K[YELLOW]
Home on keyboard	H[keyboard]
Issue Replace command	MK[R]
Type new 5-letter word	5K[word]
Terminate type-in	MK[ESC]

$$T_{execute} = 2t_M + 8t_K + 2t_H + t_P = 6.2 \text{ sec.}$$

21.5 Applications of the Keystroke Level Model

Comparison of text-editing systems on various tasks
 11 systems (Figure 21.11)
 14 tasks (Figure 21.12)

*Figure 20.11: Systems Measured in the Experiment
 (CMN, 1983, p. 271; BB, p. 198)*

System	Description
<i>Text-Editors</i>	
POET	Line-oriented, with relative line numbers.
SOS	Line-oriented, with absolute line-numbers.
BRAVO	Display-oriented; full-page; uses mouse for pointing.
<i>Graphics Systems</i>	
MARKUP	Uses mouse to draw and erase lines and areas on a display; commands selected from a hidden menu, which must be redisplayed each time.
DRAW	Lines defined by pointing with mouse to end points; commands selected with mouse from a menu.
SIL	Lines defined by pointing with mouse to end points; boxes defined by pointing to opposite vertices; commands selected by combinations of mouse buttons.
<i>Executive Subsystems</i>	
LOGIN	TENEX command for logging in.
FTP	Program for transferring files between computers.
CHAT	Program for establishing a "teletype" connection between two computers.
DIR	TENEX command for printing a file directory; has a subcommand mode.
DELVER	TENEX command for deleting old versions of a file.

Figure 20.12: *Tasks for the Experiment (CMN, 1983. p. 272; BB, p. 198)*

Editing Tasks (used for POET, SOS, BRAVO)

- T1. Replace one 5-letter word with another (one line from previous task).
- T2. Add a fifth letter to a 4-letter word (one line from previous task).
- T3. Delete a line of text (eight lines from previous task).
- T4. Move a 52-character sentence, spread over two lines, to the end of its paragraph (eight lines from previous task).

Graphics Tasks (used for MARKUP, DRAW, SIL)

- T5. Add a rectangular box to a diagram.
- T6. Add a 5-character label to a box.
- T7. Disconnect a 2-segment line from one box and reconnect it to another box.
- T8. Delete a box, but keep an overlapped line.
- T9. Copy a box to another part of the diagram.

Executive Tasks

- T10. Phone computer and log in (4-character name, 6-character password).
 - T11. Transfer a file to another computer, renaming it.
 - T12. Connect to another computer.
 - T13. Display a subset of the file directory and show file lengths.
 - T14. Delete old versions of a file.
-

Results predicting performance times
RMS error is 21% of average predicted execution time
(Figs. 21.13, 21.14)

Figure 21.13: Predicted vs. Observed Execution Times in the Experiment
(CMN, 1983, p. 277; BB, p. 201)

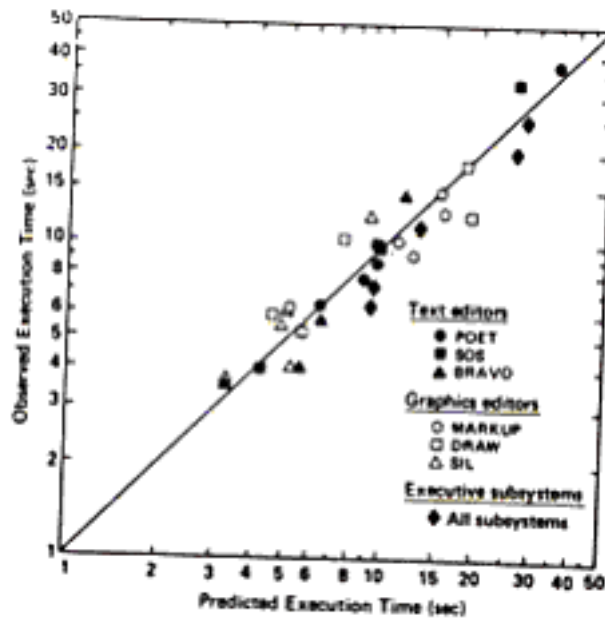


Figure 21.14: Calculated and Observed Execution Times in the Experiment (CMN, 1983, p. 276; BB, p. 200)

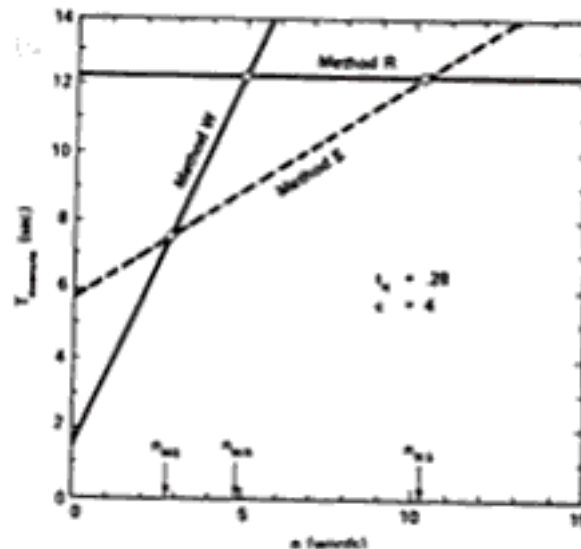
Task-System	Calculated										Observed		Pred. Error
	t_k	n_M	n_K	n_M	n_P	n_D	t_D	T_R	$T_{transfer}$	$T_{transfer}$	$M \pm SE(N)$		
	(sec)						(sec)	(sec)	(sec)	(sec)	(sec)	(sec)	
T1-PCET	.23	4	15	--	--	--	--	--	8.8	7.8 ± 0.9(27)	11%		
T1-SOS	.22	4	19	--	--	--	--	--	9.6	9.6 ± 0.8(31)	1%		
T1-BRAVO	.23	2	8	2	1	--	--	--	6.4	5.7 ± 0.3(31)	11%		
T2-PCET	.28	4	14	--	--	--	--	--	9.4	8.9 ± 0.7(17)	5%		
T2-SOS	.23	4	18	--	--	--	--	--	9.5	9.7 ± 0.8(32)	-3%		
T2-BRAVO	.24	2	4	2	1	--	--	--	5.6	4.1 ± 0.3(32)	26%		
T3-PCET	.19	3	12	--	--	--	--	--	6.3	6.3 ± 0.4(24)	0%		
T3-SOS	.23	2	7	--	--	--	--	--	4.3	4.0 ± 0.3(37)	8%		
T3-BRAVO	.23	1	2	1	1	--	--	--	3.3	3.5 ± 0.2(38)	-7%		
T4-PCET	.19	13	92	--	--	--	--	--	35.3	37.1 ± 4.3(20)	-6%		
T4-SOS	.23	12	47	--	--	--	--	--	26.8	32.7 ± 1.8(16)	-22%		
T4-BRAVO	.24	2	6	1	3	--	--	3.8	11.6	14.3 ± 1.1(33)	-23%		
T5-MARKUP	.25	--	3.2	--	2.5	4	24.9	--	11.1	10.5 ± 1.1(27)	6%		
T5-DRAW	.25	7.6	12.6	--	5	--	--	--	18.9	12.5 ± 3.0(22)	34%		
T5-SL	.27	1	4	0.4	2	--	--	--	4.8	5.4 ± 0.7(32)	-12%		
T6-MARKUP	.26	1	7	2	1	--	--	--	5.0	6.2 ± 0.4(34)	-23%		
T6-DRAW	.25	1	7	1	1	--	--	--	4.6	5.9 ± 0.4(34)	-29%		
T6-SL	.27	--	6	1.4	1	--	--	--	3.3	3.6 ± 0.3(19)	-9%		
T7-MARKUP	.24	--	8.6	--	4.8	6	13.6	--	15.1	15.0 ± 2.1(29)	2%		
T7-DRAW	.19	5	13	--	8	--	--	--	16.0	16.2 ± 1.8(9)	-1%		
T7-SL	.28	1	8	--	5	--	--	--	9.1	12.3 ± 2.1(23)	-36%		
T8-MARKUP	.26	--	8	--	8	1	4.0	--	12.3	9.3 ± 0.4(22)	24%		
T8-DRAW	.21	1	5	--	3	--	--	--	5.7	5.3 ± 0.3(25)	7%		
T8-SL	.27	1	5	0.7	2	--	--	--	5.2	4.1 ± 0.2(33)	20%		
T9-MARKUP	.25	2	8	--	6.5	--	--	3.5	15.4	13.0 ± 2.5(26)	15%		
T9-DRAW	.22	--	5.7	--	5.7	--	--	--	7.5	10.5 ± 1.0(25)	-40%		
T9-SL	.28	--	5	0.3	3	--	--	--	4.8	6.0 ± 1.0(26)	-24%		
T10-LOGIN	.29	2	28	--	--	--	--	15.9	27.4	25.1 ± 0.7(29)	9%		
T11-FTP	.30	5	31	--	--	--	--	10.1	26.1	19.7 ± 0.7(29)	24%		
T12-CHAT	.31	1	11	--	--	--	--	8.3	13.1	11.5 ± 0.6(36)	12%		
T13-DIR	.30	2	20	--	--	--	--	0.5	9.2	6.6 ± 0.3(32)	28%		
T14-DELVER	.32	2	20	--	--	--	--	0.4	9.4	7.5 ± 0.4(33)	20%		

Other results

Calculating benchmarks for performance

Parametric and sensitivity analysis (Fig. 21.15)

Figure 21.15: Execution times of 3 methods for the misspelled-word task as a function of n , where the misspelled word is n words back (CMN, 1983, p. 290; BB, p. 203; BGBG, p. 592)



Model simplifications	RMS Error
Keystroke Level	22%
Constant Operator Level	34%
$\tau(n_M + n_K + n_P + n_H + n_D) + T_R$	
Prorated Mental Time	45%
$\mu(T_K + T_H + T_P + T_D) + T_R$	
Keystroke Only	49%
$Kn_K + T_R$	

21.6 Extensions and Limitations (of Both Theories)

Display editors, graphic systems — Extension

Issues of “mental” operators — Limitation

Doesn't deal with errors, non-expert behaviour, learning, problem solving