

**THE UNIVERSITY OF MANCHESTER - APPROVED ELECTRONICALLY
GENERATED THESIS/DISSERTATION COVER-PAGE**

Electronic identifier: 28639

Date of electronic submission: 27/08/2019

Thesis format: Traditional

The University of Manchester makes unrestricted examined electronic theses and dissertations freely available for download and reading online via Manchester eScholar at <http://www.manchester.ac.uk/escholar>.

This print version of my thesis/dissertation is a TRUE and ACCURATE REPRESENTATION of the electronic version submitted to the University of Manchester's institutional repository, Manchester eScholar.

**Duration related effects in reading
performance in individuals
with visual impairment**

A thesis submitted to The University of Manchester for the degree of
Doctor of Philosophy
in the Faculty of Biology, Medicine and Health
School of Health Sciences
Division of Pharmacy and Optometry

Year: 2019

Chee Yong Kingsley, Muk
School of Health Sciences
Division of Pharmacy and Optometry

Table of Contents	i
List of Figures	vii
List of Tables	x
Abstract	xiii
Declarations	xiv
Copyright Statement	xv
Acknowledgements	xvi

Table of Contents

1. LITERATURE REVIEW	1
1.1 INTRODUCTION	1
1.1.1 <i>Nature of the reading process</i>	2
1.1.2 <i>An overview of reading performance indicators as outcome measures</i>	2
1.1.3 <i>The use of reading speed as a reading performance indicator</i>	15
1.2 EXPERIMENTAL DESIGN INFLUENCE ON READING SPEED ASSESSMENT	15
1.2.1 <i>Text presentation modes</i>	15
1.2.2 <i>Typeface</i>	18
1.2.3 <i>Contextual vs random words</i>	21
1.2.4 <i>Singles sentences vs paragraphs</i>	23
1.2.5 <i>Oral vs silent reading</i>	24
1.3 VISUAL FACTORS INFLUENCING READING PERFORMANCE	26
1.3.1 <i>Acuity reserve</i>	26
1.3.2 <i>Contrast reserve</i>	27
1.3.3 <i>Perceptual span</i>	28
1.3.4 <i>Presence and size of central scotoma</i>	29
1.4 TYPES OF READING ASSESSMENT TESTS	30
1.4.1 <i>Bailey-Lovie word-reading chart</i>	30
1.4.2 <i>International Reading Speed Text (IReST)</i>	31
1.4.3 <i>Minnesota Reading (MNREAD) acuity chart</i>	32

1.4.4	<i>Morgan Low Vision Reading Comprehension Assessment (LVRCA)</i>	35
1.4.5	<i>Pepper Visual Skills for Reading Test (VSRT)</i>	35
1.4.6	<i>Rate of Reading (RoR)</i>	37
1.4.7	<i>Precision of reading assessment</i>	37
1.5	THE INFLUENCE OF AGE ON READING SPEED PERFORMANCE	42
1.5.1	<i>Age and normal vision</i>	42
1.5.2	<i>Age and low vision</i>	46
1.6	SIMULATING VISUAL IMPAIRMENT	48
1.7	OBJECTIVES OF THIS STUDY	49
2.	TO EXPLORE THE LIMITATIONS OF READING SPEED ASSESSMENT METHODS	50
2.1	EXPLORING THE LIMITATION OF THE INCLUSION OF THE PRE-VERBALISATION TIME IN THE COMPUTATION OF READING SPEED	51
2.1.1	<i>Introduction</i>	51
2.1.2	<i>Methods</i>	53
2.1.3	<i>Statistical analysis</i>	55
2.1.4	<i>Results</i>	56
2.1.5	<i>Discussion</i>	60
2.1.6	<i>Limitations</i>	64
2.2	TO EXPLORE THE USE OF PAGE-TURN AS A TIME RECORDING METHOD IN A SILENT READING TEST	66
2.2.1	<i>Introduction</i>	66
2.2.2	<i>Methods</i>	67
2.2.3	<i>Statistical analysis</i>	68
2.2.4	<i>Results</i>	69
2.2.5	<i>Discussion</i>	73
2.3	THE USE OF PROOFREADING TO ENSURE READING COMPLIANCE IN SILENT READING TESTS	75
2.3.1	<i>Introduction</i>	75
2.3.2	<i>Methods</i>	76
2.3.3	<i>Statistical analysis</i>	77

2.3.4	<i>Results</i>	78
2.3.5	<i>Discussion</i>	82
2.3.6	<i>Limitations</i>	84
2.4	CONCLUSION	85
3.	EVALUATION OF READING ENDURANCE IN YOUNG ADULTS UNDER CONDITIONS OF SIMULATED	
	VISUAL IMPAIRMENT	87
3.1	INTRODUCTION	87
3.2	METHODS	88
3.2.1	<i>Participants</i>	88
3.2.2	<i>Reading texts</i>	89
3.2.3	<i>Procedures</i>	92
3.2.4	<i>Data analysis</i>	94
3.3	STATISTICAL ANALYSIS	97
3.4	RESULTS	98
3.4.1	<i>Calculation of reading slope in MT and SRT</i>	100
3.4.2	<i>Calculation of Normalised Reading Rate (NRR1)</i>	107
3.4.3	<i>Calculation of normalised reading rate (NRR2) using time-points</i>	108
3.4.4	<i>Factors associated with reading slopes, NRR1 and NRR2</i>	109
3.5	DISCUSSION	111
3.6	LIMITATIONS	117
3.7	CONCLUSION	117
4.	INVESTIGATION OF DURATION RELATED CHANGES IN READING SPEED PERFORMANCE IN MIDDLE-	
	AGED AND ELDERLY ADULTS	119
4.1	INTRODUCTION	119
4.2	METHODS	119
4.3	STATISTICAL ANALYSIS	121
4.4	RESULTS	121
4.4.1	<i>Calculation of reading slopes</i>	123

4.4.2	<i>Calculation of normalised reading rate (NRR1)</i>	126
4.4.3	<i>Calculation of normalised reading rate (NRR2) using time-points</i>	127
4.4.4	<i>Factors associated with reading slopes, NRR1 and NRR2</i>	127
4.4.5	<i>Sub-group (Non-completers)</i>	129
4.4.6	<i>Comparison of duration related effects parameters between the young adults and elderly group</i>	132
4.5	DISCUSSION	133
4.6	LIMITATIONS	137
5.	TO INVESTIGATE THE USE OF IREST TO ASSESS THE EFFECT OF DURATION ON READING PERFORMANCE IN A GROUP OF VISUALLY IMPAIRED INDIVIDUALS	140
5.1	INTRODUCTION	140
5.2	METHODS	141
5.2.1	<i>California Central Visual Field Test (CCVFT) (Cole, 2008)</i>	143
5.2.2	<i>Manchester Low Vision Questionnaire (MLVQ) (Harper et al., 1999)</i>	144
5.2.3	<i>International Reading Speed Texts (IReST) (Trauzettel-Klosinski et al., 2012)</i>	145
5.3	DATA ANALYSIS	146
5.3.1	<i>Reading slope</i>	146
5.3.2	<i>Normalised Reading Rate (NRR2)</i>	146
5.4	STATISTICAL ANALYSIS	147
5.5	RESULTS	147
5.5.1	<i>Reading slope</i>	147
5.5.2	<i>Normalised Reading Rate (NRR2)</i>	151
5.5.3	<i>Factors associated with reading slopes and NRR2</i>	151
5.5.3.1	Age	152
5.5.3.2	CCVFT	152
5.5.3.3	MLVQ	153
5.5.3.4	Reading speed	155
5.5.3.5	Retrace Time	156
5.6	DISCUSSION	158

5.7	LIMITATIONS	162
5.8	CONCLUSION	163
6.	THE USE OF SPONTANEOUS EYE BLINK RATE (SBR) TO CAPTURE FATIGUE EFFECT IN READING	164
6.1	LITERATURE REVIEW	164
6.1.1	<i>Introduction</i>	164
6.1.2	<i>Types of blink</i>	164
6.1.3	<i>Assessment of blinks</i>	165
6.1.4	<i>Factors affecting SBR</i>	166
6.2	METHODS	171
6.2.1	<i>Participants</i>	171
6.2.2	<i>Procedures</i>	172
6.2.3	<i>Conditions</i>	173
6.2.4	<i>Data analysis</i>	174
6.3	STATISTICAL ANALYSIS	175
6.4	RESULTS	176
6.4.1	<i>Intra-observer repeatability of the computation of the SBR</i>	176
6.4.2	<i>Spontaneous blink rate (SBR)</i>	177
6.4.3	<i>Normalised-SBR relative to habitual reading condition</i>	178
6.4.4	<i>Duration related effect</i>	179
6.4.5	<i>SBR Slopes</i>	181
6.4.6	<i>Normalised-SBR relative to the 2nd minute SBR</i>	182
6.4.7	<i>Reading speed performance</i>	183
6.4.8	<i>Relationship between SBR slopes and reading slopes</i>	184
6.5	DISCUSSION	185
6.6	LIMITATIONS	189
6.7	CONCLUSION	191
7.	SUMMARY OF FINDINGS AND FUTURE WORK	192
7.1	STRENGTHS AND LIMITATIONS	192

7.2	FUTURE WORK	196
8.	REFERENCES	198
9.	APPENDICES	208
9.1	SAMPLES OF READING TEXT	208
9.2	VALIDATION OF DIFFERENT VERSIONS OF MAGAZINE TEXT AND SRT	209
9.3	MULTIPLE-CHOICE QUESTIONS FOR SRT VERSION 1	210
9.4	NONLINEAR GRAPH PLOTS FOR ALL PARTICIPANTS IN SRT MAG (N = 25)	212
9.5	ABSTRACT ORAL PRESENTATIONS – OS9 PAGE 26 FROM EUROPEAN SOCIETY FOR LOW VISION RESEARCH AND REHABILITATION (ESLRR) 2013	216
9.6	R ² VALUES FOR THE ELDERLY PARTICIPANTS USING LOGARITHMIC AND EXPONENTIAL FUNCTIONS ON READING SPEED AGAINST TIME (SRT-MAG)	218
9.7	THE USE OF LOG ₁₀ TRANSFORMATION ON READING SPEED TO COMPUTE READING SLOPE FOR THE P-EVES DATA	219
9.8	GRAPH PLOTS FOR P-EVES PARTICIPANTS WITH NEGATIVE READING SLOPES IN THE FIRST VISIT (N = 13)	212
9.9	GRAPH PLOTS ON BLINKING RATES AGAINST TIME PER MINUTE IN THE ELDERLY PARTICIPANTS IN MAG (N = 26)	223
9.10	POST HOC POWER ANALYSIS USING G*POWER (VERSION 3.1.9.3)	227
9.10.1	<i>Post hoc power analysis for Chapter 2.1 – Exploring the limitation of the inclusion of the pre-verbalisation time in the computation of reading speed.....</i>	<i>227</i>
9.10.2	<i>Post hoc power analysis for Chapter 2.2 – To explore the use of page-turn as a time recording method in a silent reading test</i>	<i>228</i>
9.10.3	<i>Post hoc power analysis for Chapter 2.3 – The use of proofreading to ensure reading compliance in silent reading tests</i>	<i>229</i>
9.10.4	<i>Post hoc power analysis for Chapter 3 – Evaluation of reading endurance in young adults under conditions of simulated visual impairment</i>	<i>230</i>
9.10.5	<i>Post hoc power analysis for Chapter 4 – Investigation of duration related changes in reading speed performance in middle-aged and elderly adults.....</i>	<i>231</i>
9.10.6	<i>Post hoc power analysis for Chapter 6 – The use of spontaneous eye blink rates to capture fatigue effect in reading</i>	<i>232</i>
9.11	SUMMARY OF ETHICS APPROVAL	233

List of Figures

FIGURE 1.1 AN EXAMPLE OF A MNREAD ACUITY CHART GRAPH PLOT SHOWING THE READING SPEED AS A FUNCTION OF PRINT SIZE. THE COMPUTATION OF CPS AND MRS IS HIGHLY DEPENDENT ON CLINICIAN’S INTERPRETATION.	10
FIGURE 1.2 A SAMPLE SENTENCE SHOWING THE DIFFERENT FONTS USED IN XIONG ET AL. (2018B)	20
FIGURE 1.3 A SAMPLE OF BAILEY-LOVIE WORD-READING CHART (BAILEY AND LOVIE, 1980).....	31
FIGURE 1.4 A SAMPLE OF THE INTERNATIONAL READING SPEED TEXT (IREST) (TRAUZETTEL-KLOSINSKI ET AL., 2012)	32
FIGURE 1.5 A SAMPLE OF MINNESOTA READING (MNREAD) ACUITY CHART (MANSFIELD ET AL., 1994)	34
FIGURE 1.6 A SAMPLE OF A PEPPER VISUAL SKILLS FOR READING TEST (BALDASARE ET AL., 1986)	36
FIGURE 1.7 A SAMPLE OF A SINGLE PARAGRAPH FROM THE RATE OF READING (RoR) TEST (WILKINS ET AL., 1996)	37
FIGURE 2.1 TIME TAKEN DIFFERENCE PLOT FOR THE THREE READING TESTS BETWEEN AUDIO RECORDING AND STOPWATCH: BAILEY-LOVIE WORD-READING CHART = ■ (HABITUAL) □ (BOF), IREST = ● (HABITUAL) ○ (BOF), MNREAD ACUITY CHART = ◆ (HABITUAL) ◇ (BOF).	57
FIGURE 2.2 DIFFERENCE PLOT FOR ALL THE READING TESTS BETWEEN AUDIO RECORDING AND STOPWATCH. BAILEY-LOVIE WORD-READING CHART = ■ (HABITUAL) □ (BOF), IREST = ● (HABITUAL) ○ (BOF), MNREAD ACUITY CHART = ◆ (HABITUAL) ◇ (BOF). THE SOLID BLACK LINE INDICATES THE CUT OFF READING SPEED AT 10 WPM.	58
FIGURE 2.3 GRAPH PLOTS OF THE DIFFERENCE IN READING SPEED AGAINST PAGES FOR ALL THE PARTICIPANTS AND ALL THE CONDITIONS (TOP: NOR, MIDDLE BOF, BOTTOM: MAG) SPLIT INTO POSITIVE SLOPES (LEFT COLUMN) AND NEGATIVE SLOPES (RIGHT COLUMN).....	71
FIGURE 2.4 DIFFERENCE IN READING SPEED AGAINST PAGES GRAPH PLOTS FOR PARTICIPANT JY03 (NOR), PP12 (NOR) AND CG16 (MAG). DOTTED LINES: LOGARITHMIC FUNCTION, DASHED LINES: INVERSE FUNCTION.	72
FIGURE 2.5 A SAMPLE OF MULTIPLE-CHOICE QUESTION USED (ALL THE SEVEN QUESTIONS CAN BE FOUND IN APPENDIX 9.3). 77	
FIGURE 2.6 DISTRIBUTION OF THE PARTICIPANT’S SCORE FOR BOTH THE READING AND NON-READING GROUP.....	78
FIGURE 2.7 RECEIVER OPERATING CHARACTERISTIC (ROC) CURVE FOR THE MULTIPLE-CHOICE QUESTIONS SCORES.....	79
FIGURE 2.8 DISTRIBUTION OF THE NUMBER OF PARTICIPANTS WHO ANSWERED CORRECTLY IN EACH MULTIPLE-CHOICE QUESTION	80
FIGURE 2.9 LEFT: DISTRIBUTION OF PROOFREADING SCORES. RIGHT: DISTRIBUTION OF THE NUMBER OF PARTICIPANTS WHO CORRECTLY IDENTIFY THE SPELLING ERROR.	81
FIGURE 3.1 (A) TEXT VIEWING IN BOF (B) TEXT VIEWING WITH BOF USED IN MAG	93
FIGURE 3.2 A DIAGRAM TO ILLUSTRATE THE CALCULATION OF READING SPEED AT EACH 1 MINUTE TIMEPOINT	95

FIGURE 3.3 AN EXAMPLE OF IREST TEXT SHOWING THE SELECTED SECTIONS FOR NRR2 ANALYSIS.....	97
FIGURE 3.4 AN EXAMPLE OF THE GRAPH PLOT OF GROUP MEDIAN READING SPEED AGAINST TIME IN SUCCESSIVE MINUTES (I.E. SRT MAG). A LINEAR REGRESSION FUNCTION USING THE SPSS PROGRAM COMPUTED THE GROUP MEDIAN READING SLOPE.....	100
FIGURE 3.5 EXAMPLES OF HOW THE READING SLOPE IS COMPUTED. LEFT: AN EXAMPLE OF A PARTICIPANT LS02 WITH GRADUAL REDUCTION IN READING SPEED WITH TIME, AND THIS WAS REFLECTED BY A NEGATIVE GRADIENT SLOPE (-3.6 WPM/MIN). RIGHT: AN EXAMPLE OF A PARTICIPANT JY03 WITH GRADUAL INCREASE IN READING SPEED WITH TIME, AND THIS WAS REFLECTED BY A POSITIVE GRADIENT SLOPE (7.14 WPM/MIN).....	100
FIGURE 3.6 AN EXAMPLE OF NONLINEAR CURVE FITTING GRAPH PLOT OF A PARTICIPANT IN SRT MAG. THE SOLID LINE: LINEAR FUNCTION, DASHED LINE: EXPONENTIAL FUNCTION, DASHED AND DOTTED LINE: LOGARITHMIC.....	103
FIGURE 3.7 GRAPH PLOTS ON READING SPEED AGAINST TIME FOR PARTICIPANTS WITH NEGATIVE READING SLOPES IN BOTH MT AND SRT. O: MT ☆: SRT.....	106
FIGURE 3.8 A READING SPEED AGAINST TIME GRAPH PLOT (PARTICIPANT VV07), SHOWING A SUDDEN INCREASE IN THE READING SPEED AT THE SECOND MINUTE.....	108
FIGURE 4.1 THE MEDIAN AND INTERQUARTILE RANGE OF THE READING SPEEDS FOR ALL THE PARTICIPANTS IN NOR, BOF AND MAG.....	123
FIGURE 4.2 READING SLOPE GRAPH PLOTS FOR THE THREE PARTICIPANTS WITH CONSISTENT NEGATIVE READING SLOPES IN ALL THE READING CONDITIONS. EACH ROW REPRESENTS THE GRAPH PLOTS FOR EACH PARTICIPANT IN DIFFERENT READING CONDITIONS (TOP ROW: HK08, MIDDLE ROW: CC30 AND BOTTOM ROW: 44FK). ACROSS THE ROW, GRAPH PLOTS FROM LEFT: NOR, MIDDLE: BOF AND RIGHT: MAG.....	126
FIGURE 4.3 GRAPH PLOTS OF THE READING SPEED AGAINST TIME FOR THE NON-COMPLETERS. TOP: NOR, MIDDLE: BOF, BOTTOM: MAG. NEGATIVE LINEAR READING SLOPES ARE INDICATED BY DASH LINES.....	131
FIGURE 5.1 FLOWCHART OF P-EVES STUDY (TAYLOR ET AL., 2014).....	142
FIGURE 5.2 THE READING SLOPE GRAPH PLOTS FOR 2 SUBJECTS. READING SPEED FOR EACH ROW OF WORDS IS PLOTTED AGAINST THE CORRESPONDING READING ROW NUMBER. SUBJECT 197 SHOWED AN OVERALL NEGATIVE READING SLOPE OF -0.78 SWPM/ LINE. SUBJECT 191 MISSED 5 ROWS OF READING WORDS – ROW 4, 6, 9, 12 AND 13. HENCE, THESE READING ROW NUMBERS WERE OMITTED OUT. SUBJECT 191 SHOWED AN OVERALL POSITIVE READING SLOPE OF 0.03 SWPM/ LINE.....	148

FIGURE 5.3 BLAND-ALTMAN PLOT OF THE DIFFERENCE IN READING SLOPES BETWEEN 1ST VISIT AND 2ND/ 3RD VISIT. ○ = 1 ST AND 2 ND VISIT, ▲ = 1 ST AND 3 RD VISIT.....	149
FIGURE 5.4 READING SLOPE PLOTS FOR SUBJECT 114 AND 186. ○: 1 ST EVALUATION AND ▲: 2 ND EVALUATION. SOLID LINE = 1 ST EVALUATION READING SLOPE, DOTTED LINE = 2 ND EVALUATION READING SLOPE.....	150
FIGURE 5.5 GRAPH PLOTS OF READING SPEED AGAINST ROW NUMBERS. LEFT: PARTICIPANT 153, NRR2: 0.60. RIGHT: PARTICIPANT 178, NRR2: 0.54	151
FIGURE 5.6 NUMBER OF PARTICIPANTS ABLE TO PERFORM READING TASKS. THE DOTTED-BARS AND STRIPED-BARS REPRESENT READING INSTRUCTIONS AND BOOKS RESPECTIVELY.	154
FIGURE 5.7 FREQUENCY IN USING NEAR VISION OPTICAL MAGNIFIERS IN THE LAST 4 WEEKS. 4 = USES VERY OFTEN = MANY TIMES (≥ 5) EACH DAY (5X OR MORE EACH DAY). 3 = USES OFTEN SEVERAL TIMES (1-4) EACH DAY. 2 = USES SOMETIMES (< 1 DAILY BUT AT LEAST ≥ 1 PER WEEK.....	154
FIGURE 5.8 LONGEST DURATION ON THE USAGE OF THEIR NEAR VISION OPTICAL MAGNIFIER AT ANY ONE OCCASION. 4 = ≥ 30 MINUTES. 3 = ≥ 15 MINUTES AND < 30 MINUTES. 2 = ≥ 5 MINUTES AND < 15 MINUTES. 1 = ≥ 1 MINUTE AND < 5 MINUTES. 0 = < 1 MINUTE.	155
FIGURE 5.9 GRAPH PLOTS OF PARTICIPANTS WITH NRR2 LESS THAN 0.9. LEFT: SUBJECT 103; RIGHT SUBJECT 132.....	160
FIGURE 6.1 VIDEO CAMERA ATTACHED TO THE TRIAL FRAME.....	173
FIGURE 6.2 (A) SCREEN-SHOT OF THE VIDEO CAPTURING EYEBLINK (B) USING ONLINE STOPWATCH FOR RECORDING BLINKING OCCURRENCES.....	174
FIGURE 6.3 BLAND-ALTMAN PLOT OF INTRA-OBSERVER REPEATABILITY	177
FIGURE 6.4 THE NUMBER OF BLINKS IN SUCCESSIVE MINUTES FOR YOUNG ADULTS AND ELDERLY GROUP IN NOR, BOF AND MAG FOR A READING DURATION OF 10 MINUTES. TOP ROW SHOWS THE DATA FOR THE YOUNG PARTICIPANTS FOR NOR, BOF AND MAG (LEFT TO RIGHT). THE BOTTOM ROW SHOWS THE DATA FOR THE ELDERLY PARTICIPANTS FOR NOR, BOF AND MAG (LEFT TO RIGHT).	180
FIGURE 9.1 A SAMPLE OF MT READING TEXT (FONT SIZE AND SPACING NOT UP TO SCALE)	208
FIGURE 9.2 A SAMPLE OF SRT READING TEXT (FONT SIZE AND SPACING NOT UP TO SCALE).....	208
FIGURE 9.3 SAMPLE GRAPH PLOTS SHOWING THE READING SLOPE USING READING SPEED AND LOG ₁₀ READING SPEED TRANSFORMATION	219

List of Tables

TABLE 1.1 DEFINITION OF MAXIMUM READING SPEED (MRS) AND CRITICAL PRINT SIZE (CPS) (PATEL ET AL., 2011). METHOD A WAS USED IN VIRGILI ET AL. (2004A), VIRGILI ET AL. (2004B) AND SUBRAMANIAN AND PARDHAN (2009).....	10
TABLE 1.2 VISUAL REQUIREMENTS FOR READING (WHITTAKER AND LOVIE-KITCHIN, 1993)	27
TABLE 1.3 COMPARISONS OF THE FEATURES OF THE VARIOUS READING ASSESSMENT TESTS	41
TABLE 1.4 COMPARISON OF STUDIES INVESTIGATING THE EFFECT OF AGE ON READING SPEED PERFORMANCES IN NORMALLY-SIGHTED INDIVIDUALS	45
TABLE 2.1 NUMBER OF PARTICIPANTS MAKING THE WORD ERRORS	56
TABLE 2.2 DESCRIPTIVE STATISTICS ON THE READING SPEED USING AUDIO RECORDING AND STOPWATCH MEASURES FOR THE THREE READING TESTS.....	58
TABLE 2.3 MEDIAN AND INTERQUARTILE RANGE OF READING SPEED DIFFERENCES BETWEEN AUDIO RECORDING AND STOPWATCH IN NOR AND BOF.....	59
TABLE 2.4 DIFFERENCES IN READING SPEED BASED ON SAME TIME DIFFERENCE	62
TABLE 2.5 DESCRIPTIVE STATISTICS OF THE READING SPEED BETWEEN AUDIO RECORDING AND PAGE-TURN.....	69
TABLE 2.6 R ² VALUES FOR LEAST SQUARES LINEAR, LOGARITHMIC AND INVERSE FUNCTION.....	70
TABLE 2.7 COORDINATES OF THE CURVE FOR THE MULTIPLE-CHOICE QUESTIONS SCORES.....	79
TABLE 3.1 MEDIAN READING SPEED FOR ALL READING TESTS	99
TABLE 3.2 DESCRIPTIVE STATISTICS OF THE READING SLOPES FOR MT AND SRT (N = 26).....	101
TABLE 3.3 R ² VALUES FOR MT AND SRT IN ALL THE CONDITIONS	102
TABLE 3.4 R ² VALUES FOR LINEAR, LOGARITHMIC AND EXPONENTIAL FUNCTION FOR PARTICIPANTS READING SPEED AGAINST TIME IN MAG (N = 26)	103
TABLE 3.5 NUMBER OF PARTICIPANTS SHOWING POSITIVE AND NEGATIVE READING SLOPES IN MT AND SRT (N = 26)	105
TABLE 3.6 MEDIAN NRR1 FOR MT AND SRT	107
TABLE 3.7 MEDIAN NRR2 FOR ALL READING TESTS	109
TABLE 3.8 SPEARMAN CORRELATION COEFFICIENTS BETWEEN CLINICAL MEASURES AND DURATION RELATED EFFECTS INDICATORS – READING SLOPES, NRR1 AND NRR2 (N = 26).....	110
TABLE 3.9 SPEARMAN CORRELATION COEFFICIENTS BETWEEN IREST, ROR AND MT NRR2 AND THE DURATION RELATED EFFECTS INDICATORS (TWO-TAILED) (N = 26).....	110
TABLE 4.1 DEMOGRAPHICS AND CLINICAL MEASUREMENTS BETWEEN THE TWO AGE GROUPS	122

TABLE 4.2 DESCRIPTIVE STATISTICS ON THE READING SLOPES FOR ALL THE ELDERLY PARTICIPANTS (N = 31)	124
TABLE 4.3 CLINICAL PARAMETERS OF PARTICIPANTS WHO SHOWED A CONSISTENT NEGATIVE READING SLOPES IN ALL THREE CONDITIONS	125
TABLE 4.4 DESCRIPTIVE STATISTICS OF NRR1 AND NRR2 IN THE MIDDLE-AGED AND ELDERLY ADULTS	127
TABLE 4.5 SPEARMAN CORRELATION COEFFICIENTS BETWEEN CLINICAL PARAMETERS AND DURATION RELATED EFFECTS INDICATORS (TWO-TAILED)	129
TABLE 4.6 DETAILS OF THE READING TESTS UNDERTAKEN BY PARTICIPANTS WHO DID NOT COMPLETE ALL THE READING TESTS	130
TABLE 4.7 COMPARISON OF DURATION RELATED EFFECTS PARAMETERS BETWEEN THE YOUNG ADULTS (N = 26) AND THE ELDERLY GROUP (N = 31)	132
TABLE 4.8 COMPARING THE DEMOGRAPHICS AND CLINICAL PARAMETERS BETWEEN THE CURRENT STUDY AND THE PILOT STUDY	136
TABLE 5.1 CCVFT GRADING SYSTEM.....	144
TABLE 5.2 PURPOSE OF THE NEAR VISION OPTICAL MAGNIFIER	144
TABLE 5.3 FREQUENCY OF NEAR VISION OPTICAL MAGNIFIER USAGE.....	145
TABLE 5.4 DURATION OF NEAR VISION OPTICAL MAGNIFIER USAGE	145
TABLE 5.5 PARTICIPANTS WITH POSITIVE AND NEGATIVE READING SLOPE (N = 51).....	150
TABLE 5.6 PEARSON'S CORRELATION COEFFICIENTS BETWEEN THE DURATION RELATED EFFECTS INDICATORS AND CLINICAL MEASUREMENTS OR CHARACTERISTICS (TWO-TAILED) (N = 61)	152
TABLE 5.7 THE NUMBER OF PARTICIPANTS WITH CENTRAL FIELD DEFECTS.....	153
TABLE 5.8 DESCRIPTIVE STATISTICS ON INDEPENDENT-SAMPLES T-TEST COMPARING PARTICIPANTS WITH 1) POSITIVE AND NEGATIVE READING SLOPES 2) $NRR2 \geq 1.0$ AND $NRR2 < 1.0$	157
TABLE 6.1 SUMMARY OF STUDIES COMPARING SPONTANEOUS BLINKING RATE (SBR) IN READING AND OTHER ACTIVITIES ...	168
TABLE 6.2 DESCRIPTIVE STATISTICS ON THE SBR FOR THE YOUNG AND ELDERLY GROUP	178
TABLE 6.3 MEDIAN RELATIVE RATIO OF SBR–BOF AND SBR–MAG COMPARED TO SBR–NOR.....	179
TABLE 6.4 MEDIAN BLINKING SLOPES AND THE R^2 VALUES OF THE YOUNG ADULTS AND ELDERLY PARTICIPANTS.....	182
TABLE 6.5 MEDIAN NORMALISED RELATIVE RATIOS TO THE 2ND MINUTE SBR IN NOR, BOF AND MAG FOR THE YOUNG AND ELDERLY GROUP	182

TABLE 6.6 MEDIAN RATIO COMPARING THE SECOND SEGMENT OF THE SBR RELATIVE TO THE FIRST SEGMENT SBR IN NOR, BOF AND MAG FOR THE YOUNG AND ELDERLY GROUP.....	183
TABLE 6.7 MEDIAN READING SPEED, READING SLOPE AND R ² VALUE FOR THE YOUNG ADULTS AND ELDERLY ADULTS.....	184
TABLE 6.8 POST HOC WILCOXON SIGNED-RANK TEST FOR READING SPEED AND READING SLOPE IN NOR, BOF AND MAG FOR THE YOUNG ADULTS AND ELDERLY GROUP.....	184
TABLE 9.1 COMPARISON OF READING SPEED IN ALL THE DIFFERENT VERSION OF MT AND SRT.....	209
TABLE 9.2 R ² VALUES FOR THE YOUNG PARTICIPANTS USING LOGARITHMIC AND EXPONENTIAL FUNCTIONS ON READING SPEED AGAINST TIME (SRT-MAG).....	215
TABLE 9.3 DESCRIPTIVE STATISTICS FOR THE R ² COMPUTATION USING READING SPEED AND WITH LOG ₁₀ TRANSFORMATION	220

Word Count: 64, 793 words

Abstract:

Recent evidence suggests that a decline in reading speed with time due to fatigue can be observed in an extended reading test, although only in individuals with a mild visual impairment. This effect could be relevant to the leisure reading ability in individuals with visual impairment. To date, there is no study investigating duration related effects in participants using low vision aids (LVA) for a reading task. This study aims to investigate if any negative duration related effects on the reading speed can be identified within a 15-minute reading test and whether the measurement parameters can predict the success of a LVA.

The investigation was conducted in 26 young and 31 elderly normally-sighted adults, and 61 visually impaired participants. The normally-sighted participants were tasked to perform an approximately 15-minute silent reading task with habitual vision, with simulated loss of acuity and contrast sensitivity, and with the use of a hand-held magnifier. Reading slope was computed as the linear slope of the reading speed against time; NRR1 is the median ratio of the successive minutes reading speed relative to the reading speed at the second minute; and NRR2 is the ratio comparing the reading speed at the endpoint to the beginning of the reading test. A negative reading slope or NRRs of <1.0 would indicate a decline in reading speed with time. Another measurement of fatigue using spontaneous blinking rate (SBR) was also explored. The SBR slopes was computed based on the linear slope of SBR against time.

Most of the young adults and elderly participants showed an increase or almost constant reading speed with time in a silent reading test based on the reading slopes (Young: 0.37, IQR:2.10 wpm/min; Elderly: 0.43, IQR:2.24 wpm/min), median NRR1 (Young: 1.14, IQR:0.25; Elderly: 1.01, IQR: 0.43), and median NRR2 (Young: 1.13, IQR:0.36; Elderly: 1.02, IQR: 0.52) in the condition with the use of a hand-held magnifier. In the same condition, the SBR remains relatively constant with time (Young: 0.00, IQR:0.49 SBR/min; Elderly: -0.01 , IQR: 0.31 SBR/min) and the SBR slopes were poorly associated with reading slopes for all the participants ($r=0.24$, $p=0.08$). Likewise, the visually impaired participants showed an overall mean reading slope of 1.40 ($SD \pm 1.87$) swpm/ row of words and a mean NRR2 of 1.13 ($SD \pm 0.37$) suggesting an increase in reading speed with time using their optical magnifiers. Forty-eight (79%) visually impaired participants showed a positive reading slope, with almost half of them ($n = 30$) showing repeated positive reading slopes in the second visit. The duration related effects indicators were not associated with duration of leisure reading in everyday life (reading slope: $r = 0.06$, $p= 0.62$, NRR2: $r = -0.09$, $p= 0.49$) or frequency of LVA usage (reading slope: $r = 0.06$, $p= 0.62$, NRR2: $r = -0.01$, $p= 0.96$) in the visually impaired participants.

In conclusion, negative duration related effects cannot be captured within a 15-minute reading test. The duration related effects indicators (i.e. reading slopes and NRR) were not able to predict the success of magnifier usage in a clinical setting.

Declaration

This is to certify that the work contained in this thesis represents my own work. To the best of my knowledge and belief, no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

- I. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
- II. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
- III. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
- IV. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see <http://www.library.manchester.ac.uk/about/regulations/>) and in The University's policy on Presentation of Theses

Acknowledgements

Firstly, I would like to express my most profound gratitude and heartfelt thanks to my supervisor, Professor Christine Dickinson. I am grateful to her for all the early morning skype discussions on the project and her meticulous comments on my drafts in preparation for this final thesis work. Without her patience and guidance, this thesis would not be possible.

I would like to thank the p-EVES study group – Professor Christine Dickinson, Dr Jake Taylor and Ms Rachel Bambrick, who have allowed me to use the data for analysis as part of this thesis.

Special thanks go to the former Diploma in Optometry Course Chair, Dr Narayanan Rajeev, and current Academic Resource Manager, Miss Lim Hui Shan, for their continuous support during my PhD candidature. I would like to thank my optometry students who completed their projects under my supervision from 2014 to 2017 – they have helped me to recruit the subjects during the different phases of this research work. I would also like to thank the participants in my study conducted in both Manchester and Singapore.

Last but not least, I would like to thank my family members (especially my sisters, Lai Sun and Wei Ling, and my late foster dad, Chan Weng Siew) who have supported me spiritually throughout writing this thesis. I would like to extend my gratitude to my close friends, who have given me continuous support and encouragements during these several years: Andrew Quek, Laurence McAllister, Jeff Hoe and Sharon Chua.

1. Literature Review

1.1 Introduction

Reading is one of the most important tasks with which visually impaired individuals find difficulty, and for which they request help (Elliott et al., 1997). For the past few decades, the majority of research studies on reading speed performances in low vision have been limited to short reading tests to minimise the burden on the visually impaired participants. These typical short reading assessments are adequate in the assessment of the ability of the individual to perform spot or survival reading (e.g. reading bills, price tags, restaurant menus, etc.). However, little is known about the ability to sustain reading speed over a longer reading duration, also known as reading endurance (Rubin, 2013). This would be relevant in leisure reading. To date, there has been little investigation on reading endurance and how this might affect the ability of individuals with visual impairment attempting to continue leisure reading activity. Any possible negative effects of prolonged reading such as fatigue (i.e. reading slowing with time) could affect the duration of use of their optical aids – which could be very relevant to how useful these optical aids can be. It is the purpose of this study then to investigate any duration related effects in prolonged reading, especially when the reading task is challenging (i.e. with visual impairment, and while handling a magnifier). Slowed reading could be a consequence of fatigue, but fatigue per se is very subjective and can be hard to quantify accurately. Previously, Stern et al. (1994) have shown that blinking rates can be used as a sensitive indicator of fatigue. This study, therefore, also explored the possibility of using blinking rates to objectively quantify fatigue during reading.

This literature review chapter takes the form of six sections. Section 1.1 will cover an overview of the different stages of the reading process and discuss the benefits of using reading speed as a clinical performance indicator in this study. In Section 1.2, the different experimental approaches such as the types of reading material (i.e. contextual vs random words), the assessment approach (i.e. oral vs silent reading), and how these approaches can influence on the reading speed measurements were reviewed. Followed by Section 1.3, the four visual requirements for reading will be discussed. Section 1.4 presents the details and characteristics of the various reading tests commonly used in low vision research. Section 1.5 compares the previous studies on the effect of

age on reading speed performance. Pen-ultimately, Section 1.6 will discuss the use of low vision simulation used in low vision studies. The final section will provide an overview of the experiments and analysis conducted for this thesis.

1.1.1 Nature of the reading process

Reading is a highly complex skill that allows a message to be conveyed to the reader in everyday life. People read for different purposes; it can be for leisure (i.e. a book or magazine), knowledge (i.e. textbooks or reference books) or information (i.e. a map or medicine labels). In a previous landmark study by Carver (1992), the author categorised reading into five different gears based on the intention of reading. These five reading gears can be differentiated by using reading speed, which is measured in the number of words read per minute (wpm). The five reading gears are scanning (600 wpm), skimming (450 wpm), rauding (300 wpm), learning (200 wpm) and memorising (138 wpm). The five reading gears build on each other to describe the different levels of information acquisition by the reader. Scanning, which operates at the highest reading speed, refers to looking for a specific word in a passage. Followed by skimming, which extracts meaning from a few words in relation to one another. Rauding, which is the most often used reading gear, comes from a combination of two words, reading and auditing. Rauding refers to the comprehension of the text when presented either visually or aurally. Penultimately, learning refers to thinking and continually revisiting the received ideas. The last gear, memorising, requires the reader to go through the text material several times to be able to recall orally or in writing.

1.1.2 An overview of reading performance indicators as outcome measures

Previous studies (Elliott et al., 1997, Owsley et al., 2009) have shown that reading has been a priority for low vision rehabilitation. Since low vision aids (LVAs) are commonly prescribed to improve reading performance in the low vision community (Watson et al., 1997), a definitive outcome measure would be useful to determine the effectiveness of the LVAs. The most straightforward approach is to gather feedback from the patient's experience with the prescribed LVAs (e.g. frequency and duration of usage, types of task, etc.). However, these responses are highly subjective and do not describe the effectiveness of reading. A structured and detailed assessment questionnaires on the usage profile of the LVAs such as the Manchester Low Vision

Questionnaire (MLVQ) (i.e. frequency of use, length of continuous use, ease of use of the LVAs and ability to carry out specific task with the LVAs) (Harper et al., 1999) can be used but the answers are still very specific to that particular individual and their requirements. A review of physical functioning measures by Guralnik et al. (1989) has proposed that performance-based tests can provide earlier predictors of functional decline and disability compared to self reports.

Since an outcome measure is required to evaluate if the LVA had a beneficial effect on reading performance, it is not surprising that standardised reading performance indicators such as reading acuity (RA), reading speed, critical print size (CPS), reading accessibility index (ACC), reading comprehension, reading duration and reading accuracy have been used. However, it is unclear if these measurements can provide us with a summary of the actual performance in real-world conditions. A previous survey by McIlwaine et al. (1991) reported that 33% of the participants given an aid in one specific low vision clinic did not continue to use the prescribed LVA at home. This observation was supported by Leat et al. (1994), who found that while 75% of the low vision participants could read small print during the clinical testing, only 35% of them reported being able to perform a similar reading task (i.e. same print size) at home.

Following the extensive work by Gordon Legge and collaborators in the Psychophysics of Reading paper series, a reading test, currently known as the Minnesota Reading (MNREAD) acuity chart was developed. The design and properties of the MNREAD acuity chart will be discussed in Section 1.4.4. With the development of this chart, there were initially three reading performance parameters (i.e. RA, CPS, maximum reading speed (MRS)) which can be computed using the MNREAD acuity chart (Mansfield et al., 1994). More recently, a fourth MNREAD reading performance measurement, ACC, has been reported by Calabrèse et al. (2016).

1.1.2.1 Word reading acuity (RA)

Reading acuity (RA) refers to the threshold of near reading acuity. There are a few ways in which the assessment could be done (e.g. using contextual vs random words, different scoring algorithms, etc.) and it is unclear whether the approach used has an impact on the reading acuity. The Bailey-Lovie word-reading acuity chart uses random words, with the larger print sizes (i.e.

more than 1.0 log MAR at 25 cm) having only two to three words at each logarithmic progression up to 1.6 log MAR. Although the print sizes are limited by the size of the chart, these inconsistent number of words may create an unfair scoring compared to the smaller print sizes (i.e. less than 1.0 log MAR at 25 cm). The characteristics of the Bailey-Lovie word-reading acuity chart will be discussed in detail in Section 1.4.1. Single contextual sentences in logarithmic progression were used in the MNREAD acuity charts and Radner reading charts. Mansfield et al. (1993) defined RA as the smallest print size read without making a significant error – although it was not stated what constitutes a “significant” error (Mansfield et al., 1993). Maaijwee et al. (2008) suggest that the RA should be based on a minimum of 80% correctly read words. This proportion of correctly read words was referenced from preliminary data gathered on a set of pre-determined retinal eccentricities and varying print sizes by Chung et al. (1998). However, the scoring of the MNREAD acuity chart, which is suggested by the manufacturer, is different from the proposed definition. Each MNREAD acuity chart sentence consists of 60 characters, which is equivalent to 10 standard length words (i.e. six characters). Unlike the proposed definition, the scoring system allows the RA to be measured precisely up to 0.01 log MAR by using the formula:

$$RA = 1.4 - (\text{sentences} \times 0.1) + (\text{errors} \times 0.01)$$

RA has been shown to be a better predictor compared to distance VA for CPS in the low vision community (Legge et al., 1985b). The distance VA uses letter optotypes, which can be a very different task in identifying words at near. Xiong et al. (2018a) reported that the RA was significantly worse in participants with macular dysfunction compared with participants with non-macular visual impairment despite having the same distance letter acuity. The RA is a function of magnification, and the change in RA is dependent on the capabilities of the LVA and how the LVA is used; which may limit the use of RA as an indicator for duration related effects.

1.1.2.2 Reading speed

Reading speed, which is measured in words per minute (wpm), is probably the most commonly used reading performance indicator in low vision research studies. It has been shown to be the best predictor of visual ability to perform daily living tasks (e.g. reading a newspaper) in comparison

to other clinical measurements such as distance VA, near VA and contrast sensitivity (McClure et al., 2000). Unlike other clinical measurements such as distance VA, near VA or contrast sensitivity, reading speed measurement comes from the reading task, which may potentially produce a higher association with self-reported reading difficulty. In a recent Cochrane Review by Virgili et al. (2018), the authors reported that reading speed was used as an outcome measure in assessing the reading performance in the low vision community with different LVAs across 13 studies, compared to the 23 studies which used other outcome measures such as reading comprehension and questionnaires. However, there is no consensus on the approach to how this parameter should be measured and how this measurement translates to actual real-world performance (i.e. the validity of the reading speed measurement). Although Friedman et al. (1999) reported that the majority of the participants' self-reported reading difficulty agrees with their reading speed performances (e.g. participants with a reading speed of more than 160 wpm reported no to minimal reading difficulty, etc.), a discrepancy between self-reported reading difficulty and reading speed was reported in 10% (n = 227 out of 2280) of all the participants. Their findings suggest that perceived performance does not always match up to the reading speed. For the ease of ensuring reading compliance, and to simplify the measurement of the reading speed, reading out-loud is commonly used rather than silent reading. The instructions given to the participant to read "as fast and as accurately as possible" and "not to return to missed words" forces the participant to read at their fastest reading speed (Legge et al., 1989a), and encourages a more repeatable measurement in a short reading test. However, these approaches do not mimic the actual performance in real-world conditions. There are attempts to make the reading performance more "real-world", by changing the instructions to the participants (e.g. "to read the text at a normal rate to obtain meaning from the sentence" (Lovie-Kitchin et al., 2000)).

MRS is defined as the maximum reading speed achieved by an individual when the "print size is not a limiting factor" (Mansfield et al., 1994), and this is also used to determine the CPS for an individual. By saying the "print size is not a limiting factor", it refers to the larger print sizes where the reading speed plateaus (Legge et al., 1985a, Legge et al., 1985b). Clinically, the MRS and CPS provide guidance (e.g. the amount of magnification to achieve optimal reading speed) to the clinician when prescribing LVAs. Although reading speed can also decline with very large print

sizes (i.e. larger than 2° or 3° (Legge et al., 1985a, Akutsu et al., 1991)), it is less relevant and rarely feasible to prescribe LVAs to achieve such large print sizes in a clinical setting. Using the graph paper provided by the manufacturer, the reading speed can be read off the graph by plotting the corresponding time recorded for each sentence (see Figure 1.1). Mansfield et al. (1996) have shown that the reading speed plateaus for normally-sighted individuals over a wide range of sizes. The mean reading speed obtained from this plateau across the intermediate print sizes is then computed as the MRS. However, this is not as straightforward in visually impaired individuals. The reading speed data points are more scattered, and fewer in number, in visually impaired individuals (Burggraaff et al., 2010), making the interpretation of the MRS and CPS more subjective.

The main advantage of using reading speed as an outcome measure is that it can be easily monitored continuously – a feature that would be useful in assessing for any duration related changes. Previously, Ramulu et al. (2013a) have shown that reading speed can be assessed continuously and quantifiably over an extended time of 30 minutes. Subsequently, the measurements can be charted to identify any changes in the performance throughout the reading duration. Although compared to the other reading performance parameters, there is currently no “standardised” commercially available reading text created for extended reading (i.e. as compared to having a standard reading chart such as the MNREAD acuity chart). A commercially available validated extended reading text can be seen in IReST, but it has only up to 150 words per reading text. However, the use of non-standardised text within the reader’s reading ability does not pose a problem since the purpose is to observe the trend of the reading speed changes with time. As noted earlier, reading speed measurement can vary substantially in participants with visual impairment (Burggraaff et al., 2010), and this may potentially affect the reproducibility of any possible effect.

1.1.2.3 Critical Print Size (CPS)

CPS is defined as the smallest print size that a patient can read at the MRS. Although the term CPS refers to a “print size”, it is often expressed in log MAR notation rather than the physical print size. While the log MAR notation provides information on the minimum angle of resolution, the information may be somewhat less useful for a clinician as compared to a print size notation (e.g.

N-notation), which can provide information on the type of text accessible by the low vision individual. The definition of the CPS and MRS are connected, but actual data from the low vision individuals showed that there are a few different ways in which the parameters can be defined and computed (Patel et al., 2011). A clinician can either use the CPS as a guideline on the minimum print size required for comfortable reading, or a set of fixed reference acuity reserves (i.e. the ratio of actual print size to the subject's threshold print size) for different modes of reading. For instance, a maximum or near maximum reading speed would require an acuity reserve of 3:1 (Lovie-Kitchin and Whittaker, 1999). Cheung et al. (2008) have shown that the actual requirement for CPS tends to be bigger than the fixed reference acuity reserve in a small sample size of 14 participants with central vision loss. Large differences between using fixed acuity reserve and CPS have been reported in larger scale studies ($n = 100$ (Latham and Tabrett, 2012), $n = 58$ (Xiong et al., 2018a)), and this difference in range appears to be bigger in participants with central vision loss (Xiong et al., 2018a). Latham and Tabrett (2012) reported an optimum acuity reserve can range from 1.0:1 to 9.1:1 (median: 2.0:1, IQR: 1.58 to 2.50) in a heterogenous group of low vision participants, predominantly with macular problems. An acuity reserve of 2:1 (i.e. 1.58 to 2.50:1) would suffice to achieve close to the MRS in 50% of the participants, but 25% of the participants required a higher acuity reserve (i.e. above 2.5:1) and the remaining 25% of the participants required a lower acuity reserve (i.e. below 1.58:1). Latham and Tabrett (2012) reported that the participants who required higher optimal acuity reserve tended to have a better aided reading acuity (i.e. 1.2M or smaller). While fixed acuity reserve is appropriate for a majority of low vision patients, clinicians should be aware that some patients with visual impairment may benefit from greater acuity reserves to achieve a faster reading speed.

The manufacturer of the MNREAD acuity chart suggests the use of a graph plot to determine the CPS. Figure 1.1 shows an example of a typical MNREAD acuity chart graph, which uses reading speed as a function of the print size. The reading speed remains at a plateau for a range of print sizes, thereafter, the reading speed will start to decrease after reaching the CPS. As seen from Figure 1.1, the computation of MRS and CPS can be very subjective (more so in a low vision patient where the data points can be very variable) and is dependent on how the clinician defines these parameters. Previously, Legge et al. (1985a) used the two-limb function, which comprised

of two straight lines, to estimate MRS and CPS. The MRS was represented by the best straight line across the plateau. A second best straight line through the smaller print sizes that produce a reduction in reading speed was fitted. The intersection between the two lines represents the CPS. As the CPS is influenced by how the sloped line is fitted, noisy reading speed measurements near the threshold acuity print size can have a large effect on the CPS (Mansfield et al., 1996). To rectify this problem, Cheung et al. (2008) used an exponential decay function to fit the reading speed graph in a group of normally-sighted and low vision participants. However, the average differences between the two-limb, and exponential decay function derived, CPS that supports 80% and 90% of MRS were 0.07 log MAR (0.05 to 0.08 log MAR) and 0.15 log MAR (0.13 to 0.18 log MAR) respectively. In both the CPS definitions (i.e. supports 80% or 90% of MRS), the exponential decay function generated a bigger CPS compared to the two-limb method. The reported data were only limited to normally-sighted participants group and it is uncertain if this marginal difference also applies to the low vision community. More importantly, it should be considered if such a minimal difference is clinically relevant in a low vision consultation. Although similar computation methods (i.e. CPS that supports 80% and 90% of MRS) were also reported by Patel et al. (2011) in a heterogenous group of visually impaired participants, the authors did not adopt the use of exponential decay function for curve fitting. Cheung et al. (2008) have demonstrated the use of nonlinear mixed effect (NLME) modelling to estimate reading speed and CPS even with sparse data, which is commonly seen in patients with low vision. However, the modelling required prior data of a specific population group, and Cheung et al. (2008) data were limited by the small sample size ($n = 14$). Indeed, a higher variability in CPS using NLME compared to traditional graph plotting was reported in a heterogenous group of visually impaired participants (Burggraaff et al., 2010). Although the web version of the statistical software is available, the computation of the parameters may not be practical in an outpatient clinic (Patel et al., 2011).

Another common approach would be the use of a formula to calculate MRS and CPS, but there are a few ways in which these parameters can be calculated. Patel et al. (2011) have compared and investigated the repeatability of the various methods used to compute MRS and CPS reported in literatures (Virgili et al., 2004a, Virgili et al., 2004b, Subramanian and Pardhan, 2009). Table 1.1 shows the various methods used to calculate MRS and CPS. Two separate visits were conducted

to evaluate the repeatability of the calculation methods. From the results, Method A and B2 produce the most repeatable MRS and CPS respectively. The coefficients of repeatability for all MRS calculation methods ranged from 66 to 94 wpm (Method A: 66 wpm), and all CPS calculation methods 0.44 to 0.67 log MAR (Method B2: 0.44 log MAR). However, the difference for all methods did not reach statistical significance between the two visits and the power analysis to detect a difference was not reported. It is interesting to note that the AMD participants reading speeds (mean MRS :157 to 187 wpm) in Patel et al. (2011) study were not typical of those in some previous low vision studies (e.g. reading speed between 25-30 wpm (Legge et al., 1985b, Krischer and Meissen, 1983)). The higher reading speed with short sentences can potentially be more variable due to a false start on recording the timing (Rubin, 2013), more so in this case where only a stopwatch was used instead of audio recording. Given no statistically significant difference, Patel et al. (2011) proposed a simplified calculation of MRS, which can either use the mean speed of the fastest read three sentences, or the fastest single speed; and CPS can be the smallest print that supports either 80 or 90% of the MRS.

Apart from being a function of magnification, the CPS is also dependent on the amount of acuity reserve required by the reader to achieve the MRS. The acuity reserve has been reported to be variable in individuals with visual impairment (Latham and Tabrett, 2012), and the changes in the reading speed during a reading task may be associated with changes in acuity reserve demand. For instance, it may be possible that the demand of acuity reserve may be higher in individuals experiencing fatigue effect, which can results in the change of reading speed performance with time. So far, no study has been conducted using CPS as a fatigue indicator in reading tasks and further works need to evaluate if the use of CPS is sensitive to changes with time. However, the key limitation of CPS is that the assessment do not allow a continuous measurement across the duration. For instance, the CPS can only be assessed before and after the reading text; that is, the parameter does not allow a quantifiable measurement during reading duration.

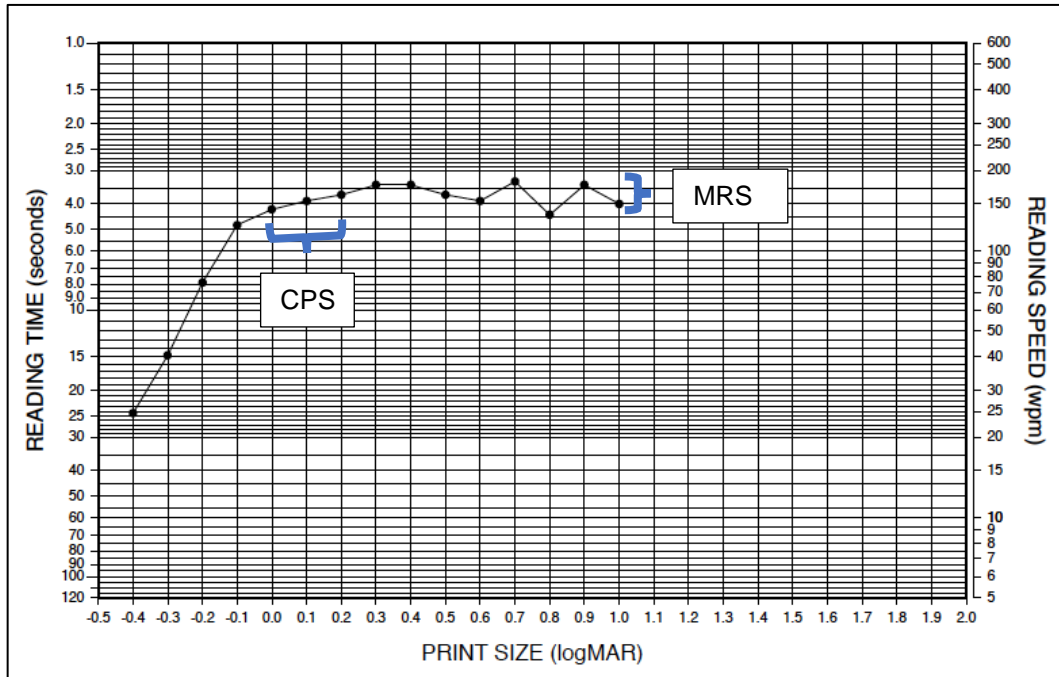


Figure 1.1 An example of a MNREAD acuity chart graph plot showing the reading speed as a function of print size. The computation of CPS and MRS is highly dependent on clinician's interpretation.

Table 1.1 Definition of maximum reading speed (MRS) and critical print size (CPS) (Patel et al., 2011). Method A was used in Virgili et al. (2004a), Virgili et al. (2004b) and Subramanian and Pardhan (2009).

Method	MRS	CPS
A	Mean reading speed for sentences in print larger than the CPS	<p>CPS defined when two of the following criteria are met:</p> <ul style="list-style-type: none"> - All smaller print size sentences than the CPS were read at a speed that was 1.96 times the standard deviation below the average of the previous larger print sizes - All the following sentences were read 5% slower than the average of the previous larger print sizes <p>If these disagreed, CPS was the print size of the largest sentence</p>
B	Mean of reading speed of the fastest 3 sentences read	<ol style="list-style-type: none"> 1. Smallest print size that supports 90% of MRS 2. Smallest print size that supports 80% of MRS
C	The speed of the fastest sentence read	

1.1.2.4 Reading Accessibility Index (ACC)

More recently, a new single-value parameter, ACC, was computed using the MNREAD acuity chart to capture an individual's ability to access a range of print sizes found in everyday life (Calabrèse et al., 2016). Although this range of print sizes (i.e. 1.3 to 0.4 log MAR at 40 cm) may be seen in newspapers (e.g. headlines, column headings, text, etc.) (Legge and Bigelow, 2011), the print sizes do not typically vary substantially across other reading tasks (e.g. a novel might maintain a standard font size of N12). The ACC is calculated using the mean reading speed across the 10 largest print sizes, which corresponds to 1.3 to 0.4 log MAR using a MNREAD acuity chart at 40 cm, and normalised by comparison to a standard mean reading speed value (i.e. 200 wpm) obtained previously from a group of normally-sighted young adults (Calabrese et al., 2016). A value of less than 1.0 indicates reduced access to the range of print text sizes encountered in daily life (Legge and Bigelow, 2011). However, it is unclear how this value is interpreted – given that the value does not directly indicate the print size that was not “accessible” to the reader. Calabrèse et al. (2016) has shown that ACC is a better predictor ($r = -0.60$, $p < 0.001$) of the five reading related daily task activities extracted from the timed instrumental activities of daily living (TIADL) (Owsley et al., 2001) tasks compared to distance VA ($r = 0.24$, $p < 0.001$) or contrast sensitivity ($r = -0.27$, $p < 0.001$). These five activities include reading ingredients on a can, reading directions on a medicine bottle, reading a newspaper article and finding a name and number in a telephone book and doing a one-touch speed dial on a telephone. Since most of these activities were related to reading, we would also expect the RA to be a better predictor compared to distance VA and contrast sensitivity. However, the RA was not included in the study investigation. More recently, Tarita-Nistor et al. (2018) have investigated the use of ACC as an outcome measure for reading rehabilitation in participants with central field loss. The ACC was found to be highly correlated with CPS ($r = -0.66$, $p < 0.001$), RA ($r = -0.78$, $p < 0.001$) and normalised MRS (i.e. reading speed was normalised using 200 wpm at any print sizes) ($r = -0.88$, $p < 0.001$). An evaluation of the impact of LVAs provision on the ACC has also been reported by Latham (2018). The use of LVA shows a statistically significant improvement of ACC from 0.31 to 0.47 in 92% of the participants. Although a statistically significantly improved ACC of 0.16 with LVA suggest that the prescribed LVA offers an improvement in reading accessibility, it was not established whether the level of improvements indicates the usefulness of these LVA to the patient (e.g. does a higher difference

in ACC indicates that the patient will more likely to use the LVA?). While the single value measure allows an easier computation compared to MRS and CPS, the ACC was calculated based on a limited range of print sizes (i.e. 1.3 to 0.4 log MAR at 40 cm). In other words, greater improvement in reading acuity than 0.4 log MAR will not be reflected by the ACC (Latham, 2018). Also, Tarita-Nistor et al. (2018) have reported that the ACC does not provide details on low vision patients with different reading abilities. For instance, the same ACC can be achieved by either someone who can read fast over a smaller range of print sizes or someone who can read the full range of print sizes slowly. Further work needs to be extended to develop a clearer interpretation of the ACC as an outcome measure for reading rehabilitation, and to determine how this parameter can provide a better understanding of the reading performance in the low vision community.

1.1.2.5 Reading comprehension

In contrast to the other reading performance indicators, the assessment of reading comprehension is perhaps the most valid measurement considering that the objective of reading is to extract information from the text. Reading comprehension in low vision related studies has been assessed through the use of multiple-choice questions (Dickinson and Rabbitt, 1991, Legge et al., 1989b), free recall (i.e. the individual was asked to describe what they have read from the reading passage, and the number of propositions was counted (Dickinson and Rabbitt, 1991)) and cloze procedure (Watson et al., 1996). The use of multiple-choice questions allows the benefit of easy scoring but still allows a certain percentage of scoring through guessing. Also, to set a question with valid options can be extremely tricky (e.g. the answer options must only be picked up through reading the text, and not by general knowledge or logical choices). It is difficult and extremely time-consuming to analyse free recall. Cloze procedure does provide some form of “quantification” of the scores, but a higher score may result from better linguistic skills rather than the visual ability. The assumption usually has to be made that the individual had good comprehension skills before the onset of their visual impairment. Early works by Baldasare and Watson (1987) postulated that the reduced reading speed in individuals with central vision loss could possibly contribute to the difficulty in reading comprehension. Subsequently, Watson et al. (1992) have shown that practice and training in reading comprehension can help individuals with macular loss to improve their level of reading comprehension. On the other hand, Legge et al. (1989b) reported that there was no

significant difference in the level of comprehension between the normally-sighted and low vision participants using multiple-choice questions, despite the markedly reduced reading speed in the latter. In the Watson et al. (1992) study, specific lessons to improve cloze passage techniques (e.g. how to use the text information to make guesses) were conducted for the participants. A Woodcock test (i.e. a cloze procedure test) assessment was done before and after the practice and training interventions, and an improvement in score was observed in the participants with the given practice and training. Although the reading speed in the visual skills reading test revealed similar reading speed before and after the practice and training (i.e. suggesting no further improvement on their visual skills using their LVA), one can argue that the better score was a result of the improvement of the cloze procedure skill rather than comprehension. More recently, there was evidence to suggest an improvement of reading comprehension in a group of visually impaired individuals in one arm of a randomised controlled trial to assess the effectiveness of undergoing a 6-weeks eccentric viewing training (EVT) programme (Brown and Rubin, 2016). Brown and Rubin (2016) reported an improvement of 1.1 grade level using the Morgan Low Vision Reading Comprehension Assessment (LVRCA), 6 months after baseline, in the group of visually impaired participants who used their Preferred Retinal Locus (PRL) during three 45-minute sessions over 6 weeks. No improvement was observed in the control (i.e. no visual training) and the trained retinal locus (TRL) group, but a decline of 1.1 grade was reported in the “supervised reading” control group. However, a subsequent report on the 12-month follow-up did not identify any measurable benefits of EVT, and no further details were provided on the comprehension scores (Rubin, 2017). Although the comprehension assessment seems to be a logical approach to measure reading performance, there is no strong evidence to suggest that the assessment of comprehension would be an appropriate outcome measure. The use of “non-standardised” (e.g. newspaper article) (Goodrich and Kirby, 2001, Ortiz et al., 1999, Stelmack et al., 1991) and standardised (e.g. IReST paragraphs) (Morrice et al., 2017) has been reported, but the validity of these multiple-choice questions is questionable. More importantly, reading comprehension can remain unaffected in individuals with visual impairment (Legge et al., 1989b), which suggest that it may not be sensitive enough to show any definitive changes as an outcome measure.

1.1.2.6 Reading duration

The assessment of reading duration can be obtained either through a self-reported questionnaire or the physical measurement of the maximum reading duration. Depending on the types of reading task, the length of the reading duration would indicate a different type of outcome (Taylor et al., 2017). For instance, in leisure reading, a longer reading duration would suggest that the user is more comfortable using the LVA, and therefore, able to read comfortably for a longer duration. On the other hand, in a spot or survival reading, a shorter duration would indicate a positive outcome because this would suggest that the user is able to use the LVA more efficiently. It is, therefore, important to determine the expected positive outcome based on the nature of the reading task. The use of a self-reported questionnaire can be a more valid option compared to measuring the reading duration clinically or experimentally because it translates to the actual performance under real-world conditions. Alternatively, researchers can choose to measure the maximum reading duration physically. Goodrich and Kirby (2001) and Stelmack et al. (1991) defined the reading duration as the participant's reading time without visual discomfort. One of the key limitations of measuring reading duration is the feasibility of adopting the method in a clinical setting. A previous study by Goodrich and Kirby (2001) have shown that the measured reading duration could stretch up to more than 40 minutes in a group of participants with visual impairment performing a reading task with a LVA – suggesting that the assessment duration in a clinical setting may not be practical.

1.1.2.7 Reading accuracy

Reading accuracy refers to an individual's ability to identify the words correctly and is commonly coupled with other reading performance indicators such as reading acuity and reading speed. Although the reading accuracy is not scored independently, poor reading accuracy can affect the coupled reading performance parameters. For instance, reading speed measurements are based on the number of words read correctly, and poor reading accuracy (e.g. missed or wrong words) can cause a reduction in the overall reading speed measurement. The comprehension of meaningful reading text would also be likely to be affected in individuals with poor reading accuracy. The Pepper Visual Skills for Reading Test (VSRT) (Baldasare et al., 1986) is one reading assessment test that can allow the reading accuracy to be scored separately (i.e. the mean percentage of correct words for each completed line of words). Although poor reading accuracy

can be caused by poor visual function, an improvement in reading accuracy may not be indicative of an improved reading performance (e.g. can be due to the cognitive influence). Given the ambiguity interpretation of “improved” reading accuracy, the assessment would not be appropriate as an outcome measure for low vision rehabilitation.

1.1.3 The use of reading speed as a reading performance indicator

In this study, the key objective was to investigate any possible duration related effects in prolonged reading duration. As such, a quantifiable performance-based measure would be required to allow any changes to be detected throughout the test duration. Unlike using a questionnaire to report reading behaviour (i.e. reading duration, ease and frequency of use), reading speed provides a more objective measurement. Although other reading parameters such as RA, CPS and ACC were also objective and quantifiable measurement units, these parameters would be difficult to monitor continuously throughout the duration of a reading test – making it difficult to track the changes with time. In addition, reading speed has been shown to be sensitive to changes within a relatively short time (i.e. 15 minutes in Ramulu et al. (2013a)) Therefore, reading speed was chosen as a measure of functional vision, and any changes in reading speed with time would indicate a duration related effect that may be linked to fatigue.

1.2 Experimental design influence on reading speed assessment

Different approaches in the experimental methods can affect the reading speed assessment. These include presentation methods (i.e. drifting vs static text), context (i.e. contextual vs random words), format (i.e. single sentence vs paragraphs), mode of assessment (i.e. oral vs silent reading), print size, contrast, etc. In relation to the current study, only context, format and mode of assessment effects will be discussed in detail. The print size and contrast will be covered in Section 1.3: Visual factors influencing reading performance.

1.2.1 Text presentation modes

In the earlier works of Legge et al. (1985a) on the psychophysics of reading, the drifting method of text presentation was used to measure reading speed. The words were automatically drifted from right to left of the monitor at a drift rate measured in the number of words per minute. Reading

speed is then computed as the drifting rate multiplied by the percentage of words read correctly (Legge et al., 1989a) (e.g. if 50% of the text is read correctly at a drift rate of 100 wpm, then a reading speed of 50 wpm would be recorded). The drifting text presentation may appear “artificial” as it does not represent the typical presentation of everyday reading material in inkprint. Reading requires a systematic sequence of eye movements which consists of fixation pauses (i.e. averaging 225 ms), followed by forward saccades (i.e. averaging 8 characters in length) and occasionally regressive eye movements (Rayner, 1998). A return sweep is then required to move the eyes from the end of the completed line to the next line of words. In the static presentation used in Legge et al. (1989a), each complete sentence was presented on a television monitor for a fixed time and this fixed time was gradually reduced until 50% word errors were made. The reading speed is computed from the total number of words read correctly divided by the exposure time. Unlike the drifting method, all the words of a single sentence are presented at one time in the static method. As such, a change in viewing distance is required to increase the range of character sizes using the static method. Legge et al. (1989a) have reported an average of 15% faster reading in the drifting method compared to the static method in a heterogeneous group of 27 low vision participants (i.e. central scotoma: $n = 16$, peripheral vision loss: $n = 11$). Also, the regression line suggests that a larger improvement in reading speed with drifting text would be observed in participants with lower reading speed (i.e. less than 30 wpm) using an angular character size of 6° to allow all the participants to achieve their MRS. However, there were only three participants with reading speed less than 30 wpm to support such an observation. Legge et al. (1989a) suggested that the improvement in the reading speed using drifting text in the low vision group could be due to the more erratic eye movements when reading static text in patients with central scotoma. However, there was no significant difference in reading speed between participants with central scotoma ($n = 16$) and those with peripheral vision loss ($n = 11$) in the drifting text mode. It is possible that patients with extreme peripheral vision loss (i.e. tunnel vision) may also benefit from the drifting text. However, the visual fields of the low vision group were not reported by Legge et al. (1989a).

Contrary to Legge et al. (1989a) findings, Bowers et al. (2004) did not find a significant difference in maximum oral reading for drifting text and static text in both normally-sighted ($n = 14$) and

visually impaired participants ($n = 35$, predominantly central field loss). The median age was 74 years (range 34 to 87 years), and the duration of the visual impairment was 13 years (0.5 to 37 years). The participants were younger in Legge et al. (1989a) (median age: 37 years, range 21 to 83 years; the duration of the visual impairment was not specified in the study). Previously, White and Bedell (1990) have reported that the duration of the disease and saccadic latencies were statistically significantly correlated ($r = -0.71$, $p < 0.001$) when observing their eye movements at different fixation points. It may be possible that the duration of the visual impairment could affect how effectively they used eccentric viewing for reading. A more recent study by Walker et al. (2016) showed that reading speed and comprehension were not statistically significantly different between drifted text and static text in a group of 26 participants with longstanding visual impairment (average length of diagnosis: approximately nine years). The drifting text does, however, produced a statistically significantly lower proportion of reading errors compared to static text (i.e. the percentage of text without errors using drifting: 77.9% vs using static text: 72.7%, $p = 0.002$).

The Rapid Serial Visual Presentation (RSVP) method was first used by Forster (1970) in syntactic cognitive studies and was subsequently described by Rubin and Turano (1992) as a method for measuring reading speed without saccadic eye movements. Unlike the static method used by Legge et al. (1989a) where each complete sentence (e.g. ten words) was shown on the screen, only one word at a time in the centre of the monitor screen was presented in RSVP. A random exposure time was initially selected, and each word regardless of length is presented at this "fixed" duration until the end of the sentence. If all words are read correctly, this exposure duration will be 50% shorter. On the other hand, if the words are misread, the exposure duration will then be increased by 50%. The step size was reduced to 25% after the second reversal, and the whole testing period will end after three reversals. The last word duration will then be used to compute the reading speed (i.e. wpm). As there is no or minimal eye movement involved, the reading speed can reach approximately 1000 wpm in an individual with normal vision, which is 3-4 times greater than a typical reading speed using static text (Rubin and Turano, 1992). Hypothetically, this method would seem to benefit patients with central field loss because it may help to eliminate the effect of erratic eye movements that might impede the reading speed. However, Rubin and Turano (1994) observed only a 1.5 times improvement on the reading speed in participants with central field loss

compared to the presentation of a complete sentence on the screen. Although the saccadic eye movements were significantly reduced, their findings suggest that the eye movement difficulties only partly explain the slower reading speed in individuals with central field loss. Harland et al. (1998) and Fine and Peli (1995) did not find a significantly higher reading speed using RSVP compared to drifting text in patients with central field loss. Fine and Peli (1998) reported that the RSVP gain (i.e. approximately 1.3 times) in reading speed compared to drifting text, only occurs when the text is at least 8x of their acuity threshold. A possible reason for the limited benefits of RSVP for individuals with central field loss could be due to the difference in visual span characteristics at different retinal locations. Visual span is defined as the number of recognizable letters at a single glance. Typically, the visual span in normally-sighted individuals consists of 10 letters. Previous studies (Legge et al., 1997a, Legge et al., 1997b) have suggested that there is a reduction in visual span with increasing retinal eccentricities (e.g. 1.7 letters at 15° eccentricities (Chung et al., 1998)), which can contribute to the reduction in reading speed in patients with central field loss.

A very similar test, the elicited sequential presentation (ESP), was first described by Arditi (1999). Unlike RSVP where words are presented sequentially with the same presentation time of each word, regardless of its length, the ESP is fully controlled by the participant using a click button. Reading speed was on average 47% faster compared to RSVP in a heterogeneous group of 15 low vision participants. However, this improvement (i.e. range 4% to 112% faster) was reduced with increasing reading speed. The ESP can provide a relatively “faster” reading speed because readers can allocate appropriate time for the different length of words. In other words, a more “effective” time allocation for each word presented helps to improve the reading speed. These results suggest that words of different lengths require different time duration to recognise them. Therefore, when constructing a reading speed test, the variation of the word length should be considered.

1.2.2 Typeface

The most significant characteristics of different typefaces which have been compared are between serifs or san serifs, and fixed-width or proportional width font.

1.2.2.1 Serifs vs Sans serifs

Serif refers to the extra detail at the end of the strokes of the letters, and sans serif refers to letters without these extra details at the end of the stroke (Figure 1.2). Most of the studies did not find an advantage of the type of font on reading speed performance. Yager et al. (1998) reported minimal differences in reading speed between sans serif font (Swiss similar to Helvetica: 531 wpm) and serif font (Dutch similar to Times: 540 wpm) using RSVP in a group of normally-sighted individuals. However, when the luminance was reduced from 146 cd/m² to 0.146 cd/m², an improvement on the reading speed in the Swiss font (i.e. 479 wpm) compared to the Dutch font (i.e. 429 wpm) was observed. The authors proposed that the improvement of the reading speed could be due to the higher acuity reserve in participants using the Swiss font (0.30 log units) compared to the participants using the Dutch font (0.27 log units). Similarly, Arditi and Cho (2005) did not find a difference in reading speed between serif and sans serif fonts in four normally-sighted participants and two participants with macular degeneration. Tiresias PC (TPC), is a font developed by the Royal National Institute of the Blind (RNIB), through a study on the preference of font characteristics by the low vision community (Perera, 2004). The preferred characteristics were sans serif font, with increased spacing, bold typeface, and medium-size punctuation marks. However, these desirable characteristics did not lead to improved reading performance. Although Rubin et al. (2006) reported an improvement in reading speed using TPC compared to other fonts, only a marginal improvement was observed (i.e. TPC was 8 wpm faster compared to other fonts such as Times New Roman). In fact, such improvement disappeared when the vertical and horizontal spacings were adjusted. As such, the improved reading rate from the TPC font was likely due to the associated inter-letter spacing rather than the design of the font typeface. More recently, Xiong et al. (2018b) investigated two new fonts which were designed specifically for low vision individuals – Eido and Maxular Rx. The design of Eido font was based on the hypothesis that reducing the physical similarity between letters may improve letter recognition. For instance, similar shape appearance of the commonly confusing letters (e.g. e, o, c) were modified to improve differentiation (Bernard et al., 2016) – the letter “e” was changed to “E” and the stroke width of the letter “o” was thickened. For the Maxular Rx, the fonts were generally bolder and have rounded serifs. The rounded serifs were claimed to enhance recognition of the terminals and the direction of the line of reading, none of which has been tested and found to improve reading performance.

Bernard et al. (2013) have shown that increasing stroke width did not help to improve reading speed. In fact, excessive “bolding” was found to impede reading speed in a group of six normally-sighted participants (Bernard et al., 2013). The benefits of the extra spacing in Maxular Rx between letters and between lines, which can help to reduce the crowding effect, remains questionable. Figure 1.2 shows the comparison of the Eido and Maxular Rx to the other commonly used fonts (i.e. Helvetica, Times-Roman and Courier). Despite attempts at font modification to enhance reading performance in patients with low vision, the results comparing the two new fonts to the three commonly used fonts (i.e. Helvetica, Times-Roman and Courier) were disappointing. There was no difference in reading speed performance between the Eido and the Courier fonts in a group of normally-sighted participants and participants with macular dysfunction (Xiong et al., 2018b). Although better scores (i.e. smaller RA and CPS, or faster MRS) were observed using the Eido and Maxular Rx compared to Helvetica and Times-Roman, the differences were rarely clinically significant (RA and CPS improved by approximately 0.1 log MAR using Eido and Maxular compared to Helvetica and Times-Roman, and the largest improvement in MRS using the new fonts was 0.1 log wpm (equivalent to 1.25 wpm)).

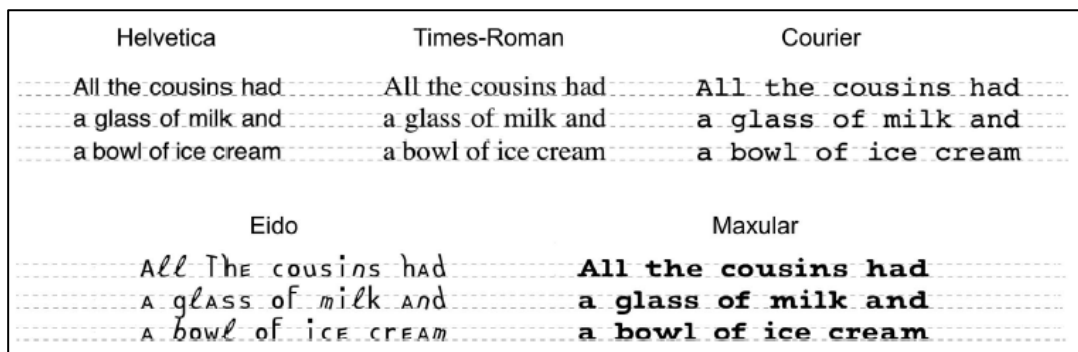


Figure 1.2 A sample sentence showing the different fonts used in Xiong et al. (2018b)

1.2.2.2 Fixed-width fonts vs proportional width fonts

There are two types of fonts width used for text printing. In fixed-width fonts such as the Courier font, the spacing is the same for every letter, despite differences in their width (i.e. “l” and “m”). On the other hand, proportional width fonts such as the Times New Roman font has variable space allocated for each letter depending on its horizontal extent (e.g. “l” takes less space than “m”). As the spacings in proportional width fonts are narrower compared to fixed-width fonts, the former

could induce a crowding effect, and this effect is more apparent in the peripheral retina (Bouma, 1970). In other words, patients with central field loss may read slower with proportional width text compared to fixed-width text. Indeed, Arditi et al. (1990) found that faster reading speed using fixed-width font at the threshold acuity limit. On the other hand, the authors also observed that the proportional width font produced a faster reading speed in medium to large characters compared to the fixed-width font, but this difference in reading speed was not seen when using RSVP. It is possible that the tighter spacing in proportional width font allows more characters to fit into the participant's visual span, and therefore, permits less saccadic eye movements. Benefits in using fixed-width font were also reported by Mansfield et al. (1996) on a group of normally-sighted participants and participants with macular degeneration. The improvement in reading speed was slightly more in the group of participants with macular degeneration (i.e. 10% faster in the fixed-width font) compared to the normally-sighted group (i.e. 5% faster in the fixed-width font). Although the above findings are suggestive of using increasing space to improve reading speed, Chung (2002) has shown that when increasing the space beyond the standard letter spacing, there is no improvement in reading speed for either the central or peripheral retina. When deciding the font in a newspaper or magazine, proportional width fonts are perhaps more practical as the space saving is approximately 40% using the Times New Roman compared to the Courier (Mansfield et al., 1996). In the context of this study, typical reading material was used to provide a more representative of real-world reading tasks. Therefore, Times New Roman, which is one of the common fonts used in newspapers (DeMarco and Massof, 1997), was adopted in our experiments.

1.2.3 Contextual vs random words

There are two main approaches in choosing the context of the reading performance tests: contextual and random words. These contextual words can be presented either in the form of contextual sentences with no thematic link between successive sentences or with a thematic link (i.e. a single paragraph). Most reading tests adopt the use of contextual reading text, as it allows a more realistic representation of daily leisure reading materials. As the words are thematically linked, reading speeds will be faster compared to random words text (Legge et al., 1989a). Unless the assessment is done with scrolling text, which forces the reader to read at MRS, readers (i.e. especially with high reading speed) can choose to read at their preferred reading rate. This

“unconstrained reading”, a term used by Whittaker and Lovie-Kitchin (1993), can result in a more variable individual reading speed (Bullimore and Bailey, 1995). On the other hand, Bailey and Lovie-Kitchin (1980) have suggested that the use of random words might be more sensitive in identifying any changes in reading speed (e.g. when a visually impaired individual was trying out different LVAs options for reading). The use of random words would minimise the influence of other reading components (e.g. cognitive) on the reading performance (Baldasare et al., 1986) by reducing the likelihood that the reader can guess the word correctly based on context rather than the visual ability to recognise the word. The impact of the use of the contextual reading material in low vision reading studies has been controversial. Bullimore and Bailey (1995) found that the benefit of contextual cues was more significant in the age-related macular degeneration (AMD) group (contextual text mean reading speed = 95.32 wpm vs random words mean reading speed = 33.1 wpm) compared to the normally-sighted group (contextual text mean reading speed = 276.92 wpm vs random words mean reading speed = 161.0 wpm).

On the other hand, Fine and Peli (1996) did not find a statistically significant benefit of contextual sentences between the normally-sighted participants (contextual text mean reading speed = 207.5 wpm vs random words mean reading speed = 121.0 wpm) and participants with central field loss (contextual text mean reading speed = 87.3 wpm vs random words mean reading speed = 46.0 wpm). The discrepancy in the relative difference between the two studies may be attributed to the use of thematic sentences in Bullimore and Bailey (1995) compared to Fine and Peli (1996), where independent contextual sentences were used. Given the more visually challenging reading conditions experienced by individuals with visual impairment, it is not surprising that the thematic content would provide extra help to guess the following word.

Baldasare and Watson (1987) suggested that individuals with visual impairment do not benefit from contextual cues as they have limited processing capacity to decode degraded visual information. Sass et al. (2006) found that the normal vision group (contextual text reading speed for young = 510 wpm, elderly = 365 wpm; continuous text reading speed increased by a factor of 1.87 in combined age group) had a higher dependency on continuous text compared to the low vision group (contextual text reading speed for young = 212 wpm, elderly = 106 wpm; contextual

text reading speed increased by a factor of 1.51 in combined age group). Previous studies (Chung et al., 1998, Latham and Whitaker, 1996) have shown that the increase in reading speed in contextual sentences over random words was the highest in the foveal region, and this effect was reduced as the retinal eccentricity increases. The results suggest that normally-sighted individuals may benefit more in contextual over random words reading material compared to individuals with central field loss who were using eccentric viewing.

We have considered the studies comparing contextual and random words reading assessment materials. The main drawback of creating a contextual reading material is that the level of difficulty between different versions (e.g. IReST) or sentences (e.g. Radner or MNREAD acuity chart) must be consistent. The assessed reading speed is commonly used to determine the success of low vision rehabilitation intervention (e.g. improved reading speed with a magnifier). As such, the variation in text difficulty should be minimal to avoid changes in the reading speed. Although considerable efforts have been expended to create “standardised” reading material, Brussee et al. (2015) have shown that there is a discrepancy in reading speed assessment using different versions of the same reading test. On the other hand, it would be straightforward to create multiple versions of equivalent difficulty in a test which used words of standard lengths presented in random order.

1.2.4 Singles sentences vs paragraphs

The difference text structures such as sentences and paragraphs have been adopted by manufacturers to design their reading speed test. In some of these reading tests which include the assessment of RA and CPS, the use of single sentences in logarithmic size progression is a practical approach (Radner et al., 2002, Mansfield et al., 1994). The use of single sentences in reading speed assessment is less demanding in individuals with visual impairment, as they might have difficulty going through long reading texts. For the assessment of spot or survival reading (i.e. reading price tags or mail), this assessment would be sufficient. However, the use of short sentences may not be adequate to provide information on page navigation and capture any potential duration related effects in leisure reading performance. Indeed, a recent study by Ramulu et al. (2013b) using an extended reading text (i.e. approximately 7300 to 7600 words) found a

duration related effect in a 30-minute silent reading task. Also, the use of a longer reading text may average out the random variability induced by a short reading test. Also, when reading speed assessment is conducted in short sentences, technical glitches (e.g. reaction time or false starts by the examiner when using a stopwatch) may affect the precise measurement of the reading speed (Hahn et al., 2006, Trauzettel-Klosinski et al., 2012). However, it remains questionable if this imprecise time recording would produce a clinically meaningful difference in the reading speeds. Besides, the false start in time recording when using the stopwatch would be more likely to occur in individuals with fast reading speed. However, most of these reading performance tests are more commonly used in individuals with visual impairment (i.e. slow reading speed).

1.2.5 Oral vs silent reading

Oral reading refers to the out-loud reading of a presented text, which can be either recorded by a voice-recorder or timed by a stopwatch for reading performance analysis. The benefits of using oral reading assessment include allowing the examiner to identify the start and end of the reading process easily, being able to pick up any mistakes made by the reader and to be almost certain that the reading task is being carried out (i.e. when comprehension is not assessed). However, oral reading has been criticised as not representative of normal daily reading. Unlike oral reading, silent reading is a more natural reading behaviour. In research studies requiring reading of lengthy reading text passages, the use of the silent reading mode might be more appropriate as it can be frustrating or tiring for the participant to perform oral reading. On the other hand, the main drawback of silent reading is the difficulty in assessing the reader's reading compliance. While infrared eye movement tracking equipment can be used, the set-ups are usually complicated and not feasible in a clinical setting, especially while using a magnifier or with a short viewing distance. To tackle this problem, researchers (Ramulu et al., 2013b) have used post-reading questions to ensure reading compliance. However, setting questions for the reading passage can be challenging. Firstly, these questions must be validated to ensure they can only be answered after reading the passage and not by general knowledge or logical thinking. Secondly, the use of multiple-choice questions can still allow a certain probability of getting the answers correct by guessing. Thirdly, it can be hard to determine an appropriate "pass" score. Even when participants were instructed to read the reading text like they normally do with their daily reading materials such as a magazine

or newspaper (i.e. rauding), they may not achieve a “perfect” score in the multiple-choice questions. To achieve this “perfect” score, a higher level of reading gear (i.e. memorising, which requires the participant to re-read several times) may be required.

In general, silent reading assessment produces faster reading speed performances compared to oral reading assessment (Ramulu et al., 2013b, Lovie-Kitchin et al., 2000, Bouma and De Voogd, 1974). This is because in oral reading, apart from the assessment of an individual’s ability to process the visual information, the speed of how fast the reader can articulate the words aloud can limit the reading speed compared to silent reading. In normally-sighted people, the silent reading speed can be up to 50% higher compared to the oral reading speed (Bouma and De Voogd, 1974). It was not clearly indicated if the participants were instructed to read as they would in their daily reading in the Bouma and De Voogd (1974) study. However, the reduction in fixation duration and increase in saccade length observed (oral reading: 7-9 characters vs silent reading: 9-14 characters) seems to suggest that the participants may have switched from rauding (i.e. during oral reading) to skimming (i.e. during silent reading) reading gear. This may have further contributed to the faster reading speed in silent reading compared to oral reading.

Previous studies on patients with visual impairment have shown a moderate ($r = 0.59$ to 0.68 , (Ramulu et al., 2013b)) to strong correlation ($r = 0.9$, (Lovie-Kitchin et al., 2000)) between oral and silent reading speed. The difference in the findings between the two studies can be attributed to the differences in methodology. Lovie-Kitchin et al. (2000) compared both oral and silent reading speed assessment using MNREAD acuity charts, and the participants were instructed to read with understanding (i.e. rauding). In Ramulu et al. (2013a) study, the oral reading tests were done using MNREAD acuity charts and IReST, while the silent reading was assessed using their designed sustained reading text (approximately 7500 words). Unlike Lovie-Kitchin et al. (2000) study, Ramulu et al. (2013b) instructed their participants to read the MNREAD acuity chart “as quickly and accurately as possible”. The instructions for IReST and silent reading was not explicitly mentioned in their paper, and it is unclear if these participants were aware beforehand that they are required to answer multiple-choice questions to test their comprehension without referring to the original text. If the participants were aware that they would be tested, then a different reading

gear (i.e. rauding or memorising) might be adopted by the participants during the silent reading test.

From the literature, as mentioned above, silent reading speed is faster compared to oral reading, with a lesser impact on individuals with slower reading speed. Since a moderately high correlation was observed between silent and oral reading speed, silent reading will be adopted in this study. Also, it is a more appropriate assessment for a reading test duration of approximately 15 minutes.

1.3 Visual factors influencing reading performance

Whittaker and Lovie-Kitchin (1993) gathered evidence from numerous studies which had been done on reading performance and highlighted four major requirements to define spot reading (i.e. reading the price tags, 40 wpm) and high fluent reading (i.e. reading a book, 160 wpm). These four components are acuity reserve, contrast reserve, the field of view and size of the central scotoma. They are not independent of each other (i.e. trying for increased acuity reserve for a patient with visual impairment may lead to a reduced field of view) and thus prescribing a LVA needs to consider a balance of each component.

1.3.1 Acuity reserve

When performing a reading task, the amount of acuity reserve can affect reading speed performance. Acuity reserve is defined as the ratio of actual print size to the subject's threshold print size. Table 1.2 shows the summarised visual requirements for reading. Spot reading, also known as survival reading, refers to short periods of reading, such as reading the price tags or a restaurant menu. Based on the clinical experience of Whittaker and Lovie-Kitchin (1993), a reading speed of approximately 40 wpm would be adequate for these daily living tasks. Leisure reading, such as reading a book or magazine, requires a higher reading speed. The fluent and high fluent reading speeds were derived from rauding reading speed for 2nd grade and 6th-grade level, respectively from Carver (1990).

Table 1.2 Visual requirements for reading (Whittaker and Lovie-Kitchin, 1993)

	Spot	Fluent	High Fluent	Optimum
Reading speed	40 wpm	80 wpm	160 wpm	
Acuity reserve	1:1	1.5:1	3:1	6-18:1
Contrast reserve	3:1	4:1	10:1	>30:1
Field of view (characters)	With scrolled text			
	1	2-5	4-6	4-6
Field of view (characters)	With stationary text			
	2	5	12	16-20
Scotoma Diameter	>30°	22°	4°	none

Although a maximum reading rate was achieved by a print size 6 to 18 times bigger than the threshold acuity in normally-sighted individuals (Legge et al., 1985a), Whittaker and Lovie-Kitchin (1993) found that the amount of optimal acuity reserve in the low vision individuals is highly variable to generate a cut-off value. Instead, the authors proposed an acuity reserve of 3:1 for high fluent reading as individuals showed a decrease in reading speed with acuity reserve less than 3:1. In a subsequent paper by the same authors (Lovie-Kitchin and Whittaker, 1999), the acuity reserve was revised to 2:1 and 1.3:1 for fluent reading (80 wpm) and spot reading (40 wpm) respectively. This was to provide an allowance for optical aberrations and poor hand control of magnifiers. The revision was done on the fluent reading (80 wpm) instead of high fluent reading (160 wpm) because most of the low patients cannot achieve the latter, and a reading speed of 80 wpm would allow a satisfactory leisure reading.

1.3.2 Contrast reserve

Contrast reserve is defined as the ratio of the print contrast to the subject's contrast threshold. Gathering findings from previous studies (Brown, 1981, Legge et al., 1987, Rubin and Legge, 1989), Whittaker and Lovie-Kitchin (1993) analysed the relationship of contrast sensitivity with reading speed. To achieve spot reading and high fluent reading, the contrast reserve ratios required were 3:1 and 10:1, respectively (see Table 1.2). To increase the acuity reserve, the print size can be enlarged. However, in contrast reserve, there is a limitation on how much the contrast of the reading material (e.g. black on white is approximately 90% (Rumney, 1995)) can be improved.

An early study by Legge et al. (1987) has shown that reading speed is reduced when the text contrast falls below 10% in normally-sighted individuals. In a subsequent study in the series of Psychophysics of Reading by Gordon E Legge and collaborators, Legge et al. (1997b) conducted a series of experiments to show that the reduction in reading speed in reduced text contrast was due to the shrinking of the visual span and prolonged viewing of the text. Visual span, a term introduced by O'Regan (1990), refers to the number of characters recognised at each glance without moving the eye. Legge et al. (1997b) have shown that the visual span of 10.6 characters at a nominal 100% text contrast in a normally-sighted individual can be reduced to 2.32 characters at 2% text contrast. When investigating the eye movements, Legge et al. (1997b) found an increase in the number of saccades and longer fixation duration when the text contrast was reduced in the normally-sighted participants. These observations support the hypothesis that the slower reading was attributed to the shrinking of the visual span and prolonged viewing of the text. These findings were also confirmed by Mohammed and Dickinson (2000), who investigated the relationship between acuity reserve and contrast reserve by evaluating the reading speed with a range of magnifiers (3x, 5x, 7.5x) at different contrast levels (43%, 20%, 9%, 5%). They found shorter forward saccades in participants with low contrast reserve when using the magnifier, which leads to the reduced reading rate. Also, Mohammed and Dickinson (2000) have also demonstrated that given an optimally corrected VA, low contrast reserve could not be compensated by increasing magnification to increase reading speed.

1.3.3 Perceptual span

Perceptual span refers to the number of characters in the field of view, which prepares the eye for the next saccade (McConkie and Rayner, 1975). This is measured in terms of characters, as the number of characters for each saccade in the reading text remains relatively the same over an extended range of viewing distances (Rayner, 1998). The perceptual span allows up to 14-15 characters to the right (McConkie and Rayner, 1975) and four letters to the left (Rayner et al., 1980), but not all characters can be discriminated at one fixation. The term perceptual span should not be confused with the visual span, which refers to the numbers of characters recognised at each glance without moving the eye. As noted earlier, there are approximately ten recognisable

characters in each fixation in optimal reading condition (i.e. 100% text contrast) (Legge et al., 1997b).

Previous experimental results (Legge et al., 1985a, Legge et al., 1985b) have shown that fast reading speed can be achieved with a very small field of view of 4 to 6 characters (see Table 2). This finding, however, is achieved by presenting a text that is automatically scrolled across the screen. It appears that 15 (Lovie-Kitchin and Woo, 1988) or more characters (Lowe and Drasdo, 1990) are necessary to produce an optimal reading performance when the individual is required to perform page navigation; that is, the tracking of a line of words with a magnifier from the end of the line to the beginning of the next line. Beckmann and Legge (1996) have shown that the higher number of characters reported in Lovie-Kitchin and Woo (1988) and Lowe and Drasdo (1990) was due to the use of the CCTV where page navigation is involved. Lower window width findings in Lovie-Kitchin and Woo (1988) compared to Lowe and Drasdo (1990) can be explained by the difference in reading tasks in their experiments. Participants in Lovie-Kitchin and Woo (1988) study were required to read one line of text (21-30 words) compared to Lowe and Drasdo (1990), where the text had 21 lines (200 words) which required more page navigation.

1.3.4 Presence and size of central scotoma

In healthy normally-sighted individuals, the macula region, where the cone photoreceptors are the densest, is responsible for clear central vision. When an individual suffers from macular disorders, such as macular degeneration, they present with central scotoma and are forced to use an alternative area for reading – the PRL. Being forced to use the eccentric retina areas for visual tasks, the use of PRL will always accompanied with decreased visual acuity and contrast sensitivity. Nonetheless, Whittaker and Lovie-Kitchin (1993) have included central scotoma size as a separate independent visual impediment factor for fast reading because even with a given sufficient acuity reserve, the participants' reading speed were reduced when performing reading tasks using eccentric retina (Legge et al., 1992).

1.4 Types of reading assessment tests

There are many reading assessment tests available, and it is not possible to discuss all of them in this review. In this report, only English reading tests that are reported in low vision research papers will be discussed. The selection of appropriate reading tests can be based on the purpose of the assessment (i.e. comprehension, reading rate, etc.).

1.4.1 Bailey-Lovie word-reading chart

The Bailey-Lovie word-reading chart (Bailey and Lovie, 1980) was developed to investigate the characteristics of the visual performance in individuals with macular diseases. The print size ranges from N80 to N2, which allows an extensive RA measurement from 1.6 to 0.0 log MAR at 25 cm in 0.1 log unit progression (Figure 1.3). The reading chart has fewer words in larger print sizes to allow the reading test to fit in one A4 reading chart size (i.e. 26 x 20.5 cm). The top three lines consist of two words (1.6 to 1.4 log MAR), followed by three words (1.3 to 1.1 log MAR) in the next three lines and six words (<1.0 log MAR) in the subsequent lines. Out of the six words, there are two 4-letter, two 7-letter and two 10-letter words. Some individuals with central field loss may find shorter words more manageable than longer words. This mixture of word length provides information on the reader's ability to cope with the different length of words. Non-contextual words were used to minimize the possibility of the cognitive component inflating the reading performance. In addition to RA, Bailey and Lovie (1980) also suggested that the chart could be used to assess reading speed in each print size. This provides information such as the optimal print size required to produce desirable reading speed, and if increasing the magnification can continue to improve the reader's reading speed.



Figure 1.3 A sample of Bailey-Lovie word-reading chart (Bailey and Lovie, 1980)

1.4.2 International Reading Speed Text (IReST)

The International Reading Speed Text (IReST) was created to provide a standardised longer reading text for reading speed assessment. Unlike most of the low vision reading tests which have relatively short sentences, each IReST paragraph (Figure 1.4) consists of approximately 150 words in each text with uniform vocabulary and sentence structures across the languages. The text contents were taken from sixth grade (10-12 years) encyclopaedia. The paragraphs use Times New Roman typeface and come in standard font size (N8) similar to the newspaper print size (DeMarco and Massof, 1997). The IReST consists of 10 different text paragraphs in each of 17 different languages (Arabic, British English, Chinese, Finnish, French, German Hebrew, Italian, Japanese, Dutch, Polish, Portuguese, Swedish, Slovenian, Spanish, Russian and Turkish) (Trauzettel-Klosinski et al., 2012). Although reading speed is still used as the reading performance indicator in IReST, the use of longer reading text has the benefit of averaging out the random blips in reading speed measurement compared to single reading sentences. However, it is questionable if the time difference in recording would be of any clinical significance. Also, the single paragraph

is more representative of leisure reading task, which can potentially highlight the difficulty (e.g. page navigation) experienced by a low vision individual using a LVA.

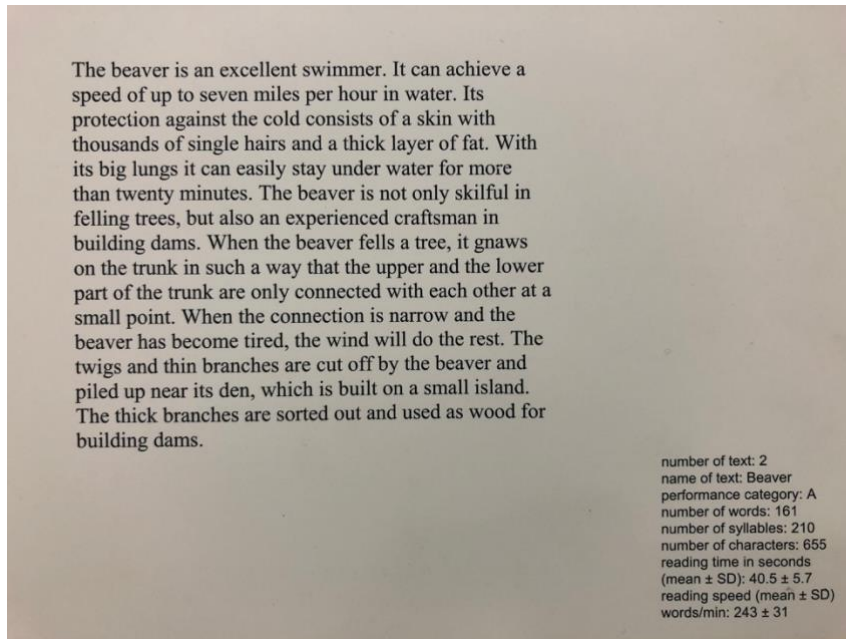


Figure 1.4 A sample of the International Reading Speed Text (IReST) (Trauzettel-Klosinski et al., 2012)

1.4.3 Minnesota Reading (MNREAD) acuity chart

The first version of this reading test, The Minnesota Low-Vision Reading Test, was reported in Legge et al. (1989a) with the use of a computer monitor. To increase the usage of the test in a clinical setting, the test was then developed onto printed cards (Ahn et al., 1995). Each card consists of one sentence, and the letters subtended 6° at a standard viewing distance of 19 cm (i.e. can be used in individuals with visual acuity up to 20/1440 or 1.86 log MAR). However, this test was not commonly used – possibly because of the development of Minnesota Reading (MNREAD) acuity chart (Mansfield et al., 1993), which allows the assessment of other parameters. The MNREAD acuity chart (Mansfield et al., 1994) measures RA, MRS (i.e. the fastest reading speed independent of character size) and the CPS (i.e. smallest print size that gives the maximum speed) and ACC (i.e. accessibility index of print sizes).

The chart has 19 sentences in decreasing sizes (range 1.3 to -0.5 log MAR), in steps of 0.1 log units (Figure 1.5). The sentences are split into 3 lines to include the possibility of vertical crowding.

Each sentence consists of 60 characters (including the spacing in between the words). Although these are contextual sentences, the sentences are not linked thematically. This is to avoid overestimation of reading performance by thematic content. Times New Roman typeface was used to give a better representation of current reading materials such as newspapers and magazines. The test is conducted through oral reading, and the reader is instructed to read the sentence as quickly and accurately as possible.

More recently, Calabrèse et al. (2018) developed the use of the MNREAD iPad application to assess the reading performance parameters in a group of 165 normally-sighted and 43 low vision participants. The benefits of using a tablet computer includes 1) a more simplified and testing procedure; 2) scoring algorithms are incorporated into the application which provides a “standardised” method in computing all the reading performance parameters (i.e. RA, MRS, CPS and ACC); 3) allows portability by allowing different variety of the testing charts (e.g. different language, contrast polarity and different versions) inside a single computer tablet. The application replicates a standard MNREAD acuity chart text structure with inbuilt timing and reading measurement features. However, the standard 19 sentences have been reduced to 14 sentences, which is equivalent to angular print sizes from 1.2 to -0.1 log MAR at 40cm. The omission of the largest print size was due to unable to fit into the screen size of the iPad, and the other four smallest sizes were also omitted due to the limited resolution of the screen. For the normally-sighted group, iPad MRS measurements were statistically significantly slower compared to the printed MNREAD acuity charts. This difference is significantly more in the faster readers (i.e. 3% slower at 100 wpm to 12% slower at 250 wpm). The authors suggested that this difference in timing between the printed copy and the iPads was due to a “procedural flaw”. When using the printed copy of the MNREAD acuity chart, the investigator is required to uncover the text and start the stopwatch simultaneously. However, the authors observed that the participants often started reading before the start of the stopwatch – resulting in shorter reading time and higher reading speed. On the iPad, the text presentation and timer were paired together and were initiated simultaneously by a finger tap on the screen. For the low vision group, only the RA measurement was statistically significantly ($p = 0.013$) smaller by 0.03 log MAR using the iPad, but this small difference would not be clinically relevant. The MRS and CPS was computed based on the original algorithm

described by Legge (2007) – MRS is estimated as the mean reading speed for the range of print sizes where reading speed is at least 1.96 standard deviations faster compared to the other print sizes, and CPS is defined as the smallest print size within this range. It was unclear the reason for this calculation approach compared to the definition recommended by Patel et al. (2011). Although automated scoring allows a more “standardised” method in computing the reading performance parameters, it is unclear how much data input is required to generate a “good” estimation of these parameters. Errors in computing the reading performance parameters were reported in 9% of the low vision participants when using the iPad scoring algorithms, albeit it is still lower compared to using the MNREAD acuity chart (i.e. 11%). It is, therefore, important to combine the interpretation of the scores with the visual inspection of these graph plots. As with any electronic devices, the disadvantages include the requirement to charge the device and update the software. Although the electronic version shows promising results to provide alternatives for reading performance assessment, further research needs to be conducted to investigate the repeatability and reproducibility of the app to assess reading performance.

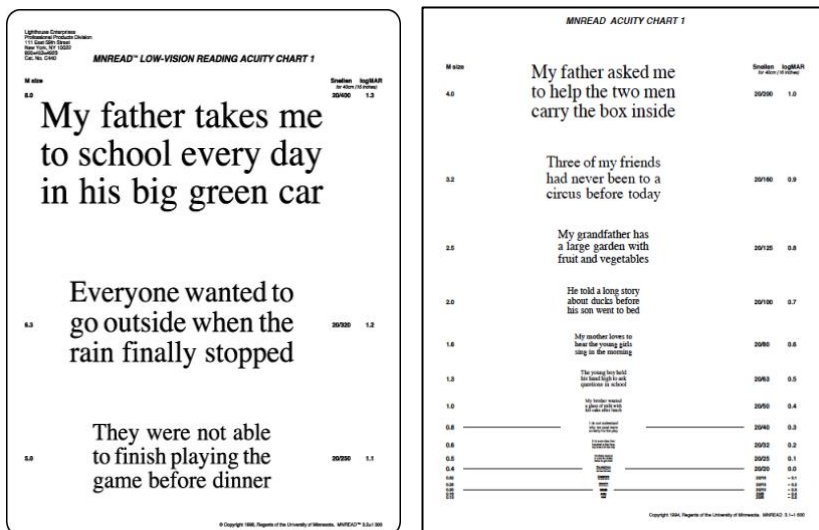


Figure 1.5 A sample of Minnesota Reading (MNREAD) acuity chart (Mansfield et al., 1994)

1.4.4 Morgan Low Vision Reading Comprehension Assessment (LVRCA)

Previously known as the LVRCA (Watson et al., 1996), the Morgan LVRCA (Watson et al., 2017), is probably the only comprehension reading test designed for individuals with visual impairment, specifically with central field loss. The Morgan LVRCA consists of 18 cloze sentences. Subjects are required to read and understand so that they could fill in the missing word. An example from Watson et al. (1996): “What is born with fins is born to _____” and the suggested answer was “swim”. Based on the given example, it may appear that the test does not solely assess on the comprehension, instead, on an individual’s vocabulary level too. There is a range of different sizes (1M, 1.5M, 2M, 3M), and the authors suggested using print size, which is two sizes larger than the subject’s predetermined acuity threshold. The objective of this reading test is to aid in the planning of interventions that can help individuals with macular disease to improve their reading comprehension ability. Earlier on in Section 1.1.2.5, we have discussed on the limitation of using comprehension assessment (i.e. the approach of the assessment, how closely is comprehension associated with visual performance, etc.) as an outcome measure in low vision rehabilitation. The assessment of comprehension is uncommon in low vision clinical practice, possibly because there are better and more quantifiable measurements (e.g. reading speed) available. Hence, the use of a comprehension test, such as the Morgan LVRCA, is relatively uncommon in low vision clinical and research settings.

1.4.5 Pepper Visual Skills for Reading Test (VSRT)

The VSRT (Baldasare et al., 1986), as the name has suggested, was designed to focus on the assessment of visual skills (i.e. the reading speed and accuracy) in individuals with central field loss. This visual skill assessment could then be used to aid in the planning of a low vision training program, and for assessing progress during rehabilitation (Stelmack et al., 1987). The test consists of 13 lines of letters and non-contextual words (Figure 1.6). The use of the various length of words from single letters to compound words (e.g. otherwise – can be read as “other” or “wise” in individuals with central field loss) were used to create different levels of difficulty for the individual to decipher the words. For example, single letters can be easier to see compared to longer words. However, longer words can also be made easier by guessing - using the first and last few letters of the word. As such, the test adopts the use of non-contextual reading material to minimise the

use of other reading abilities (e.g. cognitive component) to inflate the reading performance by guessing through the use of thematic material. The print size comes in 1M (corresponding to the size of newspaper print), 1.5M, 2M, 3M, and 4M (newspaper paragraph headings) print sizes. Near visual acuity has to be predetermined before the use of the VSRT. The manufacturers suggest using a font size two times the near acuity threshold for best performance.

There are three key assessment results of the VSRT: reading accuracy, reading speed, and the errors made by the reader. Reading accuracy is scored by the mean percentage of correct words over the total number of words (excluding the “dummy” lines: Line 3, 8 and 13). As noted earlier, VSRT attempts to assess the ability of the reader to navigate through the reading text. Hence, different line spacing was created, and lines 3, 8 and 13 were used to set the correct spacing of the subsequent lines. Reading rates can be calculated by using the total number of correctly identified words (including the “dummy” lines) divided by the total time taken in minutes. However, the reading speed performance can be misleading as the inclusion of the single letters in the calculation can inflate the reading speed. Apart from omitting the errors made by the reader, Baldasare et al. (1986) created a list of error notations (e.g. misidentification, spells words, etc.), which can help to indicate the types of difficulty experienced by the reader.

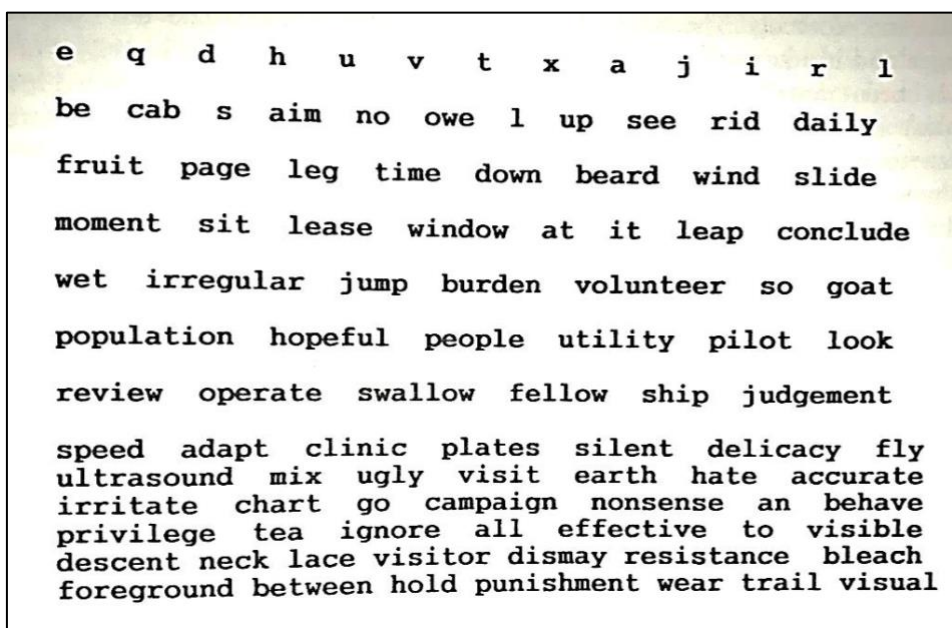


Figure 1.6 A sample of a Pepper Visual Skills for Reading Test (Baldasare et al., 1986)

1.4.6 Rate of Reading (RoR)

The Rate of Reading (RoR) (Wilkins et al., 1996) is not a low vision assessment test. However, it is included in this literature review because this test would be used in our subsequent experiments due to its unique design feature – random words. Meares-Irlen syndrome is a condition where the person can experience visual discomfort and perceptual distortion, also commonly termed as pattern glare or visual stress (Hollis and Allen, 2006). This visual stress can cause symptoms such as blurring and movement of lines. The RoR was designed to investigate the impact of reading lines (i.e. mimic spatial gratings), which can cause visual stress to readers who suffer from Meares-Irlen syndrome. The test comprises of paragraphs of 15 unrelated words with minimal spacing (see Figure 1.7). The words have been chosen to be short and simple to allow ease of reading by any reader aged seven years and above (Kriss and Evans, 2005): the influence of any cognitive difficulties with the reading process is therefore reduced. The text uses Times New Roman 9-point font size. The x-height and width are 1.6mm and 1.53mm, respectively. This size was chosen to be visually challenging to allow visual discomfort to become apparent in a short time. Reading speed can be calculated by the number of correct words read per minute. There are no normative values as the test strives to compare the performance of an individual between different reading conditions (i.e. with and without a coloured tint or overlay).

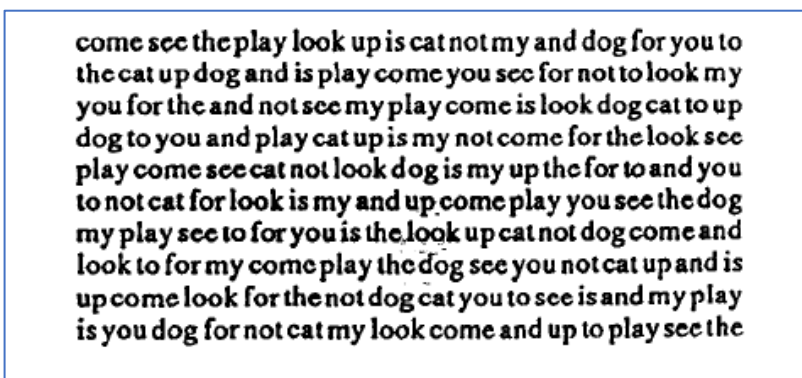


Figure 1.7 A sample of a single paragraph from the Rate of Reading (RoR) test (Wilkins et al., 1996)

1.4.7 Precision of reading assessment

So far, the purpose, design features and properties of some English reading tests used in low vision research studies have been discussed. There are many reading performance tests available commercially, but most of these reading performance tests do not provide reports on the precision

of their assessments. Rubin (2013) commented that using a poor reliability reading test for outcome measures would require a larger sample size to identify a treatment effect. McAlinden et al. (2011) defined precision as the repeatability and reproducibility of the reading test. Repeatability refers to the variability of the repeated measurements on a participant when all the conditions (e.g. investigators or testing methods) are identical. The repeated assessments should be conducted within a short time frame such that the participant's condition should also be constant. In this way, the observed measurement variability would then truly reflect the measurement error itself. On the other hand, reproducibility refers to the variability of the measurements on the same participant under different conditions. This can be interpreted as measurements taken at a considerable "time frame", where possible changes such as different investigator, different lab, etc. may occur. However, there is no consensus on the appropriate "time frame", and it would be dependent on the field of study. Unlike objective clinical measurements where repeated measures can be achieved easily (e.g. taking five repeated measurements of the intraocular pressure using a tonometer within one occasion), repeated reading assessments within a short interval can involve a learning effect, or a fatigue effect on the assessment, which can potentially introduce bias in the repeated measurements. In a MNREAD acuity chart repeatability study on a group of low vision participants, Subramanian and Pardhan (2009) conducted two assessments within the same day. Two versions of the MNREAD acuity chart were used, and the sequence was randomised among the readers. The rationale was that if a learning effect was present, all the second testing charts would show a faster reading speed. The paired t-test did not show a significantly faster reading performance in the second testing charts, meaning there was no evidence to suggest a learning effect. While this approach by Subramanian and Pardhan (2009) appeared logical, there may be other confounding bias such as fatigue (although a break was given between the testing sessions) which can potentially cause the opposite effect (e.g. slower reading speed). These two effects may not be picked up using the paired t-test as there would be an averaging out in the final result. Therefore, the use of an interval period of several days, as suggested by Brussee et al. (2014) would seem appropriate in reading performance repeatability tests.

Ideally, reporting studies on the repeatability or reproducibility of the reading tests should include detailed study design and potentially factors that can affect measurements such as the illumination

levels, number of investigators, procedures, testing distance, etc. (Brussee et al., 2014). A recent comprehensive review by Brussee et al. (2014) found that publications on the measurement properties of the commercially available reading tests were lacking. Adopting and modifying the items from the COnsensus-based Standard for the selection of health status Measurement INstruments (COSMIN) (Mokkink et al., 2010) checklist for reading tests, Brussee et al. (2014) gathered the available reports from the reading tests on the five relevant criteria; content validity (i.e. comparable text difficulty in different versions), internal consistency (i.e. the evaluation of the experiment methodology), reliability (i.e. inter-chart, inter-rater repeatability and reproducibility), cross-cultural validity (i.e. comparable sentences in different languages) and generalisability (i.e. study population). The authors only found three reading tests (i.e. the MNREAD acuity chart, Radner reading chart and IReST) with reported studies on their measurement properties. However, none of these reading tests met all the testing requirements according to the measurement properties described by Brussee et al. (2014). There was no repeatability and reproducibility data available for the IReST, and no repeatability studies had been reported on the Radner reading chart. Although there was one study on the reproducibility of the MNREAD acuity chart (Patel et al., 2011), the information on the number of investigators and procedures for the measurement was not explicitly explained. Likewise, detailed information and appropriate reporting literature on the precision assessment of the random words reading tests were not available (i.e. Bailey-Lovie word reading chart, VSRT and Rate of Reading). There was only repeatability report on reading acuity measurement (Lovie-Kitchin and Brown, 2000), but no reliability tests were available on reading speed measurement for Bailey-Lovie word reading chart (Bailey and Lovie, 1980). The VSRT (Watson et al., 1990) has been tested in a group of 48 low vision participants with central field loss. The reading speed was assessed twice with an interval of 20-minute rest. However, only the correlation coefficient was reported as a mean for "reliability" testing, which was questionable as a high correlation does not equate to a good agreement in measurements. The Rate of Reading reliability testing also reported using the Pearson correlation coefficient ($r = 0.8$) between two visits, which were eight weeks apart. As mentioned earlier, the Rate of Reading was not a test designed for low vision. As such, the test was only conducted in a group of children ($n = 77$, aged 8 to 11 years). Further research is needed to ensure that these reading assessment tests have appropriate reliability. Therefore, researchers should consider the properties of the reading

test and the reading parameters when using reading performance as an outcome measure, as this could potentially affect the pre- and post-intervention results.

Table 1.3 summarises the features of the reading assessment tests reviewed. Most of these clinical reading performance tests were relatively short (in terms of words). Although using short texts may be sufficient for the assessment of reading performance for spot or survival reading, it may not provide information (i.e. reading endurance) on reading over a longer duration. In this study, the use of longer reading texts (i.e. one or more paragraphs) will be explored to investigate any duration related effects in a leisure reading task.

Table 1.3 Comparisons of the features of the various reading assessment tests

Reading Tests	Range	Purpose	Font	Context	Features	Text length
Bailey-Lovie word-reading Chart (Bailey and Lovie, 1980)	0.0 to 1.6 log MAR at 25cm	Near Acuity Allows calculation of magnification	TNR	Non-contextual	Top 3 line of 2 words, 3 words in the next 3 lines, 6 words in the subsequent lines. Of the 6 words, there are 2 4-letter, 2 7-letter and 2 10-letter words	2 to 6 words
International Reading Speed Texts (Hahn et al., 2006)	0.4 log MAR or N8 at 40cm	Reading speed	TNR	Contextual within paragraphs	10 paragraphs of each language	~150 words
Morgan Low Vision Reading Comprehension Assessment (Watson et al., 2017)	1M, 1.5M, 2M, 3M	Assess comprehension	Palatino	Cloze sentence	18 sentences	NS
Minnesota Reading Acuity chart (Mansfield et al., 1994)	-0.5 to 1.3 log MAR at 40cm	Reading Acuity Maximum reading speed Critical print size Reading Accessibility Index	TNR	Contextual within sentences	19 sentences 3 lines, 60 characters	10 Standard Word Lengths
Pepper Visual Skills for Reading Test (Baldasare et al., 1986)	1M, 1.6M, 2.0M, 3.2M, 4.0M	Print recognition Eye movement control Navigation with scotoma	NS	Non-contextual	13 lines Line 3, 8, 13 are "dummy lines"	13 lines of approximately 6-8 words (inclusive of single letter)
Rate of Reading (Wilkins et al., 1996)	9 point x-height 1.6mm width ~ 1.53mm	Evaluating visual stress in Meares-Irlen's syndrome	TNR	Non-contextual	10 lines of the same 15 unrelated words with minimal spacing	150 words

TNR: Times New Roman, NS: Not specified

1.5 The influence of age on reading speed performance

Acquired visual loss commonly affects the older population. It is, therefore, important to understand the effects of age on reading performances. As noted earlier, reading is a highly complex task that involves various processing components. Physiological changes such as motor processing, cognitive behaviour (Balota et al., 2000), visual structures (Weale, 1963) and visual functions (Owsley, 2011) can happen even in healthy elderly individuals. These changes can influence the reading performances in the elderly community compared to young adults. A considerable amount of literature has been published on the effect of age on reading speed performances in both normally-sighted and visually impaired individuals. In this section, the literature review will be split into normal vision and low vision.

1.5.1 Age and normal vision

Several studies have explored the relationships between age and reading speed in normally-sighted individuals. Some studies (Sass et al., 2006, Hahn et al., 2006, Bowers, 2000) compared the reading speed between young adults and elderly, and reported increasing age leads to slower reading speed. Table 1.4 tabulated the demographics of the participants, the reading speed performances in the young and elderly group, and the reading speed test used in the respective studies. However, some of these sources did not provide the details explicitly (e.g. reading speeds in Bowers (2000) were reported in a histogram). It is interesting to note that the reading speeds in both the young and elderly in the Sass et al. (2006) study were unusually high. The authors argued that this unusually high reading speed was due to the assessment of MRS, which can influence of up to 50% on the reading speed, and the use of short words which could additionally inflate the reading speed. In Akutsu et al. (1991) study, the reading speeds between the young and elderly group (no eye disease) did not differ significantly for character sizes range 0.3° to 1.0°. This may be because a more stringent approach was used to divide the elderly subjects into two groups: normal and mild disease group (i.e. 4 cases of cataract and 1 case of retinal detachment), albeit all participants had a minimum visual acuity of 6/7.5 or better. When combining the two subgroups, significantly slower reading speed was observed in the elderly group compared to the young adults' group. On the other hand, a more recent study by Dickinson and Shim (2007) did not observe a difference in reading speed between the young and elderly participants.

In a large-scale study by Lott et al. (2001), the team found slower reading speed was associated with increased age in a group of 544 healthy elderly participants (i.e. minimum visual acuity of 6/9 and better). Although there was no comparison of reading speed to a younger adult group, a linear decrease in reading speed was observed in the different elderly age groups. The mean age of the participants was 72.8 years, and the age ranges were split into six groups: 58-64 years (mean reading speed: 100.4 wpm), 65-69 years, 70-74 years, 75-79 years, 80-84 years and above 85 years (mean reading speed: 68.0 wpm). Apart from the common visual assessments such as visual acuity and contrast sensitivity, Lott et al. (2001) also obtained other clinical measurements such as spatial function under different luminance, attentional field and general motor skills (e.g. walking). Based on their findings, Lott et al. (2001) found that subtle changes in the visual, motor and cognitive factors in an ageing individual can affect the reading speed. However, when the clinical assessment factors such as low contrast acuity, motor ability and attention field measures were taken into account using multiple regression analysis, age was not a significant predictor of reading speed. Indeed, Akutsu et al. (1991) only found an effect of ageing on reading speed when the elderly group included the participants with mild disease (i.e. 4 cases of cataract and 1 case of retinal detachment) but with a relatively good VA of 6/7.5. This finding suggests that the reduction in reading speed observed in elderly participants could be due to the physiological changes of ageing eyes (e.g. cataract). In a more recent retrospective study by Calabrese et al. (2016), the authors analysed the reading speed using MNREAD acuity chart in a group of 645 normally-sighted individuals, aged 8 to 81 years. From the available data, the authors observed an increase in reading speed from 8 years (MRS: 137 wpm) to 16 years (MRS: 202 wpm), at an approximate increment of 8.13 wpm/ year. Thereafter, the reading speed remains relatively constant (estimated gradient slope of the change in reading speed: -0.07 wpm/ year) until the age of 40 (MRS: 200 wpm). The reading speed at 40 years starts to decline gradually (estimated gradient slope of the change in reading speed: -0.6 wpm/ year), reaching 175 wpm by 81 years. Although the sample size in Calabrese et al. (2016) is the largest and had the most extended age range (i.e. aged 8 to 81 years) to date, the number of participants in each age group was not as equally distributed. The median age in this studied sample group was 21 years, with a high interquartile range of 19 to 29 years. The dissimilar number of participants in each age group (i.e. 8 to 12 years ($n = 60$), 13 to 17 years ($n = 40$), 18 to 39 years ($n = 433$), 40 to 59 years ($n = 68$), and 60 to 81 years ($n = 34$))

was due to the combining of data from different studies with different objectives. With the given data, the elderly aged group (i.e. 60 to 81 years) only contributed five per cent of the total sample size, and may not provide a fair representation of the elderly group.

Table 1.4 Comparison of studies investigating the effect of age on reading speed performances in normally-sighted individuals

Authors	Description of the reading test used	Young			Elderly			P value‡
		Mean age (years)	N	Mean reading speed (wpm)	Mean age (years)	N	Mean reading speed (wpm)	
Akutsu et al. (1991)	Drifting text (80 characters line)	21.6*	16	343.21*	68.7*	14	291.90*	<0.05
Bowers (2000)	Neale Analysis of Reading Ability (~120 words)	21.9 (SD ± 3.8)	10	~225*	66.7 (SD ± 6.3)	10	~175*	<0.05
Lott et al. (2001)	PVSRT				72.8*	544	<65 years: 100.4* ≥85 years: 60.4*	<0.001
Sass et al. (2006)	MNREAD sentences (Legge et al., 1989a)	25.0 (SD ± 5.3)	10	510 (SD ± 180)	75.8 (SD ± 5.5)	10	365 (SD ± 158)	<0.01
Hahn et al. (2006)	IReST x 10 passage (~150 words per passage)	18-35*	100	205.67 (SD ± 24.50)	60-85*	100	158.50 (SD ± 16.17)	0.05
Dickinson and Shim (2007)	Reading passage (300 words)	24.75 (SD ± 3.08)	12	161.99 (SD ± 40.78)	70.73 (SD ± 7.30)	15	153.06 (SD ± 33.47)	0.53
Calabrese et al. (2016)	MNREAD acuity chart	8 – 12 13 – 17 18 – 39	60 50 433	8 yrs: 137 (8.13 wpm/year) 16 yrs: 202 (-0.07 wpm/year)	40 – 59 60 – 81	68 34	40 yrs: 200 wpm 81 yrs: 175 wpm (-0.6 wpm/year)	

*Standard deviation for the specific parameter not reported in the respective paper

‡This P value refers to the level of statistically significant difference in reading speed between the young and elderly group

1.5.2 Age and low vision

In one of the earlier works of Gordon Legge and co-workers on reading speed performance studies, Legge et al. (1992) investigated the reading speed in a group of low vision individuals with central field loss. The diagnoses consist of 11 juvenile macular degeneration (JMD), 41 AMD and 45 Others mixed conditions with central field loss. The mean ages were 78.3 years and 37.7 years in the AMD and the Others group respectively; mean age for the JMD group was unavailable. At 6° character size, the reading speed in the JMD and Others group was almost two-fold faster than the AMD group. Based on the difference in reading speed between the elderly and young visually impaired participants, Legge et al. (1992) proposed that reading in visually demanding conditions (i.e. individuals with visual impairment) requires “extra cognitive capacity”. However, the elderly may not possess an adequate amount of this “extra cognitive capacity” to perform reading efficiently.

On the other hand, Lovie-Kitchin et al. (2000) did not observe any difference in reading speed between the young and elderly group of participants with macular degeneration. Their study consists of 13 AMD participants (mean age = 76.0 years) and 9 JMD participants (mean age = 36.0). Lovie-Kitchin et al. (2000) proposed that the difference in the results could be due to the different reading gear employed by the participants in the two studies. In contrast to Legge et al. (1992) where they instructed their participants to read as fast as possible (i.e. skimming), Lovie-Kitchin et al. (2000) adopted the emphasis on rauding (Carver, 1992) (i.e. to read with comprehension). Previously, Salthouse (1996) reported that the speed of information processing slows down with age, and the amount of reduction is dependent on the nature of the task. When the task gets more complex (e.g. the level of word knowledge during reading increases), the extent of the processing speed also reduced. Lovie-Kitchin et al. (2000) proposed that rauding has a higher complexity compared to skimming, and therefore, no age-related difference in reading speed was found compared to Legge et al. (1992).

In the previous section, we highlighted some of the studies (Calabrèse et al., 2016, Sass et al., 2006, Hahn et al., 2006, Akutsu et al., 1991) conducted comparing young and elderly reading

speed performances. Most of these studies “force” an MRS either through the instructions (i.e. to read as fast and accurately as possible) or the use of drifting text on a monitor. These assessment methods may cause participants to skim through rather than trying to understand the text (Lovie-Kitchin et al., 2000). Unlike most of the previous studies (Calabrese et al., 2016, Sass et al., 2006, Hahn et al., 2006, Akutsu et al., 1991), Dickinson and Shim (2007) did not attempt to force an MRS performance. Instead, the participants were instructed to read at their normal reading speed with the emphasis on the rauding rather than skimming reading gear (Carver, 1992, Lovie-Kitchin et al., 2000). The difference in reading intention might have contributed to the difference in the findings between the studies. Also, when reading assessment was conducted through out-loud reading, the speed to articulate the words can play a role in the reading speed measurement. There is evidence (Rodríguez-Aranda, 2003, Smith et al., 1987) to suggest that speech articulation speed can be reduced in elderly individuals, which can potentially result in a slower reading speed measured in an out-loud reading test.

So far, the gathered studies (Calabrese et al., 2016, Hahn et al., 2006, Sass et al., 2006, Lott et al., 2001, Akutsu et al., 1991) have strongly suggested that reading speed slows down with increasing age. However, when the speed of speech articulation is not limiting factor (i.e. the use of rauding in the reading test (Dickinson and Shim, 2007, Lovie-Kitchin et al., 2000)), this difference appears to be reduced.

1.6 Simulating visual impairment

In studies investigating reading performances in relation to low vision, researchers can either consider the use of participants with genuine visual impairment, or healthy individuals with some form of simulation. Although the use of visually impaired participants in studies is a better representation of the community, this approach is not without problems. The nature of eye diseases, such as the severity of the eye condition, and other possible confounding factors can affect reading performance to an unpredictable extent (Fine and Rubin, 1999). For example, in visually impaired individuals with macular disease, the different size and location of the scotoma can have a different impact on the visual acuity. Hence, it is not uncommon for researchers to use simulation on normally-sighted individuals, as the associating factors such as visual acuity, contrast sensitivity, age, etc. can be controlled more easily using simulation.

In this study, the objective was to investigate the reading speed performances in individuals under visually demanding conditions (i.e. to simulate similar visual difficulty to that experienced by individuals with visual impairment). It would be more meaningful to simulate central scotoma defects as AMD is one of the leading cause of visual impairment globally (Flaxman et al., 2017), but the process of central scotoma simulation is more complex (i.e. also have to consider the impact of eye movements under the simulation) (Fine and Rubin, 1999). Another approach to create a visual impairment simulation would be the use of Bangerter occlusion foils. Bangerter occlusion foils, which were initially designed to be used for amblyopia treatment, allow simulation of reduced contrast sensitivity and visual acuity in a normally-sighted individual. Unlike simulation for central scotoma, these foils are straightforward to use (i.e. placing on top of the ophthalmic lens) and it allows a more predictable visual acuity level. Therefore, Bangerter occlusion foils will be used to create visually demanding conditions in our subsequent experiments.

1.7 Objectives of this study

The main aim of this thesis was to investigate duration related effects under visually challenging conditions.

The specific objectives of this research were to:

- I. explore the limitations of reading speed measurement methods (Chapter 2):
 - a. the inclusion of the pre-verbalisation time in the computation of reading speed.
 - b. the use of page-turn as a time recording method in a silent reading test.
 - c. the use of proofreading to ensure reading compliance in silent reading.
- II. investigate duration related effects in normally-sighted young adults (Chapter 3) and elderly (Chapter 4) with and without simulated impairment:
- III. explore the ability of commercially available reading tests such as IReST and RoR to capture reading duration related effects (Chapter 3).
- IV. devise different analysis methods to investigate duration related effects in reading performance tests (Chapter 3, 4, and 5).
- V. report any duration related changes in a group of individuals with visual impairment when reading with optical magnifiers (Chapter 5).
- VI. explored the possibility of using blinking rates to index mental fatigue objectively during reading (Chapter 6).

The overall hypothesis of this study was that individuals with visual impairment suffer from duration related fatigue effects, that is, reduction in reading speed with time. This poor reading endurance performance may partially contribute to individuals giving up on leisure reading activity.

2. To explore the limitations of reading speed assessment methods

Reading has been raised as a major concern in low vision rehabilitation by the visually impaired community (Elliott et al., 1997, Owsley et al., 2009). Therefore, reading performance tests are frequently used as an outcome measure to determine the success of an intervention. Although reading speed can be used to evaluate the reading ability of a visually impaired person, there has not been a consensus on how reading speed should best be measured, in terms of the type of test (e.g., random words, single meaningful sentences, or extended paragraphs) or the calculation of reading speed (e.g., words per minute, standard words per minute, or characters per minute) or the timing of the test (e.g., by manual stopwatch (Altpeter et al., 2015), audio recording (Subramanian and Pardhan, 2009, Burggraaff et al., 2010) or video recording (Maaijwee et al., 2008)). More recently, the choice of using silent reading for an extended testing duration may pose challenges in determining the time recording effectively. The precise and repeatable measurement of reading speed is important and clinically relevant because it can be used to evaluate the effectiveness of a prescribed LVA.

2.1 Exploring the limitation of the inclusion of the pre-verbalisation time in the computation of reading speed

2.1.1 Introduction

The use of audio, video or stopwatch recording methods has been reported in studies (e.g. manual stopwatch (Altpeter et al., 2015, Subramanian and Pardhan, 2006), audio recording (Subramanian and Pardhan, 2009, Burggraaff et al., 2010) or video recording (Maaijwee et al., 2008)) as a time measurement method to calculate reading speed. The choice of time recording methods often depends on whether it is in a research or a clinical setting. In contrast to the stopwatch method, the use of audio or video recordings to track timing allows retrospective analysis and encourages a more repeatable measurement (Brussee et al., 2014). As such, audio or video recordings are usually more preferred in a research setting. However, it may not be practical to adopt audio recording in a routine low vision consultation. Stopwatch recordings are more commonly used in a clinical setting (West et al., 1997) as they are more straightforward to compute the reading speed in real time.

Regardless of the approaches, most of the investigators do not describe the timing process explicitly. Therefore, it is unclear “what” is included in the timing recorded from the reading test, which is subsequently used to compute the reading speed. For instance, the investigator can choose to initiate the timing either once the reading text is uncovered, or when the first utterance or word is verbalised. If the timing were chosen to be initiated when the reading text was uncovered, it would also include the time taken by the reader to locate and mentally process the first word on the reading passage. Although both the stopwatch and audio recording methods can have the option to include the pre-verbalisation time, it may encourage a more consistent measurement for the stopwatch to initiate the timing once the text is uncovered rather than initiating at the first utterance; as the former may reduce the variation that can be caused by false starts due to rater’s reaction to the first utterance or word verbalised. This is clinically important because the variation in time measurements could lead to interpretation errors when using the reading speed as a performance indicator in a pre and post-intervention. However, the impact of such variation on the reading speed measurements

remains unclear. On the other hand, false start errors do not occur in audio recording because a sound editor software can be used to determine the length of time of the verbalised text.

In a recent review, Rubin (2013) commented that the imprecise reading speed measurements may be attributed to various factors. One of the factors was the reaction time or false starts by the examiner when using a stopwatch. Indeed, Calabrèse et al. (2018) reported a difference of up to 12% in reading speed measurements between using the stopwatch and the iPad method. The objective of their experiment was to compare the printed copy and the electronic version (i.e. iPad) of the MNREAD acuity chart. A reading speed difference ranging from 3% at 100 wpm to 12% at 250 wpm was reported, and Calabrèse et al. (2018) concluded that the observed differences were due to the result of reaction time on operating the stopwatch. The authors used a stopwatch for the printed MNREAD acuity chart, and a time delay was observed between uncovering the text and clicking the stopwatch button. As there was a time delay, the shorter reading time captured using the stopwatch produced a faster reading speed compared to the iPad method. The timer and the flashing of the text were initiated by a single touch on the screen using the iPad. Their findings suggested that technical glitches such as reaction time can result in a more variable reading speed measurement, especially in participants with higher reading speed (i.e. 250 wpm). A lag timing of approximately 0.3 seconds using the stopwatch compared to audio recordings has also been previously reported by Xu and Bradley (2015) in a group of normally-sighted participants. To compare the reading speed measurement by the digital audio recordings and stopwatch method, the stopwatch was initiated upon the first utterance by the participant (i.e. instead upon uncovering the text) and only the verbalised text was used in digital audio for the computation of the reading speed. The average coefficients of repeatability were 0.02 seconds and 0.4 seconds for the audio and stopwatch recordings respectively, suggesting that the rater's reaction to the first utterance from the stopwatch method can produce a more variable time recording compared to the audio recordings. Previously, Trauzettel-Klosinski et al. (2012) have suggested that the use of longer reading texts can help to address the impact of the technical glitches, which can affect the precise measurement of the reading speed. There are two possibilities why the use of longer reading text can provide more advantages over short sentences for overcoming imprecise

reading speed measurements. The effect of this time difference between the start of the timer and uncovering the text would be diluted with the increased number of words (e.g. 100 words vs 10 words) in a longer reading text. Although the impact of the differences between stopwatch and audio recording has been reported a 12% difference in a reading speed of 250 wpm using short sentences (Calabrèse et al., 2018), the use of an extended reading text to mitigate the time difference has yet to be investigated.

Currently, there is no consensus on how the time recording should be conducted and what should be “included” in the timing to compute the reading speed parameter. This is more significant when using reading speed as an outcome measure is that the comparison should be under equivalent conditions, and be as reproducible as possible. The audio recording method was used in reading speed measurements in the portable-Electronic Vision Enhancement System (p-EVES) study. As such, the “pre-verbalisation” would include the time locating the starting point in the text with different magnifiers, which might compromise the comparison of the reading speed between the two visits. The purpose of this experiment was to find out the difference in results obtained using the audio recordings compared to the manual usage of a stopwatch to compute reading speed.

2.1.2 Methods

2.1.2.1 Participants

Thirty-eight normally-sighted participants were enrolled in this study. All were students with an age range of 18-40 years from the Singapore Polytechnic. The main inclusion criteria required participants who spoke fluent English and did not have reading difficulties (i.e. dyslexia). Near VA was tested with their optimised refractive correction using a Bailey-Lovie word-reading chart at a distance of 25 cm. All of them had a minimum near VA of at least 0.1 log MAR. Given that many reading speed measurements are done in patients with low vision, the participants were also simulated with reduced vision using the Bangerter occlusion foils. The Bangerter occlusion foils are plastic diffuser film of different grades, that can produce image degradation (i.e. reduced contrast and visual acuity) with predictable visual acuity (Odell et al., 2008). When visual impairment was simulated with the Bangerter occlusion foil, a minimum

near VA of 0.4 log MAR was achieved to allow testing using IReST which comes only in a single font size. For easy referencing of the two conditions, the normal and Bangerter occlusion foil simulated conditions will be abbreviated as NOR and BOF respectively in this report. It was hypothesised that under less than optimal viewing condition, participants would take more time to decipher the words and this time lag would not be captured in the audio recording. Therefore, the difference in timing would be more apparent in the BOF.

2.1.2.2 Test Procedures

The order of the reading tests was randomised, and different versions of the charts were used for NOR and BOF to avoid a learning effect. As the Bailey-Lovie word-reading chart was used initially in the eligibility assessment, IReST and MNREAD acuity chart were also shown to the participants to familiarise them with the layout before the actual test began. For both the MNREAD acuity chart and Bailey-Lovie word-reading chart, each sentence (MNREAD acuity chart) and the row of words (Bailey-Lovie word-reading chart) were only uncovered one at a time during the test. This was performed to prevent prior viewing of the following reading text of the testing charts. Participants started with the 1.0 log MAR line for both the logarithmic reading charts (i.e. MNREAD acuity chart and Bailey-Lovie word-reading chart) and continued until they were unable to read any word in the next sentence or row of words.

The reading distance was monitored to ensure it remained at 40 cm throughout the tests. The reading tests were placed on a reading stand approximately 30° below the eye level. Participants were instructed to read the reading texts out-loud, and as fast and accurately as possible. If mistakes were encountered, they were to continue reading without correction. Stopwatch timing commenced once the sentence was flashed and stopped at the last word verbalised. No attempts at modification (i.e. anticipating the movement of the lips (Radner et al., 2002)) of the stopwatch method was made. Concurrently with the stopwatch timing, the oral reading was also recorded using a voice recorder as a reference (Sony Digital Voice Recorder ICD-UX533F). As there was no cue to the revealing of the reading text, the first utterance or word verbalised by the participant marks the starting point for the audio recording. The oral recordings were subsequently analysed by WavePad Sound Editor (Version 6.64 by

NCH software). This study was approved by the Singapore Polytechnic Ethics Research Committee (Protocol no: 201501-12) and written informed consent was obtained from all participants.

2.1.2.3 Data analysis

The initial part of the data analysis identifies the time difference between the two methods of time recording in the three reading tests. However, the clinical relevance of the time difference was hard to interpret. Hence, reading speed would be a more appropriate assessment parameter, since this would allow comparison of the final calculated parameter of all the reading tests. Reading mistakes were taken into account for reading speed analysis using retrospective analysis from the audio recording. As the IReST text comes only in a standard text size equivalent to 0.4 log MAR (i.e. N8 at 40 cm), only the same text size in both MNREAD acuity chart and Bailey-Lovie word-reading charts were selected for comparative analysis. This was to remove the influence of font size on the reading speed measurement.

Reading speed in words per minute (wpm) was calculated based on the number of correct words read divided by the time taken in minutes to complete the reading test. Any misread or missed words were omitted from the calculation. The 10 IReST reading paragraphs (i.e. 136 to 165 words) were randomised among the participants. For the Bailey-Lovie word-reading chart, acuity at 0.4 log MAR line consists of 6 words: two 4-letter words, two 7-letter words and two 10-letter words (Bailey and Lovie, 1980). For MNREAD acuity chart, each sentence consists of 10 standard words (i.e. 6 characters in each word). Each word missed or misread was considered as an error. The reading speed was then calculated based on the formula recommended by the manufacturer:

Reading speed = $60 \times (10 - \text{errors}) / (\text{time in seconds})$.

2.1.3 Statistical analysis

All data were analysed using Statistical Package for the Social Sciences (SPSS) (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.).

Shapiro-Wilk test indicated a non-normal distribution for the time recordings. Therefore, median and interquartile range measurements were reported. Non-parametric statistical analysis Friedman test was used to compare the time differences between the different reading tests. Bland-Altman (Bland and Altman, 1986) analysis was then used to explore the differences in reading speed between audio recording and stopwatch methods. A probability of less than 0.05 indicated statistical significance.

2.1.4 Results

The data for all 38 participants were used for the analysis. Table 2.1 shows the number of word errors made by the participants in each reading test under the NOR and BOF conditions. In contrast to the IReST and MNREAD acuity chart, the Bailey-Lovie word reading acuity chart had the largest number of participants making word errors. However, only 4 of them had more than 3 (i.e. more than 50%) misread errors. All word errors were accounted for (i.e. omitted for the calculation of the reading speed).

Table 2.1 Number of participants making the word errors

Reading Test	NOR (n = 38)		BOF (n = 38)	
	Number of participants	Number of word errors made	Number of participants	Number of word errors made
Bailey-Lovie word reading acuity chart	1	1	15	Range: 1-3
MNREAD acuity chart	0	Nil	5	Range: 1-5
IReST	0	Nil	0	Nil

An average reading time taken was calculated between audio recording and stopwatch: these average reading times taken for IReST, MNREAD acuity chart and Bailey-Lovie word-reading chart ranged from 35.62 to 159.70 seconds, 1.98 to 12.40 seconds and 2.41 to 21.39 seconds, respectively. Time differences between stopwatch and audio measurement were calculated. Figure 2.1 shows the plot of the individual participants' time differences for each reading test. Data for one of the participants was not shown in Figure 2.1 plot as the time difference between audio recording and stopwatch was approximately 12 seconds. However, the data was included in all the analyses.

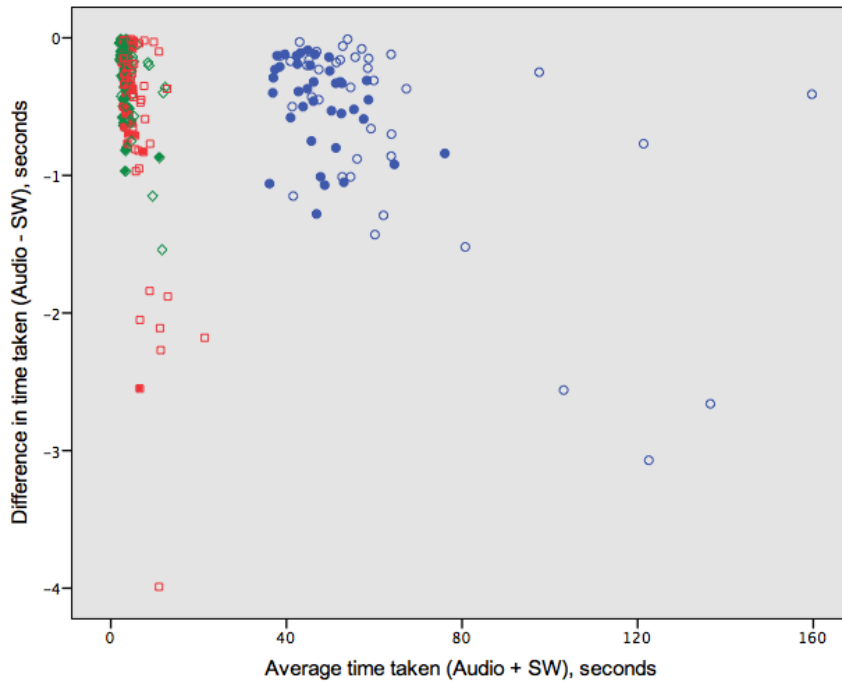


Figure 2.1 Time taken difference plot for the three reading tests between audio recording and stopwatch. Bailey-Lovie word-reading chart = ■ (habitual) ◻ (BOF), IReST = ● (habitual) ◊ (BOF), MNREAD acuity chart = ◆ (habitual) ◊ (BOF).

The median time differences (audio - stopwatch) were – 0.37 seconds (IQR: 0.62 seconds, range: – 0.01 to – 3.07 seconds), –0.23 seconds (IQR: 0.32 seconds, range: – 0.01 to – 1.54 seconds) and – 0.34 seconds (IQR: 0.52 seconds, range: – 0.01 to – 11.93 seconds) for IReST, MNREAD acuity chart and Bailey-Lovie word-reading charts, respectively. Among the three tests, the MNREAD acuity chart showed the least median difference in timings, followed by Bailey-Lovie word-reading chart and IReST. All the audio time measurements were shorter compared to stopwatch timing due to the time lapse between uncovering the text and verbalising of the first word. A Friedman test was conducted to compare the time differences between the reading tests. There was a significant difference between the audio recording and stopwatch measurements in the three reading tests, $\chi^2(2) = 13.84, p = 0.001$.

Since reading speed is the primary measurement determined by these reading tests, comparing the differences regarding reading speed would thus provide a better indication if such differences were clinically significant. Therefore, the reading speeds for all the three

reading tests were calculated. Table 2.2 compares the median reading speed calculated using the audio recording and stopwatch measures for the three reading tests.

Table 2.2 Descriptive statistics on the reading speed using audio recording and stopwatch measures for the three reading tests

Reading Tests	Median reading speed (IQR), wpm Range	
	Audio	Stopwatch
IReST	187.42 (51.82) 54.92 to 266.74	186.33 (51.11) 54.78 to 264.64
MNREAD acuity chart	198.5 (79.50) 27.50 to 306.12	175.96 (77.05) 24.10 to 300.00
Bailey-Lovie word-reading chart	82.92 (56.78) 2.96 to 155.17	75.39 (49.33) 2.67 to 144.00

Interquartile range (IQR) values are shown in the parentheses

The median reading speed differences between audio recording and stopwatch for IReST, MNREAD acuity chart and Bailey-Lovie word-reading chart were 1.31 wpm (IQR: 1.48 wpm, range: 0.02 to 7.82 wpm), 12.13 wpm (IQR: 16.22 wpm, range: 0.07 to 61.83 wpm) and 5.64 wpm (IQR: 6.92 wpm, range: 0.10 to 33.94 wpm), respectively. As the reading time was shorter using the audio recording measurement, the audio recording gave a higher reading speed measurement compared to the stopwatch timing method.

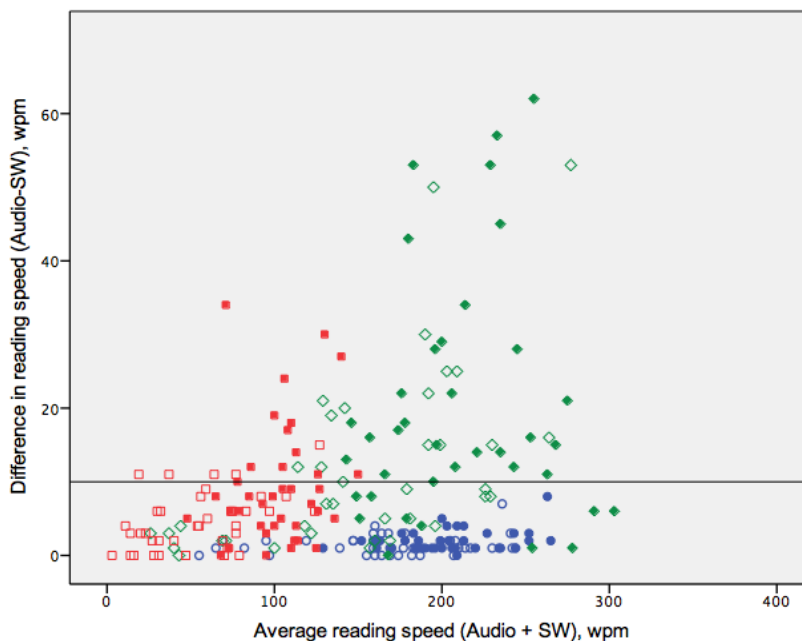


Figure 2.2 Difference plot for all the reading tests between audio recording and stopwatch. Bailey-Lovie word-reading chart = ■ (habitual) □ (BOF), IReST = ● (habitual) ○ (BOF), MNREAD acuity chart = ◆ (habitual) ◇ (BOF). The solid black line indicates the cut off reading speed at 10 wpm.

Figure 2.2 shows the difference plot in reading speed measurements for all the three reading tests. The solid black line represented a cut-off value of 10 wpm. Differences in reading speed of more than 10 wpm have been proposed as clinically different (Trauzettel-Klosinski et al., 2012, Altpeter et al., 2015). From the graph, larger differences were observed in participants using Bailey-Lovie word-reading chart and MNREAD acuity chart compared to IReST. The maximum differences were 61.83 wpm for MNREAD acuity chart and 33.94 wpm for Bailey-Lovie word-reading chart. In Bailey-Lovie word-reading chart, a difference in reading speed of more than 10 wpm was observed in most participants with reading speed above 100 wpm. Similarly, in the MNREAD acuity chart, majority of the participants with faster reading speed (i.e. > 150 wpm) showed a larger difference of more than 10 wpm. Among the three reading tests, IReST showed the least differences of up to 7.8 wpm.

Table 2.3 Median and interquartile range of reading speed differences between audio recording and stopwatch in NOR and BOF

Reading test	Audio recording - Stopwatch			
	median time difference (IQR), seconds range		reading speed difference (IQR), wpm range	
	NOR	BOF	NOR	BOF
IReST	-0.38 (0.42) -1.28 to -0.09	-0.36 (0.76) -3.07 to -0.01	1.56 (1.71) 0.43 to 7.82	1.00 (1.54) 0.02 to 6.59
MNREAD acuity chart	-0.20 (0.31) -0.97 to -0.01	-0.24 (0.34) -1.54 to -0.03	14.78 (20.00) 0.47 to 61.83	7.60 (13.26) 0.07 to 53.03
Bailey-Lovie word-reading chart	-0.23 (0.34) -2.55 to -0.02	-0.39 (0.76) -11.93 to -0.01	7.44 (8.20) 0.38 to 3.94	4.04 (5.20) 0.10 to 15.03

Interquartile range (IQR) values are shown in the parentheses

Table 2.3 shows the median time and reading speed differences between audio and stopwatch. In this table, all the reading tests under each condition were tabulated separately. Apart from IReST, the time differences were slightly higher in the BOF, especially in Bailey-Lovie word-reading chart.

2.1.5 Discussion

The current experiment used a convenience sample of 38 participants. Hence, a post hoc power analysis was conducted using the G*Power (Version 3.1.9.3). Two-tailed t-tests for repeated measures was used on the sample of 38 participants, which gave a power of 99% to detect a difference of 1.13 wpm between the audio and stopwatch method (see Appendix 9.10.1).

In this experiment, the differences in reading speed computation were determined by comparing the audio recording to a manual stopwatch on three different reading text lengths. Unlike the audio recording, which computed the reading speed based on the verbalised text, the manual stopwatch method described here included the pre-verbalisation time in addition to the verbalised text to calculate the reading speed. As noted earlier, there was no consensus if the pre-verbalisation time should be included in the calculation of the reading speed. The disadvantage of capturing only the verbalised text using the stopwatch was a false start on the timer due to rater's reaction. Alternatively, the stopwatch timer could be initiated once the text is uncovered by the investigator to avoid false starts and maintain consistency between measurements. However, the pre-verbalisation time would then be reflected in the reading speed. The inclusion of the pre-verbalisation time could be subjected to the potential variability of the measurement. For instance, the unfamiliarity of the test or in individuals with low vision who are using their LVAs; the individuals may need some time to locate the first word using the LVAs. Although the pre-verbalisation time information may be useful to understand the ability of the reader to identify the first word, especially in visually impaired individuals when they are using their LVAs, arguably, this may not truly represent the reading speed per se and potentially create a source of variability between measurements. Therefore, a more valid measurement would be to assess only the verbalised text to compute the reading speed. Here, we attempt to find out how different would the measurement be if the pre-verbalisation time (i.e. the stopwatch method) was included. It is not the intention of this experiment to determine which method works the best, but rather, to investigate the extent of the differences between the two measurements in the three different reading texts.

The inclusion of the pre-verbalisation time on the reading speed computation has a different effect with different reading text lengths. Our results showed that the time difference was the highest in IReST compared to the MNREAD acuity chart and Bailey-Lovie word-reading chart. On the other hand, the MNREAD acuity chart produced the least time difference relative to the other reading tests. Although the time differences were observed to be statistically significant in the three reading tests, these time differences appeared clinically insignificant (i.e. < 1.0 second). In contrast to Calabrèse et al. (2018) and Xu and Bradley (2015), all our stopwatch recorded timings were consistently longer than the audio recorded timings. This was because our stopwatch method includes the pre-verbalisation time and the verbalised text, but the audio recordings were only limited to the verbalised text. Also, when participants were subjected to simulated Bangerter occlusion foils, a greater difference between the audio recording and stopwatch was observed. The bigger time lapse when using the stopwatch could be explained due to more time required to locate and process the first word of the sentence.

Although IReST had the largest time difference among the three reading tests, the IReST produced the least variation in reading speed because it is the longest test. In the current study, one of the participants showed a large difference (approximately 12 seconds) only in the Bailey-Lovie word-reading chart. The large difference was due to the time taken for the subject to guess the first word on the Bailey-Lovie word-reading chart after uncovering the words. It was initially suspected that the Bangerter occlusion foil simulation might have caused the patient to be unable to read 0.4 log MAR line. However, the participant did not show a significant time difference in the other two reading tests. It was likely that the syntactic structure of IReST and MNREAD acuity chart had helped the reader to guess the first word of the sentence, unlike with the Bailey-Lovie word-reading chart where random words were used.

Assuming the reading speed was calculated based on an average reading time of each reading test, reading speed was then calculated for the Bailey-Lovie word-reading chart, MNREAD acuity chart and IReST. The reading speeds for all the three reading tests were tabulated in Table 2.4 with the same time differences for two different reading speed profiles. From Table 2.4(A), with a given time difference of 1.0, the difference in reading speed

generated were 10 wpm, 30 wpm and 4 wpm for Bailey-Lovie word-reading chart, MNREAD acuity chart and IReST respectively. Unlike IReST, both the Bailey-Lovie word-reading chart and MNREAD acuity chart consist of a lesser number of words in their reading test; which explained why a human technical error of one second would not result in a significant change in reading speed using IReST. On the other hand, as seen in Table 2.4 (B), if the reading speed is generally slower (i.e. below 100 wpm), the impact on the reading speed of having a lower number of words to calculate the value is lower. Although the Bailey-Lovie word-reading chart has only 6 words, the reading speed difference was less significant compared to the MNREAD acuity chart. The reduced reading speed observed in the Bailey-Lovie word-reading chart was due to the usage of random words, which made the reading less fluent. On the other hand, the MNREAD acuity chart, which uses continual text sentences had the least time difference; but this produced the greatest reading speed difference (see Table 2.4 (A)).

Table 2.4 Differences in reading speed based on same time difference

(A) Participant with high reading speed in NOR

Difference in time (seconds)	Bailey-Lovie word-reading chart 6 words		MNREAD acuity chart 10 standard words		IReST ~ 150 words	
	0.5		6.0 secs (60 wpm)		4.5 secs (133 wpm)	
1.0	5.5 secs (65 wpm)	6.5 secs (55 wpm)	4 secs (150 wpm)	5.0 secs (120 wpm)	56 secs (161 wpm)	57.0 secs (158 wpm)
1.5		7.0 secs (51 wpm)		5.5 secs (109 wpm)		57.5 secs (157 wpm)

(B) Participant with lower reading speed in BOF

Difference in time (seconds)	Bailey-Lovie word-reading chart 6 words		MNREAD acuity chart 10 standard words		IReST ~ 150 words	
	0.5		11.0s (33 wpm)		10.0s (60 wpm)	
1.0	10.5 secs (34 wpm)	11.5s (31 wpm)	9.5 secs (63 wpm)	10.5s (57 wpm)	159.5 secs (56 wpm)	160.5s (56.07 wpm)
1.5		12.0s (30 wpm)		11.0s (54 wpm)		161s (55.90 wpm)

As reading speed is highly variable between individuals, there is no exact cut-off value to indicate how much difference in reading speed would be clinically relevant; this is likely to be markedly different for fast and slow readers. There were different clinical interpretations as to how much difference in repeated reading speed test was clinically relevant (e.g. is the difference in repeated measurements due to a true difference or measurement error?).

Previously, Subramanian and Pardhan (2006) reported a 4% reading speed variability (coefficient of repeatability = ± 8.6 wpm, mean reading speed = 210 wpm) in a group of healthy individuals using MNREAD acuity chart. Subsequently, Subramanian and Pardhan (2009) furthered their investigation on the repeatability of the MNREAD acuity charts in a group of low vision participants. As the range of reading speed in low vision patients can be highly variable, reporting the variability based on words per minute only does not relate the impact of the change on a participant (e.g. 20 wpm would be a 20% change for a reader with 100 wpm, but 50% for a reader with 40 wpm). To counter this problem, Subramanian and Pardhan (2009) log transformed the reading speed measurements and reported the coefficient of repeatability as a change in log units. A minimal mean difference of the reading speed between the two visits (0.02 log wpm (confidence interval: 0.12 to - 0.09 log wpm)) and a small coefficient of repeatability of 0.10 log wpm was reported. The coefficient of repeatability reported by Subramanian and Pardhan (2009) was much lower (i.e. better repeatability) compared to Patel et al. (2011) (i.e. coefficient of repeatability: 0.22 to 0.25 log wpm based on different scoring methods). The larger variation was probably due to the measurements being conducted on the same day by the same rater in Subramanian and Pardhan (2009) as compared to Patel et al. (2011), where the test-retest was conducted over a period of 6-week with different raters. Also, the use of stopwatch as a timing device in Patel et al. (2011) as compared to audio recording used in Subramanian and Pardhan (2009) could potentially contribute to the higher coefficient of repeatability. However, explicit details of how the stopwatch were used to record the timing was not reported (i.e. does the timing commence when the rater says "start" or the first word verbalised?). Using the largest change of up to 25% in Subramanian and Pardhan (2009), the authors suggested that the repeated reading speed measurement differences which were less than 25% were most likely due to measurement errors. More recently, Alpeter et al. (2015) investigated the difference in reading speed measurement using an extended reading text (i.e. IReST) and single sentences (i.e. Radner reading chart). Based on an earlier study from the same group (Trauzettel-Klosinski et al., 2012), they proposed a reading speed variation of more than 10 wpm as clinically relevant. However, it was not clearly explained how this value was derived in the former study by Trauzettel-Klosinski et al. (2012). Since the current experiment involved healthy participants and both Subramanian and Pardhan (2006)

and Trauzettel-Klosinski et al. (2012) reported a similar variability range, the use of 10 wpm would be a reasonable cut-off value in this experiment. If a difference in reading speed of more than 10 wpm is clinically relevant and reading speed within 10 wpm was considered to be of the same performance level (Trauzettel-Klosinski et al., 2012, Altpeter et al., 2015), then both time recording measurements could be considered similar in both IReST and Bailey-Lovie word-reading chart. However, the difference in the time recording measurements for the MNREAD acuity chart produced a reading speed difference of more than 10 wpm.

Based on the current study, except for the MNREAD acuity chart (i.e. median difference of approximately 15 wpm), the differences in reading speed measurements with the inclusion of the pre-verbalisation time were not clinically relevant (i.e. less than 10 wpm). However, a slightly higher variation of up to 15 wpm could be expected in participants with higher reading speed (i.e. more than 200 wpm) and using a short reading text (e.g. one sentence). It is unclear if this also applies to participants using a LVA for reading speed assessment. When determining the use of reading speed measure as an outcome measure in a clinical or research setting, it is therefore important to consider the factors such as the time recording method and reading text length. As there were no clinically significant differences when including the pre-verbalisation time in the longer reading text (i.e. IReST), the reporting of reading speeds determined by audio recording in our subsequent experiments would seem appropriate and any clinically significant effects observed would be equally evident when using a stopwatch.

2.1.6 Limitations

There were a few limitations that should be mentioned. As noted earlier, it was not the intention of this experiment to determine the best method for time recording (e.g. stopwatch or audio recording) but how the audio measurements were comparable to the stopwatch method. If the objective of the experiment was to determine which method produces a “better” measurement, then both methods would have to measure the “same” thing. For instance, the audio recording should include the pre-verbalisation time (i.e. the investigator would verbalise “start” when he or she uncovered the text so that it can provide verbal cues for audio recording to capture the

pre-verbalisation time). Subsequent work should also include the assessment of the repeatability and reproducibility of the measurement methods. Although we have simulated a group of healthy participants with reduced vision, we did not include participants using a LVA where they would potentially require a longer time to search for the first word. Also, we have investigated the extent of the variation but not the reproducibility of the differences. The reproducibility of the reading speed is clinically important because it can impact on the outcome when it is assessed pre and post intervention. Further work could be done to investigate the reproducibility of the reading speed by having repeated testing, preferably at an interval of 10 days (McMonnies, 2001).

2.2 To explore the use of page-turn as a time recording method in a silent reading test

2.2.1 Introduction

In reading performance assessment tests, the participants can be tasked to perform either oral or silent reading for a given text. Comparatively, oral reading has been a more common approach in reading assessments (Legge et al., 1985a, Legge et al., 1985b, Lovie-Kitchin et al., 2000). This method requires no extra equipment and is straightforward to ensure the participant's reading compliance in reading performance tests. More importantly, the oral reading enables the researcher to determine the beginning and the end of the reading test easily to calculate the reading speed. Silent reading, albeit a more natural reading behaviour, requires strategies such as infrared eye movement tracker (Lovie-Kitchin et al., 2000) or using page-turns (Ramulu et al., 2013b) to identify the length of reading duration. In a recent investigation on reading speed over a duration of 30 minutes, Ramulu et al. (2013b) created a sustained silent reading test which is made up of approximately 7500 words. There are 100 or 200 words on each page, and the participants turn the page on their own after completing reading each page. As there was no verbal cue during the reading test, the investigator relies on each page-turn to determine the time taken to read the corresponding page of words. Given that the objective of Ramulu et al. (2013b) study was to investigate the reading speed changes with time, the time taken for each page is then used to chart the reading speed over a duration of 30 minutes. In the current study, the participants were required to use the hand-held magnifier for reading, and they may struggle to flip the pages while manipulating the magnifier. Since silent reading text would be used in our subsequent experiments, it is important to establish if there is any limitation in using the page-turn to track the reading time for each page. Technical glitches could potentially occur when flipping the pages, especially when participants have to juggle between flipping the pages and using the magnifier. Since the page-turn method would be used in our subsequent experiments to document the time duration for silent reading, it is essential to know the limitations of the reading speed documented using the page-turn method. Therefore, the objective of this experiment was not intended to determine which would be a better method to track recording time per se but to explore if there

is any limitation in using the page-turn method to determine the reading time during silent reading.

In the audio recording method, only the verbalised text was taken into account for the calculation of reading speed. As for the page-turn method, the start of the page flipping movement was used as an indicator of the end time for that particular page, and the start time for the next page. Thus, the recorded time includes the time taken to flip the page and locate the first word of the second page. As the page-turn method included the time to flip the pages and locate the first word, there could be an element of variation which would be reflected by large time differences between the page-turn and audio recording timing over the 10 reading pages. On the other hand, it could be that the differences between the two methods remain the same throughout the 10 reading pages test.

2.2.2 Methods

Thirty normally-sighted young participants with a mean age of 21.77 (\pm 2.71) years were recruited from the University of Manchester. They were either native English speakers or had English as their first language. All of them had a near VA of at least 0.0 log MAR with their current spectacles using the Bailey-Lovie word-reading chart. For the purpose of comparing audio recording and page-turn in reading speed assessment, a reading test was crafted using reading materials from the New Scientist magazine. The texts were reproduced in Times New Roman 9 point font size and printed on 11 A4 pages. Each page had a paragraph which was 100-words in length and there were approximately 50-word on the 11th (final) page. The interval between page-turns was used to document the reading time. The 11th page served to allow the first 10 page-turns to be observed and analysed. Participants were instructed to read the text out-loud and at their comfortable speed as they would in their daily reading of a magazine or newspaper. After finished reading each page, participants were instructed to turn the page by themselves. To ensure consistency, the start of the page flipping movement was used as an indicator of the end time for that particular page, and the start time for the next page. The reading test was conducted in three different conditions to create diversified reading performances; habitual or normal (NOR), using Bangerter occlusion foils simulation for a

degraded image (BOF) and the use of magnifier with a denser Bangerter occlusion foil (MAG). All participants followed the same sequence of conditions (i.e. NOR → BOF → MAG) to allow a gradual increase in the difficulty of the reading conditions. As this group participants had no prior experience in using a hand-held magnifier, the investigator went through a practice session with the participants. No practice sessions were conducted prior to the NOR and BOF. The participants tried out using the hand-held magnifier for silent reading on 3 to 5 pages (i.e. performed a minimum of 3 page-turns) of sample reading text prior to the experiment.

The audio recording method used sound editing software to obtain the time taken to read from the first word till the last word of the page verbalised. As for the page-turn method, each page-turn indicates the time taken to read the corresponding page. This is done by using an online stopwatch to lapse the time taken to turn each page. Both methods then used the number of words on each page (100 words) divided by the time taken for each page, to derive the corresponding reading speed for each page. The overall average reading speed was derived using the total number of words (i.e. 10 pages = 1000 words) divided by the total time taken. The study was approved by the University of Manchester Research Ethics Committee (Reference no: 08046) and written informed consent was obtained from all participants.

2.2.3 Statistical analysis

All data were analysed using Statistical Package for the Social Sciences (SPSS) (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.). Shapiro-Wilk test showed a non-normal distribution for the reading speed measurements. Therefore, median and interquartile range measurements were reported. Non-parametric statistical analysis Wilcoxon Signed-Rank test was used to compare the reading speed between audio and page-turn methods. A probability of less than 0.05 indicated statistical significance.

2.2.4 Results

Twenty-six participants' data were used for the analysis; four participants' data were not available due to technical problems encountered while recording the magazine text oral reading. The overall median reading speed using audio recording and page-turn was tabulated in Table 2.5. Wilcoxon Signed-Rank test showed that the use of audio recording produced a statistically significant higher reading speed compared to the page-turn ($z = -7.67$, $p < 0.001$). The overall median difference in reading speed between audio recording and page-turn for all the reading conditions was 2.31 wpm (IQR: 4.86 wpm, Range: 0.10 to 16.03 wpm), indicating a higher reading speed was obtained using audio recording compared to the page-turn method.

Table 2.5 Descriptive statistics of the reading speed between audio recording and page-turn

	Median reading speed, wpm		
	NOR	BOF	MAG
Audio	156.02 (19.45)	156.29 (23.93)	114.58 (21.33)
Page-turn	155.52 (22.38)	154.16 (24.49)	108.00 (21.84)

Interquartile range (IQR) values are shown in the parentheses

The objective of this experiment was to explore the potential limitation of using page-turn to document time recording in a silent reading test. The difference in reading speed between the audio and page-turn methods was computed for the 10 pages in each condition for all the participants. The median differences in reading speed (audio – page-turn) were 1.31 wpm (IQR: 5.61, range: – 29.20 to 48.54 wpm), 1.57 wpm (IQR: 4.85, range: – 14.47 to 22.95 wpm) and 5.81 wpm (IQR: 8.15 wpm, range: – 9.65 to 72.99 wpm) in NOR, BOF and MAG respectively. There was a procedural glitch for one of the participants (i.e. CG16) in MAG, where the reading text papers fell off the reading stand, resulting a large difference in reading speed of 72.99 wpm. The elimination of the outlier reduced the range to – 9.65 to 34.00 wpm. To determine the reading speed difference changes, the differences in the reading speed between audio and page-turn were plotted against the corresponding ten pages. Presumably the difference in reading speed between audio and page-turn would show a decreasing trend (i.e. learning effect), then the use of the linear function gradient slope would indicate if there

was an overall increase or decrease in the difference in reading speed. However, visual inspection of the scatterplot suggested that alternative curve fittings may provide a better fit of the data points. Alternative curve fittings were explored through the use of the SPSS regression curve estimation (e.g. logarithmic, inverse, exponential and power). Table 2.6 shows the R^2 values of the linear and curve fittings for least squares linear, logarithmic and inverse function; exponential and power functions were unable to generate a fitting curve for most data points, and hence are not presented in Table 2.6. The Kruskal Wallis test shows that the R^2 values were not statistically significantly different between the least squares linear, logarithmic and inverse function. There was a small group of participants who showed a better curve fitting with the inverse function (i.e. $R^2 > 0.5$); but the number of participants reduced from NOR ($n = 5$), BOF ($n = 3$) to MAG ($n = 1$).

Table 2.6 R^2 values for least squares linear, logarithmic and inverse function

	Linear	Logarithmic	Inverse	Kruskal-Wallis Test
NOR	0.14 (0.20) Range: 0.00 to 0.47	0.17 (0.24) Range: 0.00 to 0.61	0.20 (0.38) Range: 0.00 to 0.73	$X^2 = 2.74$ $p = 0.26$
BOF	0.09 (0.20) Range: 0.00 to 0.57	0.08 (0.22) Range: 0.00 to 0.66	0.08 (0.20) Range: 0.00 to 0.77	$X^2 = 0.20$ $p = 0.90$
MAG	0.08 (0.16) Range: 0.00 to 0.52	0.06 (0.13) Range: 0.00 to 0.70	0.07 (0.18) Range: 0.00 to 0.70	$X^2 = 0.10$ $p = 0.95$

Since there was no statistically significant difference between the use of the different function fitting curves, the gradient of the least square linear could then be used to denote any overall changes in the reading speed differences between audio and page-turn method. The median slopes were 0.00 wpm/ page (IQR: 0.88, range: – 3.86 to 2.61 wpm/ page), – 0.14 wpm/ page (IQR: 0.82, range: – 0.90 to 0.95 wpm/ page) and 0.13 wpm/ page (IQR: 0.62, range: – 2.50 to 0.86 wpm/ page). The linear slopes were further split into positive and negative gradient slopes. Figure 2.3 shows the graph plots of all the participants in the three conditions. Three extreme data points (i.e. CG16: 72.99 wpm at Page 3; and JY03: 48.54 wpm at Page 1, and 45.23 wpm at Page 2) were not shown on the graph plots in Figure 2.3, but were included in the analysis. There were 13, 15 and 10 participants who showed a negative gradient slope in the NOR, BOF and MAG conditions respectively. Although there were almost half the number of participants who showed a negative difference slope, the gradient was rarely more than 1.0

wpm/ page except for three participants (i.e. JY03-NOR: -3.86 wpm/ page, PP12-NOR: -1.44 wpm/ page and CG16-MAG: -2.50 wpm/ page).

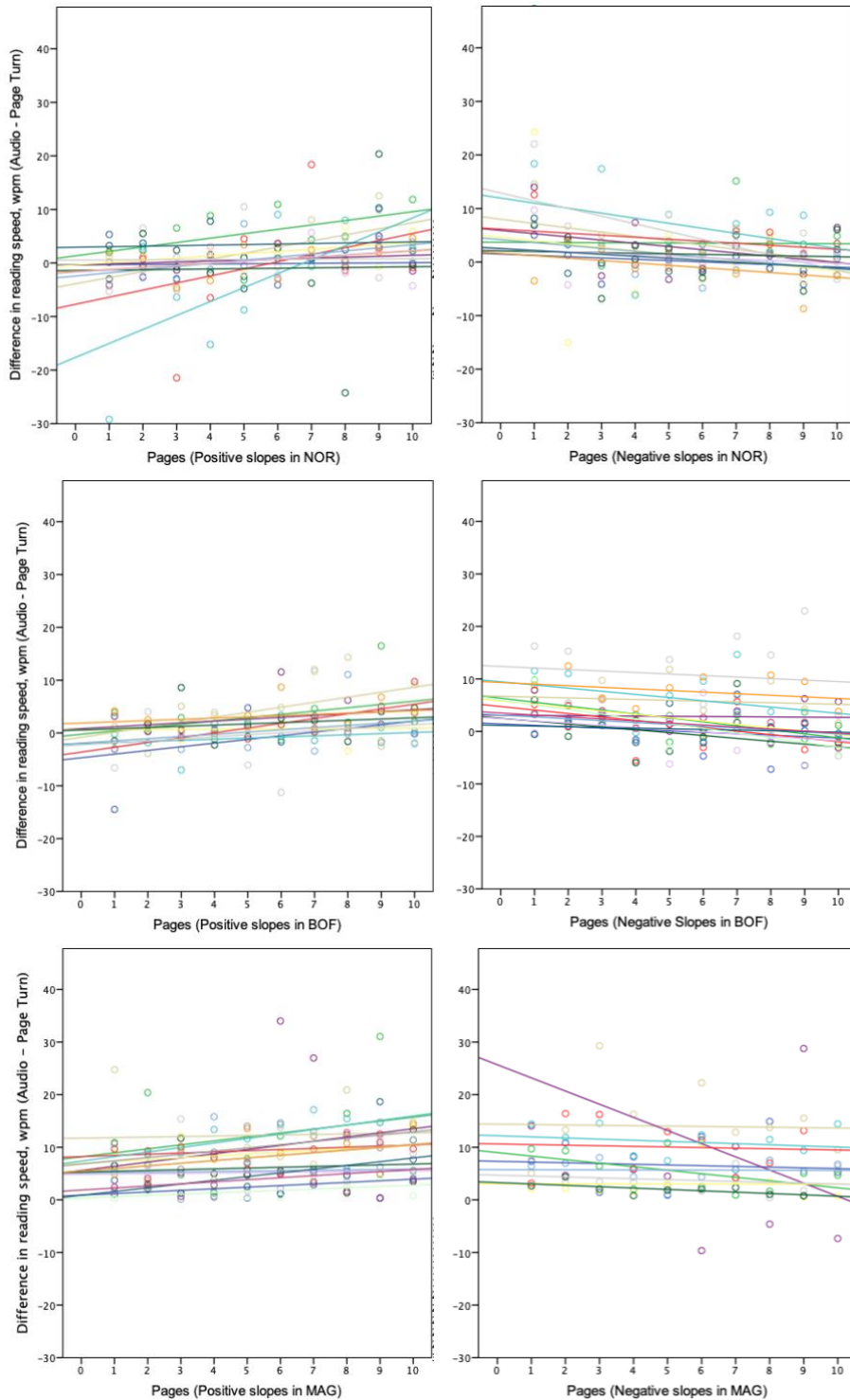


Figure 2.3 Graph plots of the difference in reading speed against pages for all the participants and all the conditions (Top: NOR, middle BOF, bottom: MAG) split into positive slopes (left column) and negative slopes (right column)

Figure 2.4 shows the graph plots for the three participants with negative difference slopes > 1.0 wpm/ page: Participant JY03 (NOR), PP12 (NOR) and CG16 (MAG). As seen from the graph, there was a large difference in the initial pages of reading speed; JY03: 48.54 wpm at Page 1, and 45.23 wpm at Page 2, PP12: 22.04 wpm at Page 1. These differences caused a large difference slope of $- 3.86$ wpm/ page and $- 1.44$ wpm/ page in Participant JY03 and PP12. The data from the participants with negative difference slopes typically showed a larger difference spike on the first page, which resulted in a negative difference slope. This observation can be seen in Figure 2.3, more evidently in the NOR (i.e. top right hand corner graph plot: negative difference slopes in NOR). As noted earlier, the outlier in CG16 in MAG has been highlighted, which was due to a procedural glitch. However, there was only one such case and on one page of the reading test.

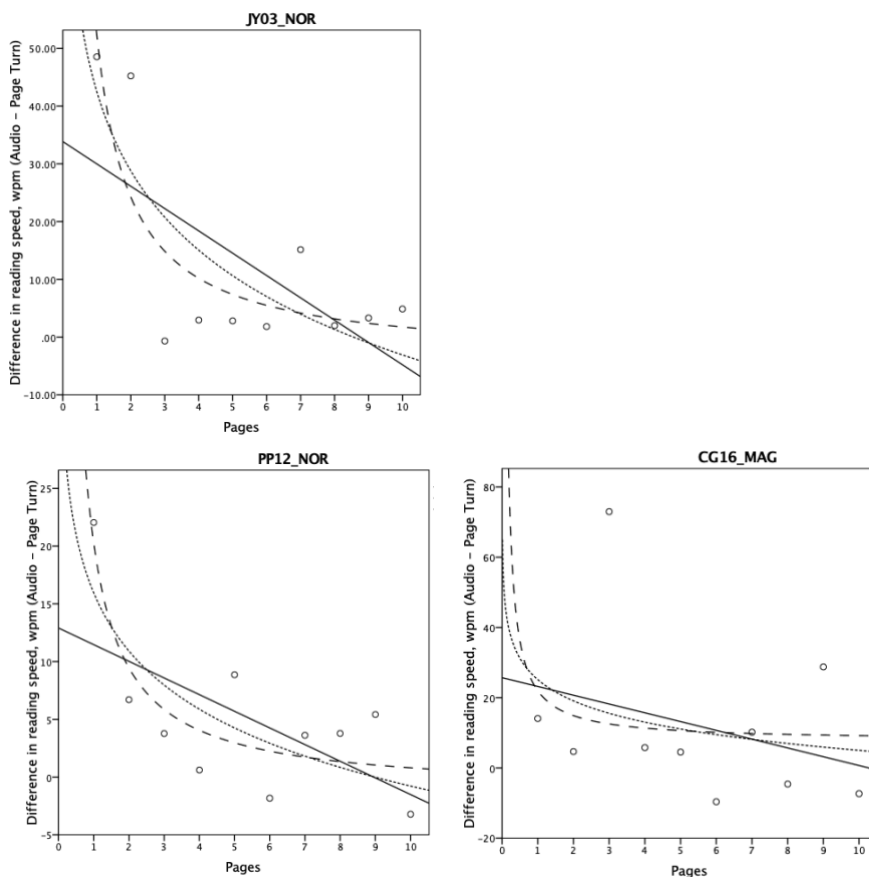


Figure 2.4 Difference in reading speed against pages graph plots for participant JY03 (NOR), PP12 (NOR) and CG16 (MAG). Dotted lines: logarithmic function, Dashed lines: Inverse function.

2.2.5 Discussion

The current experiment used a convenience sample of 26 participants. Hence, a post hoc power analysis was conducted using the G*Power (Version 3.1.9.3). Two-tailed t-tests for repeated measures was used on the sample size of 26, which gives a power of 97% to detect a difference of 0.79 wpm between the reading speed using audio and page-turn methods (see Appendix 9.10.2).

The purpose of the experiment was to explore the limitation of the page-turn method on tracking reading time to compute reading speed which has been previously reported by Ramulu et al. (2013a). It is, therefore, not the intention of this experiment to explore other possible methods to document reading time in silent reading (e.g. investigator can verbalise a cue so the page-turn can be used to capture the reading time using the audio recording in a silent reading test). As the page-turn method would be used to document time recording in subsequent experiments using a silent reading test, it is important to know the limitations of the method used to chart the reading speed with time. For instance, a learning effect may occur by the readers who struggled with the experiment setup initially (e.g. manipulating the magnifiers and flipping the pages). In this experiment, the reading speed changes using the page-turn were compared to the audio recording method. The participants performed the reading in three conditions (i.e. NOR, BOF and MAG), which were designed to represent different levels (and speeds) of reading difficulties. As expected, our results showed that the overall audio recording produced a higher reading speed compared to the page-turn method as the assessment only accounts for the words verbalised. Apart from the one of the participant in MAG (i.e. CG16), where a procedural glitch caused a large difference in reading speed (i.e. 72.99 wpm), a maximum difference of up to 48.54 wpm was observed comparing the audio and page-turn methods in all the reading conditions. Unlike in the previous section, where all the audio measurements produced higher reading speed compared to stopwatch method as the former only includes the verbalised words as the time duration, there was a mixture of higher and lower reading speed using the page-turn compared to audio in this current experiment. During the data collection, some participants were seen to flip the page before verbalising the last word on the prior page despite instructed to only flip the page after finished

reading the last word on the page. This has caused a shorter recorded time using the page-turn method compared to the audio recording, which has translated to a slower reading speed documented by the audio method.

More importantly, the graphs showing the difference in reading speed between the audio and page-turn against the pages were plotted to track if the differences are consistent throughout the 10 pages reading test. Although the R^2 values in Table 2.6 did not suggest a linear function would be the most appropriate fitting line, the gradient does provide indication on the overall changes in reading speed performance with time. The linear function slopes were used to determine if there is an overall increase or decrease in the difference in reading speed measurements. The negative slope graph plots were then further checked for any consistent changes in reading speed. Some of the participants who showed a negative reading slope had an improved curve fitting using the inverse function fitting curves – indicating that there was an initial large difference between the audio and page-turn method, and plateau after 2-3 pages of reading. The number of these occurrences diminished from NOR to MAG, suggesting that the participants got familiar with the reading test progress from NOR to MAG. In addition, participants only had the practice session in the last condition (i.e. MAG), where they have to use the hand-held magnifier for reading. When removing the first page analysis on the NOR data, the number of participants with negative difference slopes reduced from initial 13 participants to 6 participants (range – 0.13 to – 1.13 wpm/ page). This finding suggest that prior familiarisation of the test or omitting the first page for assessment would be necessary to minimise learning effects on the collected data in the subsequent experiments.

2.3 The use of proofreading to ensure reading compliance in silent reading tests

2.3.1 Introduction

In Section 1.2.3, we have discussed the pros and cons of oral and silent reading assessment methods adopted by different researchers in low vision studies. As the nature of the current study is to investigate duration related effects, it would be more appropriate to employ silent reading mode in our subsequent experiments. One of the major drawbacks in using silent reading mode is ensuring reading compliance during the test. There are various options adopted by researchers to ensure reading compliance in silent reading tests. One of the common approaches is to create multiple-choice questions to assess comprehension at the end of the reading test (Ramulu et al., 2013b, Bowers, 2000). These multiple-choice questions must be created in a way that they can only be answered by reading the text and not by general knowledge or logical thinking. One of the multiple-choice question approaches used by Dickinson and Rabbitt (1991) was to create a sentence recognition option. Participants read the passage and were tasked to choose the exact statement from four different options crafted in a structured way: 1) exactly the same as the reading text, 2) same meaning but different structure, 3) similar words and word order but with a different meaning and 4) a statement related to the story theme but was not found in the original story. An advantage of using the multiple-choice questions over cloze passages or free recalling of reading text is that it is easier to compute scores. However, it can be hard to determine the “passing” score for a reading test. In an ideal situation, a “full” score in the multiple-choice questions would indicate reading has taken place and “zero” score, if no reading has occurred. Practically, this would never be the case. Ramulu et al. (2013b) have shown that scores range from 0% to 39% can still be achieved even though the participants did not read the text, and 65% to 100% for participants who read the passage. When using multiple-choice questions, there is a possibility of getting a correct answer without reading the text (e.g. four choices equivalent to 25% chance). Despite the extensive efforts to develop and validate these multiple-choice questions, it is still questionable if the use of these multiple-choice questions is truly effective in ensuring reading compliance. The objective of this experiment is to explore the limitation of using multiple-choice questions scoring to ensure reading compliance. Seven multiple-choice questions were

crafted from a reading text abstracted from a science encyclopaedia with approximately 3500 words. To determine if the use of multiple-choice questions is an effective approach to ensure reading compliance, the scores between the readers and non-readers would be significantly different. On the other hand, if the scores are relatively similar, then it would suggest that the use of multiple-choice questions may not be appropriate as a mean to ensure reading compliance. Moreover, it may be challenging to craft out all the questions of a similar level of difficulty. As such, we would like to establish if a definitive score can be used to set the benchmark between readers and non-readers (e.g. is there a definitive high score to identify readers? or a very variable score?). Concurrently, a second method proofreading, was explored as an alternative approach to ensure reading compliance, since this would be more straightforward and easy to implement.

2.3.2 Methods

Forty-eight normally-sighted young participants with a median age of 18 years were recruited from the Singapore Polytechnic. All of them had a near VA of at least 0.0 log MAR with their current spectacles using the Bailey-Lovie word-reading chart. An abstract from a science encyclopaedia with an approximately 3500 words was used as the reading text. This reading text was created to mimic the extended period of reading, which would be used in our subsequent experiments. There were 100 words on each page, and the last page consisted of a random number of words to ensure that subjects turned over the page after they finished reading page 35. The participants were instructed to read the text silently as they would do with their daily reading materials such as a magazine or newspaper. In addition, they were asked to spot any typographical errors (e.g. “seen” spelt as “nees”) in the text. If they spotted an error on a page, they were to put that page on a separate tray after they finished reading the last word on that corresponding page. Within the reading text, there were seven typographical errors spread across the 35 pages. However, the number or frequency of the errors was not made known to the participants. After completing the reading text, the participants were also required to complete seven multiple-choice questions (see Figure 2.5 for a sample question. Appendix 9.3 consists of all the questions used in this experiment) related to the reading text. Each multiple-choice question consists of four choices, and the

participant has to choose one correct answer out of the four choices. There was no given time frame to complete the reading, but most of the participants completed the reading text within 15 minutes. To assess the extent of these multiple-choice questions that can be answered by general knowledge, the same questions were given out to a separate group of forty-eight participants without reading the text. Participants were told to try to attempt the questions based on their general knowledge. This study was approved by the Singapore Polytechnic Ethics Research Committee (Protocol no: 201501-12) and written informed consent was obtained from all participants.

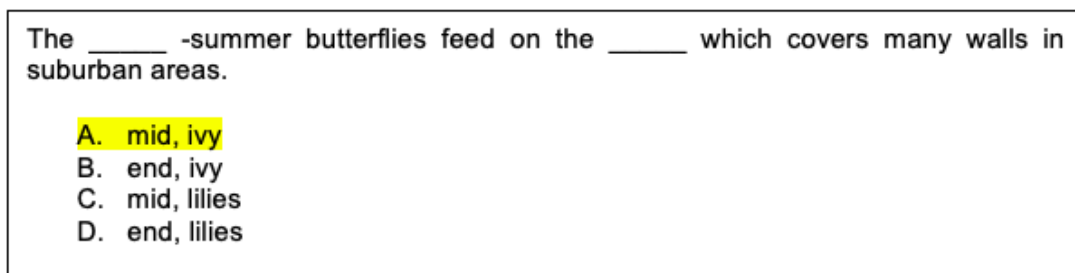


Figure 2.5 A sample of multiple-choice question used (All the seven questions can be found in Appendix 9.3).

2.3.3 Statistical analysis

All data were analysed using Statistical Package for the Social Sciences (SPSS) (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.). Shapiro-Wilk test showed a non-normal distribution for the multiple-choice questions ($p = 0.017$) and proofreading ($p = 0.01$) scores. Therefore, median and interquartile range measurements were reported. Non-parametric Mann-Whitney U tests were used to compare the scores of readers and non-readers. A probability of less than 0.05 indicated statistical significance.

2.3.4 Results

Using the multiple-choice questions, the participants who read the reading text had a median score of 4 out of 7 questions answered correctly (IQR: 2.75). Out of the 48 participants, 25 participants scored at least 50% (i.e. 4 out of 7 questions correct). In the group of participants who did not read the text and attempted the multiple-choice questions, a median score of 3 out of the 7 questions were answered correctly (IQR: 1.0). Figure 2.6 shows the distribution of the number of participants in each score for both the readers and non-readers. Thirty-nine participants scored less than 50% (i.e. ≤ 3 out of the 7 questions). Mann-Whitney U test indicated a statistically significant difference in the multiple-choice questions scores between the participants who read and those who did not read the text ($Z = -2.73$, $p = 0.006$).

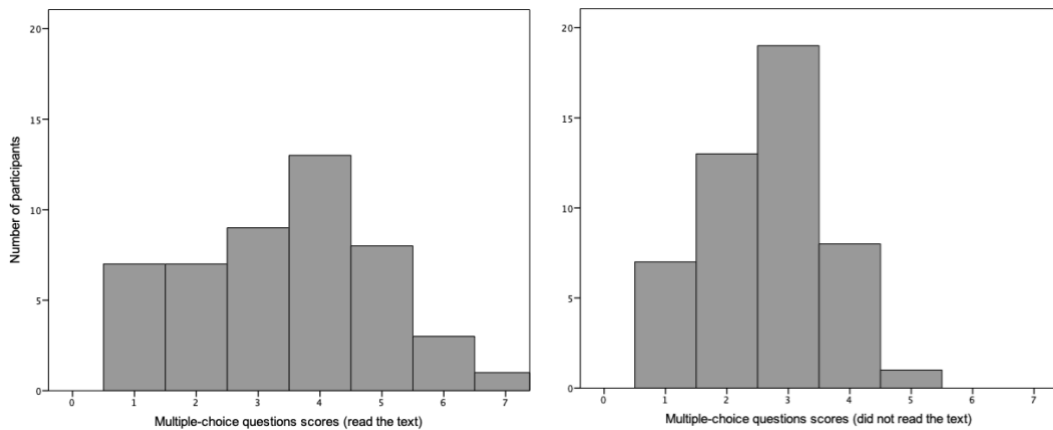


Figure 2.6 Distribution of the participant's score for both the reading and non-reading group

To determine the sensitivity of the multiple-choice questions scores to differentiate readers and non-readers, a Receiver Operating Characteristic (ROC) curve was used. Figure 2.7 shows the ROC curve with a small area under the ROC curve – AUROC: 0.65 ($p = 0.009$, 95% CI: 0.54 to 0.77), indicating that the multiple-choice questions scores have a poor discriminative ability to differentiate between readers and non-readers.

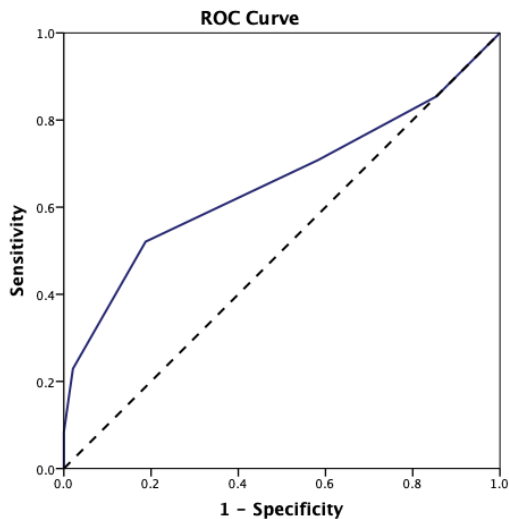


Figure 2.7 Receiver Operating Characteristic (ROC) curve for the multiple-choice questions scores

Table 2.7 shows the coordinates of the ROC curve for the multiple-choice questions scores. Using a mid-point cut-off value of 3.5, this would suggest 52% of the participants were correctly classified as readers, but 19% of the participants would still produce such a score without reading the text.

Table 2.7 Coordinates of the curve for the multiple-choice questions scores

reading occurs if multiple-choice questions score greater than or equal to	Sensitivity	1-specificity
1.5	0.85	0.85
2.5	0.71	0.58
3.5	0.52	0.19
4.5	0.23	0.02
5.5	0.08	0.00

So far, the individual total scores indicated an overall performance from these multiple-choice questions. It is, however, possible that the varying difficulty of the questions may limit the use of the multiple-choice questions. Even though the overall scores are similar, are there specific questions that can help to differentiate between the readers and non-readers? If this is the case, then we would expect these questions to have a high number of participants who read and answered correctly, and a low number of non-readers participants who answered

correctly. Figure 2.8 shows the number of participants who answered correctly for each multiple-choice question. Hypothetically, the multiple-choice questions should be answered correctly and scored higher in the readers' group. However, there were two questions where more non-readers were getting a correct answer (i.e. Question 2 and 6). The histogram did not show a distinctive peak for any multiple-choice question. The most considerable difference was the multiple-choice question Q1, where the difference between the number of readers who answered correctly ($n = 36$) and the number of non-readers who answered correctly ($n = 16$) was 20 participants. The Mann-Whitney U test was conducted to compare the difference in the number of participants between the readers and the non-readers who answered the multiple-choice questions correctly. There was a non-statistically significant difference in the number of participants who answered the multiple-choice questions correctly between the readers and non-readers group ($Z = -1.54$, $p = 0.123$).

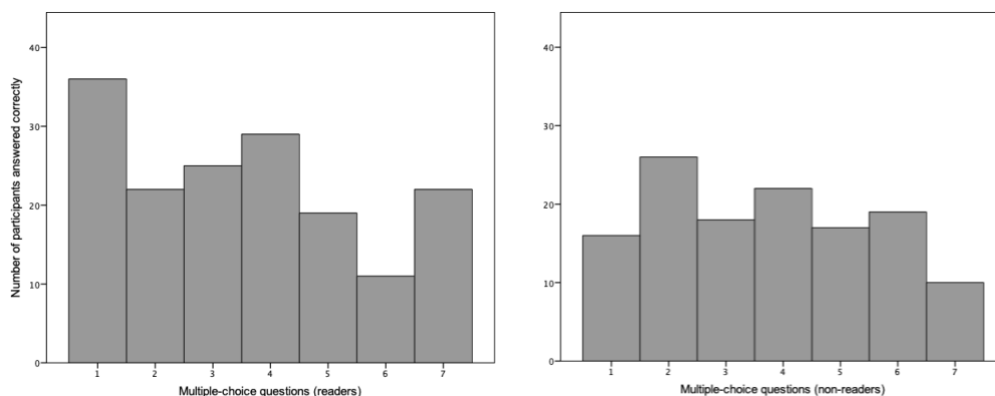


Figure 2.8 Distribution of the number of participants who answered correctly in each multiple-choice question

Using the proofreading method, the participants achieved a median proofreading score of 5 correctly identified errors out of the 7 typographical errors (IQR: 2.0). Figure 2.9 (Left histogram) shows the distribution of the scores for the proofreading. If using a 50% score as a cut-off value, there were 10 participants who scored less than 50% (i.e. did not manage to spot 4 or more spelling mistakes). To find out if all the proofreading errors were of equal difficulty, the number of participants who correctly identify the spelling errors were shown in Figure 2.9 (Right histogram). There were two pages with less than 30 participants who proofread correctly (i.e. Page 9, $n = 28$ and Page 29, $n = 23$). The large variation in the

proofreading scores suggests that it may not serve as an indicator if reading has occurred, and this large variation is unlikely due to the differences in the difficulty of the spelling error created. If a cut-off value cannot indicate whether the participants are reading, a possible observation would be that a “non-compliant reader” would complete the reading in a very short time. In other words, we would expect a negative correlation between reading speed and proofreading scores. However, reading speed measurements were not investigated in this group of participants.

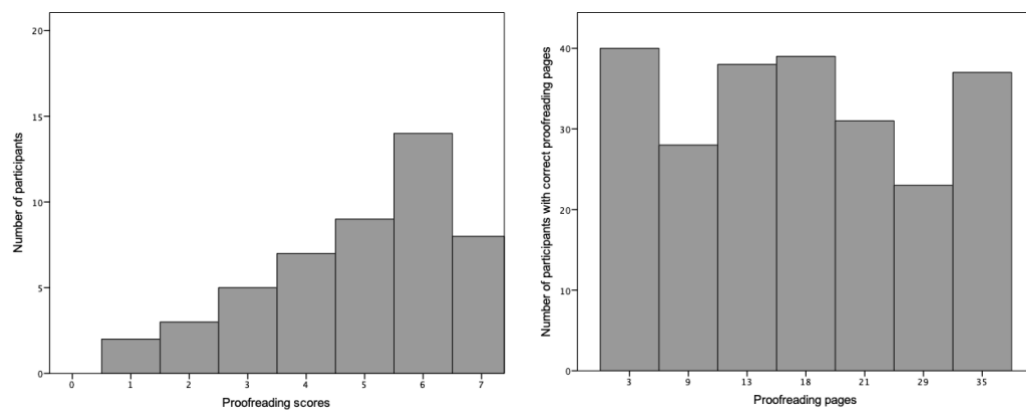


Figure 2.9 LEFT: Distribution of proofreading scores. RIGHT: Distribution of the number of participants who correctly identify the spelling error.

The proofreading errors were also subsequently incorporated in our silent reading test used in Chapter 3. These data allowed us to investigate the association between reading speed performance and proofreading scores. The data was presented here instead of Chapter 3 because it sets the context on the rationale for incorporating the proofreading errors in our reading text, and yet did not serve as a deciding factor to reject the participant’s data.

The same reading text, except the multiple-choice questions, was presented to a separate group of 29 participants from the University of Manchester. In addition to the proofreading procedures (i.e. to identify any spelling errors and out on a separate tray), the reading time was also recorded to compute the average reading speed. The median reading speed and proofreading scores were 233.09 wpm (IQR: 59.79 wpm) and 100% (IQR: 14.29%), respectively. Spearman’ rho correlation showed no correlation between the reading speed and

the proofreading scores ($r = 0.039$, $p = 0.86$). Specifically, we would like to identify readers with extreme reading speed, and to determine the cut-off value for “extreme” reading speed would seem arbitrary. As such, participants with reading speed outside the 95% confidence interval (i.e. $\text{mean} \pm 1.96 \text{ SD}$) would be termed as “extreme” high reading speed. There were three readers with reading speed outside the 95% confidence interval range (JK01: 394.70 wpm; TT22: 389.08 wpm; AM26: 425.01 wpm), but all three of them also had the same high proofreading score of 85.71% (i.e. 6 out of 7 correctly identified proofreading errors).

2.3.5 Discussion

The current experiment used a convenience sample of 48 participants. Hence, a post hoc power analysis was conducted using the G*Power (Version 3.1.9.3). Independent two-tailed t-tests were used on a sample of 48 participants gave a power of 84% to detect a difference score of 0.61 in the multiple-choice questions scores between the readers and non-readers (see Appendix 9.10.3).

In a silent reading test, ensuring reading compliance has always been a challenging factor. Various measures such as eye movement tracking (Lovie-Kitchin et al., 2000), multiple-choice questions (Ramulu et al., 2013b), and improvised assessment modes (e.g. verbalising only the last word in the sentence) (Legge et al., 1989a) can be adopted during the reading test. The development of multiple-choice questions to ensure reading compliance requires extensive efforts to set questions that can only be answered by reading the passage. The purpose of this experiment was to explore the limitations of using multiple-choice questions as a reading compliance approach. If there were no cut-off score which clearly distinguished a reader from a non-reader, then perhaps the use of proofreading by creating deliberate spelling mistakes in the reading passage would potentially provide an alternative approach.

The multiple-choice questions were given to two groups of participants: a) participants who read the passage and b) participants who did not read the passage. The purpose of recruiting the participants to answer the multiple-choice questions without reading the passage was to find out if these questions could be easily answered by general knowledge, and if a cut-off

score can be generated to differentiate between readers and non-readers. Our results show that a median score of 3 out of 7 questions can be answered correctly without reading the passage compared to a median score of four out of seven correctly answered by those who read the passage. Although there was a statistical difference between the two groups of participants (i.e. readers vs non-readers), the score difference between the groups was only by one question. The poor AUROC value suggests the limitation of our multiple-choice questions scores to differentiate between readers and non-readers. The best balance between sensitivity and specificity of a cut-off score of 3.5 would only have 52% of the participants be correctly classified as readers, but 19% of the participants would still produce such a score without reading the text. If a scoring criterion of less than 50% (i.e. less than 4 out of 7 correctly answered questions) is used to differentiate between readers and non-readers was apply to our readers' group, this would mean 23 of the participant's data (i.e. ~48%) would not be usable due to poor reading compliance. Our results differ from Ramulu et al. (2013b), who reported a large difference in median score of 79% between the non-readers (median score: 13%, range 0 to 39%) and readers (median score: 92%, range 65 to 100%) in a set of 16 multiple-choice questions. The difference in results compared to our current experiment could possibly be explained by the different nature of the reading text. Our current experiment used an abstract of a science encyclopaedia. Although the science encyclopaedia would theoretically have less gender bias (e.g. a romantic fictional story may appeal more to the female sample group), it also creates a higher possibility of guessing the answer through general knowledge.

Concurrently, we have also explored the use of proofreading as a reading compliance method. The large variation of the proofreading scores suggests that the ability to identify the spelling errors does not necessarily indicate that reading has occurred. Since the large variation in proofreading scores cannot provide us with an indicator if reading has occurred, a possible observation would be that the participants with poor proofreading skills may have significantly higher reading speed. This is because the participants would be merely flipping pages rather than reading the text. However, no correlation between reading speed and proofreading scores was observed. Using a 95% confidence interval, three "outliers" were identified, but all

of them correctly identified 6 out of 7 proofreading errors. It could either be that these participants were fast readers, or they were purely skimming for spelling errors.

The use of these reading compliance measures is somewhat artificial. Multiple-choice questions rely on one's memory ability, and proofreading relies on the reader's ability to pick up errors – and the experimental results suggest that these skills do not directly indicate that reading has occurred (e.g. poor scores in the multiple-choice questions or proofreading probably do not equate to not reading the text). However, by using any compliance measure, the approach sends the message to the participants that their reading is being monitored and thus will make them more likely to comply.

As such, the use of such reading compliance tests should serve as a reference rather than a definitive score to determine if reading has taken place. In the previous study by Ramulu et al. (2013b), the highest non-readers score of up to 34% on the multiple-choice questions was reported, but the authors only excluded participants with less than 50% score. In summary, neither of the two techniques for checking compliance provides a clear cut-off score which can be used to exclude a participant's data from the analysis. As proofreading is more straightforward and easier to implement, it will be used in our subsequent experiments.

2.3.6 Limitations

There were a few problems in the current experiment that can be improved. As discussed earlier, an ideal set of multiple-choice questions should not be answered by logical thinking. Although the initial idea was to create a "neutral" genre of reading text for both genders, this has also encouraged questions to be easily answered through prior science knowledge. The multiple-choice questions used in the current experiment were created by switching out words used in the original reading sentences, and the missing word will be used to create different multiple-choice options. As such, these questions may not be appropriate to check for comprehension as it is possible that readers who skim through the text could still identify the correct answer through flashback memory of the "familiar" words. A more appropriate and robust approach to elicit comprehension can be adopted to design the multiple-choice

questions, as suggested by Dickinson and Rabbitt (1991). The use of structured design multiple-choice questions has only been used in studies reporting comprehension assessment in participants simulated visual impairment. Given a more stringent type of multiple-choice questions, it is possible that a portion of the readers (e.g. skimming through the reading text) might not score well, albeit reading has occurred. Depending on the nature of the experiment, the researcher would then have to decide if these group of “readers” should be included in the study.

2.4 Conclusion

In the context of this thesis, the main objective of this study is to investigate any duration related effects using extended reading texts. We have explored some of the method limitations which will be used in our subsequent experiments. In the first section, the difference in reading speed measurements with the inclusion of the pre-verbalisation time was explored. Since the use of audio recording will be adopted in the subsequent experiments, the results of this experiment serve to explore the limitations of including the pre-verbalisation time when computing reading speed. The findings suggest that the inclusion of the pre-verbalisation time to compute reading speed was not significantly different from using only the audio recordings (i.e. verbalised text only). As such, any reported effects observed using the audio recordings would be equally clinically evident when using a stopwatch. However, this generalisation was only limited to habitual reading (i.e. without the use of a LVA). In the second section, the use of the page-turn method for time recording was explored. As page-turn will be used in the subsequent silent reading test to determine the reading time, it is important to know the limitation of this novel time recording method. The results showed that while some learning effects were seen in the initial page of the silent reading test, the effects dissipated on subsequent reading conditions. The findings suggest that prior familiarisation of the test to the participants would help to minimise the learning effect, which otherwise could produce inaccurate data. Lastly, as silent reading would be adopted as part of the assessment on duration related effect, a key challenge would be to ensure reading compliance for the reading tests. The preliminary findings on the use of multiple-choice questions or proofreading to ensure reading compliance did not suggest that either of the two approaches was a sensitive

method. Although the findings remain inconclusive, the use of proofreading was used in the subsequent experiments to reinforce participants' compliance when performing the silent reading task. However, it is not appropriate to use a specific score result as a threshold to reject any of the participants' data.

3. Evaluation of reading endurance in young adults under conditions of simulated visual impairment

3.1 Introduction

Reading endurance, a term coined by Rubin (2013), refers to the ability to sustain reading speed over time. The clinical assessment in this area is important as it may provide information on the ability of individuals with visual impairment to undertake leisure reading tasks. In people using LVAs for reading tasks, the duration that the person can read with their magnifiers could be highly relevant to how useful the magnifiers are, particularly for leisure reading. One of the limiting factors may be reduced reading endurance, which could be a primary factor or secondary to other causes such as visually challenging conditions.

To date, there are limited studies investigating duration related effects on reading speed performance in participants with normal vision (Brussee et al., 2016) or visual impairment (Ramulu et al., 2013a, Stangler-Zuschrott, 1990). As early as 1990, Stangler-Zuschrott (1990) investigated the reading speed performances in a group of 20 participants with pre-existing eye conditions. In a short testing duration of approximately two minutes, he found reading speed decreased in the participants with maculopathy. In a recent study on participants with visual impairment, Ramulu et al. (2013a) observed a marginal decrease (i.e. – 0.37 wpm/ min) in reading speed in a group of elderly glaucoma participants using an extended reading assessment duration of 30 minutes. In contrast to the previous two studies, Brussee et al. (2016) did not observe duration related effects in a group of normally-sighted participants using an average testing duration of 20 minutes.

There is some evidence to suggest that patients with visual impairment are susceptible to both mental and physical fatigue. A recent survey by Schakel et al. (2017) reported that 15 out of 16 participants perceived fatigue when involving “high cognitive load” such as memorising information and processing information. Eleven out of the 16 participants also reported fatigue when making an effort to “get a sharp image”. This perceived fatigue could have implications for a visually impaired individual attempting to perform a leisure reading task. So far, no study

has been carried out to investigate duration related effects on reading speed in patients using magnifiers for reading. Any effect could be clinically meaningful because it would be highly relevant for the duration of the LVAs usage, that is, how useful the LVAs are in supporting leisure reading. Also, if there are observed duration related changes in reading speed assessment, then one might question how appropriate are the commercially available short reading texts (e.g. Bailey-Lovie word-reading chart (Bailey and Lovie, 1980), MNREAD acuity chart (Mansfield et al., 1994), etc.) which are used in assessing reading speed performance. In this chapter, the experiment was conducted in a group of young and normally-sighted individuals as this group of participants have minimal confounding factors (e.g. age, physical limitations, etc.). By using simulated visual impairment, its extent can be carefully controlled. The primary objective of the experiment reported in this chapter was to investigate any duration related changes in reading speed under conditions of simulated visual impairment in young, and normally-sighted participants. It was hypothesised that the reading speed will decrease over time due to reading fatigue under visually challenging conditions and that these effects might be sufficiently strong to be measured during a relatively short reading test. The secondary objective was to find out if the use of a commercially available reading test such as IReST and RoR, would be able to capture and replicate the same duration related effects observed in a more natural reading environment.

3.2 Methods

3.2.1 Participants

Thirty participants with a mean age of 21.77 (\pm 2.71) years participated in this study. Whittaker and Lovie-Kitchin (1993) suggested that the reading ability of an individual must be a minimum of three levels above that of the materials used for the reading test to avoid a limitation of reading speed due to the technical difficulty of the text. As the study required participants to read out long passages of text, they were either native English speakers or had English as their first language. Hence, the participants were recruited from the university. All of them had a near VA of at least 0.0 log MAR with their current spectacles using the Bailey-Lovie word-reading chart (Bailey and Lovie, 1980). The Mars letter contrast sensitivity test (Arditi, 2005) was used to assess the contrast sensitivity. The test was conducted at 50 cm, and individuals

read letters of reducing contrast until two consecutive mistakes were made. Any misreading of 'C' for 'O' and vice versa was scored as correct, and 5 to 10 seconds was given for each letter to be visible due to temporal summation (Elliott et al., 1990). All participants had normal contrast sensitivity of at least 1.62 log units (Haymes et al., 2006). Also, it was confirmed that no participants had a self-reported history of dyslexia, neurological or physical problems that may limit eye-hand coordination as the study involved the use of a hand-held magnifier.

3.2.2 Reading texts

Four different types of reading texts: IReST, RoR, magazine text (MT), and sustained reading text (SRT) were used in this study. Examples of the MT and SRT reading tests can be found in Appendix 9.1.

3.2.2.1 International Reading Speed Text (IReST)

The IReST (Trauzettel-Klosinski et al., 2012) is a standardised extended text (i.e. approximately 150 words) used to measure reading speed. In this experiment, two IReST paragraphs were combined using the original format (i.e. column width, typeface and font size) reproduced on an A4 single page. Three versions were produced; Version 1 (original IReST paragraphs 1 and 2: 317 words), Version 2 (original IReST paragraphs 3 and 4: 320 words) and Version 3 (original IReST paragraphs 5 and 6: 291 words). Although the previous experiments (Trauzettel-Klosinski et al., 2012, Hahn et al., 2006) evaluating IReST employed the use of maximum oral reading speed (i.e. participants were instructed to read as fast and as accurately as possible), the current experiment did not attempt to force the participants to work at their fastest reading speed. Instead, participants were instructed to read as they would do in their daily reading materials such as magazine or newspaper.

3.2.2.2 Rate of Reading (RoR)

The RoR test (Wilkins et al., 1996) was developed in 1996 to investigate the effects of coloured overlays on reading disability. Unlike most reading tests (e.g. IReST, MNREAD acuity chart, etc.) which use contextual reading materials, the RoR uses 15 random words commonly found in children's reading books. Using contextual reading materials is laborious as the inventor

needs to create a standardised level of difficulty throughout the reading sentences or paragraphs. Also, most reading tests have different versions to allow multiple testing on the same person without the influence of learning effects. As such, the inventors have to ensure that these different versions of the same test are of an equivalent level of reading difficulty. The benefit of using random words is that it allows the reading texts to be reproduced easily in many different versions. If similar effects which are documented using contextual text can be replicated using random words text, then this approach would be more straightforward and easier to reproduce in different equivalent versions. The original RoR paragraph consists of 10 lines each of 15 random words. Likewise, as with the replicated IReST, two paragraphs of RoR were combined and printed on a single A4 page. As advocated by Wilkins et al. (1996), participants were instructed to read the paragraph of words out-loud as fast as possible. The different versions were generated by randomising the line sequence.

3.2.2.3 Magazine text (MT)

Although the developers of reading speed tests have attempted to create a standardised reading difficulty in different versions of their reading tests, the consistency of the reading difficulty in various reading texts versions remains questionable. If it is impossible to create a standardised reading text, then the use of daily reading materials such as the newspaper or magazines would seem to be a plausible approach to use as a reading test. General reading materials from the magazine "New Scientist" were used to represent actual reading material. The texts were printed on 11 A4 papers; each page was 100-words in length and approximately 50-words on the 11th (final) page. The page-turn method, as reported in Chapter 2, was used to document the reading time. The 11th page served to allow the first ten page-turns to be observed and analysed. Similarly to IReST, participants were instructed to read the text out-loud as they would do in their daily reading materials such as a magazine or newspaper.

3.2.2.4 Sustained Reading Text (SRT)

The sustained reading text (SRT) consisted of approximately 3500 words abstracted from a science encyclopaedia. Analogous to the MT, there were 100 words on each page, and the

last page of a random number of words was to ensure subjects continued to flip the page after they finished reading page 35. The SRT represents a realistic reading setting (i.e. silent leisure reading over an approximate period of 15 minutes), which allows us to investigate any duration related changes in this reading process. The participants were instructed to complete reading all the pages or were stopped after 15 minutes, whichever came first. As this was a silent reading test, reading speed was calculated based on the page-turn to indicate the reading duration for that corresponding page. Also, participants were asked to spot deliberate typographical errors (i.e. "seen" spelt as "nees") in the text. There were seven typographical errors (one error in each page) out of the 35 pages. Upon detecting the error, they were expected to continue reading until the last word of the page, after which this page was then placed in a separate tray.

All the reading tests were reproduced in three different versions to ensure no learning effect across the three conditions (i.e. normal; simulated for reduced contrast and VA, and the use of a hand-held magnifier). For SRT, validation was carried out to ensure a similar difficulty level of the reading materials for all the three versions (see Appendix 9.1). Instructions were given to the participants to read the text as they would normally read a paperback novel or magazine except for the RoR. In the case of RoR, no comprehension was needed as the text is made up of random words, so participants were asked to read the words as fast and as accurately as possible. Flesch-Kincaid Grade Level (Kincaid et al., 1975) is a formula used to test the readability of a reading text. This formula calculates the average sentence length and an average number of syllables per word to derive a reading grade level (i.e. a score of 8.0 means the passage can be comprehended by a United States 8th Grade student, equivalent to UK year 8). As the MT and SRT were created specifically for this study, the Flesch-Kincaid formula was applied to the two reading texts. The MT and the SRT had Flesch-Kincaid grades ranging from 8.8 to 10.2, which were well within our participants' reading ability (i.e. all the participants were from the university, which achieved at least four levels above the Flesch-Kincaid grades). All the texts were printed in Times New Roman 9.0 point size (i.e. equivalent to a N8 newspaper print size), which has an approximate x-height of 1.5 mm, close to that of

the text print in magazines (i.e. N8 print size with a x-height of 1.4mm) (Legge and Bigelow, 2011, DeMarco and Massof, 1997) and the IReST.

3.2.3 Procedures

All the reading tasks were conducted in three different conditions. In the first condition, termed as the normal condition (NOR), participants used their optimal spectacle correction. In the second condition, termed as the Bangerter occlusion foil condition (BOF), participants were then simulated with a Bangerter occlusion foil to give a degraded image (i.e. reduced contrast and VA). A range of different intensities of the Bangerter occlusion foil was tried with each participant, and the one which achieved a threshold near acuity of 0.3 log MAR was chosen. Near VA and contrast sensitivity were measured binocularly since the reading tasks were done binocularly. Using the formula (Acuity Reserve = Print Size/ Threshold RA) by Whittaker and Lovie-Kitchin (1993, 1999), a minimum acuity reserve of 2:1 was achieved with a near threshold acuity of 0.3 log MAR (N4 at 25 cm) on our reading materials (N8 at 25cm). This acuity reserve would be expected to support a high fluent reading speed. The participant's contrast threshold can be calculated (Mantjarvi and Laitinen, 2001) from the Mars letter contrast sensitivity.

For example, the subject scored 1.76 log units on Mars chart.

$$1/10^{1.76} = 0.0174 \text{ (1.74\% contrast threshold)}$$

The text samples were printed using Hewlett Packard Laser Colour Printer, suggestive of an approximate contrast of 90% (Rumney, 1995, Whittaker and Lovie-Kitchin, 1994). To calculate the contrast reserve ratio, the contrast of the sample text (i.e. 90%) was divided by the contrast threshold (i.e. 1.74%), which gave a ratio of 51.79:1. Different Bangerter foils were used for each participant in an attempt to achieve a standardised (and just sufficient) acuity reserve among the participants. However, this will result in a variably reduced contrast threshold, and therefore, a variable contrast reserve between individuals was expected.

In the last condition, termed as the Bangerter occlusion foil with magnifier condition (MAG), a higher density of Bangerter occlusion foil was used to further degrade the near VA to approximately 1.0 to 0.9 log MAR. Figure 3.1 compares the view of the text through the two BOF used. The extent of this simulation forces the participants to rely on a hand-held magnifier to perform the reading task. To achieve fluent reading at an N8 print size at 25 cm (with an acuity reserve of 2:1), a 0.3 log MAR threshold acuity would be required. A 4x hand-held magnifier was therefore issued to the participant to perform the reading task when viewing through the Bangerter occlusion foil. Before the experiment, instructions on how to use the magnifier and the time to allow participants to familiarise themselves with the utilisation of a hand-held magnifier were given. Participants were instructed to keep the distance between the magnifier lens and their fogging lens as close as possible to optimise the field of view.

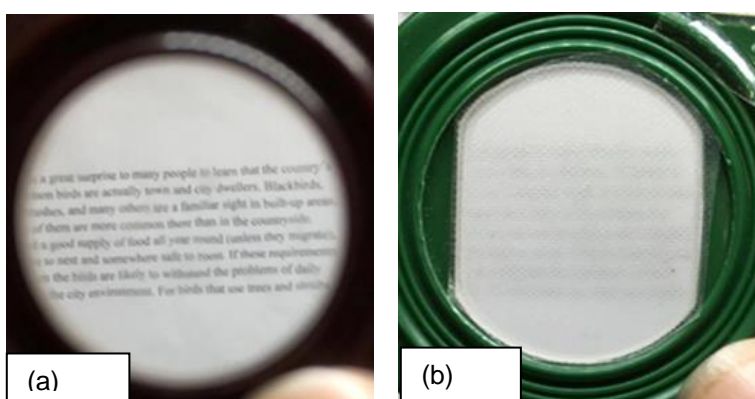


Figure 3.1 (a) text viewing in BOF (b) text viewing with BOF used in MAG

All the reading materials were positioned on a reading stand. The participants used a standard viewing distance of 25 cm except when using the hand-held magnifier, where they were instructed to keep the distance from their eye to the magnifier as close as possible. The sequence of reading tests was randomised to minimise the systematic influence of practice from order effects. However, the same sequence of conditions (i.e. NOR, BOF and MAG) was conducted for all participants. The reading out-loud tests (i.e. IReST, RoR and MTs) were recorded using a voice recorder (Olympus Digital Voice Recorder VN-5500PC). The time taken to verbalise the words was analysed by WavePad Sound Editor (Version 6.64 by NCH software) to calculate the reading speed. The time for each page-turn was recorded by an

online stopwatch lap timer (<http://www.timeanddate.com/stopwatch/>) for the MT and the SRT. The study was approved by the University of Manchester Research Ethics Committee (Reference no: 08046) and written informed consent was obtained from all participants.

3.2.4 Data analysis

3.2.4.1 Calculation of average reading speed

For the IReST and RoR, the average reading speeds in wpm were calculated by using the number of words read correctly divided by the time taken in minutes on each page. For the MT and SRT, average reading speed was calculated using the total number of words read through the pages divided by the total time taken in minutes. While there is no definitive way to examine duration related effects in a reading passage, four approaches were devised for the analysis of the gathered data. The type of approach depends on the information available; therefore, some analyses were not applicable to all the reading tests.

3.2.4.2 Calculation of reading slope in MT and SRT

The use of reading slope as a measure of duration related effects in reading performance was first reported by Ramulu et al. (2013a). The authors investigated the reading speed in successive minutes using an extended reading passage (approximately 7000 words). Like the Ramulu et al. (2013a) study, the current study adopted the use of reading slope to determine any changes in reading speed with time. Figure 3.2 shows an illustration of how the reading speeds for successive minutes were calculated. Each page-turn (indicated as the blue turning arrow) was used to document the time taken to read each page. To ensure consistency, the start of the flipping page was used as an indicator of the end time for that particular page, and the start time for the next page. The reading speed on that corresponding page was then calculated using the total number of words (i.e. 100 words) divided by the time taken in minutes between each page-turn. In the example shown in Figure 3.2, the end of the first minute of the reading test occurred on the second page. The reading speed calculated from the second page (i.e. 213.83 wpm) would then be used for the reading speed in the first minute, and so on until the last minute. By calculating the slope of the best straight line from these data points, the gradient can be used to represent any increase (i.e. positive gradient) or decrease (i.e.

negative gradient) in reading speed during the reading duration. However, the use of such analysis required an extended reading time to provide a meaningful number of values to be used in the calculation. Hence, this analysis was only used in the MT and SRT.

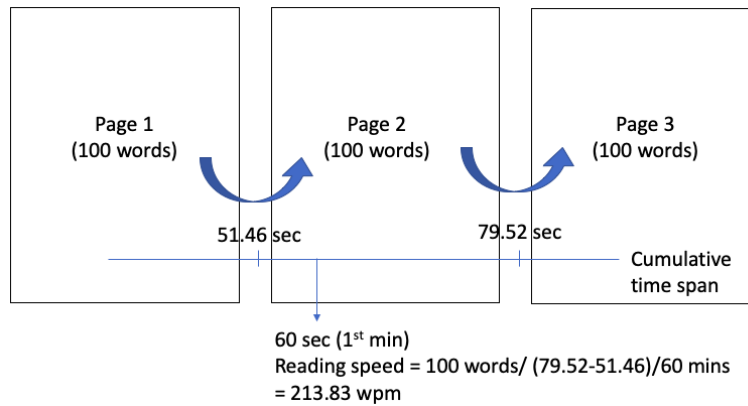


Figure 3.2 A diagram to illustrate the calculation of reading speed at each 1 minute timepoint

3.2.4.3 Calculation of normalised reading rate (NRR1)

To date, only a linear relationship was used by Ramulu et al. (2013a) as a measure of a change in reading speed with time. A similar linear analysis (i.e. linear mixed model) has also been reported by Brussee et al. (2016). In this study, we explored a new approach to report changes in reading speed with time using different computing methods based on normalised reading rate. Normalised reading rate (NRR1) compares the reading speed in each successive minute to the baseline reading speed (i.e. in the second minute). The second minute was used instead of the first minute to reduce any possible timing inaccuracy when the first page was revealed. The NRR1 was calculated by dividing the reading speed in every subsequent minute (3rd, 4th, ..., etc.) by the reading speed in the second minute. A median NRR1 was used to calculate each participant's overall NRR1 instead of the average value to minimise the effect of outliers. A NRR1 value of 1.0 would indicate no overall change in reading speed with time; a value of more than 1.0 would indicate an increased reading speed, and a value of less than 1.0 would indicate a decreased reading speed.

3.2.4.4 Calculation of normalised reading rate (NRR2)

Normalised reading rate (NRR2) compared the reading speed at the end relative to the beginning of the text paragraph. Two-time-points at the start and end of the test were selected

in IReST and RoR, each of which consisted of two short phrases made up of approximately 180 characters (inclusive of spacing) as seen in Figure 3.3. In the case of IReST, the phrase in each case consisted of four lines and involved three return sweep eye movements (moving from the end of the line to the beginning of the subsequent line). By using a standard word length (SWL) of six characters (Carver, 1992), 180 characters are equivalent to 30 standard words. Similarly, in RoR, three lines at the beginning (3rd to 5th line) and towards the end of the test (16th to 18th line) were selected for the two time-points. To achieve a similar number of characters as the IReST, three lines were chosen as the total characters ranged from 174 to 177, approximating to 180 characters. However, in this case, only two return sweeps were involved. The NRR2 was then calculated by using the second time-point reading divided by the first time-point reading speed. A NRR2 value of 1.0 would thereby indicate no change in reading speed for the second segment of the reading paragraph; a value of more than 1.0 would indicate an increase in reading speed and a value of less than 1.0 would indicate a decrease in reading speed.

Similarly, reading speeds at two time-points were also determined for both MT and SRT. However, instead of using sections of the text to denote the time-points, time intervals were used. As the MT is a shorter reading text compared to SRT, the 2nd and the 5th minute were chosen as the time-points for all conditions. In SRT, the selected time intervals were 2nd and 10th minute in NOR and BOF, and 2nd and 14th minute in MAG. NRR2 was then calculated by using the second time-point reading speed divided by the first time-point reading speed. The aim was to maximise the time difference between the time intervals chosen under SRT to try to demonstrate a maximum effect.

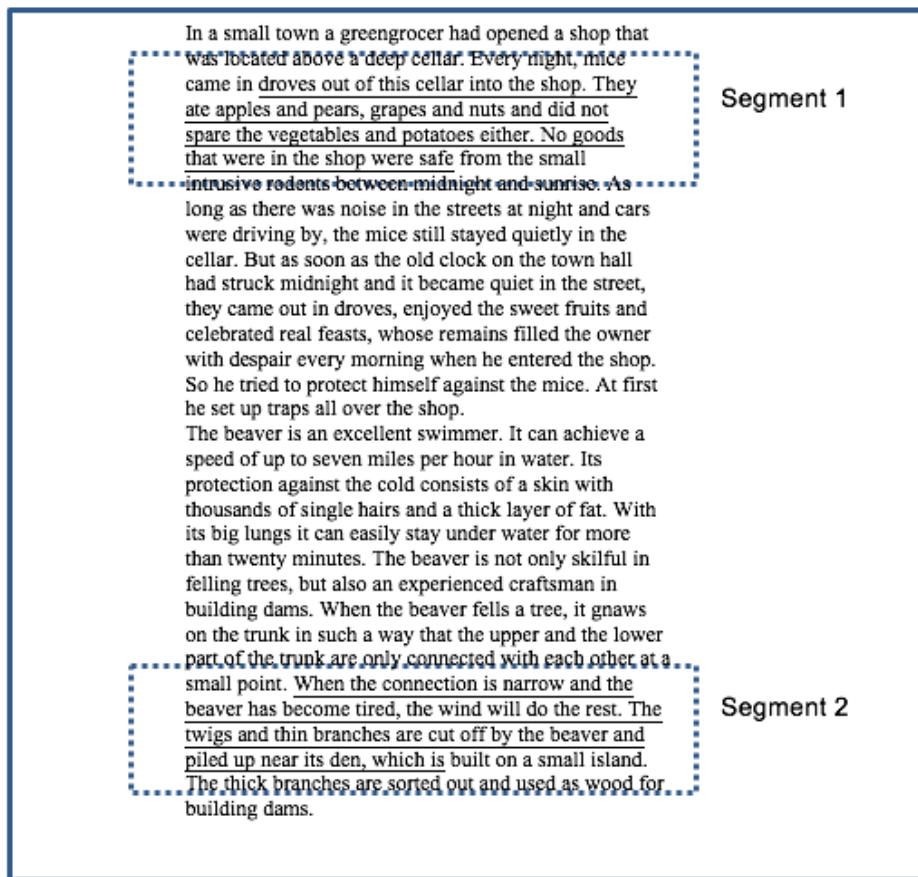


Figure 3.3 An example of IReST text showing the selected sections for NRR2 analysis.

3.3 Statistical analysis

All data were analysed using Statistical Package for the Social Sciences (SPSS) (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.). Shapiro-Wilk test showed a normal distribution of the reading speeds in IReST ($p = 0.31$), RoR ($p = 0.43$) and MT ($p = 0.86$), but not in the SRT ($p = 0.004$). \log_{10} transformation of SRT data ($p = 0.003$) did not improve the distribution. Therefore, the median and interquartile range measurements were reported. Non-parametric statistical analyses (Friedman and post hoc Wilcoxon Signed-Rank tests) were used to explore the differences in reading speed among the various reading tests and reading conditions. Spearman correlation was used to investigate the association between reading slopes and normalised reading rates among the different reading tests. A probability of less than 0.05 indicated statistical significance.

3.4 Results

The data from 26 participants were used for analysis. The data for four participants were rejected; three participants had an extremely high reading speed (i.e. 400 wpm), which resulted in insufficient reading time (i.e. less than 8 minutes) for analysis, and there was a procedural error for one participant. The median near VA and contrast sensitivity were 0.0 log MAR and 1.76 log units, respectively. Out of the 26 participants, four were myopes (spherical equivalent range: -0.50 to -4.00 DS), one hyperope (spherical equivalent: $+7.00$ DS) and the rest were emmetropes. As mentioned in the data analysis Section 3.2.3, the acuity reserve was controlled at 6:1 in NOR, and 2:1 in BOF and MAG (i.e. with the help of the hand-held magnifier). The resulting median contrast reserves were 51.79 (IQR: 0.0, range: 47.23 to 51.79):1, 35.83 (IQR: 3.46, range: 24.79 to 47.23):1, and 5.66 (IQR: 3.49, range: 2.72 to 8.21):1 in NOR, BOF and MAG, respectively. The latter is, therefore, below the limit suggested for high fluent reading (10:1) by Whittaker and Lovie-Kitchin (1993). The overall median reading speeds and interquartile ranges were shown in Table 3.1. The SRT produced the fastest reading speed, followed by IReST, RoR and MT, in NOR and BOF. The order was slightly different in MAG: RoR had the fastest reading speed, followed by IReST, SRT and MT. Friedman test showed a statistically significant difference in reading speed between the reading tests in NOR ($\chi^2(3, 26) = 51.46, p < 0.001$), BOF ($\chi^2(3, 26) = 42.69, p < 0.001$) and MAG ($\chi^2(3, 26) = 25.38, p < 0.001$). Within the same reading test, Friedman test showed a statistically significant difference in reading speed for IReST ($\chi^2(2, 26) = 39.69, p < 0.001$), RoR ($\chi^2(2, 26) = 36.00, p < 0.001$), MT ($\chi^2(2, 26) = 39.69, p < 0.001$) and SRT ($\chi^2(2, 26) = 39.07, p < 0.001$) between the three conditions. Post hoc test using the Wilcoxon Signed-Rank test showed that only the reading speed in MAG ($p < 0.001$) was statistically significantly different from NOR and BOF in all the reading tests.

Table 3.1 Median reading speed for all reading tests

Reading tests	Median reading speed, wpm		
	NOR	BOF	MAG
IReST	190.89 (24.47) Range: 158.38 to 236.19	193.33 (45.43) Range: 124.50 to 225.83	125.09 (33.98) Range: 90.49 to 176.73
MT	156.55 (19.47) Range: 123.27 to 195.78	153.31 (25.03) Range: 113.81 to 192.47	106.97 (23.78) Range: 61.22 to 143.85
RoR	174.91 (41.46) Range: 129.58 to 154.99	170.76 (38.83) Range: 110.75 to 232.95	130.80 (32.33) Range: 95.93 to 210.45
SRT	227.68 (90.62) Range: 154.99 to 283.66	226.11 (81.15) Range: 130.11 to 343.12	111.06 (34.56) Range: 73.77 to 185.68

Interquartile range (IQR) values are shown in the parentheses

To find out an overall change in reading speed with time, the group median reading speeds were computed at each minute time point. The reading speed was plotted against time duration for the reading task. The duration of the reading test ranged from a 5-minute out-loud reading task (i.e. MT), a 10-minute (i.e. NOR and BOF) and a 15-minute (i.e. MAG) silent reading task (i.e. SRT). The difference in reading duration for NOR and BOF compared to the MAG in the SRT was because most of the participants finished the reading tests before 15 minutes in the NOR and BOF. A linear regression function was then used to determine the gradient slope, also termed as the reading slope. Figure 3.4 shows the graph plots and how the group median reading slope was computed. The group median reading slopes for MT were 2.85 wpm/ min (R^2 : 0.84), 3.35 wpm/ min (R^2 : 0.84) and 1.87 wpm/ min (R^2 : 0.41) in NOR, BOF and MAG, respectively; the group median reading slopes for SRT were 5.55 wpm/ min (R^2 : 0.70), 2.90 wpm/ min (R^2 : 0.39) and 0.68 wpm/ min (R^2 : 0.68) in NOR, BOF and MAG, respectively. Both the MT and SRT group median reading slopes were positive, suggesting that there was no overall reduction in reading speed with time; in fact, the reading speed appeared to increase with time. However, this “increment” was reduced in the MAG compared to the NOR and BOF conditions in both the MT and SRT reading tests.

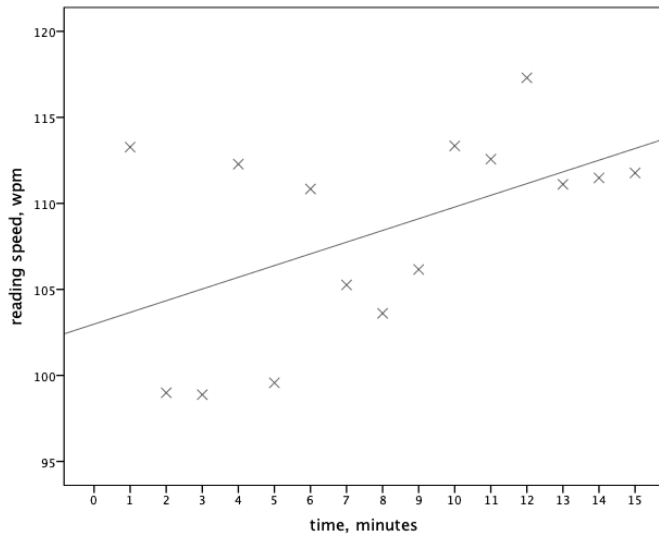


Figure 3.4 An example of the graph plot of group median reading speed against time in successive minutes (i.e. SRT MAG). A linear regression function using the SPSS program computed the group median reading slope.

3.4.1 Calculation of reading slope in MT and SRT

The reading slope for each individual participant was computed using the linear regression function in the SPSS program for both the MT and SRT reading tests. Figure 3.5 shows an example of how the reading slope plots for two participants; one with a negative reading slope (LEFT) and another with a positive reading slope (RIGHT) in the SRT.

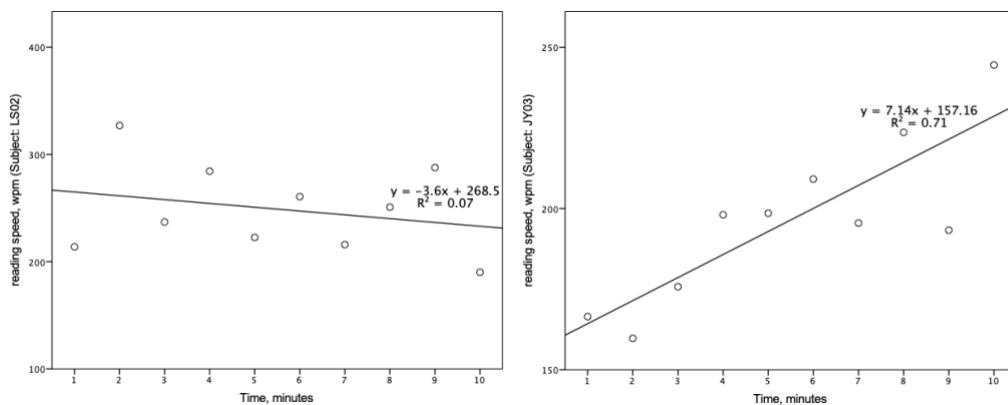


Figure 3.5 Examples of how the reading slope is computed. LEFT: An example of a participant LS02 with gradual reduction in reading speed with time, and this was reflected by a negative gradient slope (-3.6 wpm/min). RIGHT: An example of a participant JY03 with gradual increase in reading speed with time, and this was reflected by a positive gradient slope (7.14 wpm/min).

The median slopes of all the individual participant's reading slopes for both MT and SRT were shown in Table 3.2. An overall median positive reading slope gradient indicated an increase in reading speed with time for MT and SRT in all reading conditions. However, this increment was markedly reduced in the MAG condition. Both the MT and SRT showed that the reading slopes were lowest in MAG but still maintained at a positive gradient slope. The Friedman test indicated that there was a statistically significant difference in reading slopes between the three different conditions in SRT ($\chi^2(2, 26) = 8.54, p = 0.01$) but not in MT ($\chi^2(2, 26) = 2.85, p = 0.24$). Post hoc Wilcoxon Signed-Rank test showed only a statistically significant difference in the reading slopes between NOR and MAG ($z = -3.14, p = 0.002$). Wilcoxon Signed-Rank test showed no statistically significant difference in reading slopes between MT and SRT in NOR ($z = -1.11, p = 0.27$), BOF ($z = -1.12, p = 0.26$) and MAG ($z = -0.66, p = 0.51$) conditions.

Table 3.2 Descriptive statistics of the reading slopes for MT and SRT (n = 26)

	median reading slope (wpm/min)		
	NOR	BOF	MAG
MT	1.89 (6.65) Range: -12.85 to 10.42	2.44 (3.99) Range: -3.86 to 6.81	0.57 (4.66) Range: -6.55 to 11.31
SRT	4.46 (5.24) Range: -4.13 to 16.42	1.85 (5.34) Range: -6.77 to 11.43	0.37 (2.10) Range: -1.26 to 3.09

Interquartile range (IQR) values are shown in the parentheses

Therefore, the goodness of fit, R^2 value, was computed for the linear reading slopes in MT and SRT. Assuming we used a mid-point cut-off value for $R^2 = 0.5$, Table 3.3 shows that the overall median R^2 values in both MT and SRT were less than 0.50, with higher median R^2 values observed in MT. Wilcoxon Signed-Rank test showed a statistically significant overall difference in the R^2 values between MT and SRT ($z = -3.04, p = 0.002$). Out of the number of participants with R^2 value more than 0.5, only three participants had negative reading slopes in MT; none of the participants with R^2 value more than 0.5 in SRT had negative reading slopes.

Table 3.3 R² values for MT and SRT in all the conditions

MT	R ² of the reading slope		Number of participants with R ² ≥ 0.5		
	median	range	Positive reading slopes	Negative reading slopes	Total
NOR	0.24 (0.53)	0.01 to 0.74	7	1	8
BOF	0.20 (0.42)	0.00 to 0.94	6	0	6
MAG	0.24 (0.62)	0.00 to 0.86	7	2	9
SRT	median	range	Positive reading slopes	Negative reading slopes	Total
NOR	0.16 (0.34)	0.00 to 0.76	5	0	5
BOF	0.05 (0.23)	0.00 to 0.73	3	0	3
MAG	0.07 (0.18)	0.00 to 0.54	1	0	1

Interquartile range (IQR) values are shown in the parentheses

Given that most of the participants with negative reading slopes have low R² values, this suggested that using a linear function fitting might not be the optimum representation of the data trend. Since the SRT MAG had the lowest number of participants with R² values more than 0.5, it may be possible that the use of alternative curve fittings would be more appropriate, and allows an indication of any overall increase or decrease in the reading speed with time. Scatterplots of the individual participants were created and visually inspected; the purpose was to determine if there were specific trends on the reading speed changes with time. However, the visual inspection of the graph plots did not suggest a distinctive graph plot, although some participants did show an exponential function. Hence, curve fittings using the automated logarithmic and exponential functions in SPSS were applied to all the participants' results in SRT MAG (see Appendix 9.4 for the nonlinear curve fitting for all the participants). Figure 3.6 shows an example of the additional nonlinear curve fittings compared with the linear least square function slope. The attempt to categorise specific curve fittings was unsuccessful, given that the data points for each participant were largely variable (see Appendix 9.4). Table 3.4 shows the R² values for the linear, logarithmic and exponential function for the participants reading speed against time in MAG. Although the logarithmic function showed a slightly higher R² value, the Friedman test showed that there was no statistically significant ($\chi^2(2,26) = 0.98$, $p = 0.61$) difference in the R² values between the three function fitting graphs. Given the finding that the R² values did not suggest that alternative curve fittings such as logarithmic and exponential functions would provide a better representation of the data points, the subsequent experiment would only report the linear gradient slopes.

Table 3.4 R² values for linear, logarithmic and exponential function for participants reading speed against time in MAG (n = 26)

	Linear function	Logarithmic function	Exponential function
Median R ²	0.07 (0.17)	0.08 (0.17)	0.07 (0.21)
Range	0.00 to 0.54	0.00 to 0.58	0.00 to 0.55

Interquartile range (IQR) values are shown in the parentheses

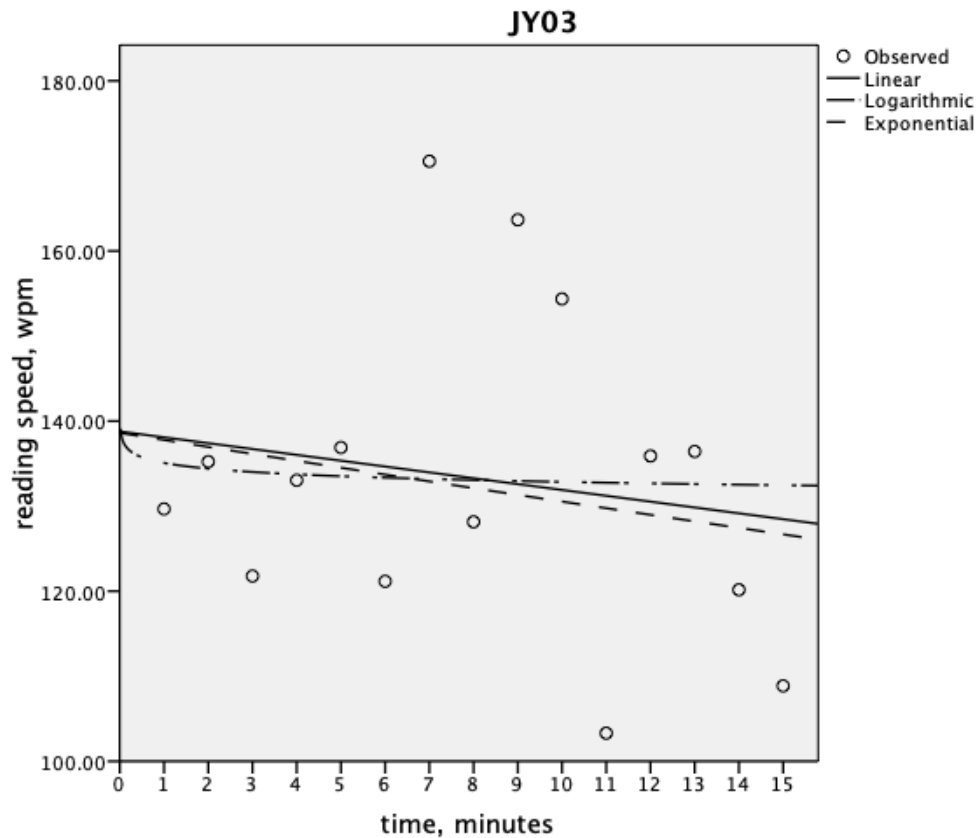


Figure 3.6 An example of nonlinear curve fitting graph plot of a participant in SRT MAG. The solid line: linear function, dashed line: exponential function, dashed and dotted line: logarithmic

So far, only the overall median linear gradient slope results have been reported, and the large interquartile range observed in Table 3.2 suggests that the individual participant's reading linear gradient slopes were highly variable. As such, the number of participants was split into those with positive and negative slopes. Table 3.5 shows the number of participants and the descriptive statistics of the group of participants showing positive or negative reading slopes. Most of the participants showed a positive reading slope in both the MT and SRT reading tests. The large range of the reading slopes was primarily due to the large variability of the data

points within the 10 to 15 minutes time frame. Interestingly, Spearman correlation did not show any correlation between the SRT and MT reading slopes ($r = 0.085$, $p = 0.46$) when combining all the conditions (i.e. NOR, BOF and MAG). There was only one participant (i.e. same participant: VV07) who showed consecutive negative reading slopes in both MT and SRT in NOR and BOF; four participants (i.e. JY03, JF03, LP25 and OB29) showed a consecutive negative reading slope in MT and SRT in the MAG. Figure 3.7 shows the graph plots of the participants with negative reading slopes in both the MT and SRT in the NOR, BOF and MAG. All the five participants were emmetropes, and all had a near VA and contrast sensitivity of 0.0 log MAR and 1.76 log units, respectively. A Mann-Whitney Signed-Rank test compared the habitual reading speed in SRT (i.e. NOR) and the contrast reserve in the MAG between the five participants and the remaining participants. There was no statistically significant difference in the reading speed ($z = -0.81$, $p = 0.42$) and contrast reserve ($z = -0.76$, $p = 0.45$) between the five participants and the rest of the 21 participants. Apart from the graph plots from VV07-NOR (Top Left) in Figure 3.7, which showed a gradual drop in reading speed in the SRT, most of the data points were highly variable in successive minutes. Despite the MT and SRT being very similar tests, the two reading tests produced very different reading speed changes with time.

Table 3.5 Number of participants showing positive and negative reading slopes in MT and SRT (n = 26)

NOR		MT	
		Positive (n = 19) Median: 0.70 (0.59) Range: 0.37 to 10.42	Negative (n = 7) Median: -7.32 (10.26) Range: -1.16 to -12.85
SRT	Positive (n = 21) Median: 2.13 (1.78) Range: 0.36 to 16.42	15	6
	Negative (n = 5) Median: -3.15 (1.51) Range: -1.37 to -4.13	4	1

BOF		MT	
		Positive (n = 22) Median: 0.64 (0.59) Range: 0.55 to 6.81	Negative (n = 4) Median: -1.72 (1.86) Range: -1.45 to -3.86
SRT	Positive (n = 17) Median: 0.89 (1.20) Range: 0.32 to 11.43	14	3
	Negative (n = 9) Median: -3.28 (4.36) Range: -0.35 to -6.77	8	1

MAG		MT	
		Positive (n = 15) Median: 0.57 (1.01) Range: 0.13 to 11.31	Negative (n = 11) Median: -4.78 (2.76) Range: -0.02 to -6.55
SRT	Positive (n = 16) Median: 0.18 (0.38) Range: 0.02 to 3.09	9	7
	Negative (n = 10) Median: -1.02 (0.35) Range: -0.01 to -1.26	6	4

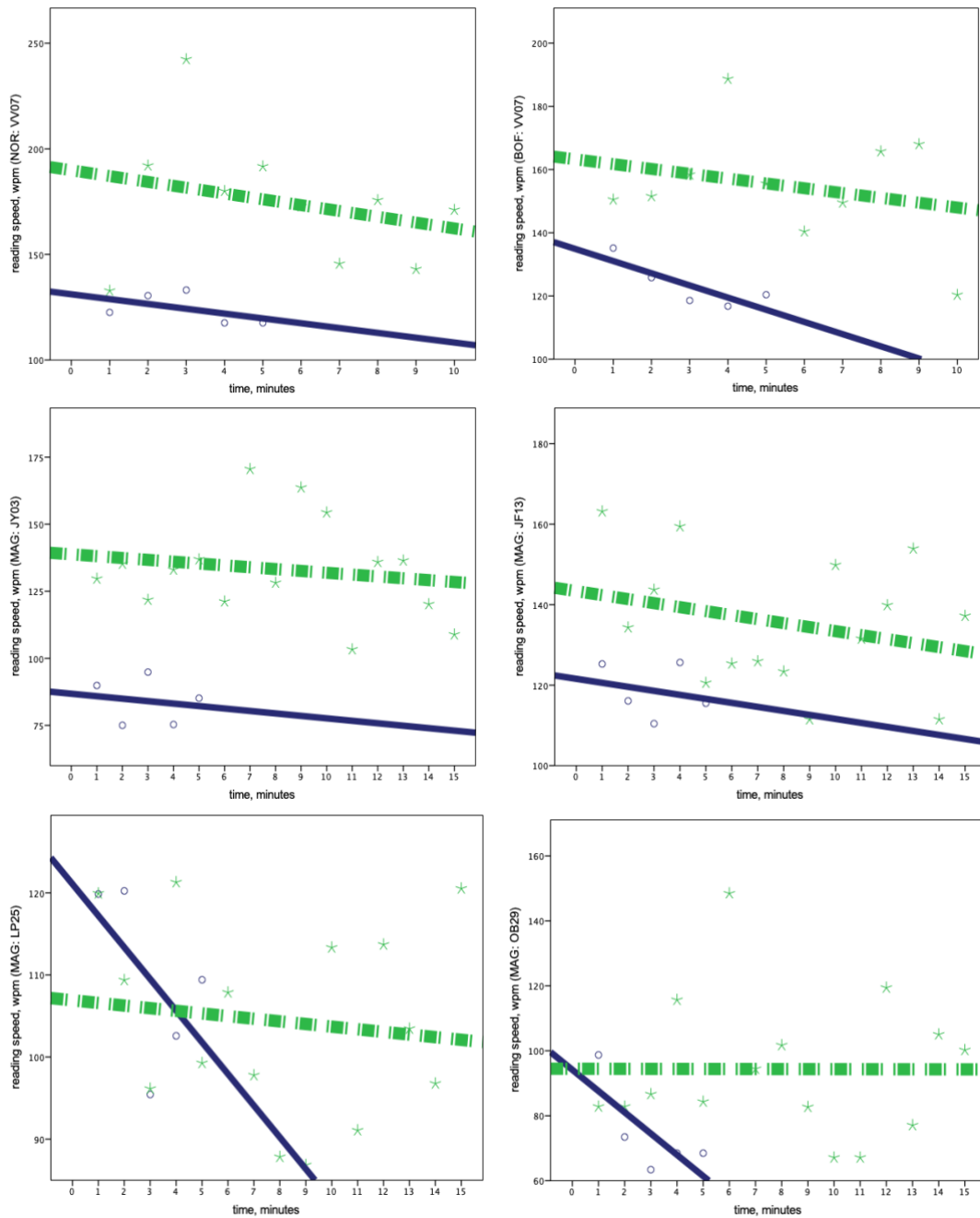


Figure 3.7 Graph plots on reading speed against time for participants with negative reading slopes in both MT and SRT. ○: MT ☆: SRT

Top left: NOR VV07, Top Right BOF, VV07; Middle left: MAG, JY03, Middle Right JF13; Bottom Left MAG, LP25; Bottom Right: MAG, OB29 (the OB29 SRT slope was -0.01 wpm/ min, which explains the almost zero gradient slope)

3.4.2 Calculation of Normalised Reading Rate (NRR1)

In the second analysis, the median NRR1 was calculated for all participants. In MT NRR1, the time span was plotted up to 5 minutes. Table 3.6 shows the median and range of NRR1 in for both MT and SRT in all the conditions. All the median NRR1 had a minimum value of more than 1.0, indicating no decrease in reading speed within the testing duration (i.e. 5 minutes, or 10 to 15 minutes). Table 3.6 also shows the breakdown of the number of participants with a NRR1 value of less than 1.0 in MT and SRT in the respective conditions. The participants showing a NRR1 value of less than 1.0 in MT in all conditions ranged from 0.80 to 0.99, indicating that the reading speed reduction ranged from 1% to 20%. The participants with NRR1 values less than 1.0 in SRT showed a larger range (i.e. 0.62 to 0.98), which suggests a decrease in reading speed of up to 40%. Although we have mentioned that a NRR1 value of less than 1.0 indicates a decline in reading speed, the analysis assumed that the second minute was representative of the baseline reading speed. Henceforth, we looked at the reading speed calculated on each successive minute. Figure 3.8 shows an example of the graph plot from a participant (i.e. VV07) in SRT-MAG. Typically, the participants who showed a NRR1 value of less than 1.0 had an increase in the reading speed at the second minute, which resulted in a NRR1 value of less than 1.0. Apart from the increase in the reading speed at the second minute, observation of the graph plots did not suggest that there was a declination of reading speed with time.

Table 3.6 Median NRR1 for MT and SRT

	NOR		BOF		MAG	
MT	1.04 (0.09) Range: 0.90 to 1.14		1.04 (0.16) Range: 0.84 to 1.21		1.00 (0.22) Range: 0.80 to 1.94	
	NRR1<1.0 n = 7	NRR1>1.0 n = 19	NRR1<1.0 n = 11	NRR1>1.0 n = 15	NRR1<1.0 n = 13	NRR1>1.0 n = 13
Median range	0.97 (0.08) 0.90 to 0.99	1.07 (0.06) 1.00 to 1.14	0.93 (0.09) 0.84 to 0.98	1.09 (0.09) 1.02 to 1.21	0.91 (0.08) 0.80 to 0.99	1.13 (0.13) 1.00 to 1.94
SRT	1.06 (0.22) Range: 0.65 to 1.41		1.02 (0.23) Range: 0.76 to 1.37		1.14 (0.25) Range: 0.62 to 1.47	
	NRR1<1.0 n = 6	NRR1>1.0 n = 20	NRR1<1.0 n = 11	NRR1>1.0 n = 15	NRR1<1.0 n = 9	NRR1>1.0 n = 17
Median range	0.79 (0.18) 0.65 to 0.91	1.02 (0.20) 1.00 to 1.41	0.82 (0.11) 0.76 to 0.98	1.12 (0.15) 1.01 to 1.37	0.91 (0.10) 0.62 to 0.98	1.18 (0.18) 1.02 to 1.47

Interquartile range (IQR) values are shown in the parentheses

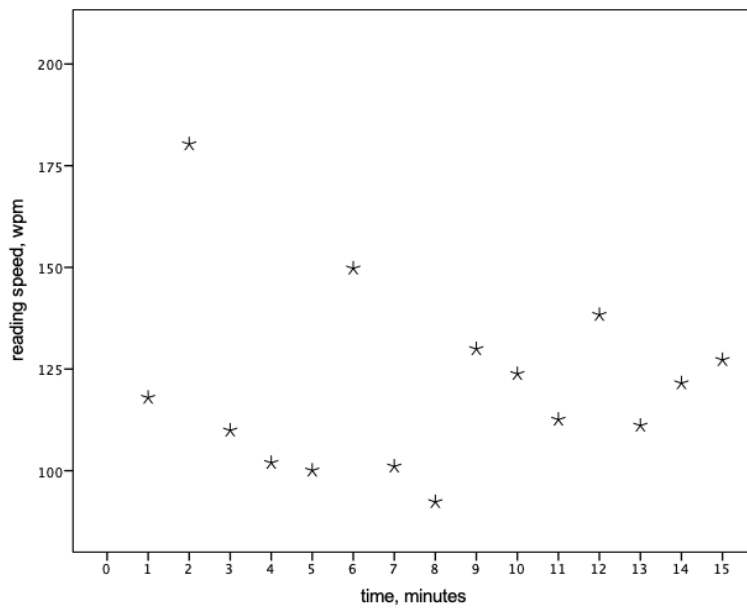


Figure 3.8 A reading speed against time graph plot (Participant VV07), showing a sudden increase in the reading speed at the second minute.

3.4.3 Calculation of normalised reading rate (NRR2) using time-points

NRR2 is another novel method to investigate duration related effects by comparing the reading speeds at two time-points – at the beginning and the end of the reading text. Table 3.7 shows the NRR2 for all the reading tests in each condition. The RoR had the lowest median NRR2 compared to the other reading tests, and the median NRR2 was less than 1.0 in all conditions. This suggested that the participants showed a marginal reduction in reading speed in the second segment of the reading text. The Friedman test indicated a statistically significant difference in NRR2 between the different reading tests in NOR ($\chi^2(3, 26) = 14.18, p = 0.003$) and BOF ($\chi^2(3, 26) = 14.65, p = 0.002$), but not in MAG ($\chi^2(3, 26) = 4.15, p = 0.24$). Post hoc Wilcoxon Sign Ranked test revealed a statistically significant difference in NRR2 between the IReST and RoR ($z = -3.45, p = 0.001$), the RoR and MT ($z = -2.77, p = 0.006$), and the RoR and SRT ($z = -2.66, p = 0.008$) in NOR; IReST and RoR ($z = -2.91, p = 0.004$), RoR and MT ($z = -3.03, p = 0.002$), and RoR and SRT ($z = -2.66, p = 0.008$) in BOF.

Table 3.7 Median NRR2 for all reading tests

	NOR	BOF	MAG
IReST	1.08 (0.17) Range: 0.80 to 1.28	1.09 (0.23) Range: 0.65 to 1.35	1.14 (0.31) Range: 0.69 to 1.93
RoR	0.95 (0.18) Range: 0.66 to 1.28	0.94 (0.21) Range: 0.73 to 1.23	0.97 (0.20) Range: 0.55 to 1.20
MT	1.06 (0.15) Range: 0.90 to 1.33	1.08 (0.20) Range: 0.83 to 1.26	1.02 (0.22) Range: 0.82 to 1.94
SRT	1.12 (0.29) Range: 0.58 to 1.62	1.08 (0.28) Range: 0.77 to 1.58	1.13 (0.36) Range: 0.67 to 1.78

Interquartile range (IQR) values are shown in the parentheses

3.4.4 Factors associated with reading slopes, NRR1 and NRR2

The Spearman correlation was used to investigate if there are any clinical measures that were associated with the duration related effects indicators. It was hypothesised that a bigger effect on the reading slopes would be observed in a more visually demanding condition. To determine the level of difficulty experienced by the participant, a relative reading difficulty ratio was computed by comparing the overall average reading speed from the BOF and MAG to the NOR. The median relative reading difficulty ratios were 1.02 (IQR: 0.21, range: 0.65 to 1.33) and 0.53 (IQR: 0.20, range: 0.35 to 0.74) in BOF and MAG, respectively. Unlike in MAG, 15 participants in BOF had a reading difficulty ratio of more than 1.0, suggesting that the BOF reading condition was not visually demanding. As such, only the reading difficulty ratios in the MAG were used in the correlation analysis. Table 3.8 shows that a moderately high correlation was observed between the reading slopes and contrast reserve ($r = 0.42$, $p < 0.001$). Interestingly, the reading speed had a positive correlation with reading slopes ($r = 0.32$, $p = 0.004$), suggesting that participants with faster reading have a higher tendency to increase their reading speed with time. However, a negative correlation between the reading speed and NRR1 ($r = -0.31$, $p = 0.006$), and reading speed and NRR2 ($r = -0.23$, $p = 0.04$) indicated otherwise. These contradictory findings of the relationship between the reading speed and the duration related effects indicators were due to the different approaches and measures used by the indicators to denote a duration related effect. The Spearman correlations between the reading duration related effects indicators (reading slopes vs NRR1 ($r = 0.39$, $p < 0.001$), reading slopes vs NRR2 ($r = 0.53$, $p < 0.001$) and NRR1 vs NRR2 ($r = 0.75$, $p < 0.001$)) suggested that there was only a moderate to moderately high correlation between the reading duration related effects indicators. There were negative low correlation between the relative

reading difficulty and the reading slopes ($r = -0.20$, $p = 0.33$), and moderate correlation between the relative reading difficulty and the NRR2 ($r = -0.38$, $p = 0.05$).

Table 3.8 Spearman correlation coefficients between clinical measures and duration related effects indicators – reading slopes, NRR1 and NRR2 (n = 26)

Factors	Duration related effects indicators		
	SRT Reading slopes	SRT NRR1	SRT NRR2
Refractive error ^{a,b}	$r = 0.03$ $p = 0.89$	$r = 0.16$ $p = 0.45$	$r = 0.12$ $p = 0.57$
Relative reading difficulty ^a	$r = -0.20$ $p = 0.33$	$r = -0.13$ $p = 0.54$	$r = -0.38$ $p = 0.05$
Contrast Reserve	$r = 0.42$ $p < 0.001$	$r = -0.02$ $p = 0.86$	$r = -0.08$ $p = 0.47$
Reading speed	$r = 0.32$ $p = 0.004$	$r = -0.31$ $p = 0.006$	$r = -0.23$ $p = 0.04$
Proofreading scores	$r = 0.13$ $p = 0.25$	$r = 0.02$ $p = 0.83$	$r = 0.05$ $p = 0.64$

^a Given that the refractive errors do not change with different reading conditions, the correlation test between the reading slopes, NRR1 and NRR2 were only limited to the MAG condition.

^b Refractive error is based on the worse eye spherical equivalent.

The inclusion of the IReST and RoR was to evaluate if the two much shorter commercially-available reading tests could replicate the effects of the SRT. Concurrently, we have also included another short reading text (i.e. MT) to assess if the duration related could be captured using non-standardised reading materials. Table 3.9 shows the association between the IReST, RoR, MT NRR2, and the SRT duration related effects indicators. Spearman correlation did not show any correlation between the SRT duration related reading speed performance indicators and the NRR2 in the respective reading tests. This finding suggests that the changes in reading speed performance with time could vary between reading tests, and the use of shorter reading test did not seem to be able to replicate the reading speed changes effect relative to a leisure reading task (i.e. SRT).

Table 3.9 Spearman correlation coefficients between IReST, RoR and MT NRR2 and the duration related effects indicators (two-tailed) (n = 26)

Duration related effects indicators	IReST	RoR	MT
SRT reading slope	$r = 0.02$ $p = 0.84$	$r = -0.01$ $p = 0.95$	$r = 0.17$ $p = 0.14$
SRT NRR1	$r = 0.09$ $p = 0.44$	$r = -0.01$ $p = 0.96$	$r = 0.02$ $p = 0.84$
SRT NRR2	$r = 0.10$ $p = 0.41$	$r = 0.05$ $p = 0.67$	$r = 0.05$ $p = 0.67$

3.5 Discussion

The primary objective of this study was to investigate if there were any changes in reading speed performance in a group of normally-sighted participants under simulated visual impairment as they read for a 10 to 15 minutes duration – whether 1) this was more likely to occur in visually challenging conditions, 2) and equivalent effects could be shown over shorter reading periods. Here, the use of linear reading slopes and novel alternative analysis approaches of the data were described. Reading slopes and NRR1 were devised to quantify and provide insights on individual reading speed performance with time. Reading slopes utilised the reading speed on each successive minute to derive an overall performance based on the gradient of the slope. Based on the overall median reading slopes, most of the participants showed an increased reading speed with time in both the MT and SRT. However, when reading under visually challenging condition (i.e. using a hand-held magnifier under simulation), the increases were markedly reduced.

To date, there are not many studies investigating duration related effects in reading performance assessments. The use of reading slope (i.e. wpm/min) to capture a duration related effect was first reported by Ramulu et al. (2013a). Using a simple linear regression model, a negative gradient slope would suggest a trend that the reading speed decreases with time. Although sample reading slopes plots were presented in Ramulu et al. (2013a), explicit details on the overall linear model fit (i.e. R^2) were not reported. In the current study, it was observed that the linear function did not provide a suitable fit for most of the participants based on the R^2 values. By observation of the scatter plots (i.e. reading speed against time) was used to predict an appropriate curve fitting. However, visual inspection of the scatter plots of all the participants did not suggest a typical non-linear relationship curve. It was hypothesised that if the reading speed declined with time, it might plateau before a drop in reading speed. If this is the case, an exponential function may then provide a closer fit of the data points in the latter scenario. As such, non-linear curve fitting (i.e. logarithmic and exponential) was applied to all the individual graph plots in the SRT MAG to determine whether there is any overall increase or decrease in reading speed with time. However, there was no significant improvement in the curve fitting using the logarithm or exponential function based on the R^2 values, and on the

observation of the number of actual data points which did not coincide with the logarithmic or exponential curve. Since the use of other function curves did not seem to improve the graph fitting, the use of the gradient slope from the linear function to represent the overall reading speed changes with time will be used.

Although the use of reading slope values was one method to assess reading speed performance changes with time, the calculation of the slopes can be affected by the presence of outliers. In the second analysis, the NRR1 compares the reading speed in the subsequent minutes to the initial baseline reading speed (i.e. 2nd-minute reading speed). The median NRR1 was then calculated from the duration (e.g. 10-15 minutes for SRT) to denote if there is any increase or decrease in reading speed with time. A value of more than 1.0 will represent an increase in reading speed with time, and a value of less than 1.0 will represent a consistent decrease in reading speed with time. The main advantage of using NRR1 over reading slopes was the reduction of the effect of outliers by averaging the results. However, there are limitations in using the NRR1 analysis method. Firstly, the NRR1 uses the second minute as a reference point to reduce possible timing accuracy when the first page was revealed. However, it is unclear how representative is the second minute as the baseline reading speed, and this "reference point" can potentially affect the NRR1 values. Secondly, the NRR1 does not provide detailed information on the changes in reading performance throughout the reading test. Instead, the NRR1 quantifies an overall performance. For example, an NRR1 value of 0.5 could indicate a gradual decrease in reading speed with time or an initial sudden drop in reading speed, followed by a plateau. Lastly, any duration related effects that happen towards the end of the reading test cannot be picked up using the median values. For example, the NRR1 can be 1.0 until the 7th minute but decrease from the 8th to 10th minute. However, the use of NRR2, which compares the beginning and end of the reading test, would capture such changes. The median values were used instead of mean values because the latter can produce an averaging out effect (i.e. an average value of 1.0 could be due to initial increase and subsequent decrease in the reading speed with time). Although the reading slopes and NRR1 were moderately correlated (MT: $r = 0.30$, SRT: $r = 0.39$), both analyses showed that the reading speeds did not slow down with time. It was also noted that the SRT reading slope

analysis showed that the increase in reading speed reduces with the more visually demanding conditions. On the other hand, the NRR1 value was the highest in MAG, which means the reading speed increase was more apparent in MAG compared to NOR and BOF. It was initially thought that the difference could be due to the inclusion of the first minute in the reading slope analysis compared to NRR1, which uses the second minute as the baseline. The recalculation of the reading slope, that is, exclusion of the first minute in the reading slope analysis (NOR-2nd min: 2.18 wpm/min, BOF-2nd min: 0.97 wpm/min, MAG-2nd min: 0.53 wpm/min) still did not produce agreement with the NRR1, where the change was the highest in MAG. This supported our discussion earlier on the limitation of NRR1; that is, the effects that happened towards the end of the reading test cannot be picked up by NRR1. In this case, the NRR2 might provide more information on the reading speed performance towards the end of the reading test.

The current study used simulated conditions to increase the visual difficulty of the reading tasks. It was hypothesised that the reading speed would progressively reduce during the test with the increase in visual demand of the reading task in BOF and MAG. Based on the findings of the current study, the analyses did not suggest any trend that the reading speed is declining in a group of healthy participants with and without simulation within a reading duration of 10 to 15 minutes. The reading speed appears to increase with time, albeit the increment was significantly decreased in the more visually demanding condition (i.e. simulation with a denser Bangerter occlusion foil and using the magnifier). Although participants were also simulated with the Bangerter occlusion foil in the BOF condition, the simulation was not adequately challenging for the participants – the reading speeds between the NOR and BOF were not significantly different compared to MAG, which was statistically significantly slower (see Table 3.1). There was, however, a moderate correlation between contrast reserve and reading slopes ($r = 0.42$, $p < 0.001$), and reading speed and reading slopes ($r = 0.32$, $p = 0.004$). As the acuity reserve was controlled using the Bangerter occlusion foils, this has inevitably resulted in a variable contrast reserve. Whittaker and Lovie-Kitchin (1993) have reported that a minimum contrast reserve of 10:1 is required for fluent reading. Here, a moderate correlation was found, suggesting that a higher contrast reserve was associated with an increase in

reading speed within the 10 to 15 minutes reading duration. Also, participants with faster reading speeds were associated with increased reading slopes. However, the reading speed does not indicate the level of difficulty experienced by the reader. For instance, the slow reading speed in the MAG could indicate that the participant is struggling with the reading, or it could also be possible that the reader is a slow reader. As such, the reading difficulty ratios (i.e. the reading speed in the MAG relative to the reading speed in NOR) were computed. Contrary to the hypothesis that participants who experienced with more visual difficulty will be more susceptible to duration related effects, our findings did not show that participants who experienced more visual difficulty were more predisposed to negative duration related effects observed from the reading slopes, NRR1 and NRR2 (see Table 3.8). Interestingly, a negative but weak and non-statistically significantly correlation was observed between the reading difficulty and the reading slopes, NRR1 and NRR2 (i.e. participants who experienced lesser reading difficulty showed a greater tendency to have negative duration related effects). The weak “opposite” effect could due to a lesser number of participants with negative duration related effects. The current study was compared to a similar investigation done by Ramulu et al. (2013a), who found a marginal decrease in reading speed in the glaucoma participants group. The glaucoma participants (mean age: 71.6 years, median reading slope: -0.37 wpm/min) were found to be more susceptible to a decrease in reading speed compared to their normally-sighted group (mean age: 67.0 years, median reading slope: -0.02 wpm/min). In the Ramulu et al. (2013a) study, the decline in reading speed was associated with participants with larger visual field mean deviation. The visual field mean deviation refers to the overall value of the total amount of visual field loss, which can also be affected by other conditions such as cataract. Normal values typically range from 0 to -2 dB, but all the participants had a mean deviation worse than -3 dB in Ramulu et al. (2013a) study. Also, Ramulu et al. (2013a) reported that participants who had significant cataracts or posterior capsular opacification were associated with more negative reading slopes suggesting that the negative duration effects may be more obvious in visually demanding conditions. Perhaps, the lack of visually challenging condition in Brussee et al. (2016) may help to explain why a duration related effect was not observed in a group of normally-sighted participants (mean age = 55 years) performing a 20-minute out-loud reading test. Another possible explanation based

on the current study and Brussee et al. (2016), could be that a shorter time duration (i.e. 10 to 15 minutes in the current study and 20 minutes in Brussee et al. (2016)) was not long enough to capture a negative duration related effect. Ageing has been commonly associated with reduced reading speed performances (Akutsu et al., 1991, Lott et al., 2001, Sass et al., 2006, Hahn et al., 2006, Ramulu et al., 2013a, Brussee et al., 2016), although there were also studies (Lovie-Kitchin et al., 2000, Dickinson and Shim, 2007) showing no difference in reading speed with age. In the Ramulu et al. (2013a) study, slower reading speeds were reported in the older age participants (i.e. the glaucoma group, mean age: 71.6 years vs glaucoma suspect group, mean age: 67.0 years) but age per se was weakly associated with decreasing reading speed with time. This may be due to a narrow age range group (normal participants: 58.6 to 75.4, glaucoma participants: 62.5 to 80.7) used in the study. Normally-sighted young individuals were used in our study because elderly individuals may show more confounding effects (e.g. age, physical limitations, etc.) on the results, and we wanted to investigate only possible visual effects on performance.

In this study, the use of two commercially available reading tests, IReST and RoR, were included. The reading tests were chosen as they were of a reasonable length to allow a possibility to investigate duration related effects. The main difference between the two reading tests was the use of contextual materials in IReST and non-contextual materials in RoR. To create a standardised reading text requires extensive efforts, and even so, it may be futile to attempt to develop a so-called "standardised" text. In a recent investigation on reading text difficulty and reliability in five Dutch language reading tests (i.e. Colenbrander, IReST, Laboratory of Experimental Ophthalmology, and 'de Nederlanders') by Brussee et al. (2015), the authors have found that the reading speed performances differ significantly among individual reading test versions. As such, if the use of random words paragraph can elicit similar results as contextual material, then the use of random words would provide an easier alternative for reading endurance assessment. This is because the passage of random words can be easily reproduced in many "standardised" versions by rearranging the sequence of the words.

As these two reading texts are comparatively shorter than MT and SRT, an analysis method was devised to allow comparison across all reading tests. Two segments of reading speed were compared to calculate the NRR2 to determine any increase or decrease in reading speed at the end of the reading test. All reading tests except RoR presented a median NRR2 of approximately 1.0; only the RoR had an NRR2 value less than 1.0 in all the conditions and was statistically significantly lower compared to the other three reading tests (i.e. IReST, MT and SRT) in NOR and BOF. The difference in the instructions given to the participants to read the RoR could be the primary reason for the lower NRR2 values compared to the other reading tests. For the IReST, MT and SRT, participants were asked to read the text as they would when reading their daily materials. It is possible that if the participants were to read at their maximum reading speed (i.e. instructing them to read as fast as possible), they would be more likely to slow down in the remaining part of the reading test. On the other hand, this would force an unnatural reading speed on a contextual reading test. The initial fast reading speed was observed in the RoR, where the reading test uses random words with no contextual meaning. When participants were instructed to read the text as fast and accurately as possible in the RoR, participants tended to max out their breath to read as much as possible. As a result, the first segments were more likely to show a faster reading speed compared to the second segments.

Nevertheless, the reading speed performances using NRR1 were consistent (i.e. no duration related effects) in all the reading tests. Also, a correlation was conducted between the SRT reading slopes and IReST, RoR and MT. The purpose was to find out if the use of a shorter reading test can capture and replicate the same duration related effects observed in a more natural reading environment (i.e. SRT). The weak correlation did not suggest the ability of the shorter reading test to capture and replicate the results seen in SRT. However, it can also be argued that the reading speed changes were too subtle for the reading tests to pick up any effects.

3.6 Limitations

To investigate possible duration related effects, the use of silent reading mode and instructing participants to read as they would do in their daily reading materials such as magazine or newspaper were employed. This emphasis was to create a more real-world reading performance. However, silent reading and instructing participants to read as they would do in their daily reading materials can inevitably contribute to a larger variation in reading speed across the 10 to 15-minute testing time. Given the large variability of the data, a reproducibility assessment on the changes of reading speed with time may be required to ascertain if the observed duration related effects are true effects. However, the current experiment did not capture a second visit assessment. The portable-Electronic Vision Enhancement System (p-EVES) – The p-EVES study (Taylor et al., 2014) presents an opportunity to look into the assessment of reproducibility and will be discussed in Chapter 5.

The current experiment used a convenience sample of 30 participants, and only 26 participants' data were used. Post hoc power analysis using the G*Power (Version 3.1.9.3) for a sample of 26 participants gave a limited power of 48% to detect a difference of 0.39 wpm/min from zero in the reading slopes (see Appendix 9.10.4). Based on the existing data, a sample of 65 participants would be ideal to produce a minimum power of 80% to detect the difference and allows a 10% (approximately 6 participants) data rejection due to any technical or procedural errors.

3.7 Conclusion

We have explored the use of different analyses to identify any possible duration related effects in a 10 to 15-minute reading test. Although an earlier study by Ramulu et al. (2013a) has reported the use of linear reading slope as an indicator of group reading speed changes, our findings suggest that more assessment approaches would be required to characterise individual performance, given the large inherent variability of the reading speed with time. The multiple approaches should include, but not limited to, the visual inspection of the graph plots, reading slopes, and normalised reading rate to determine any possible duration related effects. Despite the multiple analysis approaches, no consistent negative duration related effects were

observed in the current study. However, we are certain that there is no sign of a reduction in reading speed in MT and SRT in a group of young participants. Instead, an overall increase in reading speed was observed. However, when participants were subjected to visually demanding conditions, the amount of reading speed increment was reduced.

4. Investigation of duration related changes in reading speed performance in middle-aged and elderly adults

4.1 Introduction

In the study reported in Chapter 3, the duration related changes in reading speed performances were investigated in a group of young adults in less than optimal reading conditions (i.e. BOF and MAG). Overall, the reading slopes showed that there was no decrease in reading speed with time in different conditions. In fact, the reading speed increased with time, although this increment was markedly reduced in MAG. In the Ramulu et al. (2013a) study, the duration related effects were minimal in elderly adults (mean age = 67 years, median reading slope: - 0.02 wpm/min), but were apparent in the group with glaucoma (mean age = 71.6 years, median reading slope: - 0.37 wpm/min). However, the difference failed to reach statistical significance marginally ($p = 0.06$). On the other hand, Brussee et al. (2016) did not observe a duration related effect in a group of 71 healthy participants (mean age: 55 ± 20 years; range 18 to 86 years). They suggested the effect was not seen because the participants were not subjected to visually challenging conditions. Seemingly, age and visually challenging reading conditions are considered as possible predisposing factors to show a duration related effect in reading speed performance. The experiment was therefore repeated on a group of six elderly participants, and it was found that reading speed decreased with time (ESLRR 2013 Abstracts Oral Presentations – OS9 Page 26 <https://www.eslrr.org/2013-abstracts> See Appendix 9.4). In this chapter, further study with an expanded sample size was reported. It was hypothesised that the reading speed would decrease with time, with the effect being more significant in the more visually challenging condition (i.e. MAG).

4.2 Methods

Thirty-eight participants, aged more than 40 years old, were recruited for this study. As this part of the study was conducted in Singapore (i.e. English was not necessarily the first language of the participants), the inclusion criterion required all participants to have at least an “O” level English education (i.e. secondary school). Although no comprehensive eye

examination or eye reports were requested from the participants, the participants were generally healthy and did not have any reported major ocular disease. Hence, a minimum best-corrected distance VA of 6/9 was required for each eye.

Unlike the previous experiment with the young adults, this group of participants were only required to perform silent reading using the SRT reading test. The full details of the SRT text can be found in Section 3.2.2. Briefly, the text is extracted from an encyclopaedia to create a reading material of general interest and an extended reading passage of approximately 3500 words to allow assessment of any duration related effect. As this is a silent reading procedure, each page-turn of the reading test indicates the time taken to read the previous page. The time for each page-turn was recorded by using an online stopwatch lap timer (<http://www.timeanddate.com/stopwatch/>). Using the time taken and the number of words on each page, reading slope, that is reading speed over a period of time can be calculated to determine any duration related changes.

To create different levels of reading difficulty, the participants performed the reading task in different reading conditions: NOR, BOF and MAG. The full details of the respective conditions can be found in Section 3.2.3. Briefly, NOR refers to the habitual reading condition. As this is a group of older individuals, they used their reading correction. In the case of a progressive lens or multifocal lens wearer, the reading prescription was transferred to a trial frame. In BOF, a reduced vision was simulated with the use of a Bangerter occlusion foil to degrade the image contrast but still achieved a threshold near acuity of 0.3 log MAR. In MAG, participants were simulated with a denser Bangerter occlusion foil to degrade the near VA to approximately 1.0 to 0.9 log MAR. With a denser Bangerter occlusion foil, they are “forced” to use the 4x hand-held magnifier to perform the reading task. Before the experiment, participants were taught how to use the magnifier. They were also given time to familiarise themselves with the use of the hand-held magnifier on five pages of reading text, that was not used in the actual experiment. All the reading materials were positioned on a reading stand. The participants were instructed to read at a standard viewing distance of 25 cm in NOR and BOF except when using the hand-held magnifier; where they were instructed to keep the distance from their eye

to the magnifier as close as possible. The study was approved by the Singapore Polytechnic Ethics Committee (Reference no: 201402-04) and written informed consent was obtained from all the participants.

4.3 Statistical analysis

All data were analysed using Statistical Package for the Social Sciences (SPSS) (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.). Shapiro-Wilk test showed a normal distribution for the reading speed in NOR ($p = 0.40$), BOF ($p = 0.08$) but not in MAG ($p = 0.002$). As such, median and interquartile range measurements were reported. Non-parametric statistical analysis Friedman and post hoc Wilcoxon Signed-Rank tests were used to explore the differences in reading speed and reading slopes among the reading conditions. Spearman correlation was used to investigate the association between reading slopes and normalised reading rates among the different reading tests. A probability of less than 0.05 indicated statistical significance.

4.4 Results

Out of the 38 participants, the data for only 31 were used for this analysis. Seven participants were unable to complete the reading tests in either one or two of the conditions. As such, the analyses were separated into the main group (i.e. participants who completed all the reading tests for all the three conditions) and the subgroup (i.e. participants who did not manage to complete the reading tests in either one or two conditions). Table 4.1 shows the demographics of all the participants. All the 31 participants who completed all the reading tests for all the conditions had a median age of 58.0 years (IQR = 13.0, range: 41 to 68 years). The median habitual binocular near VA and contrast sensitivity were 0.10 log MAR (IQR: 0.0; range 0.0 to 0.2 log MAR) and 1.76 log units (IQR: 0.0, range: 1.68 to 1.76 log units) respectively. The median contrast reserves were 51.79 (IQR: 0.0, range: 43.08 to 51.79):1, 35.83 (IQR: 6.61, range: 24.79 to 39.29):1, and 4.72 (IQR: 2.90, range: 3.27 to 9.00):1 in NOR, BOF and MAG respectively. The elderly group was split into two groups: middle-aged (40 to 60 years) and elderly adults (> 60 years). Out of the 31 participants, there were 17 middle-aged adults and 14 elderly adults. Table 4.1 shows the comparison of the clinical measurements between the

middle-aged group and the elderly group of participants. Mann-Whitney U test did not show a statistically significant difference in the clinical measurements between the two age group of participants.

Table 4.1 Demographics and clinical measurements between the two age groups

Demographics	Middle-aged (40 to 60), n = 17	Elderly (> 60 years), n = 14
Age	median: 50 years (11.50) range: 41 to 58 years	median: 63 years (3.25) range: 61 to 68 years
Refractive error	Myope (n = 11) Range: -0.50 to -8.75DS* Emmetropes (n = 6)	Myopes (n = 6) Range: -0.50 to -6.75DS Emmetropes (n = 5) Hyperopes (n = 3) Range: +0.25 to +2.25DS
Education	O levels (n = 7) Diploma (n = 8) Degree (n = 2)	O levels (n = 8) Diploma (n = 3) Degree (n = 3)

Clinical measurements	Middle-aged (40 to 60), n = 17	Elderly (> 60 years), n = 14	Mann-Whitney U test
Distance VA	0.1 log MAR (0.00) Range: 0.0 to 0.1 log MAR	0.1 log MAR (0.02) Range: 0.0 to 0.1 log MAR	Z = -0.26 p = 0.86
Near VA	0.1 log MAR (0.05) Range: 0.0 to 0.2 log MAR	0.1 log MAR (0.10) Range: 0.0 to 0.10 log MAR	Z = -0.16 p = 0.92
Contrast Sensitivity	1.76 log units (0.00) Range: 1.68 to 1.76 log units	1.76 log units (0.00) Range: 1.72 to 1.76 log units	Z = -0.46 p = 0.83
Reading speed (NOR)	182.29 wpm (94.53) Range: 118.51 to 282.45 wpm	183.40 (76.31) Range: 86.61 to 265.49 wpm	Z = -0.20 p = 0.86

Interquartile range (IQR) values are shown in the parentheses

*spherical equivalent based on the worse eye

Figure 4.1 shows the distribution of the reading speed in the three different conditions. The overall median reading speeds were 182.29 wpm (IQR: 90.53, range: 86.61 to 282.45 wpm), 158.08 wpm (IQR: 63.22, range: 89.23 to 305.10 wpm) and 63.71 wpm (IQR: 43.09, range 20.81 to 147.61 wpm) in NOR, BOF and MAG respectively. As expected, MAG produced the lowest reading speed among the reading conditions. The Friedman test indicated a statistically significant difference in reading speed between the reading conditions ($\chi^2(2, 31) = 48.45, p < 0.001$). Post hoc Wilcoxon Signed-Rank test showed that the reading speeds were statistically significantly different between NOR and BOF ($z = -2.82, p = 0.005$), BOF and MAG ($z = -4.86, p < 0.001$), and NOR and MAG ($z = -4.86, p < 0.001$). Our previous experience on the analysis of duration related effects suggested that a combination of assessments (i.e. linear reading slopes, visual inspection of the graph plots, normalised reading rates) was required to comprehensively evaluate any duration related effects. Therefore, subsequent analyses (i.e.

reading slopes and normalised reading rate) were replicated from the previous chapter and were described in Chapter 3.2.4.

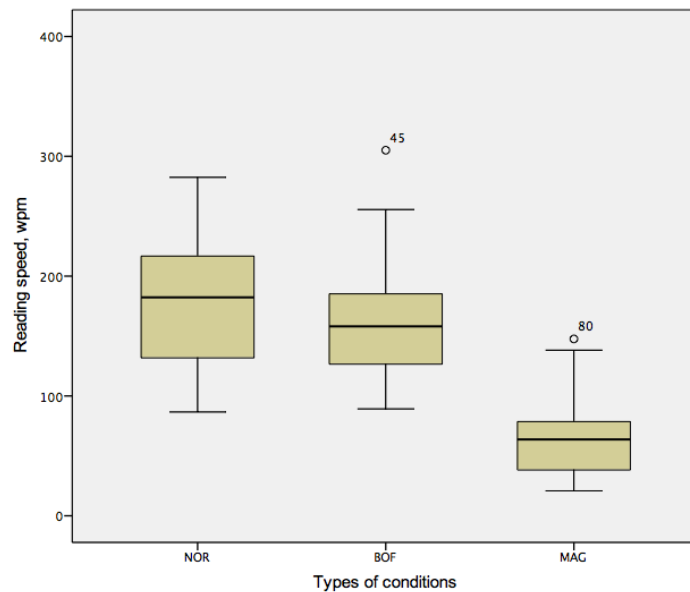


Figure 4.1 The median and interquartile range of the reading speeds for all the participants in NOR, BOF and MAG

4.4.1 Calculation of reading slopes

The reading slopes were captured for the first 10 minutes in both the NOR and BOF, and 15 minutes in the MAG. Similarly to Chapter 3, the group median reading slopes for the SRT were computed based on the median reading speed of all participants at each successive minute. The group median reading slopes were 1.61 wpm/ min (R^2 : 0.35), 0.86 wpm/ min (R^2 : 0.09) and -0.39 wpm/ min (R^2 : 0.12) in NOR, BOF and MAG, respectively. Unlike the young adults' group from the previous chapter, there was a negative duration related group effect observed in the elderly MAG condition. The median reading slopes based on the linear regression function in the SPSS program from all the participant's individual reading slopes were computed; the median slopes were 0.45 wpm/min (IQR: 5.47, range: -7.94 to 10.13 wpm/min), 0.40 wpm/min (IQR: 5.33, range: -8.65 to 7.70 wpm/min) and 0.43 wpm/min (IQR: 2.24, range: -6.28 to 4.61 wpm/min) in NOR, BOF and MAG, respectively. Given the large R^2 values in the linear fitting function (median: 0.29, IQR: 0.60, range: 0.00 to 0.86) in the MAG, alternative curve fittings logarithmic and exponential functions were attempted (Appendix 9.6). The overall median R^2 values for logarithmic and exponential functions were 0.17 (IQR: 0.47, range: 0.00

to 0.82) and 0.30 (IQR: 0.60, range: 0.00 to 0.86). The Friedman test did not show that both the logarithmic and exponential functions were statistically significantly better fitting than the linear function (based on the R^2 values; $\chi^2(2, 31) = 0.735$, $p = 0.69$). Therefore, the report on using linear regression function would be appropriate. Since the previous study by Ramulu et al. (2013a) and the previous pilot study have reported observed duration related effects in a group of participants with an older age group, we have split the analysis into middle-aged and elderly participants. Table 4.2 shows the descriptive statistics for the reading slopes in all the reading conditions. All, except for the elderly participants in the BOF, had overall positive median reading slopes. The positive reading slopes indicated that there was no observed decrease in reading speeds with time in almost all the reading conditions; the close to zero reading slopes suggest that the reading speeds were almost consistent throughout the 10 to 15 minutes testing time. Mann-Whitney U test showed no statistically significant difference in the reading slopes between the middle-aged and elderly participants in all the reading conditions.

Table 4.2 Descriptive statistics on the reading slopes for all the elderly participants (n = 31)

Conditions	Reading slopes (wpm/ min)				Mann-Whitney U test
	40 – 60 years (n = 17)		> 60 years (n = 14)		
	median	range	median	range	
NOR	0.49 (6.51)	-7.94 to 9.90	0.40 (3.40)	-2.58 to 10.13	Z = -0.12 p = 0.91
BOF	0.75 (5.35)	-1.56 to 7.70	-0.67 (6.70)	-8.65 to 6.85	Z = -1.47 p = 0.14
MAG	0.08 (2.20)	-3.12 to 2.26	0.46 (2.54)	-6.28 to 4.61	Z = -0.52 p = 0.61

Conditions	R^2 of the reading slope				Mann-Whitney U test
	40 – 60 years (n = 17)		> 60 years (n = 14)		
	median	range	median	range	
NOR	0.09 (0.38)	0.00 to 0.84	0.03 (0.20)	0.00 to 0.36	Z = -0.96 p = 0.33
BOF	0.05 (0.15)	0.00 to 0.23	0.12 (0.31)	0.00 to 0.78	Z = -1.55 p = 0.12
MAG	0.13 (0.19)	0.00 to 0.78	0.47 (0.67)	0.00 to 0.86	Z = -1.03 p = 0.30

Interquartile range (IQR) values are shown in the parentheses

Out of the 31 participants, 12 (38%), 13 (42%) and 12 (38%) of them had negative reading slopes in NOR, BOF and MAG respectively; only three participants (08HK, 30CC and 44FK) had consistently negative reading slopes in all the three conditions. Table 4.3 tabulated the clinical parameters for the three participants who showed negative reading slopes in all the reading conditions. All the three participants had the same near, VA and contrast sensitivity

as the median values reported in the main group participants. The reading speeds reported in Table 4.3 were based on the habitual reading speed (i.e. NOR). The sample size was too small to allow the use of multiple regression for analysis. From Table 4.3, the only observed trend was that all the three participants were from the elderly adult age group (i.e. > 60 years).

Table 4.3 Clinical parameters of participants who showed a consistent negative reading slopes in all three conditions

Clinical Parameters	Participants ID		
	08HK	30CC	44FK
Age	63 years	63 years	61 years
Distance VA	RE: 6/6 LE: 6/7.5	RE: 6/7.5 LE: 6/7.5	RE: 6/6 LE: 6/7.5
Near VA	N2.5 @ 25cm (0.1 log MAR)		
Contrast Sensitivity	1.76 log units		
Refractive errors	RE: -6.25/-1.25*100 LE: -4.50/-0.75*110	RE: -0.75DS LE: PL	Emmetrope
Reading speed (NOR)	180.18 wpm	168.48 wpm	220.43 wpm

Figure 4.2 shows the graph plots for the three participants who showed negative reading slopes in all the reading conditions. Similarly to our previous analysis with the young adults' group, attempts to find an appropriate model fitting or grouping of model fitting for the graph plots were unsuccessful due to the large variation presented in the participants' graph plot. As seen in Figure 4.2 (HK08 MAG – top right corner graph plot), the reading speed started high initially but started to drop at around the 5th minute. After which, the changes within subsequent minutes was lesser, but the trend suggests that the reading speed reduced gradually with time. In such cases, the exponential curve would thus provide a better fitting compared to a linear fitting slope. In the graph plots of participant 44FK (Figure 4.2 44FK – bottom right corner), the reading speed dropped by “clusters”. For instance, the reading speed was the same for 6th, 7th, 8th minutes, then the same reading speed at 9th, 10th, 11th. This happens because the reading speed was extremely slow, and the participant spent approximately 3-4 minutes reading on the same page. As such, the reading speed for several consecutive minutes was extracted from the same page, and so remained constant across that time. As the reading duration was only carried out up to a maximum of 15 minutes, the exponential and linear fitting produced a similar fitting curve. However, it is unclear how this may change if the reading tasks were extended to a longer duration.

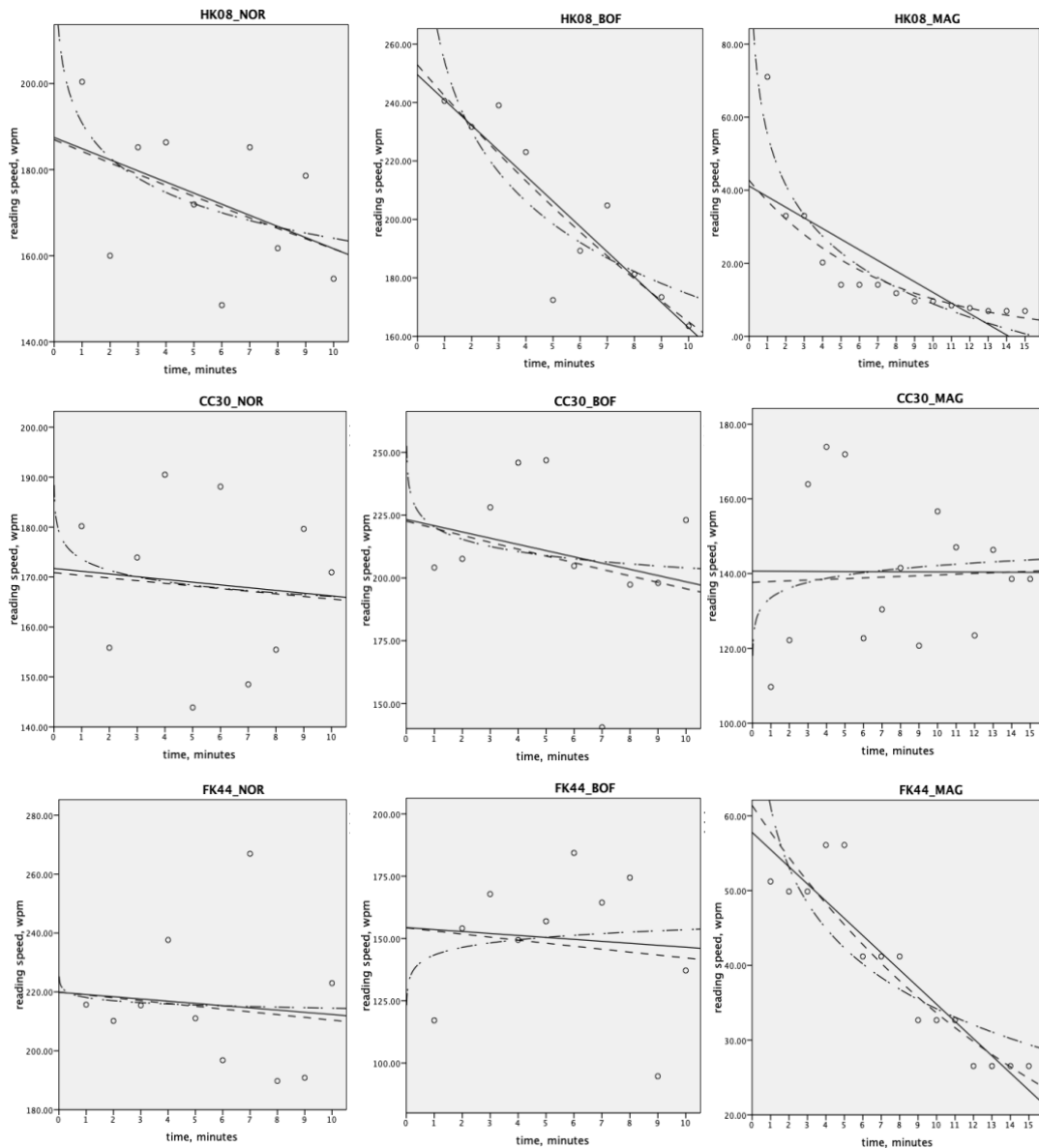


Figure 4.2 Reading slope graph plots for the three participants with consistent negative reading slopes in all the reading conditions. Each row represents the graph plots for each participant in different reading conditions (top row: HK08, middle row: CC30 and bottom row: 44FK). Across the row, graph plots from left: NOR, middle: BOF and right: MAG.

4.4.2 Calculation of normalised reading rate (NRR1)

The NRR1 compares the reading speed in each successive minutes to the baseline reference reading speed (i.e. 2nd minutes reading speed). The overall median NRR1 for all the participants were 1.02 (IQR: 0.17, range: 0.81 to 1.66), 1.02 (IQR: 0.24, range: 0.69 to 1.71) and 1.01 (IQR: 0.43, range: 0.15 to 1.81) in NOR, BOF and MAG respectively. The Friedman test showed no statistically significant difference in the NRR1 between the conditions ($\chi^2 = 2$,

31) = 1.33, $p = 0.51$). Table 4.4 compares the NRR1 analysis between the middle-aged and elderly adults group, and there was no statistically significant difference in NRR1 between the two age groups. The median NRR1s in all the conditions were almost 1.0 in all conditions, suggesting that the reading speeds were almost constant.

4.4.3 Calculation of normalised reading rate (NRR2) using time-points

NRR2 was an alternative method to investigate duration related effects by comparing the reading speed at the beginning of the reading test (i.e. 2nd minute) and the end of the reading test (i.e. 10th/ 15th minute). The overall median NRR2 for all the participants were 1.06 (IQR: 0.36, range: 0.49 to 1.71), 1.05 (IQR: 0.42, range: 0.50 to 1.78) and 1.02 (IQR: 0.52, range: 0.10 to 2.04) in NOR, BOF and MAG, respectively. The Friedman test showed no statistically significant difference in the NRR2 between the conditions ($\chi^2 = (2, 31) = 1.82, p = 0.40$). Table 4.4 shows the descriptive statistics of the NRR2 split into the middle-aged and elderly group participants. In line with reading slopes and NRR1 observation, there were no indications of decreased reading speed with time. Instead, reading speed remained relatively constant throughout the reading duration, even in visually challenging conditions.

Table 4.4 Descriptive statistics of NRR1 and NRR2 in the middle-aged and elderly adults

Conditions	NRR1				Mann-Whitney U test
	40 – 60 years (n = 17)		> 60 years (n = 14)		
	median	range	median	range	
NOR	1.03 (0.42)	0.84 to 1.66	1.00 (0.13)	0.81 to 1.17	Z = -0.88 p = 0.38
BOF	1.00 (0.18)	0.69 to 1.71	1.04 (0.24)	0.82 to 1.60	Z = -0.78 p = 0.44
MAG	1.01 (0.58)	0.66 to 1.81	1.05 (0.27)	0.15 to 1.72	Z = -0.30 p = 0.77

Conditions	NRR2				Mann-Whitney U test
	40 – 60 years (n = 17)		> 60 years (n = 14)		
	median	range	median	range	
NOR	1.09 (0.63)	0.49 to 1.71	1.06 (0.23)	0.83 to 1.63	Z = -0.49 p = 0.62
BOF	0.95 (0.40)	0.72 to 1.71	1.09 (0.45)	0.50 to 1.78	Z = -0.18 p = 0.86
MAG	1.00 (0.45)	0.51 to 1.90	1.08 (0.66)	0.10 to 2.04	Z = -0.54 p = 0.59

4.4.4 Factors associated with reading slopes, NRR1 and NRR2

From the analysis of the reading slopes and normalised reading rates, there was a minority group of participants who presented with negative duration related effects. Spearman

correlation was conducted to investigate if any predisposing factors might be associated with the participants that showed a negative duration related effect in the reading test. Table 4.5 shows the correlation between the proposed associated factors and duration related effects indicators.

It was hypothesised that the participants who struggled more in the reading tasks would be more susceptible to show a decreased reading speed with time. To find out the degree of difficulty experienced by each participant, a relative reading difficulty ratio comparing the overall average reading speed in the MAG to the NOR was calculated. The median relative reading difficulty ratios were 0.36 (IQR: 0.16, range 0.17 to 0.53) and 0.34 (IQR: 0.34, range 0.11 to 0.82) in the middle-aged and elderly adults group respectively. This means the reading speed drops by approximately 65% when reading in MAG. The correlation between the fixed factors (i.e. age, education level, relative reading difficulty ratio and refractive errors) and the duration related effects indicators (i.e. reading slopes, NRR1 and NRR2) was only limited to the MAG condition. This was because these fixed factors did not vary across the different reading conditions. The other clinical parameters (i.e. near VA, contrast sensitivity and reading speed) that were used for the correlation analysis were based on the measurements in the respective reading conditions (i.e. NOR, BOF and MAG). Table 4.5 shows that most of the associated factors were minimally correlated with the duration related effects indicators, and none of these factors has reached statistical significance. The highest correlation observed was between the relative reading difficulty rate and the reading slopes ($r = -0.21$, $p = 0.24$) and NRR2 ($r = -0.24$, $p = 0.19$). The negative correlation suggests that participants with higher relative reading difficulty ratio (i.e. lesser reading speed declination when a switch from NOR to MAG), have smaller reading gradient slopes. However, no relationship was observed between relative reading difficulty rate and NRR1 ($r = -0.08$, $p = 0.65$).

Table 4.5 Spearman correlation coefficients between clinical parameters and duration related effects indicators (two-tailed)

Factors	Reading slopes	NRR1	NRR2
Age ^a	r = -0.006 p = 0.98	r = -0.17 p = 0.37	r = 0.02 p = 0.91
Education ^{a,b}	r = -0.05 p = 0.80	r = 0.16 p = 0.39	r = -0.02 p = 0.92
Relative reading difficulty ^a	r = -0.22 p = 0.24	r = -0.08 p = 0.65	r = -0.24 p = 0.19
Refractive error ^{a,c}	r = 0.11 p = 0.57	r = 0.21 p = 0.25	r = 0.16 p = 0.41
Near VA ^d	r = -0.14 p = 0.17	r = -0.05 p = 0.61	r = -0.08 p = 0.40
Contrast Sensitivity ^d	r = 0.04 p = 0.71	r = 0.01 p = 0.91	r = 0.05 p = 0.60
Contrast Reserve	r = 0.14 p = 0.17	r = 0.05 p = 0.65	r = 0.07 p = 0.53
Reading speed	r = 0.17 p = 0.11	r = -0.04 p = 0.73	r = -0.04 p = 0.72
Proofreading scores	r = 0.02 p = 0.82	r = 0.04 p = 0.74	r = 0.04 p = 0.74

^a Given that these were fixed parameters (i.e. unchanged with different reading conditions), the correlation tests between the reading slopes, NRR1 and NRR2 were only limited to the MAG condition.

^b Education level: 1 = O levels, 2 = Diploma, 3 = Degrees, 4 = Masters or higher

^c Refractive error is based on the worse eye spherical equivalent.

^d The near VA and contrast sensitivity is based on the corresponding Bangerter occlusion foils simulation in the respective conditions

4.4.5 Sub-group (Non-completers)

The subgroup consists of seven participants who did not complete the full set of the reading test. The median age of the subgroup was 67 years (IQR: 12 years, range: 57 to 70 years), with a median near VA and contrast sensitivity of 0.1 log MAR (IQR: 0.10, range 0.1 to 0.2 log MAR) and 1.72 log units (IQR: 0.08, range 1.68 to 1.76 log units) respectively. Mann-Whitney U test showed a statistically significantly older age ($z = -2.21$, $p = 0.03$), poorer near VA ($z = -2.69$, $p = 0.007$) and lower contrast sensitivity ($z = -3.65$, $p < 0.01$) in the subgroup compared to the main group. The median reading speed was slower in the subgroup (i.e. 128.16 wpm (IQR: 135.73, range: 66.12 to 249.78 wpm)) compared to the main group, but the difference did not reach a statistically significant level ($z = -1.00$, $p = 0.34$). Table 4.6 shows the full details of the reading tests that were either incomplete or not done. Among the seven participants, some participants attempted the reading tests but did not manage to finish the reading test of up to 15 minutes. There were also participants who either did not meet the desired near VA when the vision is simulated with the Bangerter occlusion foils simulation or struggled with the use of the hand-held magnifier (e.g. maintaining the distance between the magnifier and the reading text) during the practice session. As such, the participants did not perform the test in that specific condition.

Table 4.6 Details of the reading tests undertaken by participants who did not complete all the reading tests

Conditions	Test completed	Number of participants	Remarks
NOR	Complete	3	Participants 10AL, 31WL and 32KC
	Incomplete	4	Participants 16PH, 24WH, 33LC and 41LJ All had managed to read at least 10 minutes before giving up the reading test
	Not done	0	NA
BOF	Complete	1	Participant 10AL
	Incomplete	3	Participants 24WH, 31WL and 32KC All had managed to read at least 10 minutes before giving up on the reading test
	Not done	3	Participants 16PH, 33LC & 41LJ Unable to achieve the desired near VA (i.e. 0.3 log MAR)
MAG	Complete	0	NA
	Incomplete	3	Participant 32KC: 4 mins Participant 33LC and 41LJ: 10mins
	Not done	4	10AL, 16PH, 24WH and 31WL Participants having difficulty in keeping the text in focus using the magnifier during the practising session

Figure 4.3 shows the reading speed against time graph plots for the non-completers. In the NOR condition, all except Participant 33LC showed an overall increased in reading speed with time, based on the positive linear reading slopes through the data points. Participant 33LC showed a decline in reading from the 3rd to 4th minute but remains relatively consistent up till the 10th minute. There is only one participant who showed a negative reading slope in BOF (i.e. 31 WL). Three participants performed the reading test in MAG but did not complete the reading test in MAG (i.e. up to 15 minutes). Participant 32KC only managed to perform the reading test up to four minutes, which suggests that the reading slope based on these four data points may not be representative of a 15-minute reading test; the remaining two participants have also shown an overall gradual increase in reading speed with time. Although there are participants with negative reading slopes, the changes in reading speed trend were not consistent in the BOF or MAG conditions. This suggests that the negative duration related effects could be a random effect, and may not be reproducible in a repeated assessment.

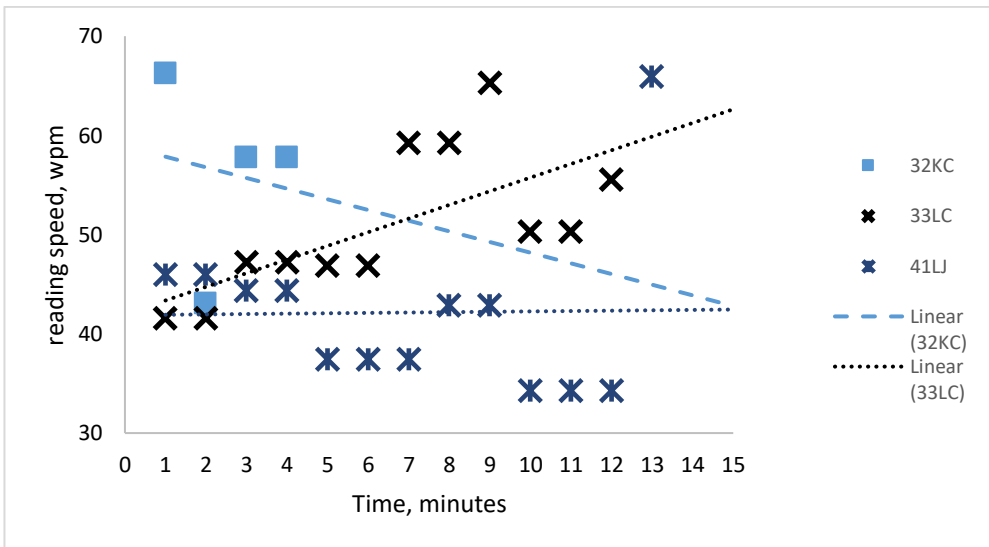
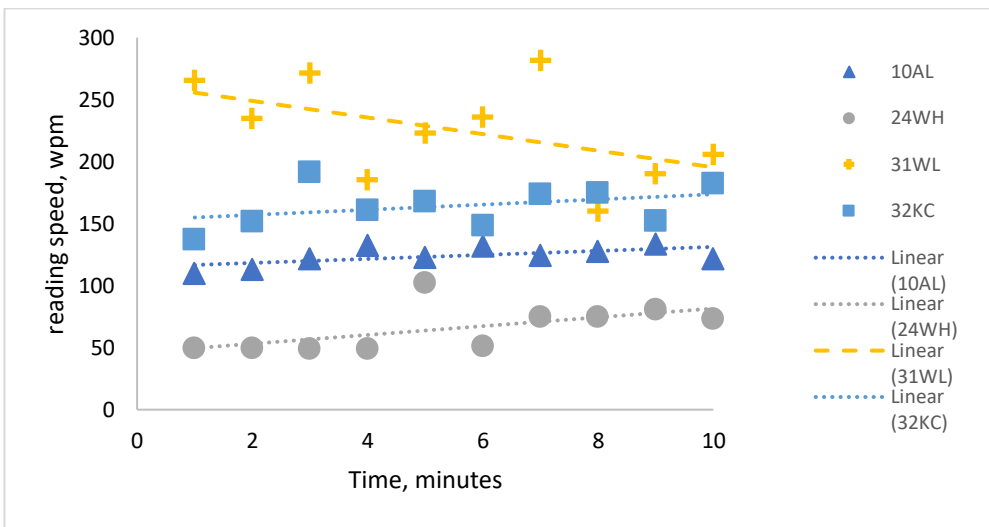
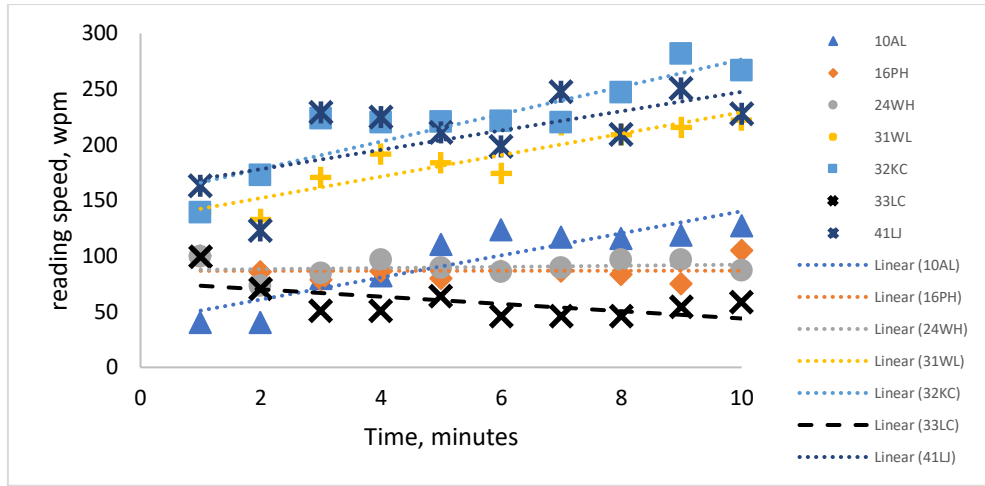


Figure 4.3 Graph plots of the reading speed against time for the non-completers. Top: NOR, middle: BOF, bottom: MAG. Negative linear reading slopes are indicated by dash lines.

4.4.6 Comparison of duration related effects parameters between the young adults and elderly group

The duration related parameters between the previous young adults (Chapter 3) and the current elderly group was compared. As mentioned earlier, the use of the Bangerter occlusion foil created the desired acuity reserve but a variable contrast reserve. In comparison to the contrast reserve in the young adult group, a slightly larger interquartile range in BOF and MAG, and a lower contrast reserve in MAG was observed in the elderly group. Despite the differences in contrast reserve between the young adults and elderly group, Mann-Whitney U test did not show any statistically significant difference in the contrast reserves between the young adult and the elderly group in NOR ($z = -0.18$, $p = 0.86$), BOF ($z = -0.22$, $p = 0.83$) and MAG ($z = -0.69$, $p = 0.49$). Mann-Whitney U tests were used to compare the reading slopes, NRR1 and NRR2 between the young and elderly group. Table 4.7 shows the comparison Mann-Whitney U test analysis results. Apart from a marginal statistically significant difference in the reading slope-NOR, there was no statistically significant difference in the reading slopes, NRR1 and NRR2 between the young adults and elderly group for all the conditions.

Table 4.7 Comparison of duration related effects parameters between the young adults (n = 26) and the elderly group (n = 31)

	Reading slopes (wpm/ min)		Mann-Whitney U test	Median NRR1		Mann-Whitney U test	Median NRR2		Mann-Whitney U test
	Young	Elderly		Young	Elderly		Young	Elderly	
NOR	4.46 (5.24)	0.45 (5.47)	Z = - 2.09 p = 0.04	1.06 (0.22)	1.02 (0.17)	Z = - 0.37 p = 0.71	1.12 (0.29)	1.06 (0.36)	Z = - 0.22 p = 0.83
BOF	1.85 (5.34)	0.40 (5.33)	Z = - 0.75 p = 0.45	1.02 (0.23)	1.02 (0.24)	Z = - 0.78 p = 0.44	1.08 (0.28)	1.05 (0.42)	Z = - 0.22 p = 0.82
MAG	0.37 (2.10)	0.43 (2.24)	Z = - 0.87 p = 0.39	1.14 (0.25)	1.01 (0.43)	Z = - 0.93 p = 0.35	1.13 (0.36)	1.02 (0.52)	Z = - 0.95 p = 0.34

Interquartile range (IQR) values are shown in the parentheses

4.5 Discussion

Ageing can cause normal physiological changes to the visual performance (e.g. spatial contrast sensitivity, visual processing speed, (Owsley, 2011) etc.) and cognitive ability (e.g. information processing (John and Cole, 1986), working memory (Salthouse, 2004), etc.) which are potential factors that can affect reading performance compared to young adults. As such, elderly adults may be more susceptible to duration related effects when reading. Therefore, based on the decreased reading slopes observed in a pilot study of six elderly participants, this experiment increased the sample size for the investigation on duration related effects. Based on the previous study by Ramulu et al. (2013a) and the previous pilot study, duration related effects were reported in a group of “older” participants (Ramulu et al. (2013a) mean age: 71.6 years; pilot study median age: 62.5 years). As such, the analyses were split into two different age groups: middle-aged and elderly group.

Based on the reading slopes, NRR1 and NRR2 findings, the reading speeds were almost constant throughout the 10 to 15 minutes silent reading in habitual and simulated visually challenging conditions. These findings were consistent with Brussee et al. (2016), who also did not find any duration related effects in a group of 71 normally-sighted adults (mean age = 55 years). Their participants were presented with five different reading tests in random order (Colenbrander, IReST, Laboratory of Experimental Ophthalmology (LEO), de Nederlanders (NED) and Radner), which allows a total testing duration of approximately 20 minutes. However, the assessment on reading fatigue was only limited to IReST, which took approximately 8 minutes to complete all the ten versions of the paragraphs. Reading fatigue assessment was determined by using the mixed linear effect to find the association between each IReST paragraph and reading speed. As the sequence of all the five reading tests was arranged in random, the 8-minute IReST could be in any sequence of the total testing duration of 20 minutes; For example, the testing sequence could be 1. Colenbrander, 2. IReST, 3. Laboratory of Experimental Ophthalmology (LEO), 4. de Nederlanders (NED) and 5. Radner, in this sequence the IReST assessment would be the beginning of the total testing duration of 20 minutes. In other words, there could be a possibility that the assessment might be too early (i.e. IReST as the first assessment test) to capture any duration related effects. Brussee et al.

(2016) suggested that the healthy normally-sighted participants were not visually challenged in their experiment, leading to the difference in results compared to the previous studies (Ramulu et al., 2013a, Stangler-Zuschrott, 1990) where participants with visual impairment were recruited.

Our experiment was designed to create possible visually challenging conditions in the elderly community. To our knowledge, this is the first study undertaken to investigate duration related effects in elderly adults with simulated visual impairment and the use of hand-held magnifiers. The main advantage of inducing simulation on normally-sighted participants allows a comparison to their habitual reading speed performances and how differently they performed when subjected to visually demanding conditions. On the other hand, the lack of knowledge of using the magnifiers may create a potential source of a learning effect (i.e. reading speed increases with time). It was hypothesised that the participants who showed large relative reading speed changes in MAG compared to reading in NOR were maximally visually challenged as denoted by low relative reading ratio would be more susceptible to reading fatigue (i.e. decreased reading speed with time). However, only a low correlation was observed between the relative reading difficulty ratio vs the reading slope ($r = -0.21$, $p = 0.24$); and relative reading difficulty ratio vs NRR2 ($r = -0.24$, $p = 0.19$). The association between reading difficulty ratios and reading slopes was similar to our previous study on a group of young adult participants ($r = -0.20$, $p = 0.33$). The negative and low correlation suggests that the participants who experienced less reading difficulty relative to the habitual reading speed (i.e. higher reading difficulty ratios) have a greater tendency to decreasing reading slopes.

In the Ramulu et al. (2013a) study, a decrease in reading speed with time denoted by the overall median reading slope was reported in the glaucoma participants group (-0.37 wpm/min). Despite an overall median group performance, the negative duration related effects were relatively subtle. In the current study, a negative group median reading slope was observed (-0.39 wpm/min) only in the elderly group under the MAG. However, when the analyses were extended to individual reading performance, the data were too variable to reveal any effect. On the contrary, the individual reading slopes indicated that most of the

young participants showed an increase in reading speed with time, and the elderly reading speeds were almost constant with time in all the conditions. As both groups of participants had no prior experience with the use of hand-held magnifiers, there is potentially an overlying practice effect (despite having practice sessions were conducted) which tends to improve the reading speed with time. This “practice” effects may override any possible subtle duration related effects, which can show up as an overall increase or consistent reading speed with time. Perhaps, investigations on the visually impaired participants who are more familiar with the use of the LVAs may reduce this “practice” effect during the assessment. The data from the p-EVES study may then allow us to investigate if the duration related effects can be identified in a group of experienced LVA users.

An earlier pilot study had found negative duration related effects in a group of six elderly participants, which motivated the main study to expand the sample size. The overall median SRT reading slopes were 0.28 wpm/ min (IQR: 11.50, range: – 8.53 to 12.25 wpm/ min), – 2.56 wpm/ min (IQR: 9.72, range: – 12.07 to 3.93 wpm/ min) and – 1.38 wpm/ min (IQR: 2.78, range: – 2.85 to 0.18 wpm/ min) in NOR, BOF and MAG, respectively. In contrast to our earlier pilot study, the current experiment did not manage to replicate the same findings, that is, a decrease in reading speed with time. We compared the clinical measures between the pilot study group and the elderly age main group (i.e. > 60 years). Table 4.8 shows the comparison of the clinical measurements between the pilot study and the main study using Mann-Whitney U test. All clinical measurements, except contrast sensitivity, did not show a statistically significant difference between the pilot study and the main study – suggesting that the difference in findings (i.e. negative duration related effects observed in the pilot study group) was not due to a very different profile group. Given the small sample size, the negative duration related observed in the pilot study could be a one-time effect that may not be reproducible.

Table 4.8 Comparing the demographics and clinical parameters between the current study and the pilot study

	Pilot Study n = 6 median: 62.5 (8.75) range: 56 to 67 years	Elderly (> 60 years) n = 14 median: 63 (3.25) range: 61 to 68 years	Mann-Whitney U test
Distance VA	0.1 log MAR (0.00) Range: 0.0 to 0.1 log MAR	0.1 log MAR (0.02) Range: 0.0 to 0.1 log MAR	Z = -0.26 p = 0.86
Near VA	0.1 log MAR (0.02) Range: 0.0 to 0.10 log MAR	0.1 log MAR (0.10) Range: 0.0 to 0.10 log MAR	Z = -0.24 p = 0.81
Contrast Sensitivity	1.72 log units (0.08) Range: 1.68 to 1.76 log units	1.76 log units (0.00) ^a Range: 1.72 to 1.76 log units	Z = -2.82 p = 0.005
Reading speed (NOR)	207.07 wpm (49.65) Range: 183.92 to 280.91 wpm	183.40 (76.31) Range: 86.61 to 265.49 wpm	Z = -1.15 p = 0.25
Relative reading difficulty	0.44 (0.17) Range: 0.36 to 0.56	0.34 (0.34) Range: 0.11 to 0.82	Z = -1.32 p = 0.19

Interquartile range (IQR) values are shown in the parentheses

^a the IQR was 0 because there is only 1 participant who had a score of 1.72 log units, the rest of the participants achieved a score of 1.76 log units

A considerable amount of research has been published on the effect of age on reading speed performances with mixed findings of decreased (Akutsu et al., 1991, Lott et al., 2001, Sass et al., 2006, Hahn et al., 2006), or no difference in reading speeds (Lovie-Kitchin et al., 2000, Dickinson and Shim, 2007) with older age. In studies reporting a difference in reading speed between different age groups, the researchers (Akutsu et al., 1991, Lott et al., 2001, Sass et al., 2006, Hahn et al., 2006) instructed the participants to read as fast and as accurately as possible. Although the use of MRS may encourage a more reproducible measurement as it forces the participant to read at their maximal performance (Legge, 2007), it may not be very representative of real-world reading. On the other hand, in those studies (Lovie-Kitchin et al., 2000, Dickinson and Shim, 2007) that reported no difference in reading speed between the young and elderly participants, it was observed that the emphasis on the instructions given to the readers was to encourage reading rather than skimming through the text (e.g. “read accurately at their normal speed” (Dickinson and Shim, 2007) and “read at a normal rate to obtain meaning from the sentence or passage” (Lovie-Kitchin et al., 2000)). When participants are instructed to read as fast and as accurately as possible, they are challenged both on their ability to process visual information and ability to speak fast. Smith et al. (1987) have reported that elderly adults speaking durations were 20 to 25% longer than young adults. This slower speaking ability in elderly adults may have possibly limited their ability to reach a higher

reading speed compared to younger adult participants. On the other hand, when participants were instructed to read at their comfortable speed, one's ability to speak fast is not maximally challenged. As such, the difference in reading speed between the young and elderly groups would not be significant. Dickinson and Shim (2007) only observed a difference in reading speed when their participants were using the magnifiers, suggesting that the difference was probably due to the different manual dexterity between the young adults and elderly group. Lovie-Kitchin et al. (2000) reported no significant correlation between age and reading speed but it could be that the poor correlation was limited by a small sample size of participants (AMD: $n = 13$, JMD: $n = 9$) and a small age range (i.e. only 4 in the young group were aged less than 40 years). Since verbal articulation is not a limiting factor in silent reading, we would expect the reading speeds to be similar between the elderly and young participants in the current study. Here, we found that the reading speeds in elderly adults were statistically significantly lower compared to the young adults in all the conditions – NOR, BOF and MAG. The findings were similar to most of the previous studies (Akutsu et al., 1991, Lott et al., 2001, Sass et al., 2006, Hahn et al., 2006), and suggested that the difference use of reading modes (i.e. oral or silent) was unlikely to be the reason for the similar reading speeds between the young and elderly groups observed in Lovie-Kitchin et al. (2000) and Dickinson and Shim (2007). Previously, it has been reported that older participants had a higher tendency to execute more regressions than the young adults during silent reading (Rayner et al., 2006), which may lead to a significantly lower reading speed compared to the young adults. However, this can only be determined through the use of eye movements tracking during the reading process.

4.6 Limitations

One of the key limitations in the current study was the age of our participants. The current study recruited a group of participants aged 40 and above, which includes the middle-aged (i.e. 40 to 60 years) and the elderly group (i.e. > 60 years). Experimental studies (i.e. Ramulu et al. (2013a), the previous pilot study) that reported a duration related effect were mainly consisted of older age participants (i.e. > 60 years). Since most of the analyses were split into two separate groups, this caused a reduced sample size, which further limits the power to

detect a statistically significant difference or a change in reading speed performance with time. On the other hand, our data analysis did not show a consistent trend of a decline in reading speed within the 10 to 15 minutes of our testing duration in the elderly participants' group. The recruitment criteria of participants above 60 years with at least an "O" level English education (i.e. secondary school) poses a great challenge in Singapore and the sample size shrank further when seven participants were unable to continue the reading test due to inability to focus the reading material using the hand-held magnifier. Future work could consider the use of Chinese reading materials to reach out to a more elderly adults group in Singapore. One of the potential limitations of using the Chinese reading text is to determine whether the level of the reading material was within the reader's ability, since speaking Chinese as their first language may not necessarily equate to their ability to recognise the Chinese characters. Although obtaining the education level can indicate the English proficiency level, it would be useful to find out the reading language of the reader since English is not the only language used in Singapore. An assessment of the participants reading profile (e.g. reading language, frequency, duration, etc.) would be useful to establish the inclusion criteria for future related studies.

This investigation did not find any strong correlation between the proposed associated factors (e.g. age, near VA, contrast sensitivity, etc.) and duration related effect indicators (i.e. reading slopes, NRR1 and NRR2 (Table 4.5)). There may be other possible predisposing factors that may cause some participants to be more susceptible to duration related effects. Recently, Schakel et al. (2017) conducted a fatigue symptoms evaluation interview on a group of 16 visually impaired adults using the Fatigue Assessment Scale (FAS). The assessment on the level of fatigue was based on the participant's daily experience. Out of these 16 visually impaired participants, 11 of them reported that the symptoms of fatigue seemed to be caused by "high effort to establish visual perception". The current study did not capture information on the patient's subjective response on fatigue, and it may be worthwhile to consider the use of additional fatigue assessment test before and after the reading task in future studies.

Post hoc power analysis using the G*Power (Version 3.1.9.3) for a sample of 31 participants gave a limited power of 12% to detect a difference of 0.14 wpm/ min from zero in the reading slopes (see Appendix 9.10.5). The extremely low power was primarily due to the large variation in the reading slopes, which was used to calculate the post hoc power analysis. The existing data suggest a sample of 300 participants would be ideal to provide a power of 80% to detect a difference. However, consistent findings of negative duration related effects in all the three conditions were only seen in three participants, and there was no observable trend in the graph plots. Hence, it is unlikely that the expansion of the sample would yield a different result. In conclusion, we found no consistent decrease in reading speed with time over a period of 10 to 15 minutes of silent reading under visually challenging conditions in elderly adults.

5. To investigate the use of IReST to assess the effect of duration on reading performance in a group of visually impaired individuals

5.1 Introduction

In the previous chapters, we have reported the investigation of duration related effects in a group of normally-sighted young and elderly participants under conditions of simulated visual impairment. Normally-sighted young participants were chosen as they have less confounding factors such as physiological changes in the eye (e.g. cataract, miotic pupils, etc.). In Chapter 4, we have expanded the experiment to a group of elderly participants based on a declined in reading speed with time, observed in a pilot study.

In Chapter 3, we have briefly highlighted the key findings in the previous studies (Stangler-Zuschrott, 1990, Ramulu et al., 2013a) reporting a decrease in reading speed with time in participants with visual impairment. Although Stangler-Zuschrott (1990) was the first to report a duration related effect on reading speed performance in visual impairment participants, all reported information in this report was based solely on the available English abstract. The full details of the experiment design were not available as the publication was in German. In Stangler-Zuschrott (1990) study, the author recruited a group of 20 participants with a mix of pre-existing eye conditions (i.e. cataract, glaucoma and maculopathy). Using a short testing duration of approximately two minutes, he reported that participants with maculopathy showed a decrease in reading speed with time and participants with cataract showed an increase in reading speed with time. Although Stangler-Zuschrott (1990) reported using infrared oculography, it was unclear what parameters were used to determine how the changes in reading speed performance occurred and how the assessment was done within the short testing time of two minutes. Also, the small sample size ($n = 20$) may limit the ability of the study to be a good representation of the observed effect from the general low vision community. In a more recent study by Ramulu et al. (2013a), where a large sample size was used, the authors reported a decrease in reading speed in both the glaucoma participants group ($n = 64$, $- 0.37$ wpm/min) and the glaucoma suspect participants group ($n = 59$, $- 0.02$ wpm /min).

So far, no study has been carried out to investigate duration related effects in patients with visual impairment using magnifiers for reading. It would be meaningful to investigate any duration related effects in reading performances in this community as this may allow us to relate their ability to perform leisure reading. The p-EVES study presented a perfect opportunity to investigate reading duration effects in a large group of participants with visual impairment. The primary objective of this chapter is to report any duration related changes in reading speed in a group of individuals with visual impairment when reading with magnifiers.

5.2 Methods

Data were available from an existing study on the use of p-EVES – The p-EVES study (Taylor et al., 2014). The objective of the p-EVES study was to investigate the potential benefit of adding electronic magnifiers to the range of optical devices available. In this two-arm crossover randomised controlled trial, 100 low vision individuals with experience in using optical vision aids were recruited from Manchester Royal Eye Hospital (MREH). Figure 5.1 shows the flow chart of the p-EVES Study design. The complete detailed list of visual (e.g. distance VA, threshold near acuity, reading speed, etc.) and non-visual assessments (i.e. Addenbrooke's Cognitive Examination, Brief Resilience Scale, Finger-to-nose test, etc.) can be found in Taylor et al. (2014). In this report, only measures relevant to the objective of this study were presented.

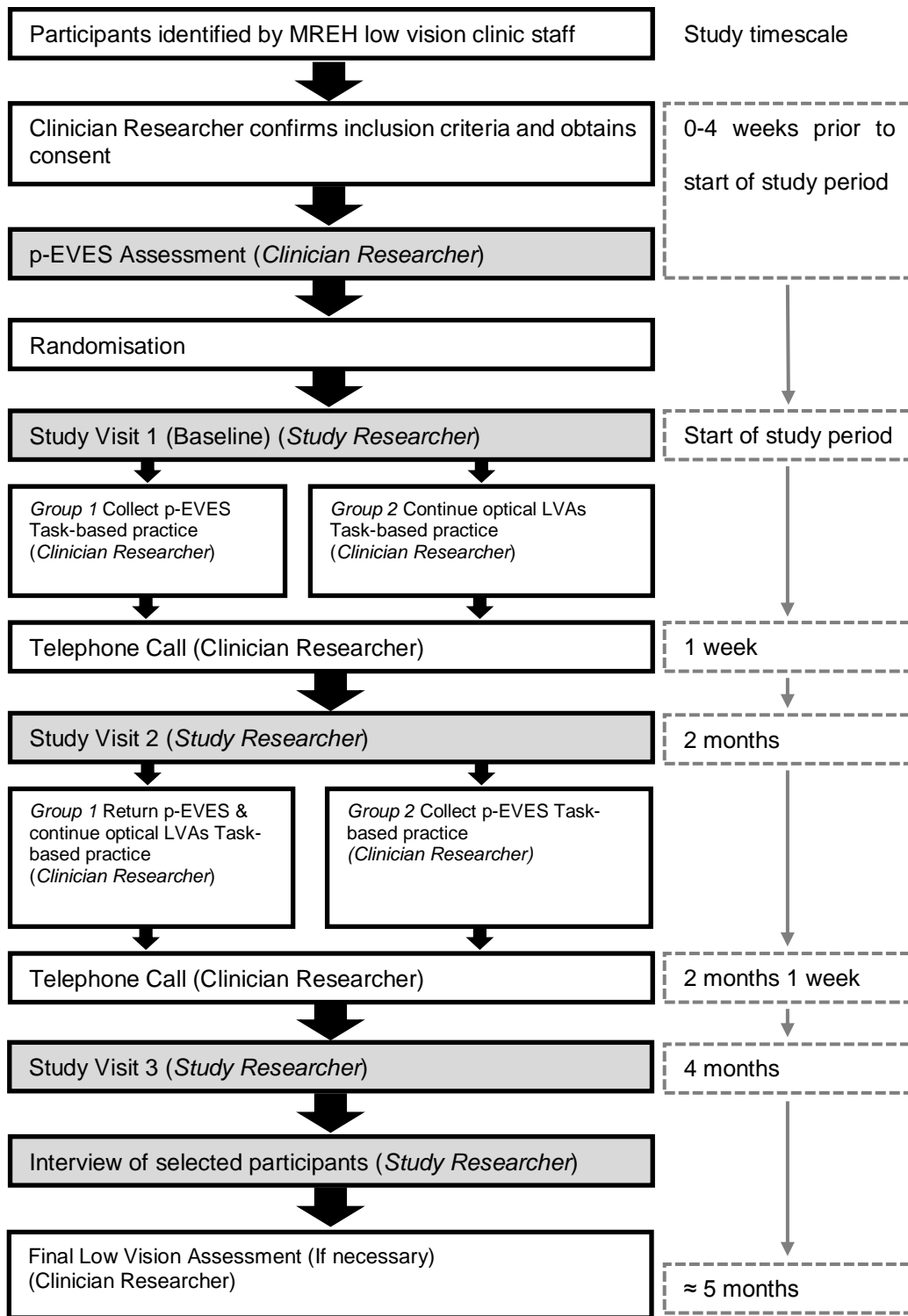


Figure 5.1 Flowchart of p-EVES Study (Taylor et al., 2014)

5.2.1 California Central Visual Field Test (CCVFT) (Cole, 2008)

The participants in the p-EVES study had a diversified range of eye conditions. The nature of the eye conditions in individuals with similar visual acuity can affect reading in varying degrees (Fine and Rubin, 1999); in particular, patients with central scotoma experience more difficulties with reading (Whittaker and Lovie-Kitchin, 1993). The California Central Visual Field Test (CCVFT) (Cole, 2008) was used to delineate participants with or without macular scotomas. We hypothesised that participants with macular scotomas, that is, with more difficulty in reading, would be more susceptible to duration related effects.

The CCVFT was used to quantify the extent and location of the binocular scotomas within the central 15° of the visual field. A recording chart was placed in the plastic holder, which was placed upright on a table, 57 cm from the subject. Three different charts, each with a different size central fixation spot, were used. The chart with the smallest target, which could be seen comfortably by the participant, was chosen. There were three laser pointers which provide the variable red spot stimulus: small (1mm) and dim (laser #1), small and brighter (#2) and large (3mm) and bright (#3). However, only lasers #2 and #3 were used in the study, since #1 was rarely visible to participants.

The participants viewed the central spot throughout the test, and the researcher briefly flashed the laser in numerous locations onto the back of the recording chart, asking the observer to tap the table when they see the target. "Seen" and "not seen" locations can be pencilled on the back of the chart, and transferred to the front at the end of the test. Trials for false positives can be introduced by pointing the laser towards the floor intermittently; false negatives were determined by repeating previously-seen locations. For the ease of statistical analyses, a simple grading system was formulated by the p-EVES Study group, as seen in Table 5.1.

Table 5.1 CCVFT Grading System

Grading Scores	Location of defects
0	no defect within central 15°
1	only a defect (with either laser 2 or laser 3) outside central 10°
2	a defect with laser 2 between 5°-10°
3	a defect with laser 3 between 5°-10°
4	a defect with laser 2 inside 5°
5	a defect with laser 3 inside 5°

5.2.2 Manchester Low Vision Questionnaire (MLVQ) (Harper et al., 1999)

The purpose of the Manchester Low Vision Questionnaire (MLVQ) (Harper et al., 1999) is to evaluate the effectiveness of low vision rehabilitation. Information extracted includes the purpose (see Table 5.2), frequency (see Table 5.3), and duration (see Table 5.4) of their current near vision optical magnifier usage. The questionnaire was administered in all their three visits. However, only information collected during the first visit was presented in this report as we were interested to find out if the usage profile was associated with any duration related effects. We hypothesised that participants who have been using their near vision optical magnifiers for leisure reading might have less duration related effects compared to participants who only use their near vision optical magnifiers for spot reading. The rationale was that those who were only doing spot reading might have given up leisure reading because they have experienced such a duration related effect.

Table 5.2 Purpose of the near vision optical magnifier

How do you perform the task: reading instructions (e.g. spot or survival reading)? reading books and newspapers (e.g. leisure reading)?	
Grades	Description
5	Does not perform task
4	Cannot perform task
3	No aid required
2	p-EVES device
1	Optical aid

Table 5.3 Frequency of near vision optical magnifier usage

How often has the most-used near vision magnifier been used in the last 4 weeks?	
Grades	Description
4	Uses very often = Many times (≥ 5) each day (5x or more each day)
3	Uses often = Several times (1-4) each day
2	Uses sometimes = (< 1 daily but at least ≥ 1 per week) (used more than once a week but not every day)
1	Uses occasionally = (< 1) per week (less than once a week)
0	Never (not used at all in the last 4 weeks)

Table 5.4 Duration of near vision optical magnifier usage

What is the longest length of time (on any one occasion) that a single near vision magnifier has been used in the last 4 weeks?	
Grades	Description
4	≥ 30 minutes
3	≥ 15 minutes but < 30 minutes
2	≥ 5 mins but < 15 minutes
1	≥ 1 minute but < 5 minutes
0	< 1 minute

5.2.3 International Reading Speed Texts (IReST) (Trauzettel-Klosinski et al., 2012)

The IReST is a standardised text used to measure reading speed. There are ten different text paragraphs, but only three were used in the p-EVES study. IReST paragraphs 1, 2 and 3, which consists of 156 words (830 characters), 161 words (824 characters) and 156 words (825 characters) respectively were used to assess reading speed. Each paragraph is made up of 16 rows of words. Participants were instructed to read one IReST paragraph by using their existing optical magnifiers. Their oral readings were audiotaped for analysis. All mistakes and missing words were omitted for reading speed calculation. As there were three different IReST text paragraphs used across the follow-up visits, standard words per minute (swpm) to represent reading speed was used. A standard word length of six characters (Carver, 1992) mitigates the effects of variable word length among the three IReST text paragraphs. The average reading speed was calculated using the number of words read correctly divided by the total time in minutes. The benefits of having an extended reading paragraph compared to single sentences allow a more repeatable reading speed assessment since it is not so

influenced by single random errors. In this study, the use of an extended text also gives the opportunity to investigate reading speed changes over time.

5.3 Data analysis

5.3.1 Reading slope

In our previous analysis in Chapter 3, reading speed was plotted against successive minutes to determine any duration related effect when using an extended reading text (i.e. approximately 3500 words in each reading text). Reading slope, the gradient of the best straight line plotted using the reading speed against successive minutes throughout the reading duration, was used as a measure of any duration related effects. This was possible and meaningful because the participants in the previous study took an average reading time of approximately 10 to 15 minutes. In this experiment, the participants were given only a single IReST paragraph to read with their magnifiers. The total time taken by the participants to complete the single IReST paragraph ranged from 47.99 seconds to 596.85 seconds (approximately 9 minutes). As such, the assessment of reading speed based on individual minutes may not be appropriate because it is hard to make a comparison of the reading performances between a 1-minute vs a 10-minute task duration. Instead, an alternative way is to break down the paragraph into an individual line of words and calculate the reading speed in each line of text. In this way, reading slope will be based on the reading speed for each successive lines of words. All the measurements of the time taken were done using the WavePad Sound Editor (Version 6.64 by NCH software).

5.3.2 Normalised Reading Rate (NRR2)

Apart from using the reading slope method, a faster way of assessing duration related effects is to compare the reading speed at the beginning and end of the reading paragraph. Normalised reading rate (NRR2), which compares the reading speed between 2 reading segments (beginning and end), was used to provide an alternative way to analyse the data. These two chosen reading segments denote the start (Line 2, 3, 4) and end (Line 13, 14, 15) of the reading process. Line 1 was omitted in the analysis to avoid inaccuracy in time recording between uncovering the text and the first word verbalised; Line 16 was also omitted, as there

were too few words in the line (i.e. 3-4 words) for a fair comparison. If the reading speed is greater in the second segment, then NRR2 will be > 1.0 . If there is slowing down, then the NRR2 will be < 1.0 .

5.4 Statistical analysis

All data were analysed using Statistical Package for the Social Sciences (SPSS) (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.). Kolmogorov Smirnov test showed a normal distribution for the reading speed ($p = 0.06$). Therefore, the mean and standard deviation measurements were reported. Parametric statistical analysis included the use of Pearson correlation to explore the association between reading slope and the various parameters, and Bonferroni correction was applied to determine the p-value for statistical significance. Paired t-test was used to compare the reading speed between the two segments. A probability of less than 0.05 indicated statistical significance.

5.5 Results

Out of the 100 participants, 6 participants defaulted the first appointment, and 33 participants' data were rejected because either they had a poor reading accuracy (i.e. less than 50% of words read correctly) or technical errors such as the audio having interruptions (i.e. talking halfway through the experiment). As such, the data for the remaining 61 were used for this analysis. The mean binocular visual acuity of included participants was 0.90 log MAR ($SD \pm 0.24$), and they had a mean contrast sensitivity of 0.86 log units ($SD \pm 0.34$). All analyses done in this chapter were only limited to near vision optical magnifiers, and reading performances with the use of p-EVES were excluded.

5.5.1 Reading slope

The average reading speed for all the 61 participants ranged from 9.18 to 168.37 swpm. Thirteen participants (21%) of the group missed one or more rows of words, up to a maximum of five rows of words. In this group, the missing rows of words were omitted from the reading slope (swpm /line) analysis (e.g. if a participant only reads five rows of words, the best straight line fit, and the slope will be calculated based on these five data points). Figure 5.2 provides

two examples of how these individual reading slopes were computed. The overall mean reading slope was 1.40 swpm /line (range: – 2.28 to 6.80 swpm/ line, SD \pm 1.87 swpm/ line). The mean for the overall model fit using R^2 was 0.14 (range: 0.00 to 0.51, SD \pm 0.14).

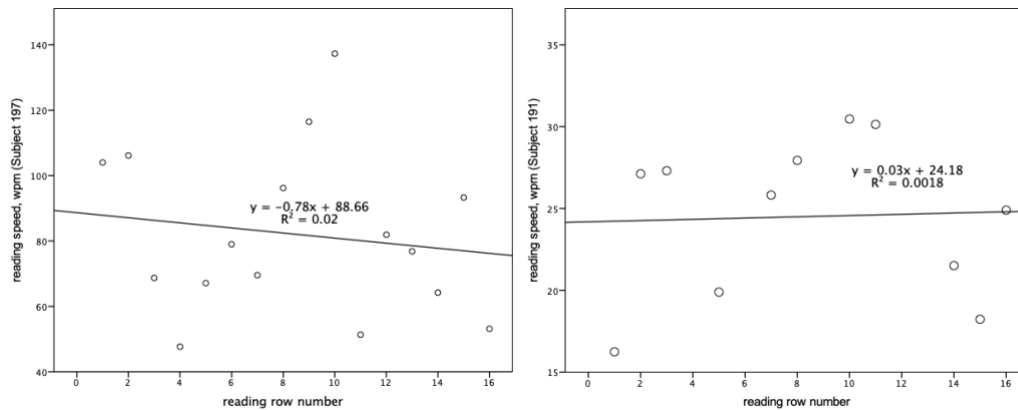


Figure 5.2 The reading slope graph plots for 2 subjects. Reading speed for each row of words is plotted against the corresponding reading row number. Subject 197 showed an overall negative reading slope of -0.78 swpm/ line. Subject 191 missed 5 rows of reading words – row 4, 6, 9, 12 and 13. Hence, these reading row numbers were omitted out. Subject 191 showed an overall positive reading slope of 0.03 swpm/ line.

Overall, there was no reduction in reading speed with time in 48 participants (79%) based on their individual reading slope (mean: 1.98, SD \pm 1.67; range: 0.04 to 6.8 swpm/ line). In fact, they showed an increase in reading speed within a single paragraph. Only 13 participants (21%) showed a negative reading slope (mean: -0.74, SD \pm 0.58; range: – 0.19 to – 2.28 swpm/ line), that is, reduced reading speed with time. To investigate if the duration related effects observed were reproducible, the reading performance data on either their second or third visit (i.e. whichever visit occurred when they were using the optical magnifiers) was used. The second and third visit assessments took place after periods of 2 months and 4 months, respectively. Out of the 61 participants, the data from 10 of these were unavailable due to dropouts from the study, corrupted audio recording or poor reading performance (e.g. two participants only managed to read three lines of text). The mean reading slope in the 26 participants who read with their optical magnifier at the first and second visits were 1.45 swpm /line (SD \pm 1.73), mean R^2 = 0.16 (SD \pm 0.14, range: 0.00 to 0.51) and 0.88 swpm /line (SD \pm 1.39), mean R^2 = 0.11 (SD \pm 0.12, range: 0.00 to 0.46) respectively. The remaining 25

participants who read with their optical magnifier at first and third visits had a mean reading slope of 1.83 swpm /line (SD \pm 2.08), mean $R^2 = 0.16$ (SD \pm 0.15, range: 0.00 to 0.47) and 1.14 swpm /line (SD \pm 1.71), mean $R^2 = 0.09$ (SD \pm 0.13, range: 0.00 to 0.61) respectively.

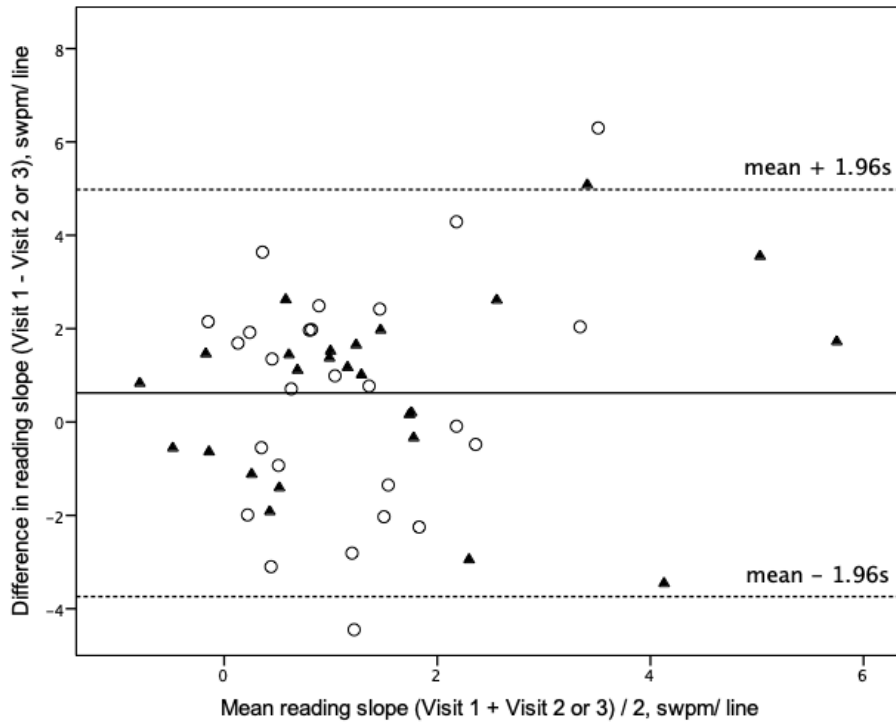


Figure 5.3 Bland-Altman plot of the difference in reading slopes between 1st visit and 2nd/ 3rd visit.
 ○ = 1st and 2nd visit, ▲ = 1st and 3rd visit.

Figure 5.3 shows a Bland-Altman plot of the individual reading slopes for all the participants in both visits. The Bland-Altman plot shows a mean difference in the reading slope of 0.62 swpm/ line (Visit 1 – Visit 2 or 3), with a 95% confidence interval of – 3.74 to 4.98 swpm/ line. Table 5.5 shows the breakdown of the 51 participants with positive and negative reading slopes in the two assessment visits. Overall, 30 participants showed positive reading slopes, and two participants showed negative reading slopes consistently in both visits. The remaining 19 participants showed a mixture of positive and negative in different assessment visits.

Table 5.5 Participants with positive and negative reading slope (n = 51)

	First Visit	
	Positive	Negative
Second/ Third visit Positive	N = 30	N = 8
Second/ Third visit Negative	N = 11	N = 2

Due to the large reading speed variability, \log_{10} transformation on the reading speed was attempted (see Appendix 9.7). However, the approach did not suggest an improvement in a linear model regression fit. Given that the low R^2 value (i.e. less than 0.50) in most of the participants, perhaps, the use of linear regression may not be the most appropriate model fitting to denote changes in reading performance with time. The use of non-linear curve fitting (i.e. logarithmic and exponential) for all the participants. The mean R^2 values for the linear, logarithmic, and exponential fittings were 0.14 (SD \pm 0.13, range: 0.00 to 0.51), 0.11 (SD: \pm 0.13, range: 0.00 to 0.55) and 0.13 (SD \pm 0.13, range: 0.00 to 0.51), respectively. Since the non-linear curve fittings did not provide a better fitting of the data points, the use of the linear fitting would be used to represent the change in reading with each subsequent rows of words. Figure 5.4 shows the reading slope graph plots for two participants with negative reading slopes in both reading slopes evaluation.

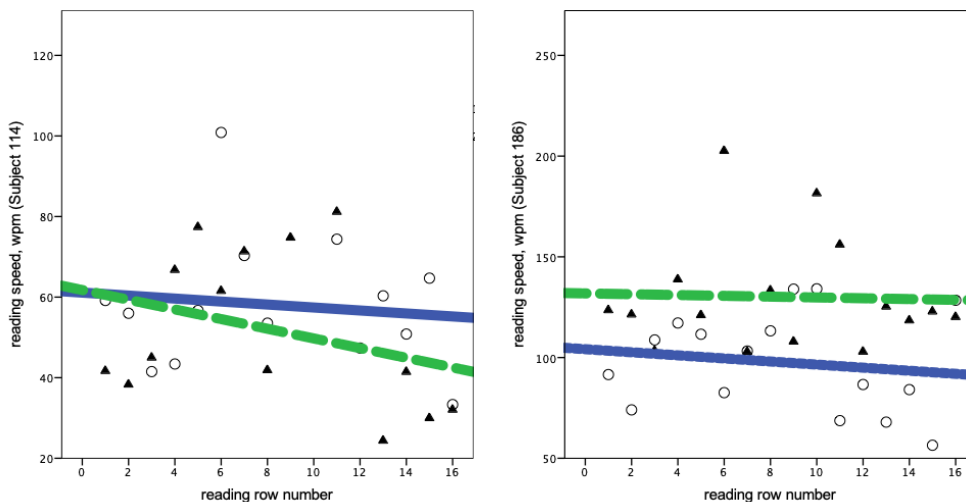


Figure 5.4 Reading slope plots for Subject 114 and 186. O: 1st evaluation and ▲: 2nd evaluation. Solid line = 1st evaluation reading slope, dotted line = 2nd evaluation reading slope.

5.5.2 Normalised Reading Rate (NRR2)

Overall, the mean normalised reading rate (NRR2) from the 61 participants was 1.13 (SD \pm 0.37, range: 0.54 to 2.50). This means that the reading speed generated from the second segment was faster compared to the first segment. Out of the 61 participants, 35 participants had NRR2 more than or equal to 1.0 (mean NRR2: 1.31 SD \pm 0.38, range: 1.00 to 2.50) and 26 participants had an NRR2 of less than 1.0 (mean NRR2: 0.87 (SD \pm 0.13, range: 0.54 to 0.99). Figure 5.5 shows two sample graph plots of the participants who had a low NRR2 value. Although the NRR2 values were low, the graph plots did not suggest that a trend that the reading speed is declining.

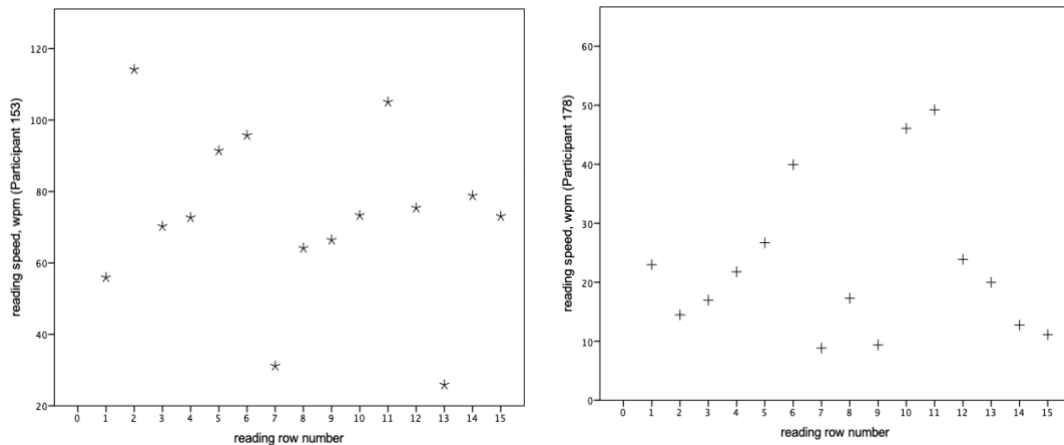


Figure 5.5 Graph plots of reading speed against row numbers. Left: Participant 153, NRR2: 0.60. Right: Participant 178, NRR2: 0.54

A paired-sample t-test was conducted for the 61 participants to evaluate the difference in reading speed between the first and the second segments. There was no statistically significant difference in the reading speed between the two segments ($t(60) = -0.69$, $p = 0.49$). There was no statistically significant correlation between the NRR2 and reading slopes ($r = 0.09$, $p = 0.49$).

5.5.3 Factors associated with reading slopes and NRR2

To investigate if there is an association between the clinical measures and characteristics and the duration related indicators, Pearson's two-tailed correlation was conducted. Table 5.6

shows the relationships between the duration related effects indicators (i.e. reading slopes and NRR2) and the clinical measurements or characteristics using Pearson's two-tailed correlation. The p-value was adjusted accordingly based on Bonferroni correction. Therefore, a p-value of less than 0.002 will indicate a statistically significant difference.

Table 5.6 Pearson's correlation coefficients between the duration related effects indicators and clinical measurements or characteristics (two-tailed) (n = 61)

	Age	CCVFT	MLVQ (duration)	MLVQ (frequency)	Reading speed	Retrace time
Reading slope	r = 0.07 p = 0.58	r = -0.11 p = 0.38	r = 0.06 p = 0.62	r = 0.06 p = 0.63	r = 0.16 p = 0.23	r = -0.18 p = 0.16
NRR2	r = 0.14 p = 0.28	r = 0.26 p = 0.05	r = -0.09 p = 0.49	r = -0.01 p = 0.96	r = -0.39 p = 0.002	r = 0.17 p = 0.19
Age						
CCVFT	r = 0.10 p = 0.47					
MLVQ (duration)	r = -0.15 p = 0.24	r = -0.28 p = 0.03				
MLVQ (frequency)	r = 0.19 p = 0.13	r = -0.08 p = 0.51	r = 0.36 p = 0.005			
Reading speed	r = -0.33 p = 0.01	r = -0.30 p = 0.02	r = 0.33 p = 0.01	r = 0.25 p = 0.06		
Retrace time	r = 0.24 p = 0.06	r = 0.25 p = 0.06	r = -0.38 p = 0.003	r = -0.38 p = 0.002	r = -0.64 ^a p < 0.001	

^a p < 0.002. Given the number of comparisons (n = 27), a Bonferroni-corrected p value of 0.05/27 = 0.002 would be appropriate.

NRR2, normalised reading rate 2; CCVFT, California Central Visual Field Test: Grading Scores 1 to 5; MLVQ: Manchester Low Vision Questionnaire: Grading scores 0 to 4 for duration of near vision optical magnifier usage

5.5.3.1 Age

The mean age of the 61 participants was 68.37 years (SD ± 18.90, range 20 to 94 years). Six participants were below 40 years, 11 participants aged between 40 and 60 years, and the remaining 44 participants were aged above 60 years. Given the diversified age range in our participant's group, there was a moderate negative correlation between reading speed and age (r = - 0.33, p = 0.01). There was, however, no correlation between the age and reading slope (r = 0.07, p = 0.58) or the age and NRR2 (0.14, p = 0.28).

5.5.3.2 CCVFT

Out of the 61 participants, 55 of them had central scotoma (i.e. defects within central 15°). Table 5.7 shows the degree of defects in these 55 participants. A moderate correlation (r = - 0.30, p = 0.02) was observed between reading speed and CCVFT grading scores. We

hypothesised that participants with macular scotomas, that is, with more difficulty in reading, would be more susceptible to duration related effects. No correlation was observed between CCVFT grading scores and the reading slopes ($r = -0.11$, $p = 0.38$), but a low correlation was seen between the CCVFT grading scores and the NRR2 ($r = 0.26$, $p = 0.05$).

Table 5.7 The number of participants with central field defects

Grading Scores	Location of defects	No. of participants (n = 61)
0	no defect within central 15°	6
1	only a defect (with either laser 2 or laser 3) outside central 10°	9
2	a defect with laser 2 between 5° -10°	16
3	a defect with laser 3 between 5° -10°	12
4	a defect with laser 2 inside 5°	12
5	a defect with laser 3 inside 5°	6

5.5.3.3 MLVQ

During their first visit, the MLVQ was administered to find out the usage profile of near vision optical magnifiers in the participant's daily life. Figure 5.6 shows the distribution of the number of participants engaged in reading tasks. Out of the 61 participants, 85% of the participants were able to use their optical magnifiers for reading instructions (i.e. spot or survival reading), and 73% managed to use the magnifiers to read a book or newspaper (i.e. leisure reading). Specifically for leisure reading, 14 participants were unable to perform the task, and two participants did not perform any reading task. Figure 5.7 shows the frequency usage of the optical magnifiers by all participants, with 70% of them using their magnifiers at least five times each day. Participants were also asked about the longest length of time that a single near vision magnifier had been used. As seen in Figure 5.8, 32% of the participants had used their near vision optical magnifier for 30 minutes on at least on one occasion.

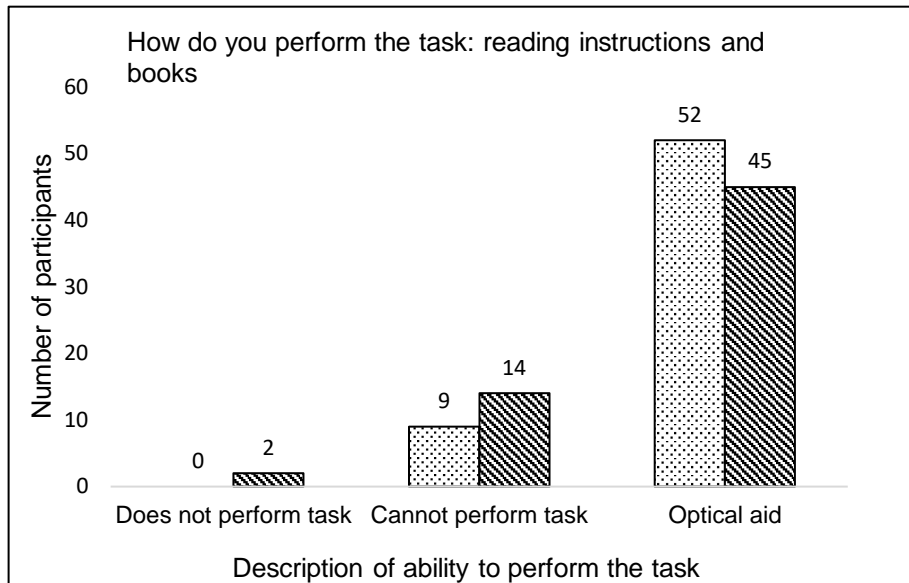


Figure 5.6 Number of participants able to perform reading tasks. The dotted-bars and striped-bars represent reading instructions and books respectively.

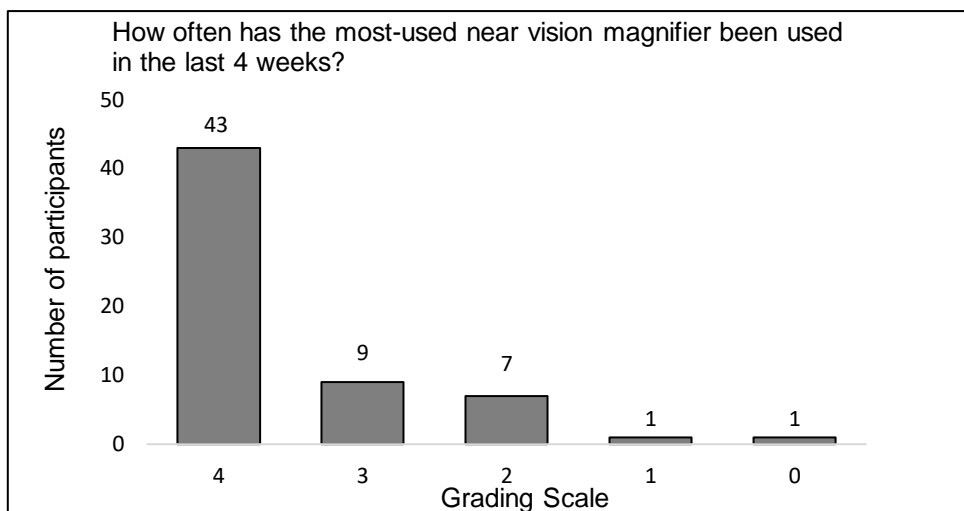


Figure 5.7 Frequency in using near vision optical magnifiers in the last 4 weeks. 4 = Uses very often = Many times (≥ 5) each day (5x or more each day). 3 = Uses often Several times (1-4) each day. 2 = Uses sometimes (< 1 daily but at least ≥ 1 per week

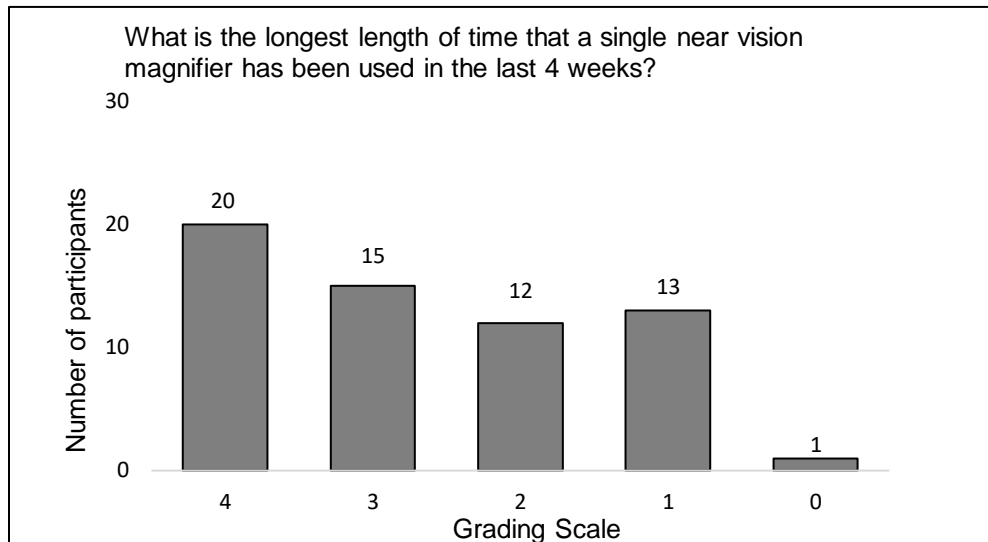


Figure 5.8 Longest duration on the usage of their near vision optical magnifier at any one occasion. 4 = ≥ 30 minutes. 3 = ≥ 15 minutes and < 30 minutes. 2 = ≥ 5 minutes and < 15 minutes. 1 = ≥ 1 minute and < 5 minutes. 0 = < 1 minute.

Earlier on, we hypothesised that participants who have been using their near vision optical magnifiers for leisure reading might have less duration related effects compared to participants who only use their near vision optical magnifiers for spot reading. Also, we would expect that these participants who were comfortable using their near vision optical magnifiers for leisure reading would achieve a higher reading speed. However, as seen in Table 5.8, there was no correlation between the duration and frequency usage of their LVAs and the duration related effects indicators (i.e. reading slopes and NRR2).

5.5.3.4 Reading speed

The reading speed was assessed using the IReST, with the participant's near vision optical magnifier. The average reading speed was 55.18 swpm (SD \pm 35.35 swpm, range: 9.18 to 168.37 swpm). A moderate and marginally statistically significant negative correlation was observed between the reading speed and NRR2 ($r = -0.39$, $p = 0.002$), indicating that participants who had higher reading speed had a lower NRR2. The correlation between the reading speed and reading slope was lower ($r = 0.16$, $p = 0.23$).

5.5.3.5 Retrace Time

Apart from the reading speed in each row of words, retrace time (Bowers and Ackerley, 1994), which is the time taken to move from the end of the row of words to the beginning of the next row of words, is also responsible for the reduced reading speed when reading with a magnifier. Overall, the average retrace time was 3.04 seconds (SD \pm 3.42, range: 0.13 to 17.31 seconds), and the time accounts for approximately 20% of the total reading time. Although the retrace time only accounts for 20% of the total reading time, a moderately high negative correlation ($r = -0.64$, $p < 0.01$) was observed between the participant's reading speed and average retrace time. The usage profile of their optical magnifiers for reading task could influence the retrace timing. We would expect users who are currently comfortable using their magnifiers for leisure reading activity to be more proficient in page navigation with the magnifier and that this would result in shorter retrace time. There was almost a statistically significant negative moderate correlation between participant's reported duration ($r = -0.38$, $p = 0.003$) and frequency usage ($r = -0.38$, $p = 0.002$) on their magnifiers and the average retrace time.

To investigate any changes in retrace time in the single paragraph, we calculated the gradient of the best straight line plotted using the retrace time against the successive rows of words. Out of the 61 participants, 31 participants showed an improvement in retrace time (i.e. shorter) within a single text paragraph suggesting an improvement in page navigation. The average retrace time slope was close to zero: -0.08 seconds /line (SD \pm 0.40, range: -2.15 to 1.53 seconds/ line). This is because half of the group got faster and the rest of them got slower, which produced an averaging out effect. It must be noted that the mean R^2 value of the linear function was 0.09 (SD \pm 0.11, range: 0.00 to 0.47), and neither the attempted use of logarithmic (mean R^2 : 0.10, SD \pm 0.11, range: 0.00 to 0.40) or exponential (mean R^2 : 0.10, SD \pm 0.12, range: 0.00 to 0.54) functions showed an improvement on the fitting on the data points. We hypothesised that participants with decreasing retrace timing slope would be associated with an increase in reading slope or NRR2. However, there was neither a correlation between the retrace time slope and reading slope ($r = 0.12$, $p = 0.37$), nor a correlation between the retrace time slope and NRR2 ($r = -0.09$, $p = 0.48$).

To investigate if participants who showed negative duration related effects shared a common characteristic, we conducted independent-samples t-tests to compare the characteristics between the participants with 1) positive reading slopes and negative reading slopes, and 2) $NRR2 \geq 1.0$ and $NRR2 < 1.0$. Table 5.8 shows the results of the independent-samples t-test on the reading slopes and NRR2. All factors, except age, did not show a statistically significant difference between participants with positive reading slopes and negative reading slopes. For the NRR2, reading speed was the only clinical measurement that was statistically significantly different between the $NRR2 \geq 1.0$ group and the $NRR2 < 1.0$ group.

Table 5.8 Descriptive statistics on independent-samples t-test comparing participants with 1) positive and negative reading slopes 2) $NRR2 \geq 1.0$ and $NRR2 < 1.0$

	Participants with Positive reading slopes n = 48		Participants with Negative reading slopes n = 13		Independent-samples t-test			
	mean	SD	mean	SD	t(59)	p	mean difference	CI
Age (years)	71.60	18.01	56.44	17.86	-2.69	0.01	15.16	3.91 to 26.41
CCVFT	2.47	1.51	2.76	1.30	-0.62	0.53	-0.29	-1.21 to 0.63
MLVQ Duration	2.62	1.19	2.76	1.23	-0.38	0.70	-0.14	-0.89 to 0.60
MLVQ Frequency	3.50	0.92	3.53	0.77	-0.13	0.89	-0.03	-0.59 to 0.52
Reading speed (wpm)	55.98	37.12	52.24	28.93	0.34	0.74	3.73	-18.55 to 26.00
Retrace timings (seconds)	3.08	3.59	2.84	2.77	0.22	0.82	0.24	-1.90 to 2.39

	Participants with $NRR2 \geq 1.0$ n = 35		Participants with $NRR2 < 1.0$ n = 26		Independent-samples t-test			
	mean	SD	mean	SD	t(59)	p	mean difference	CI
Age (years)	70.92	19.02	64.94	18.53	-1.23	0.22	-6.00	-15.73 to 3.75
CCVFT	2.71	1.41	2.31	1.54	-1.07	0.29	-0.41	-1.16 to 0.35
MLVQ Duration	2.60	1.16	2.74	1.25	0.42	0.68	0.13	-0.49 to 0.75
MLVQ Frequency	3.54	0.74	3.46	1.06	-0.35	0.73	-0.08	-0.54 to 0.38
Reading speed (wpm)	52.80	31.58	79.21	48.61	2.57	0.01	26.41	5.84 to 46.98
Retrace timings (seconds)	3.39	3.16	2.56	3.74	-0.94	0.35	-0.83	-2.60 to 0.94

5.6 Discussion

The objective of this analysis was to determine whether any duration related effect in reading performance could be identified in a group of visually impaired individuals using near vision optical magnifiers when reading a short paragraph. Recently, researchers (Brussee et al., 2016, Ramulu et al., 2013a, Stangler-Zuschrott, 1990) have begun to investigate the possibility of duration related effects changes on reading speed. Rubin (2013) has termed these effects as reading endurance, that is, the ability to sustain reading speed over a period of time. This is the first study to investigate duration related changes in reading speed in a group of visually impaired individuals using optical magnifiers for a reading task. In our study, participants with visual impairment read a single paragraph from the IReST text, which was intended to simulate a leisure reading task. Our results did not show a decline in reading speed for the majority of visually impaired individuals using their optical magnifiers. Although a small group of participants did present a reduced reading speed with time, these findings were rarely consistent during the second visit.

A previous study by Ramulu et al. (2013a) has investigated reading endurance over 30 minutes using sustained silent reading in a group of glaucoma suspect and glaucoma participants. Reading slope, which is the reading speed (words per minute) over individual minutes, was used as an indicator to denote any duration related changes in reading speed within a 30-minute period. Their results have shown a reading speed reduction with time (-0.37 wpm/min) in the group of subjects with glaucoma but not in the normally-sighted group. For the participants in the Ramulu et al. (2013a) study, the duration related changes in reading speed were found to be clinically insignificant after 15 minutes. Unlike Ramulu et al. (2013a) study, our participants were only given a single IReST paragraph, and the maximum reading duration for our group of participants lasted for approximately 9 minutes. It can be argued that the shorter reading duration may not be sufficient to elicit any duration related effects in our study. Our participants had a diversified range of reading speeds (i.e. 9.18 swpm to 168.37 swpm). Many of the visually impaired participants struggled with just a single reading IReST paragraph (i.e. missing rows of words and poor word reading accuracy). Almost 40% of the participants' data were rejected for this analysis as they were unable to carry out meaningful

leisure reading (i.e. accuracy <50%). Although the use of an extended reading text (e.g. 7300 to 7600 words in Ramulu et al. (2013a)) would be more appropriate to investigate duration related effects, this would not be practical in a group of visually impaired participants where the average reading speed is around 55.2 wpm in the current study. This is in contrast to Ramulu et al. (2013a) study, where all their participants generally had good VA (mean VA for subjects with glaucoma was 0.09 log MAR) and high reading speeds ranging from 147 to 163 wpm. In the current study, we observed that the changes in reading speed within the time duration trend investigated were in the opposite direction, suggesting an improvement in reading speed as the participant became familiar with the task. Therefore, if there is any potential duration related effects that can be identified, the assessment would need to be long enough for this “learning” effect to dissipate and be overridden by the effect of fatigue. As noted earlier, these negative reading slopes were rarely reproducible in the second visit; suggesting that the duration related effects were very subtle.

As with our analyses used in the previous chapters, the NRR2, which is a comparison of reading speed between two reading segments within a single paragraph, was included. Our results show a mix of participants with $NRR2 \geq 1.0$ (57%) and $NRR2 < 1.0$ (43%). There was a statistically significant difference in reading speed between the two groups; with participants having a lower average reading speed (52.80 swpm, $SD \pm 31.58$ swpm) in $NRR2 \geq 1.0$ compared to participants with an $NRR2 < 1.0$ (79.21 swpm, $SD \pm 48.61$ swpm). The overall negative correlation ($r = -0.39$, $p = 0.002$) suggests that the participants with higher reading speed had lower NRR2. We identified those readers who had “fast” reading speed (i.e. more than 100 swpm, $n = 13$), majority of the participants had NRR2 almost close to 1.0 (i.e. range 0.93 to 0.99); only 2 of them had NRR2 lower than 0.90 (i.e. Participant 106, reading speed: 168.37 swpm, NRR2: 0.80; Participant 132, reading speed: 160.96 swpm, NRR2: 0.85). Due to the difficulty to define what NRR2 value would indicate a clinically significant decrease in reading speed, we have adopted the use of any values less than 1.0 to indicate a drop in reading speed with time. For example, with NRR2 values of 0.50 and 0.99, both would have indicated a negative duration related effect. However, such a marginal NRR2 value (i.e. 0.99) could be contributed by the human errors (e.g. misread) and does not clearly indicate a fatigue

effect. A closer look at the graph plots (as seen in Figure 5.9) of the two participants with NRR2 value less than 0.90 did not show an overall reduction in reading speed with time within the 16 rows of words. The reduced NRR2 in these participants can be explained by the extreme lower reading speed data point, which coincides at the segment where the NRR2 was computed. In fact, the graph plots suggest that the reading speed appeared to show an overall gradual increase in reading speed within the 16 rows of words.

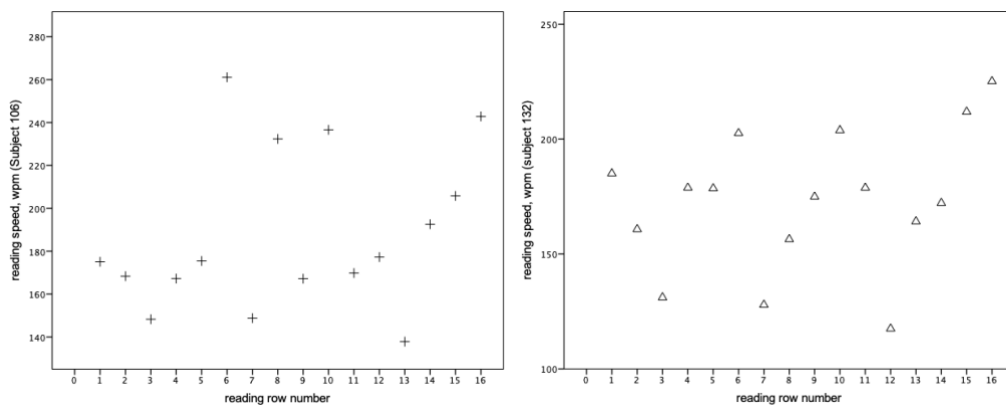


Figure 5.9 Graph plots of participants with NRR2 less than 0.9. Left: Subject 103; Right Subject 132

We did not find any association between the reading slope measurements and the various clinical measurements or characteristics. There is, however, a moderate negative correlation between reading speed and NRR2 – suggesting participants with slower reading speeds showed a larger increase in the reading speed in the second segment. A negative correlation was observed between the CCVFT grading scores and reading speed (i.e. higher the CCVFT grading scores, slower the reading speed). As such, it is not surprising that a similar negative correlation between the CCVFT grading scores and NRR2 was also observed; where the participants with higher CCVFT grading scores (i.e. more severe defect) had a faster reading speed in the second segment of the IReST as denoted by NRR2. These findings contradict our hypothesis that participants who experienced more difficulty in reading (i.e. slower reading speed) are more predisposed to duration related effects. Although the participants were experienced in using their magnifiers, it is possible that the increase in reading speed with time could be due to a learning effect from how the reading speed assessment was conducted. However, it is unclear why the effect was larger in participants with slower reading speed.

Apart from the duration effects analyses, we have also found a moderate negative correlation ($r = -0.33$, $p = 0.01$) between reading speed and age. The objective of this data analysis in this study was to investigate potential duration related effect; and there might be other confounding factors (e.g. onset of the disease, VA, nature of the visual impairment, etc.) that can contribute to the reduction of reading speed with age, but these factors were not investigated. The nature of the eye conditions with similar visual acuity can affect reading speed to varying degrees (Fine and Rubin, 1999), and patients with central scotoma generally experience more difficulties with reading (Whittaker and Lovie-Kitchin, 1993). Similarly, we have also observed that higher CCVFT gradings were moderately correlated with the reading speed. The CCVFT grading scores used in this study mainly helps to determine the location of the defects, but no information on the size and depth of scotoma were included in the grading scores. This may explain why the correlation between the scores was only moderately strong.

As noted earlier, 76% of our participants showed that when using the magnifiers, their reading speed improves within a single paragraph. When using hand-held magnifiers for reading, the forward saccade and retrace timing are the two components that determine the reading speed. From our data analysis, the improvement in the reading speed was not limited to the reduced time to saccade along the line, but also in the reduction in retrace timing. Out of the 61 participants, 31 participants showed a negative retrace timing slope (i.e. reduced retrace timing with consecutive sentences), and the overall average retrace timing slope was -0.08 secs /line. Although it was not the intention of this study to investigate the usage behaviour of their magnifiers and reported questionnaire, a moderate correlation was observed between the longest duration usage ($r = -0.38$, $p = 0.003$) and the average retrace timing. This suggests that participants with a reported history of using their magnifiers for longer periods (i.e. around 30 minutes) tend to navigate their magnifiers more efficiently between rows of words, resulting in reduced retrace timing. This observation on the negative correlation between reported usage of magnifiers and retrace time was also in line with a previous study by Bowers et al. (2007). In their investigation on the influence of page navigation using the hand-held and stand magnifiers, they reported a non-significant negative correlation between

self-reported navigation difficulties and retrace time ($r = -0.20$, $p > 0.1$) in a group of 43 visually impaired participants.

5.7 Limitations

In this study, due to a shorter reading duration captured using a single IReST paragraph, the gradient slope from the best straight line analysis was modified using reading speed against rows of reading text. In Chapter 3, we have described the limitations of using the reading slope to indicate changes in reading speed with time. Here, we extended the analysis on the reproducibility of the observed duration effects in a group of visually impaired individuals. Given a large coefficient of repeatability (CR: 4.36, 95% CI: -3.74 to 4.98 swpm/line) of the reading slopes values may suggest a possibility that these reading slopes were random values. If this is true, a similar correlation between the first and second assessment would be the same even though the reading slope values on the first visit were jumbled up. Although Pearson showed a low correlation ($r = 0.23$, $p = 0.10$) between the first and second reading slopes assessment, a completely different correlation ($r = -0.12$) was observed when these reading slope numbers were deliberately jumbled up within each visit. The difference in correlation suggested that these reading slopes were not just random values, and the values calculated from these reading slopes were representative of a true effect. Likewise with the analysis in Chapter 3, the presented graph plots in Figure 5.4 from the two participants with consecutive negative reading slopes in both visits did not suggest a better curve fitting would be more representative in our findings. The visual inspection of the remaining graph plots showed no typical curve fittings for the data points (see Appendix 9.8). The assessment of reading performance parameters from the low vision community can be challenging due to a potentially large variation in the measurements (Burggraaff et al., 2010). Indeed, the majority of our participants showed a low R^2 value suggesting that the reading speed performance varies extensively, and most of each data points did not fall close to the linear function slope. Also, we have explored other possible analysis such as the use of NRR2 (i.e. comparison of reading speed between the beginning and end segments of a reading text). If the exponential function graph is true, then the use of NRR2 would provide an indication (i.e. less than 1.0) on reading speed changes. However, this was not seen in our NRR2 analysis. Perhaps, the

testing duration was not adequate for any effect to be detected. Although Ramulu et al. (2013a) suggested a testing duration of 15 minutes may be adequate, this is hard to achieve for patients with more severe vision impairment (i.e. macular dysfunction). Given that no negative duration related effect was observed in the analysis, it was difficult to determine which analysis approach would be the most appropriate. Nonetheless, the data was comprehensively analysed, and no findings were observed to suggest a negative duration related effect was present.

5.8 Conclusion

In this study, the duration related effects in a group of visually impaired participants using optical near vision magnifiers for reading using the reading slopes and NRR2 assessment parameters were investigated. Although the reading slopes allow the assessment to be quantifiable, the use of reading slopes may be limited by the large variation in reading speed in visually impaired individuals and large limits of agreement on reproducibility testing. As such, a comprehensive assessment of the reading changes should include the visual inspection of the graph plots. Based on the analyses, we found no duration related decrease in reading speed based on the assessment of reading slopes and NRR2 parameters. In fact, these parameters suggest that the reading speed appears to increase with time.

6. The use of spontaneous eye blink rate (SBR) to capture fatigue effect in reading

In the previous chapters, we have investigated the duration related effects on reading speed. Although most of the participants from the simulated visual impairment and the true visual impairment group did not show a reduction in reading speed with time, there was a small number of participants in both groups that did show a decreased reading speed with time. As noted earlier in the introduction in Chapter 1 - Section 1.1, the decrease reading speed with time could be a consequence of mental fatigue. In an early study investigating the effect of time-on-task on blinking patterns, Stern et al. (1994) defined mental fatigue as a decline in the performance on tasks requiring cognitive involvement. From their study, the authors infer that the increase in blink rate could be used as a sensitive indicator of mental fatigue. Fatigue per se is very subjective and can be hard to quantify accurately. This experiment, therefore, explored the possibility of using blinking rates to index mental fatigue objectively during reading.

6.1 Literature review

6.1.1 Introduction

Blinking, the swift motion of the eyelids that closes and opens the palpebral aperture plays an essential role in maintaining corneal integrity (Carney and Hill, 1982). The movement involves the activation of the orbicularis oculi muscle to lower the upper eyelids, followed by the activation of the levator palpebrae superioris and Muller's muscle to elevate and hold the upper eyelids. There are three different types of blinks; voluntary blinks and involuntary blinks, and involuntary blinks can be subdivided into reflex and spontaneous blinks.

6.1.2 Types of blink

Voluntary blinks are self-initiated, resulting from a decision by the individual. On the other hand, involuntary blinks are uncontrolled blinks which can be either triggered by an approaching stimulus (i.e. reflex blinking) or in the absence of an external stimulus (i.e. spontaneous eye blinks (Zaman and Doughty, 1997)). Reflex blinking acts as a protective mechanism to prevent or reduce any incoming hazards to the eye. Unlike reflex blinking, the spontaneous eye blink occurs when there is no identifiable direct stimulus.

6.1.3 Assessment of blinks

Expensive and complex instruments such as electro-oculography (EOG) and electromyography (EMG) have been used by researchers to assess blinking rates. The principle of EOG (Stern et al., 1984) is based on the different charges on the cornea, which is more positively charged compared to the fundus. The closing of the eyelid generates less positivity than the retraction of the eyelid, resulting in the change of charge, which is registered as deflections on the display screen. The deflections can then be used to calculate the number of blinks during the experiment. On the other hand, EMG (Sheedy et al., 2005, Nahar et al., 2007, Gowrisankaran et al., 2012) records the activity of the orbicularis oculi muscles, which are responsible for lid closure. The electrodes are taped to the lower eyelid for better comfort with minimal restriction on the blink pattern (Sheedy et al., 2005) to detect frequency signals. By differentiating the frequency, the investigators can compute the number of blinks in the testing session. Apart from the use of such highly technical equipment, other simpler alternatives to capture the number of blinks have also been used in other research studies. This includes a kymograph (York et al., 1971) – a rotating cylindrical drum use to track timing events by a stylus during real-time observation, or simply by counting the blinks through real-time observation (Luckiesh and Moss, 1939, Cho et al., 2000) or video recordings (Doughty, 2016, Sheedy et al., 2005, Doughty, 2002, Doughty, 2001, Cho et al., 2000)

Blinking can be assessed in terms of the amplitude, velocity, interblinking interval, blinking rate, etc. (Cruz et al., 2011). Specifically, spontaneous blinking rate (SBR), which is the number of blinks in a minute, is the blinking parameter that has been consistently reported to be associated with mental fatigue (Martins and Carvalho, 2015). Therefore, the investigation of SBR will be our primary focus of blinking assessment in this study. The video camera method is chosen to capture SBR because it has been used in previous research studies on SBR assessment (Doughty, 2016, Sheedy et al., 2005, Doughty, 2002, Doughty, 2001, Cho et al., 2000). The experiment set-up is more straightforward to implement (i.e. mount a small video camera on the trial frame) and it can be done without the participant being aware that their blinks are being monitored. Therefore, there is less likely to be a stimulus to blink voluntarily.

6.1.4 Factors affecting SBR

There are intrinsic and extrinsic factors that can affect SBR. The intrinsic factors include age, dry eyes, sleep deprivation, anxiety, depression, psychosis, etc.; the extrinsic factors include temperature, humidity, types of activities, duration of the task, etc. In relation to our current study, only age, type of activities, visual and cognitive tasks, and duration of the reading task will be discussed in this section.

6.1.4.1 Age

To date, there are limited studies investigating the relationship between age and SBR. Zametkin et al. (1979) investigated the SBR in a group of participants aged 2-months to 60 years. The authors reported that the blinking rates increase from 2 months old to 20–25 years old, and remains constant after 25 years old. However, the biased number of participants in each age group (i.e. 20–25 years, $n = 27$; >60 years, $n = 7$) does not provide a good representation of the older age group. A better distributed number of participants was seen in the study of Sun et al. (1997). The authors observed a non-statistically significant increase in SBR between the different age groups (age: 40 to 49, mean SBR: 23.5 blinks/min vs age: 80 to 89 years, mean SBR 31.3 blinks/min). Although the number of participants per age group was evenly distributed, the small sample size ($n=8$ per age group) may have limited the statistical power of the study to find a difference between the groups.

A more recent study by Doughty and Naase (2006) investigating SBR in primary gaze had observed a trend of decrease SBR with age. However, this age-related effect did not reach statistical significance. Although the sample size for their study was large ($n=100$), most of their participants' ages ranged between 20-40 years old (i.e. only about 12 participants were above 40 years old). From the perspective of the anatomical structure of the eye, age-related ptosis can occur in elderly individuals, resulting in a smaller exposed ocular surface compared to a young adult. This can potentially contribute to a reduced blinking rate in elderly individuals. The blinking investigation studies mentioned above were limited to general activities such as waiting in the clinics, sitting in church, Sunday school or waiting rooms (Zametkin et al., 1979), casual conversation (Sun et al., 1997) or primary gaze (Doughty and

Naase, 2006). In addition to the anatomical structure difference between ages, the type of activities involved in the assessment can indirectly cause a change in SBR due to the interaction of the task with the age differences. For example, an additional component such as the cognitive processing ability involved in a reading task may be different between the young and elderly (Wood et al., 2010, Wood et al., 2009), contributing a confounding factor in the change in SBR.

6.1.4.2 Types of activities

The influence of different activities such as reading, watching videos, primary gaze and casual conversation, on SBR has been reported in numerous studies (Gowrisankaran et al., 2012, Cho et al., 2000, Bentivoglio et al., 1997, Tsubota and Nakamori, 1993, Telford and Thompson, 1933, Hall, 1945). Table 6.1 provides detailed information from several studies on the comparison of SBR between different activities and reading task. When comparing activities involving visual inputs (i.e. reading, watching videos, etc.) to non-visual activities (e.g. listening to an audio story, conversation, etc.), SBRs were reduced in the former activities. Large individual SBR variation has been reported previously in primary gaze (6.9 to 21.8 blinks per minute in Doughty (2016)). Likewise, a large variation in SBR-reading tasks was also observed amongst the studies in Table 6.1, ranging from 2.8 blinks/min to 18 blinks/min.

Reading involves two key elements, the visual and cognitive process, which are not mutually exclusive. From the visual processing perspective, it is not surprising to understand why SBR can be affected in a reading task. Blinking involves eyelid movements, which can disrupt the visual information input process. The term cognitive process encompasses the need for the reader to devote their attention and process the visual information during reading (Britton et al., 1978). These two components could contribute to the inhibition of the blinking mechanism simultaneously or independently.

Table 6.1 Summary of studies comparing spontaneous blinking rate (SBR) in reading and other activities

Authors	Methods	Age	N	Mean SBR (blinks / minute)			
		(years)		Reading	SD	Other activities	SD
Hall (1945)	Observe and event marker	NA	16	2.8	NA	Conversation: 19.8	NA
Telford and Thompson (1933)	Observe and counting	NA	36	2.9	0.54	Mental Arithmetic: 10.9 Conversation: 8.9	0.80 0.32
Tsubota and Nakamori (1993)	NA	20-69	104	Hardcopy: 10 VDT: 7	6 7	Relaxed: 22	9
Bentivoglio et al. (1997)	Videography	mean= 36.9	150	4.5 (95% CI: 0.7 to 22; 90% CI: 1 to 16)	NA	Conversation: 26 (95% CI: 11 to 53; 90% CI: 13 to 47) Rest: 17 (95% CI: 4 to 48; 90% CI: 6 to 40)	NA
Cho et al. (2000)	Videography	19-25	49	PG: 17.5 DG: 9.4	12.8 11.1	Conversation: 28.6 VA taking: 14.2 Sitting on the consultation chair: 21.2	11.9 7.6 11.6
Gowrisankaran et al. (2012)	EMG	18-35	33	16 to 18	NA	Video: ~23	NA

NA: Not available – data or information not reported in the publication

VDT = Video Display Terminal

PG: Primary Gaze; DG: Downward gaze

CI: Confidence Interval

6.1.4.3 Visual and cognitive tasks

In the earlier section, we have discussed that reading tasks can cause a reduction in blinking rates compared to other non-visual activities. It appears that the decrease in the blink rates in reading is not restricted to visual tasks per se, but an additional factor of cognitive load. Cognitive load is defined as the amount of attention required to perform a specific task (Maffei and Angrilli, 2018). Studies are suggesting that the combination of visual processing and cognitive load tasks can further reduce blinking rates. The reduction in blinking was reported in tasks requiring higher mental concentration, such as performing a surgery vs casual conversation (Wong et al., 2002). Other studies (Poulton and Gregory, 1952, Drew, 1951) suggested that in a more “difficult” visual tracking task (e.g. visual tracking of straight line vs oscillating curve), participants were presumably to focus and devote more attention in the task and result in lower SBR. More recently, Maffei and Angrilli (2018) conducted a vigilance test using the Mackworth Clock test - a psychological vigilance test where a light signal moves in a circle and the participant has to press a button when there is a double jump of the light signal. The difficulty of the task was increased by reducing the time intervals between the moving light, as it requires the participant to be more alert to pick up the “double jump” signal with a shorter time interval. The participants were tested over three rounds using different time intervals (i.e. easy: 500 msec, medium: 350 msec and hard: 200 msec). Their results showed a statistically significant reduction in blinking rate in the “hard” condition compared to the “easy” and “medium” condition, which suggest that the blinking rates can reduce with increasing attention demand. Although these studies reported an associated increase in SBR with the visual tasks, the role of the “attention demand” on the visual demanding tasks remains poorly understood.

So far, the studies described above have been limited to performing a single visual task. Recarte et al. (2008) conducted an experiment investigating SBR under dual tasks condition. In their study, participants were engaged in each task (i.e. listening, talking about what they had listened to, calculating, and a control condition – no cognitive involvement) lasting for 2 minutes, together with and without an additional visual detection task (i.e. detecting a specific letter from randomly presenting letters on a computer screen). Their results showed that the

SBR decreased in dual tasks condition (i.e. participant performed both the cognitive and visual demand tasks simultaneously) compared to the single-task condition (i.e. the participant only performed the cognitive task such as listening, talking or calculating). The dual tasks could be analogous to visually impaired individuals who are required to manipulate and navigate with their LVAs when performing reading tasks. As such, we hypothesised that the participants would show a significantly lower SBR with the use of a hand-held magnifier under reduced vision simulated conditions when compared to their habitual reading conditions.

6.1.4.4 Duration of the task

Several studies have attempted to investigate the effect of time-on-task on SBR involving different activities, and an increase in SBR with time has been reported. These activities include 1 to 4 hours of reading (Luckiesh, 1947, Hoffman, 1946), a 30-minute driving (Stern et al., 1976, Pfaff et al., 1976), and a 2-hours Mackworth Clock test (Carpenter, 1948). More recently, Maffei and Angrilli (2018) have also investigated the blinking rates using a simplified version of the Mackworth Clock test, where the test lasted for 7 minutes. Their results showed an increase in blink rates in the second half of the testing duration (i.e. 4th minute of the test). The authors proposed that a “fatigue” effect could have caused an increase in the blinking rates. However, the study did not provide information on whether the participants reported fatigue during or after the 7-minute test. There are very limited studies on the effect of duration related on reading tasks. In a 4-hours reading duration, Hoffman (1946) tracked the number of blinks occurring in nine 5-minute samples spread across the four hours (i.e. first 5-minute, then last 5-minute of each 30-minute). The authors reported a gradual increase in SBR with time from 6.8 blinks/min to 11 blinks/min. Similarly, increased SBR was also observed in Luckiesh (1947) (i.e. 1-hour reading – from 7.2 to 9 blinks/min) and Tinker (1946) (i.e. 20-minute reading – from 4.8 to 6.6 blinks/min). However, these observations of increased SBR with time were limited by the statistical analyses used, which were not explicitly reported in their papers. On the other hand, Kim et al. (1984) did not observe any change in SBR in a 1-hour reading task. The difference in findings could be due to the difference in methodology. The total duration of the reading task lasted for an hour, but a 1-minute rest was given to the participant with every 5-minute reading. Also, the reading texts used in their experiment had a

different reading material presentation (i.e. blurry text, clear text and clear text which participants were instructed to read slowly) within the 1-hour testing duration. Although a recent study by Chu et al. (2014) investigating SBR in a reading task showed a graph on SBR plotted as a function of time, it was not indicated if there are any changes in SBR after a 20-minute out-loud reading session. Their objective was to compare the mean SBR between the computer screen and hardcopy reading. Therefore, no information on duration related effects were reported. However, based on only the visual inspection from the graph, it appears that the SBR remains relatively constant throughout the 20-minute out-loud reading task.

In an investigation on the use of SBR as an indicator of mental fatigue conducted by Stern et al. (1994), the authors defined mental fatigue as the decline in task performance involving cognitive activity. Previously, Summala et al. (1999) have reported that an increase in SBR was associated with self-reported mental fatigue during an overnight driving task. Some of the above studies (e.g. 30-minute driving (Stern et al., 1976, Pfaff et al., 1976), a 2-hour Mackworth Clock test (Carpenter, 1948)) reporting increased SBR with time-on-task were repetitive tasks, which can inevitably be attributed to both boredom and fatigue. It is possible that the use of a reading task might be more appropriate to provide a more engaging task compared to a repeated vigilance test (e.g. reaction to press a button). In studies (Luckiesh, 1947, Hoffman, 1946) that showed an increase in SBR on extended reading duration, there was no indication if there was any decline in the performance of the task. The purpose of this experiment is to investigate if there is any consistent change pattern in SBR with time. If this does occur, some of the participants may be experiencing mental fatigue in the reading task – which in such cases, we can investigate if there is an association between the SBR and reading slope. The finding of a negative reading slope (i.e. a slowing of reading speed with time) could then be suggested as an indication of fatigue.

6.2 Methods

6.2.1 Participants

Sixty participants were recruited for this study, 30 participants in both the young adults' group (age range 18-25 years) and the elderly group (age range 40-70 years). All participants were

generally healthy with no self-reported serious eye conditions. The data collection for this experiment was done in Singapore (i.e. different data compared to the results used in Chapter 3), and the reading speed data for the elderly have already been presented in Chapter 4. The study was approved by the Singapore Polytechnic Ethics Research Committee (Reference no: 201402-04) and written informed consent was obtained from all participants.

6.2.2 Procedures

The sustained reading text (SRT) was used in this experiment (The details of the SRT can be found in Chapter 3.2.2). The 36-pages of the SRT created an approximately 15-minute silent reading task. So far, the effect of increase SBR with time has only been reported in longer reading durations (i.e. 1 hour to 4 hours (Luckiesh, 1947, Hoffman, 1946)). The objective of this experiment was to explore the use of SBR as an indication of fatigue effects, and an assessment period of 10 to 15 minutes would seem to be more feasible in a clinical setting. Previously, Ramulu et al. (2013a) have reported that a decrease in reading speed can be observed within the first 15 minutes of the reading test. Since the current reading tests were approximately 10 to 15 minutes, the assessment of any SBR changes was also limited to the same duration of time. The reading text was placed 25 cm away from the participant on a reading stand, and approximately 30° below the eye level to simulate normal reading in downward gaze. Variable gaze position has previously been shown to affect the blinking patterns due to the area of cornea exposed (Cho et al., 2000, Doughty, 2014). Therefore, the use of the reading stand was to provide a consistent and typical reading condition (i.e. downward gaze) for all the participants.

The temperature and humidity of the room were not measured (Zaman and Doughty, 1997), but were kept constant for all the participants throughout the tests. Previous studies (De Padova et al., 2009, Barbato et al., 1993, Yap, 1991) have advocated allocating adaptation time for the participant before any blinking assessment. Although there was no specific allocated time for adaptation, the process of explaining to the participant on the nature of the experiment and procedures (i.e. approximately 10 minutes) would have been sufficient for the participants to adapt to the environment.

Blinking movements were recorded by attaching a small video camera (Yong Shi Mini Camera YS-203CA-1) to the right inferior temporal side of the trial frame, as seen in Figure 6.1. Blinking action of only one eye was captured as most of the blinks occur identically in both eyes (Becker and Fuchs, 1988). As the awareness of one's eye blink measurements may (Collins et al., 1987, Collins et al., 1989, Drew, 1951) influence the blinking patterns, participants were only informed that the camera was used to track their eye movements during the reading test.



Figure 6.1 Video camera attached to the trial frame

6.2.3 Conditions

Participants used their optimal refractive correction and were instructed not to wear contact lenses during the experiment. The elderly group used their reading correction. All participants performed the reading task in the three different reading conditions in the sequence of normal (NOR), simulated with Bangerter occlusion foil (BOF) and simulated with a denser Bangerter occlusion foil with the use of a hand-held magnifier (MAG). In the attempt to create a visually degraded image, Bangerter occlusion foil was used to create a reduction in print contrast and yet keeping a high fluent RA reserve of 2:1 in BOF, and in MAG with the help of a hand-held magnifier. However, maintaining this acuity reserve inevitably creates a variable reduced contrast reserve between individuals. The text samples were printed using the Hewlett Packard Laser Colour Printer and had an approximate contrast of 90% (Rumney, 1995, Whittaker and Lovie-Kitchin, 1994). The participant's contrast threshold can then be calculated (Mantjarvi and Laitinen, 2001) from the Mars letter contrast sensitivity.

For example, subject scored 1.76 log units on Mars chart.

$$1/10^{1.76} = 0.0174 \text{ (1.74\% contrast threshold)}$$

To calculate the contrast reserve ratio, contrast of the sample text (i.e. 90%) is divided by the contrast threshold (i.e. 1.74%), which gives a ratio of 51.79:1. The full details of the respective conditions can be found in Chapter 3.2.3.

6.2.4 Data analysis

6.2.4.1 Spontaneous blink rate (SBR)

Blinking activity captured by the video camera was projected on a monitor screen, and the number of blinks was counted by lapsing the time using an online stopwatch as seen in Figure 6.2. The number of “splits” generated at the end of the reading reflects the number of blinks which occurred in successive minutes throughout the reading session – this is defined as the SBR. In this study, an eye blink is defined as any downward movement and upward movement of the eyelids discernible by the examiner (Abelson and Holly, 1977), which includes full and partial blinking. A single blink was counted if it had a time interval of one second between the second blink. In the event of blink flurries (i.e. multiple blinks happening consecutively), this was recorded as a single blink. To minimise the variation between different observers, the computation of the number of blinks was only performed by one observer.

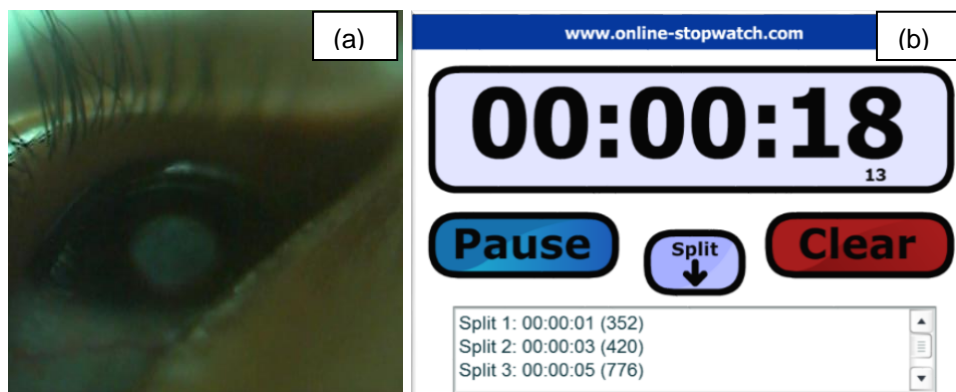


Figure 6.2 (a) Screen-shot of the video capturing eyeblink (b) using online stopwatch for recording blinking occurrences

To identify any potential duration related changes in SBR, three analytical approaches were used in this experiment: SBR slopes and normalised SBR.

6.2.4.2 Spontaneous Blinking Rate (SBR) Slopes

SBRs were calculated based on each successive minute, that is, the number of blinks occurred within the 1st minute, 2nd minute, etc. A best linear straight line was plotted based on SBR in each successive minute. The gradient of this slope would then be used as an indication for any duration related changes within the 15 minutes of reading. For example, a positive slope indicates that there is a general increase in SBR over the 10 minutes reading time and vice versa. This gradient slope would be termed as SBR slope in this study.

6.2.4.3 Normalised SBR

As SBR can potentially be highly variable among different individuals (Doughty, 2016), an alternative way of analysing the data would be to normalise the SBR. To normalise individual SBR, the second minute of the SBR was used as a baseline value; and the SBR of the successive minutes were compared to the second-minute SBR. A median normalised SBR can then be calculated.

6.2.4.4 Reading speed and reading slopes

As noted earlier, mental fatigue would involve a decline in task performance. In this case, the assessment of the reading speed performance using reading slope (full details in Chapter 3.2.4) would give an indication if the task ability is affected.

6.3 Statistical analysis

All data were analysed using Statistical Package for the Social Sciences (SPSS) (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp). Shapiro-Wilks test showed a normal distribution of the blinking rates for the young group ($p = 0.19$) but not for the elderly group ($p = 0.002$) and using log units to achieve a normal

distribution was unsuccessful. As such, non-parametric statistical analyses were used. A probability of less than 0.05 indicated statistical significance.

6.4 Results

Out of the 60 participants, 54 participants' data were used for the analysis. The remaining 6 participants' data were not used due to the misalignment of the video camera during the reading process, resulting in an incomplete assessment of the blinking movements. There were 28 participants in the young group (median age: 19.0 years, IQR: 0.0 years) and 26 participants in the elderly group; Out of the 26 elderly participants, there were 15 middle-aged (median age: 50.00 years, IQR: 11.0; range 41 to 58 years) and 11 elderly adults (median age: 63 years, IQR: 3; range 62 to 68 years). The median near VA and contrast sensitivity were the same for both the middle-aged and elderly adults group (i.e. near VA: 0.10 log MAR (IQR: 0.10; contrast sensitivity: 1.76 log units (IQR: 0.00)). As expected, Mann-Whitney U test did not show a statistical significant difference in the near VA ($z = -0.17$, $p = 0.87$) and contrast sensitivity ($z = -0.37$, $p = 0.71$) between the middle-aged and elderly adults group. Although the reading duration was longer in MAG (i.e. 15 minutes), all the blinking assessment was done only for the first 10 minutes in each reading condition; because all participants had a minimum reading time of at least 10 minutes in all the reading conditions (i.e. NOR, BOF and MAG). The contrast reserve for the young participants were 51.79 (IQR: 0.0):1, 41.03 (IQR: 3.61):1 and 5.41 (IQR: 3.09):1 in the NOR, BOF and MAG respectively; For the elderly participants, the contrast reserve were 51.79 (IQR: 0.0):1, 35.83 (IQR: 3.46):1, 5.68 (IQR: 1.89):1 in NOR, BOF and MAG respectively.

6.4.1 Intra-observer repeatability of the computation of the SBR

Although the assessment of the blinking rates has been limited to one rater in the current study, the large variation in the SBR could be attributed due to the way that the blinks were counted. To assess the intra-observer repeatability of the computation of the SBR, a second attempt on counting the number of blinks was conducted by the same rater on the 28 young participants in NOR. The second attempt was conducted more than a month after the first attempt on the counting of the number of blinks. Figure 6.3 shows a Bland-Altman plot of the

SBR tally on two attempts by the same rater for the 28 young participants in NOR. The Bland-Altman plot shows a mean difference of 8.62 blinks (Attempt 2 – Attempt 1), with a 95% confidence interval of 20.93 to 37.36 blinks. There were three data points outside the 95% confidence interval with a higher total number of blinks (i.e. > 200 blinks within the 10 minutes).

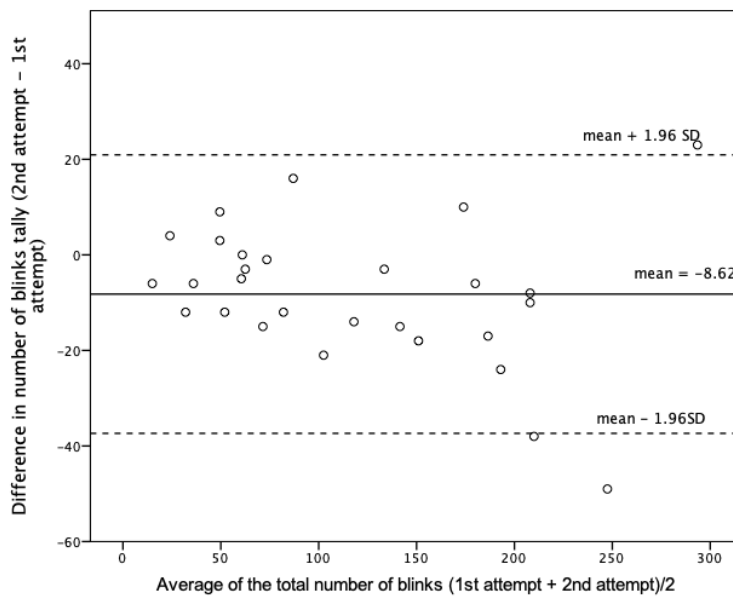


Figure 6.3 Bland-Altman plot of intra-observer repeatability

6.4.2 Spontaneous blink rate (SBR)

Table 6.2 shows the descriptive statistics on the SBR for the young, middle-aged (40 to 60 years) and elderly adults (> 60 years) group. The large individual variation of the SBR causes a large interquartile range, with the largest range of SBR (range: 0.60 to 29.20) observed in the elderly adults group under NOR. Mann-Whitney U test shows no statistically significant difference in the SBR between the middle-aged and elderly adults' group in NOR ($z = -0.02$, $p = 0.98$), BOF ($z = -0.65$, $p = 0.52$) and MAG ($z = -0.68$, $p = 0.50$). Since there was no difference in SBR between the middle-aged and elderly group, the subsequent analyses will combine the participants from the middle-aged and elderly group.

Table 6.2 Descriptive statistics on the SBR for the young and elderly group

	Conditions		
	NOR	BOF	MAG
Young (n = 28)			
Median	8.35 (12.43)	11.00 (11.93)	4.40 (7.08)
Range	1.80 to 28.20	3.20 to 34.70	1.90 to 26.0
Middle-aged (40 to 60) (n = 15)			
Median	5.30 (8.50)	7.10 (8.20)	4.50 (6.10)
Range	0.50 to 23.60	1.00 to 28.10	0.30 to 14.80
Elderly (>60) (n = 11)			
Median	6.90 (11.0)	5.80 (11.30)	2.60 (5.40)
Range	0.60 to 29.20	0.50 to 26.40	0.50 to 8.10

The Friedman test shows a statistically significant difference in SBR between the three different reading conditions in both the young ($\chi^2 = (2, 28) = 22.57, p < 0.001$) and elderly ($\chi^2 = (2, 26) = 25.56, p < 0.001$) group. The post hoc Wilcoxon Signed-Rank did not show a statistically significant difference in the SBR between NOR-BOF condition in young ($z = -1.17, p = 0.24$) and elderly ($z = -1.58, p = 0.12$). However, a statistically significant difference in SBR was observed between the NOR-MAG condition in the young ($z = -3.86, p < 0.001$) and elderly group ($z = -3.46, p < 0.001$). When comparing the SBR between the young and elderly group in each condition, the Mann-Whitney U test showed a statistically significant difference only in the NOR ($z = -2.12, p = 0.04$) but not in the BOF ($z = -1.84, p = 0.06$) and MAG ($z = -1.62, p = 0.11$).

6.4.3 Normalised-SBR relative to habitual reading condition

Due to the considerable variation in the SBR data, the comparison between individuals SBR across the young and elderly in the three reading conditions may not be the most appropriate approach. Another analysis approach would be to use the habitual-SBR (NOR) as a reference point, and the relative ratio would represent the amount of change in BOF and MAG. Table 6.3 shows the median relative ratios of the SBR-BOF and MAG. The Wilcoxon Signed-Rank test shows that all, except the young adults' group in BOF, showed a statistically significant difference in SBR relative ratios compared to a ratio of 1.0. The results showed a median increase of 10-25% in SBR-BOF, and a median reduction of up to 40% in the SBR-MAG for both the young and elderly group. A Mann-Whitney U test revealed no statistically significant

difference in the relative ratios between the young and elderly in the BOF ($z = -0.70$, $p = 0.49$) and the MAG ($z = -0.23$, $p = 0.82$).

Table 6.3 Median relative ratio of SBR–BOF and SBR–MAG compared to SBR–NOR

	NOR-BOF	Mann-Whitney U test	NOR-MAG	Mann-Whitney U test
Young	1.12 (0.75) Range: 0.5 to 2.8	$Z = -1.62$ $p = 0.11$	0.61 (0.68) Range: 0.1 to 1.3	$Z = -3.67$ $p < 0.001$
Elderly	1.23 (0.90) Range: 0.6 to 3.4	$Z = -2.36$ $p = 0.02$	0.59 (0.40) Range: 0.2 to 3.0	$Z = -3.35$ $p < 0.001$

Interquartile range (IQR) values are shown in the parentheses

6.4.4 Duration related effect

Figure 6.3 shows the box plot of the number of blinks in successive minutes for both the young adults and elderly participants in NOR, BOF and MAG. From the graphs in Figure 6.3, there was no obvious changes (i.e. general increase or decrease) in SBR throughout the 10-minute in NOR, BOF and MAG. To capture any changes in SBR over the 10-minute period quantitatively, SBR slopes were calculated.

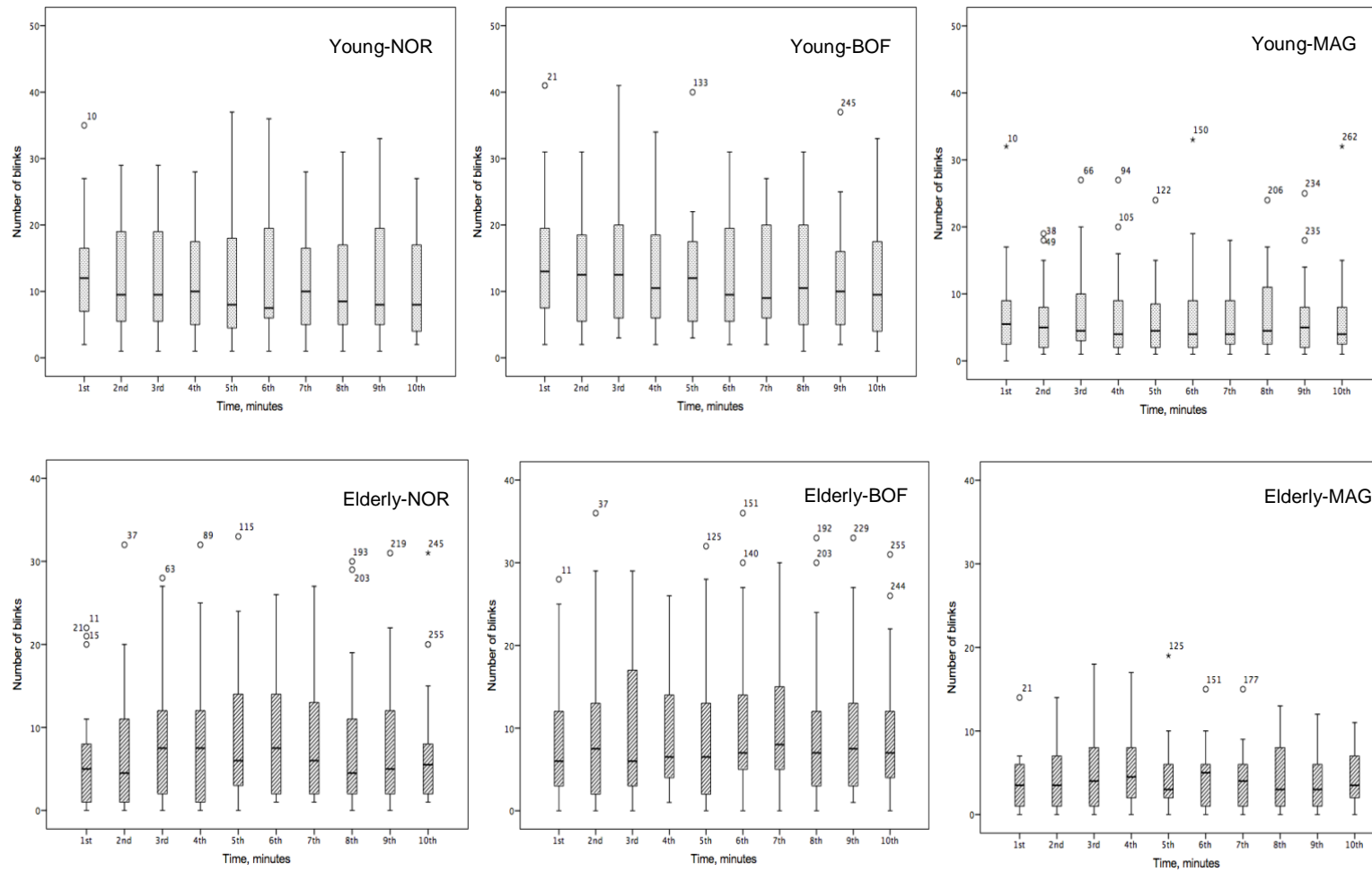


Figure 6.4 The number of blinks in successive minutes for young adults and elderly group in NOR, BOF and MAG for a reading duration of 10 minutes. Top row shows the data for the young participants for NOR, BOF and MAG (left to right). The bottom row shows the data for the elderly participants for NOR, BOF and MAG (left to right).

6.4.5 SBR Slopes

Table 6.4 shows the descriptive statistics for the SBR slopes for both the young and elderly group. Based on our hypothesis, a “fatigue” effect would cause an increasing trend in SBR with time, and this would suggest a positive SBR slope. The median blinking slopes in the young participants were – 0.14 SBR/min in NOR, – 0.24 in BOF SBR/min and 0.00 SBR/min in MAG. Out of the 28 participants, 9, 7 and 14 participants showed positive SBR slopes in NOR, BOF and MAG, respectively. The young participants who had negative SBR slopes ranged from – 0.13 to – 2.44 SBR/ min in NOR, –0.03 to – 1.75 SBR/ min in BOF and – 0.03 to – 1.02 SBR/ min in MAG. Although the majority of the participants had negative SBR slopes, they were rarely more than 1.0 except for one participant (08GT: – 2.44 SBR/min) in NOR, three participants (01JQ: – 1.08 SBR/min, 13EN: – 1.75 SBR/min and 35HM: – 1.35 SBR/min) in BOF and one participant (36AC: – 1.02 SBR/min) in MAG. The median SBR slopes in the elderly were 0.05 SBR/min in NOR, 0.05 SBR/min in BOF, and –0.01 SBR/min in MAG. Out of the 26 participants, 16, 17 and 14 participants had positive SBR slopes. In all the participants with negative SBR slopes in the different conditions, none of the elderly participants had a negative SBR slope of more than 1.0. The elderly participants who had negative SBR slopes ranged from – 0.02 to – 0.61 SBR/ min in NOR, – 0.06 to – 0.76 SBR/ min in BOF and – 0.01 to – 0.48 SBR/ min in MAG. The Friedman test revealed no statistically significant difference in the SBR slopes in the three different conditions in both the young adults ($\chi^2 = (2, 28) = 0.08$, $p = 0.96$) and elderly group ($\chi^2 = (2, 26) = 2.20$, $p = 0.33$). As there was no precedent study on how the SBR changes within a continuous time frame, an assumption was made that the SBR would change in a linear pattern. However, the low R^2 values strongly indicated a poor linear function fit. Visual inspection of the graph plots did not allow a straightforward classification of the curve fitting (the graph plots for all the elderly participants group in MAG were presented in Appendix 9.9). From the graph plots, the SBR did not show a consistent trend or pattern (e.g. exponential increase or decrease).

Table 6.4 Median blinking slopes and the R² values of the young adults and elderly participants

Conditions	Young (n = 28)		Elderly (n = 26)	
	Blinking Slope SBR/ min	Blinking Slope R ²	Blinking Slope SBR/ min	Blinking Slope R ²
NOR	-0.14 (0.44) Range: -2.44 to 0.85	0.11 (0.22) Range: 0.00 to 0.59	0.05 (0.28) Range: -0.61 to 0.84	0.11 (0.33) Range: 0.00 to 0.65
BOF	-0.24 (0.73) Range : -1.75 to 0.68	0.09 (0.26) Range: 0.00 to 0.60	0.05 (0.33) Range: -0.92 to 1.06	0.08 (0.27) Range: 0.00 to 0.84
MAG	0.00 (0.49) Range : -1.02 to 0.78	0.15 (0.33) Range: 0.00 to 0.72	-0.01 (0.31) Range: -0.48 to 0.62	0.12 (0.19) Range : 0.00 to 0.67

Interquartile range (IQR) values are shown in the parentheses

6.4.6 Normalised-SBR relative to the 2nd minute SBR

Table 6.5 shows the median SBR relative ratios to the 2nd minute in NOR, BOF and MAG. In the young adults' group, there were 16, 15 and 19 participants with an SBR relative ratio of more than 1.0 in the NOR, BOF and MAG, respectively. For the elderly adults' group, there were 20, 15, and 15 participants with SBR relative ratio of more than 1.0 in the NOR, BOF and MAG, respectively. Based on the hypothesis that a more visually challenging condition is more likely to induce a "fatigue" effect, we would expect a higher number of participants showing a significant increase in SBR in the MAG compared to the NOR and BOF. However, the Friedman test revealed no statistically significant difference in normalised SBR among the three different conditions in both the young adults ($\chi^2 = (2, 28) = 0.17, p = 0.92$) and elderly ($\chi^2 = (2, 26) = 4.82, p = 0.09$) participants.

Table 6.5 Median normalised relative ratios to the 2nd minute SBR in NOR, BOF and MAG for the young and elderly group.

	NOR	BOF	MAG
Young	1.00 (0.49) Range: 0.33 to 2.00	1.00 (0.38) Range: 0.40 to 2.00	1.00 (0.51) Range: 0.40 to 2.00
Elderly	1.08 (0.51) Range: 0.67 to 3.50	1.00 (0.30) Range: 0.00 to 6.00	1.00 (0.42) Range: 0.00 to 3.00

Interquartile range (IQR) values are shown in the parentheses

In the current analysis, the second minute was used as the "reference" because it allows more comparison from the successive minutes. For instance, a 10-minute would then allows eight relative ratios (e.g. 3rd/2nd, 4th/2nd, 5th/2nd ... 10th/ 2nd minute). However, it was unclear how "representative" is the use of second minute SBR as a reference SBR. As such, we extended

the analysis by splitting the testing duration into two segments; the first segment (i.e. 1st to 5th minute) and the second segment (6th to 10th minute). The average SBR over the 5 minutes was computed, and the relative ratio (i.e. average SBR 2nd segment/ average SBR 1st segment) would indicate if there is an increase (i.e. > 1.0) or decrease in SBR (i.e. < 1.0) between these two segments. Table 6.6 shows the median relative rates comparing the second segment average SBR to the first segment average SBR. Wilcoxon Signed-Rank test showed no statistically significant difference between the ratio of SBR using relative to the 2nd minute and relative to the average 5 minutes ($z = -0.702$, $p = 0.48$) for all participants.

Table 6.6 Median ratio comparing the second segment of the SBR relative to the first segment SBR in NOR, BOF and MAG for the young and elderly group.

	NOR	BOF	MAG
Young	0.93 (0.28) Range: 0.43 to 1.38	0.89 (0.21) Range: 0.45 to 1.53	0.95 (0.29) Range: 0.55 to 1.42
Elderly	1.01 (0.62) Range: 0.61 to 2.50	1.00 (0.34) Range: 0.57 to 3.18	0.93 (0.79) Range: 0.43 to 2.00

Interquartile range (IQR) values are shown in the parentheses

6.4.7 Reading speed performance

Table 6.7 shows the median reading speeds and reading slopes for the young adults and elderly group (the elderly group reading speed data has been presented previously in Chapter 4). Table 6.8 shows the post hoc Wilcoxon Signed-Rank test comparing the reading speed and reading slopes between the conditions. Friedman test shows a statistically significant difference in the reading speed between the three conditions in young ($\chi^2 = (2, 28) = 43.14$, $p < 0.001$) and elderly ($\chi^2 = (2, 26) = 39.69$, $p < 0.001$). The reading speed in MAG was statistically significantly lower compared to NOR and BOF in both the young adults and elderly group. Friedman test showed a statistically significant difference in reading slopes in all the three conditions in the young group ($\chi^2 = (2, 28) = 11.64$, $p = 0.003$) but not in the elderly group ($\chi^2 = (2, 26) = 1.46$, $p = 0.48$). For the young adults' group, post hoc Wilcoxon Signed-Rank test showed a statistically significant difference in the reading slopes between NOR and MAG ($z = -3.56$, $p < 0.001$), and BOF and MAG ($z = -2.66$, $p = 0.008$).

Table 6.7 Median reading speed, reading slope and R² value for the young adults and elderly adults

Young (n=28)	NOR	BOF	MAG
Reading speed (wpm)	207.23 (62.49) Range: 151.63 to 366.84	210.95 (106.65) Range: 116.15 to 332.81	103.57 (47.94) Range: 58.90 to 164.51
Reading slope (wpm/min)	4.09 (6.56) Range: -7.56 to 24.64	1.51 (7.72) Range: -5.66 to 18.84	-0.10 (1.86) Range: -4.10 to 4.95
Reading Slope R ²	0.19 (0.31) Range: 0.00 to 0.81	0.10 (0.17) Range: 0.00 to 0.50	0.14 (0.26) Range: 0.00 to 0.42
Elderly (n=26)	NOR	BOF	MAG
Reading speed (wpm)	174.25 (80.99) Range: 86.61 to 282.45	154.07 (61.32) Range: 89.23 to 305.10	64.00 (46.63) Range: 20.81 to 147.61
Reading slope (wpm/min)	0.41 (6.19) Range: -7.94 to 10.13	0.48 (5.68) Range: -8.65 to 7.70	0.44 (1.95) Range: -6.28 to 4.61
Reading Slope R ²	0.10 (0.27) Range: 0.00 to 0.84	0.09 (0.18) Range: 0.00 to 0.78	0.13 (0.60) Range: 0.00 to 0.78

Interquartile range (IQR) values are shown in the parentheses

Table 6.8 Post hoc Wilcoxon Signed-Rank test for reading speed and reading slope in NOR, BOF and MAG for the young adults and elderly group.

	Reading speed (wpm)		Reading slopes (wpm/ min)	
	NOR-BOF	NOR-MAG	NOR-BOF	NOR-MAG
Young	z = - 0.97 p = 0.32	z = - 4.62 p < 0.001	z = - 1.70 p = 0.08	z = - 3.56 p < 0.001
Elderly	z = - 1.99 p = 0.04	z = - 4.45 p < 0.001	z = - 6.99 p = 0.48	z = - 0.80 p = 0.42

6.4.8 Relationship between SBR slopes and reading slopes

Although the median blinking slopes indicated almost no changes in the SBR with time in both the young and elderly group, there was a mixture of both positive and negative blinking slopes within the group. As such, the Spearman rho correlation was conducted to investigate the relationship between SBR slopes and reading slopes for all the participants. There was no correlation between SBR slopes and reading slopes in NOR ($r = 0.04$, $p = 0.76$); a weak and opposite correlation in BOF ($r = - 0.24$, $p = 0.07$), and a weak but positive correlation in MAG ($r = 0.24$, $p = 0.08$) between the SBR slopes and reading slopes was observed. The correlation suggests that the SBR reduced as the reading speed increased in BOF, and an increased in SBR was associated with an increased in reading speed in MAG.

6.5 Discussion

Overall, our results showed that the SBRs were statistically significantly lower in the middle-elderly group (SBR-NOR: 6.10 blinks/min) compared to the young adults' group (SBR-NOR: 8.35 blinks/min) in the habitual reading condition. The SBR was also lower in the elderly group compared to the young adults' group in the BOF and MAG, did not reach statistical significance. Previous studies on SBR in elderly participants have shown mixed findings of increase SBR (Sun et al., 1997) and decrease SBR (Doughty and Naase, 2006) with age. One possible explanation of the reduced SBR in elderly can be attributed to the smaller palpebral aperture with increasing age (i.e. senile ptosis (Zaman et al., 1998, van den Bosch et al., 1999)). Tsubota and Nakamori (1995) have reported that an increase in evaporation can occur in individuals with a larger exposed ocular surface, resulting in more frequent blinking. Indeed, Cho et al. (2000) have demonstrated that the different positions of gaze which changes the exposed ocular surface area can affect the SBR. For instance, lower SBR were reported in participants in downward gaze compared to the primary gaze. This would suggest that individuals with narrow palpebral aperture height would have a lower SBR. However, Doughty and Naase (2006) and Zaman et al. (1998) did not find any significant correlation between the participant's palpebral aperture height and SBR in primary gaze. Although our study shows a significantly lower SBR in the elderly group compared to the young adults' group, it is unlikely that the lower SBR was due to a smaller palpebral aperture in the elderly participants. Our study did not attempt to measure the palpebral height of our participants during downward gaze. However, our elderly participants did not show visible signs of age-related ptosis based on direct observation. So far, the studies mentioned above investigating an age effect on SBR were limited to simple activities or tasks (e.g. casual conversation, awaiting in rooms, etc.). In this study, the participants were involved in a reading task, a cognitive task which could potentially influence the blinking mechanism (Holland and Tarlow, 1975, Holland and Tarlow, 1972). Previous studies (Wood et al., 2010, Wood et al., 2009) have reported that despite similar visual degradation using visually degrading filters, the impact on cognitive performance can vary between the young and elderly adults. This cognitive effect could exert its effect on the blinking mechanism. However, it is unclear at this moment how this would impact on the blinking performance.

Our results have shown that a significant reduction of SBR (i.e. up to 40%) was seen in MAG compared to NOR. The reduction in SBR was consistent with the study of Nahar et al. (2007), showing that increased visual demand (i.e. reduction in contrast reserve) can cause a decrease in SBR. During blinking, the temporary eye closure may interfere with visual information acquisition. Previously, Orchard and Stern (1991) and Hall (1945) have shown that blinking actions did not occur randomly during reading. The blinks happened during the return sweep (i.e. moving the eyes from the end of one line of text to the beginning of another line (Orchard and Stern, 1991)). The findings suggest that when visual information is available, the blinking initiation reduces to avoid disturbance to visual input. Therefore, SBRs are reduced in visually demanding conditions to optimise visual information acquisition (Holland and Tarlow, 1972, Holland and Tarlow, 1975).

Apart from the visual demand, another possible contributing factor could be the reading task. As a denser Bangerter occlusion foil was used in MAG, a hand-held magnifier was used to support the reading task. Our results show that the SBR was statistically significantly reduced in MAG, where participants were required to read the text and to maintain the focus using the hand-held magnifier simultaneously – dual tasks which can have the potential to reduce SBR (Recarte et al., 2008). Although dual tasks can have the capacity reduce SBR in the MAG, it is hard to determine which of the above two factors (i.e. reduced text contrast or the dual tasks) was the main factor causing the reduction in SBR. A subsequent experiment can be planned to assess the SBR using a high contrast text and a low contrast text. If the percentage of the relative reduction in SBR in the low contrast text is similar (i.e. approximately up to 40%) to the current study, then it can be concluded that the main cause of the SBR reduction is primarily due to the low text contrast instead of the dual tasks. Also, our results showed that there was a 10-20% increase in SBR in both the young adults and elderly groups of participants in BOF. Although participants were also simulated with Bangerter occlusion foils in BOF, they were of lesser intensity compared to the ones used together with the hand-held magnifier. Our current results on the reading assessment showed no statistically significant difference in reading speed between NOR and BOF but reached statistically significant difference when the reading speed in NOR was compared to MAG. This finding suggests that the reading in BOF may not

be adequately challenging to induce higher visual demand. Therefore, no decrease in SBR was observed.

To date, the use of SBR has only been reported as an indicator of fatigue in extended reading duration (i.e. 1 hour to 4 hours (Luckiesh, 1947, Hoffman, 1946)). However, such testing durations were not practical to be assessed in a clinic setting. Recent evidence suggests that fatigue effects may potentially be elicited within a 15-minute reading duration (Ramulu et al., 2013a), and this is the first study to explore the use of SBR as an indicator of fatigue effect in a 10-minute testing duration. We hypothesised that the SBR would increase with time due to mental fatigue, especially in the elderly participants. To identify changes in the SBR with time, we explored two analyses methods, SBR slope and normalised SBR, were used. Both the SBR slopes and normalised SBR suggested that the SBR remains relatively constant with time in the elderly group. Although the reading slopes were slightly negative in the NOR and BOF in the young participants, the difference did not reach a statistical significance. There are, however, limitations on using SBR slopes to denote changes in SBR with time. The number of blinks was tracked for every minute, and a best straight line was fitted across the data points to denote any indication of an increase or decrease in SBR with time. As such, the assumption was that any of these changes would occur linearly. Our experience in analysing the data shows that the SBR can be highly variable and inconsistent within successive minutes, making it unlikely to confine to a linear trend. In this thesis, we presented the graph plots for the participants who, presumably, have the most effect (i.e. elderly participants > 60 years, in MAG). The graph plots on SBR with time for each participant were presented in Appendix 9.9. As noted earlier, visual inspection from the individual graph plots suggests that the number of blinks per minute were highly variable within the 10 minutes time span within an individual. As such, monitoring the blinking rates within a small time interval (i.e. every minute) presented difficulty in identifying a blinking pattern. A second approach to identify changes in SBR was to compare to the SBR in each successive minutes to the SBR calculated at the second minute. Our results on the normalised SBR showed relatively consistent blinking rates across the time duration. As mentioned earlier, large SBR variation occurs within a minute, and an assumption on this analysis method was that the SBR calculated at the second minute was representative

of the participant “normal” blinking rate. As such, the third approach was to compare the average SBR from the second half of the 5-minute relative to the average SBR from the first half of the 5-minute. The relative SBR ratio would then suggest if the SBR has increase (i.e. > 1.0) or decrease (i.e. <1.0). By averaging the SBR for 5-minute, this approach reduces the random blinking inconsistency. Likewise from the previous analyses, the relative ratio SBRs were almost 1.0 indicating that the average SBR (i.e. from the last 5 minutes) did not change significantly in the second half of the blinking assessment. There was also no statistically significant difference in the SBR ratios between the normalised SBR relative to the 2nd minute, and the ratio obtained from averaging the SBR in the second half of the test relative to the first 5 minutes. The analyses show that the SBR generally remains consistent for a testing duration of 10-minute. Arguably, a longer duration may be required to identify blinking pattern changes, but one also has to consider how practical this can be done in a clinical setting. Previously Ramulu et al. (2013a) have reported a “fatigue” effect in a silent reading testing duration of 10 to 15-minute. Although our assessment was only limited to 10-minute, the visual inspection on the graph plots did not suggest a consistent trend that the SBR is increasing gradually.

So far, no study has been undertaken to investigate the relationship between duration related reading speed performances and SBR. In this current study, the reading slopes were used as an indicator for duration related effect on the reading speed. If the decline of the reading speed is associated with mental fatigue, then we would expect the SBR to increase with time. Although most of the participants did not show a reduction in reading speed with time in our previous results, there was a minority of participants showing a decline in reading speed with time. As all the participants displayed at least one or more negative reading or SBR slopes in either NOR, BOF or MAG, the correlation analysis was performed on all participants. Our results showed that a weak correlation was observed between SBR slopes and reading slopes in the BOF and MAG. The correlation in BOF shows opposite changes in SBR (i.e. increasing reading speed with time is associated with decreasing SBR with time) whereas MAG positive correlation indicates an increase in SBR with increased reading speed. The inconsistent correlations found in BOF and MAG did not support our initial hypothesis, that is, the

participants who showed a reduction in reading speed with time, presumably due to mental fatigue, will show an increase in SBR.

6.6 Limitations

There are some limitations in this study that should be mentioned. One of the major limitations was the SBR assessment method used in this study. The use of video recording to document blinking patterns in the current study has been reported previously (Cho et al., 2000, Bentivoglio et al., 1997). The definition of a single blink used by Abelson and Holly (1977) was adopted in the current study, which includes complete, incomplete, forced or a twitch. Blink flurries, which is defined as three or more blinks within a 3-seconds window period by Stern et al. (1994), can also occur occasionally during the reading task. However, the use of this definition may not be as straightforward, based on our video observations. There are other variations, in our view, should be defined as blink flurries. For instance, multiple partial blinks that were less than 3 seconds or multiple blinks of partial and full blinks that lasted for more than 3 seconds, etc. The occurrence of the blink flurries can make it hard to demarcate if the subsequent blinks were considered as part of the blink flurries, or counted as a few individual blinks. Although we have attempted to standardise the way the blinks were counted, the variation in the blinking patterns could complicate the determination of a single blink. As such, this can affect the precision and repeatability of the SBR measurements and limits its usage as an outcome measure indicator. From our observation, participants with higher SBR have a higher tendency to execute more blink flurries, which explained that the data points were outside the 95% confidence interval in the Bland-Altman plot in Figure 6.3.

Apart from the limitations of the repeatability in counting the blinking rates, there were challenges when recording blinking activities in a typical reading position (i.e. slightly lower gaze). The video camera has to be anchored on the bottom part of the trial frame, and the camera has to be attached “before” the participant’s refractive correction trial lens to avoid reflections from the trial lenses (see Figure 6.1). Previously, Doughty (2016) defined a blink when the upper eyelid occludes the pupil partially or completely. As the testing duration in our experiment can last from 10 to 15 minutes, and the participant’s palpebral aperture height can

vary during this testing duration. Therefore, it is tricky to use the occlusion of the pupil to differentiate between full and partial blinks. Some participants showed a narrowing of the palpebral aperture height with time, presumably due to “boredom” or sleepy. Perhaps, the use of an eye tracker may help to streamline the types of “countable” blinks or the development of a software algorithm to compute the blinks. The benefit of having a software algorithm may provide a more consistent approach to tabulate the number of blinks. Unless a more robust method can be used to determine the blinking patterns clearly, the large intra-observer variability and the inherent individual variability observed from our results suggest that SBR may not be sensitive to identify “fatigue” effect within a short time. The elderly participants recruited in our study were of a younger age group. As discussed in our previous chapter, the data would be more meaningful in a group of individuals above 65 years old. The recruitment of participants above 65 years old poses a challenge in Singapore because participants were required to have reached at least “O” level education and be English literate. However, at the time where the older participants would have received a formal education, English was not the first language in Singapore.

The current experiment used a convenience sample of 54 participants (young and elderly). Post hoc power analysis using G*Power (Version 3.1.9.3) for a total sample of 54 participants gave a limited power of 28% to detect a difference of 0.19 SBR/ min from zero in the blinking slopes (see Appendix 9.10.6). As the primary objective was to investigate the changes of SBR within the 10 minutes duration, the SBR slopes were used to compute the post hoc power analysis. Given the large variability of the data, a larger sample of 230 participants would be required to give a power of 80%. Apart from the limitations of the data analysis using the video camera, another considering factor is the reproducibility of the blinking rates of the participants. The current experiment was only limited to a single visit assessment, and it is unclear if the observed effect would be reproducible in a second assessment. So far, no studies have been reported in the investigation on the reproducibility of SBR in specific activities. The current experiment assumed a constant temperature and humidity had been maintained, and it is unlikely to cause a difference in the current study results because - Firstly, the participants performed all the reading tests within a single session. Secondly, the SBRs were compared

relative to their own habitual reading SBR, rather than between participants in different sessions, which may potentially have a different temperature and humidity. However, further works on the reproducibility should include the temperature and humidity measurements. In addition, it is necessary to determine if the SBR (i.e. in terms of large inherent variability, measurement criteria, reproducibility) would be a robust indicator for any observed effects (e.g. increase SBR indicates mental fatigue (Martins and Carvalho, 2015)).

6.7 Conclusion

In conclusion, elderly participants have lower SBRs compared to the young adults' group in a silent reading task. Also, both groups shows a reduced SBR under visually challenging conditions. The objective of this experiment was to explore the possible use of SBR to index mental fatigue within a short testing duration. Our results suggest that the SBR remains relatively constant within a 10-minute silent reading task. Although Martins and Carvalho (2015) have suggested the use of SBR as a possible indicator of mental fatigue, investigators should consider the following factors: Firstly, a short reading duration (i.e. 10 minutes) may not be sufficient to pick up "fatigue" effects. Secondly, the use of video recording to document blinking pattern requires a more robust and precise definition of the required blinking patterns. Lastly, further works would need to establish if any of the observed effects are reproducible.

7. Summary of findings and future work

Reading is one of the most important concerns for which visually impaired individuals seek help in low vision rehabilitation (Elliott et al., 1997). Therefore, it is not surprising that reading assessments are included in a typical low vision consultation. To minimise the burden of the reading assessments on individuals with visual impairment, most of the reading speed assessments have been limited to short reading tests (i.e. a few words or single sentences). This reading assessment may be adequate when considering spot or survival reading performance. However, information such as reading endurance, which requires the ability of an individual to sustain the reading performance during an extended duration in leisure reading is lacking, and may potentially be important. To date, there are limited studies investigating reading performance over an extended time. Although Ramulu et al. (2013a) have previously reported a decline in reading speed in a group of glaucoma participants over a 30-minute reading task, this is the first study undertaken to investigate the duration related effects on reading speed in a group of individuals using magnifiers for reading. Using a LVA for leisure reading can be a challenge for some visually impaired individuals, and it was hypothesised that the reading speed would progressively slow as the task progressed due to reader fatigue. If this was found to be the case, it might be useful to determine if some individuals were more susceptible to “fatigue”; and could potentially be used as a predicting factor for the success of a LVA for leisure reading. Thus, the purpose of this research was to investigate if any negative duration related effects could be picked up in an extended reading assessment test (i.e. 10 to 15 minutes), and (if so) also identified in much shorter reading tests. However, our study did not find any consistent evidence to suggest that the reading speed was progressively slower with time in either the normally-sighted participants with simulations or the visually impaired participants.

7.1 Strengths and Limitations

In this study, two different approaches were used to identify the presence of fatigue: one was to monitor the reading speed to see if it would progressively slow as the task progressed. The second method was to monitor the SBR, to see if that would progressively increase with the duration of the reading task. Previous researchers (Martins and Carvalho, 2015) have

suggested the possible use of SBR as an indicator of mental fatigue. Apart from an overall reduction in SBR when participants were using the hand-held magnifiers in the simulated impairment group, a consistent increase in SBR with reading time was not observed in most of the participants. There were, however, some participants who did show an increase in SBR with time. As such, a correlation analysis was conducted, but no association was found between the reading speed performance and the SBR. The limitations of the methodology of using a video camera for SBR computation have been discussed in detail in Section 6.5, which presented an ambiguity in the interpretation of the current results. While no association between SBR and reading slopes has been observed in the current study, other performance measures on fatigue, such as the Fatigue Assessment Scale (FAS) should be considered in future works.

The reading performance was investigated under visually demanding conditions through two approaches. Firstly, the use of Bangerter occlusion foils to create a visually demanding condition for the normally-sighted individuals so that the simulated conditions were controlled and predictable, and could be induced in groups of a selected age. Secondly, the investigation was extended to a group of visually impaired patients from the p-EVES study. Most participants from the p-EVES study had binocular central scotoma, which is typical of the general low vision community, and something which could not be investigated in our simulated impairment group. In the simulated impairment study group, we have limited the investigation to the use of a hand-held magnifier. In a previous survey conducted by Watson et al. (1997), most of the visually impaired individuals reported that the hand-held or stand magnifiers only allow a short reading duration (i.e. a few minutes). There were, however, some participants who reported more than one hour of leisure reading with video magnifiers and spectacle magnifiers, but none of them was able to do so with either hand-held or stand magnifiers. These findings suggest that the difficulty in performing leisure reading was not due to their visual impairment per se, but the types of LVA used for leisure reading. Indeed, a recent study by Taylor et al. (2017) has shown that the users reported a significant increase in reading duration when switching from optical magnifiers to electronic video magnifiers. One of the possible reasons for restrictions when using optical magnifiers such as the hand-held magnifiers could be due to a fatigue effect

during a prolonged period of reading in a physically demanding condition (i.e. holding the magnifier and maintaining the position). When comparing stand magnifiers and hand-held magnifiers, the latter poses a more “ergonomically” demanding vision aid, as it requires the reader to maintain the correct distance between the reading material and the magnifier for an extended time during leisure reading. Therefore, if there was any duration related effect, the choice of magnifier chosen in this study would have most likely to induce the biggest effect compared to other magnifiers.

The limitations of the experiments and analysis have been discussed in details in the respective chapters. Apart from Chapter 5, the other experiments used a convenience sample size of approximately 30 participants (some experiment sample sizes were further reduced due to limited data available), and the post hoc power analysis has shown that most of the experiments were limited in their power to detect a difference. The experiment was designed to mimic real world habitual reading such as giving instructions to participants to read like they would when reading a newspaper and the use of silent reading rather than out-loud reading. It is, however, unclear if the instructions and the use of silent reading in the current study could contribute to a more variable reading speed in successive minutes observed in the 10 to 15 minutes reading duration, which produced a more variable reading slope. Given the large variability of the reading slopes (which was not known before the experiment), a larger sample size would be required to achieve a minimum standard power of 80%. Based on the existing findings, most of the participants did not show a negative duration related effects. Even in the group of participants who did show a negative duration related effect, this was rarely consistent in all the three reading conditions. There was no systematic trend in the reading speed changes (e.g. reading speed decline in a more visually challenging conditions compared to the habitual condition). As such, it is unlikely that a larger sample size would have produced a statistically significant effect. This was also supported by a similar observation on the analyses on a group of visually impaired individuals from the p-EVES study; where there was also no evidence on the reading speed changes to suggest a negative duration related effect. On the other hand, it could also be argued that because there is also an element of learning effect (i.e. shown as an

increase in reading speed with time), this effect could have outweighed the negative duration related effect. As a result, an increase or consistent reading speed with time was observed.

As the current study aims to investigate the duration related effects in reading performance, one of the key consideration factors would be the assessment duration. Although a minimum reading duration of 15 minutes was adopted in our simulated study group, the reading duration for the participants in the p-EVES study was shorter, ranging from 47.99 seconds to 596.85 seconds (approximately 9 minutes). Would the shorter reading duration in this group of participants have prevented the occurrence of any fatigue effect? The data from the participants from the p-EVES study was collected from a single reading paragraph (approximately 150 words), but some participants took several minutes to complete the task. The use of an extended reading assessment (in terms of the number of words) is not practical in a clinical setting. However, it seems unlikely that the use of a single paragraph has prevented the identification of a fatigue effect, since the trend for the reading speed was seen to be increasing rather than slowing down. Although some visually impaired individuals reported that prolonged reading is challenging for them even with the appropriate magnifiers, there was no evidence that this was due to a fatigue effect – which could be measured in a clinical setting and used to predict unsuccessful magnifier use.

So, why do visually impaired individuals report difficulty in leisure reading despite having an appropriate LVA? An early study by Leat et al. (1994) reported a marked difference in LVAs performance between the clinical assessment and the actual home usage (i.e. 75% of the participants reported being able to read the newspaper in the clinic, but only 35% of them could read books at home). While the clinical assessment of reading performance is always in ideal conditions (e.g. high contrast text, optimal lighting at the proper location, maintaining the distance between the hand-held magnifiers and the reading material, encouragement from the clinician, etc.), this may not be the same as what the patients experienced at home. However, this is unlikely to affect the patients for spot or survival reading as the duration is relatively shorter compared to leisure reading. The short reading assessment in the clinic can be problematic when the actual reading duration is longer because the individual may have

difficulty sustaining the optimal “ergonomics” at home (e.g. maintaining the distance between the hand-held magnifier and the reading material or having difficulty to navigate a small paperback novel using a stand magnifier, etc.). Also, the different nature of survival and leisure reading must be considered. It can be argued that survival reading is what has to be done – either by the visually impaired individual or by someone else on his or her behalf. This forms an essential component of many instrumental of activities of daily living (IADLs) and is a key factor in maintaining independent living. On the other hand, leisure reading is an “optional” activity, and it has to be enjoyable; if it is not, the individual will not be motivated and will restrict leisure reading activity. If the lack of motivation is the cause for the individuals to limit their leisure reading time, then this would be difficult to quantify or predict.

In conclusion, the findings in the current study did not suggest that a negative duration related effect can be picked up using an extended reading duration in the visually impaired individuals (i.e. up to 9 minutes) and the healthy participants with simulation (i.e. 10–15 minutes). The negative duration related effect, if any, would require a long testing duration (i.e. 30 minutes (Ramulu et al., 2013a)) which is not practical in a routine low vision consultation. Even in participants who did show a negative duration related effect, there was no clinical measurement indicators that can predict which group of participants might be more susceptible to such effects. Also, the lack of correlation between the duration related effects indicators (i.e. reading slopes and NRR2) and the reported duration and frequency use of the magnifiers from the participants in the p-EVES study suggests that these duration related effects indicators were not predictive of their ability to undertake leisure reading. The current study adopted the assessment of reading speed changes with time (i.e. reading slope), where a negative duration related effect has been reported by Ramulu et al. (2013a). However, our findings have shown that reading speed measurements can be very variable in individuals with visual impairment, which can make it difficult to reveal any systematic patterns in the data.

7.2 Future work

Future work on the investigation of duration related effects could consider the use of other parameters. The most straightforward approach would be to measure the reading duration

directly. However, this would not be a practical approach. Another possible parameter could be the assessment of CPS before and after an extended reading period; higher acuity reserves are required for high fluent reading (i.e. 160 wpm) compared to spot or survival reading (i.e. 40 wpm), and any changes in the acuity reserves would be reflected from the CPS.

Although the current study did not suggest any evidence of a fatigue effect within 10 to 15 minutes of reading speed assessment tests in a group of healthy participants with simulated degraded vision and visual impaired participants, the study can possibly be extended to individuals who might be more susceptible to fatigue. An example of this particular group might be people suffering from post traumatic brain injury (PTBI) (Cantor et al., 2008). PTBI is a neurological disorder that can be seen in people with history of road-traffic accident, stroke, etc. Cognitive fatigue is the fifth most common symptom experienced by such patients. It is hard to quantify fatigue per se and usually doctors and clinicians can only pick up the symptom by asking the patient. The assessment of reading endurance, which involves a cognitive skill, may provide a possible objective way to quantify such cognitive fatigue. If the use of reading endurance assessment can prove to index this cognitive fatigue, then it could provide alternative to asking patients - which can be highly subjective, or to using complex equipment (i.e. Functional Magnetic Resonance Imaging) (Cantor et al., 2013) to pick up a fatigue effect in this group of patients.

8. References

- ABELSON, M. B. & HOLLY, F. J. 1977. A tentative mechanism for inferior punctate keratopathy. *American journal of ophthalmology*, 83, 866-869.
- AHN, S. J., LEGGE, G. E. & LUEBKER, A. 1995. Printed cards for measuring low-vision reading speed. *Vision Res*, 35, 1939-44.
- AKUTSU, H., LEGGE, G. E., ROSS, J. A. & SCHUEBEL, K. J. 1991. Psychophysics of reading--X. Effects of age-related changes in vision. *J Gerontol*, 46, P325-31.
- ALTPETER, E. K., MARX, T., NGUYEN, N. X., NAUMANN, A. & TRAUZETTEL-KLOSINSKI, S. 2015. Measurement of reading speed with standardized texts: a comparison of single sentences and paragraphs. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 253, 1369-1375.
- ARDITI, A. 1999. Elicited sequential presentation for low vision reading. *Vision Research*, 39, 4412-4418.
- ARDITI, A. 2005. Improving the design of the letter contrast sensitivity test. *Investigative Ophthalmology & Visual Science*, 46, 2225-2229.
- ARDITI, A. & CHO, J. 2005. Serifs and font legibility. *Vision Research*, 45, 2926-2933.
- ARDITI, A., KNOBLAUCH, K. & GRUNWALD, I. 1990. Reading with fixed and variable character pitch. *Journal of the Optical Society of America a-Optics Image Science and Vision*, 7, 2011-2015.
- BAILEY, I. L. & LOVIE, J. E. 1980. The design and use of a new near-vision chart. *Am J Optom Physiol Opt*, 57, 378-87.
- BALDASARE, J. & WATSON, G. R. 1987. Observations from the psychology of reading relevant to low vision research. *Low Vision*. Springer.
- BALDASARE, J., WATSON, G. R., WHITTAKER, S. G. & MILLERSHAFFER, H. 1986. The development and evaluation of a reading test for low vision individuals with macular loss. *Journal of Visual Impairment & Blindness*, 80, 785-789.
- BALOTA, D. A., DOLAN, P. O. & DUCHEK, J. M. 2000. Memory changes in healthy older adults. *The Oxford handbook of memory*, 395-409.
- BARBATO, G., MOUL, D. E., SCHWARTZ, P., ROSENTHAL, N. E. & OREN, D. A. 1993. Spontaneous eye blink rate in winter seasonal affective disorder. *Psychiatry research*, 47, 79-85.
- BECKER, W. & FUCHS, A. F. 1988. Lid-eye coordination during vertical gaze changes in man and monkey. *Journal of neurophysiology*, 60, 1227-1252.
- BECKMANN, P. J. & LEGGE, G. E. 1996. Psychophysics of reading--XIV. The page navigation problem in using magnifiers. *Vision Res*, 36, 3723-33.
- BENTIVOGLIO, A. R., BRESSMAN, S. B., CASSETTA, E., CARRETTA, D., TONALI, P. & ALBANESE, A. 1997. Analysis of blink rate patterns in normal subjects. *Mov Disord*, 12, 1028-34.
- BERNARD, J.-B., AGUILAR, C. & CASTET, E. 2016. A new font, specifically designed for peripheral vision, improves peripheral letter and word recognition, but not eye-mediated reading performance. *PloS one*, 11, e0152506.
- BERNARD, J.-B., KUMAR, G., JUNGE, J. & CHUNG, S. T. 2013. The effect of letter-stroke boldness on reading speed in central and peripheral vision. *Vision research*, 84, 33-42.
- BLAND, J. M. & ALTMAN, D. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, 327, 307-310.
- BOUMA, H. 1970. Interaction effects in parafoveal letter recognition. *Nature*, 226, 177-178.
- BOUMA, H. & DE VOOGD, A. 1974. On the control of eye saccades in reading. *Vision Research*, 14, 273-284.

- BOWERS, A. & ACKERLEY, R. 1994. Reading Characteristics of Normal Observers Using Low-Vision Aids. *Low Vision: Research and New Developments in Rehabilitation*, 11, 235-238.
- BOWERS, A., CHEONG, A. M. & LOVIE-KITCHIN, J. E. 2007. Reading with optical magnifiers: page navigation strategies and difficulties. *Optom Vis Sci*, 84, 9-20.
- BOWERS, A. R. 2000. Eye movements and reading with plus-lens magnifiers. *Optom Vis Sci*, 77, 25-33.
- BOWERS, A. R., WOODS, R. L. & PELI, E. 2004. Preferred retinal locus and reading rate with four dynamic text presentation formats. *Optometry and Vision Science*, 81, 205-213.
- BRITTON, B. K., PIHA, A., DAVIS, J. & WEHAUSEN, E. 1978. Reading and cognitive capacity usage: Adjunct question effects. *Memory & Cognition*, 6, 266-273.
- BROWN, B. 1981. Reading performance in low vision patients: Relation to contrast and contrast sensitivity. *American Journal of Optometry and Physiological Optics*, 58, 218-226.
- BROWN, G. M. & RUBIN, G. S. Investigating The Impact of Eccentric Viewing Training On reading Performance In People With AMD Participating In The EFFECT Trial. 2016. Annual Meeting of the Association-for-Research-in-Vision-and-Ophthalmology
- BRUSSEE, T., NISPEN, R., KLERKX, E. M., KNOL, D. L. & RENS, G. H. 2015. Comparison of reading performance tests concerning difficulty of sentences and paragraphs and their reliability. *Ophthalmic and Physiological Optics*, 35, 324-335.
- BRUSSEE, T., VAN NISPEN, R. & VAN RENS, G. H. 2016. Visual and personal characteristics are associated with reading performance in normally sighted adults. *Clinical and Experimental Optometry*.
- BRUSSEE, T., VAN NISPEN, R. M. & VAN RENS, G. H. 2014. Measurement properties of continuous text reading performance tests. *Ophthalmic Physiol Opt*, 34, 636-57.
- BULLIMORE, M. A. & BAILEY, I. L. 1995. Reading and eye movements in age-related maculopathy. *Optom Vis Sci*, 72, 125-38.
- BURGGRAFF, M. C., VAN NISPEN, R. M., HOEK, S., KNOL, D. L. & VAN RENS, G. H. 2010. Feasibility of the Radner Reading Charts in low-vision patients. *Graefes Arch Clin Exp Ophthalmol*, 248, 1631-7.
- CALABRESE, A., CHEONG, A. M., CHEUNG, S.-H., HE, Y., KWON, M., MANSFIELD, J. S., SUBRAMANIAN, A., YU, D. & LEGGE, G. E. 2016. Baseline MNREAD measures for normally sighted subjects from childhood to old age. *Investigative ophthalmology & visual science*, 57, 3836-3843.
- CALABRÈSE, A., OWSLEY, C., MCGWIN, G. & LEGGE, G. E. 2016. Development of a reading accessibility index using the MNREAD acuity chart. *JAMA ophthalmology*, 134, 398-405.
- CALABRÈSE, A., TO, L., HE, Y., BERKHOLTZ, E., RAFIAN, P. & LEGGE, G. E. 2018. Comparing performance on the MNREAD iPad application with the MNREAD acuity chart. *Journal of vision*, 18, 8-8.
- CANTOR, J. B., ASHMAN, T., GORDON, W., GINSBERG, A., ENGMANN, C., EGAN, M., SPIELMAN, L., DIJKERS, M. & FLANAGAN, S. 2008. Fatigue after traumatic brain injury and its impact on participation and quality of life. *The Journal of head trauma rehabilitation*, 23, 41-51.
- CANTOR, J. B., GORDON, W. & GUMBER, S. 2013. What is post TBI fatigue? *NeuroRehabilitation*, 32, 875-83.
- CARNEY, L. G. & HILL, R. M. 1982. The nature of normal blinking patterns. *Acta ophthalmologica*, 60, 427-433.
- CARPENTER, A. 1948. The rate of blinking during prolonged visual search. *Journal of Experimental Psychology*, 38, 587.
- CARVER, R. P. 1990. *Reading rate: A review of research and theory*, Academic Press.

- CARVER, R. P. 1992. Reading Rate- Theory, Research, and Practical implications. *Journal of Reading*, 36.
- CHEUNG, S.-H., KALLIE, C. S., LEGGE, G. E. & CHEONG, A. M. 2008. Nonlinear mixed-effects modeling of MNREAD data. *Investigative ophthalmology & visual science*, 49, 828-835.
- CHO, P., SHENG, C., CHAN, C., LEE, R. & TAM, J. 2000. Baseline blink rates and the effect of visual task difficulty and position of gaze. *Curr Eye Res*, 20, 64-70.
- CHU, C. A., ROSENFELD, M. & PORTELLO, J. K. 2014. Blink patterns: reading from a computer screen versus hard copy. *Optometry & Vision Science*, 91, 297-302.
- CHUNG, S. T. 2002. The effect of letter spacing on reading speed in central and peripheral vision. *Investigative Ophthalmology & Visual Science*, 43, 1270-1276.
- CHUNG, S. T., MANSFIELD, J. S. & LEGGE, G. E. 1998. Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision research*, 38, 2949-2962.
- COLE, R. J. 2008. Modifications to the Fletcher Central Field Test for patients with low vision. *Journal of Visual Impairment & Blindness*, 102, 659-660.
- COLLINS, M., HERON, H., LARSEN, R. & LINDNER, R. 1987. Blinking patterns in soft contact lens wearers can be altered with training. *Am J Optom Physiol Opt*, 64, 100-3.
- COLLINS, M., SEETO, R., CAMPBELL, L. & ROSS, M. 1989. Blinking and corneal sensitivity. *Acta Ophthalmol (Copenh)*, 67, 525-31.
- CRUZ, A. A., GARCIA, D. M., PINTO, C. T. & CECHETTI, S. P. 2011. Spontaneous eyeblink activity. *The ocular surface*, 9, 29-41.
- DE PADOVA, V., BARBATO, G., CONTE, F. & FICCA, G. 2009. Diurnal variation of spontaneous eye blink rate in the elderly and its relationships with sleepiness and arousal. *Neuroscience letters*, 463, 40-43.
- DEMARCO, L. M. & MASSOF, R. W. 1997. Distributions of print sizes in U.S. newspapers. *Journal of Visual Impairment and Blindness*, 91, 9-13.
- DICKINSON, C. & RABBITT, P. 1991. Simulated visual impairment: Effects on text comprehension and reading speed. *Clinical vision sciences*, 6, 301-308.
- DICKINSON, C. & SHIM, M. 2007. The influence of manual dexterity on reading speed with a hand-held magnifier. *Invest Ophthalmol Vis Sci*, 48, 4368-74.
- DOUGHTY, M. J. 2001. Consideration of three types of spontaneous eyeblink activity in normal humans: during reading and video display terminal use, in primary gaze, and while in conversation. *Optom Vis Sci*, 78, 712-25.
- DOUGHTY, M. J. 2002. Further assessment of gender-and blink pattern-related differences in the spontaneous eyeblink activity in primary gaze in young adult humans. *Optometry & Vision Science*, 79, 439-447.
- DOUGHTY, M. J. 2014. Spontaneous eyeblink activity under different conditions of gaze (eye position) and visual glare. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 252, 1147-1153.
- DOUGHTY, M. J. 2016. Assessment of short-term variability in human spontaneous blink rate during video observation with or without head/chin support. *Clinical and Experimental Optometry*, 99, 135-141.
- DOUGHTY, M. J. & NAASE, T. 2006. Further analysis of the human spontaneous eye blink rate by a cluster analysis-based approach to categorize individuals with 'normal' versus 'frequent' eye blink activity. *Eye & contact lens*, 32, 294-299.
- DREW, G. C. 1951. Variations in reflex blink-rate during visual-motor tasks. *Quarterly Journal of Experimental Psychology*, 3, 73-88.
- ELLIOTT, D. B., TRUKOLO-ILIC, M., STRONG, J. G., PACE, R., PLOTKIN, A. & BEVERS, P. 1997. Demographic characteristics of the vision-disabled elderly. *Invest Ophthalmol Vis Sci*, 38, 2566-75.

- ELLIOTT, D. B., WHITAKER, D. & BONETTE, L. 1990. Differences in the legibility of letters at contrast threshold using the Pelli-Robson chart. *Ophthalmic and Physiological Optics*, 10, 323-326.
- FINE, E. M. & PELI, E. 1995. Scrolled and rapid serial visual presentation texts are read at similar rates by the visually impaired. *JOSA A*, 12, 2286-2292.
- FINE, E. M. & PELI, E. 1996. The role of context in reading with central field loss. *Optometry & Vision Science*, 73, 533-539.
- FINE, E. M. & PELI, E. 1998. Benefits of rapid serial visual presentation (RSVP) over scrolled text vary with letter size. *Optometry and vision science: official publication of the American Academy of Optometry*, 75, 191-196.
- FINE, E. M. & RUBIN, G. S. 1999. The effects of simulated cataract on reading with normal vision and simulated central scotoma. *Vision Research*, 39, 4274-4285.
- FLAXMAN, S. R., BOURNE, R. R., RESNIKOFF, S., ACKLAND, P., BRAITHWAITE, T., CICINELLI, M. V., DAS, A., JONAS, J. B., KEEFFE, J. & KEMPEN, J. H. 2017. Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. *The Lancet Global Health*, 5, e1221-e1234.
- FORSTER, K. I. 1970. Visual perception of rapidly presented word sequences of varying complexity. *Perception & psychophysics*, 8, 215-221.
- FRIEDMAN, S. M., MUNOZ, B., RUBIN, G. S., WEST, S. K., BANDEEN-ROCHE, K. & FRIED, L. P. 1999. Characteristics of discrepancies between self-reported visual function and measured reading speed. Salisbury Eye Evaluation Project Team. *Investigative ophthalmology & visual science*, 40, 858-864.
- GOODRICH, G. L. & KIRBY, J. 2001. A comparison of patient reading performance and preference: optical devices, handheld CCTV (Innoventions Magni-Cam), or stand-mounted CCTV (Optelec Clearview or TSI Genie). *Optometry (St. Louis, Mo.)*, 72, 519-528.
- GOWRISANKARAN, S., NAHAR, N. K., HAYES, J. R. & SHEEDY, J. E. 2012. Asthenopia and blink rate under visual and cognitive loads. *Optom Vis Sci*, 89, 97-104.
- GURALNIK, J. M., BRANCH, L. G., CUMMINGS, S. R. & CURB, J. D. 1989. Physical performance measures in aging research. *Journal of Gerontology*, 44, M141-M146.
- HAHN, G., PENKA, D., GEHRLICH, C., MESSIAS, A., WEISMANN, M., HYVÄRINEN, L., LEINONEN, M., FEELY, M., RUBIN, G. & DAUXERRE, C. 2006. New standardised texts for assessing reading performance in four European languages. *British Journal of Ophthalmology*, 90, 480-484.
- HALL, A. 1945. The origin and purposes of blinking. *The British journal of ophthalmology*, 29, 445.
- HARLAND, S., LEGGE, G. E. & LUEBKER, A. 1998. Psychophysics of reading. XVII. Low-vision performance with four types of electronically magnified text. *Optometry and Vision Science*, 75, 183-190.
- HARPER, R., DOORDUYN, K., REEVES, B. & SLATER, L. 1999. Evaluating the outcomes of low vision rehabilitation. *Ophthalmic and Physiological Optics*, 19, 3-11.
- HAYMES, S. A., ROBERTS, K. F., CRUESS, A. F., NICOLELA, M. T., LEBLANC, R. P., RAMSEY, M. S., CHAUHAN, B. C. & ARTES, P. H. 2006. The letter contrast sensitivity test: clinical evaluation of a new design. *Invest Ophthalmol Vis Sci*, 47, 2739-45.
- HOFFMAN, A. C. 1946. Eye-movements during prolonged reading. *Journal of experimental psychology*, 36, 95.
- HOLLAND, M. K. & TARLOW, G. 1972. Blinking and mental load. *Psychol Rep*, 31, 119-27.
- HOLLAND, M. K. & TARLOW, G. 1975. Blinking and thinking. *Percept Mot Skills*, 41, 503-6.
- HOLLIS, J. & ALLEN, P. M. 2006. Screening for Meares-Irlen sensitivity in adults: can assessment methods predict changes in reading speed? *Ophthalmic and Physiological Optics*, 26, 566-571.

- JOHN, D. R. & COLE, C. A. 1986. Age differences in information processing: Understanding deficits in young and elderly consumers. *Journal of consumer research*, 13, 297-315.
- KIM, W., ZANGEMEISTER, W. & STARK, L. 1984. No fatigue effect on blink rate.
- KINCAID, J. P., FISHBURNE, R. P., ROGERS, R. L. & CHISSOM, B. S. 1975. Derivation of new readability formulas (Automated Readability Index, FogCount and Flesch Reading Ease Formula) for Navy enlisted personnel. *In: 8-75*, N. T. C. R. B. R. (ed.).
- KRISCHER, C. C. & MEISSEN, R. 1983. Reading speed under real and simulated visual impairment. *Journal of Visual Impairment & Blindness*, 77, 386-388.
- KRISS, I. & EVANS, B. J. W. 2005. The relationship between dyslexia and Meares-Irlen Syndrome. *Journal of Research in Reading*, 28, 350-364.
- LATHAM, K. 2018. Benefits of low vision aids to reading accessibility. *Vision research*, 153, 47-52.
- LATHAM, K. & TABRETT, D. R. 2012. Guidelines for predicting performance with low vision aids. *Optometry and Vision Science*, 89, 1316-1326.
- LATHAM, K. & WHITAKER, D. 1996. A comparison of word recognition and reading performance in foveal and peripheral vision. *Vision Research*, 36, 2665-2674.
- LEAT, S. J., FRYER, A. & RUMNEY, N. J. 1994. Outcome of low vision aid provision: The effectiveness of a low vision clinic. *Optometry and Vision Science*, 71, 199-206.
- LEGG, G., MANSFIELD, J. & CHUNG, S. The visual span for reading decreases in peripheral vision. *Investigative Ophthalmology & Visual Science*, 1997a. LIPPINCOTT-RAVEN PUBL 227 EAST WASHINGTON SQ, PHILADELPHIA, PA 19106, 1057-1057.
- LEGG, G. E. 2007. *Psychophysics of Reading in Normal and Low Vision*, United States of America, Lawrence Erlbaum Associates, Inc.
- LEGG, G. E., AHN, S. J., KLITZ, T. S. & LUEBKER, A. 1997b. Psychophysics of reading - XVI. The visual span in normal and low vision. *Vision Research*, 37, 1999-2010.
- LEGG, G. E. & BIGELOW, C. A. 2011. Does print size matter for reading? A review of findings from vision science and typography. *J Vis*, 11.
- LEGG, G. E., PELLI, D. G., RUBIN, G. S. & SCHLESKE, M. M. 1985a. Psychophysics of reading-- I. Normal vision. *Vision Res*, 25, 239-52.
- LEGG, G. E., ROSS, J. A., ISENBERG, L. M. & LAMAY, J. M. 1992. Psychophysics of reading. Clinical predictors of low-vision reading speed. *Invest Ophthalmol Vis Sci*, 33, 677-87.
- LEGG, G. E., ROSS, J. A., LUEBKER, A. & LAMAY, J. M. 1989a. Psychophysics of reading -- VIII. The Minnesota Low-Vision Reading Test. *Optometry and Vision Science*, 66, 843-853.
- LEGG, G. E., ROSS, J. A., MAXWELL, K. T. & LUEBKER, A. 1989b. Psychophysics of reading - VII. Comprehension in normal and low vision. *Clinical Vision Sciences*, 4, 51-60.
- LEGG, G. E., RUBIN, G. S. & LUEBKER, A. 1987. Psychophysics of reading--V. The role of contrast in normal vision. *Vision Res*, 27, 1165-77.
- LEGG, G. E., RUBIN, G. S., PELLI, D. G. & SCHLESKE, M. M. 1985b. Psychophysics of reading-- II. Low vision. *Vision Res*, 25, 253-65.
- LOTT, L. A., SCHNECK, M. E., HAEGERSTRÖM-PORTNOY, G., BRABYN, J. A., GILDENGORIN, G. L. & WEST, C. G. 2001. Reading performance in older adults with good acuity. *Optom Vis Sci*, 78, 316-24.
- LOVIE-KITCHIN, J. E., BOWERS, A. R. & WOODS, R. L. 2000. Oral and silent reading performance with macular degeneration. *Ophthalmic Physiol Opt*, 20, 360-70.
- LOVIE-KITCHIN, J. E. & BROWN, B. 2000. Repeatability and intercorrelations of standard vision tests as a function of age. *Optometry and vision science*, 77, 412-420.
- LOVIE-KITCHIN, J. E. & WHITTAKER, S. G. 1999. Prescribing near magnification for low vision patients. *Clin Exp Optom*, 82, 214-224.
- LOVIE-KITCHIN, J. E. & WOO, G. C. 1988. Effect of magnification and field of view on reading speed using a CCTV. *Ophthalmic and Physiological Optics*, 8, 139-145.

- LOWE, J. B. & DRASDO, N. 1990. Efficiency in reading with closed-circuit television for low vision. *Ophthalmic and Physiological Optics*, 10, 225-233.
- LUCKIESH, M. 1947. Reading and the rate of blinking. *Journal of experimental psychology*, 37, 266.
- LUCKIESH, M. & MOSS, F. K. 1939. Frequency of blinking as a clinical criterion of ease of seeing. *The Australasian Journal of Optometry*, 22, 501-508.
- MAAIJWEE, K., MULDER, P., RADNER, W. & VAN MEURS, J. C. 2008. Reliability testing of the Dutch version of the Radner Reading Charts. *Optom Vis Sci*, 85, 353-8.
- MAFFEI, A. & ANGRILLI, A. 2018. Spontaneous eye blink rate: An index of dopaminergic component of sustained attention and fatigue. *International Journal of Psychophysiology*, 123, 58-63.
- MANSFIELD, J. S., AHN, S. J., LEGGE, G. E. & LUEBKER, A. 1993. A new reading-acuity chart for normal and low vision. *Ophthalmic Visual Optics/ Non-Invasive Assessment of the Visual System Technical Digest*. Washington DC: Optical Society of America.
- MANSFIELD, J. S., LEGGE, G. E. & BANE, M. C. 1996. Psychophysics of reading. XV: Font effects in normal and low vision. *Invest Ophthalmol Vis Sci*, 37, 1492-501.
- MANSFIELD, J. S., LEGGE, G. E., LUEBKER, A. & CUNNINGHAM, K. 1994. MNRead Acuity Charts -Continuous-text reading acuity charts for normal and low vision. University of Minnesota -Minneapolis, MN 55455: Minnesota Laboratory for Low-Vision Research.
- MANTYJARVI, M. & LAITINEN, T. 2001. Normal values for the Pelli-Robson contrast sensitivity test. *Journal of Cataract and Refractive Surgery*, 27, 261-266.
- MARTINS, R. & CARVALHO, J. 2015. Eye blinking as an indicator of fatigue and mental load-a systematic review. *Occupational Safety and Hygiene III*, 231-235.
- MCALINDEN, C., KHADKA, J. & PESUDOV, K. 2011. Statistical methods for conducting agreement (comparison of clinical tests) and precision (repeatability or reproducibility) studies in optometry and ophthalmology. *Ophthalmic and Physiological Optics*, 31, 330-338.
- MCCLURE, M., HART, P., JACKSON, A., STEVENSON, M. & CHAKRAVARTHY, U. 2000. Macular degeneration: do conventional measurements of impaired visual function equate with visual disability? *British Journal of Ophthalmology*, 84, 244-250.
- MCCONKIE, G. W. & RAYNER, K. 1975. Span of effective stimulus during a fixation in reading. *Perception & Psychophysics*, 17, 578-586.
- MCILWAINE, G. G., BELL, J. A. & DUTTON, G. N. 1991. Low vision aids—is our service cost effective? *Eye*, 5, 607-611.
- MCMONNIES, C. W. 2001. Control of chart memory for retesting acuity. *Clinical and Experimental Optometry*, 84, 78-84.
- MOHAMMED, Z. & DICKINSON, C. M. 2000. The inter-relationship between magnification, field of view and contrast reserve: the effect on reading performance. *Ophthalmic Physiol Opt*, 20, 464-72.
- MOKKINK, L. B., TERWEE, C. B., PATRICK, D. L., ALONSO, J., STRATFORD, P. W., KNOL, D. L., BOUTER, L. M. & DE VET, H. C. 2010. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of life research*, 19, 539-549.
- MORRICE, E., JOHNSON, A., MARINIER, J. & WITTICH, W. 2017. Assessment of the Apple iPad as a low-vision reading aid. *Eye*, 31, 865.
- NAHAR, N. K., SHEEDY, J. E., HAYES, J. & TAI, Y. C. 2007. Objective measurements of lower-level visual stress. *Optom Vis Sci*, 84, 620-9.
- O'REGAN, J. 1990. Eye movements and reading. *Reviews of oculomotor research*, 4, 395-453.

- ODELL, N. V., LESKE, D. A., HATT, S. R., ADAMS, W. E. & HOLMES, J. M. 2008. The effect of Bangerter filters on optotype acuity, Vernier acuity, and contrast sensitivity. *J AAPOS*, 12, 555-9.
- ORCHARD, L. N. & STERN, J. A. 1991. Blinks as an index of cognitive activity during reading. *Integrative Physiological and Behavioral Science*, 26, 108-116.
- ORTIZ, A., CHUNG, S., LEGGE, G. E. & JOBLING, J. T. 1999. Reading with a head-mounted video magnifier. *Optometry and vision science: official publication of the American Academy of Optometry*, 76, 755-763.
- OWSLEY, C. 2011. Aging and vision. *Vision Res*, 51, 1610-22.
- OWSLEY, C., MCGWIN, G., LEE, P. P., WASSERMAN, N. & SEARCEY, K. 2009. Characteristics of low-vision rehabilitation services in the United States. *Arch Ophthalmol*, 127, 681-9.
- OWSLEY, C., MCGWIN, G., SLOANE, M. E., STALVEY, B. T. & WELLS, J. 2001. Timed instrumental activities of daily living tasks: Relationship to visual function in older adults. *Optometry and Vision Science*, 78, 350-359.
- PATEL, P. J., CHEN, F. K., DA CRUZ, L., RUBIN, G. S. & TUFAIL, A. 2011. Test-retest variability of reading performance metrics using MNREAD in patients with age-related macular degeneration. *Invest Ophthalmol Vis Sci*, 52, 3854-9.
- PERERA, S. 2004. LPfont: An investigation into the legibility of large print typefaces. *Online: www.tiresias.org/fonts [accessed May 2005]*.
- PFAFF, U., FRUHSTORFER, H. & PETER, J. Changes in eye-blink duration and frequency during car driving. PFLUGERS ARCHIV-EUROPEAN JOURNAL OF PHYSIOLOGY, 1976. SPRINGER VERLAG 175 FIFTH AVE, NEW YORK, NY 10010, R21-R21.
- POULTON, E. & GREGORY, R. L. 1952. Blinking during visual tracking. *Quarterly Journal of Experimental Psychology*, 4, 57-65.
- RADNER, W., OBERMAYER, W., RICHTER-MUEKSCH, S., WILLINGER, U., VELIKAY-PAREL, M. & EISENWORT, B. 2002. The validity and reliability of short German sentences for measuring reading speed. *Graefes Arch Clin Exp Ophthalmol*, 240, 461-7.
- RAMULU, P. Y., SWENOR, B. K., JEFFERYS, J. L., FRIEDMAN, D. S. & RUBIN, G. S. 2013a. Difficulty with out-loud and silent reading in glaucoma. *Investigative ophthalmology & visual science*, 54, 666.
- RAMULU, P. Y., SWENOR, B. K., JEFFERYS, J. L. & RUBIN, G. S. 2013b. Description and validation of a test to evaluate sustained silent reading. *Invest Ophthalmol Vis Sci*, 54, 673-680.
- RAYNER, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychol Bull*, 124, 372-422.
- RAYNER, K., REICHLER, E. D., STROUD, M. J., WILLIAMS, C. C. & POLLATSEK, A. 2006. The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and aging*, 21, 448.
- RAYNER, K., WELL, A. D. & POLLATSEK, A. 1980. Asymmetry of the effective visual-field in reading. *Perception & Psychophysics*, 27, 537-544.
- RECARTE, M. Á., PÉREZ, E., CONCHILLO, Á. & NUNES, L. M. 2008. Mental workload and visual impairment: Differences between pupil, blink, and subjective rating. *The Spanish journal of psychology*, 11, 374-385.
- RODRÍGUEZ-ARANDA, C. 2003. Reduced writing and reading speed and age-related changes in verbal fluency tasks. *The Clinical Neuropsychologist*, 17, 203-215.
- RUBIN, G. 2017. First results from the EFFECT Trial, an RCT of eccentric viewing training for patients with AMD. *Investigative Ophthalmology & Visual Science*, 58, 4766-4766.
- RUBIN, G. S. 2013. Measuring reading performance. *Vision Res*.
- RUBIN, G. S., FEELY, M., PERERA, S., EKSTROM, K. & WILLIAMSON, E. 2006. The effect of font and line width on reading speed in people with mild to moderate vision loss. *Ophthalmic and Physiological Optics*, 26, 545-554.

- RUBIN, G. S. & LEGGE, G. E. 1989. Psychophysics of reading. VI--The role of contrast in low vision. *Vision Res*, 29, 79-91.
- RUBIN, G. S. & TURANO, K. 1992. Reading without saccadic eye movements. *Vision Research*, 32, 895-902.
- RUBIN, G. S. & TURANO, K. 1994. Low vision reading with sequential word presentation. *Vision research*, 34, 1723-1733.
- RUMNEY, N. J. 1995. Using visual thresholds to establish low vision performance. *Ophthalmic and Physiological Optics*, 15, S18-S24.
- SALTHOUSE, T. A. 1996. The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103, 403-428.
- SALTHOUSE, T. A. 2004. What and when of cognitive aging. *Current directions in psychological science*, 13, 140-144.
- SASS, S. M., LEGGE, G. E. & LEE, H. W. 2006. Low-vision reading speed: influences of linguistic inference and aging. *Optom Vis Sci*, 83, 166-77.
- SCHAKEL, W., BODE, C., VAN DER AA, H. P., HULSHOF, C. T., BOSMANS, J. E., VAN RENS, G. H. & VAN NISPEN, R. M. 2017. Exploring the patient perspective of fatigue in adults with visual impairment: a qualitative study. *BMJ open*, 7, e015023.
- SHEEDY, J. E., GOWRISANKARAN, S. & HAYES, J. R. 2005. Blink rate decreases with eyelid squint. *Optom Vis Sci*, 82, 905-11.
- SMITH, B. L., WASOWICZ, J. & PRESTON, J. 1987. Temporal characteristics of the speech of normal elderly adults. *Journal of Speech and Hearing Research*, 30, 522-529.
- STANGLER-ZUSCHROTT, E. 1990. Decreased reading speed and rapid fatigue as signs of disordered visual function. *Klinische Monatsblätter für Augenheilkunde*, 196, 150-157.
- STELMACK, J., REDA, D., AHLERS, S., BAINBRIDGE, L. & MCCRAY, J. 1991. Reading performance of geriatric patients post exudative maculopathy. *Journal of the American Optometric Association*, 62, 53-57.
- STELMACK, J., STELMACK, T. R., FRAIM, M. & WARRINGTON, J. 1987. Clinical use of the Pepper Visual Skills for Reading Test in low vision rehabilitation. *Am J Optom Physiol Opt*, 64, 829-31.
- STERN, J., BEIDEMAN, L. & CHEN, S. 1976. Effect of alcohol on visual search and motor performance during complex task performance.
- STERN, J. A., BOYER, D. & SCHROEDER, D. 1994. Blink rate: a possible measure of fatigue. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 36, 285-297.
- STERN, J. A., WALRATH, L. C. & GOLDSTEIN, R. 1984. The endogenous eyeblink. *Psychophysiology*, 21, 22-33.
- SUBRAMANIAN, A. & PARDHAN, S. 2006. The repeatability of MNREAD acuity charts and variability at different test distances. *Optom Vis Sci*, 83, 572-6.
- SUBRAMANIAN, A. & PARDHAN, S. 2009. Repeatability of reading ability indices in subjects with impaired vision. *Invest Ophthalmol Vis Sci*, 50, 3643-7.
- SUMMALA, H., HAKKANEN, H., MIKKOLA, T. & SINKKONEN, J. 1999. Task effects on fatigue symptoms in overnight driving. *Ergonomics*, 42, 798-806.
- SUN, W. S., BAKER, R. S., CHUKE, J. C., ROUHOLIMAN, B. R., HASAN, S. A., GAZA, W., STAVA, M. W. & PORTER, J. D. 1997. Age-related changes in human blinks. Passive and active changes in eyelid kinematics. *Investigative ophthalmology & visual science*, 38, 92-99.
- TARITA-NISTOR, L., GONZÁLEZ, E. G., MANDELCORN, M. S., BRENT, M. H., MARKOWITZ, S. N. & STEINBACH, M. J. 2018. The reading accessibility index and quality of reading grid of patients with central vision loss. *Ophthalmic and Physiological Optics*, 38, 88-97.

- TAYLOR, J., BAMBRICK, R., DUTTON, M., HARPER, R., RYAN, B., TUDOR-EDWARDS, R., WATERMAN, H., WHITAKER, C. & DICKINSON, C. 2014. The p-EVES study design and methodology: a randomised controlled trial to compare portable electronic vision enhancement systems (p-EVES) to optical magnifiers for near vision activities in visual impairment. *Ophthalmic and Physiological Optics*, 34, 558-572.
- TAYLOR, J. J., BAMBRICK, R., BRAND, A., BRAY, N., DUTTON, M., HARPER, R. A., HOARE, Z., RYAN, B., EDWARDS, R. T., WATERMAN, H. & DICKINSON, C. 2017. Effectiveness of portable electronic and optical magnifiers for near vision activities in low vision: a randomised crossover trial. *Ophthalmic Physiol Opt*, 37, 370-384.
- TELFORD, C. & THOMPSON, N. 1933. Some factors influencing voluntary and reflex eyelid responses. *Journal of Experimental Psychology*, 16, 524.
- TINKER, M. A. 1946. Validity of frequency of blinking as a criterion of readability. *Journal of experimental psychology*, 36, 453.
- TRAUZETTEL-KLOSINSKI, S., DIETZ, K. & GROUP, T. I. S. 2012. Standardized Assessment of Reading Performance: The New International Reading Speed Texts IReST. *Invest Ophthalmol Vis Sci*, 53, 5452-5461.
- TSUBOTA, K. & NAKAMORI, K. 1993. Dry eyes and video display terminals. *N Engl J Med*, 328, 584.
- TSUBOTA, K. & NAKAMORI, K. 1995. Effects of ocular surface area and blink rate on tear dynamics. *Archives of Ophthalmology*, 113, 155-158.
- VAN DEN BOSCH, W. A., LEENDERS, I. & MULDER, P. 1999. Topographic anatomy of the eyelids, and the effects of sex and age. *British journal of ophthalmology*, 83, 347-352.
- VIRGILI, G., ACOSTA, R., BENTLEY, S. A., GIACOMELLI, G., ALLCOCK, C. & EVANS, J. R. 2018. Reading aids for adults with low vision. *Cochrane Database of Systematic Reviews*.
- VIRGILI, G., CORDARO, C., BIGONI, A., CROVATO, S., CECCHINI, P. & MENCHINI, U. 2004a. Reading acuity in children: evaluation and reliability using MNREAD charts. *Invest Ophthalmol Vis Sci*, 45, 3349-54.
- VIRGILI, G., PIERROTTET, C., PARMEGGIANI, F., PENNINO, M., GIACOMELLI, G., STEINDLER, P., MENCHINI, U., ORZALESI, N. & CHARTS, M. 2004b. Reading performance in patients with retinitis pigmentosa: a study using the MNREAD charts. *Invest Ophthalmol Vis Sci*, 45, 3418-24.
- WALKER, R., BRYAN, L., HARVEY, H., RIAZI, A. & ANDERSON, S. J. 2016. The value of Tablets as reading aids for individuals with central visual field loss: an evaluation of eccentric reading with static and scrolling text. *Ophthalmic and Physiological Optics*, 36, 459-464.
- WATSON, G., BALDASARE, J. & WHITTAKER, S. 1990. The validity and clinical uses of the pepper visual skills for reading test. *Journal of Visual Impairment & Blindness*, 84, 119-123.
- WATSON, G. R., DE L'AUNE, W., LONG, S., MAINO, J. & STELMACK, J. 1997. Veterans' use of low vision devices for reading. *Optometry and Vision Science*, 74, 260-265.
- WATSON, G. R., WRIGHT, V. & DE L'AUNE, W. 1992. The efficacy of comprehension training and reading practice for print readers with macular loss. *Journal of Visual Impairment & Blindness*.
- WATSON, G. R., WRIGHT, V., LONG, S. & DELAUNE, W. 1996. A low vision reading comprehension test. *Journal of Visual Impairment & Blindness*, 90, 486-494.
- WATSON, G. R., WRIGHT, V. & LONG, S. L. 2017. *Morgan low vision reading comprehension assessment*, Homer Printing Company.
- WEALE, R. A. 1963. *The aging eye*, Hoeber Medical Division, Harper & Row.
- WEST, S. K., RUBIN, G. S., MUNOZ, B., ABRAHAM, D. & FRIED, L. P. 1997. Assessing functional status: correlation between performance on tasks conducted in a clinic setting and

- performance on the same task conducted at home. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 52, M209-M217.
- WHITE, J. M. & BEDELL, H. E. 1990. The oculomotor reference in humans with bilateral macular disease. *Investigative ophthalmology & visual science*, 31, 1149-1161.
- WHITTAKER, S. G. & LOVIE-KITCHIN, J. 1993. Visual requirements for reading. *Optom Vis Sci*, 70, 54-65.
- WHITTAKER, S. G. & LOVIE-KITCHIN, J. E. 1994. The assessment of contrast sensitivity and contrast reserve for reading rehabilitation. *Low Vision: Research and New Developments in Rehabilitation*, 11, 88-92.
- WILKINS, A. J., JEANES, R. J., PUMFREY, P. D. & LASKIER, M. 1996. Rate of Reading Test: its reliability, and its validity in the assessment of the effects of coloured overlays. *Ophthalmic Physiol Opt*, 16, 491-7.
- WONG, K., WAN, W. & KAYE, S. 2002. Blinking and operating: cognition versus vision. *British journal of ophthalmology*, 86, 479-479.
- WOOD, J., CHAPARRO, A., ANSTEY, K., LACHEREZ, P., CHIDGEY, A., EISEMANN, J., GAYNOR, A. & LA, P. 2010. Simulated visual impairment leads to cognitive slowing in older adults. *Optometry & Vision Science*, 87, 1037-1043.
- WOOD, J. M., CHAPARRO, A., ANSTEY, K. J., HSING, Y. E., JOHNSSON, A. K., MORSE, A. L. & WAINWRIGHT, S. E. 2009. Impact of simulated visual impairment on the cognitive test performance of young adults. *British Journal of Psychology*, 100, 593-602.
- XIONG, Y.-Z., CALABRESE, A., CHEONG, A. M. & LEGGE, G. E. 2018a. Reading Acuity as a Predictor of Low-Vision Reading Performance. *Investigative ophthalmology & visual science*, 59, 4798-4803.
- XIONG, Y.-Z., LORSUNG, E. A., MANSFIELD, J. S., BIGELOW, C. & LEGGE, G. E. 2018b. Fonts Designed for Macular Degeneration: Impact on Reading. *Investigative ophthalmology & visual science*, 59, 4182-4189.
- XU, R. & BRADLEY, A. 2015. IUR ead: a new computer-based reading test. *Ophthalmic and Physiological Optics*, 35, 500-513.
- YAGER, D., AQUILANTE, K. & PLASS, R. 1998. High and low luminance letters, acuity reserve, and font effects on reading speed. *Vision Research*, 38, 2527-2531.
- YAP, M. 1991. Tear break-up time is related to blink frequency. *Acta ophthalmologica*, 69, 92-94.
- YORK, M., ONG, J. & ROBBINS, J. C. 1971. *Variation in blink rate associated with contact lens wear and task difficulty*.
- ZAMAN, M., DOUGHTY, M. & BUTTON, N. 1998. The exposed ocular surface and its relationship to spontaneous eyeblink rate in elderly Caucasians. *Experimental eye research*, 67, 681-686.
- ZAMAN, M. L. & DOUGHTY, M. J. 1997. Some methodological issues in the assessment of the spontaneous eyeblink frequency in man. *Ophthalmic Physiol Opt*, 17, 421-32.
- ZAMETKIN, A. J., STEVENS, J. R. & PITTMAN, R. 1979. Ontogeny of spontaneous blinking and of habituation of the blink reflex. *Ann Neurol*, 5, 453-7.

9. Appendices

9.1 Samples of reading text

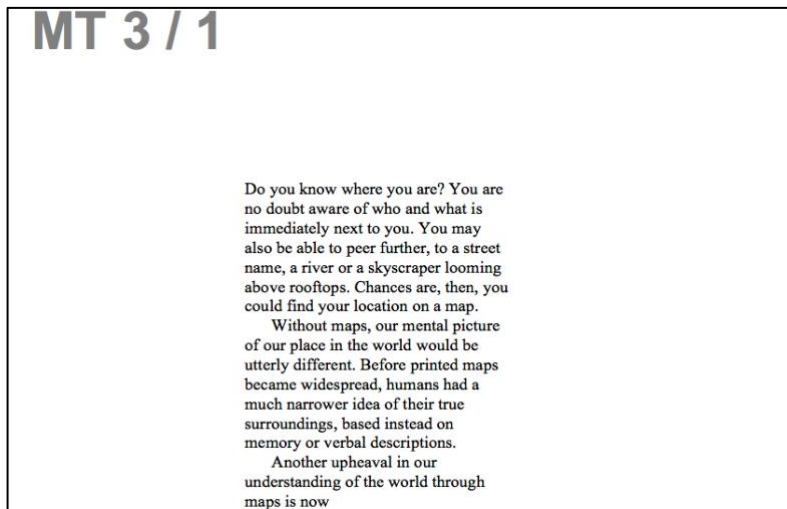


Figure 9.1 A sample of MT reading text (font size and spacing not up to scale)

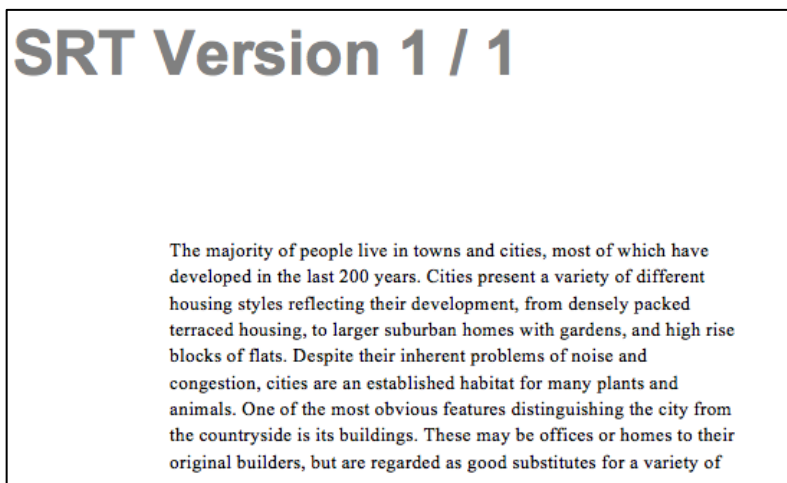


Figure 9.2 A sample of SRT reading text (font size and spacing not up to scale)

9.2 Validation of different versions of magazine text and SRT

In this study, we created two reading texts, MT & SRT, in three different versions using available reading materials from science encyclopaedia to extend the reading duration during the assessment. However, the technical difficulty of the reading material can affect the reading speed performance between the three versions. To determine if the level of text difficulty is similar for the three versions, reading speed performances were grouped according to the different versions. The full details of the procedure can be found in Chapter 3. We hypothesised if any one of the reading text version was relatively more difficult than the others, we would expect the respective version of MT or SRT to have a consistently slower reading speed compared to the other reading text versions.

Table 9.1 shows the reading speed for the participants for different versions of the MT and SRT. A Kruskal-Wallis test revealed no statistically significant differences between the three versions of test across the three conditions. In another words, any changes in the reading speed in the subsequent reading assessments using the three different versions are unlikely to be due to the level of text difficulty in this study.

Table 9.1 Comparison of reading speed in all the different version of MT and SRT

Reading speed in wpm				
MT	Version 1	Version 2	Version 3	P value
NOR	159.91, n=9	159.18, n=10	152.96, n=10	0.57
BOF	151.46, n=10	159.31, n=10	151.29, n=9	0.46
MAG	103.71, n=10	110.31, n=9	117.83, n=10	0.27
SRT	Version 1	Version 2	Version 3	P value
NOR	264.12, n=9	212.38, n=10	249.88, n=10	0.47
BOF	251.85, n=10	244.33, n=10	209.75, n=9	0.30
MAG	116.03, n=10	131.50, n=9	122.52, n=10	0.85

9.3 Multiple-choice questions for SRT Version 1

SRT Version 1

Question 1 (page 4)

The _____ -summer butterflies feed on the _____ which covers many walls in suburban areas.

- A. mid, ivy
- B. end, ivy
- C. mid, lilies
- D. end, lilies

Question 2 (page 9)

The attic, one of the _____ disturbed places in the house, can harbour a variety of _____.

- A. most, creatures
- B. least, creatures
- C. least, insects
- D. most, insects

Question 3 (page 13)

As this land in turn was developed, _____ moved into the surrounding suburbs and from there to the city centres.

- A. wolves
- B. wild dogs
- C. snakes
- D. foxes

Question 4 (page 18)

However, it is the _____ species' capacity for survival as either fast growing annuals, or almost _____ perennials, that have guaranteed their vigorous history of success

- A. plants, destructible
- B. weed, destructible
- C. weed, indestructible
- D. plants indestructible

Question 5 (Page 23)

Therefore, during the heat of the day, creatures such as earthworms, slugs and _____ _____
_____ the paving slabs.

- A. millipedes creep out above
- B. millipedes hid away under
- C. centipedes creep out above
- D. centipedes hid away under

Question 6 (page 29)

No matter how carefully the _____ is tended, these colonising _____ inevitably find a foothold.

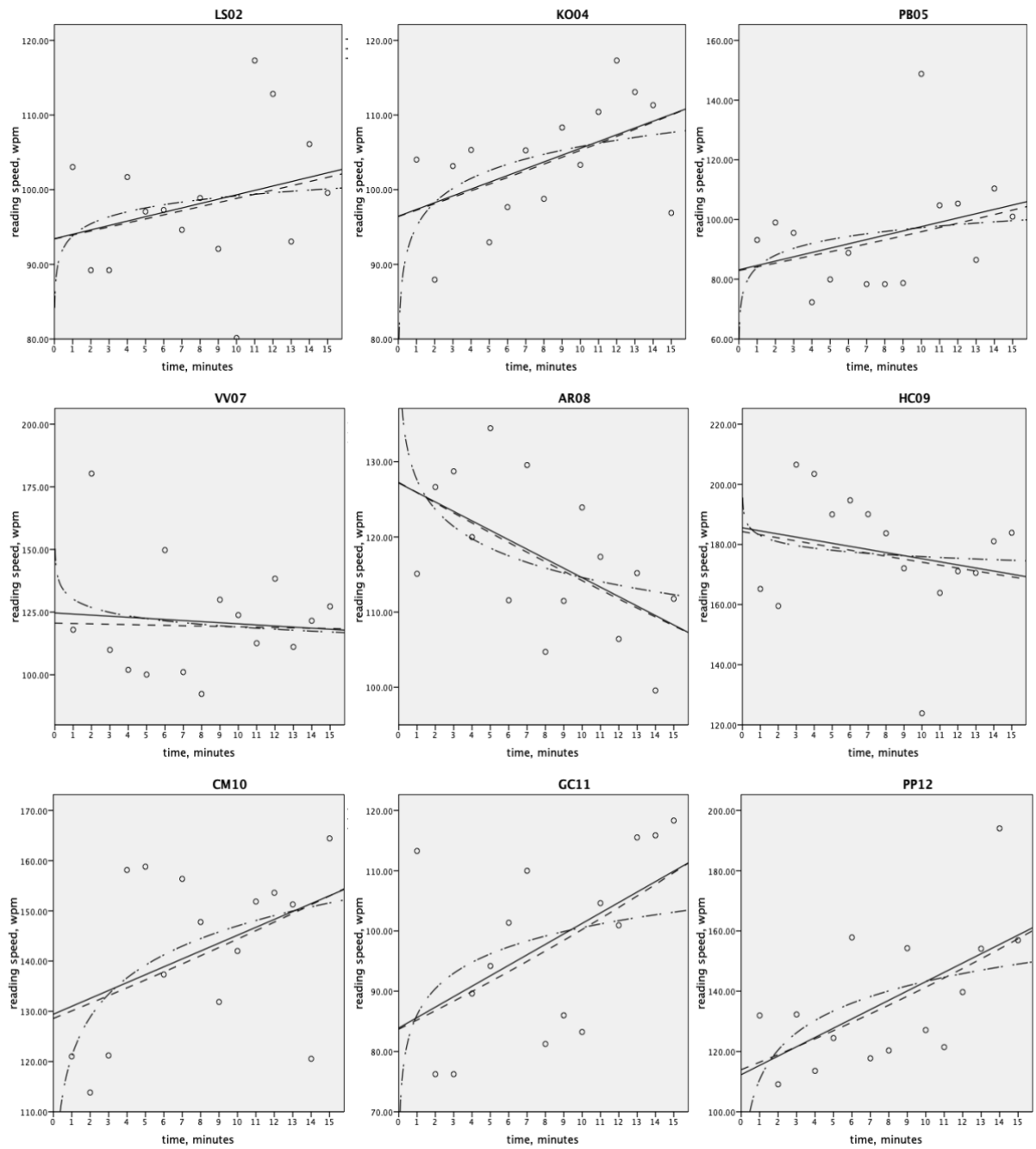
- A. garden, weeds
- B. park, plants
- C. park, weeds
- D. garden, plants

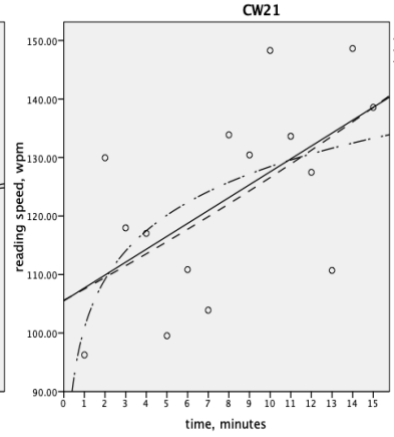
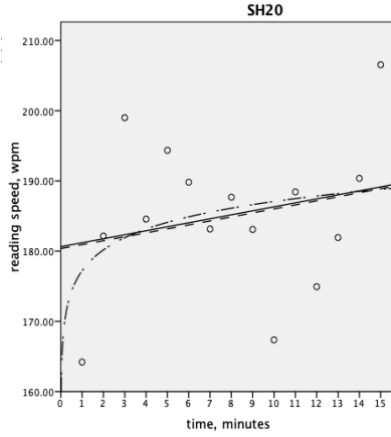
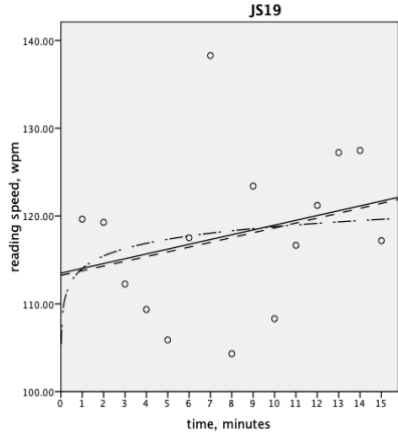
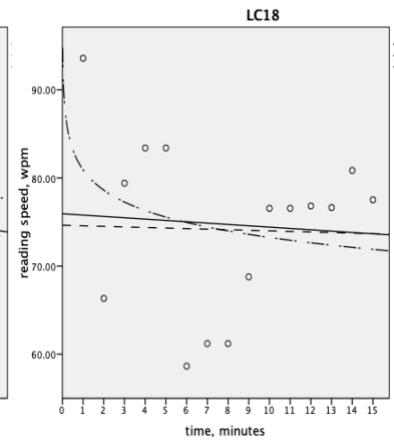
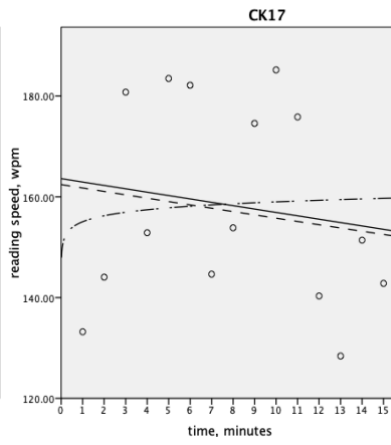
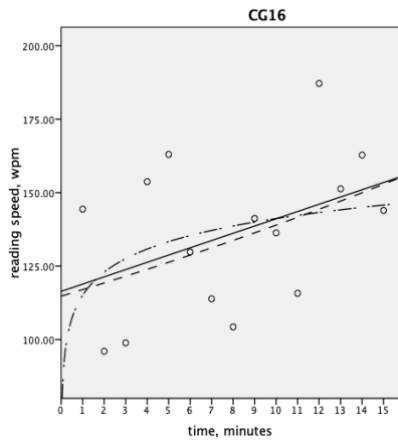
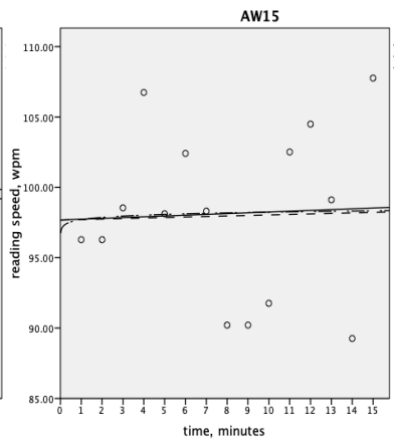
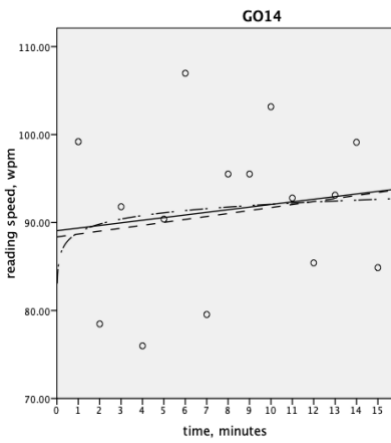
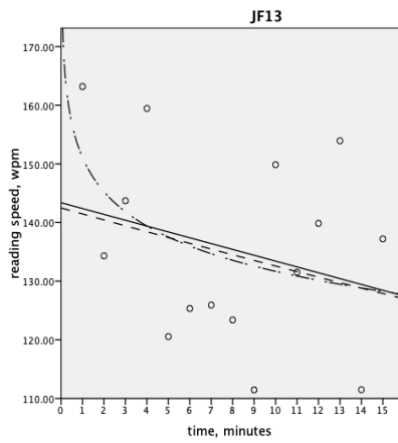
Question 7 (page 35)

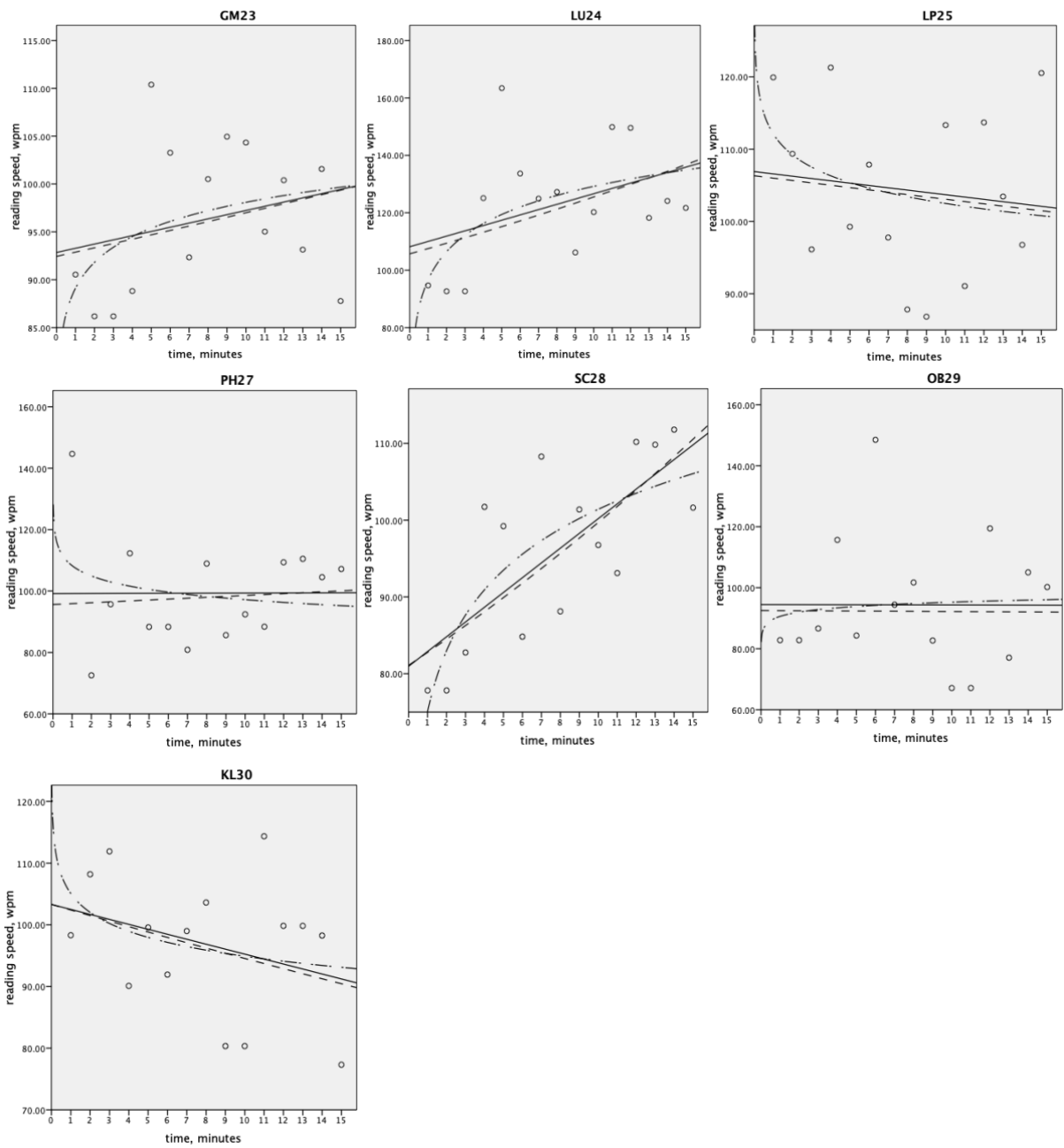
Where concrete _____ been too liberally applied, a gentle bank allows _____ to climb, or hedgehogs to escape if they fall in whilst trying to get a drink

- A. has not, ducklings
- B. has not, chicks
- C. has, ducklings
- D. has, chicks

9.4 Nonlinear graph plots for all participants in SRT MAG (n = 25)







A total of 25 participants graph plots on reading speed against time (minutes) graph plots in SRT-MAG. Participant JY03 graph plot was not reflected in this appendix as it was shown as an example in Figure 3.6.

Table 9.2 R² values for the young participants using logarithmic and exponential functions on reading speed against time (SRT-MAG)

Participants	R ² values (n = 26) in SRT-MAG		
	Linear	Logarithmic	Exponential
LS02	0.08	0.04	0.07
JY03	0.03	0.00	0.04
KO04	0.26	0.21	0.26
PB05	0.11	0.05	0.12
VV07	0.01	0.03	0.00
AR08	0.31	0.19	0.32
HC09	0.05	0.01	0.04
CM10	0.18	0.28	0.19
GC11	0.27	0.11	0.27
JF13	0.08	0.16	0.07
PP12	0.37	0.24	0.38
GO14	0.02	0.01	0.03
AW15	0.00	0.00	0.00
CG16	0.18	0.11	0.19
CK17	0.02	0.05	0.22
LC18	0.01	0.07	0.00
JS19	0.07	0.03	0.07
SH20	0.05	0.09	0.05
CW21	0.35	0.32	0.35
GM23	0.06	0.15	0.07
LU24	0.16	0.28	0.20
LP25	0.02	0.08	0.02
PH27	0.00	0.05	0.01
SC28	0.54	0.58	0.55
OB29	0.00	0.01	0.00
KL30	0.10	0.10	0.11

9.5 Abstract Oral Presentations – OS9 Page 26 from European Society for Low Vision Research and Rehabilitation (ESLRR) 2013

Title: Reading duration effects in young and elderly subjects with simulated visual impairment

Background: Most published reading tests are short and reading endurance is not considered clinically. Thus, it is difficult to determine whether a visually impaired patient can cope with leisure reading. The availability of a longer reading test such as IReST (with approximately 150 word passages), may allow determination of duration related effects. However, it has been suggested that this effect may only be apparent over a reading period of at least 15 minutes (Ramulu et al, 2013).

Methods: Thirty young normal (mean 21.77 years) and six elderly subjects (mean 61.67 years) read with fully-corrected vision, with visual impairment simulated using Bangerter occlusion foils (BOF), and with the BOF and a hand-held magnifier. IReST (audio-recording of reading out-loud) and a sustained reading text (SRT) (silent reading timed by page-turns) were used. Subjects read two consecutive IReST passages (approximately 300 words) and segments at the beginning and end were compared to calculate the relative reading rate (RRR). The SRT had 3500 words to allow longer reading duration up to 10-15 minutes. Reading speed was documented at 10 and 15 minutes interval and compared relative to the 2nd minute reading speed.

Results: In the young group, neither IReST nor SRT showed a slowing down of reading speed with time (RRR 1.04-1.10 and 1.08-1.13 respectively). Likewise, no reading speed slowing was seen using IReST in the elderly group (RRR 1.02-1.08). However, a slowing down effect was observed in SRT. The RRR for SRT was 0.84-0.96 which was statistically significantly different from IReST, $F(1, 30) = 7.69, p=0.01$.

Discussion: Reading speed decelerates with time in the elderly but not the young individuals. This effect is greater when the reading task increases in difficulty (i.e reduced contrast

sensitivity and use of a magnifier). However this effect was not reflected using the IReST, even with extended passages.

Conclusion: Reading speed slows in older subjects with simulated impairment, but short reading passages do not show this effect.

9.6 R² values for the elderly participants using logarithmic and exponential functions on reading speed against time (SRT-MAG)

Participants	R ² values (n = 31) in SRT-MAG		
	Linear	Logarithmic	Exponential
06ML	0.19	0.12	0.24
08HK	0.61	0.80	0.81
09CL	0.00	0.06	0.00
11CS	0.09	0.04	0.10
12AO	0.14	0.17	0.16
13ZB	0.01	0.07	0.02
14SS	0.63	0.63	0.56
15KC	0.67	0.59	0.67
17AN	0.13	0.35	0.16
18JL	0.00	0.02	0.01
19FW	0.67	0.52	0.63
20CK	0.06	0.09	0.05
21SL	0.04	0.00	0.02
22WL	0.01	0.04	0.09
23SH	0.05	0.19	0.03
25CC	0.04	0.21	0.06
26SL	0.24	0.14	0.28
27MY	0.12	0.05	0.09
28KN	0.03	0.01	0.03
29SH	0.05	0.21	0.02
30CC	0.00	0.02	0.00
34CK	0.14	0.17	0.13
35LC	0.78	0.82	0.76
36PC	0.33	0.21	0.48
37HS	0.73	0.46	0.76
38SM	0.71	0.73	0.72
39JC	0.64	0.37	0.63
40ST	0.20	0.11	0.22
42GA	0.72	0.77	0.73
43SR	0.04	0.03	0.04
44FK	0.86	0.69	0.86

The alternative curve fittings logarithmic and exponential functions from SPSS to fit the data points reading speed against time. Friedman test showed that there was no statistically significant difference in the R² values ($\chi^2 = 0.735$, $p = 0.692$). Visual inspection of the data also did not suggest the logarithmic and exponential had a better curve fittings in those cases where the linear R² was poor (i.e. close to zero).

9.7 The use of Log_{10} transformation on reading speed to compute reading slope for the p-EVES data

In view that the reading speeds were highly variable, an attempt on Log_{10} transformation for the reading speed was made to compute the reading slope. Figure 9.3 shows the sample graph plots of the use of Log_{10} transformation on the reading speed, and R^2 value was marginally better compared to using the reading speed for computation. Table 9.3 shows the comparison of the log_{10} reading speed and the reading speed data. Although paired t-test showed a marginally statistically significant difference ($t = -2.19$, $p = 0.032$), the range of R^2 value did not appear to improve significantly.

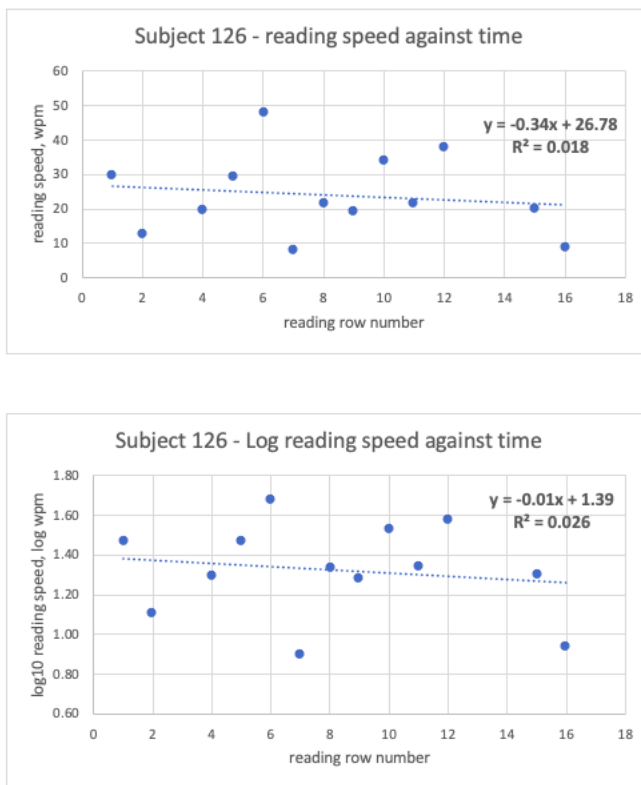
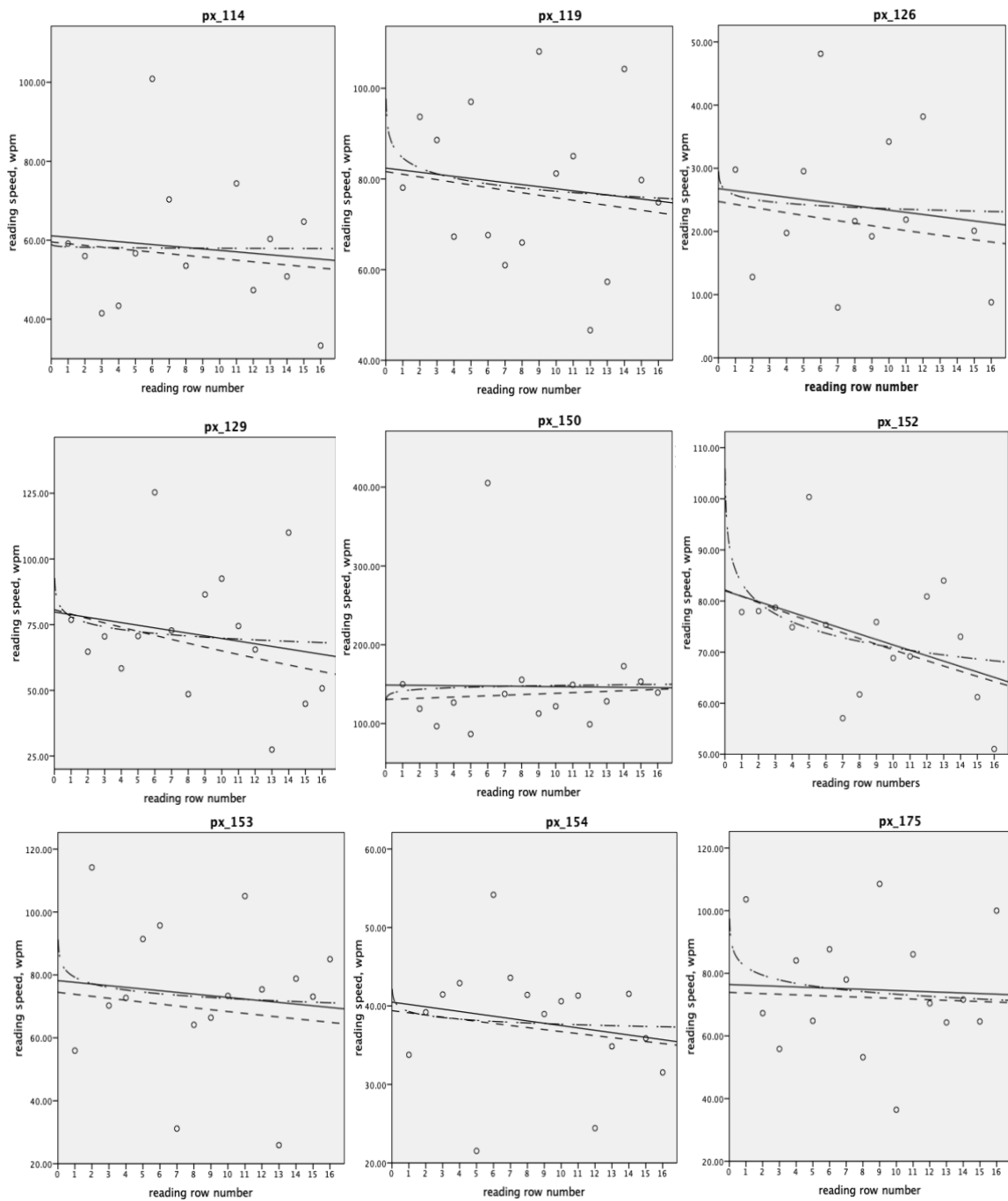


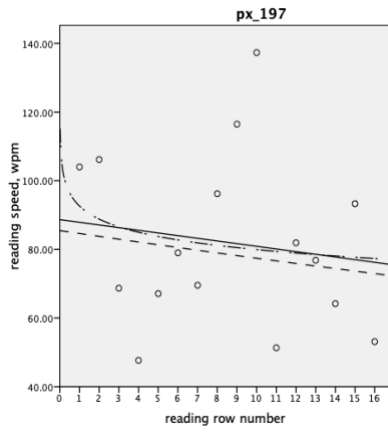
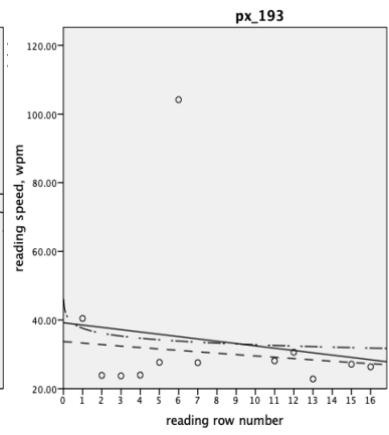
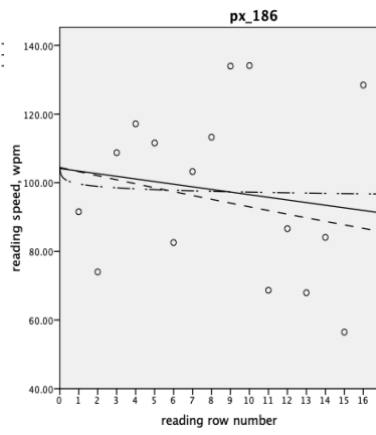
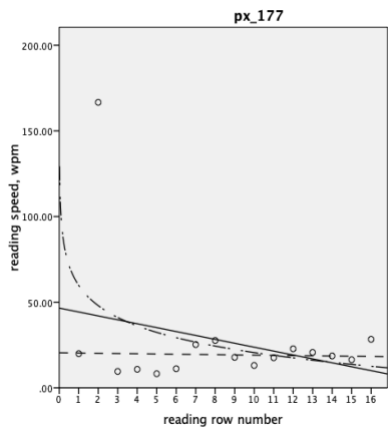
Figure 9.3 Sample graph plots showing the reading slope using reading speed and Log_{10} reading speed transformation

Table 9.3 Descriptive statistics for the R² computation using reading speed and with Log10 transformation

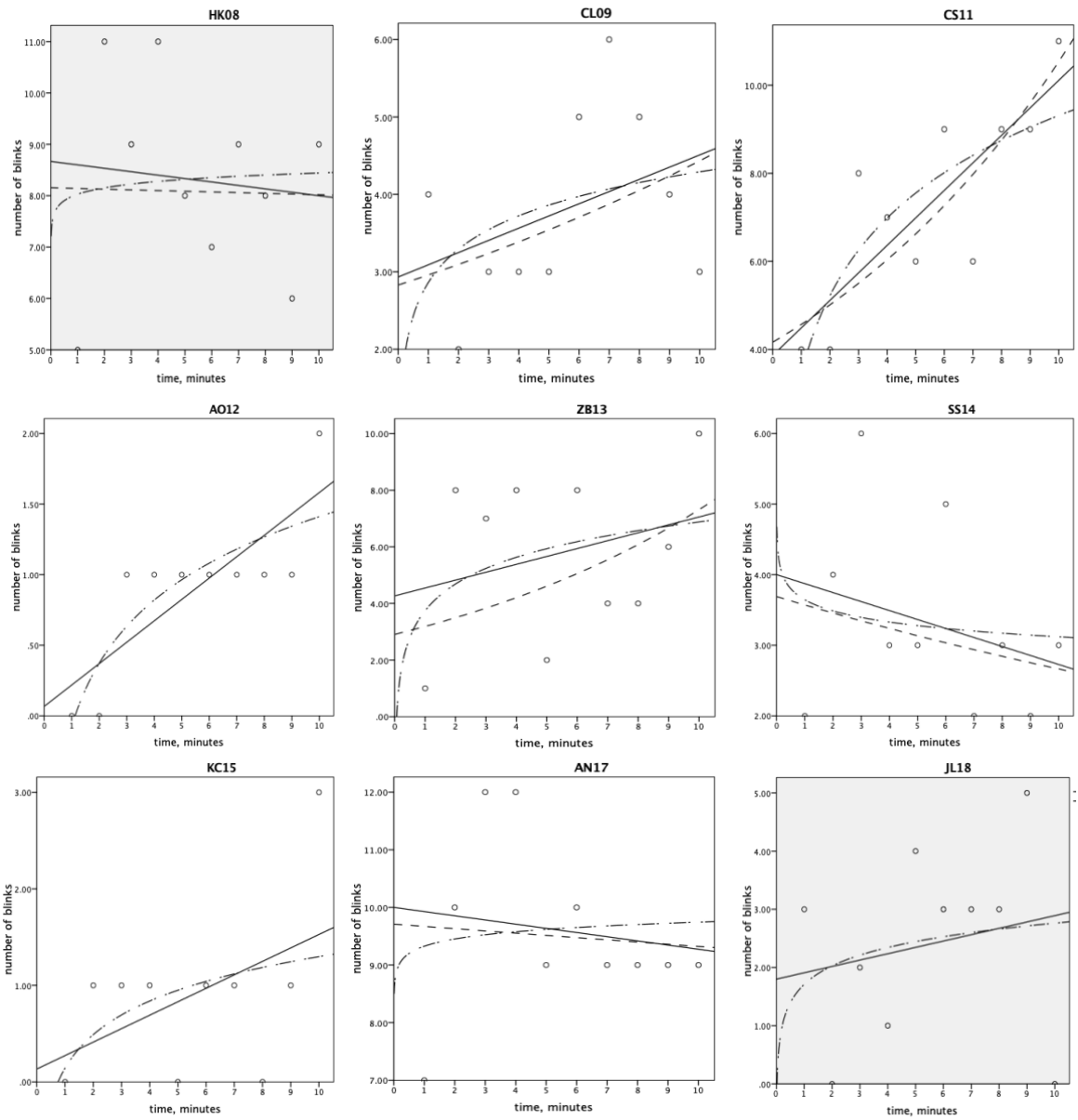
	R ² for reading slope using	
	read speed (wpm)	Log reading speed (Log wpm)
Mean	0.14	0.14
SD	0.14	0.14
Range	0.00 to 0.52	0.00 to 0.51

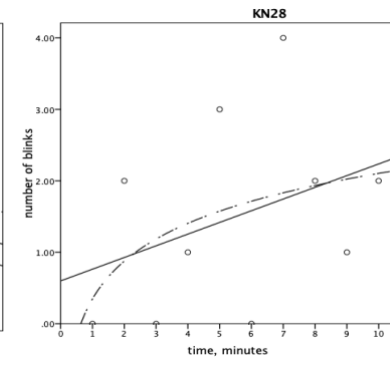
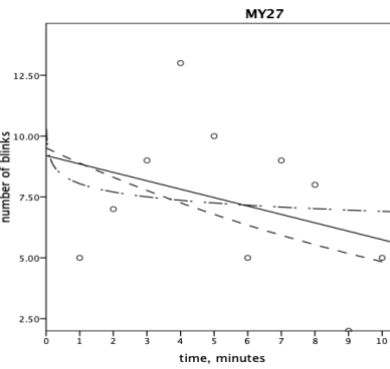
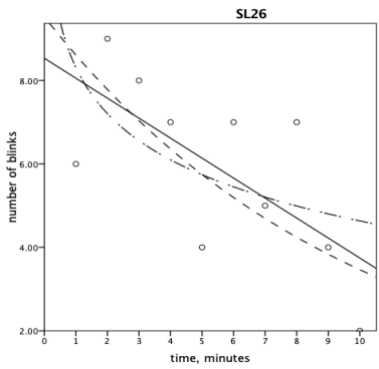
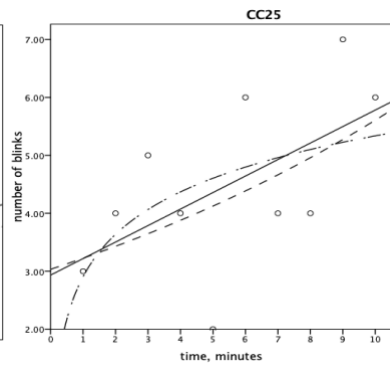
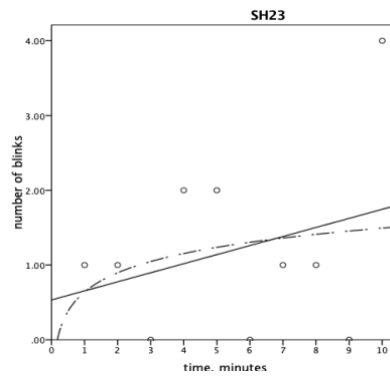
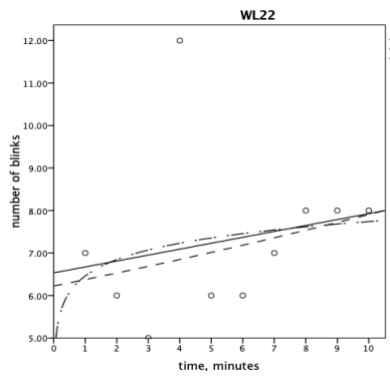
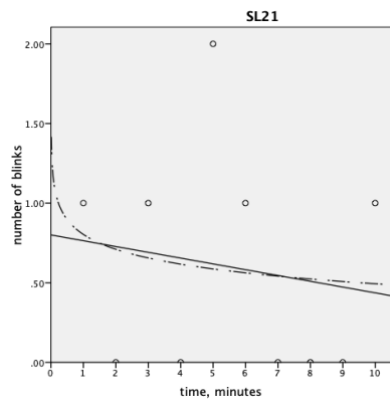
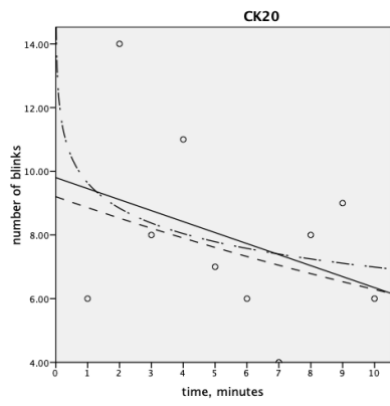
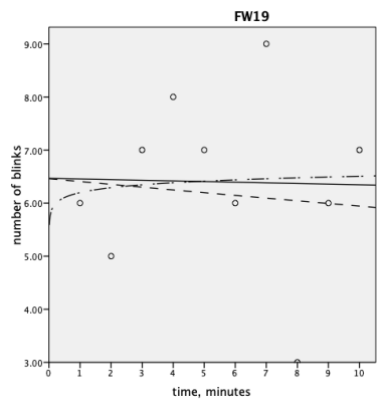
9.8 Graph plots for p-EVES participants with negative reading slopes in the first visit (n = 13)

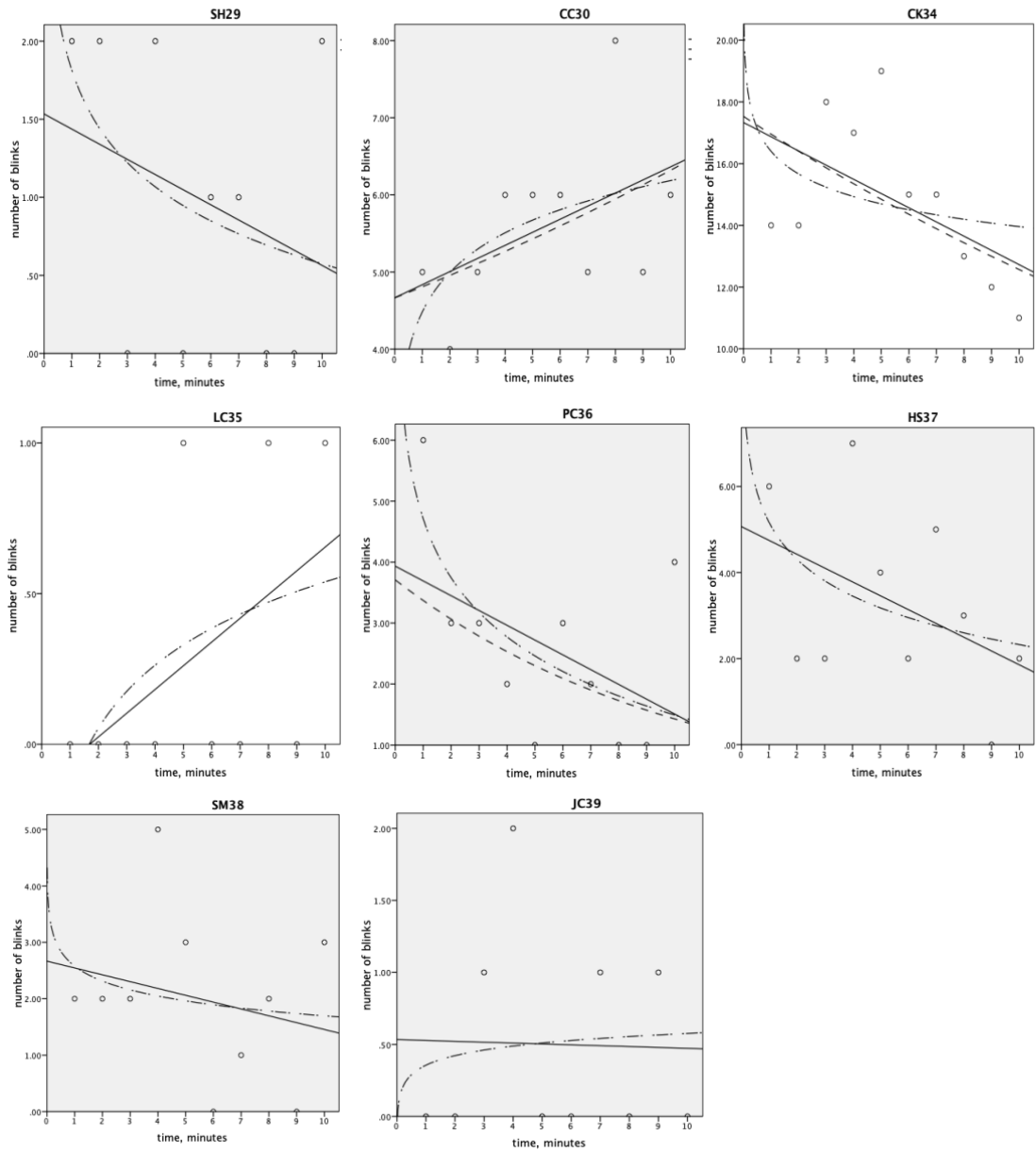




9.9 Graph plots on blinking rates against time per minute in the elderly participants in MAG (n = 26)







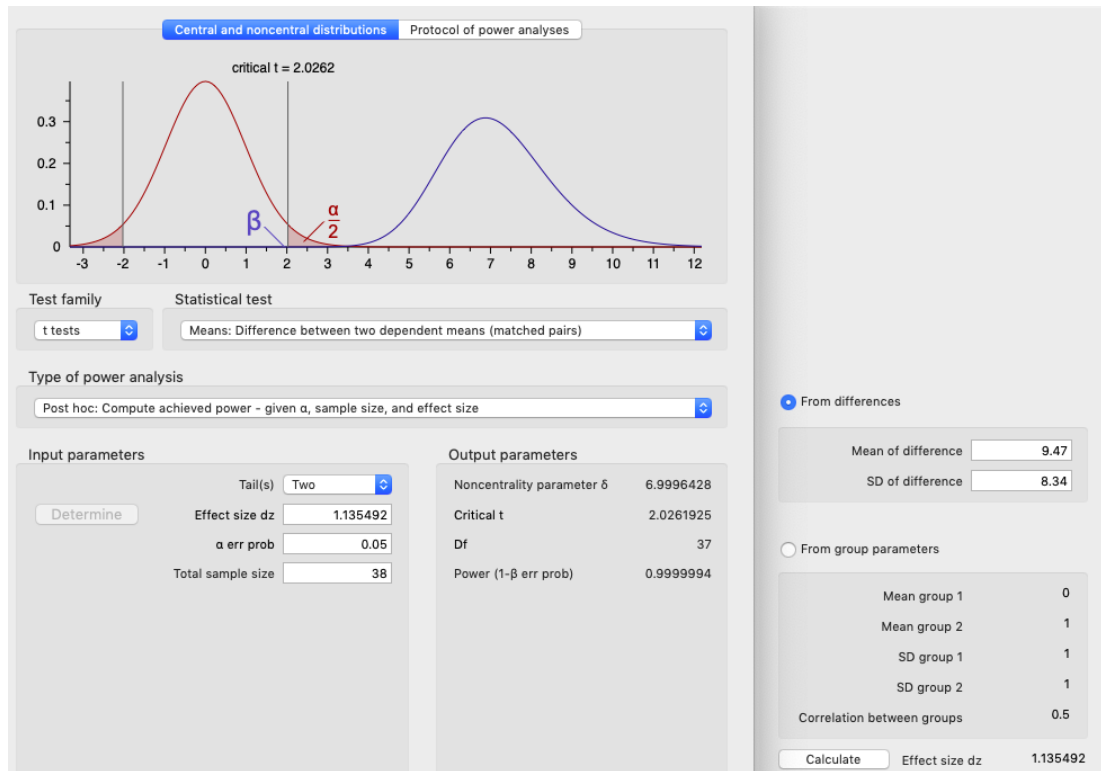
The graph plots (number of blinks against time) for elderly participants in MAG (n = 26). The grey background indicates participants aged > 60 years.

	R ² values for SBR Slopes Elderly MAG (n= 26)		
	Linear	Logarithmic	Exponential
08HK	0.01	0.01	0.00
09CL	0.15	0.14	0.17
11CS	0.67	0.65	0.66
12AO	0.65	0.69	NA
13ZB	0.08	0.12	0.15
14SS	0.08	0.02	0.07
15KC	0.23	0.18	NA
17AN	0.02	0.01	0.01
18JL	0.04	0.04	NA
19FW	0.00	0.00	0.01
20CK	0.13	0.09	0.11
21SL	0.03	0.02	NA
22WL	0.05	0.04	0.09
23SH	0.09	0.05	NA
25CC	0.33	0.26	0.26
26SL	0.46	0.30	0.47
27MY	0.11	0.01	0.15
28KN	0.13	0.17	NA
29SH	0.10	0.18	NA
30CC	0.23	0.25	0.25
34CK	0.29	0.09	0.33
35LC	0.24	0.21	NA
36PC	0.22	0.42	0.21
37HS	0.20	0.18	NA
38SM	0.06	0.04	NA
39JC	0.00	0.01	NA

R² values for SBR Slopes Elderly MAG (n= 26) for linear =, logarithmic and exponential functions. NA indicated that value was not available due to unable to fit the data points. Grey filled boxes indicated participants > 60 years.

9.10 Post hoc power analysis using G*Power (Version 3.1.9.3)

9.10.1 Post hoc power analysis for Chapter 2.1 – Exploring the limitation of the inclusion of the pre-verbalisation time in the computation of reading speed



Total sample size: $n = 38$ participants

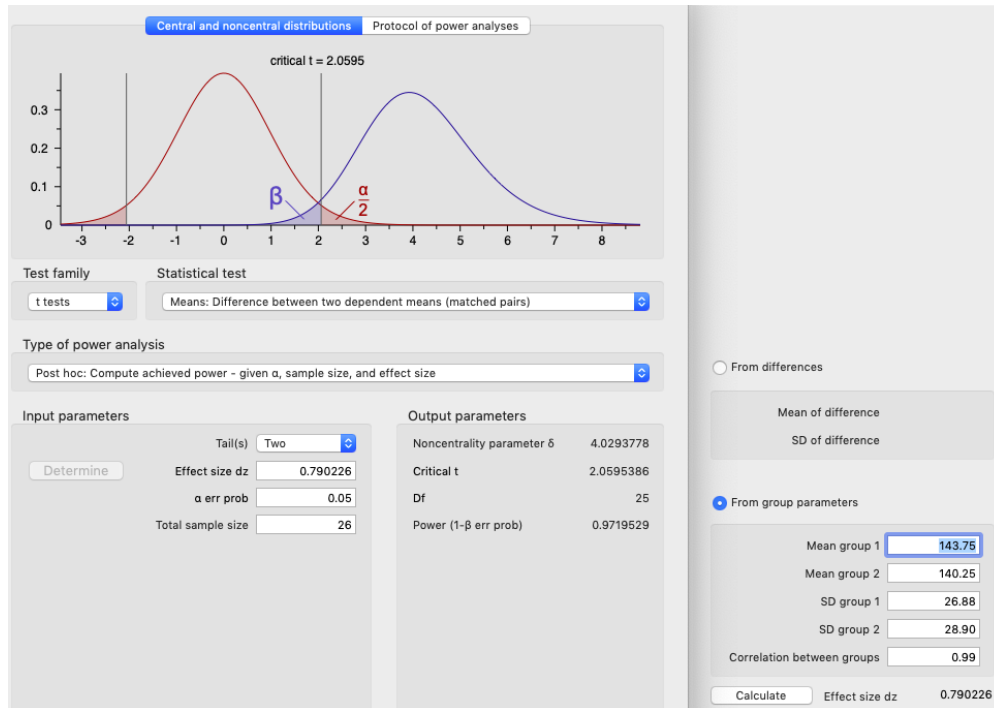
The effect size for all the 3 reading tests were calculated (i.e. IReST: 1.29, MNREAD acuity chart: 1.19 and Bailey-Lovie word-reading chart: 1.13). Since Bailey-Lovie word-reading chart had the smallest effect size, the data will be used to calculate post hoc power analysis.

Mean of difference: mean audio reading speed – mean stopwatch reading speed = 9.47 wpm

SD of difference: 8.34 wpm

Therefore, the post hoc power analysis = **0.99**

9.10.2 Post hoc power analysis for Chapter 2.2 – To explore the use of page-turn as a time recording method in a silent reading test



Total sample size: $n = 26$ participants

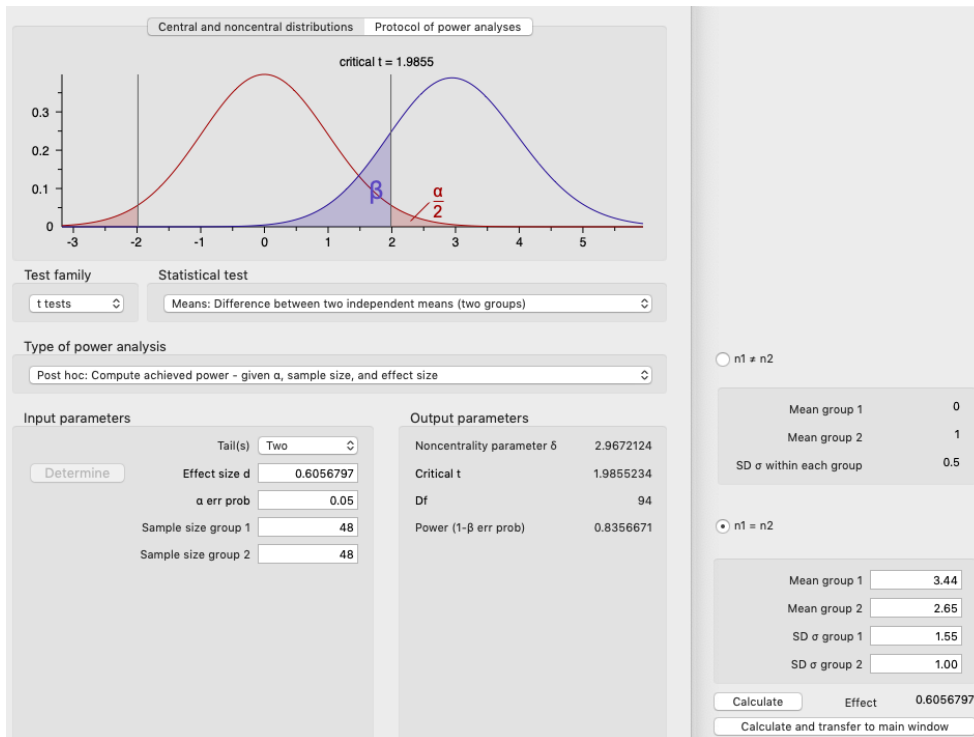
Mean group 1: 143.75 (audio average reading speed for NOR, BOF and MAG)

Mean group 2: 140.25 (page-turn average reading speed for NOR, BOF and MAG)

Number of groups: 2 groups (comparing page-turn vs audio recording)

Therefore, the post hoc power analysis = **0.97**

9.10.3 Post hoc power analysis for Chapter 2.3 – The use of proofreading to ensure reading compliance in silent reading tests



Total sample size: $n = 48$ participants in each group (readers vs non-readers)

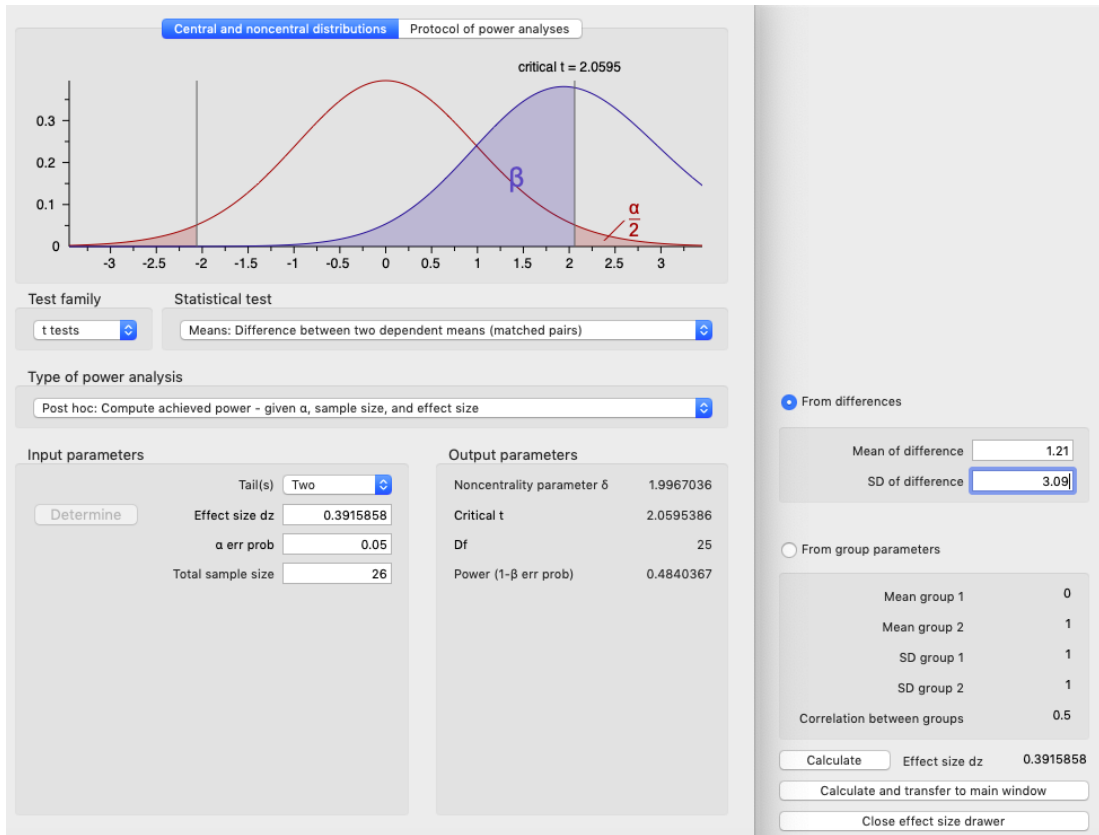
Mean group 1: 3.44 (average readers multiple-choice question scores)

Mean group 2: 2.65 (average non-readers multiple-choice question scores)

Effect Size = 0.61

Therefore, the post hoc power analysis = **0.84**

9.10.4 Post hoc power analysis for Chapter 3 – Evaluation of reading endurance in young adults under conditions of simulated visual impairment



Total sample size: $n = 26$ participants

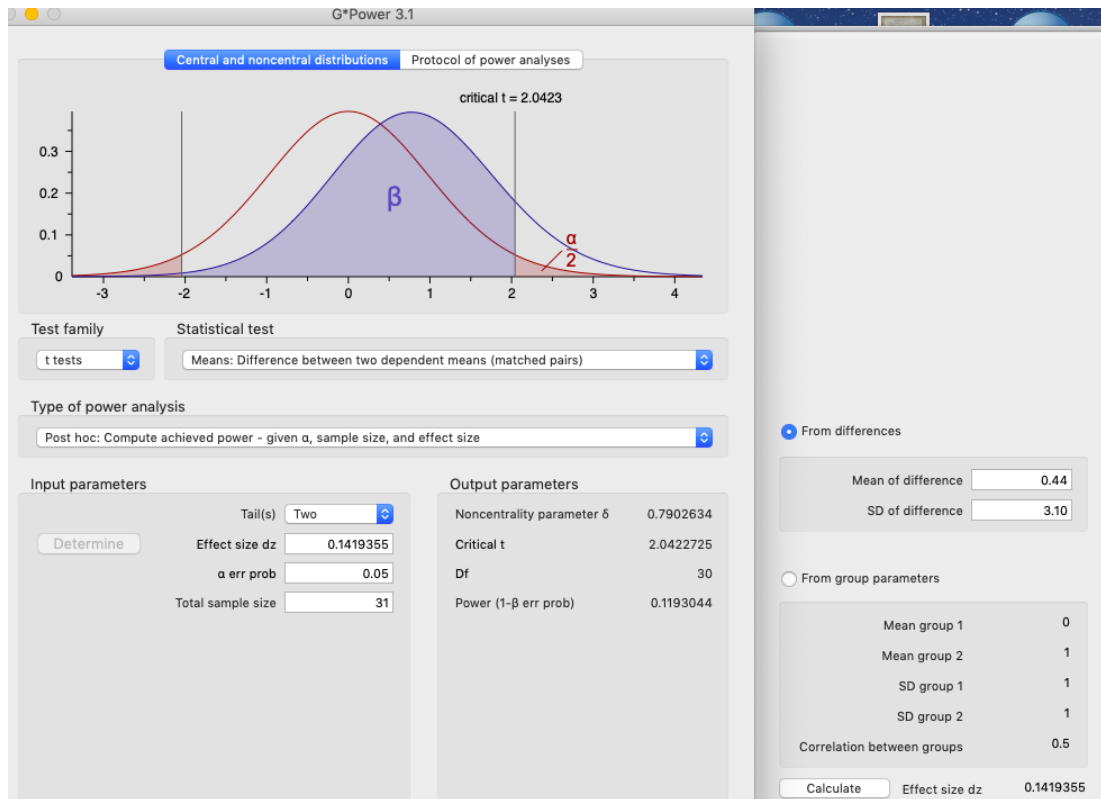
Based on the hypothesis that the reading slopes would be different from "0",

Mean of difference: $0 - (\text{mean reading slope BOF} + \text{MAG}) = 1.21$

SD of difference: $(\text{SD of reading slopes BOF} + \text{MAG}) = 3.09$

Therefore, the post hoc power analysis = **0.48**

9.10.5 Post hoc power analysis for Chapter 4 – Investigation of duration related changes in reading speed performance in middle-aged and elderly adults



Total sample size: $n = 31$ participants

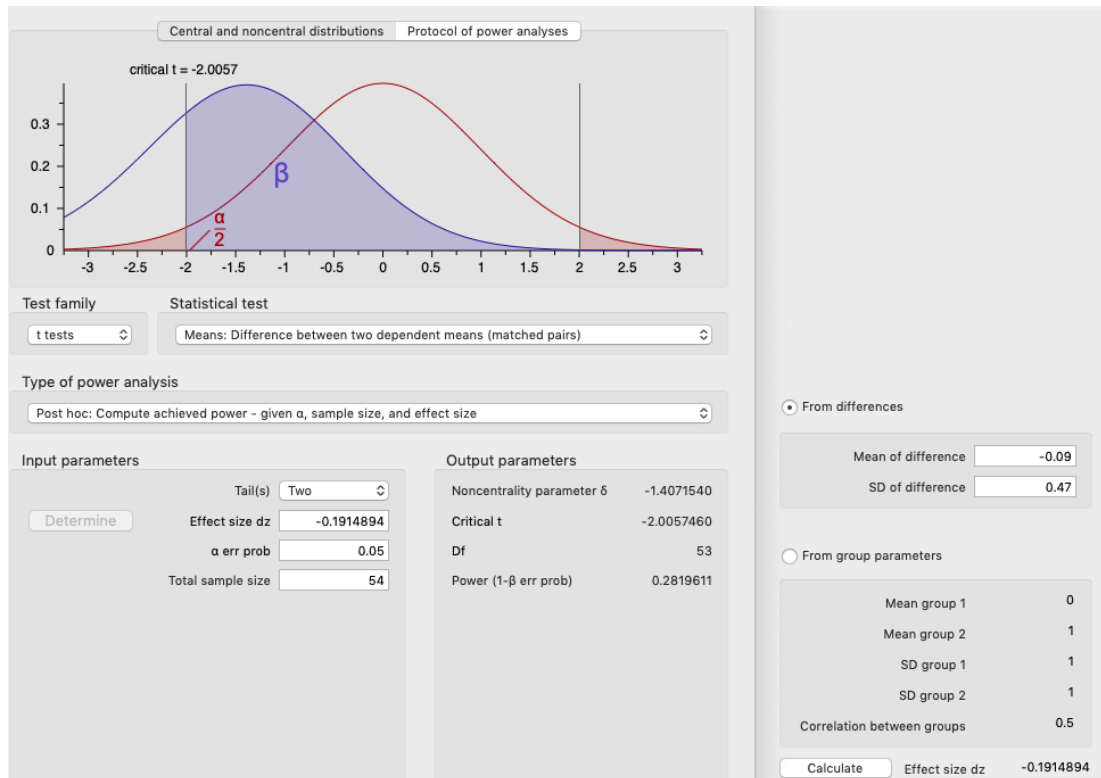
Based on the hypothesis that the reading slopes would be different from "0",

Mean of difference: $0 - (\text{mean reading slope BOF} + \text{MAG}) = 0.44$

SD of difference: $(\text{SD of reading slopes BOF} + \text{MAG}) = 3.10$

Therefore, the post hoc power analysis = **0.12**

9.10.6 Post hoc power analysis for Chapter 6 – The use of spontaneous eye blink rates to capture fatigue effect in reading



Total sample size: $n = 54$ participants (young, $n = 28$ + elderly, $n = 26$)

Based on the hypothesis that the SBR slopes in BOF and MAG would be different from "0",

Mean of difference: $0 - (\text{mean SBR slope BOF} + \text{MAG}) = -0.09$

SD of difference: $(\text{SD of SBR slopes BOF} + \text{MAG}) = 0.47$

Therefore, the post hoc power analysis = **0.28**

9.11 Summary of ethics approval

Ethics Approval	Sample size	Chapters	Remarks
The p-EVES study was registered with clinical trials.gov (Identifier: NCT01701700), received National Research Ethics Service approval (NRES)	N = 100	Chapter 5: To investigate the use of IReST to assess the effect of duration on reading performance in a group of visually impaired individuals	N = 61
University of Manchester Research Ethics Committee (Reference no: 08046)	N = 30	Chapter 2.2: To explore the use of page-turn as a time recording method in a silent reading test Chapter 3: Evaluation of reading endurance in young adults under conditions of simulated visual impairment	Same subjects used in both sections N = 26
Singapore Polytechnic Ethics Committee (Reference no: 201402-04)	N = 38 N= 60	Chapter 4: Investigation of duration related changes in reading speed performance in middle-aged and elderly adults Chapter 6: The use of spontaneous eye blink rate (SBR) to capture fatigue effect in reading	elderly = 31 (both reading speed and SBR captured simultaneously) elderly subjects (n = 26) young subjects (n = 28)
Singapore Polytechnic Ethics Research Committee (Protocol no: 201501-12)	N = 38 N = 48	Chapter 2.1: Exploring the limitation of the inclusion of the pre-verbalisation time in the computation of reading speed Chapter 2.3: The use of proofreading to ensure reading compliance in silent reading tests	N = 38 N = 48 (readers) N = 48 (non-readers)