

USER MODELING FOR INDIVIDUALS WITH DISABILITIES

by

Abhishek Agarwal

BE, Instrumentation and Control, Pune University, India, 2001

Submitted to the Graduate Faculty of
The School of Engineering in partial fulfillment
of the requirements for the degree of
Master of Science in Bioengineering

University of Pittsburgh

2005

UNIVERSITY OF PITTSBURGH

SCHOOL OF ENGINEERING

This thesis was presented

by

Abhishek Agarwal

It was defended on

November 22nd, 2005

and approved by

Dr. Rory A. Cooper, FISA/PVA Chair and Distinguished Professor, Rehabilitation Science and

Technology, Professor, Department of Bioengineering

Dr. Michael McCue, Associate Professor, Rehabilitation Science and Technology

Thesis Advisor: Dr. Richard C. Simpson, Assistant Professor, Rehabilitation Science and

Technology and Department of Bioengineering

Copyright © by Abhishek Agarwal

2005

USER MODELING FOR INDIVIDUALS WITH DISABILITIES

Abhishek Agarwal, M.S.

University of Pittsburgh, 2005

Clinicians have a limited amount of time for performing computer access and augmentative and alternative communication (AAC) assessments. In addition, they do not have access to all of the computer access devices that could potentially be useful for each client. An accurate modeling technique would help clinicians to identify the most appropriate kind of devices and device configurations for their clients. It would also be able to provide accurate prediction of performance, learning and fatigue.

Investigators are using word prediction (WP) as a test-bed for user modeling techniques. The goal is to develop accurate models that will form the basis for clinical assessment tools. The focus of this research was to observe users' interaction with WP in great detail, in preparation for future studies and developing a model.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	XXI
1.0 INTRODUCTION.....	1
2.0 BACKGROUND	3
2.1 GOMS.....	3
2.2 PREVIOUS RESEARCH ON WORD PREDICTION	5
2.3 MOTIVATION FOR DOING THIS RESEARCH	6
3.0 RESEARCH QUESTIONS	7
4.0 INSTRUMENTATION.....	8
4.1 HARDWARE	8
4.2 SOFTWARE	10
4.2.1 WP Java testbed.....	10
4.2.2 Key-Capture	13
4.2.3 Atomic clock sync.....	13
4.2.4 Biopac software	14
4.3 PROCEDURE	15
4.3.1 System Initialization	15
4.3.2 System Synchronization	15
4.3.3 Configure Subjects' Computer	16

4.3.4	Activating and Initializing the Iscan System	19
4.3.5	Calibrating the I-Scan System	22
4.3.6	Setting up the Biopac computer.....	24
4.3.7	Starting the experiment.....	29
4.3.8	Saving data	30
4.3.9	Create a boundary file for the Iscan software.....	31
4.3.10	Iscan Data Analysis.....	31
4.3.11	Data integration using Matlab codes.....	34
4.3.12	Data analysis using Matlab codes.....	42
5.0	METHODOLOGY.....	50
5.1	HUMAN SUBJECTS.....	50
5.2	PROTOCOL	50
6.0	DATA ANALYSIS	52
6.1	INDEPENDENT VARIABLES.....	52
6.2	DEPENDENT VARIABLES	56
6.2.1	Keystrokes	56
6.2.2	Keystroke Savings.....	57
6.2.3	Text Entry Rate.....	58
6.2.4	Keystroke Rate	59
6.2.5	Time between keystrokes	59
6.2.6	List Searches.....	59
6.2.7	List Search Time	61
6.2.8	Transcribed and Target Text Views	61

6.2.9	Sequences of Events	62
6.2.10	Test Duration.....	63
6.3	ERRORS ENCOUNTERED AND HANDLED.....	63
6.3.1	Systematic errors	64
6.3.2	Random errors	64
6.3.3	Instrumentation failure	65
6.3.4	Human error and unexpected occurrences	65
6.4	STATISTICAL TESTS - PAIR-WISE COMPARISON (T-TEST).....	66
6.4.1	Why do we need pair-wise comparison?.....	66
6.4.2	T-test analysis for Text Entry Rate	68
6.4.3	T-test analysis for Keystroke Rate	69
6.4.4	T-test analysis for % Correct Keystrokes	70
6.4.5	T-test analysis for %Error Keystrokes.....	70
6.4.6	T-test analysis for Time between Keystrokes.....	71
6.4.7	T-test analysis for List Search Time.....	72
6.4.8	T-test analysis for Successful Predictions	72
6.4.9	T-test analysis for Unsuccessful Predictions.....	73
6.4.10	T-test analysis for Target Views	74
6.4.11	T-test analysis for Transcribed Views.....	75
6.5	MULTI-VARIATE ANALYSIS (ANOVA)	75
6.5.1	Effect of List Length	76
6.5.2	Effect of Minimum Word Size	77
6.5.3	Effect of Show and Hide	79

7.0	DISCUSSION	82
7.1	SUBJECT ‘G’	82
7.2	SUBJECT ‘A’	83
7.3	SUBJECT ‘D’	84
7.4	REMARKS	84
7.5	IMPLICATIONS FOR FUTURE RESEARCH	86
	APPENDIX A	89
	APPENDIX B	117
	BIBLIOGRAPHY	128

LIST OF TABLES

Table 1: The instruments used, what they measure and their resolution	9
Table 2: Sample output of ‘output.txt’ file	34
Table 3: Sample output of ‘hand events only’	36
Table 4: Sample output of ‘event frequency and time’ file	37
Table 5: Sample output of ‘transition frequency’	38
Table 6: Sample output from ‘finaloutput_1.txt’ file.....	39
Table 7: Sample output from ‘finaloutput_11.txt’ file.....	40
Table 8: Format of the list-search table	42
Table 9: Format for data_by_condition.txt	42
Table 10: Format for data_by_sequence.txt.....	43
Table 11: Layout of table for 1 way ANOVA - condition wise	44
Table 12: Layout of table for 1 way ANOVA - sequence wise.....	44
Table 13: Layout of table for ANOVA for factors word size and actual length	45
Table 14: Layout of table for ANOVA for factors word size and trial number	46
Table 15: Layout of table for ANOVA for factors trial number and actual length	46
Table 16: Layout of table for ANOVA for factors max list length and actual length	47
Table 17: Layout of table for ANOVA for factors max list length and trial number	47
Table 18: Layout of table for ANOVA for factors actual length and trial number	47

Table 19: Layout of table for ANOVA for 2 factors arranged in row and column	49
Table 20: Subjects	50
Table 21: Independent variable values which define various configurations	53
Table 22: The values of variables for all the conditions	54
Table 23: Showing sequence order in which the conditions were presented to each subject	55
Table 24: Summary of missing data of the subjects	65
Table 25: Shows sample output from the T-test	67
Table 26: T-test analysis for text entry rate	68
Table 27: T-test analysis for keystroke rate	69
Table 28: T-test analysis for % Correct Keystrokes	70
Table 29: T-test analysis for % Error Keystrokes	70
Table 30: T-test analysis for time between keystrokes	71
Table 31: T-test analysis for list search time	72
Table 32: T-test analysis for successful predictions	73
Table 33: T-test analysis for unsuccessful predictions	73
Table 34: T-test analysis for target views	74
Table 35: T-test analysis for transcribed views	75
Table 36: Result for effect of list length from ANOVA2	76
Table 37: Results for effect of minimum word size from ANOVA2	78
Table 38: Result for effect of show and hide from ANOVA2	80
Table 39: T-test results for text entry rate for 5th trial only	117
Table 40: T-test results for text entry rate for average of all trials	118
Table 41: T-test results for keystroke rate for 5th trial only	118

Table 42: T-test results for keystroke rate for average of all trials	119
Table 43: T-test results for % correct keystrokes for 5th trial only	119
Table 44: T-test results for %correct keystrokes for average of all trials	120
Table 45: T-test results for %error keystrokes for 5th trial only	120
Table 46: T-test results for %error keystrokes for average of all trials	121
Table 47: T-test results for time between keystrokes for 5th trial only	121
Table 48: T-test results for time between keystrokes for average of all trials	122
Table 49: T-test results for list search time for 5th trial only	122
Table 50: T-test results for list search time for average of all trials	123
Table 51: T-test results for successful prediction for 5th trial only	123
Table 52: T-test results for successful prediction for average of all trials	124
Table 53: T-test results for unsuccessful prediction for 5th trial only	124
Table 54: T-test results for unsuccessful prediction for average of all trials	125
Table 55: T-test results for target text views for 5th trial only	125
Table 56: T-test results for target text views for average of all trials	126
Table 57: T-test results for transcribed text views for 5th trial only	126
Table 58: T-test results for transcribed text views for average of all trials	127

LIST OF FIGURES

Figure 1: Schematic of experimental setup.....	8
Figure 2: Screen shot of WP interface	10
Figure 3: Sample output from WP test bed.....	11
Figure 4: Sample of Key-Capture output file	13
Figure 5: Sample output from Biopac data acquisition system	14
Figure 6: Screen shot of atomic clock interface.....	16
Figure 7: Position of monitor for subjects' computer.....	16
Figure 8: Screen shot of the WP interface showing option for configuration selection	18
Figure 9: Screen shot of key capture interface.....	19
Figure 10: Screenshot of I-Scan system interface.....	20
Figure 11: Scene and Eye monitors	21
Figure 12: Cross marks used to calibrate the eye tracking system	22
Figure 13: Sample output from eye tracking data file	23
Figure 14: Biopac MP 150 hardware	24
Figure 15: Biopac MP 150 system.....	25
Figure 16: Screen shot showing system calibration I	26
Figure 17: Screen shot showing system calibration II	26
Figure 18: Screen shot showing system calibration III.....	27

Figure 19: Screen shot showing setting time in real time format	28
Figure 20: Picture of an accelerometer	28
Figure 21: Screen shot showing outputs of X, Y and Z movements	29
Figure 22: Screen shot of Point of Regard (POR) data analysis.....	32
Figure 23: Sample output of eye fixation file	32
Figure 24: Screen shot of eye fixations.....	33
Figure 25: Compliance, for the 5th trial in all the conditions.....	89
Figure 26: Compliance, for the average of all trials in all the conditions.....	89
Figure 27: Percentage correct keystrokes for 5th trial only in all the conditions. The data is presented in order of condition.	90
Figure 28: Percentage correct keystrokes for the average of all trials in all the conditions. The data is presented in order of condition.	90
Figure 29: Percentage correct keystrokes for 5th trial only in all the conditions. The data is presented in order of sequence.....	90
Figure 30: Percentage correct keystrokes for the average of all trials in all the conditions. The data is presented in order of sequence.	90
Figure 31: Percentage error keystrokes for 5th trial only in all the conditions. The data is presented in order of condition.	91
Figure 32: Percentage error keystrokes for the average of all trials in all the conditions. The data is presented in order of condition.....	91
Figure 33: Percentage error keystrokes for 5th trial only in all the conditions. The data is presented in order of sequence.....	91
Figure 34: Percentage error keystrokes for the average of all trials in all the conditions. The data is presented in order of sequence.	91
Figure 35: Text entry rate for 5th trial only in all the conditions. The data is presented in order of condition.	92

Figure 36: Text entry rate for the average of all trials in all the conditions. The data is presented in order of condition.....	92
Figure 37: Text entry rate for 5th trial only in all the conditions. The data is presented in order of sequence.....	92
Figure 38: Text entry rate for the average of all trials in all the conditions. The data is presented in order of sequence.....	92
Figure 39: Keystroke rate for 5th trial only in all the conditions. The data is presented in order of condition.	93
Figure 40: Keystroke rate for the average of all trials in all the conditions. The data is presented in order of condition.....	93
Figure 41: Keystroke rate for 5th trial only in all the conditions. The data is presented in order of sequence.....	93
Figure 42: Keystroke rate for the average of all trials in all the conditions. The data is presented in order of sequence.....	93
Figure 43: Time between keystrokes for 5th trial only in all the conditions. The data is presented in order of condition.....	94
Figure 44: Time between keystrokes for the average of all trials in all the conditions. The data is presented in order of condition.	94
Figure 45: Time between keystrokes for 5th trial only in all the conditions. The data is presented in order of sequence.....	94
Figure 46: Time between keystrokes for the average of all trials in all the conditions. The data is presented in order of sequence.....	94
Figure 47: User selected the word when it was displayed in the list, for the 5th trial in all the conditions.....	95
Figure 48: User selected the word when it was displayed in the list, for the average of all trials in all the conditions.	95

Figure 49: User did not select the word when the target word was displayed in the list, for the 5th trial in all the conditions.	95
Figure 50: User did not select the word when the target word was displayed in the list, for the average of all trials in all the conditions.	95
Figure 51: User selected the word when the target word was not displayed in the list, for the 5th trial in all the conditions.	96
Figure 52: User selected the word when the target word was not displayed in the list, for the average of all the trials in all the conditions.	96
Figure 53: User did not search the word when the target word was there in the displayed list, for the 5th trial in all the conditions.	96
Figure 54: User did not search the word when the target word was there in the displayed list, for the average of all the trials in all the conditions.	96
Figure 55: User did not search the word when the target word was not there in the displayed list, for the 5th trial in all the conditions.	97
Figure 56: User did not search the word when the target word was not there in the displayed list, for the average of all the trials in all the conditions.	97
Figure 57: User did not search the word when the list was not displayed, for the 5th trial in all the conditions.	97
Figure 58: User did not search the word when the list was not displayed, for the average of all trials in all the conditions.	97
Figure 59: User searched the word when the list was not displayed, for the 5th trial in all the conditions.	98
Figure 60: User searched the word when the list was not displayed, for the average of all trials in all the conditions.	98
Figure 61: Successful anticipation, for the 5th trial in all the conditions.	98
Figure 62: Successful anticipation, for the average of all trials in all the conditions.	98

Figure 63: List search time when the target word is in the list, for 5th trial only in all the conditions. The data is presented in order of condition.	99
Figure 64: List search time when the target word is in the list, for the average of all trials in all the conditions. The data is presented in order of condition.	99
Figure 65: List search time when the target word is in the list, for 5th trial only in all the conditions. The data is presented in order of sequence.	99
Figure 66: List search time when the target word is in the list, for the average of all trials in all the conditions. The data is presented in order of sequence.....	99
Figure 67: List search time when the target word is not in the list, for 5th trial only in all the conditions. The data is presented in order of condition.	100
Figure 68: List search time when the target word is not in the list, for the average of all trials in all the conditions. The data is presented in order of condition.	100
Figure 69: List search time when the target word is not in the list, for 5th trial only in all the conditions. The data is presented in order of sequence.	100
Figure 70: List search time when the target word is not in the list, for the average of all trials in all the conditions. The data is presented in order of sequence.....	100
Figure 71: List search time – total (word present + word not present), for 5th trial only in all the conditions. The data is presented in order of condition.	101
Figure 72: List search time – total (word present + word not present), for the average of all trials in all the conditions. The data is presented in order of condition.	101
Figure 73: List search time – total (word present + word not present), for 5th trial only in all the conditions. The data is presented in order of sequence.	101
Figure 74: List search time – total (word present + word not present), for the average of all trials in all the conditions. The data is presented in order of sequence.	101
Figure 75: List search time when the search was successful, for 5th trial only in all the conditions. The data is presented in order of condition.	102

Figure 76: List search time when the search was successful, for the average of all trials in all the conditions. The data is presented in order of condition.	102
Figure 77: List search time when the search was successful, for 5th trial only in all the conditions. The data is presented in order of sequence.	102
Figure 78: List search time when the search was successful, for the average of all trials in all the conditions. The data is presented in order of sequence.	102
Figure 79: List search time when the search was not successful, for 5th trial only in all the conditions. The data is presented in order of condition.	103
Figure 80: List search time when the search was not successful, for the average of all trials in all the conditions. The data is presented in order of condition.	103
Figure 81: List search time when the search was not successful, for 5th trial only in all the conditions. The data is presented in order of sequence.	103
Figure 82: List search time when the search was not successful, for the average of all trials in all the conditions. The data is presented in order of sequence.....	103
Figure 83: Target text views for 5th trial only in all the conditions. The data is presented in order of condition.	104
Figure 84: Target text views for the average of all trials in all the conditions. The data is presented in order of condition.	104
Figure 85: Target text views for 5th trial only in all the conditions. The data is presented in order of sequence.....	104
Figure 86: Target text views for the average of all trials in all the conditions. The data is presented in order of sequence.....	104
Figure 87: Transcribed text views for 5th trial only in all the conditions. The data is presented in order of condition.....	105
Figure 88: Transcribed text views for the average of all trials in all the conditions. The data is presented in order of condition.	105

Figure 89: Transcribed text views for 5th trial only in all the conditions. The data is presented in order of sequence.	105
Figure 90: Transcribed text views for the average of all trials in all the conditions. The data is presented in order of sequence.	105
Figure 91: Still frequency for 5th trial in all the conditions.	106
Figure 92: Still frequency for average of all trials in all the conditions.	106
Figure 93: Still duration for 5th trial in all the conditions.	106
Figure 94: Still duration for average of all trials in all the conditions.	106
Figure 95: Movement frequency for 5th trial in all the conditions.	107
Figure 96: Movement frequency for average of all trials in all the conditions.	107
Figure 97: Movement duration for 5th trial in all the conditions.	107
Figure 98: Movement duration for average of all trials in all the conditions.	107
Figure 99: Keystrokes frequency for 5th trial in all the conditions.	108
Figure 100: Keystrokes frequency for average of all trials in all the conditions.	108
Figure 101: Keystrokes duration for 5th trial in all the conditions.	108
Figure 102: Keystrokes duration for average of all trials in all the conditions.	108
Figure 103: Total hand frequency for 5th trial in all the conditions.	109
Figure 104: Total hand frequency for average of all trials in all the conditions.	109
Figure 105: Total hand duration for 5th trial in all the conditions.	109
Figure 106: Total hand duration for average of all trials in all the conditions.	109
Figure 107: Target text views duration for 5th trial in all the conditions.	110
Figure 108: Target text views duration for average of all trials in all the conditions.	110
Figure 109: Transcribed text views duration for 5th trial in all the conditions.	110
Figure 110: Transcribed text views duration for average of all trials in all the conditions.	110

Figure 111: WP List views frequency for 5th trial in all the conditions.....	111
Figure 112: WP List views frequency for average of all trials in all the conditions.	111
Figure 113: WP List views duration for 5th trial in all the conditions.	111
Figure 114: WP List views duration for average of all trials in all the conditions.	111
Figure 115: Keyboard views frequency for 5th trial in all the conditions.	112
Figure 116: Keyboard views frequency for average of all trials in all the conditions.....	112
Figure 117: Keyboard views duration for 5th trial in all the conditions.....	112
Figure 118: Keyboard views duration for average of all trials in all the conditions.	112
Figure 119: Total eye views frequency for 5th trial in all the conditions.....	113
Figure 120: Total eye views frequency for average of all trials in all the conditions.....	113
Figure 121: Total eye views duration for 5th trial in all the conditions.	113
Figure 122: Total eye views duration for average of all trials in all the conditions.	113
Figure 123: Unsuccessful Prediction for 5th trial only in all the conditions. The data is presented in order of condition.....	114
Figure 124: Unsuccessful Prediction for the average of all trials.....	114
Figure 125: Unsuccessful Prediction for 5th trial only in all the conditions. The data is presented in order of sequence.	114
Figure 126: Unsuccessful Prediction for the average of all trials.....	114
Figure 127: Successful Prediction for 5th trial only in all the conditions. The data is presented in order of condition.....	115
Figure 128: Successful Prediction for the average of all trials in all the conditions. The data is presented in order of condition.	115
Figure 129: Successful Prediction for 5th trial only in all the conditions. The data is presented in order of sequence.	115

Figure 130: Successful Prediction for the average of all trials in all the conditions. The data is presented in order of sequence..... 115

Figure 131: Duration of tests in each condition (in minutes). (Note that the time is for four trials for subject ‘G’ in condition 15). 116

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr Richard Simpson for giving me an opportunity to work on this project. It was truly a wonderful experience working with him. I sincerely thank him for his understanding, support and willingness to share and help.

I express my appreciation to Dr Rory Cooper and Dr Michael McCue for their advice and for serving as members of the thesis committee.

I am thankful to the participants who volunteered themselves for this research study.

Finally, I would like to express my gratitude to my family. The continued encouragement from my parents has always motivated me to attain higher goals in life.

1.0 INTRODUCTION

Modeling is the process of creating a representation of real objects or environments, which simulate processes to predict results. Models are useful in testing the effect of changes of system variables on the overall performance of the user and help in better understanding of a user's interaction with the system. The user model defines the type of users and their relevant attributes to the interface. In this project, we are trying to predict user performance based on different system configurations so that in the future we may classify the need and capabilities of the user and predict the right configuration for them.

Word prediction (WP) is used here as a broad term for word completion and prediction. WP is often presented as a means of (1) reducing the number of keystrokes required to generate text and (2) increasing the rate at which text can be generated, and is particularly useful for individuals with cognitive or motor impairments. While it is clear that WP will reduce keystrokes, empirical evidence [1] has demonstrated that WP does not necessarily increase text generation rate (TGR) and may, in fact, reduce TGR. This may still be useful for subjects with motor or cognitive disabilities for whom a substantial reduction in motor movements is more important than TGR alone.

An estimated 19.4% of non-institutionalized civilians in the United States, totaling 48.9 million people have a disability. Almost half of these people (an estimated 24.1 million people)

can be considered to have a severe disability [2]. A national survey on disabilities in the US found that two-thirds of working-age adults with disabilities were not employed and that nearly 80 percent of them wanted to work [3]. In a study, employment in individuals after rehabilitation rose from 18.8% to 85.4% [4]. These figures motivate our efforts to develop innovative technologies to help individuals with disabilities lead an independent life.

An important goal in selecting a suitable assistive technology is often to reduce motor movement and fatigue. One way to reduce hand movement for written communications is to use WP. Not only is WP useful for people with disabilities but it also has applications in text entry in mobile devices such as hand held devices and mobile phones, where a miniature form of keyboard is provided for text entry. It also has use for individuals with dyslexia who have trouble learning difficult words and spelling.

2.0 BACKGROUND

2.1 GOMS

A widely used modeling technique in human computer interaction is Goals, Operators, Methods and Selection Rules (GOMS). It is based on the general theory of human information processing called the Model Human Processor [5]. According to this theory, human activity is governed by Cognitive, Perceptual and Motor Processors. The user is assigned a task to achieve a specific 'Goal'. This can be done through certain defined rules called 'Operators', which are low level tasks (perceptual, cognitive and motor) that bring about changes and are packaged into subtasks called 'Methods'. 'Selection Rules' are applied when the task can be accomplished by several methods. Proper specification of operators is crucial in determining the accuracy of the model [5].

GOMS provides both quantitative and qualitative analysis of a system. Qualitatively, it provides a picture of the users' interaction with the modeled system. Quantitatively it predicts features of the user's performance such as execution time, learning time and memory loads. These predictions can be used to design an appropriate system or strategy for the user [5].

There are at least four different modeling techniques based on GOMS. These four modeling techniques vary in complexity and are used to model different types of activities.

- 1) Keystroke-Level Model (KLM) [6][7] - KLM is based on a serial model of human information processing in which one activity is done at a time. It divides a task into small

operations (generally keystrokes) and totals the execution time for the individual operations. Its operators include four motor control operators (key-press (K), homing (H), pointing (P), drawing (D)), one mental operator (M) and one system response operator for time during execution (R). The execution time is the sum of all these operators:

$$T_{\text{EXECUTE}} = t_K + t_P + t_H + t_D + t_M + t_R$$

- 2) Card, Moran and Newell GOMS (CMN-GOMS) [6][7] – CMN-GOMS follows a strict goal hierarchy. When a certain task is given, it predicts both operator sequence and execution time. Based on the system design, a user's activity can be predicted. It is a serial stage architecture like KLM, meaning each action the user takes (searching the list, pressing a key) happens in sequence, with no overlap.
- 3) Natural GOMS Language (NGOMSL) [6][7] – NGOMSL provides predictions of operator sequence, execution time, and time to learn methods. It is based on a cognitive architecture called Cognitive Complexity Theory (CCT). Because it follows CCT, it can give estimations for execution time as well as for learning time. It also uses a serial model of cognition.
- 4) CPM-GOMS [6][7][8]- CPM stands for both “Cognitive Perceptual Model” and “Critical Path Method”. CPM-GOMS is also based directly on the model human processor, but does not make the assumption that the user's interaction is a serial process. The tasks are first joined together serially and then examined to see which actions can be overlapped so that they happen in parallel. This technique facilitates representation of overlapping and very efficient ‘chunks’ of activity characteristic of expert users. Models are developed using PERT charts and execution time is derived from the critical path.

2.2 PREVIOUS RESEARCH ON WORD PREDICTION

Koester developed and evaluated a model for word prediction using KLM, and performed two studies [11] [12]. The first study required two simulations (models 1A and 1B) and the second study involved an additional model (model 2). Model 1A and 1B represented performance as a linear combination of keypress and list search time, while model 2 used a revised model for list search time. Models 1B and 2 used parameter values derived from subjects' data.

Two interfaces were developed: a letters only system (in which text is entered by typing each letter) and a letters + WP system in which single letter entry was augmented by WP. When using WP, each subject was asked to use one of two strategies. Strategy 1 was to search the list before every selection and Strategy 2 was to choose the first two letters of a word without searching the list.

Koester's work resulted in the following conclusions:

- 1) Strategy of use can have a substantial effect on performance with WP.
- 2) It might help to display the list only when it is supposed to be searched, providing less distraction during keypress-only selections.
- 3) Use of fixed WP lists, as compared to dynamic lists, may be important in allowing the user to gradually develop expectations about the lists.
- 4) Presenting words in alphabetical order might decrease reliance on serial search.

2.3 MOTIVATION FOR DOING THIS RESEARCH

According to Koester, the simple structure of Model 1 led to a mismatch between the model's predictions and subjects' actual behavior [11]. Koester's model relied heavily on assumptions that led to a mismatch between empirical data and the model's predictions:

- 1) Keypress time was the same for all letters and was independent of a key's position.
- 2) The time required to search the word prediction list was independent of the list's contents and sequential position of the item in the list.
- 3) The user's interaction with the word prediction system was a serial process; in other words, each action the user took (searching the list, pressing a key) happened in sequence, with no overlap.
- 4) Users were experts: There were no errors encountered during the trials and they had consistent times for actions.
- 5) Subjects used the same strategy for each keystroke and word.
- 6) Search time was the same for all individuals even after performing several trials.
- 7) Performance of able-bodied (AB) individuals typing with mouth sticks was taken as an approximate time for typing for spinal cord injured (SCI) individuals.

3.0 RESEARCH QUESTIONS

Many potential questions arose from Koester's research findings:

1. How is typing accuracy affected by different configurations and user characteristics?
2. How can cognitive and perceptual time and effort be evaluated?
3. What determines how long it takes a person to search the word prediction list?

4.0 INSTRUMENTATION

To develop an accurate model we must have a detailed understanding of how people interact with WP software. The main interaction components examined in our project are key presses, eye fixations and hand movement.

4.1 HARDWARE

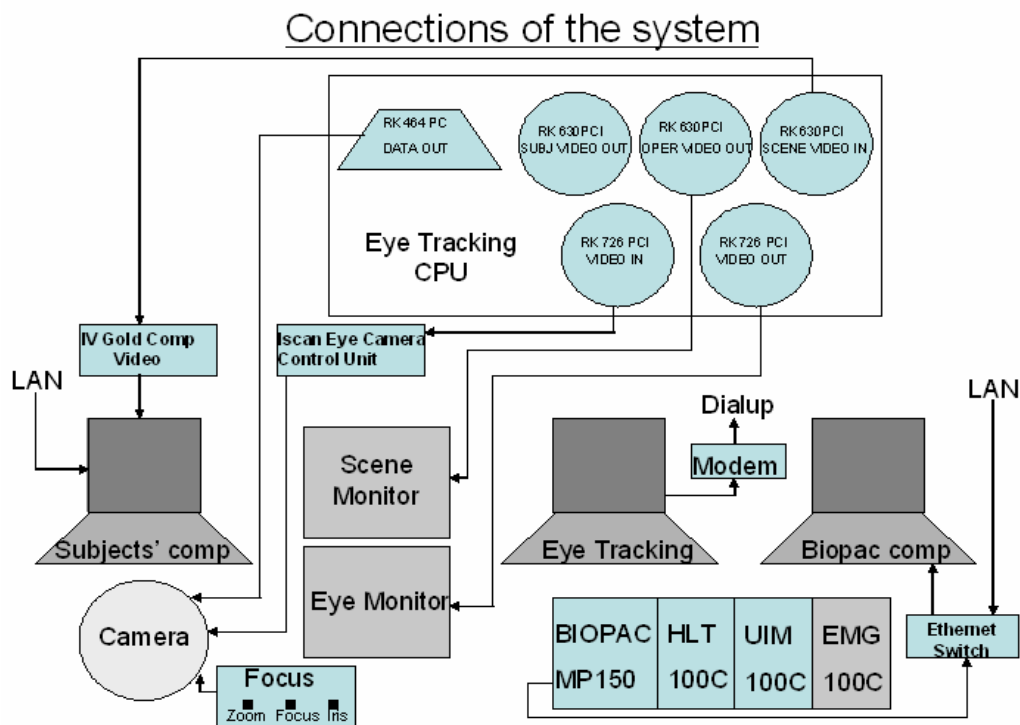


Figure 1: Schematic of experimental setup

As shown in Figure 1, the experimental setup involves three computers and associated hardware. The subjects' computer runs the WP test bed and key-capture software.

The Biopac computer runs the physiological measurement data acquisition software packaged with the Biopac MP150 system. The MP150 is used to record hand movement using an accelerometer placed on the subject's wrist.

The eye-tracking computer runs the Iscan eye tracking software. The system tracks the center of the subject's pupil by measuring the differential reflection between the cornea and pupil. Low level infrared (IR) light, from an IR source attached to the camera, are made incident on the eyes. The differential reflection is passed through IR filters and captured by the camera. The Eye monitor shows the subjects eye as seen by the camera and the Scene monitor mirrors the display of the subject's computer and superimposes a cross-hatch where the subject is looking.

Table 1: The instruments used, what they measure and their resolution

<u>Instrument</u>	<u>Measure</u>	<u>Unit</u>	<u>Frequency/Resolution</u>
Accelerometer	Hand motion in the X, Y and Z planes	cm/s ²	50 Hz
Eye Tracking System	Visual Fixations	(X, Y) pixel coordinates and time duration	60 Hz
Key Capture Software	Key Up and Key Down Events	N/A	Milliseconds
WP testbed	Contents of WP list, WP list displayed (yes/no)	N/A	N/A

4.2 SOFTWARE

4.2.1 WP Java testbed

The user interface with which subjects interact is a word prediction program written in Java. The testbed allows investigators to record a time-stamped log of each keystroke and provides several configuration options. As shown in Figure 2, two text boxes are displayed at the center of the screen: the target text and the transcribed text. The WP list shows up at the top left corner of the monitor. The user's task is to type the contents of the target text box into the transcribed text box.

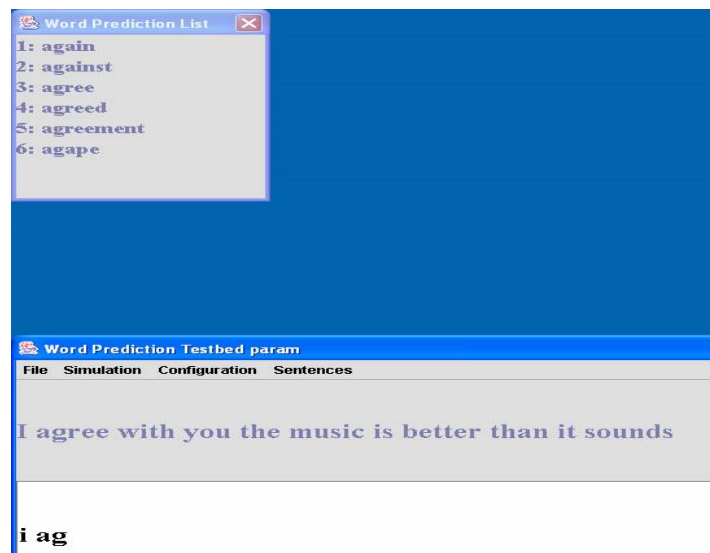


Figure 2: Screen shot of WP interface

The output from this software (shown in Figure 3) provides information about the system parameters of the condition, sentence group used, what the subject typed and the contents of the list displayed.

Figure 3: Sample output from WP test bed

Subject Name = G

End Sentence = 20

Show Word Prediction List when 0 letter(s) is/are typed

Hide Word Prediction List when 4 letter(s) is/are typed after the list is displayed

Max List size = 2

Min Word Size = 3

Target Sentences are :

Sentence number 16 is -> starlight and dewdrop if at first you do not succeed

Sentence number 17 is -> you are not a jedi yet circumstances are unacceptable

Sentence number 18 is -> love means many things physics and chemistry are hard

Sentence number 19 is -> one heck of a question great disturbance in the force

Sentence number 20 is -> breathing is difficult prayer in schools offends some

Figure 3 (Continued)

Simulation parameters are :

Type	Text	WPL Len	WPL Disp	Total Time	WPL Contents
7		0	false	7286299	
8		0	false	7291337	
9		0	false	7291357	
4	s	0	false	7291708	
9		0	false	7291898	
4	t	8	true	7291708	the that you they this take time think
9		0	false	7291898	
4	a	2	true	7292018	the that
9		0	false	7292198	
4	r	2	true	7292248	take talk
9		0	false	7292419	
4	l	2	true	7292719	tardiness tardy
9		0	false	7292829	
4	i	2	false	7292920	tardiness tardy
9		0	false	7293010	
4	g	2	false	7293140	tardiness tardy
9		0	false	7293240	
4	h	2	false	7293340	tardiness tardy
9		0	false	7293430	
4	t	2	false	7293541	tardiness tardy
9		0	false	7293651	

4.2.2 Key-Capture

Key-Capture is used to record the time when the user presses or releases a key. Key-Capture is used in addition to the word prediction interface because it provides more accurate event timing. A sample from an output file generated by Key-Capture is shown in Figure 4.

```
001 GeneralMessage Version(1.3) Subject(a)
002 TimeMessage 20041209-172114-712 Counter(378670428066)
Frequency(3579545Hz) Correction(0ms)
003 KeyPress 20041209-172125-829 "S" Status=(down) Key(83) Extra(0x1f)
004 KeyPress 20041209-172125-926 "S" Status=(up) Key(83) Extra(0xc01f)
005 KeyPress 20041209-172126-288 "T" Status=(down) Key(84) Extra(0x14)
006 KeyPress 20041209-172126-385 "T" Status=(up) Key(84) Extra(0xc014)
007 KeyPress 20041209-172126-590 "A" Status=(down) Key(65) Extra(0x1e)
008 KeyPress 20041209-172126-688 "A" Status=(up) Key(65) Extra(0xc01e)
009 KeyPress 20041209-172126-828 "R" Status=(down) Key(82) Extra(0x13)
010 KeyPress 20041209-172126-917 "R" Status=(up) Key(82) Extra(0xc013)
011 KeyPress 20041209-172127-292 "L" Status=(down) Key(76) Extra(0x26)
```

Figure 4: Sample of Key-Capture output file

4.2.3 Atomic clock sync

Atomic Clock Sync is used on all three computers to synchronize their internal clocks to an atomic clock through the internet. This ensures that the data acquired from all the systems can be integrated together using a common reference time.

4.2.4 Biopac software

The Biopac software is used in combination with the Biopac data acquisition hardware to record hand movement. The output file has four columns, one showing time and the other three showing movement along the X, Y and Z axes. A sample of an output file from the Biopac software is shown in Figure 5.

Time	X	Y	Z
62477	-57.6288	0.117624	7.21387
62477	-57.0184	0.131525	7.70215
62477	-58.468	0.127734	6.54248
62477.1	-56.7896	0.135316	6.2373
62477.1	-57.3999	0.111306	7.88525
62477.1	-57.2473	0.122679	7.0918
62477.1	-57.3236	0.14037	5.99316
62477.1	-57.2473	0.11257	6.96973
62477.2	-56.7133	0.120152	7.45801
62477.2	-58.0103	0.125206	6.42041
62477.2	-57.171	0.11257	6.72559

Figure 5: Sample output from Biopac data acquisition system

4.3 PROCEDURE

4.3.1 System Initialization

Switch on all the computers, the Iscan camera, the focus control box, the eye and scene monitors, the Iscan eye camera control unit and the Biopac system. Make sure that all of the systems (including external modem and Ethernet switch) are plugged in properly.

4.3.2 System Synchronization

Before starting the experiment make sure that the atomic clock software is running on all the three computers and they are synchronized. Since the eye tracking computer runs on Windows ME (which is not supported by University LAN), it has to be connected to the internet through a dialup connection using an external modem. The atomic clock runs automatically when each computer is connected to Internet. Synchronize the clocks on all three computers by clicking 'ping now' on each computer, as shown in Figure 6.

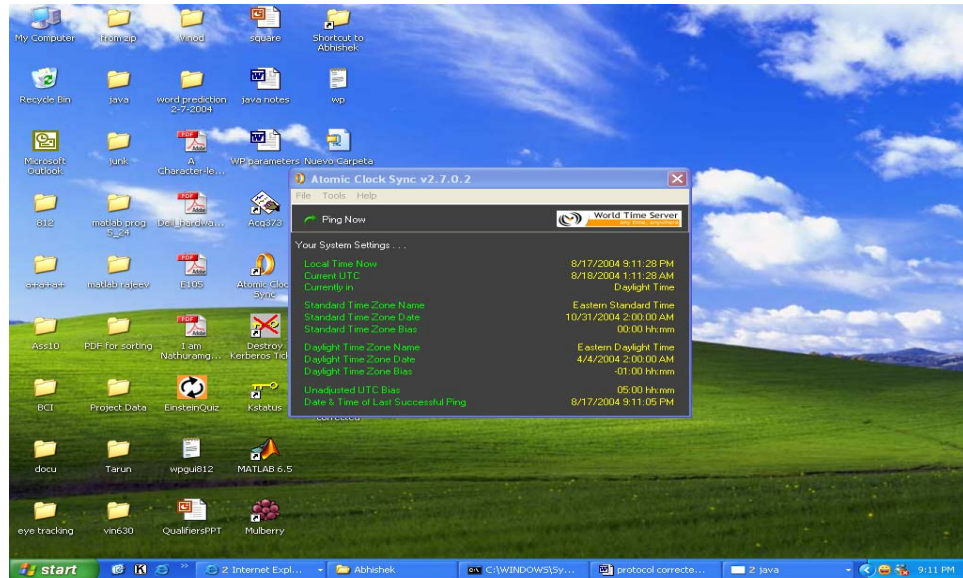


Figure 6: Screen shot of atomic clock interface

4.3.3 Configure Subjects' Computer

As shown in Figure 7, the subjects' computer monitor is raised by placing thick books underneath it to raise the base of the monitor to Iscan camera level. This ensures that the camera placed in front of the monitor does not block the view of the subject.



Figure 7: Position of monitor for subjects' computer

Ask the subject to sit in front of the monitor in the chair to which a headrest has been attached. Adjust the position of the headrest and chin pads to support the subject's head and prevent the subject's eye from moving outside the field of view of the camera.

To run the word prediction software:

- Go to start -> run
- type 'command'
- change directory to C:\Abhishek\wpsim
- type 'java wpgui'.

As shown in Figure 8, the experimenter can choose the condition from the 'configuration' drop down menu. There is currently a choice of 16 different conditions to choose from with the 17th being letters only. The experimenter can also select the sentence group in the test bed before starting the trial. A pre-sentence pause of five seconds is added by default to give the subject time to adjust and the investigator to start other systems.

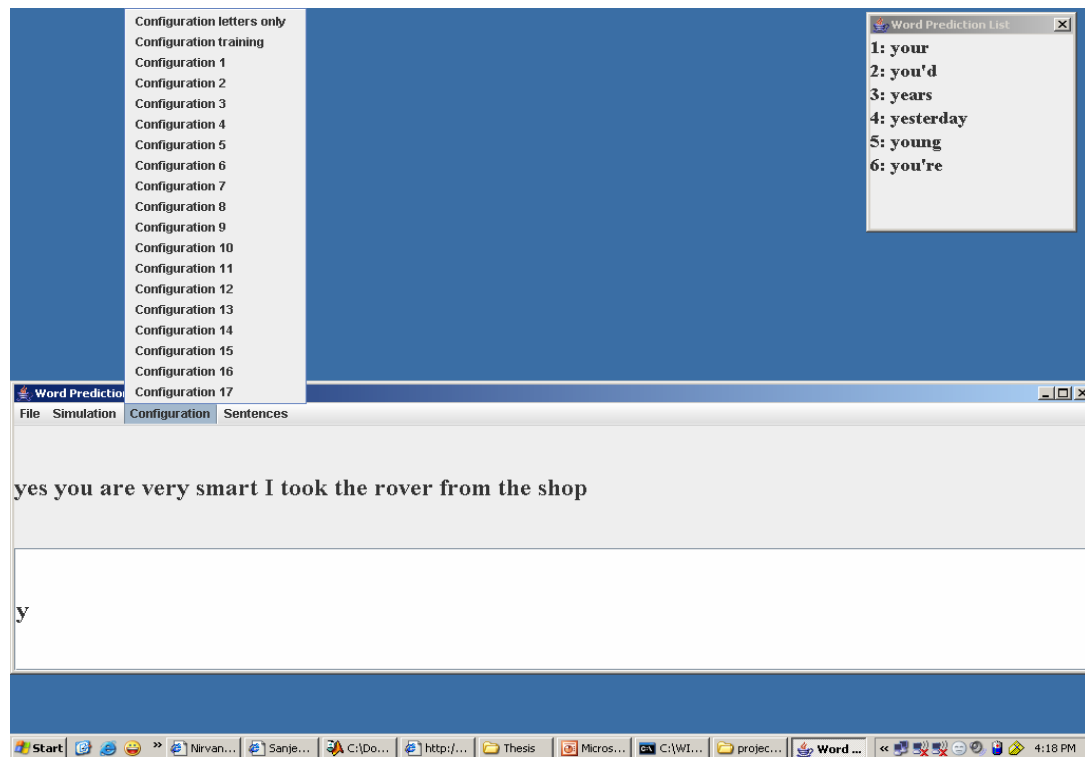


Figure 8: Screen shot of the WP interface showing option for configuration selection

Run the Key Capture software by clicking on the icon on the Windows quick launch bar. As shown in Figure 9, clear old data and configure the software to 'capture keystroke data' & 'capture keystroke release data'. Enable the automatic data save option and enter the appropriate time and size of file values. Check the 'Enable Key Capture' option.

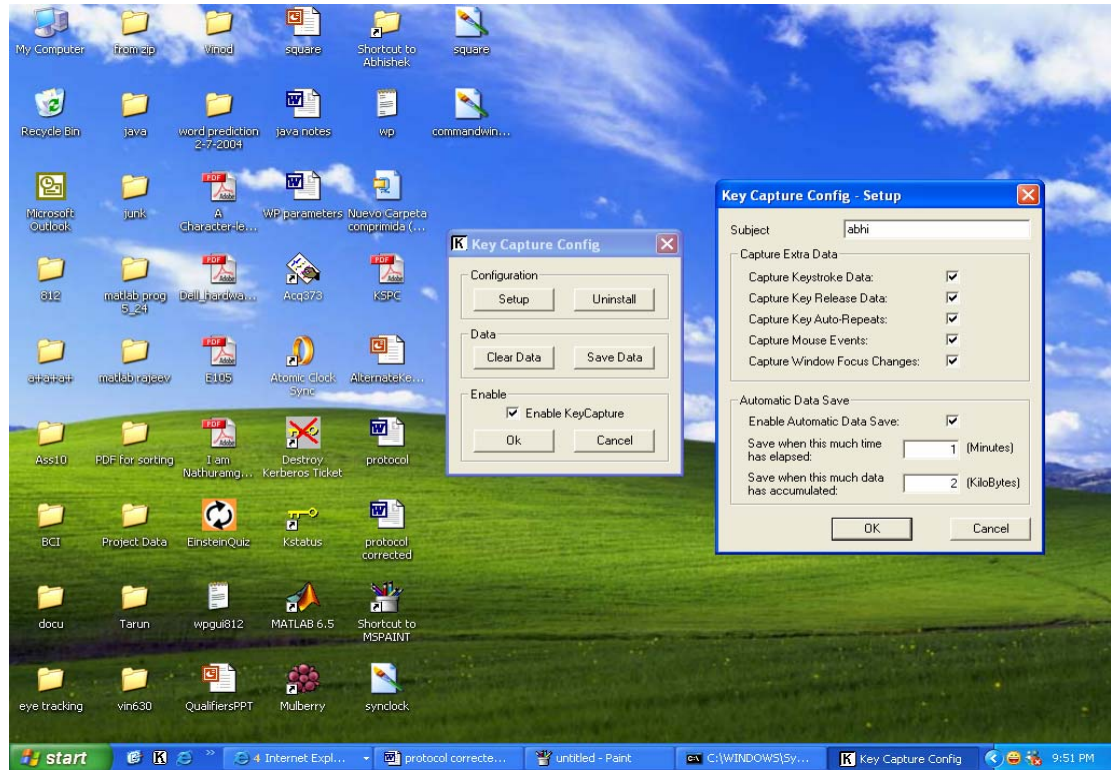


Figure 9: Screen shot of key capture interface

4.3.4 Activating and Initializing the Iscan System

Run the eye tracking software by going through the DOS command line. Note that if you run the software by clicking on the icon then you may encounter some problems later on, such as not being able to save data or completely freezing the computer. At the DOS prompt, change the directory to where the executable file is located and run the file. In our case the path is C:\Iscan\Abhishek. Now type the name of the executable file (daq3_58e).

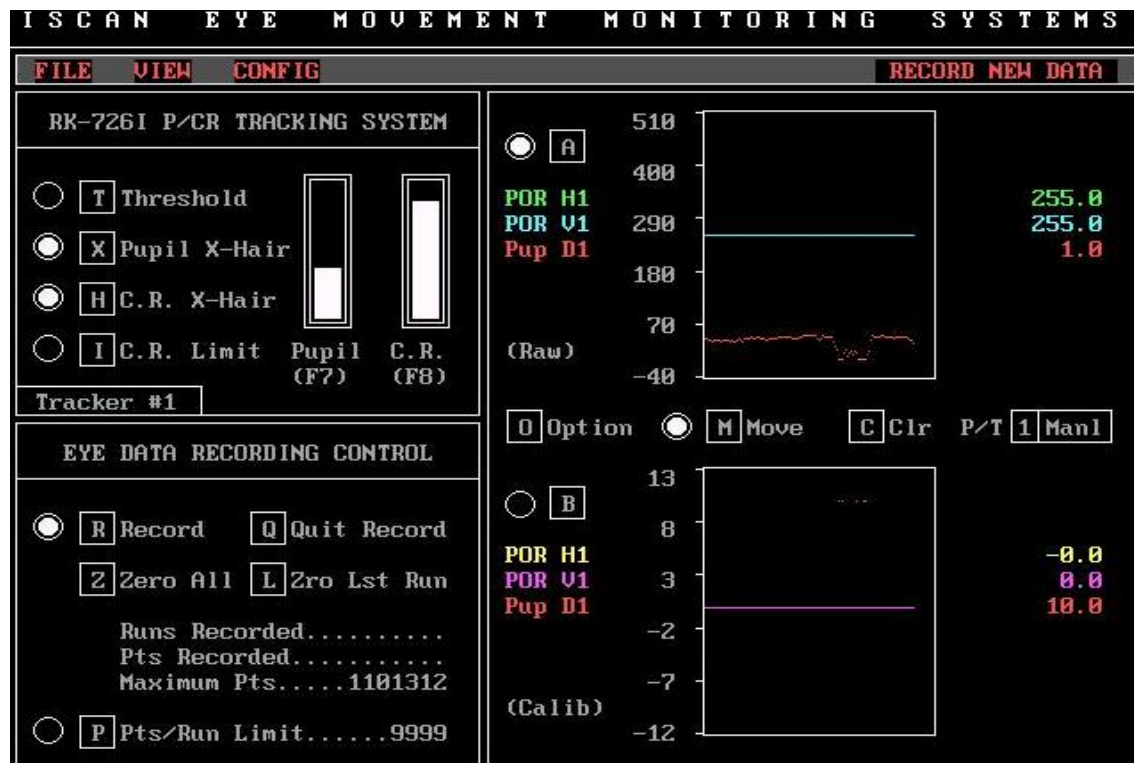


Figure 10: Screenshot of I-Scan system interface

The camera must be adjusted so that the eye is in focus on the eye monitor. You can use Focus manual box for zooming in/out the lens for proper focus and sharp image. As shown in Figure 10, the Pupil X-Hair and the Corneal Reflection (C.R.) X-Hair should always be activated. The Pupil X-Hair is the “+” that is displayed over the pupil. The C.R. X-Hair is the “+” displayed over the white dot at the edge of the pupil.

Activate threshold by pressing ‘T’ and then press ‘F7’. Now increase/decrease threshold by up/down arrow keys until the pupil is filled white. If the rest of the eye turns white, lower the threshold value. Deactivate threshold by pressing ‘T’ again.

Repeat the above procedure for the corneal reflection limit by pressing 'I' key and then 'F8'. This is done to avoid multiple reflections for people who wear contact lens or glasses.

Now turn the camera to automatic mode by pressing 'F1'. You should be able to see two distinct crosses superimposed on the eye on the Eye monitor. The camera now tracks the eye movement. A dot can be seen on Scene monitor that moves in accordance with the eye movement (see Figure 11). The scene monitor also shows the horizontal & vertical position of the eye.

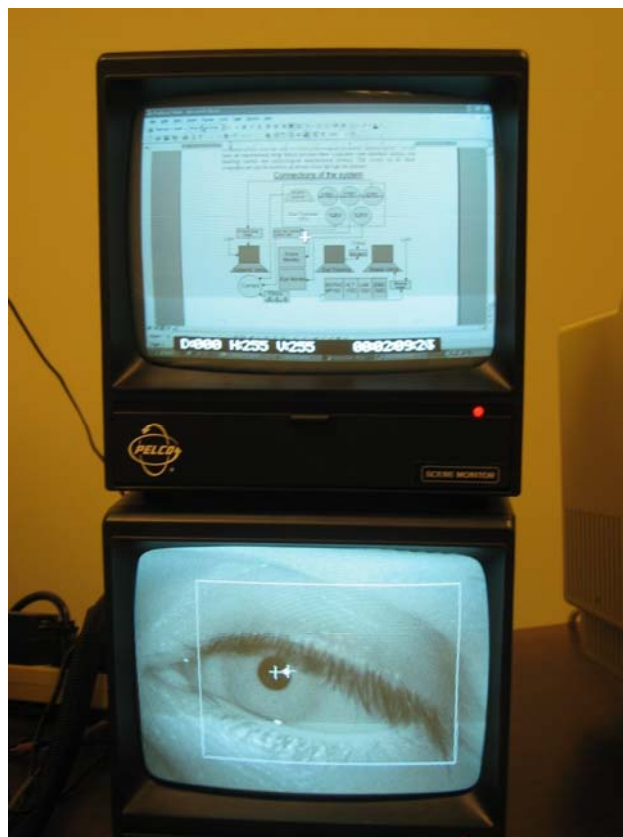


Figure 11: Scene and Eye monitors

4.3.5 Calibrating the I-Scan System

Run the PowerPoint file 'Square' saved on the desktop of the Subject's computer. As shown in Figure 12, this has four cross at the corners of the monitor for calibrating the extreme points.

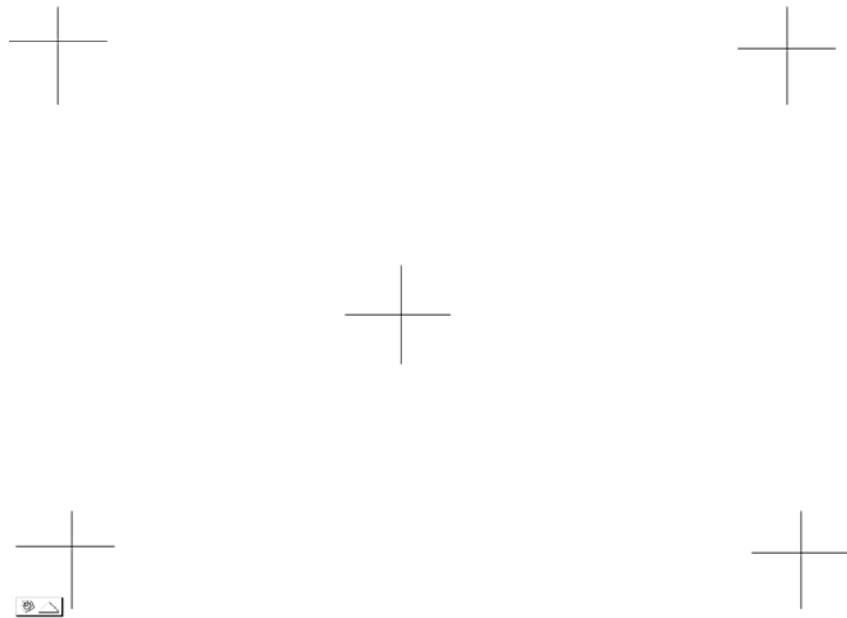


Figure 12: Cross marks used to calibrate the eye tracking system

Press 'O' until you get point of regard (POR) calibrator #1. Press 'M' for move and then 'R' for reset. Press 'L' for calibrating the equipment. Now ask the subject to look at the four corners of the Subjects' computer monitor. Press 'enter' each time he/she looks at those points. The system is now calibrated.

Press 'R' again and then 'P'. You can see the point moving on the small screen in POR calibrator #1. To confirm that everything is working, see if the vertical & horizontal coordinates are changing on the scene monitor when the subject's pupil moves.

Press 'O' until you get the eye data recording control. Zero all previous readings by pressing 'Z' and press 'R' when you want to start recording data. Sample output from the I-Scan file is shown in Figure 13. The number in the rectangle are the start time while the circled number is the total sample numbers.

```
ISCAN DATA FILE 2.06
NONE
NONE
0504 second
1
0 10449 0 60 17 21 19
36 1 256 0.000000
0 -10.000000
512 10.000000
0.039063 0.039063
37 1 256 0.000000
0 10.000000
512 -10.000000
-0.039063 -0.039063
```

Figure 13: Sample output from eye tracking data file

4.3.6 Setting up the Biopac computer

Connect the HLT 100C module to the Biopac MP150 system, as shown in Figure 14. Connect the ethernet port to the ethernet switch box and the LAN wire to MDI of the ethernet switch. Connect the accelerometer to the analog channels of the HLT 100C. Make sure that the Biopac system is switched on, otherwise you will get a message to retry or work offline. If the system is working then the activity (yellow) light will be blinking. Click on the 'Acq373' icon on the Biopac computer desktop.



Figure 14: Biopac MP 150 hardware



Figure 15: Biopac MP 150 system

If there is a problem with the connection, or if the computer is not able to identify the Biopac system, then the activity and busy lights will blink alternatively, as shown in Figure 15. If this occurs, make sure that the DLC protocol in internet connections is activated. If it is not installed then you have to install it by following the instructions in the Biopac manual.

If the computer is still not able to recognize the system, then check the network adaptor card. If you have more than 1 network card then select the appropriate one through the Biopac software. In our case it is 3Com Ethernet link XL 10/100 PCI for complete PC Management.

Finally go to MP150 -> MP 150 Serial Number. Refresh the list and if the system is identified then the serial number of the MP 150 system being used will be added to the list.

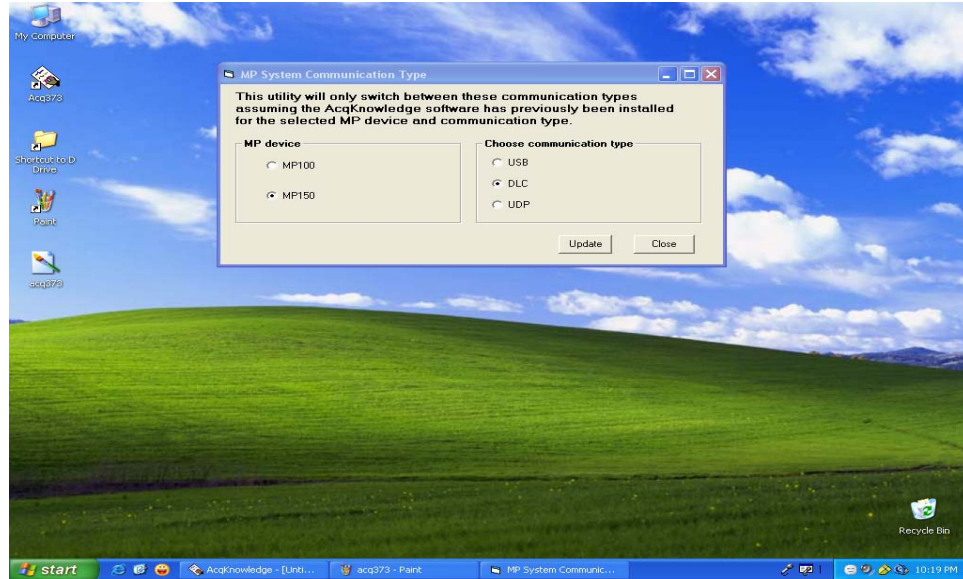


Figure 16: Screen shot showing system calibration I

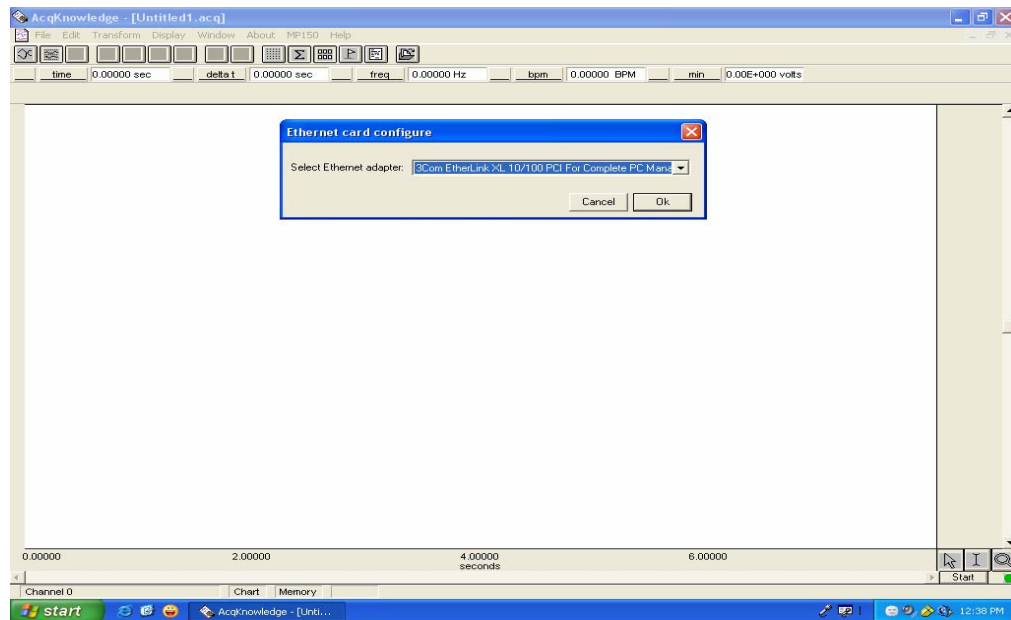


Figure 17: Screen shot showing system calibration II

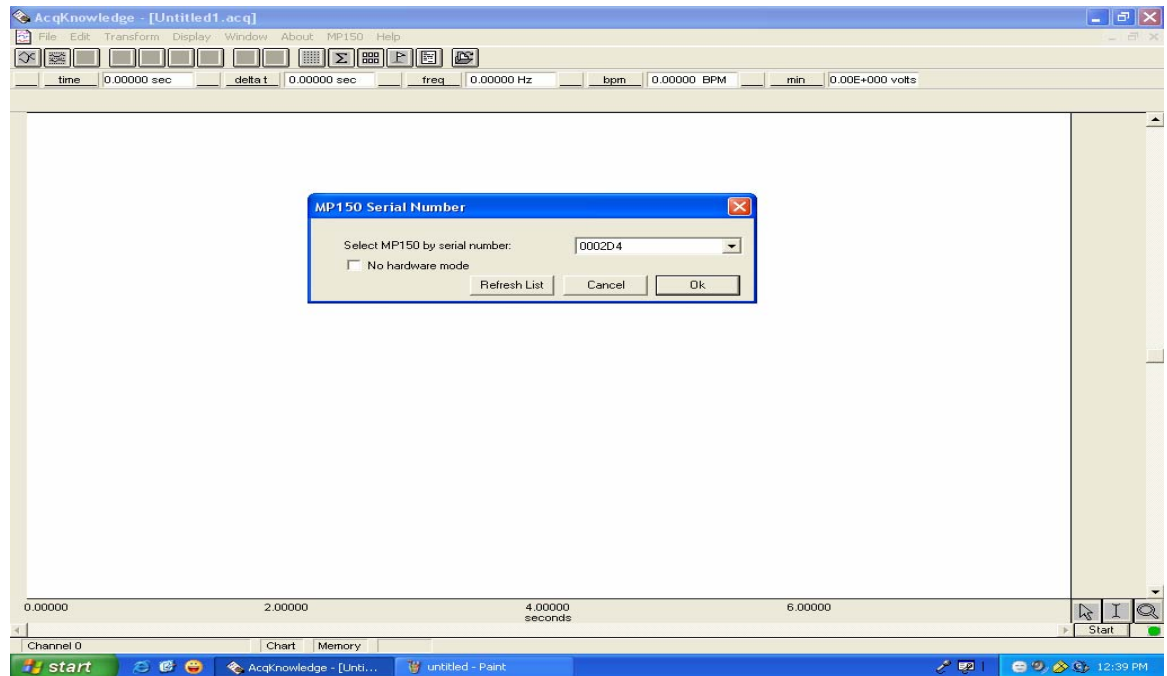


Figure 18: Screen shot showing system calibration III

Note that the Biopac graph is plotted with time in X-axis starting at zero. In order to display actual clock time set the horizontal axis display in HH:MM:SS format by selecting this option through display-> horizontal axis->time (HH:MM:SS).

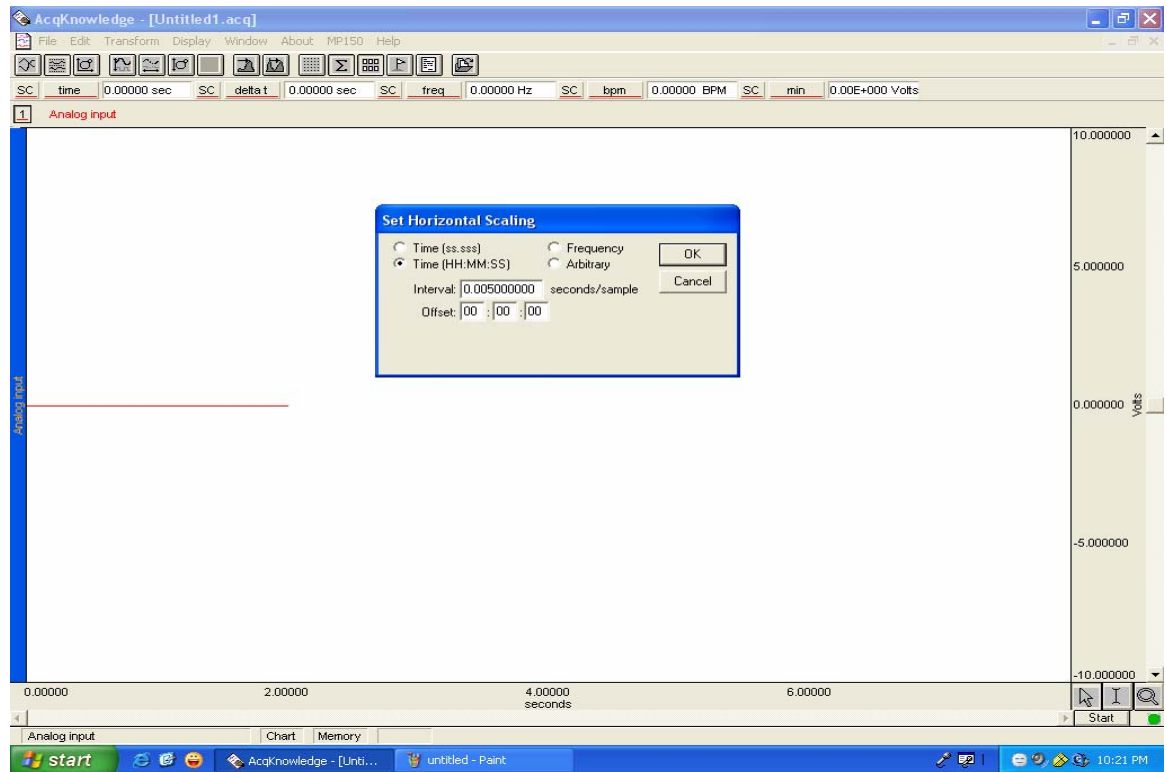


Figure 19: Screen shot showing setting time in real time format

Go to MP150 -> setup acquisition. The default value for data recording is 30 seconds, but it should be changed to at least 5 minutes. Before starting the experiment one should take some trial readings from the accelerometer to verify that it is connected properly and recording data.



Figure 20: Picture of an accelerometer

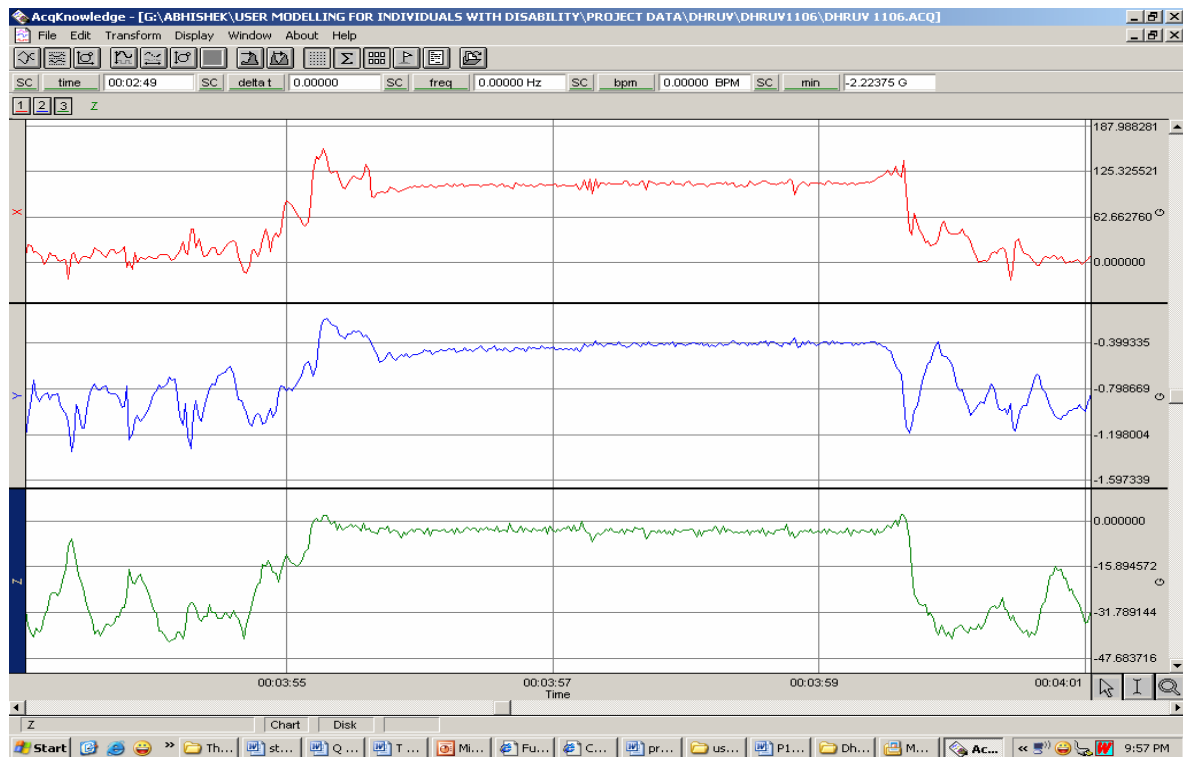


Figure 21: Screen shot showing outputs of X, Y and Z movements

4.3.7 Starting the experiment

Once all the systems are calibrated and the subject is ready for the experiment, tell him/her all the instructions and do not disturb once the experiment starts. The subject should perform the experiment in as natural way as possible so as to avoid any additional cognitive load.

Start the word prediction trial. It takes five seconds before the cursor appears in the transcribed text area. Meanwhile, press 'Ok' in Key Capture, 'R' in eye tracking software & 'start' in Biopac software.

4.3.8 Saving data

At the end of the 5th trial, the subject will press the Enter key to indicate that he/she is done entering text. Immediately stop the eye tracking, Biopac and the Key Capture software. The Java WP test bed stops by itself at the end of the trials.

Now save all the files. A good suggestion when saving data is to name each file with the name of the subject followed by name of the software and the date. Saving in this format will help in distinguishing files when analysis includes all the files together.

Another good suggestion to follow is to stop all the software together and then name and save the data. Stopping and naming a file in one software while the other programs are still running will accumulate unnecessary data which might be difficult to filter out. Save the Biopac file in .acq format as well as in text format.

Place all these files in a new folder named after the subject followed by the condition number and the sentence group chosen.

4.3.9 Create a boundary file for the Iscan software

The boundary file defines the boundaries of important screen regions (e.g., word prediction list, target text and transcribed text). We can see the fixations in these boundaries during analysis by calling on this boundary file. To create a boundary file:

- 1) In the command window open c:\Iscan\Abhsihek. Type pfa1_00b to run 'Analyze Point of Regard'.
- 2) Choose option '6'.
- 3) Type in 'C' to input object boundaries.
- 4) Enter name of the object (like target text, WPL etc).
- 5) Type in the boundary points of the object. First point means upper left corner & second point means lower right corner of the box. (Note that every vertical arrow key press corresponds to 5 pixels & horizontal arrow key press corresponds to 10 pixels).
- 6) Type '*' to stop entering object boundaries.
- 7) Select option '6' and press 'S' to save the file.
- 8) Type in the name with which you want to save the file.

4.3.10 Iscan Data Analysis

To begin Iscan data analysis, exit the daq3_58e software. Type pfa1_00b at the DOS prompt (the folder should be C:\Iscan\Abhishek). Press '1' to assign which file to open and give the full path. Press '2' to see the point of regard data.

```

ANALYZE POINT-OF-REGARD DATA

COMMAND MENU

FIXATION ANALYSIS
<1> Read Raw Point-of-Regard Data File
<2> Review Raw Point-of-Regard Data
<3> Run Fixation Analysis on Raw Data
<4> Fixation Display Options
<5> Save Fixation Data

CORRELATION OF FIXATIONS TO SCENE ELEMENTS
<6> Enter/Recall/Save Scene Element Boundaries
<7> Correlate Fixations to Scene Elements
<8> Correlation Display Options

GENERAL
<9> Save/Recall Analysis Data File
<10> Exit Program

Enter Desired Command Number 1 2 3 4 5 6 7 8 9 or 10 >

```

Figure 22: Screen shot of Point of Regard (POR) data analysis

Make a fixation file (see Figure 23) from the data recorded during the experiment.

```

ISCAN P.O.R. Fixation Data File, Version 1.0

1 23 2005

105

0 20 233 401

31 48 261 368

52 84 171 377

90 120 156 332

121 155 53 298

156 243 41 325

246 284 43 327

361 375 78 288

```

Figure 23: Sample output of eye fixation file

Recall the boundary file and correlate fixations to the boundaries you have defined. Now you can see the eye fixations in the defined regions, as shown in Figure 24.

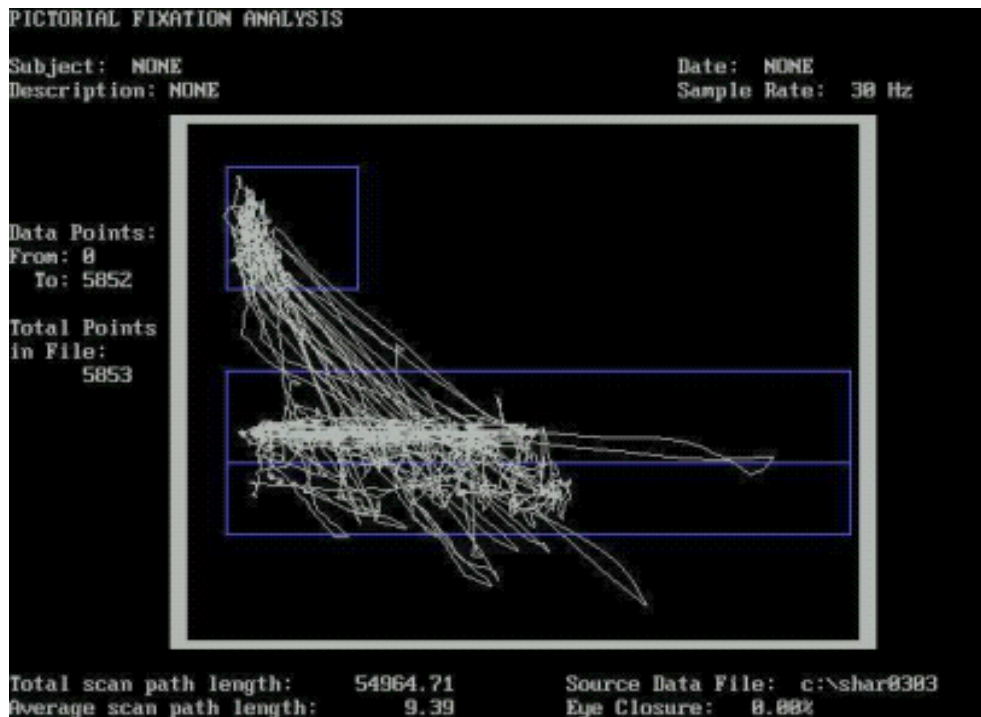


Figure 24: Screen shot of eye fixations

Place the following Matlab files in the same folder where all the data files are saved.

1. final.m
2. wp_read_new.m
3. compare.m
4. list_search.m

The other files already present in this folder are

1. Biopac .acq file
2. Biopac text file
3. Iscan original text file
4. Iscan eye fixation text file
5. Key capture text file
6. Word prediction text file
7. .bmp file of the fixations in the defined boundaries
8. Excel file

4.3.11 Data integration using Matlab codes

Final.m calls all the files described above (key-capture file, Biopac file, eye tracking raw data file and eye fixation file) and reads them in ASCII format. Useful information is recognized and extracted using this code. All the data is aligned chronologically in the ‘output.txt’ file (see Table 2) indicating when the event occurred, the duration of the event and what key was typed when ever the keypress event took place.

Table 2: Sample output of ‘output.txt’ file

Event	When	Duration	Letter typed
Keypress	62218.32	0.08	I
Still	62218.42	0.04	

Table 2 (Continued)

Movement	62218.48	0.08	
WPList	62218.56	0.2	
Keypress	62218.58	0.12	N
Still	62218.72	0.78	
Keyboard	62218.78	1.26	
Movement	62219.52	0.34	
Keypress	62219.88	0.1	1
Keypress	62233.36	0.08	BACK
Still	62233.46	0.1	
Keypress	62233.58	0.06	BACK
Still	62233.66	0.3	
Keyboard	62233.98	1.38	
Keypress	62234.02	0.1	M
Still	62234.14	0.04	
Movement	62234.2	0.08	
Keypress	62234.3	0.1	I
Still	62234.42	0.62	
Movement	62235.06	0.32	
WPList	62235.38	0.18	
Keypress	62235.4	0.08	C
Movement	62235.5	0.18	

This Matlab code also creates a file called ‘event.txt’. This file contains all in-between keypress events. The file called ‘event_eye_only.txt’ contains information about all eye fixation events between keystrokes while ‘event_hand_only.txt’ (see Table 3) contains information about all hand movement events between keystrokes.

Table 3: Sample output of ‘hand events only’

Event		From	To
Keypress->	1	62205.28	62205.74
Still			
Movement			
Keyboard			
Keypress->	2	62205.86	62206.4
Still			
Movement			
Keyboard			
Keypress->	3	62206.52	62206.82
Movement			
Transcribed			

The file ‘event_frequency_and_duration.txt’ (see Table 4) contains information about the frequency of occurrence and the total duration of events for every trial. The events defined are keypress, hand movement, still (when there is no hand movement and no key-presses), target text fixations, transcribed text fixations, WP-list fixations and keyboard fixations (when there is no target, transcribed, or WP-List fixations).

Table 4: Sample output of ‘event frequency and time’ file

freq-S(5)-	40
freq-M(5)-	53
freq-KS(5)-	52
freq-TH(5)-	145
freq-P(5)-	4
freq-T(5)-	2
freq-WP(5)-	1
freq-KB(5)-	6
freq-TE(5)-	13
time_S(1)	20.38
time_M(1)-	13.06
time_KS(1)-	4.48
time_TH(1)-	37.92

Table 4 (Continued)

time_P(1)-	4.84
time_T(1)-	1.9
time_WP(1)-	1.98
time_KB(1)-	34.06
time_TE(1)-	42.78

The file 'sequence.txt' (see table 5) contains information about the frequency of transitions from one event to another within each trial. There were 52 such transitions defined with complexity of as much as 5 event transitions.

Table 5: Sample output of 'transition frequency'

S(1)	41
M(1)	45
KS(1)	49
S-KS(1)	4
M-KS(1)	32
S-M-KS(1)	20
M-S-KS(1)	0
S-M-S-KS(1)	0
M-S-M-KS(1)	0
S-M-S-M-KS(1)	0
M-S-M-S-KS(1)	0

Wp_read_new.m reads the file generated by the WP interface and generates an output file ('WP_stats_output.txt') which contains all the statistics as calculated by the java interface. The output file also contains other calculated data like time between key-presses, text entry and keystroke rates.

'Error file.txt' identifies the missed key-presses. For example every key down event should be followed by key-up, and the subject was not supposed to type in the pause period between two trials. Unexpected occurrences are identified and displayed in this file so that investigators can rectify those to make future analysis easier.

Table 6: Sample output from 'finaloutput_1.txt' file

Event	When	Duration	Letter typed	List displayed	Length	Searched	#	Selected
Keypress	62205.16	0.12	A	f	2	not_srch	0	not_Sltd
Still	62205.3	0.06						
Movement	62205.38	0.36						
Keypress	62205.76	0.1	SPACE	f	2	not_srch	0	not_Sltd
Still	62205.88	0.16						
Movement	62206.06	0.34						
Transcribed	62206.42	0.46						
Keypress	62206.42	0.1	P	f	2	not_srch	0	not_Sltd
Movement	62206.54	0.28						
Keypress	62206.84	0.12	R	f	2	not_srch	0	not_Sltd

A ‘word_final.txt’ file is generated which has all the target text words for all the trials.

‘Word_WP.txt’ has all the words displayed in the WP list. This file is used to identify if the target word was displayed when the subject was transcribing text.

The letters typed by the subject are joined together to form words. This may include numbers which represented the use of WP list and errors that might have occurred during typing. ‘word_sub_typed_final.txt’ contains all the words typed by the subject.

‘len_output.txt’ contains the letters typed by the subject along with true or false, which tells whether the WP list was on or off when that character was typed. It also contains information about the length of the list that showed up when each letter was typed.

Compare.m determines if the target word was displayed in the list when the subject was transcribing text. This is used to classify searches for every letter typed.

Table 7: Sample output from ‘finaloutput_11.txt’ file

Event	When	Duration	Letter typed	List displayed	Length	Searched	#	Word Displayed	Selected
Keypress	62222.68	0.1	E	t	2	not_srch	0	not_disp	not_Sltd
Still	62222.8	0.44							
Movement	62223.26	0.26							

Table 7 (Continued)

Keypress	62223.54	0.1	P	t	2	not_srch	0	not_disp	not_Sltd
Movement	62223.66	0.22							
WPList	62223.82	0.24							
Keypress	62223.9	0.1	H	t	2	SEARCHED	1	DISPLAYED	not_Sltd
Still	62224.02	0.88							
Keyboard	62224.08	1.06							
Movement	62224.92	0.32							
WPList	62225.16	0.28							
Keypress	62225.26	0.12	2	t	2	SEARCHED	1	DISPLAYED	SELECTED
Still	62225.4	0.66							

list_search.m fills in the ‘list_search’ table in an Excel spreadsheet (see Table 8). The spreadsheet contains data about the various types of searches.

Table 8: Format of the list-search table

List is displayed	Yes								No	
User searches list	Yes								Yes	No
Word is in list	Yes		No		Yes		No			
User selects from list	Yes	No	Yes	No	Yes	No	Yes	No		

4.3.12 Data analysis using Matlab codes

ttest_by_condition.m performs a t-test for each dependent variable. The data is arranged in order of conditions and the input file is 'data_by_condition.txt'. For each dependent measure, it compares the mean of two conditions. Based on the p value calculated, it tells whether the two conditions are significantly different.

Table 9: Format for data_by_condition.txt

Subject	Sequence A	Sequence B
G		

Table 9 (Continued)

A		
D		
S		

ttest_by_sequence.m performs a t-test for each dependent variable in which the data is arranged in order of sequence in which the conditions were produced. The input file is 'data_by_sequence.txt'. For each dependent measure, it compares the mean of two conditions and based on the p-value calculated, determines whether there is a significant difference for those two conditions.

Table 10: Format for data_by_sequence.txt

Subject	Sequence A	Sequence B
G		
A		
D		
S		

test_duration.m calculates the time required to complete each condition for every subject.

anova_by_condn.m performs a one-way ANOVA to compare all conditions for each dependent measure. The data is arranged in order of conditions.

Table 11: Layout of table for 1 way ANOVA - condition wise

Subject	Condition																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
G																		
A																		
D																		
S																		

anova_by_seq.m performs a one-way ANOVA to compare between all the conditions for each dependent measure. The data is arranged in order of sequence in which the conditions were produced.

Table 12: Layout of table for 1 way ANOVA - sequence wise

Subject	Sequence																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

Table 12 (Continued)

G																		
A																		
D																		
S																		

anova_word_size.m generates output in a text file that is used to fill the ‘word_size’ table in the excel data analysis sheet. The output has 3 factors: word size, actual list length and trial number. Each factor is compared against another one at a time, thus producing 3 tables.

Table 13: Layout of table for ANOVA for factors word size and actual length

Min Word Size		3		5	
Actual List Length		2	8	2	8
Measure	Subject				
Text Entry Rate	1				
	2				
	3				
	4				

Table 14: Layout of table for ANOVA for factors word size and trial number

Min Word Size		3					5				
Trial		1	2	3	4	5	1	2	3	4	5
Measure	Subject										
Text Entry Rate	1										
	2										
	3										
	4										

Table 15: Layout of table for ANOVA for factors trial number and actual length

Actual list length		2					8				
Trial		1	2	3	4	5	1	2	3	4	5
Measure	Subject										
Text Entry Rate	1										
	2										
	3										
	4										

anova_list_length.m generates a text file that is used to fill the ‘list_length’ table in the excel data analysis sheet. The output has 3 factors: maximum list length size, actual list length and trial number. Each pair of factors is compared, making 3 tables.

Table 16: Layout of table for ANOVA for factors max list length and actual length

Max list length		2			8								
Actual List Length		0	1	2	0	1	2	3	4	5	6	7	8
Measure	Subject												
Text Entry Rate	1												
	2												
	3												
	4												

Table 17: Layout of table for ANOVA for factors max list length and trial number

Max list length		2					8				
Trial		1	2	3	4	5	1	2	3	4	5
Measure	Subject										
Text Entry Rate	1										
	2										
	3										
	4										

Table 18: Layout of table for ANOVA for factors actual length and trial number

Actual list length		0							8
--------------------	--	---	--	--	--	--	--	--	---

Table 18 (Continued)

Trial																		
Measure	Subject																	
Text Entry Rate	1																	
	2																	
	3																	
	4																	

list_length.m runs a two-way repeated measures ANOVA for the independent variable list length. One factor is written in columns while the other factor is written in rows. The data for 4 subjects are written in rows and the value for reps is chosen as 4 (= subject number), which tells the number of observations per cell. It tells that value of the factor in row is varying after every 4 rows.

show_hide.m and **remove_nan.m** run two-way repeated measures ANOVA for the independent variables show and hide. One factor is written in columns while the other factor is written in rows. **remove_nan** identifies which cell doesn't have a value and removes the complete row because we don't want to compare data for an unequal number of subjects for dependent variables.

Table 19: Layout of table for ANOVA for 2 factors arranged in row and column

		Factor 1 – value 1	Factor 1 – value 2
Factor 2 – value 1	Subject 1		
	Subject 2		
	Subject 3		
	Subject 4		
Factor 2 – value 2	Subject 1		
	Subject 2		
	Subject 3		
	Subject 4		

word_size.m runs a two-way repeated measures ANOVA for the independent variable word size. One factor is written in columns while the other factor is written in rows.

5.0 METHODOLOGY

5.1 HUMAN SUBJECTS

Four able-bodied users (male and female) between the ages of 18 and 65 participated. The study did not differentiate based on gender, race or ethnicity. Subjects had no prior experience with WP, but were familiar with a standard (QWERTY) keyboard.

Table 20: Subjects

Subject	Gender	Ethnicity
G	Male	Indian
A	Female	American
D	Male	Indian
S	Male	Indian

5.2 PROTOCOL

Each subject entered text under the 17 experimental conditions shown in Table 22. Each interface parameter has two possible values producing a total of 16 conditions. The 17th experimental condition was without word prediction (referred to as “letters only”). For each experimental condition, subjects were asked to enter five phrases. Subjects were asked to type with a single finger to simulate the typing speed of a physically disabled person.

A corpus of 50 test phrases was compiled from a collection of 500 phrases created by MacKenzie [14]. Each sentence was between 50 and 55 characters long and did not contain any special characters such as capital letters or punctuation marks. These sentences were divided into 10 groups of 5 sentences each. The group of sentences used with each experimental condition was chosen randomly for each participant, but the order of the 5 sentences within each group was fixed.

Subjects were expected to complete 9 conditions in Session 1 and 8 conditions in Session 2. Conditions were randomized and counterbalanced. Each session lasted approximately two hours. Breaks were given between conditions or when requested by the subject.

6.0 DATA ANALYSIS

Five successive trials were conducted for each condition. Each dependent variable was calculated on a per trial basis, to identify any improvement (learning) within a condition. Data was also evaluated to determine whether improvement reached a peak after certain trials. This will help determine how many trials per condition are sufficient to conduct in future studies.

6.1 INDEPENDENT VARIABLES

The independent variables are the parameters which define various configurations of the WP test-bed. The independent variables identified were:

1. Show – The subject has to type in a predetermined number of letters before the word list shows up on the screen. The two values defined for this variable were (1) always display and (2) display when two letters of a word have been typed.
2. Hide – For one setting of this variable we never hide the WP but for the other setting we hide the list after four letters of the target word are typed.
3. Max Length – The length of the list directly affects the list search time. Two values were chosen so that in one case the length is short (2 words) and the other long (8 words).

4. Min Size – Showing the WP list for smaller words may reduce text generation rate (TGR). It might, therefore, be better to type short words without searching the list. Two values were chosen for this variable to study the effect of minimum word size which should be displayed in the list.

Table 21: Independent variable values which define various configurations

Independent Variables	System Configuration	Group 1	Group 2
Show	The number of keystrokes entered per word before the word prediction list appears	0 (word prediction list is always present)	2
Hide	The number of keystrokes entered per word before the word prediction list hides	Never (word prediction list never disappears)	4
Max Length	Maximum number of words in the prediction list	2	8
Min Size	Minimum number of letters in each word in the prediction list	3	5

Based on two values for each of the four independent variables, sixteen conditions were defined. Condition 1 and 18 are without WP and were used to compare the results generated from use of WP system to without WP system use.

Table 22: The values of variables for all the conditions

Condition	Show	Hide	Max Length	Min Size
1 (letters only)	Never	N/A	N/A	N/A
2	0	N	2	3
3	0	N	2	5
4	0	N	8	3
5	0	N	8	5
6	0	4	2	3
7	0	4	2	5
8	0	4	8	3
9	0	4	8	5
10	2	N	2	3
11	2	N	2	5
12	2	N	8	3
13	2	N	8	5
14	2	4	2	3
15	2	4	2	5
16	2	4	8	3
17	2	4	8	5
18 (letters only)	Never	N/A	N/A	N/A

Two other independent variables that are of interest are:

5. Trial – Each condition has five trials, with each trial having a different sentence to transcribe. Testing the same condition with multiple trials provides insight about the effect of learning on subject performance. Previous research [11][12] has shown that

keeping words fixed (static) in the dictionary and practiced produce better results on the same set of conditions.

6. Sequence – The order of conditions can have a confounding effect on results. Suppose, if the best result was achieved with condition number fifteen, then it would be hard to tell whether the result is because of the chosen value of the independent variables for that condition or is it because of enough practice the subject has after having gone through so many trials and conditions. The sequences of conditions were randomized across subjects to limit the effect of practice for a condition which is going to appear towards the end of the test.

Table 23: Showing sequence order in which the conditions were presented to each subject

Sequence order				
Subjects/Sequence	G	A	D	S
1 (letters only)	1	1	1	1
2	15	12	12	3
3	13	11	11	5
4	16	13	13	2
5	9	10	10	13
6	7	7	7	15
7	10	9	9	4
8	11	14	14	6
9	8	8	8	7
10	17	15	15	14
11	6	6	6	16
12	5	5	5	8
13	4	4	4	12
14	12	16	16	9

Table 23 (continued)

15	14	17	17	10
16	2	2	2	11
17	3	3	3	17
18 (letters only)	18	18	18	18

6.2 DEPENDENT VARIABLES

The dependent variables and their sub-categories whose values are calculated from all the trials are described in the following sections.

6.2.1 Keystrokes

Mackenzie and Soukoreff identified four different types of keystrokes [17]:

- Incorrect fixed (IF) – characters that were identified as incorrect and were corrected.
- Incorrect not fixed (INF) – characters that were incorrect but were never corrected. Its value equals to minimum string distance (MSD), which gives the total number of substitution, insertions and deletions required to correct the transcribed text.
- Fixes (F) – arrow, backspace, delete, home and end keys entered to fix mistakes.
- Correct keystrokes (C) – keystrokes that comply with the target text.

It should be noted that these categories assume that there is a 1:1 correspondence between characters and keystrokes. This is not true when using T9 (where more than one key press is required per character) or when using word prediction (when one keystroke can result in multiple characters).

For each trial and each condition, we can then calculate:

- a. $\%CorrectKeystrokes = \frac{C}{C+INF+IF+F}$
- b. The ratio of correct keystrokes to total text - producing keystrokes $= \frac{C}{C+INF+F}$
- c. $\%Error = \frac{INF+IF}{C+INF+IF+F}$
- d. The ratio of errors to total text - producing keystrokes $= \frac{INF+IF}{C+INF+F}$

6.2.2 Keystroke Savings

Keystroke savings is the difference between the amount of text generated (measured in number of characters) and the number of keystrokes. If keystroke savings is 0, then there is a 1:1 correspondence between keystrokes and characters. In most experimental protocols, errors are not allowed, so keystroke savings is never negative.

There are three strings of interest:

- Presented (P) - The target text presented to the user.

- Input Stream (I) - The sequence of keys pressed by the user
- Transcribed (T) - The final sequence of characters that results from the keystrokes

We can calculate:

- keystroke savings compared to letters only $= \frac{T}{I}$
- keystroke savings compared to optimal letters only $= \frac{P}{I}$

6.2.3 Text Entry Rate

Text entry rate is measured as the total number of characters in the transcribed string over the total amount of time. We can calculate the following measures:

- Average text entry rate per trial
- Maximum/minimum text entry rate per trial
- Standard deviation of text entry rate per trial
- Average text entry rate per condition

6.2.4 Keystroke Rate

Keystroke rate is measured as the total number of keystrokes entered over the total amount of time. We can calculate the same measures for keystroke rate as for text entry rate.

6.2.5 Time between keystrokes

The time between keystrokes is measured as the time elapsed between the end of a key-up event and the beginning of the next key-down event. This is an indirect measure of the cognitive and perceptual effort between keystrokes. We expect the time between keystrokes to increase with the number of cognitive and perceptual events and the time per event.

6.2.6 List Searches

We are interested in when the person does or does not search the word prediction list. We can classify the events between each keystroke as follows:

1. Is the list displayed on the screen?
2. Is the target word in the list?
3. Did the user select a word from the list?

Based on these questions, we can calculate the following measures:

- 1) Positive searches: User searched list when word was in the list.
 - a) Successful searches: User selected the target word during a search in which the target word was displayed in list.
 - b) Unsuccessful searches: User did not select the target word during a search in which the target word was displayed in list.
- 2) False Positive searches: User searched the list when the target word was not in the list.
 - a) False word selected: User selected a word during a search when the word was not in the list.
 - b) False word not selected: User did not select a word during a search when the target word was not in the list.
- 3) Negative searches: User did not search list when the target word was not in list.
- 4) False Negative searches: User did not search list when the target word was in list.
- 5) Missing WP
 - a) List searched: User search list when the list is not displayed.
 - b) List not searched: User did not search the list when the list is not displayed.
- 6) Successful anticipation: $(\text{Positive searches} + \text{Negative searches}) / (\text{total searches})$
- 7) Strategy compliance: $(\text{User looked for list when list was displayed}) / (\text{total keystrokes})$

6.2.7 List Search Time

We want to see the time taken for list search during each trial for different factors like

List search time

- a. when the word is present in the list
- b. when the word is not present in the list
- c. when the word is (present + not present) in the list
- d. when the search is successful
- e. when the search is not successful

List search time is important in assessing the cognitive and perceptual load of using WP and the effect of using different configurations (effect of independent variables).

6.2.8 Transcribed and Target Text Views

The number of times a subject looks at the transcribed or target text may provide insight into the cognitive demands of the task. Target text views should increase as cognitive load increases. For both transcribed and target views, we can calculate the following measures:

- a. Average views per trial
- b. Maximum/minimum views per trial
- c. Standard deviation of views per trial

- d. Average views per condition

6.2.9 Sequences of Events

We are interested in two types of events:

- Hand Events
 - Movement between keys
 - Keystroke
 - Still
 - Total hand events
- Eye Events
 - Looking at presented text
 - Looking at transcribed text
 - Looking at word prediction list
 - Looking at the keyboard
 - Total eye events

We assume that only one hand event can happen at one moment. From this assumption it follows that when a keystroke or hand movement is not happening, then the hand must be still. Keystrokes are identified by the Key-Capture software. Hand movements are identified by accelerometer outputs exceeding a predetermined level.

For eye events, we also assume that one, and only one, eye event can be happening at once. We further simplify our calculations by assuming that (1) the “transition time” between fixations on different parts of the screen is minimal and that (2) when the user is not looking at the screen his or her visual attention is on the keyboard. The eye tracking system identifies when the user is looking at the presented (target) text, the transcribed text or the word prediction list.

6.2.10 Test Duration

This is the time taken to conduct all of the trials in each condition.

6.3 ERRORS ENCOUNTERED AND HANDLED

Some errors were encountered during testing. The errors were classified under four categories:

- 1) Systematic errors
- 2) Random errors
- 3) Instrumentation failure
- 4) Human error and unexpected occurrences

6.3.1 Systematic errors

Factors such as practice, prior knowledge and subject variability may confound other independent variables. These unidentified variables account for this category of errors. To eliminate these, careful analysis of the cause for each was identified before the tests were run. To limit subject variability, able-bodied subjects were used. No subject had prior knowledge of WP. To eliminate variable typing speeds, all the subjects were asked to type with a single finger on their dominant hand. Equal practice was given to all the subjects before data was collected. The order in which the conditions were presented to the subjects was randomized.

6.3.2 Random errors

Examples of this kind of error include irregular breaks as requested by the subject and fatigue level of the subjects when they completed the protocol. Tests conducted at early hours of the day as compared to tests conducted during late hours may differ because of varying fatigue levels of the subjects during the day. No remedial action could be taken for these kinds of errors.

6.3.3 Instrumentation failure

This kind of error resulted from instrumentation failure. Due to drift in the Iscan camera, the tests produced erroneous data which was rendered useless when examined against other events. This led to a loss of eye tracking data for the fourth subject (Table 24).

6.3.4 Human error and unexpected occurrences

An interesting incident that occurred during testing was when subject ‘G’ pressed enter key twice for condition 15, thus skipping trial 5 (see Table 24). For subject ‘D’, due to human error the file from the WP interface data could not be saved which led to a loss of condition 4 data for this subject. All the missing data is reported in Table 24.

Table 24: Summary of missing data of the subjects

G		A		D		S	
Condition	Trial	Condition	Trial	Condition	Trial	Condition	Trial
15	5	11	All	4	All WP data	2 to 17	All Eye Tracking data
				18	All		

6.4 STATISTICAL TESTS - PAIR-WISE COMPARISON (T-TEST)

Since this is a pilot study and our main goal is to formulate research questions, evaluate protocols and implement software analysis, a small number of subjects were used. Future studies will involve a greater number of subjects. We assume that all the samples are taken from normal distribution and the samples are randomly distributed.

6.4.1 Why do we need pair-wise comparison?

A two-sample t-test was used in which the means of the conditions for all the four subjects were compared and the probability value calculated to see if the difference in means is statistically significant. The value of type I error (α) was taken as 0.05.

Table 25: Shows sample output from the T-test

Dependent measure																		
c	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
0		4>*	4>*	3>*	4>*	4>*	4>*	4>*	4>*	4>*		4>*	4>*			4>#	4>*	
1				3>@										4<@				
2						4<#		4<#		4<#	3<#			4<*				3<*
3						3<*		3<#		3<*	2<@			3<*				3<*
4						4<*		4<*		4<*	3<*			4<*				3<*
5													4>#	4<@			4>#	3<@
6														4<#				3<@
7													4>#				4>@	3<@
8														4<@				3<#
9													4>*	4<#			4>@	3<@
10														3>#			3>#	
11														4<@				3<#
12														4<*				3<*
13																4>@	4>*	
14																		3<#
15																		3<*
16																		3<#
17																		

Each cell in this table has three different symbols:

- Number – Indicates the number of subject for which the data is available (Note that due to experimental error some of the data is missing).
- < - The measure being compared between conditions is significantly less under the condition at the row than under the condition at the top of the column.
- > - The measure being compared between conditions is significantly greater under the condition at the row than under the condition in the column at the top.
- * - $p < 0.05$

- # - $0.05 < p < 0.10$
- @ - $0.10 < p < 0.15$
- Blank – $p > 0.15$

We wanted to see the variation in dependent variables for the 5th trial alone and the average of all the trials for a condition. We also wanted to see whether the order in which the conditions were presented plays any role in the value of the dependent variable along with the type of conditions provided. In all, we got four such tables for each dependent variable.

6.4.2 T-test analysis for Text Entry Rate

Text entry rate for each subject is shown in Figures 35 through 38. Results from standard t-tests (Table 39 and 40) indicates that there is a significant difference between conditions 1 (103.06 char/min), 4 (57.71 char/min) and 14 (98.49 char/min). Condition 1 is the letters-only condition. Condition 4 and 14 differ by duration and length of the WP list. Condition 4 has a maximum list length of 8 and always shows the list, while the list displays after 2 keystrokes and is removed after 4 keystrokes in condition 14 with a maximum list length of 2.

Table 26: T-test analysis for text entry rate

Text Entry Rate			
5 th		Average	
Best	Worst	Best	Worst
1,14	4	1,14	4

Results from table 26 shows that text entry rate is best for conditions 1 and 14 and worst for condition 4. Interestingly, under condition 4 the WP list was displayed for the maximum amount of time, while under condition 14 the WP list was displayed for the least amount of time among all the conditions. This suggests that the duration of display of the WP list distracts the subject and results in lower text entry rate.

6.4.3 T-test analysis for Keystroke Rate

Keystroke rate for each subject is shown in Figures 39 through 42. Results from a standard t-tests (Table 41 and 42) indicate that there is a significant difference between conditions 1 (107.07 char/min) and 5 (74.01 char/min). Condition 1 is the letters-only condition while condition 5 has the maximum list length of 8 and the list is always shown.

Table 27: T-test analysis for keystroke rate

Keystroke Rate			
5 th		Average	
Best	Worst	Best	Worst
1	5	1	5

6.4.4 T-test analysis for % Correct Keystrokes

The percentage of correct keystrokes for each subject is shown in Figures 27 through 30. Results from a standard t-tests (Table 43 and 44) indicate that condition 4 has the worst % correct keystrokes for the 5th trial (95.16%).

Table 28: T-test analysis for % Correct Keystrokes

% Correct Keystrokes			
5 th		Average	
Best	Worst	Best	Worst
-	4	-	-

6.4.5 T-test analysis for %Error Keystrokes

The percentage of erroneous keystrokes for each subject is shown in Figures 31 through 34. Results from a standard t-tests (Tables 45 and 46) indicate that there is no significant difference between conditions for this measure.

Table 29: T-test analysis for % Error Keystrokes

% Error Keystrokes

Table 29 (continued)

5 th		Average	
Best	Worst	Best	Worst
-	-	-	-

6.4.6 T-test analysis for Time between Keystrokes

The time between keystrokes for each subject is shown in Figures 43 through 46. Results from standard t-tests (Tables 47 and 48) indicate that there is a significant difference between conditions 1 (0.4485 sec), 18, 4 (0.646 sec) and 5. Conditions 1 and 18 are letters-only conditions. Conditions 4 and 5 have the maximum list length of 8 and always showed the list.

Table 30: T-test analysis for time between keystrokes

Time Between Keystrokes			
5 th		Average	
Best	Worst	Best	Worst
1, 18	4,5	1,18	4,5

6.4.7 T-test analysis for List Search Time

List search time for each subject is shown in Figures 71 through 74. Results from a standard t-tests (Tables 49 and 50) indicate that there is a significant difference for conditions 5 (1.5566 seconds) and 17 (1.5466 seconds). Condition 5 has a maximum list length of 8 and always showed the list, while condition 17 also has a maximum list length of 8 but showed the list when 2 keystrokes were typed and was removed when 4 keystrokes were typed. This suggests that list length is one of the major factors in reducing text entry rate.

Table 31: T-test analysis for list search time

List Search Time			
5 th		Average	
Best	Worst	Best	Worst
-	-	-	5,17

6.4.8 T-test analysis for Successful Predictions

Successful prediction for each subject is shown in Figures 127 through 130. Results from a standard t-test (Tables 51 and 52) indicate that condition 5 (frequency = 0.266) showed the best prediction. Condition 5 has the maximum list length of 8 and always showed the list.

Table 32: T-test analysis for successful predictions

Successful Predictions			
5 th		Average	
Best	Worst	Best	Worst
-	-	5	-

6.4.9 T-test analysis for Unsuccessful Predictions

Unsuccessful predictions for each subject are shown in Figures 123 through 126. Results from a standard t-test (Tables 53 and 54) indicate that there is a significant difference between conditions 4 (frequency = 7.466) and 14 (frequency = 6.933). Condition 4 has the maximum list length of 8 and always showed the list, while the list was displayed after 2 keystrokes and was removed after 4 keystrokes in condition 14 with the maximum list length of 2.

Table 33: T-test analysis for unsuccessful predictions

Unsuccessful Predictions			
5 th		Average	
Best	Worst	Best	Worst
14	4	14	4

6.4.10 T-test analysis for Target Views

Target views for each subject are shown in Figures 83 through 86. Results from a standard t-test (Tables 55 and 56) indicate that there is a significant difference for conditions 10, 14 and 16. Conditions 10 and 14 use a list of length 2, while condition 16 displays a list of length 8.

Table 34: T-test analysis for target views

Target Views			
5 th		Average	
Best	Worst	Best	Worst
-	16	10, 14	-

From Table 34, conditions 10 and 14 produce best result for target views. Here “best” means fewer number of views. Both these conditions have a short list, while condition 16 has a longer list length and produce more target views. From this observation, it can be inferred that a longer list length imposes more cognitive load on the subject.

6.4.11 T-test analysis for Transcribed Views

Transcribed views for each subject are shown in Figures 87 through 90. Results from a standard t-tests (Tables 57 and 58) indicate that there is a significant difference between conditions 1, 9 and 18.

Table 35: T-test analysis for transcribed views

Transcribed Views			
5 th		Average	
Best	Worst	Best	Worst
1, 18	9	1, 18	-

From Table 35, we see that transcribed views are worst (implies maximum occurrence) for condition 9, while it is good for the letters-only condition. This implies that the subject is less distracted when using the letters-only condition and so has to make fewer transcribed text fixations.

6.5 MULTI-VARIATE ANALYSIS (ANOVA)

From pair-wise comparisons we can see the effect of conditions on the dependent variables. We know that the building blocks of each condition are four independent variables

(Section 6.1). To determine which independent variables have the greatest effect on performance, a post-hoc analysis by 2-way repeated measure ANOVA (Analysis of Variances) was conducted. The test is 2-way because we are considering two factors at a time and repeated because all the subjects were tested on the same set of conditions [15].

6.5.1 Effect of List Length

From Section 6.1, we see that we are taking two values, 2 and 8 for list length. We want to see if there is any statistical relationship between list search time, maximum list length and actual number of words present in the WP list at a given time. We are comparing these factors against the trial number to see the improvement with each trial (Section 6.1.1). So, out of three factors (trials, maximum list length and actual number of words in WP) we take two factors at a time to run a 2-factor repeated measure ANOVA, thus producing three result tables. The analysis also provides the effect of interaction of both the factors together on the dependent variable.

Table 36: Result for effect of list length from ANOVA2

	Max Length (ML)	Trial (T)	Interaction (ML-T)		Actual Length (AL)	Trial (T)	Interaction (AL-T)		Max Length (ML)	Actual Length (AL)	Interaction (ML-AL)
--	-----------------------	--------------	-----------------------	--	--------------------------	--------------	-----------------------	--	-----------------------	--------------------------	------------------------

Table 36 (continued)

List search time					S				~	S	S
------------------------	--	--	--	--	---	--	--	--	---	---	---

S – Significant, ~ - Almost significant

Results from Table 36 show that list search time is dependent on actual list length when tested against trial ($p = 0.00$) as well as maximum list length ($p = 0.0298$). It is also significant for the interaction ($p = 0.0184$) between actual list length and maximum list length. However it shows trend (slight significance ($p = 0.0507$)) for maximum list length.

6.5.2 Effect of Minimum Word Size

We want to see if is there any statistical relationship between trials, word size and maximum list length for the given dependent variables (Section 6.1.2). Three 2-factor repeated measure ANOVA tests were performed using two factors at a time among trials, word size and maximum list length. The analysis also provides the effect of interaction of both the factors together on the dependent variable.

Table 37: Results for effect of minimum word size from ANOVA2

	Word Size (WS)	Max List (ML)	Interaction (WS-ML)	Word Size (WS)	Trial (T)	Interaction (WS-T)	Max List (ML)	Trial (T)	Interaction (ML-T)
Text entry rate					~		S	~	
Keystroke rate							S		
Correct keystrokes									
Incorrect keystrokes									
List search time		S					S		
Successful predictions									
Unsuccessful predictions									
Target views				S	S			S	
Transcribed views				S					

S – Significant, ~ - Almost significant

Results computed from Table 37 for the following dependent measures were:

- Text entry rate – maximum list length shows significance for text entry rate when viewed against trials ($p = 0.025$), while trial shows trend for text entry rate when viewed against both word size ($p = 0.0519$) and maximum list size ($p = 0.067$).
- Keystroke rate – maximum list length shows significance for keystroke entry rate when viewed against trials ($p = 0.027$).
- Correct keystrokes – no statistical significance.
- Incorrect keystrokes – no statistical significance.
- List search time – maximum list length shows significance for list search time when viewed against word size ($p = 0.033$) as well as trials ($p = 0.0005$).
- Successful predictions – no statistical significance.
- Unsuccessful predictions – no statistical significance.
- Target views – trial shows significance when viewed against maximum list size ($p = 0.0242$) and word size ($p = 0.0154$). Word size shows significance when viewed against trials ($p = 0.0228$).
- Transcribed views – word size shows significance for transcribed views when viewed against trials ($p = 0.007$).

6.5.3 Effect of Show and Hide

To determine if there is any statistical relationship between trials, show and hide on the dependent variables, 2-factor repeated measure ANOVA tests were performed taking two factors

at a time among trials, show and hide. The analysis also provides the effect of interaction of both factors together on the dependent variable.

Table 38: Result for effect of show and hide from ANOVA2

	Show (S)	Hide (H)	Interaction (S-H)		Show (S)	Trial (T)	Interaction (S-T)		Hide (H)	Trial (T)	Interaction (H-T)
Text entry rate										~	
Keystroke rate											
Correct keystrokes											
Incorrect keystrokes											
List search time					S						
Successful predictions											
Unsuccessful predictions					S						
Target views						S			S	S	
Transcribed views											

S – Significant, ~ - Almost significant

Results computed from Table 38 for the following dependent measures were:

- Text entry rate – no statistical significance.
- Keystroke rate – no statistical significance.
- Correct keystrokes – no statistical significance.
- Incorrect keystrokes – no statistical significance.
- List search time – ‘show’ factor shows significance for list search time when viewed against trials ($p = 0.0231$).
- Successful predictions – no statistical significance.
- Unsuccessful predictions – ‘show’ factor shows significance for unsuccessful predictions when viewed against trials ($p = 0.0252$).
- Target views – trial shows significance when tested with hide ($p = 0.0306$) as well as show ($p = 0.012$). Hide shows significance when viewed against trial ($p = 0.0075$).
- Transcribed views – no statistical significance.

7.0 DISCUSSION

7.1 SUBJECT ‘G’

Subject ‘G’ rarely used the WP list to fill in words. Figure 54 shows that subject ‘G’ had the most instances of not searching the list in spite of the target word being displayed. If we look at conditions 2, 4, 8, 11 and 16 more closely, we find that for these conditions the false negative searches (user did not search the list when the word was in the list) were almost constant for subject ‘G’. The time between keystrokes for these conditions in Figure 44 is also constant (and least among all the subjects). Figures 100 and 102 shows that subject ‘G’ had the maximum number of keystrokes and keystroke duration for all the conditions. The consistency of data for each measure shows that subject ‘G’ followed a fixed strategy of typing most of the letters without being distracted by WP.

Subject ‘G’ had the most transcribed text fixations (except subject ‘D’ – Figure 88) and keyboard fixations (except subject ‘A’ – Figure 116) which also imply that this subject found the letters-only condition more favorable. Conditions 2, 4, 8 and 16 have a minimum word size of 3. For these conditions the time between keystrokes, negative searches and keystrokes are fairly constant. But for other conditions (in which the minimum word size is 5) these variables show fluctuation in values. Figure 114 shows that for condition 4 this subject had the longest WP list fixation time in spite of having fewer eye fixations on the WP list (Figure 112). This implies that this subject experiences the highest amount of cognitive load when searching through the list. Note that condition 4 displays the WP list for the maximum amount of time, since it shows up at

the start of the event and doesn't hide until the word is completed. The length of the list is largest (8) and it appears for minimum word size of 3, thus contributing to maximum display time of the list on the subjects' computer.

This subject had the highest percentage error (Figure 32) for most of the conditions. So, this shows another disadvantage of not using WP. When using WP, the subject can be less concerned about spelling and also the fewer the keys typed the lesser the chances of committing mistakes.

7.2 SUBJECT 'A'

Contrary to subject 'G', subject 'A' used the WP list the greatest number of times. Figure 114 show that subject 'A' had the longest time duration for WP list views (except for condition 4, as discussed above). Subject 'A' searched the WP list the largest number of times, thus giving the most successful and false positive searches (Figures 48 and 60, respectively). Figure 68 shows that the list search time even when the target word was not there was greatest for subject 'A' over all the conditions, while Figures 112 and 114 show that this subject has far more WP list view fixations than any other subject for all the conditions. This shows that this subject was trying to use WP the most often at the expense of increased cognitive load and time (Figure 131). For this reason 'time between keystrokes' shows maximum fluctuation for this subject, because this subject often switched back and forth between the target text, WP list and keystrokes.

Figure 131 shows that subject ‘A’ had a longer duration of test completion in almost every condition. And in conditions without WP he/she has lower time-duration than other subjects. This supports the empirical evidence from Section 1.2 that WP does not necessarily increase text generation rate and may, in fact, reduce it.

7.3 SUBJECT ‘D’

Subject ‘D’ followed the mid path between subjects ‘G’ and ‘A’ in terms of strategy. This subject used his/her judgment to use the WP list when it was most appropriate. Figures 36 and 40 show that the text entry and keystroke rate of this subject was highest amongst all the subjects. This subject adopted a flexible strategy for text entry. For this reason he/she showed the greatest number of target and transcribed fixations and the maximum duration for these fixations (Figures 84, 88, 108 and 110).

7.4 REMARKS

Each subject followed a different strategy. One relied more on WP (subject ‘A’), another used WP to the minimum extent (subject ‘G’) and a third subject (subject ‘D’) used WP differently across conditions. This presents an opportunity to study the effect of strategy and

configuration on performance. (Note that all the subjects were given an equal amount of practice and were allowed to use the system as they wished).

Some questions that this research study addressed are:

- 1) How is typing accuracy affected by different configurations and user characteristics?

Figures 28 and 32 show that the percentage of correct keystrokes is lowest for subject 'G' while the percentage of erroneous keystrokes is highest. These same figures show that subject 'A' produced the highest percentage of correct keystrokes and the lowest percentage of erroneous keystrokes among all the subjects. Interestingly subject 'G' used WP infrequently, while subject 'A' used WP very frequently. There may, therefore be a relationship between accuracy and the use of WP. The increased accuracy associated with WP could be a result of typing fewer keystrokes.

- 2) How can cognitive and perceptual time and effort be evaluated?

In this study, we evaluated cognitive and perceptual time and effort by eye fixations. The frequency and duration of eye fixations are considered important elements for representing cognitive and perceptual loads, with longer eye fixations indicating greater load on the subject. The frequency of fluctuations from one event to another and the time between consecutive keystrokes also serve as indicators of cognitive load.

- 3) What determines how long it takes a person to search the word prediction list?

Results from paired T-tests (Section 6.3.1) show that the actual length of the list significantly affects the list search time. The length of words in the list, whether the list was displayed right away, and whether the search was successful or not did not significantly affect list search time.

7.5 IMPLICATIONS FOR FUTURE RESEARCH

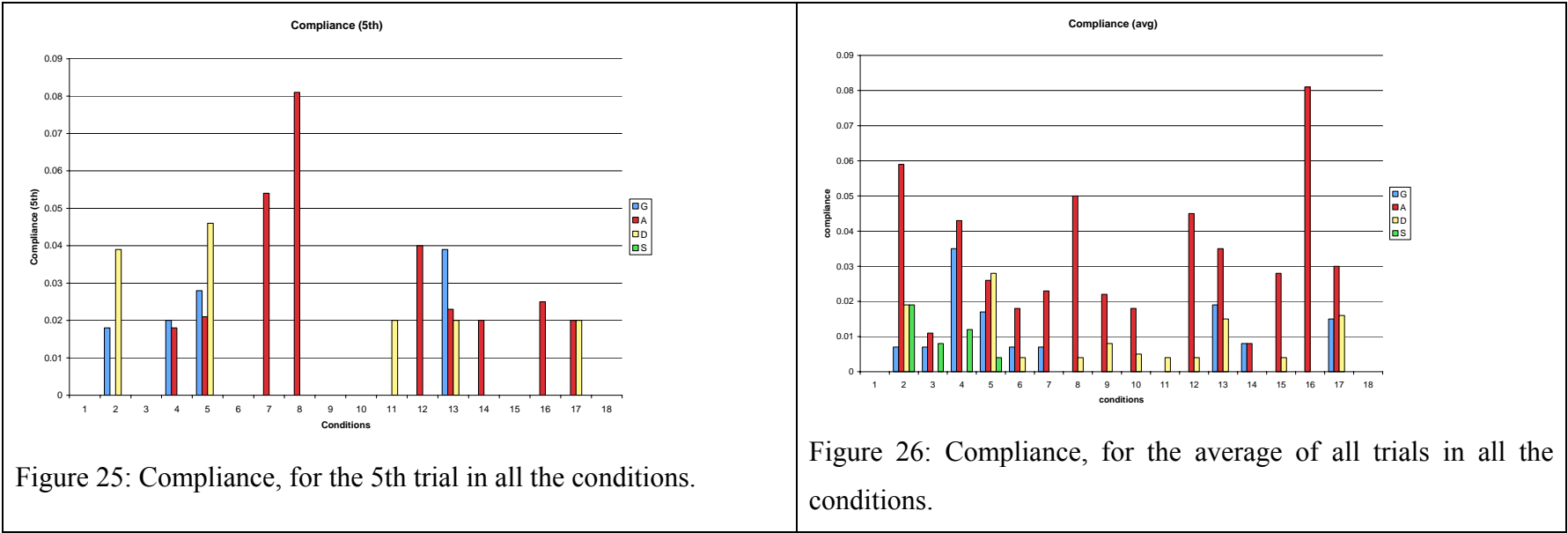
Since the main aim of this project was to prepare for future research work, some interesting observations are suggested that may help in future studies.

1. In this study, WP did not confer any advantage over letters-only typing because we were testing on AB subjects whose typing speeds are much faster than a disabled person. Although slowed the text entry rate by asking subjects to type with a single finger, the rate should be reduced further to imitate the typing speed of a disabled person more closely.
2. One of the goals of the research was to determine how many trials were sufficient for each condition. From multivariate analysis, it is clear that trials do not have any significant effect on the dependent variables (except for target views). Future tests, therefore can use 5 sentences per condition.

3. List search time is dependent on the length of the list. To evoke a difference in list search time, lists of length 2 and 8 are good choices for test conditions.
4. Subjects were unable to accurately predict when a target word would appear in the WP list. If future studies wish to explore the effect of prediction accuracy on performance, subjects will need to be given help in making accurate predictions. The help can be in the form of a visual or audible cue. For a visual cue, the target word in the target text box or in the WP list box can have its background colored or boxed to indicate that the target word is present.

APPENDIX A

FIGURES



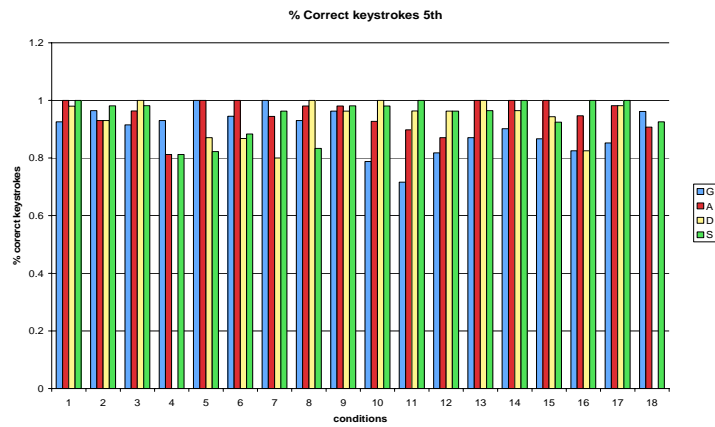


Figure 27: Percentage correct keystrokes for 5th trial only in all the conditions. The data is presented in order of condition.

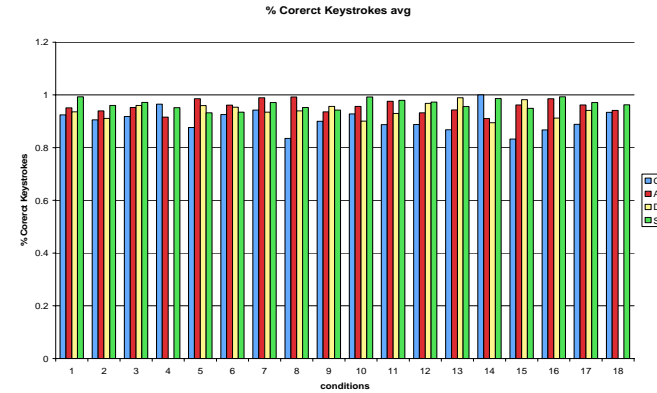


Figure 28: Percentage correct keystrokes for the average of all trials in all the conditions. The data is presented in order of condition.

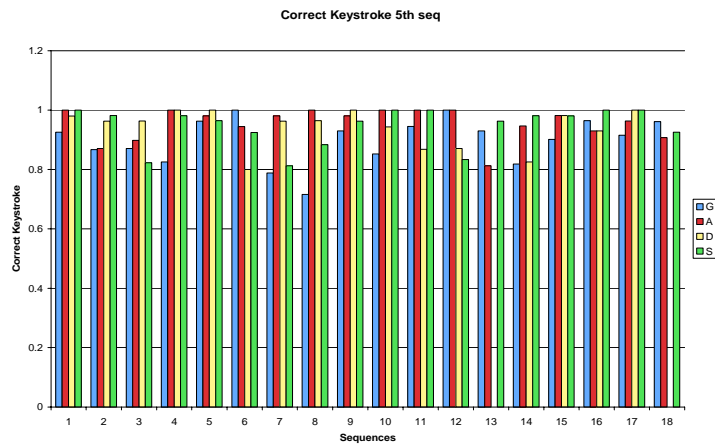


Figure 29: Percentage correct keystrokes for 5th trial only in all the conditions. The data is presented in order of sequence.

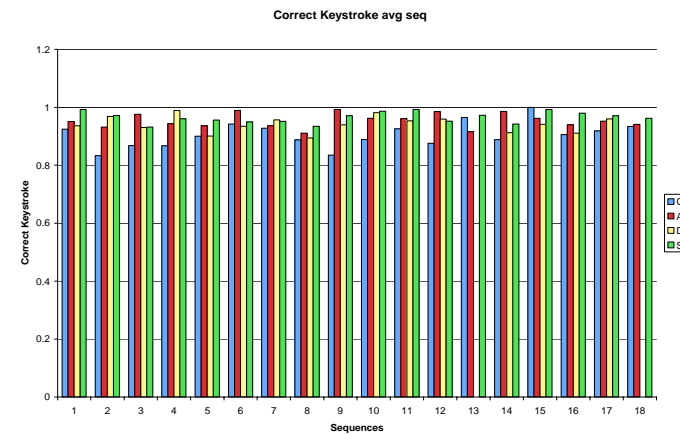


Figure 30: Percentage correct keystrokes for the average of all trials in all the conditions. The data is presented in order of sequence.

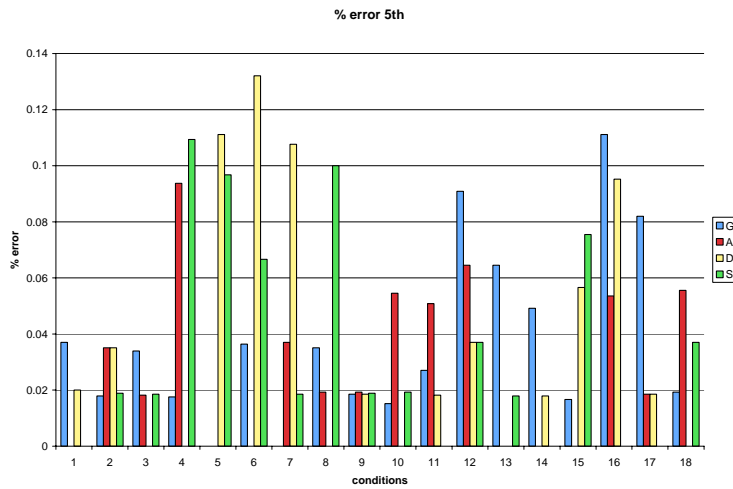


Figure 31: Percentage error keystrokes for 5th trial only in all the conditions. The data is presented in order of condition.

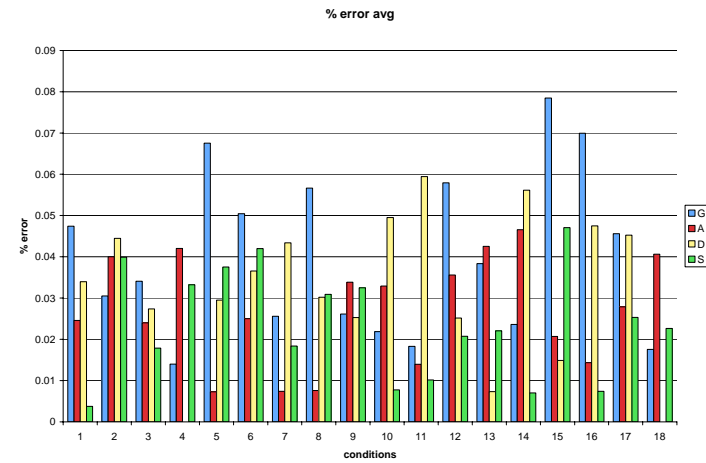


Figure 32: Percentage error keystrokes for the average of all trials in all the conditions. The data is presented in order of condition.

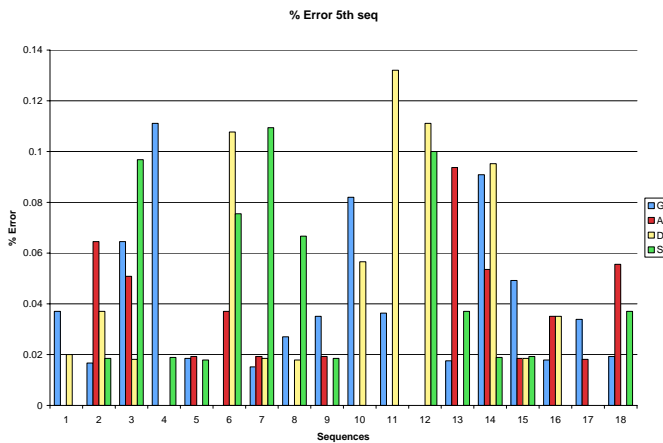


Figure 33: Percentage error keystrokes for 5th trial only in all the conditions. The data is presented in order of sequence.

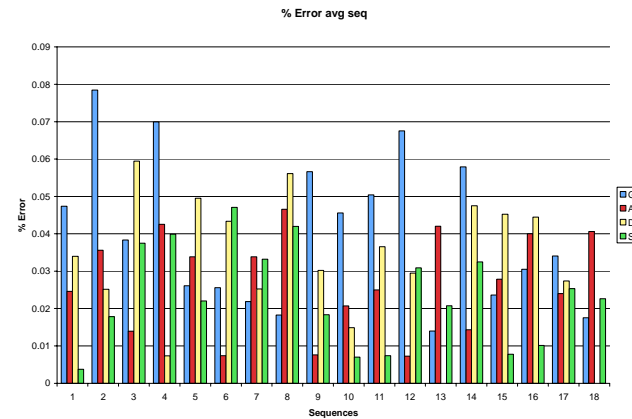


Figure 34: Percentage error keystrokes for the average of all trials in all the conditions. The data is presented in order of sequence.

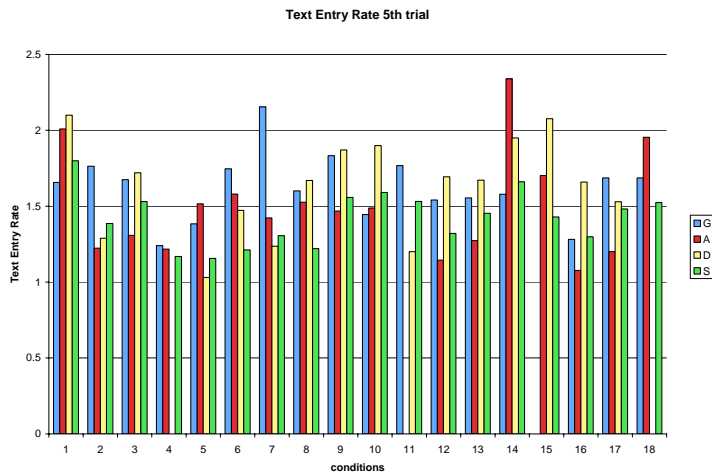


Figure 35: Text entry rate for 5th trial only in all the conditions. The data is presented in order of condition.

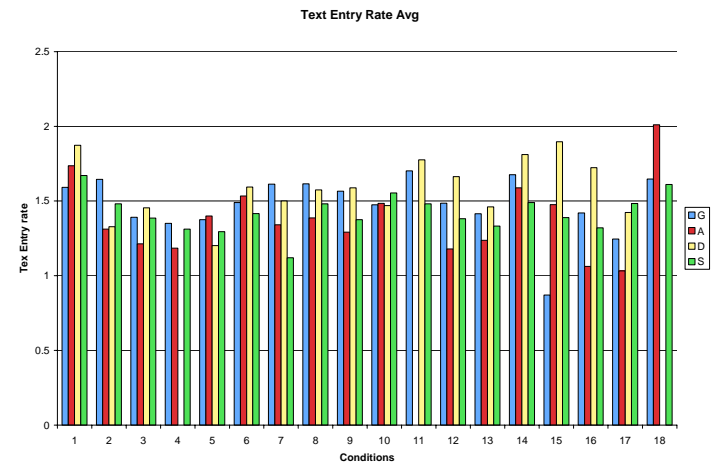


Figure 36: Text entry rate for the average of all trials in all the conditions. The data is presented in order of condition.

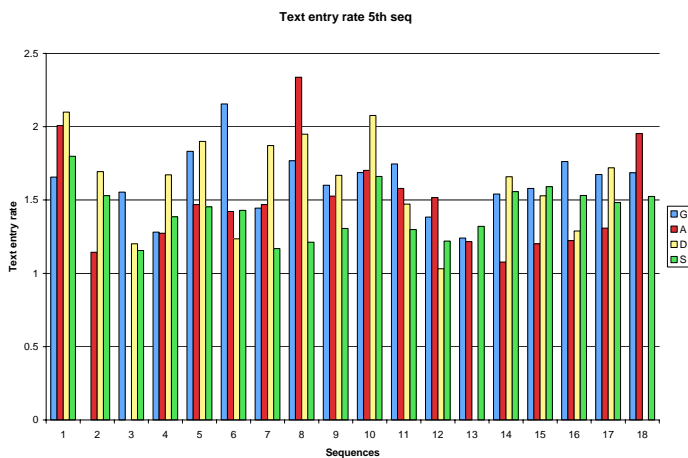


Figure 37: Text entry rate for 5th trial only in all the conditions. The data is presented in order of sequence.

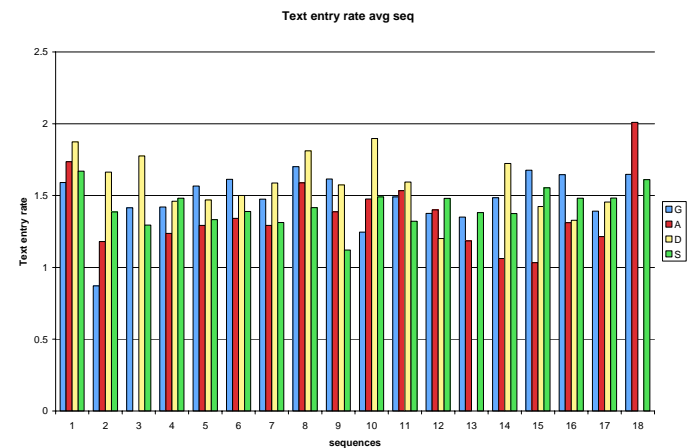


Figure 38: Text entry rate for the average of all trials in all the conditions. The data is presented in order of sequence.

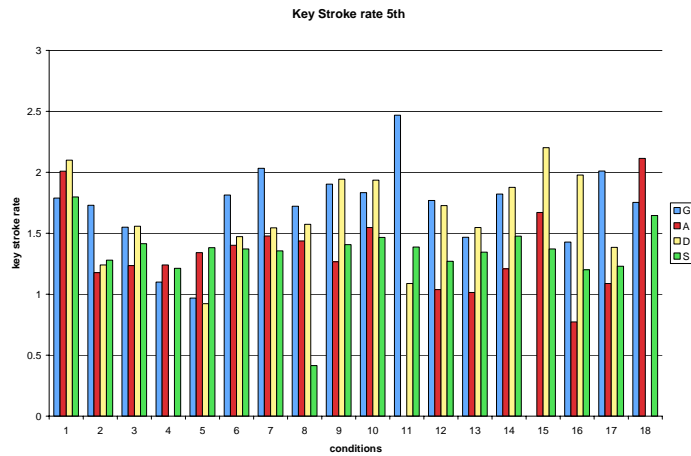


Figure 39: Keystroke rate for 5th trial only in all the conditions. The data is presented in order of condition.

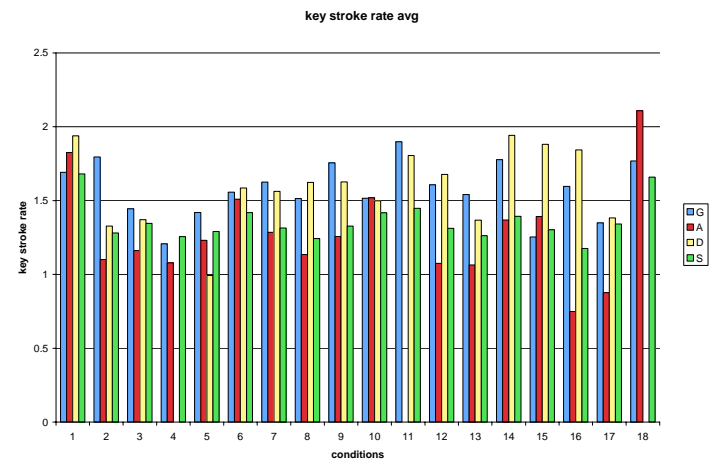


Figure 40: Keystroke rate for the average of all trials in all the conditions. The data is presented in order of condition.

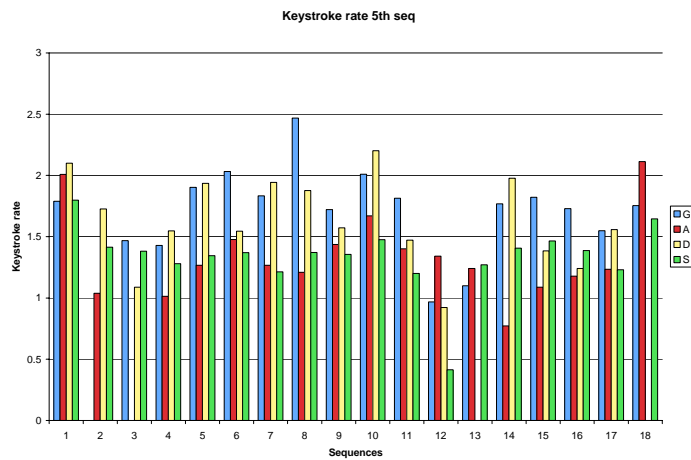


Figure 41: Keystroke rate for 5th trial only in all the conditions. The data is presented in order of sequence.

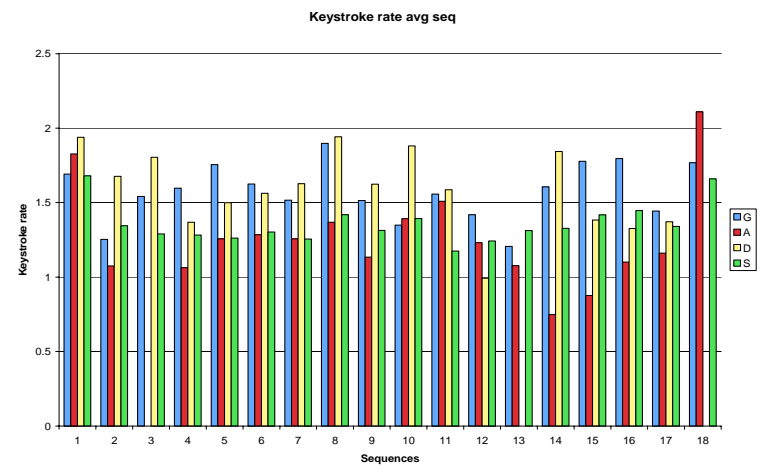


Figure 42: Keystroke rate for the average of all trials in all the conditions. The data is presented in order of sequence.

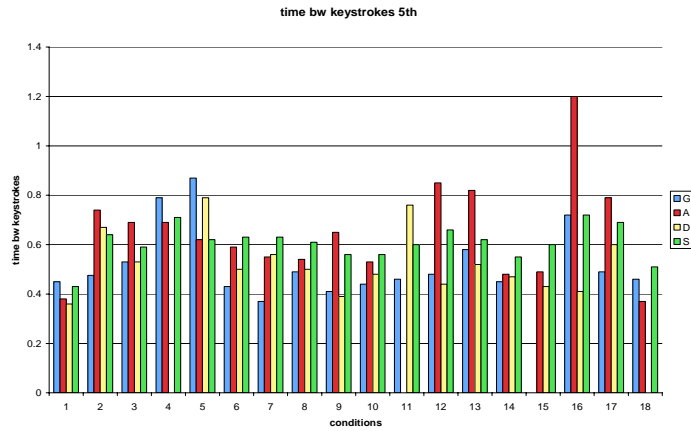


Figure 43: Time between keystrokes for 5th trial only in all the conditions. The data is presented in order of condition.

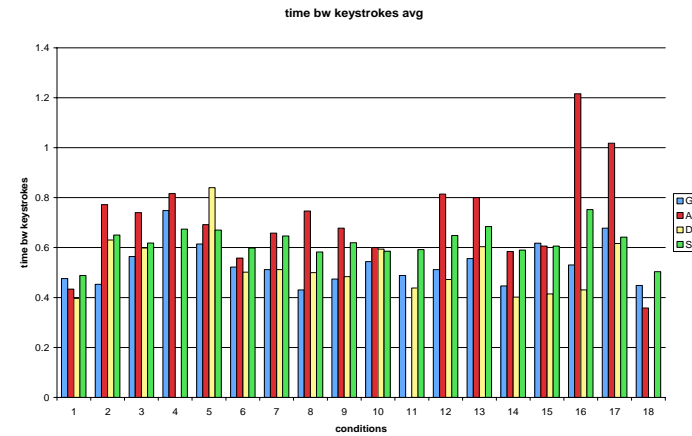


Figure 44: Time between keystrokes for the average of all trials in all the conditions. The data is presented in order of condition.

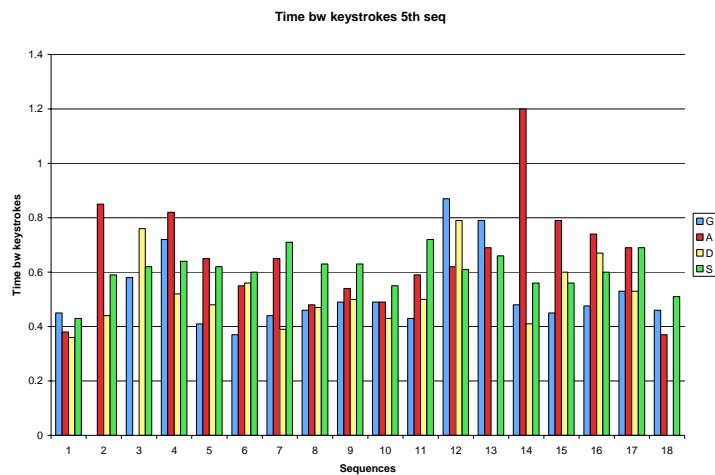


Figure 45: Time between keystrokes for 5th trial only in all the conditions. The data is presented in order of sequence.

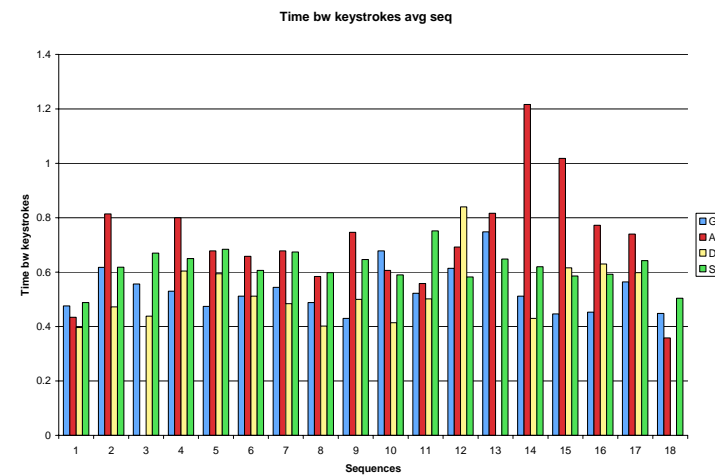


Figure 46: Time between keystrokes for the average of all trials in all the conditions. The data is presented in order of sequence.

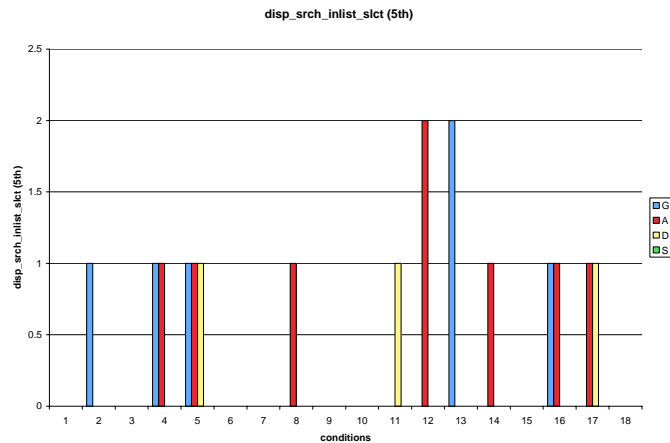


Figure 47: User selected the word when it was displayed in the list, for the 5th trial in all the conditions.

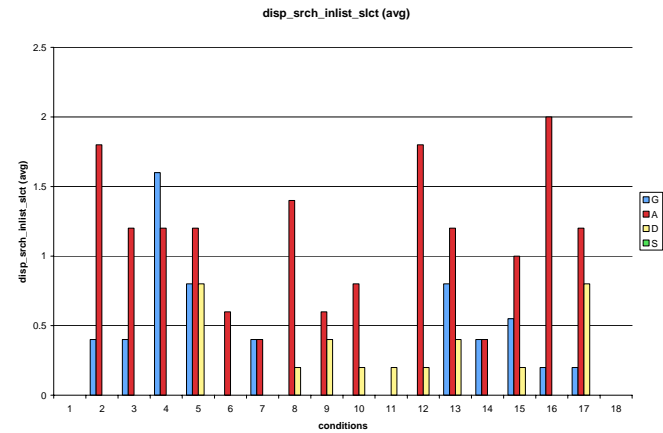


Figure 48: User selected the word when it was displayed in the list, for the average of all trials in all the conditions.

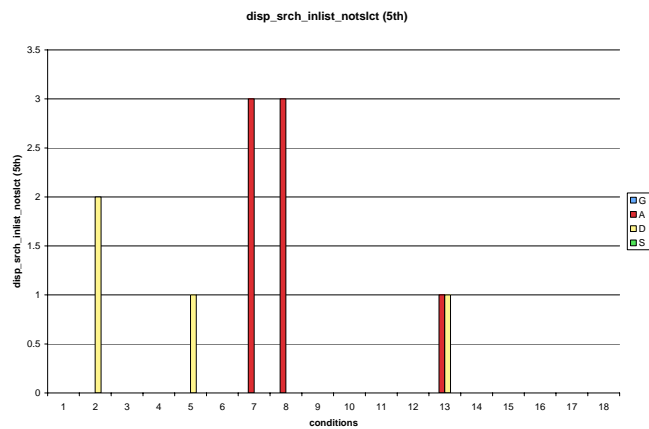


Figure 49: User did not select the word when the target word was displayed in the list, for the 5th trial in all the conditions.

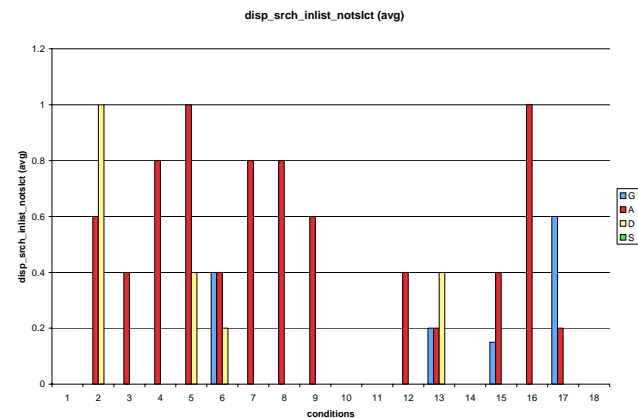


Figure 50: User did not select the word when the target word was displayed in the list, for the average of all trials in all the conditions.

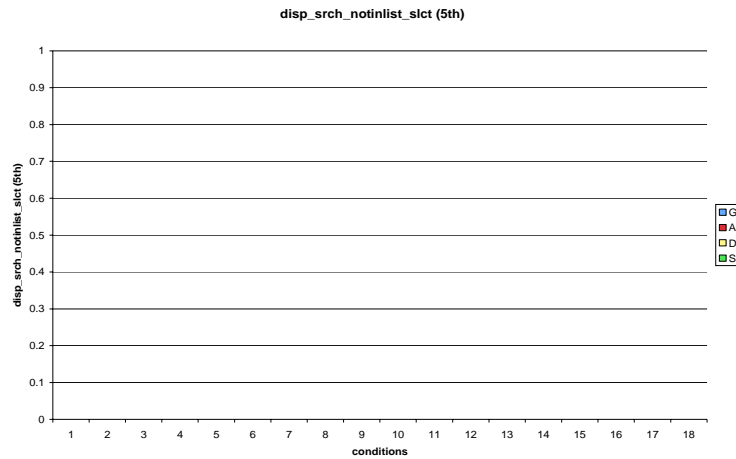


Figure 51: User selected the word when the target word was not displayed in the list, for the 5th trial in all the conditions.

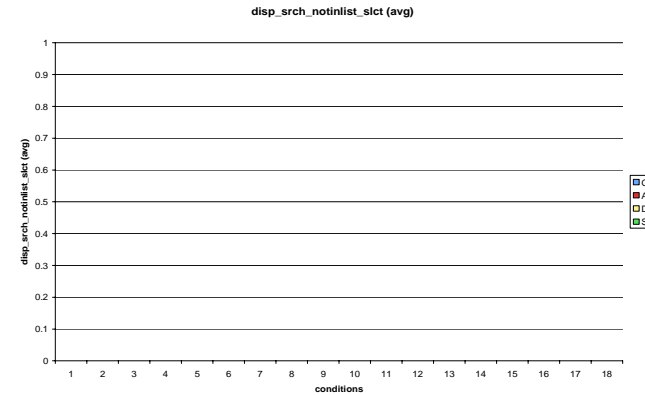


Figure 52: User selected the word when the target word was not displayed in the list, for the average of all the trials in all the conditions.

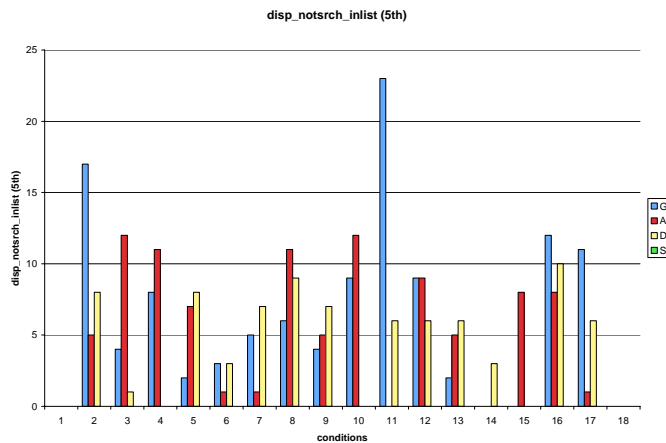


Figure 53: User did not search the word when the target word was there in the displayed list, for the 5th trial in all the conditions.

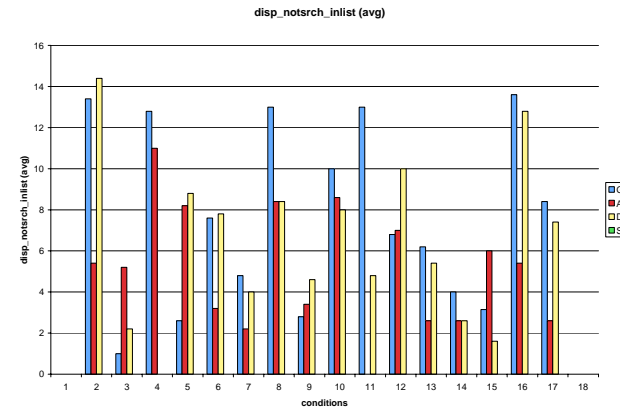


Figure 54: User did not search the word when the target word was there in the displayed list, for the average of all the trials in all the conditions.

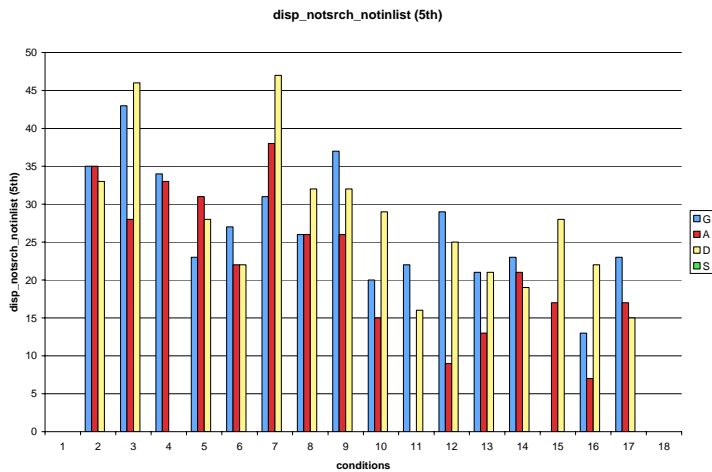


Figure 55: User did not search the word when the target word was not there in the displayed list, for the 5th trial in all the conditions.

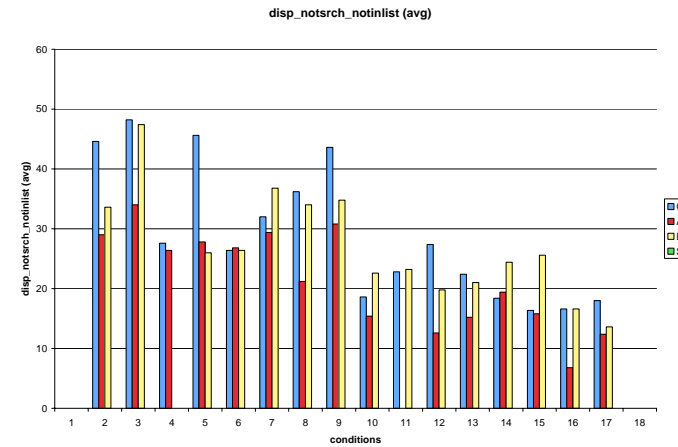


Figure 56: User did not search the word when the target word was not there in the displayed list, for the average of all the trials in all the conditions.

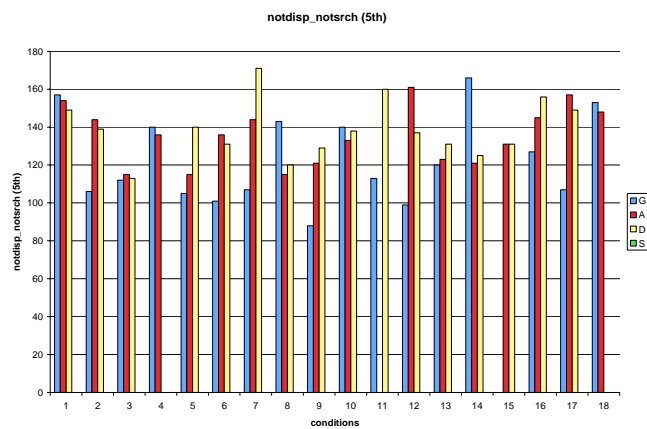


Figure 57: User did not search the word when the list was not displayed, for the 5th trial in all the conditions.

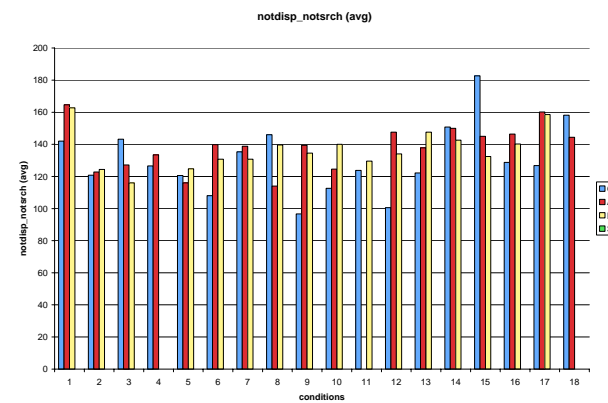


Figure 58: User did not search the word when the list was not displayed, for the average of all trials in all the conditions.

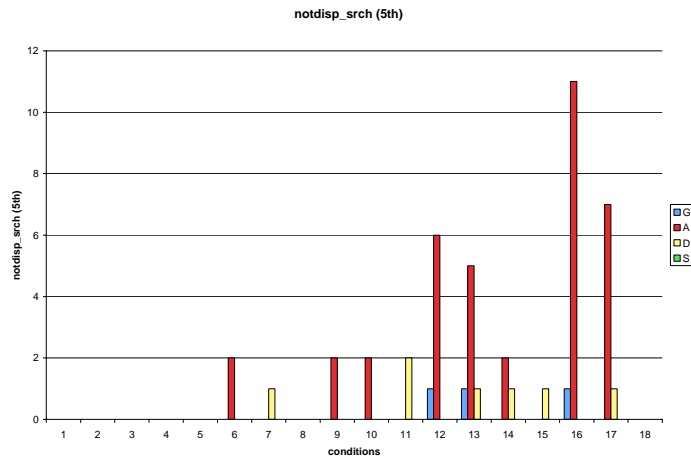


Figure 59: User searched the word when the list was not displayed, for the 5th trial in all the conditions.

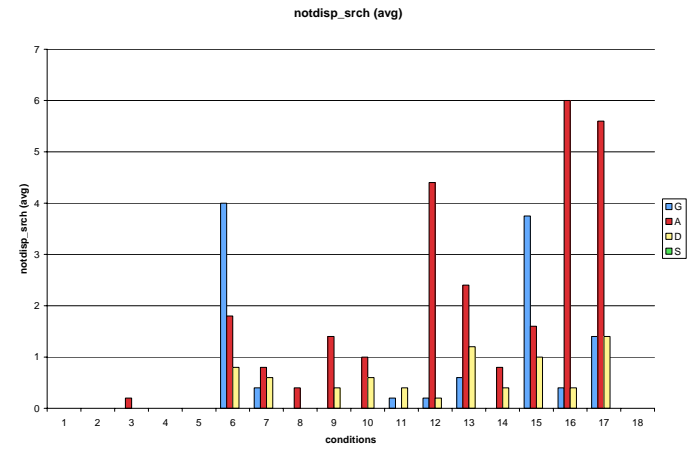


Figure 60: User searched the word when the list was not displayed, for the average of all trials in all the conditions.

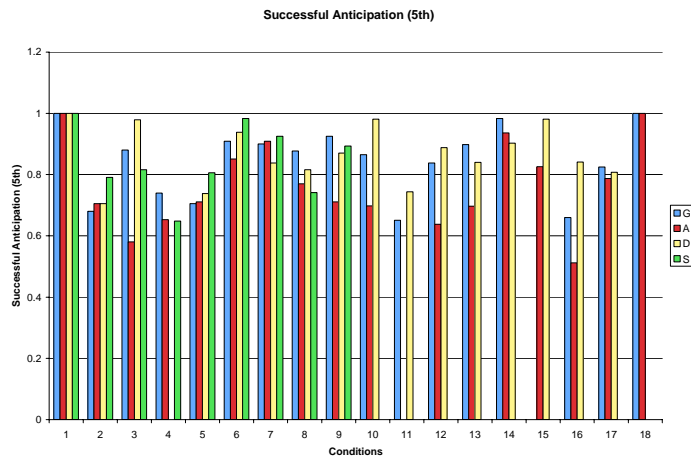


Figure 61: Successful anticipation, for the 5th trial in all the conditions.

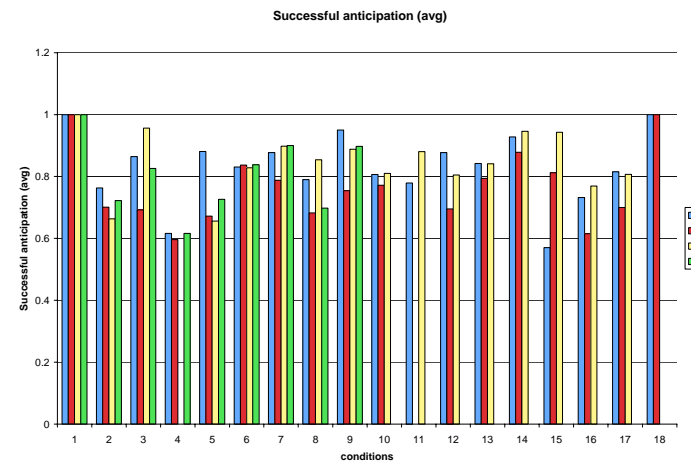


Figure 62: Successful anticipation, for the average of all trials in all the conditions.

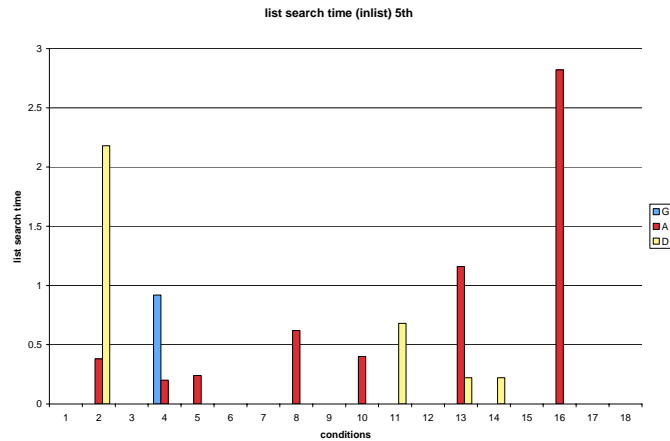


Figure 63: List search time when the target word is in the list, for 5th trial only in all the conditions. The data is presented in order of condition.

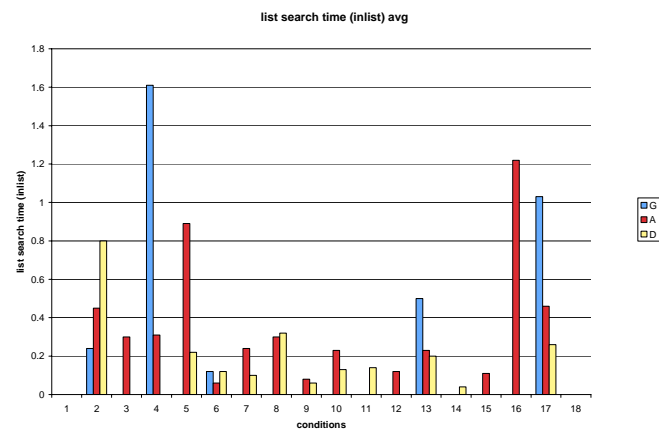


Figure 64: List search time when the target word is in the list, for the average of all trials in all the conditions. The data is presented in order of condition.

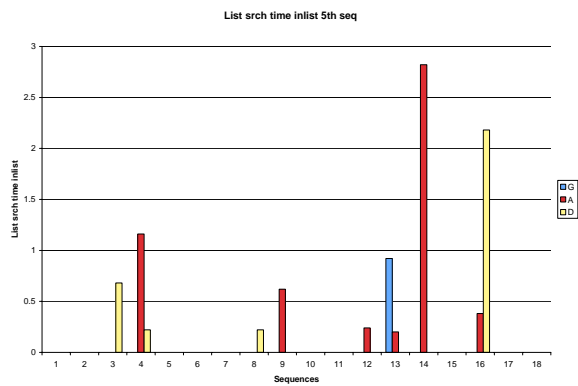


Figure 65: List search time when the target word is in the list, for 5th trial only in all the conditions. The data is presented in order of sequence.

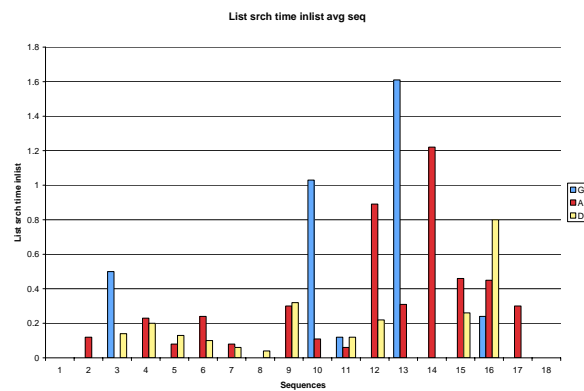


Figure 66: List search time when the target word is in the list, for the average of all trials in all the conditions. The data is presented in order of sequence.

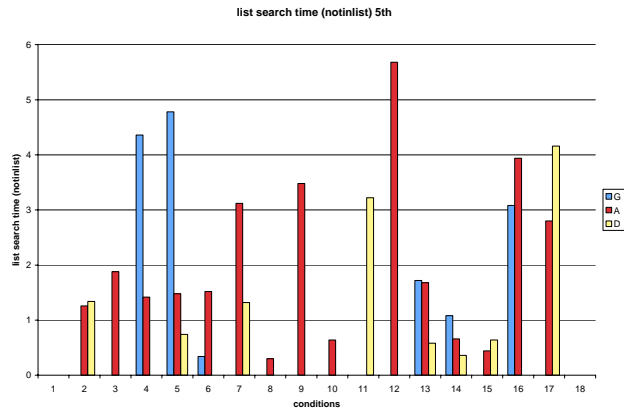


Figure 67: List search time when the target word is not in the list, for 5th trial only in all the conditions. The data is presented in order of condition.

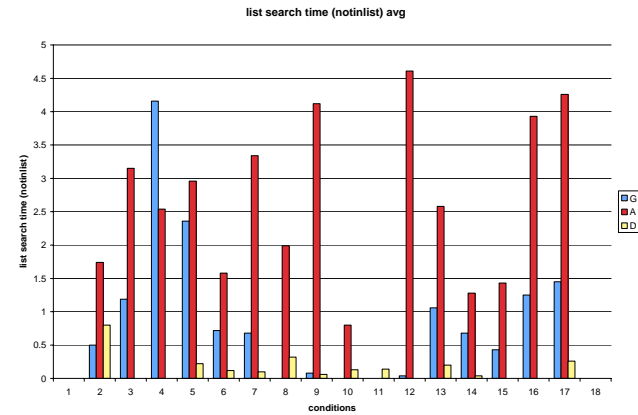


Figure 68: List search time when the target word is not in the list, for the average of all trials in all the conditions. The data is presented in order of condition.

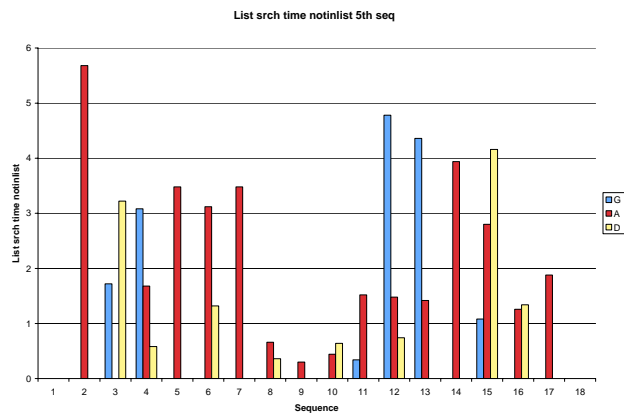


Figure 69: List search time when the target word is not in the list, for 5th trial only in all the conditions. The data is presented in order of sequence.

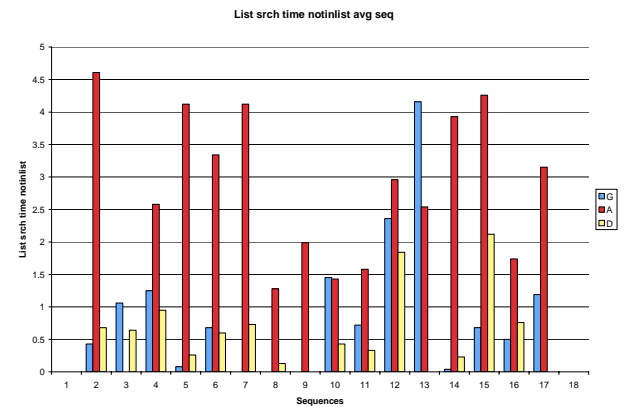


Figure 70: List search time when the target word is not in the list, for the average of all trials in all the conditions. The data is presented in order of sequence.

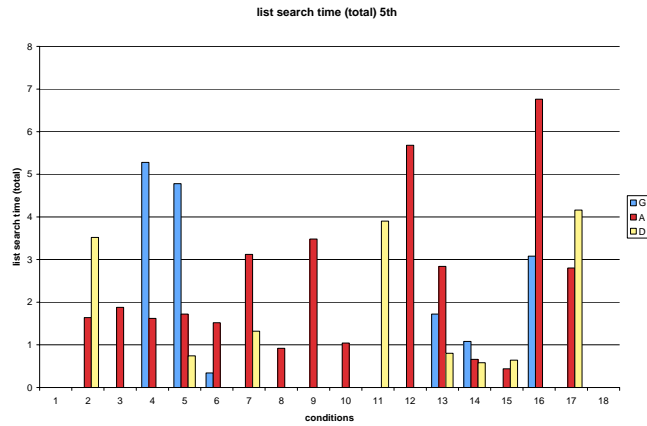


Figure 71: List search time – total (word present + word not present), for 5th trial only in all the conditions. The data is presented in order of condition.

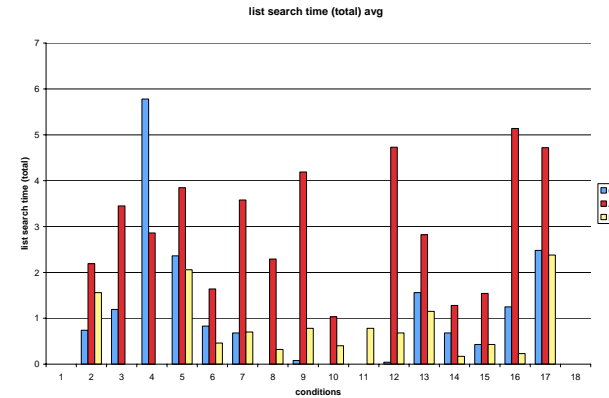


Figure 72: List search time – total (word present + word not present), for the average of all trials in all the conditions. The data is presented in order of condition.

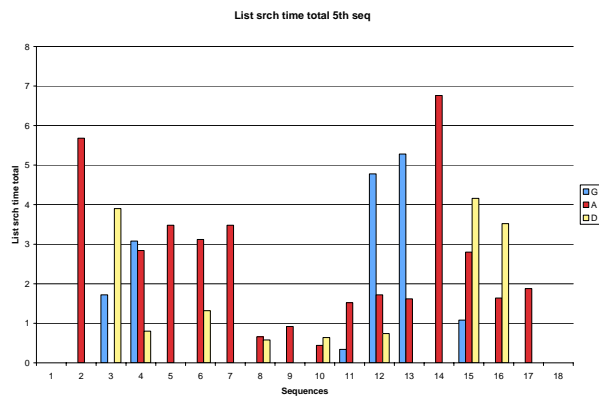


Figure 73: List search time – total (word present + word not present), for 5th trial only in all the conditions. The data is presented in order of sequence.

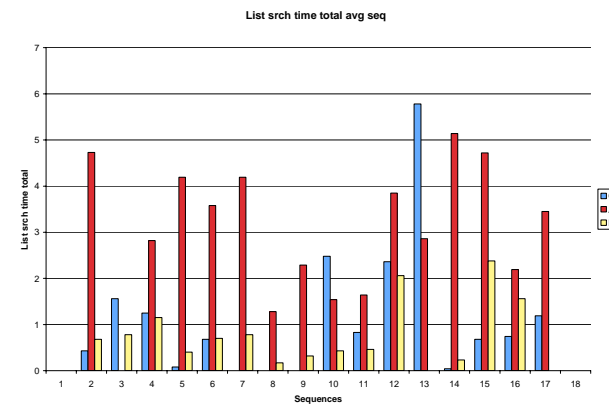


Figure 74: List search time – total (word present + word not present), for the average of all trials in all the conditions. The data is presented in order of sequence.

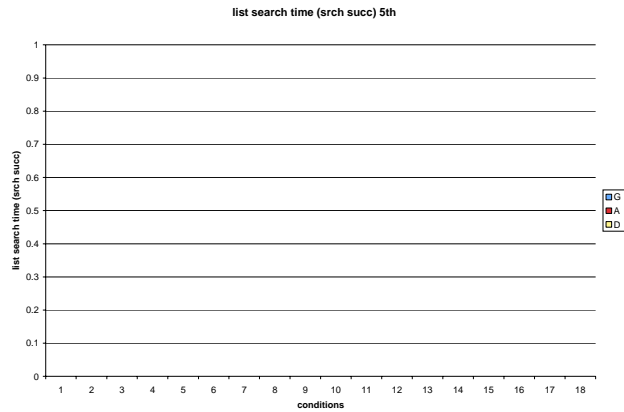


Figure 75: List search time when the search was successful, for 5th trial only in all the conditions. The data is presented in order of condition.

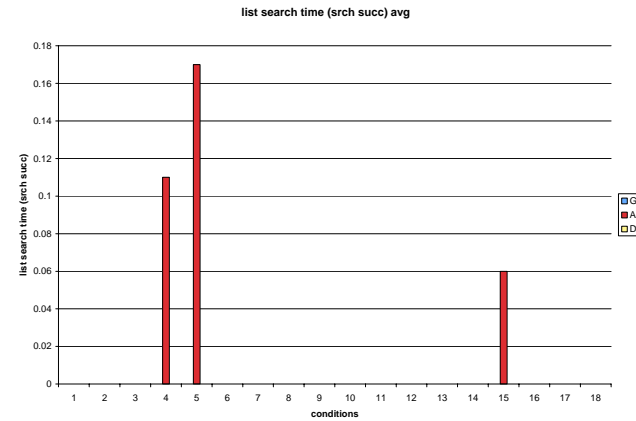


Figure 76: List search time when the search was successful, for the average of all trials in all the conditions. The data is presented in order of condition.

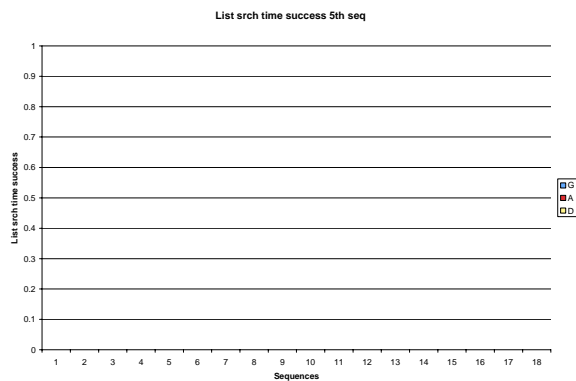


Figure 77: List search time when the search was successful, for 5th trial only in all the conditions. The data is presented in order of sequence.

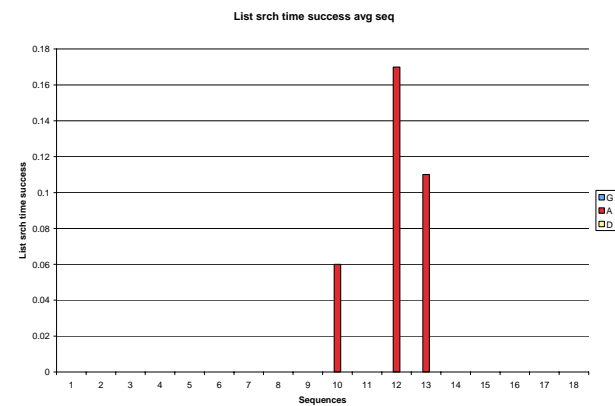


Figure 78: List search time when the search was successful, for the average of all trials in all the conditions. The data is presented in order of sequence.

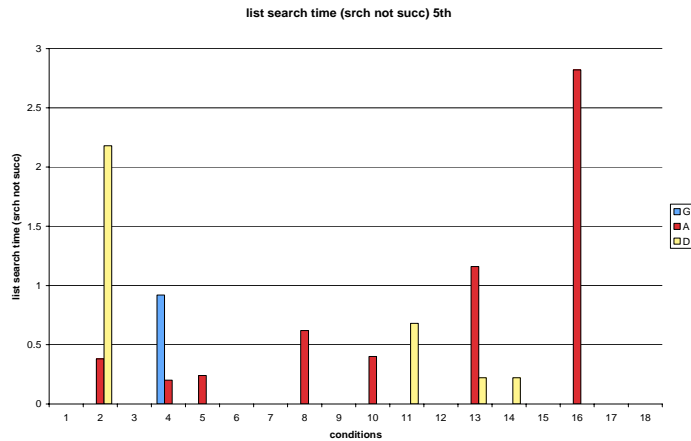


Figure 79: List search time when the search was not successful, for 5th trial only in all the conditions. The data is presented in order of condition.

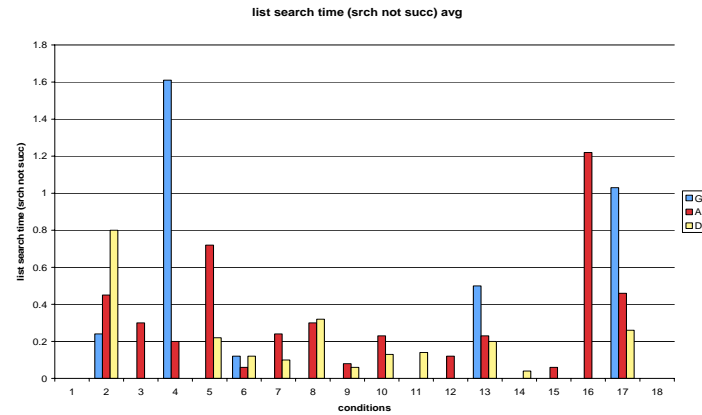


Figure 80: List search time when the search was not successful, for the average of all trials in all the conditions. The data is presented in order of condition.

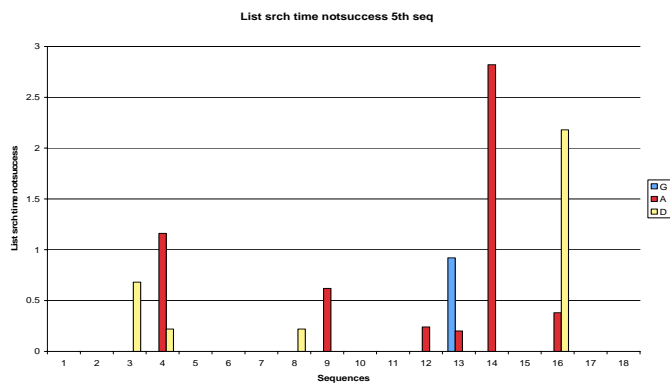


Figure 81: List search time when the search was not successful, for 5th trial only in all the conditions. The data is presented in order of sequence.

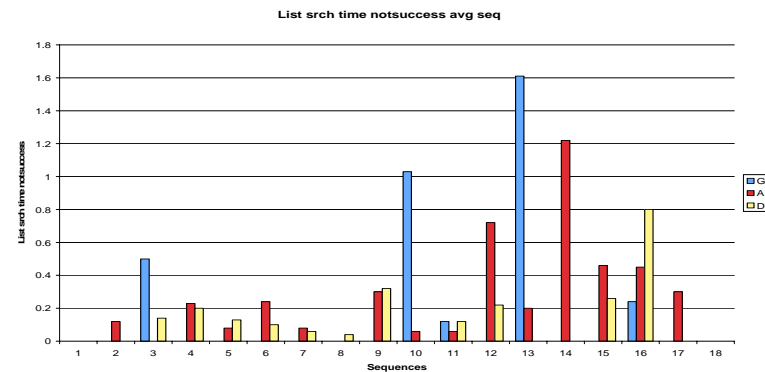


Figure 82: List search time when the search was not successful, for the average of all trials in all the conditions. The data is presented in order of sequence.

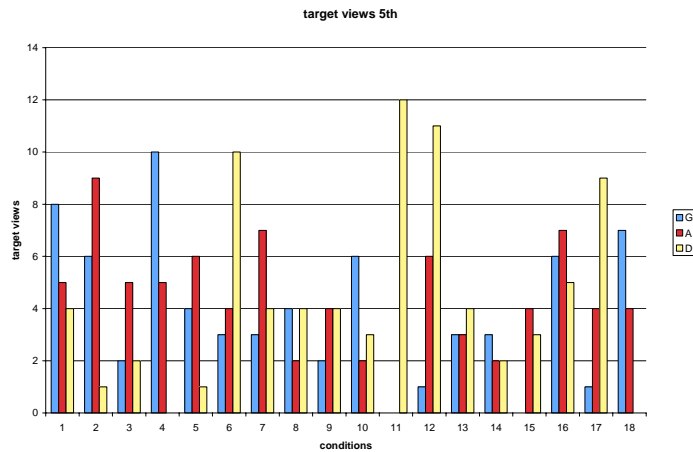


Figure 83: Target text views for 5th trial only in all the conditions. The data is presented in order of condition.

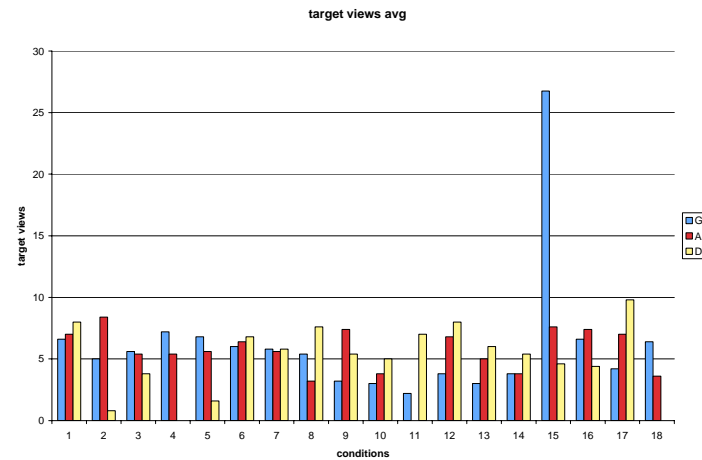


Figure 84: Target text views for the average of all trials in all the conditions. The data is presented in order of condition.

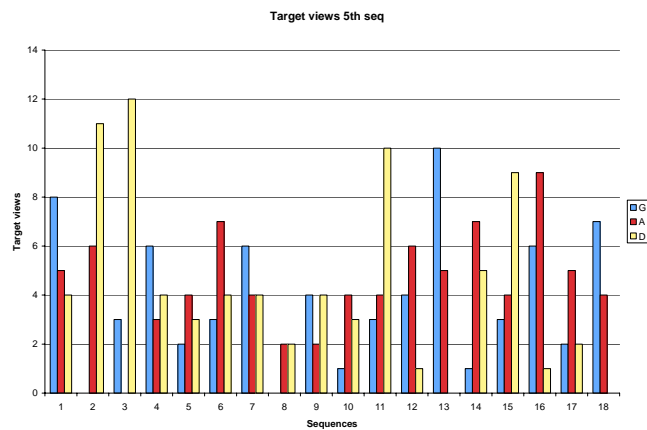


Figure 85: Target text views for 5th trial only in all the conditions. The data is presented in order of sequence.

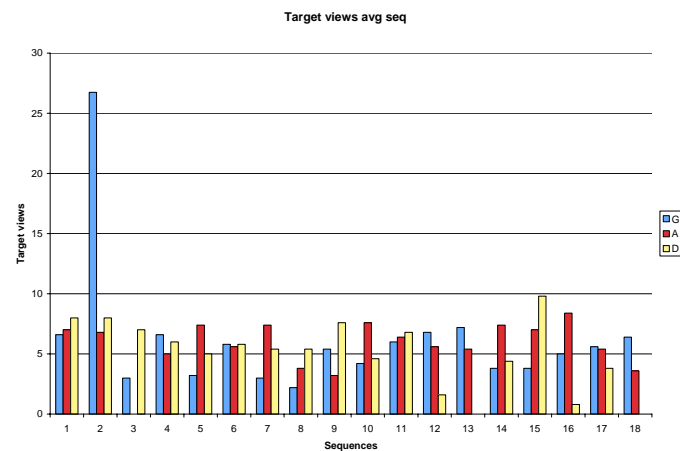


Figure 86: Target text views for the average of all trials in all the conditions. The data is presented in order of sequence.

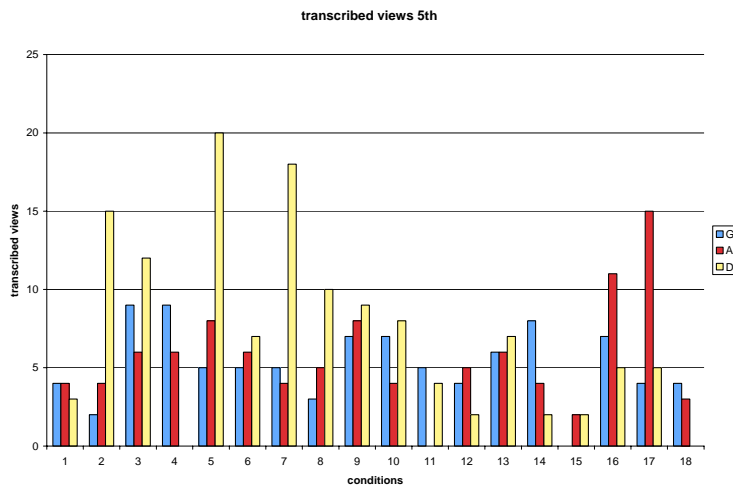


Figure 87: Transcribed text views for 5th trial only in all the conditions. The data is presented in order of condition.

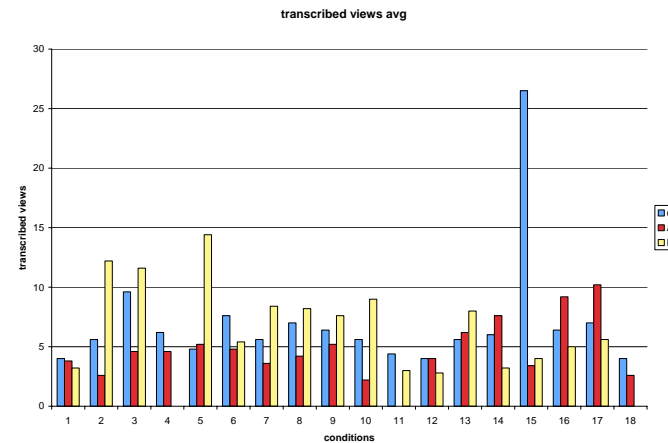


Figure 88: Transcribed text views for the average of all trials in all the conditions. The data is presented in order of condition.

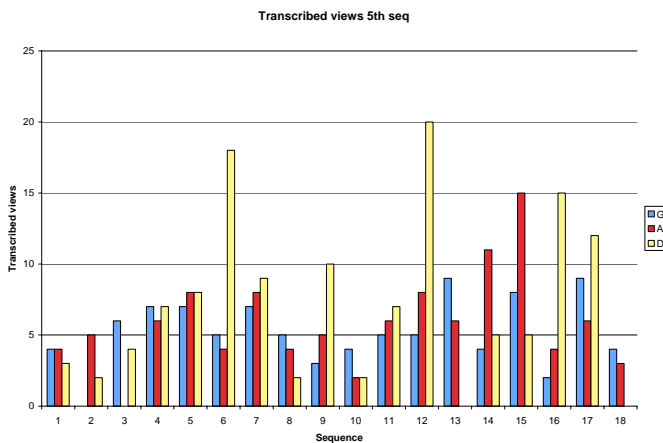


Figure 89: Transcribed text views for 5th trial only in all the conditions. The data is presented in order of sequence.

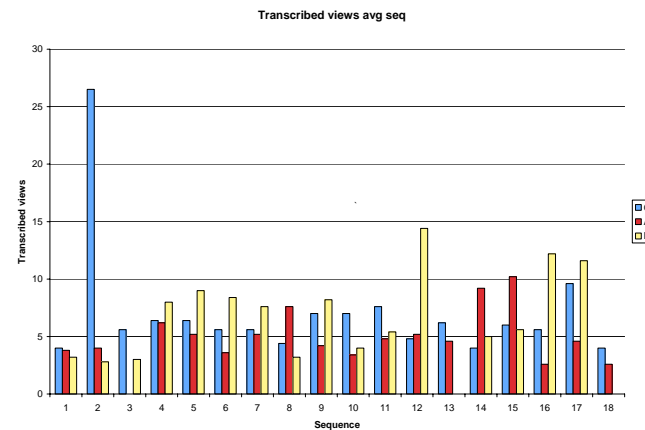


Figure 90: Transcribed text views for the average of all trials in all the conditions. The data is presented in order of sequence.

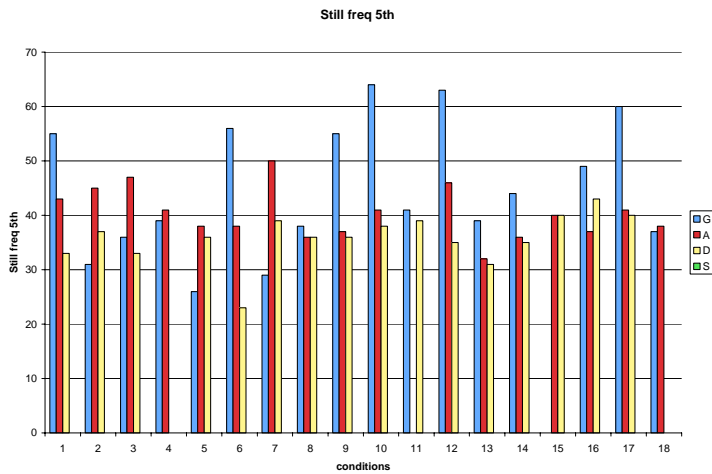


Figure 91: Still frequency for 5th trial in all the conditions.

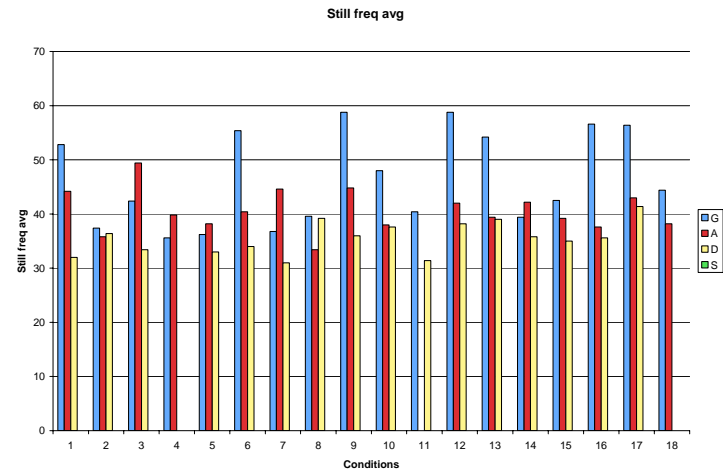


Figure 92: Still frequency for average of all trials in all the conditions.

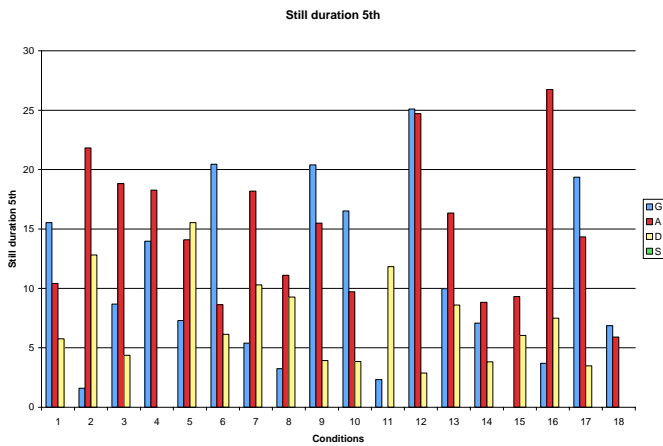


Figure 93: Still duration for 5th trial in all the conditions.

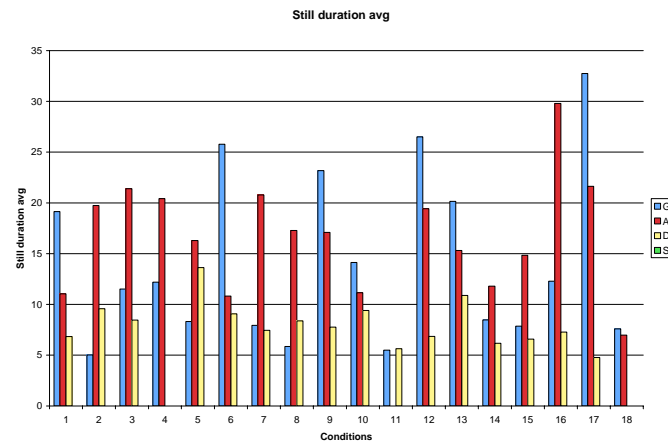


Figure 94: Still duration for average of all trials in all the conditions.



Figure 95: Movement frequency for 5th trial in all the conditions.

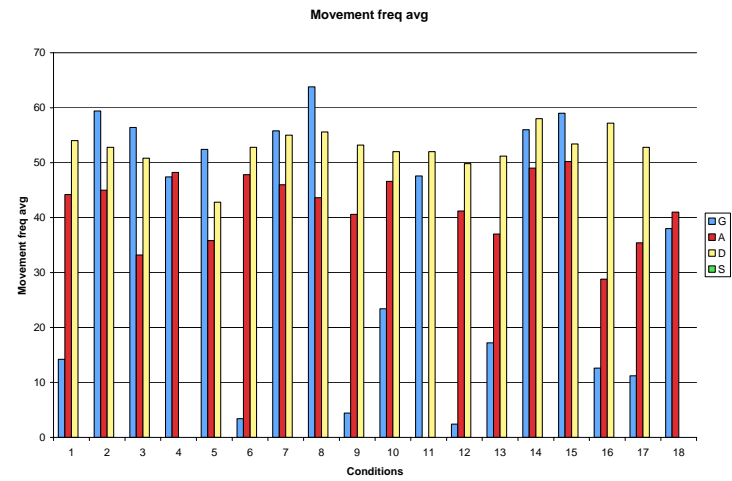


Figure 96: Movement frequency for average of all trials in all the conditions.

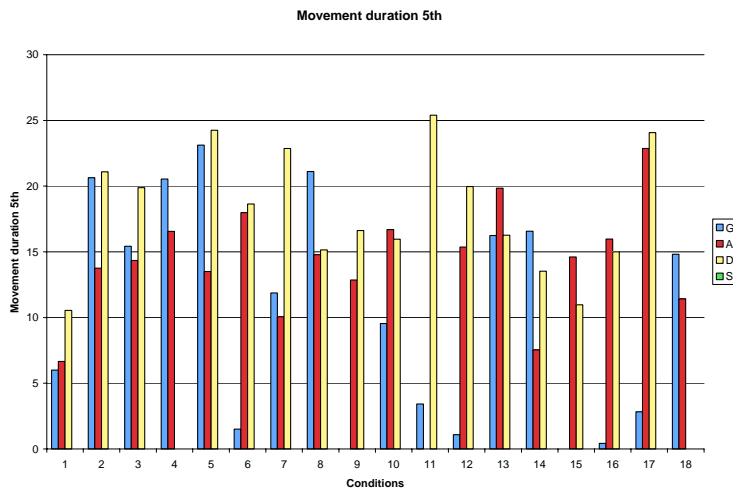


Figure 97: Movement duration for 5th trial in all the conditions.

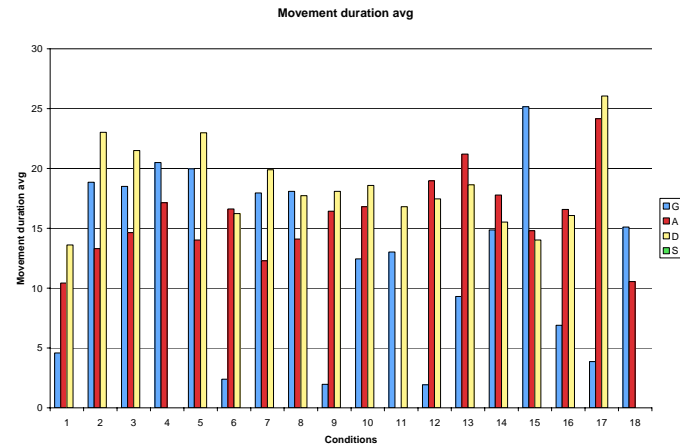


Figure 98: Movement duration for average of all trials in all the conditions.

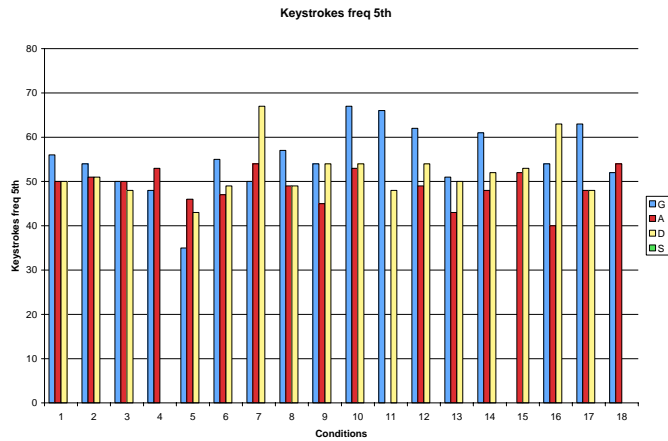


Figure 99: Keystrokes frequency for 5th trial in all the conditions.

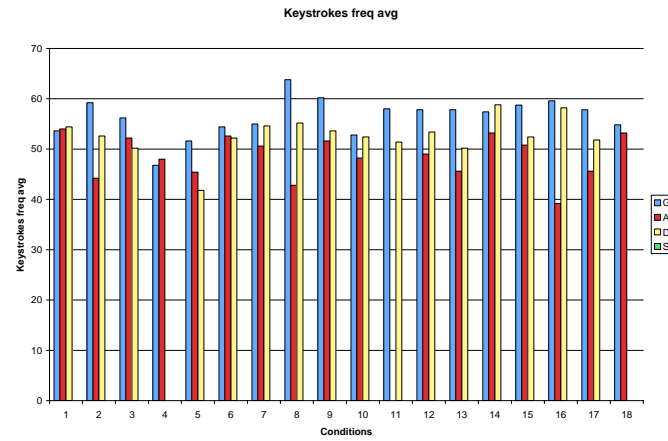


Figure 100: Keystrokes frequency for average of all trials in all the conditions.

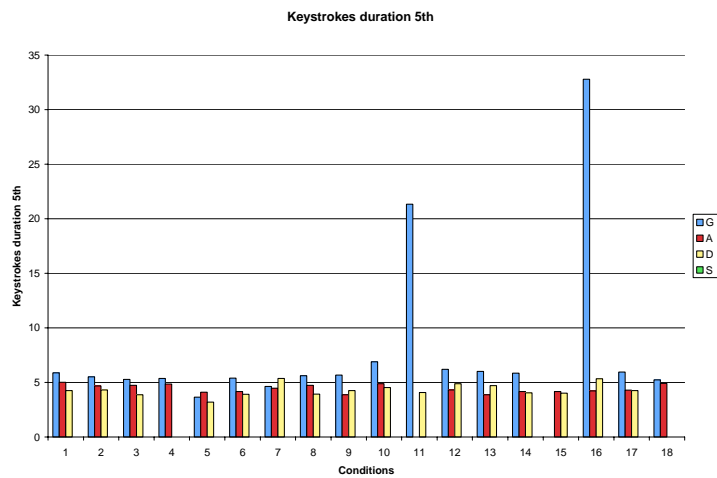


Figure 101: Keystrokes duration for 5th trial in all the conditions.

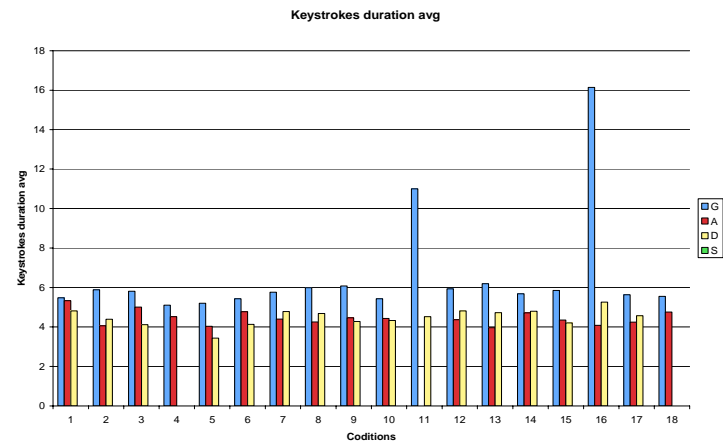


Figure 102: Keystrokes duration for average of all trials in all the conditions.

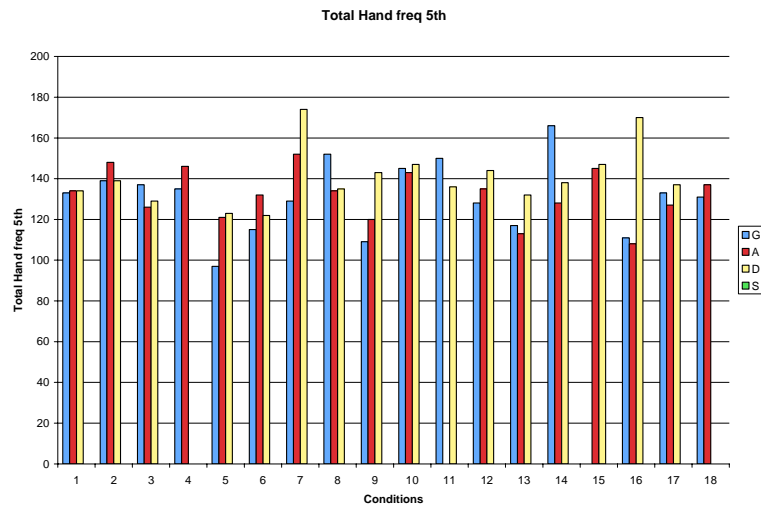


Figure 103: Total hand frequency for 5th trial in all the conditions.

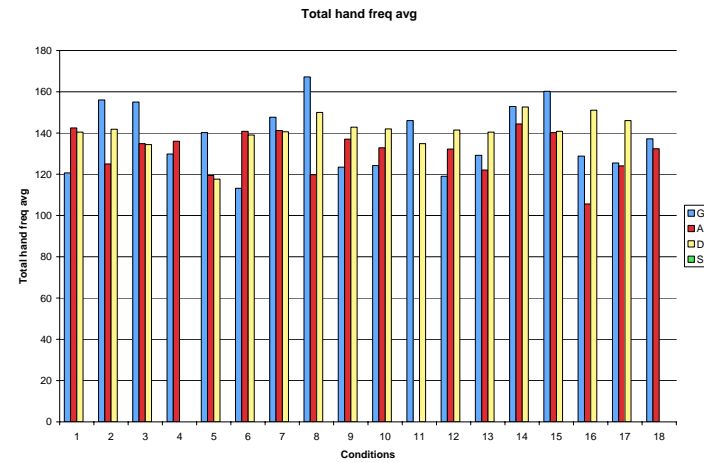


Figure 104: Total hand frequency for average of all trials in all the conditions.

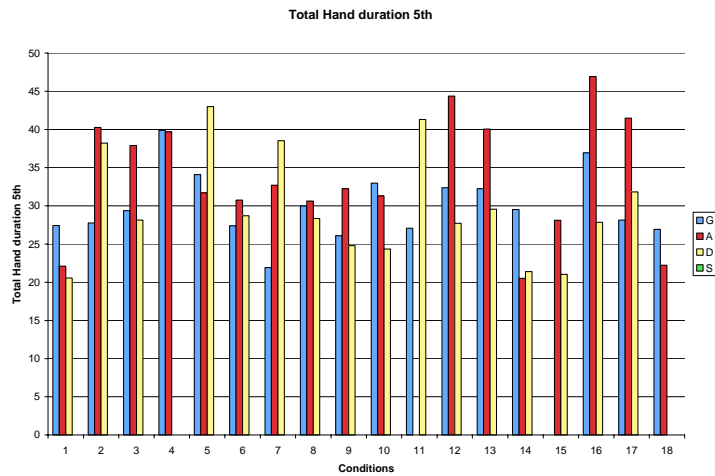


Figure 105: Total hand duration for 5th trial in all the conditions.

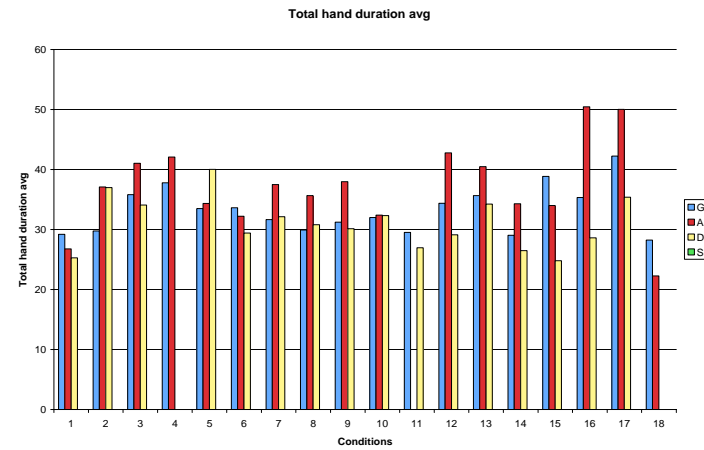


Figure 106: Total hand duration for average of all trials in all the conditions.

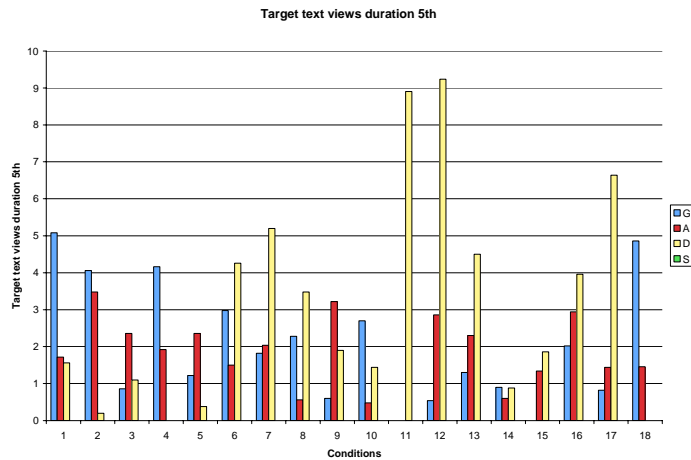


Figure 107: Target text views duration for 5th trial in all the conditions.

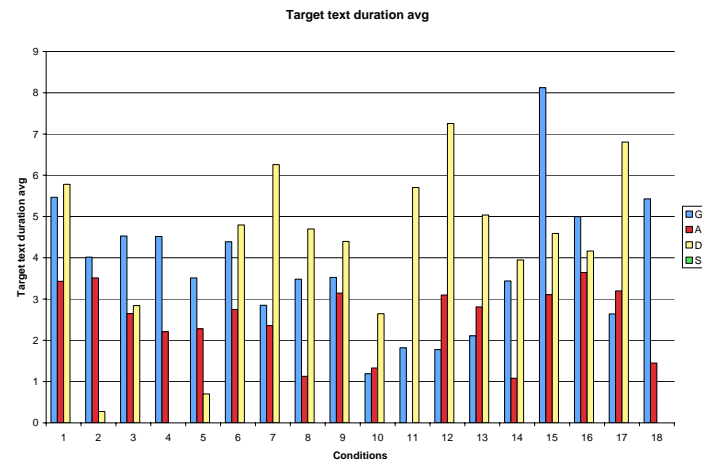


Figure 108: Target text views duration for average of all trials in all the conditions.

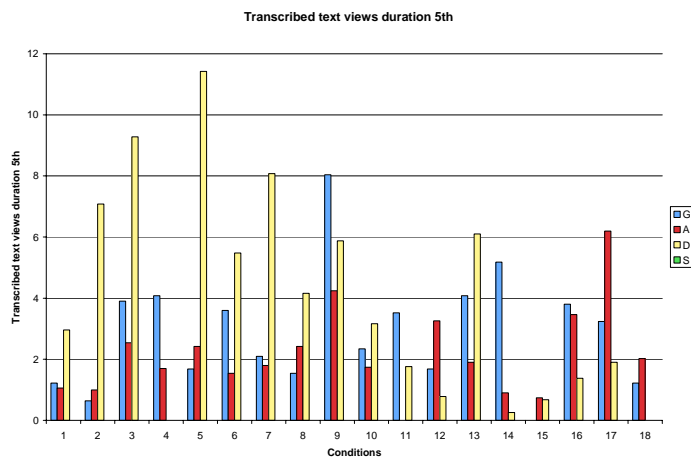


Figure 109: Transcribed text views duration for 5th trial in all the conditions.

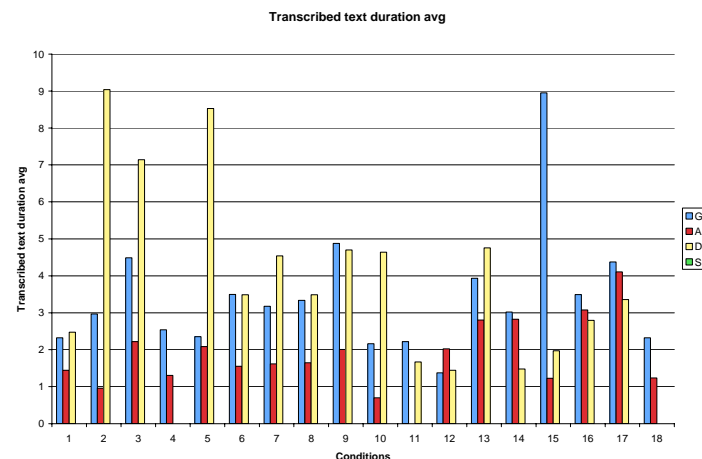


Figure 110: Transcribed text views duration for average of all trials in all the conditions.

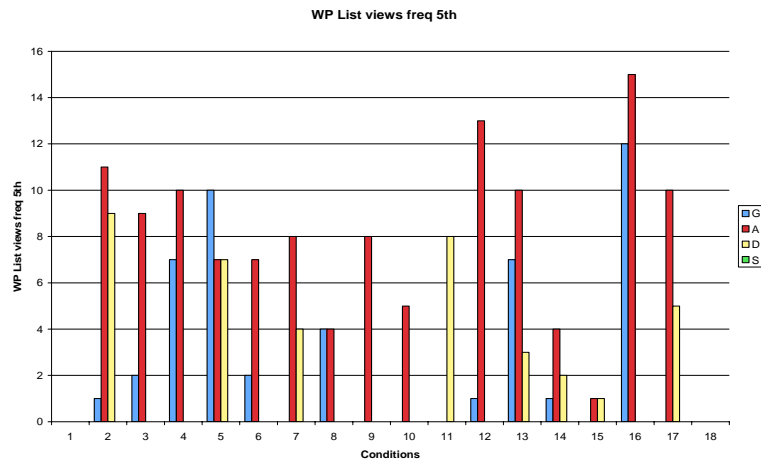


Figure 111: WP List views frequency for 5th trial in all the conditions.

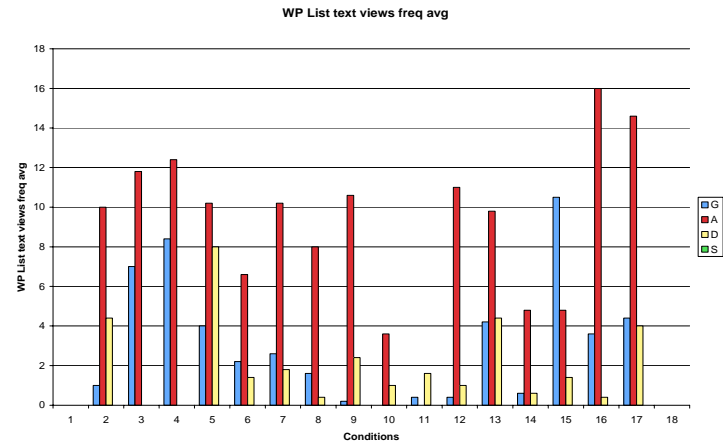


Figure 112: WP List views frequency for average of all trials in all the conditions.

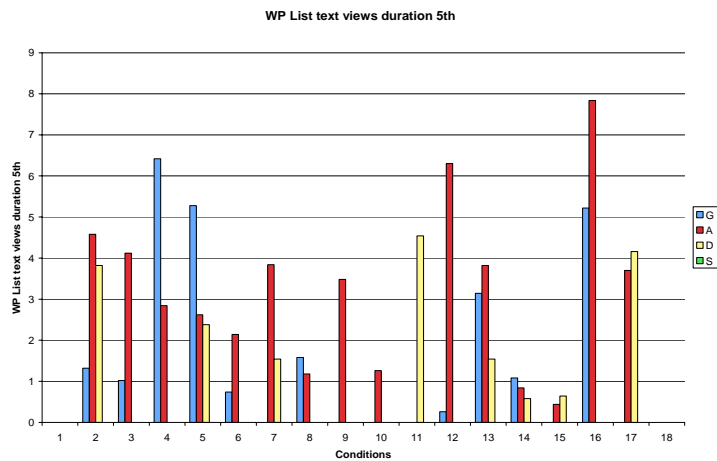


Figure 113: WP List views duration for 5th trial in all the conditions.

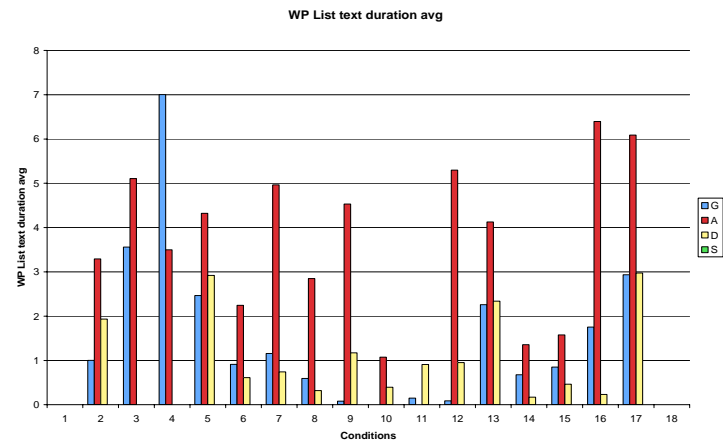


Figure 114: WP List views duration for average of all trials in all the conditions.

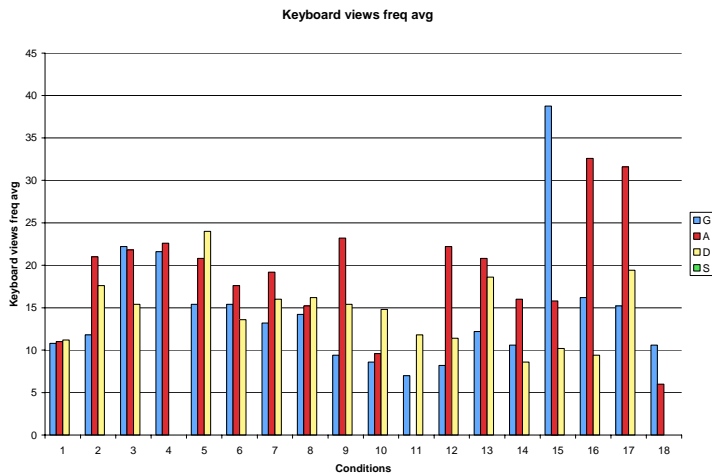


Figure 115: Keyboard views frequency for 5th trial in all the conditions.

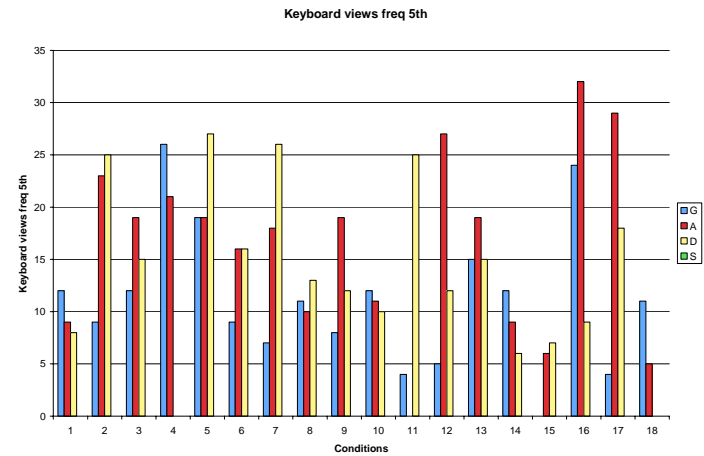


Figure 116: Keyboard views frequency for average of all trials in all the conditions.

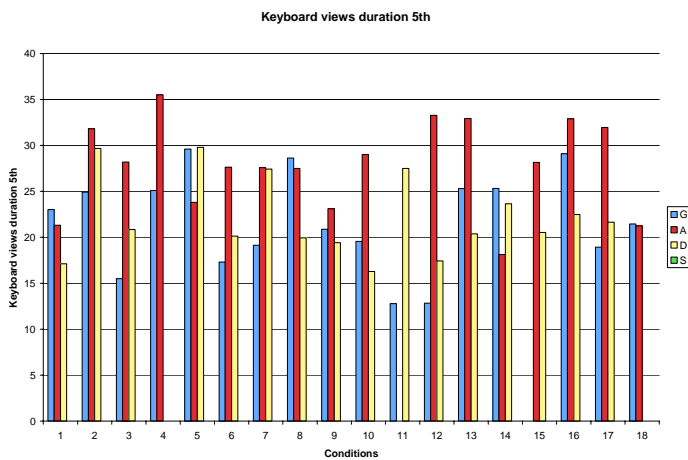


Figure 117: Keyboard views duration for 5th trial in all the conditions.

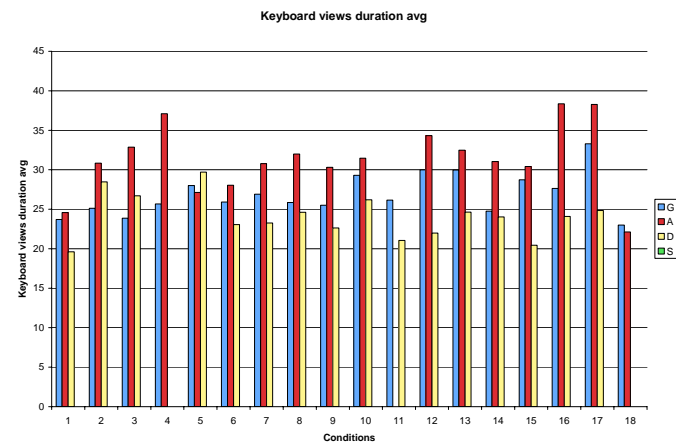


Figure 118: Keyboard views duration for average of all trials in all the conditions.

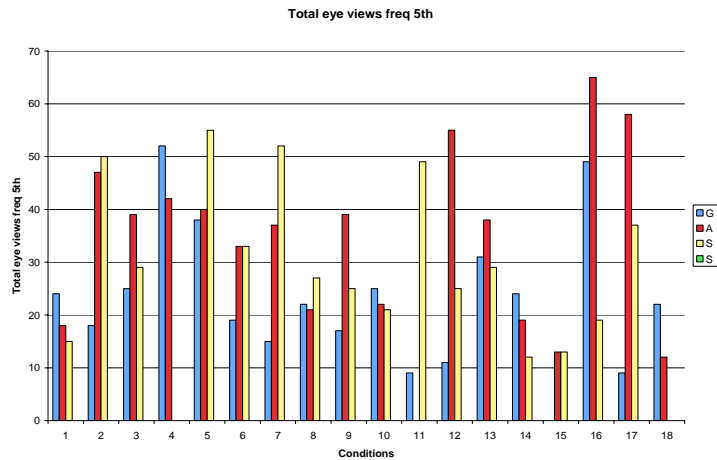


Figure 119: Total eye views frequency for 5th trial in all the conditions.

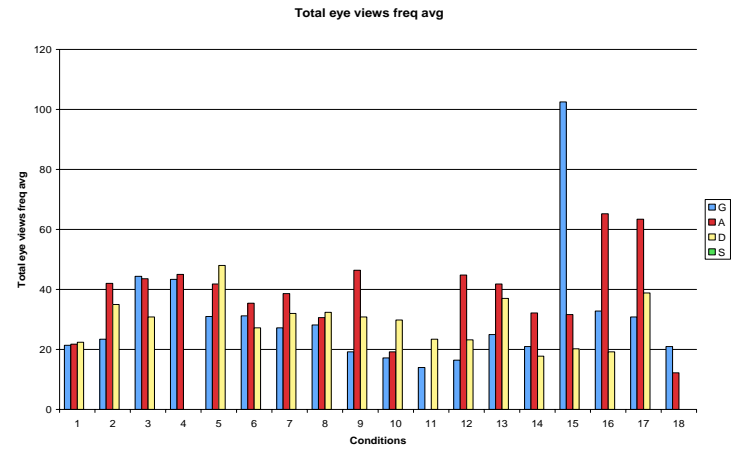


Figure 120: Total eye views frequency for average of all trials in all the conditions.

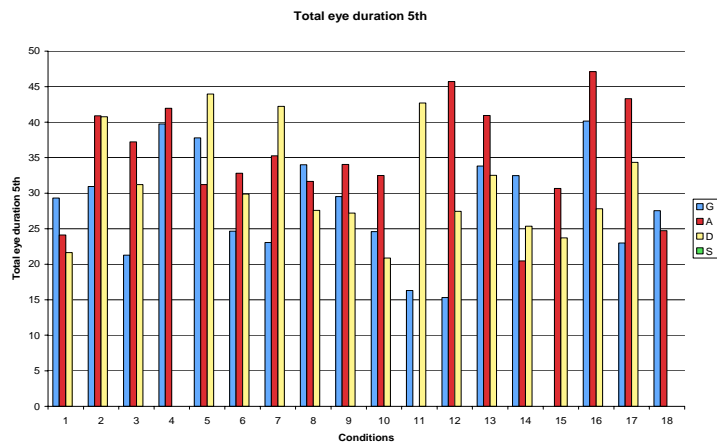


Figure 121: Total eye views duration for 5th trial in all the conditions.

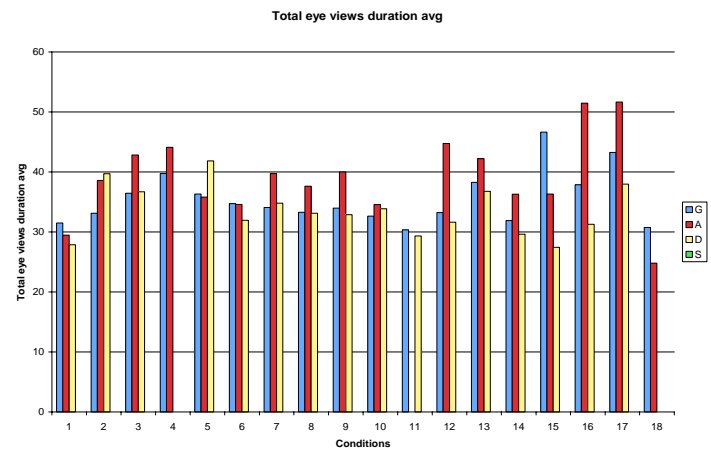


Figure 122: Total eye views duration for average of all trials in all the conditions.

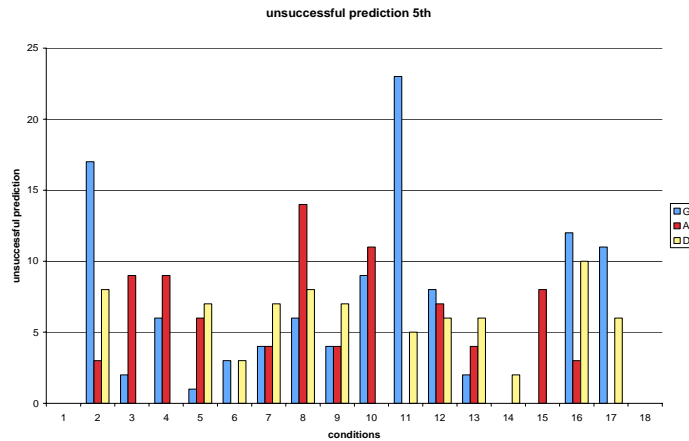


Figure 123: Unsuccessful Prediction for 5th trial only in all the conditions. The data is presented in order of condition.

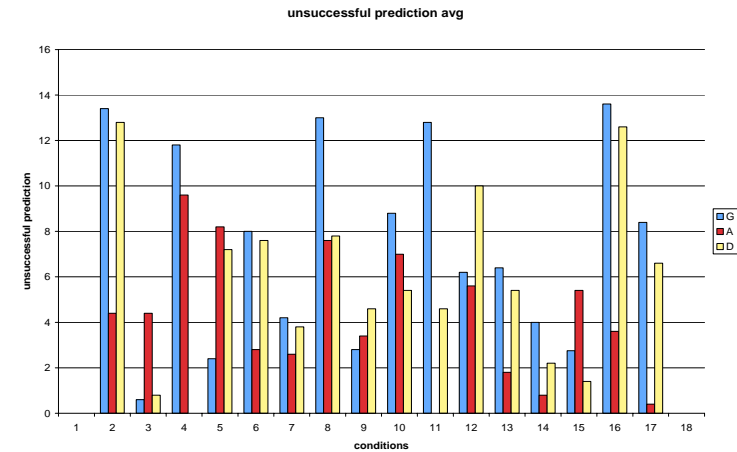


Figure 124: Unsuccessful Prediction for the average of all trials in all the conditions. The data is presented in order of condition.

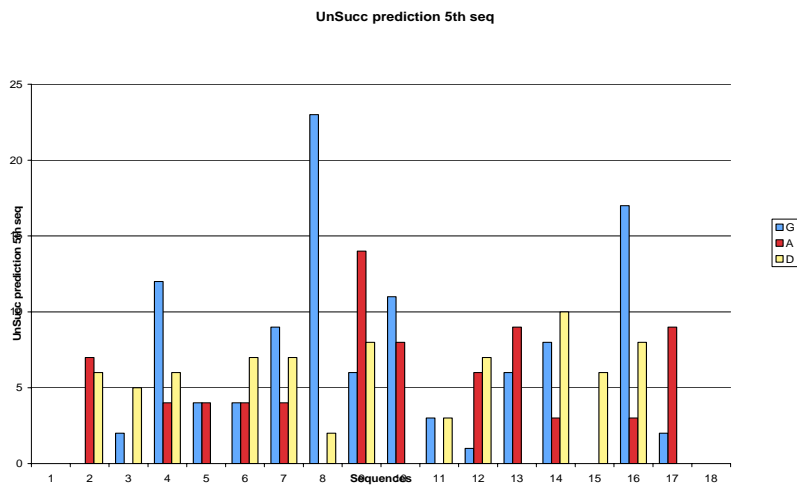


Figure 125: Unsuccessful Prediction for 5th trial only in all the conditions. The data is presented in order of sequence.

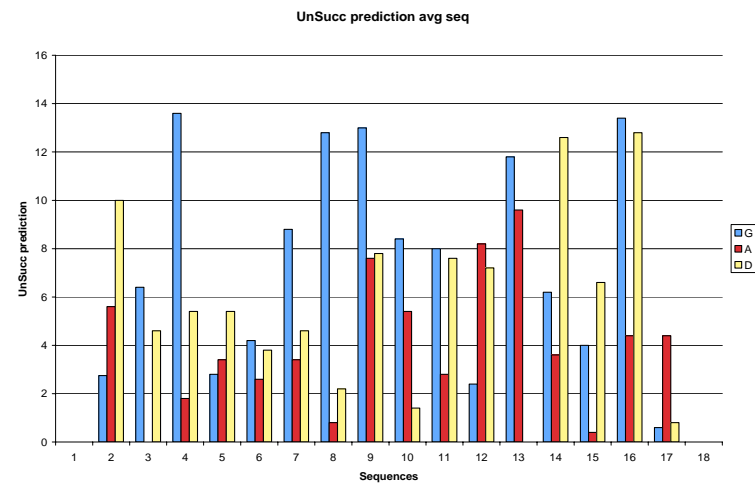


Figure 126: Unsuccessful Prediction for the average of all trials in all the conditions. The data is presented in order of sequence.

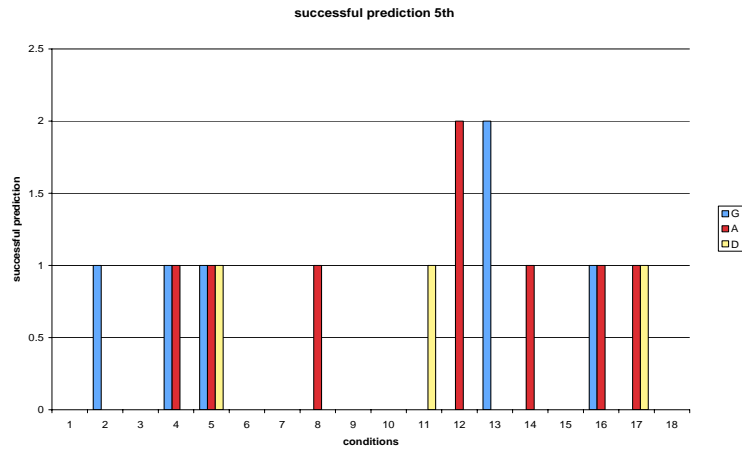


Figure 127: Successful Prediction for 5th trial only in all the conditions. The data is presented in order of condition.

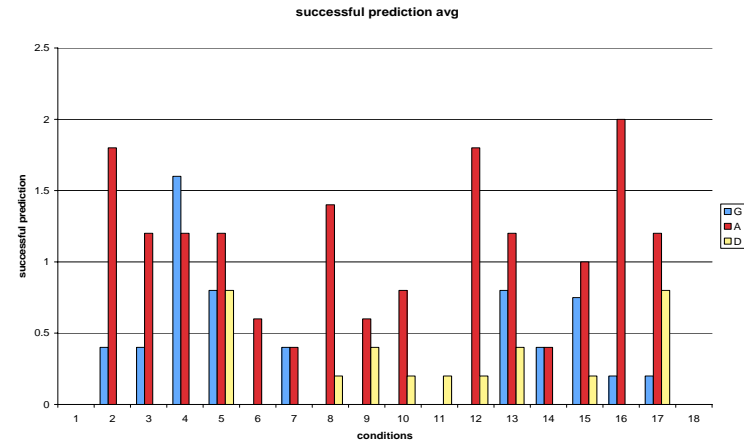


Figure 128: Successful Prediction for the average of all trials in all the conditions. The data is presented in order of condition.

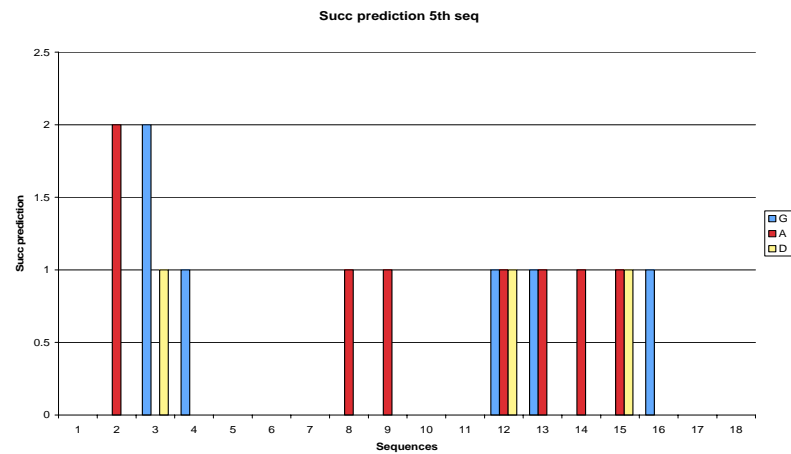


Figure 129: Successful Prediction for 5th trial only in all the conditions. The data is presented in order of sequence.

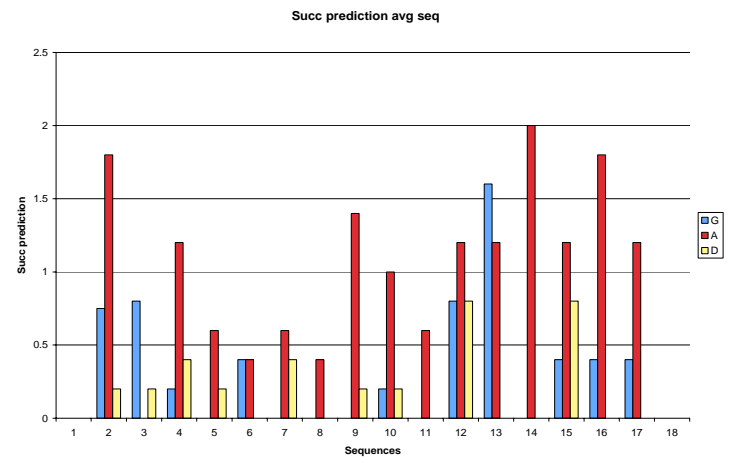


Figure 130: Successful Prediction for the average of all trials in all the conditions. The data is presented in order of sequence.

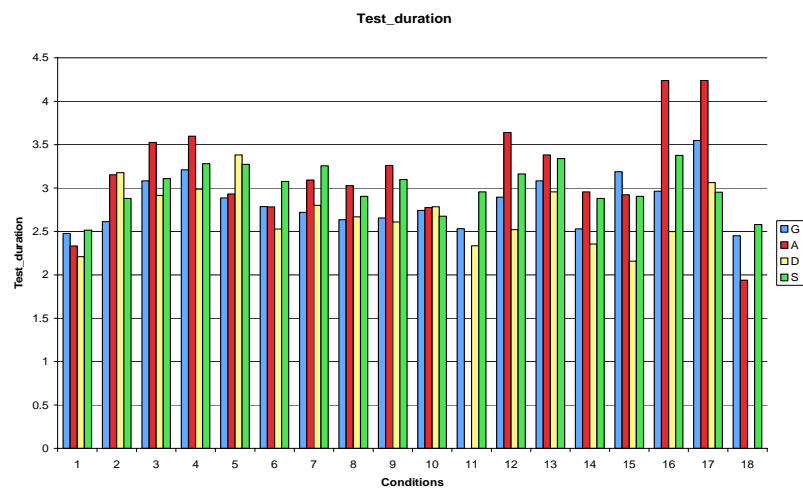


Figure 131: Duration of tests in each condition (in minutes). (Note that the time is for four trials for subject 'G' in condition 15).

APPENDIX B

TABLES

Table 39: T-test results for text entry rate for 5th trial only

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		4>*	4>#	3>*	4>*	4>*		4>*		4>#		4>*	4>*			4>*	4>*	
2									4<@					4<#	3<#			
3				3>#	4>#									4<@				
4						3<@		3<@	3<*	3<*	2<#		3<#	3<#	2<@			3<*
5									4<*	4<#				4<*	3<@			3<#
6														4<@				
7															3<@			
8														4<@				
9																4>#		
10																4>@		
11																		
12														4<#				3<#
13														4<#				3<@
14																4>*	4>#	
15																		
16																		3<*
17																		

Table 40: T-test results for text entry rate for average of all trials

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		4>*	4>*	3>*	4>*	4>*	4>*	4>*	4>*	4>*		4>*	4>*			4>#	4>*	
2				3>@										4<@				
3						4<#		4<#		4<#	3<#			4<*				3<*
4						3<*		3<#		3<*	2<@			3<*				3<*
5						4<*		4<*		4<*	3<*			4<*				3<*
6													4>#	4<@			4>#	3<@
7													4<#					3<@
8													4>#				4>@	3<@
9													4<@					3<#
10													4>*	4<#			4>@	3<@
11													3>#				3>#	
12														4<@				3<#
13														4<*				3<*
14																4>@	4>*	
15																		3<#
16																		3<*
17																		3<#

Table 41: T-test results for keystroke rate for 5th trial only

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		4>*	4>*	3>*	4>*	4>*	4>@	4>#		4>@		4>#	4>*	4>@		4>#	4>#	
2										4<#					3<@			3<@
3				3>@	4>#					4<@								3<#
4						3<#	3<@			3<*								3<*
5						4<#	4<#		4<#	4<*				4<#	3<@			3<*
6																		
7																		
8																		
9																		
10													4>#					
11																		
12																		3<@
13																		3<*
14																		
15																	3>@	
16																		3<*
17																		

Table 42: T-test results for keystroke rate for average of all trials

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		4>*	4>*	3>*	4>*	4>*	4>*	4>*	4>#	4>*		4>#	4>*		4>#	4>@	4>*	
2																		3<@
3						4<*				4<#	3<#			4<@				3<*
4						3<*	3<@			3<*				3<#	3<@			3<*
5						4<*	4<@		4<@	4<*	3<#			4<#				3<*
6													4>#				4>#	3<#
7																		3<#
8																		3<*
9																		3<@
10													4>@				4>#	3<#
11													3>@				3>#	
12																		3<#
13														4<@				3<*
14																	4>#	
15																		3<*
16																		3<#
17																		3<*

Table 43: T-test results for % correct keystrokes for 5th trial only

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1				3>#								4>@						
2				3>#														
3				3>#														
4							3<#		3<*					3<#				3<@
5																		
6																		
7																		
8																		
9												4>@						3>#
10																		
11																		
12																		
13																		
14																		
15																		
16																		
17																		

Table 44: T-test results for %correct keystrokes for average of all trials

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1																		
2							4<@											
3																		
4																		
5																		
6																		
7																		
8																		
9																		
10																		
11																		
12																		
13																		
14																		
15																		
16																		
17																		

Table 45: T-test results for %error keystrokes for 5th trial only

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1																		
2				3<@														
3																		
4							3>@		3>@									
5																		
6																		
7																		
8																		
9																		
10																		
11																		
12																		
13																		
14																		
15																		
16																		
17																		

Table 46: T-test results for %error keystrokes for average of all trials

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1																		
2			4>*															
3																		
4										3>@								
5																		
6																		
7																		
8																		
9																		
10																		
11																		
12																		
13																		
14																		
15																		
16																		
17																		

Table 47: T-test results for time between keystrokes for 5th trial only

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		4<*	4<*	3<*	4<*	4<*	4<#	4<*		4<*	3<@	4<#	4<*	4<*	3<#	4<#	4<*	
2										4>#				4>#	3>*			3>@
3				3<#	4<@					4>@				4>#				3>#
4						3>#	3>#	3>*	3>#	3>*	2>@			3>*	2>@			3>*
5						4>#	4>#	4>*	4>*	4>*				4>*	3>#			3>#
6																		
7																		
8																		3>@
9																		
10													4<@				4<#	
11																		
12																		3>@
13														4>#				3>#
14																4<@	4<#	
15																	3<#	
16																		3>#
17																		3>#

Table 48: T-test results for time between keystrokes for average of all trials

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		4<*	4<*	3<*	4<*	4<*	4<*		4<#	4<*		4<#	4<*		4<#		4<*	
2																		3>@
3						4>@					3>@			4>#				3>*
4					3>@	3>*	3>#		3>@	3>*	2>@			3>*	3>*			3>*
5						4>*	4>@	4>@	4>#	4>*	3>#			4>*	4>#			3>*
6													4<#				4<#	3>#
7																		3>#
8																		
9																		3>@
10																	4<@	3>*
11													3<@				3<*	
12																		3>#
13														4>#				3>*
14																	4<#	
15																	4<@	3>*
16																		3>@
17																		3>#

Table 49: T-test results for list search time for 5th trial only

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0
2																		all_0
3											all_0							all_0
4											all_0							all_0
5																		all_0
6											all_0							all_0
7																		all_0
8											all_0		3<#					all_0
9											all_0							all_0
10											all_0		3<@					all_0
11												all_0					all_0	all_0
12																		all_0
13															3>#			all_0
14																		all_0
15																		all_0
16																		all_0
17																		all_0

Table 50: T-test results for list search time for average of all trials

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0
2					3<@					3>@							3<@	all_0
3											all_0							all_0
4										2>@	all_0							all_0
5						3>#		3>@		3>*	2>*			3>*	3>*			all_0
6																	3<#	all_0
7																		all_0
8																	3<#	all_0
9																		all_0
10													3<#				3<*	all_0
11																	2<*	all_0
12																		all_0
13														3>@				all_0
14																	3<*	all_0
15																	3<*	all_0
16																		all_0
17																		all_0

Table 51: T-test results for successful prediction for 5th trial only

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
9																		
10																		
11																		
12																		
13																		
14																		
15																		
16																		
17																		

Table 52: T-test results for successful prediction for average of all trials

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0
2											all_0							all_0
3											all_0							all_0
4						2>#	2>*		2>#		all_0			2>*				all_0
5						3>*	3>*		3>#	3>#	2>*			3>*				all_0
6											all_0		3<@					all_0
7											all_0		3<@					all_0
8																		all_0
9																		all_0
10																		all_0
11														all_0		all_0		all_0
12																		all_0
13														3>@				all_0
14																		all_0
15																		all_0
16																		all_0
17																		all_0

Table 53: T-test results for unsuccessful prediction for 5th trial only

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0
2														3>@				all_0
3														all_0				all_0
4						2>@	2>@		2>@				2>@	all_0				all_0
5														3>@				all_0
6							3<@	3<*	3<@			3<*			all_0	3<#		all_0
7														3>*				all_0
8													3>@	3>*				all_0
9														3>*				all_0
10														all_0				all_0
11															all_0			all_0
12													3>#	3>*				all_0
13														3>#				all_0
14															all_0	3<#		all_0
15																	all_0	all_0
16																		all_0
17																		all_0

Table 54: T-test results for unsuccessful prediction for average of all trials

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0	all_0
2			3>#				3>#		3>#					3>#	3>#			all_0
3				2<#	3<@	3<@		3<*		3<*		3<*				3<#		all_0
4							2>*		2>*			2>#	2>@	2>#	2>#			all_0
5														3>@				all_0
6														3>@				all_0
7								3<*		3<*		3<#				3<@		all_0
8								3>*					3>#	3>*	3>*			all_0
9									3<*			3<#				3<@		all_0
10														3>*	3>#			all_0
11																		all_0
12														3>*	3>#			all_0
13																		all_0
14																3<#		all_0
15																3<@		all_0
16																		all_0
17																		all_0

Table 55: T-test results for target text views for 5th trial only

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1													3>@	3>#				
2																		
3																3<#		
4											all_0							
5																		
6																		
7														3>@				
8																3<*		
9																3<*		
10																		
11																		all_0
12																		
13														3>@		3<*		
14															2<#	3<*		
15																		
16																		
17																		

Table 56: T-test results for target text views for average of all trials

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1			3>*				3>*			3>*			3>#	3>*				
2																		
3						3<#												
4										2>#				2>@				
5																		
6							3>#			3>*			3>@	3>*				
7										3>*				3>#				
8																		
9																		
10																3<@	3<@	
11																		
12																		
13																		
14																		
15																		
16																		
17																		

Table 57: T-test results for transcribed text views for 5th trial only

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1			3<*	2<@		3<*			3<*	3<#			3<*		2>#	3<#		
2				2<@														
3											2>#	3>#			2>@			2>@
4																		2>@
5																		
6									3<#			3>#			2>*			2>@
7																		
8																		
9											2>#	3>*	3>#	3>@	2>*			2>*
10											2>#	3>@						
11													2<@					
12													3<*			3<@		
13															2>*			2>*
14																		
15																		
16																		2>@
17																		

Table 58: T-test results for transcribed text views for average of all trials

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1			3<#			3<#		3<#	3<*				3<*			3<#	3<*	
2																		
3											2>*	3>#						
4																		
5																		2>@
6												3>#						
7																		
8											2>#	3>#						
9											2>#	3>*						2>@
10																		
11																	2<@	
12													3<*			3<#	3<*	
13																		2>#
14																		2>#
15																		
16																		2>@
17																		2>#

BIBLIOGRAPHY

1. Koester, H. H., & Levine, S.P. (1994). Modeling the Speed of Text Entry with a Word Prediction Interface. *IEEE Transactions on Rehabilitation Engineering*, 2:3, 177-187.
2. http://www.infouse.com/disabilitydata/disability/1_1.php
3. Louis Harris (1996) - Disability Studies: What it is and Why it is Needed? - *Herald* 25(4), February 12.
4. <http://www.ed.gov/about/offices/list/osers/rsa/index.html> (1995) - administrative data from the Rehabilitation services Administration.
5. Horstmann, H. M., & Levine, S.P. (1990). Modeling of user performance with computer access and augmentative communication systems for handicapped people. *Augmentative and Alternative Communication*, 6, 231-241.
6. John, B. E., & Kieras, D. E. (1996). The GOMS family of User Interface analysis techniques: Comparison and contrast. *ACM Transactions on Computer-Human Interaction*, 3:4, 320-351.
7. http://www.cc.gatech.edu/classes/cs6751_97_fall/projects/closet2000+/FinalEssays/goms.html.
8. http://sigchi.org/chi95/Electronic/documnts/tutors/bej0_bdy.htm
9. Card S., Moran T., Newell A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
10. Koester, H.H. and Levine, S.P. (1996). Effect of a Word Prediction Feature on User Performance. *Augmentative and Alternative Communication*, 12:3, 155-168.
11. Koester H.H. & Levine S.P. (1997). Keystroke - Level Models for User Performance with Word Prediction. *Augmentative and Alternative Communication*, 13:4, 239-257
12. Koester H.H. & Levine S.P. (1998). Model simulations of User performance with Word Prediction. *Augmentative and Alternative Communication*, 14:01, 25-35.

13. Olson J. R., Olson G.M. (1990). The growth of cognitive modeling of human-computer interaction. 5,221-265.
14. MacKenzie, I. S., & Soukoreff, R. W. (2003). Phrase sets for evaluating text entry techniques. *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems – CHI 2003*, pp. 754-755 New York: ACM
15. Hinton, P. R. (1995). *Statistics Explained*. New York, Routledge.
16. Keates, S., & Robinson, P. (1997). User performance modeling and cognitive load. *Proceedings of RESNA 97*, Pittsburgh, PA. 342-344
17. MacKenzie, I. S., & Soukoreff, R. W. (2003). Metrics for Text Entry Research: An Evaluation of MSD and KSPC, and a New Unified Error Metric. *Proceedings of the ACM Conference on Human Factors in Computing Systems – CHI*, 113-120.