

# The Effects of a Peer-Mediated Synthetic Phonics Intervention with Children from a Rural Indian Town

Christopher John Counihan

**A thesis submitted in candidature for the degree of Doctor of Philosophy**

**School of Education, Communication and Language Sciences**

**Newcastle University**

**October 2017**



## Abstract

It's a truism that children living in rural areas of India suffer from a lack of quality reading instruction and guidance. Singular rote pedagogies hinder children's progress and potential for higher study. In light of this, research continues to find benefits in synthetic phonic and peer learning methods, offering a potential avenue for accelerating early reading achievement.

This study investigated the effects of a peer-mediated synthetic phonics programme operating in a rural Northern Indian town. Following a sociocultural peer cooperative approach to reading instruction, children from a low-cost private school (N = 44) were recruited as Monitors and trained in a peer/synthetic method ahead of a four-month intervention. Given localised problems associated with education access and quality, children from a range of mainstream education providers (N = 701) were recruited as Learners and randomised into control and intervention groups.

An embedded mix methods research design was incorporated into a three-phase research plan. Using an experimental random controlled trial (RCT) design, quantitative collections involved pre/post testing on all children's phonological awareness, spelling, reading skills and reading attitudes. Qualitative collections during phase two developed pre-test findings by exploring Monitors' reading attitudes ahead of the intervention.

The main findings suggest children receiving peer-mediated synthetic phonic instructions are able to improve their English phonemic awareness, spelling, and reading skills. It demonstrates capacity for accelerating reading achievement over a short intervention period, potential for further study in this area and pedagogical reform. In addition to this, factors associated with peer reading opportunities, location and pedagogical structure improved reading attitudes. Attitudes to reading is an underdeveloped area in Indian education policy, it is not referred to in any policy documents or curricula. The significance of peer reading promotion in academic and recreational settings is crucial to accelerating reading achievement. Taken as a whole, this research has major implications for international development programmes looking to target progression in early reading achievement and the promotion of positive attitudes in reading activities.

## Acknowledgements

I am extremely grateful to a number of people who have helped with the development of this study. Firstly, my supervisor, Professor Pauline Dixon has been a rock of support throughout the entire process. Her work and dedication to my study often went above and beyond the call of duty. I'm forever grateful for the training, unscheduled meetings and time invested in the development of this study.

I would also like to express thanks to Preet Grewal and the management team at Alpine Public School. If not for them, general access to the school's facilities and network would have not taken place. I'm extremely grateful for the unscheduled meetings and numerous phone/Skype calls – sometimes very late into the night.

To all the children who invested their time and energy over the study period, I would like to express how unbelievably proud you made me over the study period. Your conduct, desire and passion for teaching and learning were commendable and made a huge difference to the challenging situations you all worked in.

Special thanks goes to Chris Jolly who supported the study with generous donations of books and materials. I'm extremely grateful for your support and help.

Finally, I would like to thank my parents and wider family for all their love and support during this process. To my partner, Lisa, I love you very much and could not have completed this work without your support.

## Table of Contents

<b>ABSTRACT</b> .....	<b>I</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>II</b>
<b>LIST OF FIGURES</b> .....	<b>VIII</b>
<b>LIST OF TABLES</b> .....	<b>IX</b>
<b>GLOSSARY OF TERMS</b> .....	<b>XI</b>
<b>CHAPTER ONE - INTRODUCTION</b> .....	<b>1</b>
1.1 INTRODUCTION.....	1
1.2 ILLITERACY IN INDIA – A NATION’S ANATHEMA .....	1
1.3 ENGLISH POLICY & EDUCATION.....	2
1.4 CHALLENGES IN PRIMARY EDUCATION AND READING PEDAGOGY.....	4
1.5 THE CHOICE OF LOCATION AND STUDY CONTEXT.....	6
<b>CHAPTER TWO - LITERATURE REVIEW</b> .....	<b>14</b>
2.1 INTRODUCTION.....	14
2.2 SOCIOCULTURAL THEORY & EMERGENT LITERACY.....	14
2.2.1 <i>Relationships of L1 - L2 Reading and Emergent Readers</i> .....	16
2.3 THEORETICAL MODELS & READING DEVELOPMENT.....	17
2.3.1 <i>Stages of Reading</i> .....	17
2.3.2 <i>Reading Pedagogy</i> .....	19
2.3.3 <i>Whole Language Approach</i> .....	19
2.3.4 <i>Analytic Phonics Approach</i> .....	21
2.3.5 <i>Synthetic Phonic Approaches</i> .....	22
2.4 RECIPROCITY IN L1 AND L2 READING ATTITUDES.....	25
2.4.1 <i>Day and Bamford (1998) Model</i> .....	27
2.4.2 <i>Ruddell and Speaker (1985) Model</i> .....	28
2.4.3 <i>Mathewson (1985, 1994) Model</i> .....	28
2.4.4 <i>Mckenna (1994) Model</i> .....	29
2.5 PEER LEARNING: THEORETICAL DIMENSIONS .....	30
2.5.1 <i>Introduction</i> .....	30
2.5.2 <i>Social Interdependence Theory</i> .....	31
2.6 PEER LEARNING METHODS.....	36
2.6.1 <i>Introduction</i> .....	36
2.6.2 <i>Peer Tutoring</i> .....	37
2.6.3 <i>Peer Collaborative Learning</i> .....	37
2.6.4 <i>Peer Cooperative Learning</i> .....	39
2.6.5 <i>Peer Cooperative Learning and L2 Reading</i> .....	41
2.7 PEER COOPERATIVE TEACHING & LEARNING STRATEGIES.....	42
2.7.1 <i>Introduction</i> .....	42
2.7.2 <i>Constructive Controversy</i> .....	43
2.7.3 <i>Teams Assisted Individualisation (TAI)</i> .....	43
2.8 PEER COOPERATIVE READING META-ANALYTIC REVIEWS .....	48
2.8.1 <i>Introduction</i> .....	48
2.8.2 <i>Slavin, et al., (2009b)</i> .....	50
2.8.3 <i>Puzio &amp; Colby (2013)</i> .....	50

2.9	THEORETICAL FRAMEWORK.....	52
<b>CHAPTER THREE - METHODOLOGY .....</b>	<b>53</b>	
3.1	<i>Introduction</i> .....	53
3.1.1	<i>Philosophical Foundation</i> .....	54
3.1.2	<i>Epistemological Framework</i> .....	55
3.1.3	<i>Mix Methods</i> .....	56
3.1.4	<i>Mix Method Research Design</i> .....	58
3.1.5	<i>Mix Method Intervention</i> .....	58
3.1.6	<i>Why a Random Control Trial?</i> .....	59
3.2	QUANTITATIVE METHODS .....	60
3.2.1	<i>Introduction</i> .....	60
3.2.2	<i>Family Background Questionnaire</i> .....	60
3.2.3	<i>Reading Aspiration Survey</i> .....	61
3.2.4	<i>Phonological and Character Test Battery Scoring</i> .....	62
3.2.5	<i>Reading Test and Dictation Test Battery Scoring</i> .....	63
3.2.6	<i>Dictation Test Delivery and Scoring</i> .....	67
3.2.7	<i>Native Reading Test Battery Scoring</i> .....	68
3.3	QUALITATIVE METHODS .....	70
3.3.1	<i>Introduction</i> .....	70
3.3.2	<i>Focus Groups</i> .....	70
3.4	DATA ANALYSIS .....	71
3.4.1	<i>Quantitative Data</i> .....	71
3.4.2	<i>Qualitative Data</i> .....	74
3.5	ETHICAL CONSIDERATIONS.....	75
3.5.1	<i>School Clearance</i> .....	75
3.5.2	<i>Request for Study Location</i> .....	76
3.5.3	<i>Recruitment of Children</i> .....	77
3.5.4	<i>Monitor Recruitment</i> .....	77
3.5.5	<i>Learner Recruitment</i> .....	78
3.6	PHASE ONE –GROUND WORK .....	79
3.6.1	<i>Introduction</i> .....	79
3.6.2	<i>Setup</i> .....	79
3.6.3	<i>Project Searchlight</i> .....	81
3.6.4	<i>Towards a Peer-Mediated Trial</i> .....	82
3.6.5	<i>Establishing a Research Team</i> .....	83
3.6.6	<i>Research Team Members (RTMs)</i> .....	83
3.6.7	<i>Searching for a Location</i> .....	85
3.6.8	<i>Gurudwara</i> .....	85
3.6.9	<i>Searching for Monitors</i> .....	86
3.6.9.1	<i>Parent Communication</i> .....	87
3.6.9.2	<i>Family Background Questionnaire</i> .....	87
3.6.9.3	<i>Searching for Learners</i> .....	88
3.6.9.4	<i>Phase One Summary</i> .....	89
3.7	PHASE TWO –FOLLOW-UP AND PILOT WORK.....	90
3.7.1	<i>Introduction</i> .....	90
3.7.2	<i>Location Follow-up</i> .....	91

3.7.3	<i>Recruitment by Media</i> .....	91
3.7.4	<i>Pilot Study Development</i> .....	92
3.7.5	<i>Testing the Tests</i> .....	93
3.7.6	<i>Reading Aspiration Survey Administration</i> .....	98
3.7.7	<i>Monitor Baseline Tests and Focus Group</i> .....	98
3.7.8	<i>Developing Peer Monitors Through Initial Mentor Training (IMT)</i> .....	99
3.7.9	<i>Phase Two Summary</i> .....	104
3.8	PHASE THREE – REGISTRATION, BASELINE TESTS, AND INTERVENTION.....	105
3.8.1	<i>Programme Registration</i> .....	105
3.8.2	<i>Baseline Testing</i> .....	107
3.8.3	<i>Test Scheduling and Management</i> .....	108
3.8.4	<i>Post - Baseline</i> .....	111
3.8.5	<i>Randomisation</i> .....	112
3.8.6	<i>Group Announcements and Feedback</i> .....	115
3.8.7	<i>Orientation – Year 0</i> .....	115
3.8.8	<i>Losses and Exclusions</i> .....	118
3.8.9	<i>Phase Three Summary</i> .....	119
<b>CHAPTER FOUR RESULTS</b> .....		<b>121</b>
4.1	INTRODUCTION.....	121
4.2	MONITORS SOCIO-ECONOMIC BACKGROUND VIGNETTE (N =44) .....	122
4.2.1	<i>Introduction</i> .....	122
4.2.2	<i>Socio-Economic and Family Findings</i> .....	122
4.2.3	<i>Monitors’ Characteristics and Education Inputs</i> .....	127
4.3	LEARNER SOCIO-ECONOMIC BACKGROUND VIGNETTE (N = 701) .....	130
4.3.1	<i>Introduction</i> .....	130
4.3.2	<i>Social Economic and Family Findings</i> .....	130
4.3.3	<i>Leaners’ Educational Inputs</i> .....	131
4.3.4	<i>Summary</i> .....	133
4.4	MONITORS ATTITUDES TO READING ENGLISH IN RECREATIONAL AND ACADEMIC SETTINGS..	134
4.4.1	<i>Introduction</i> .....	134
4.4.2	<i>Monitor Pre/Post Recreational &amp; Academic Scores</i> .....	134
4.4.3	<i>Monitors’ Gender Effects on Academic and Recreational Attitudes</i> .....	139
4.5	PCA ON MONITORS’ L2 BASELINE ENGLISH READING ATTITUDES .....	143
4.5.1	<i>Introduction</i> .....	143
4.5.2	<i>English Reading Attitudes Within a Recreational Setting</i> .....	144
4.5.3	<i>Recreational Focus Group Findings</i> .....	148
4.5.4	<i>English Reading Attitudes Within an Academic Setting</i> .....	152
4.5.5	<i>Academic Focus Group</i> .....	156
4.6	DESCRIPTIVE FINDINGS OF MONITORS’ PRE/ POST TEST SCORES.....	162
4.6.1	<i>Introduction</i> .....	162
4.6.2	<i>Monitor Reading &amp; Dictation Test Battery</i> .....	163
4.6.3	<i>Monitor Character and Phonological Processing Test Battery</i> .....	165
4.6.4	<i>Monitor Native Reading Test Battery</i> .....	165
4.7	MONITORS’ GENDER AND TEST BATTERY INFERENTIAL FINDINGS .....	166
4.7.1	<i>Introduction</i> .....	166
4.7.2	<i>Gender Effects: Reading &amp; Dictation Test Battery</i> .....	167

4.7.3	<i>Gender Effects: Phonological and Character Awareness Test Battery</i>	169
4.7.4	<i>Gender Effects: of Native Reading Battery</i>	173
4.7.5	<i>Monitor Reading and Dictation Test Battery</i>	175
4.7.6	<i>Phonological and Letter Character Test Battery</i>	177
4.7.7	<i>Native Reading Test Battery</i>	180
4.8	<b>DESCRIPTIVE &amp; INFERENTIAL FINDINGS OF LEARNERS' PRE/POST TEST SCORES</b>	181
4.8.1	<i>Introduction</i>	181
4.8.2	<i>Gender Effects: Reading and Dictation Test Battery</i>	181
4.8.3	<i>Gender Effects: Phonological &amp; Character Awareness Battery</i>	183
4.8.4	<i>Gender Effects: Native Language Test Battery</i>	187
4.8.5	<i>Learners' Reading and Dictation Baseline Comparisons</i>	188
4.8.6	<i>Learners' Reading &amp; Dictation Effects</i>	190
4.8.7	<i>Learners' Phonological and Character Awareness Baseline Comparisons</i>	195
4.8.8	<i>Learners' Phonological and Character Awareness Effects</i>	200
4.8.9	<i>Miskin Tests Over Time</i>	205
4.8.9.1	<i>Learners' Native Reading Test Baseline Comparison</i>	207
4.8.9.2	<i>Learner' Native Reading Effects</i>	208
4.9	<b>LEARNERS' ATTITUDES TO READING ENGLISH IN ACADEMIC AND RECREATIONAL SETTINGS</b>	213
4.9.1	<i>Introduction</i>	213
4.9.2	<i>Control Group Pre/Post Academic and Recreational Findings</i>	213
4.9.3	<i>Intervention Group Pre/Post Academic and Recreational Findings</i>	215
4.94	<i>Chapter Conclusion</i>	217
<b>CHAPTER FIVE - DISCUSSION</b>		<b>218</b>
5.1	<b>INTRODUCTION</b>	218
5.2	<b>COMMENTARY ON FINDINGS</b>	218
5.2.1	<i>Pro-Social Recreational Reading Attitudes</i>	219
5.2.2	<i>Academic Reading Attitudes</i>	222
5.2.3	<i>Pre – Post Trial Attitudes</i>	224
5.2.4	<i>Learners' Pre/Post Trial Attitudes</i>	226
5.2.5	<i>Trial Impact</i>	228
5.2.6	<i>Assessing Monitors' Phonological, Spelling and Reading Abilities</i>	228
5.2.7	<i>Monitors as Peer Instructors</i>	229
5.2.8	<i>Learning to Read with Peer Instruction</i>	231
5.2.9	<i>The Significance of Peer Reading Instruction and Attitudes</i>	235
<b>CHAPTER SIX - CONCLUSION</b>		<b>236</b>
6.1	<i>Introduction</i>	236
6.1.2	<i>Contribution to Knowledge</i>	236
6.1.3	<i>Future Research and Implications</i>	237
6.1.4	<i>Implications for Pedagogy</i>	240
6.1.5	<i>Implications for Policy</i>	241
6.1.6	<i>Concluding Thoughts</i>	242
<b>REFERENCES</b>		<b>243</b>
<b>APPENDICES</b>		<b>275</b>
FAMILY BACKGROUND QUESTIONNAIRE		275
ENGLISH READING ASPIRATION SURVEY		277



NFER A TEST (LETTER MATCHING) .....	278
NFER TEST B (PHONEME IDENTIFICATION) AND NFER TEST C (2/3 BLENDING).....	279
MISKIN NONSENSE WORDS .....	280
DICTATION TESTS.....	281
HINDI WORD TEST .....	282
RECREATIONAL AND ACADEMIC READING ATTITUDE FOCUS GROUP ANALYSIS .....	284
LETTER OF ENQUIRY .....	285
GURUDWARA LETTER .....	288
RECRUITMENT LETTER .....	289
MEDIA RECRUITMENT POSTER.....	291
LETTER OF CONSENT AND DEBRIEFING GUIDE.....	292
FOCUS GROUP INTERVIEW GUIDE .....	298
EXTRACTS FROM SUPPORTED STUDY GUIDE.....	299
PHOTOS FROM THE INTERVENTION .....	303

## List of Figures

FIGURE 1 MAP OF CENTRAL PUNJAB, INDIA.....	7
FIGURE 2 MAP OF BHAWANIGARH (STUDY LOCATION) .....	9
FIGURE 3 INTERVENTION TEACHING AND LEARNING STRUCTURE.....	59
FIGURE 4 PROJECT SEARCHLIGHT CONCEPT MAP.....	81
FIGURE 5 PEER TEACHING IMT: BIG BOOK AND INITIAL SOUNDS .....	99
FIGURE 6 CONSORT FLOWCHART OF RCT PROFILE .....	117
FIGURE 7 CONCEPTUAL SCHEMATIC OF STUDY TRAJECTORY.....	119
FIGURE 8 FAMILY CASTE TYPE .....	123
FIGURE 9 HOME OWNERSHIP .....	124
FIGURE 10 FATHER'S JOB TYPE.....	124
FIGURE 11 MOTHER'S OCCUPATION.....	125
FIGURE 12 FATHER'S EDUCATION LEVEL ACHIEVED .....	126
FIGURE 13 MOTHER'S EDUCATION LEVEL ACHIEVED .....	127
FIGURE 14 DISTRIBUTION OF MONITORS' AGE (IN YEARS) .....	128
FIGURE 15 PERCENTAGES OF ENGLISH TUITION FEES .....	129
FIGURE 16 NUMBER OF ENGLISH TUITION (IN HOURS) .....	130
FIGURE 17 HIGHEST LEVEL OF EDUCATION (MALE CARER) .....	132
FIGURE 18 HIGHEST LEVEL OF EDUCATION (FEMALE CARER) .....	133
FIGURE 19 BOXPLOT OF RECREATION AND ACADEMIC ATTITUDE SCORES .....	138
FIGURE 20 SCREE PLOT OF RECREATIONAL READING ATTITUDE EIGENVALUES >1.....	144
FIGURE 21 SCREE PLOT OF ACADEMIC READING ATTITUDE EIGENVALUES >1 .....	153
FIGURE 22 BOXPLOTS OF MONITORS' PRE/POST READING TEST SCORES.....	164
FIGURE 23 MISKIN NONSENSE WORDS BASELINE TEST (BY GENDER).....	187
FIGURE 24 BOXPLOT OF COMPOSITE DIFFERENCES OF ACADEMIC AND RECREATIONAL READING ATTITUDES (CONTROL GROUP).....	214

## List of Tables

TABLE 1 INTRACLASS COEFFICIENTS FOR TESTS.....	95
TABLE 2 SYNOPSIS OF INITIAL MONITOR TRAINING AND SOUND AWARENESS.....	102
TABLE 3 TEST SCHEDULE.....	108
TABLE 4 STRATIFIED CONDITIONS FOR RANDOM ALLOCATION .....	115
TABLE 5 MONITOR PRE/POST RECREATIONAL AND ACADEMIC SCORES .....	134
TABLE 6 ANCOVA OF RECREATIONAL POST ATTITUDE RESULTS CONTROLLING FOR BASELINE MEASURES.....	139
TABLE 7 UNADJUSTED AND ADJUSTED RECREATIONAL ATTITUDE MEANS (BY GENDER) .....	140
TABLE 8 ANCOVA OF ACADEMIC POST ATTITUDE RESULTS CONTROLLING FOR BASELINE MEASURES .....	141
TABLE 9 UNADJUSTED AND ADJUSTED ACADEMIC ATTITUDE MEANS (BY GENDER) .....	141
TABLE 10 ANCOVA OF TOTAL POST ATTITUDE RESULTS CONTROLLING FOR BASELINE MEASURES .....	142
TABLE 11 UNADJUSTED AND ADJUSTED TOTAL ATTITUDE MEANS (BY GENDER) .....	142
TABLE 12 FACTOR LOADINGS & COMMONALITIES FOR ENGLISH RECREATIONAL READING (N = 44).....	145
TABLE 13 DESCRIPTIVE STATISTICS FOR THREE RECREATIONAL SCALE FACTORS (N = 44).....	147
TABLE 14 DESCRIPTIVE STATISTICS OF MONITORS' FOCUS GROUP INFORMATION .....	148
TABLE 15 FACTOR LOADINGS & COMMONALITIES FOR ENGLISH ACADEMIC READING (N = 44).....	154
TABLE 16 DESCRIPTIVE STATISTICS FOR THREE ACADEMIC SCALE FACTORS (N = 44).....	155
TABLE 17 DESCRIPTIVE STATISTICS OF MONITORS' READING ASPIRATIONS .....	157
TABLE 18 DESCRIPTIVE STATISTICS OF MONITORS' BASELINE READING TEST SCORES .....	163
TABLE 19 DESCRIPTIVE STATISTICS OF MONITORS' PRE/POST LETTER MATCHING AND BLENDING SCORES.....	165
TABLE 20 DESCRIPTIVE STATISTICS OF MONITORS' PRE/POST NATIVE READING TEST SCORES.....	166
TABLE 21 DESCRIPTIVE STATISTICS OF MONITORS' PRE/POST READING TEST BATTERY SCORES (BY GENDER).....	168
TABLE 22 MONITORS' PRE DICTATION MEAN RANK TEST SCORES (BY GENDER).....	168
TABLE 23 MONITORS' PHONOLOGICAL AND CHARACTER AWARENESS TEST BATTERY (BY GENDER).....	170
TABLE 24 DESCRIPTIVE STATISTICS OF MONITORS' MISKIN NONSENSE WORDS (BY GENDER)...	172
TABLE 25 NATIVE LANGUAGE READING AND WORD TESTS (BY GENDER).....	174
TABLE 26 READING AND DICTATION TESTS FOR NORMALITY .....	175
TABLE 27 SIGN TEST FREQUENCIES OF MONITORS' DICTATION PRE/POST TESTS.....	177

TABLE 28 WILCOXON SIGNED TEST ON PHONOLOGICAL AND CHARACTER AWARENESS TEST .....	178
TABLE 29 WILCOXON SIGNED TESTS OF MONITORS' PHONOLOGICAL AND LETTER CHARACTER DIFFERENCES .....	179
TABLE 30 MONITORS' MISKIN NONSENSE TESTS (BY TYPE) .....	179
TABLE 31 LEARNERS' PRE READING TEST BATTERY (BY GENDER) .....	182
TABLE 32 LEARNERS' BASELINE READING TEST SCORES.....	189
TABLE 33 ANCOVA OF BURT READING TEST CONTROLLING FOR BASELINE MEASURES .....	191
TABLE 34 UNADJUSTED AND ADJUSTED BURT READING TEST MEANS (BY GROUP) .....	191
TABLE 35 ANCOVA OF SCHONELL SPELLING TEST CONTROLLING FOR BASELINE .....	193
TABLE 36 ADJUSTED AND UNADJUSTED MEANS FOR SCHONELL SPELLING TEST MEANS (BY GROUP).....	193
TABLE 37 ANCOVA OF DICTATION TEST SCORES CONTROLLING FOR BASELINE MEASURES .....	194
TABLE 38 ADJUSTED AND UNADJUSTED DICTATION TEST MEANS (BY GROUP) .....	195
TABLE 39 BASELINE PHONOLOGICAL AND LETTER MATCHING (BY GROUP).....	196
TABLE 40 BASELINE MISKIN NONSENSE WORD (BY GROUP) .....	198
TABLE 41 ANCOVA OF LETTER MATCHING TEST SCORES CONTROLLING FOR BASELINE MEASURES .....	201
TABLE 42 ADJUSTED AND UNADJUSTED MEANS FOR LETTER MATCHING (BY GROUP) .....	201
TABLE 43 ANCOVA OF PHONOLOGICAL TEST SCORES CONTROLLING FOR BASELINE MEASURES .....	202
TABLE 44 ADJUSTED AND UNADJUSTED MEANS FOR PHONOLOGICAL TEST SCORES (BY GROUP).....	203
TABLE 45 ANCOVA OF BLENDING TEST SCORES CONTROLLING FOR BASELINE MEASURES.....	204
TABLE 46 ADJUSTED AND UNADJUSTED MEANS FOR BLENDING TEST SCORES (BY GROUP).....	204
TABLE 47 DESCRIPTIVE STATISTICS AND T-TEST RESULTS FOR MISKIN NONSENSE WORD GROUPS (1-4).....	206
TABLE 48 DESCRIPTIVE BASELINE STATISTICS FOR NATIVE LANGUAGE TESTS (BY GROUP).....	207
TABLE 49 ANCOVA OF PUNJABI POST TEST SCORES CONTROLLING FOR BASELINE MEASURES .....	209
TABLE 50 ADJUSTED AND UNADJUSTED MEANS FOR PUNJABI DICTATION TEST SCORES .....	209
TABLE 51 ANCOVA OF HINDI POST TEST SCORES CONTROLLING FOR BASELINE MEASURES .....	211
TABLE 52 ADJUSTED AND UNADJUSTED MEANS FOR HINDI TEST SCORES (BY GROUP) .....	211
TABLE 53 BASELINE RECREATIONAL, ACADEMIC AND TOTAL ENGLISH READING ATTITUDE SCORES.....	213
TABLE 54 BASELINE RECREATIONAL, ACADEMIC AND TOTAL ENGLISH READING ATTITUDE SCORES.....	216

## Glossary of Terms

ANCOVA	Analysis of Covariance
CIRC	Cooperative Integrated Reading and Composition
CONSORT	Consolidated Standards of Reporting Trials
DFE	Department of Education
EAL	English as an Additional Language
EFA	Education for All
L1	First Language
L2	Second Language
PALs	Peer Assisted Learning
RCT	Random Controlled Trial
RTMs	Research Team Members
SLA	Second Language Acquisition
ZPD	Zone of Proximal Development



## **Chapter One - Introduction**

### **1.1 Introduction**

This introductory chapter traces the significant developments of English reading policy as it affects the modern Indian classroom. It begins with the current picture before evaluating the historical factors that have shaped vital policies and frameworks. From here, the chapter considers the difficulties in reading achievement before introducing the immediate study context and research questions.

### **1.2 Illiteracy in India – A Nation’s Anathema**

Throughout India, children face an uphill struggle in becoming literate in one or more languages. By 2030, India is estimated to have a population of 1.5 billion, surpassing China as the world’s most populous country. In the same year, it is expected that all young people (15 – 24-year-olds) will be literate as set by the quality education benchmarks from the Sustainable Development Goals (SDGs) agenda (UN, 2015). While there are encouraging signs of improvement in widespread literacy, a significant challenge remains in ensuring all children and young people have an equal chance of becoming literate. In this regard, the rural aspect of illiteracy is a growing concern for educational planners and policymakers. According to the Social Economic and Caste Census (GoI, 2011,p23) report, 36% (844 million people) of India’s rural population are illiterate. More shockingly, this is higher than the predictions found in the original 2011 census. Of families that are found in rural parts, 24% have no literate adult over the age of 25, and only 5.4% have completed high school. But why is this significant? India’s rural population accounts for 70% of the population, making this a real and unprecedented National issue (Census 2011).

It is without a doubt that the gains in achieving literacy are beneficial to the economic growth of society (Barro, 2001). Indeed, achieving widespread literacy has long been the nucleus of promoting intellectual property, and the stimulus for social mobility. The underlying theoretical assumptions relate to the diffusion of mass education to increase human capital and following labour outputs (Heyneman & Stern, 2015). Literacy as a social practice and vehicle for mobility is very much part of the post-2015 development agenda; it plays a significant role in realising economic prosperity. Studies have shown in recent years an important association between literacy/learning attainment and increased earnings

in developing climates (UNESCO, 2011, Hanushek & Woessmann, 2008, 2012). While the evidence is clear on its benefits to society, there are some forces to limit its widespread acquisition. For the case of India, this includes a diverse linguistic, demographic, limited education provision and weak policies, some of which are discussed in the next section.

### **1.3 English Policy & Education**

The adoption of English as a mainstream policy for any developing country is perhaps best described by Johnstone (2009, p 33); ‘as the world’s biggest policy development in education.’ To be successful in India, you must speak at least two of your home languages and English. Some highlight the absolute importance of acquiring English (Azam & Kingdom, 2011) and the positive returns it has in bridging a common understanding, while others refer to its impeding heritage and colonial legacies (Graddol, 2010). In terms of usage, various estimates have found similarities in English language proficiency, such as, Mishra (2000) suggesting that 5% of the country knows English in some capacity, whereas, other estimates, notably from Sonalde & Vanneman (2005) who concentrated on fluency, found 4% were able to read, write and converse while 16% had a basic understanding of the language. If we used the latter estimates, of the 4% of total population, 60 million Indian’s are proficient in English, a significant number in its right but significantly low when considering its 1.2 billion inhabitants. Some would argue the low estimates are a direct representation of the 1652 indigenous languages that are used and, in some cases, are not permitted to flourish (Mitchel, 2009). A familiar narrative within the postcolonial literature is the impact English has in local languages, highlighting the colonial legacy acts as an oppressor for modern times (Bhattacharya, 2016).

Since its inception at the dawn of independence, English has been more of a bridging language than a national medium (Ramanathan, 2008). This poses a difficult and often complicated situation for children who are learning it for the first time. Stringer (2015) suggests English exposure is acquired predominately through mainstream media and the school system. It’s the latter where efforts have been expended on improving access and developing structures for multiple linguistic experiences. As a result, two movements exist – at one end you have those who believe that world languages are “under assault by the forces of globalisation, cultural, economic and political forces that work to standardize and homogenize, even as they stratify and marginalise” (McCarty, 2015, p147). Whereas, others believe that the medium of instruction should par with international standards, which includes English-medium instruction as the main contributor in forwarding the aspirations



and hopes of families in a state of social mobility (Graddol, 2006). Interestingly, both parties would agree that the driving force between their vested interests is mainly in part to the rapid movement English has in advancing the global perspective in critical disciplines, such as science, popular culture, and business.

Since the National Commission on Education 1964, a three-language formula has been recommended for school-based education (NCERT, 2012). It works by providing the right to children to learn their mother tongue (often their State language), alongside Hindi and English (Mallikarjun, 2010). The formula is used in schools and promoted as a product of language assimilation, in that, all children will advance in three languages as they progress through the education system (NCERT, 2005). However, while there is a general acceptance of the formula in schools, there is a practical concern of its usage in complex learning situations and widespread implementation (Meganathan, 2011). Throughout history, there have been efforts to target illiteracy and promote English within the education system. However, a perusal of census data suggests these policies are not far-reaching. For instance, if we take the population of illiterates in 2001, this was greater than the population of the entire country at independence in 1947 (Graddol, 2010). Policies such as the Second Five Year Plan (1957-62) took the focus away from elementary education planning, steering towards higher and technical education. It left an everlasting legacy of a malnourished primary education system with weak policies on access and pedagogical excellence (ibid, 2010). Achieving universal education in India isn't a new focus. It has been on the political radar since independence from Britain in 1947, yet it remains a significant education policy concern based on large illiterate rates (Kingdom & Muzzamil, 2009).

The shift back towards acknowledging the benefits of primary education and resourcing the language classroom came through the Sarva Shiksha Abhiyan (SSA) programme, launched in 2001 (Govinda and Bandyopadhyay, 2008). It mirrored a new age of development coinciding alongside the Millennium Development Goals and Education for All (UNESCO, 2009) strategies. Its primary focus was achieving universal primary education, permitting free and compulsory education to 6-14-year-olds. Its supporting focus was to develop children as English speakers, part of the many necessary life skills needed for modern India. Although there have been advances in access, both are yet to be realised. Further evidence of policies promoting pro-English development was stipulated by the National Knowledge Commission (2007) stating the importance of English language

teaching from the first year in primary school. It also emphasised the significance of quality teaching and media-rich classrooms as being a high priority and associated with the likelihood of progression to higher levels of study. However, it failed to acknowledge the difference in schools and their approach to English. Instead, it called for pedagogical reform with an emphasis on quality teaching, moving away from rote-memorization and ensuring access to enriching textbooks. Primary schools in rural areas face significant challenges in providing quality teaching, learning materials and general access to English, as found in the report. The shift of attention from universal access to quality teaching is further supported by the Right to Education Act (2009) with a major emphasis on getting children to school and ensuring they stay the course. While it was well intended, the Act puts pressure on schools to reserve 25% of places to poor and disadvantaged children. The main argument from a school perspective is how best to accommodate children who are unlikely to have previous exposure to English teaching. Ramanathan (2008) highlights this case by suggesting disadvantaged groups would suit English teaching as a foreign language (EFL) because children already placed in school will already be taught English as a L1 or L2. Without an EFL programme embedded into the curriculum, it's hard to see how schools can accommodate quality teaching for all.

#### **1.4 Challenges in Primary Education and Reading Pedagogy**

India has the most extensive education system in the world with 200 million children in attendance, a success mirrored by the steadily growing enrolments in primary provision across the country (Muralidharan & Sundararaman, 2016). However, access to rural schools remains a challenge with high dropout rates when children transit through to middle and higher primary years (Mukherjee, 2011). The National Curriculum Framework (NCF) (2005) stresses the importance of mother tongue instruction in the early primary years. English also has equal status. However, guidance on its introduction in the primary years is variable. Most States offer English as a taught subject from the first year of Primary school and some as the medium of instruction (NCERT, 2012). However, as children transit through the system, most lag behind, which often prohibits their educational development and contributes to their withdrawal from the system altogether (Chakraborty & Kapur, 2008).

There are some factors that contribute to the system failing children at the beginning of their educational journeys. One important factor is the presence of teachers. An alarming trend of teacher absence has emerged in recent times with estimates indicating 24% of

teachers not present in the classroom (Muralidhara et al., 2016), costing the system roughly \$1.5 billion a year. Other factors include the quality of teaching instruction and reliance of singular rote learning pedagogies as the predominant method of knowledge transfer (Vaidyanathan and Nair, 2001, Dixon, et al., 2011, Counihan, 2015). The NCF (2005) states the importance of moving away from rote methods when designing learning experiences. However, many studies continue to report its dominance and report an over-reliance of textbooks because of translation problems or a lack of subject knowledge (Jhingran, 2005, 2009, NCF, 2005). Rote learning has a significant presence in India and other similar developing nations. Some argue of its existence because of teachers' own educational experiences and lack of formal training (Hardman, 2012, Hardman, et al., 2012), while others discuss teaching time and the pressures of delivering a syllabus before preparing children for examinations (Smail, 2013, Sriprakash, 2010). Rote methods have a historical significance in India, particularly during British rule (Counihan, 2015). Modern developments have led to its widespread usage, such as the adoption of communicative pedagogies preferred to grammar instruction (Prabhu, 1987). These types of methods require children reciting words and passages of texts in chorus and implicit word memory tasks (ibid, 1987). The method gained currency during a mass teacher recruitment programme in the mid-70s and is still popular today. One of the problems is with the complex translation and sounds in English for non-native speakers, this leading to confidence and knowledge problems facing teachers in modern classrooms (He, Linden and Macleod, 2007). The quality of instruction is a redundant area for most schools, not least for the undernourished rural school. Children from rural areas are considered to be first generation learners, placing a massive burden on them to develop learning capacity with little or no support from family members or quality of instruction (Dreze and Sen, 2002).

The poverty of teaching is best acknowledged from the NCF (2005) itself through the identification of teachers who have limited English language skills or creativeness in curricula/lesson design. Specific to reading instruction, the Framework provides a “workable approach to beginning reading” (p41) requiring schools to have a print-rich environment, reading activities to include partner work and oral recitals. It also places emphasis on schools to design wall charts for children to recognise letters ‘in addition to teaching letter-sound correspondences” (p41). Instructions and guidance on activities are in abundance in the policy document. For instance, it discusses the importance of

storytelling, encouraging speaking in a variety of languages, planned time for individual and group reading opportunities. However, there's no specific guidance on how schools (and teachers) implement sound awareness training. The Primary syllabus from for Classes I-V (NCERT, 2006) is broken into two levels where the first level encourages language familiarity and introduces basic written texts alongside speaking and listening tasks. Level I recognises the importance of letter-sound-word correspondences by requiring children to “notice its components- letter(s) and the sound-values they stand for” (p64). However, it's unclear how this is implemented in the classroom. There's no guidance on assessment or the types of sounds children need to identify in their first year of primary school. Further, by the end of level II (introduced at Class III) children are expected to read basic passages of text and use “written language in meaningful contexts”(p64). If we consider a student enrolled in Class I and followed the Level I route, it means within two years the learner must be able to read small sentences. This is a big ask for children who typically follow this route but now consider those entering at Class II or III, without remedial sessions or flexibility in learning English – it becomes virtually impossible for children to catch-up. Evidence from ASER, who have conducted rural surveys in India on English reading since 2007, confirm the difficulties children face in learning to read English in the early primary years. Their latest findings suggest only 32% of children in Class III could read simple words in English (ASER, 2016, p44). This gets worse as children progress to upper primary classes where 49% of children in Class VIII were able to read basic English words, a major decline from 2009 where 60% of children were observed as beginning readers (ASER, 2016, p44). These alarming results only magnify the requirement for pedagogical and system reform. Teachers are ill-equipped in the design process of English language teaching, and this affects children's capacity to develop core-reading skills required for progression.

### **1.5 The Choice of Location and Study Context**

The research took place in the rural township of Bhawanigarh, Punjab. It is situated in the District of Sangrur and is approximately 23 miles west from Patiala as illustrated in Figure 1.

**Figure 1 Map of Central Punjab, India**



Google Maps (2017a)

According to the census GOI (2011, p24) report, Bhawanigarh has a total population of 100,507 with 61,426 recorded as being literate. Data suggest a marginal improvement in the District and slightly above the National average for rural areas. However, there remain roughly 40% of people considered to be illiterate based on 2011 estimates.

Further, data reports job type and language preferences. The majority of jobs is in agriculture and manual labour work. The local language is Punjabi, but like other Punjab townships, Hindi and English are given equal status. Similar to the National picture, local parents recognise the ability to speak English can offer increased wages and access to higher education (Heyneman & Stern, 2015). As a result, there is an increasing demand for English-medium boards of education (Tooley, 2009). While English-medium suggests teaching instructions to be in English; there tends to be a bilingual approach in teaching and learning situations. Again, similar to the national picture, English is seen as a second language and Boards of education develop curricula materials to reflect the difficulties encountered in teacher-learner partnerships. One such Board is the Central Board of Secondary Education (CBSE), which places strict demands on schools to recruit English teachers with postgraduate qualifications. This puts rural schools at a significant disadvantage due to difficulties in attracting highly skilled teachers—a problem that affects other schools across the country. In the host school's context, teachers are typically recruited as far as Patiala (see map Figure 1) and trafficked between locations on a daily basis. The school takes a pragmatic stance on hiring highly skilled early grade teachers, but this remains a significant challenge. In particular, recruiting qualified-skilled English teachers for the lower grades is virtually impossible. Instead, teachers are hired locally but with limiting English language capabilities.

To reflect local trends, Punjabi is the medium of instruction in nursery and kindergarten year groups; English is introduced but as a stand-alone subject. However, when students transit to the first year of primary education there is a shift towards English medium instructions and the introduction of Hindi as a core subject. The dynamic of learning three languages without skilled teachers following a structured curriculum arrives at a critical period in advancing achievement in all three languages. It is perhaps unsurprising to find local children learning to read in any language a challenge without language structure and support. Indeed, recent data on literacy performance from Punjab suggests only 35% of children in Class III can read a Class II text ASER (2016, p55) in any language. Moreover, only 54% of those children who can read English words can understand their meaning (p184). These trends are a major concern for educational planners in the town and for families who live in the surrounding area.

Furthermore, the local context of providing equitable educational choices for girls is a challenge concerning enrolment and capacity in reaching parity in literacy achievement. The school reflects similar trends found Nationally (and historically), that girls tend to enrol late and leave school early (Nayar, 2002), with various explanations linked to cultural views on education (Azam & Kingdom, 2011), gender roles in society (Rao, et al., 2003) and perceived dangers of travelling distances to reach establishments (PROBE, 1999). These factors are corroborated by ASER (2016, p181) who present results from Punjab, where girls' leaving school rises to 8.3% by the time they reach secondary level education. And while this has improved regionally in recent years, it still raises important questions related to best education outcomes for girls and the importance of accelerating basic skills in the primary years.

Perhaps more alarming is the rate at which gender disparities occur when comparing progression routes and achievement gaps between boys and girls literacy. White, et al. (2016) examined gender differences in reading and mathematics using data from the 2005 Human Development survey. They found stark differences in reading attainment, where girls significantly underperformed compared to boys. The findings are consistent with recent analysis from Jain (2016) who predicted reading achievement levels using ASER's (2009, 2012, 2014) data. The main findings reported a widening gap over time with predictions based on differences found in the household and family members' level of education. Both articles discuss the implication of a scarcity of resources (either in school

or the home) but also the general attitudes towards girls' education as a universal driver of why the gap is increasing. In sum, the consensus of universal access to education provision leads to more significant learning outcomes is a falsism. Indeed, locally, there is an urgent need to understand the sociocultural aspect of educative demand and trends when designing equal learning opportunities.

School infrastructure in the Town is limited; there is one Primary and Secondary State school located in the Town's centre, and two recognised low-cost private schools (affiliated to CBSE), which can be found on the outskirts. One other regional school following the Punjabi State Education Board (PSEB) is also on the town's outskirts. Figure 2 illustrates the location of the host school (see  $\alpha$ ) and the densely populated housing in the main town.

**Figure 2: Map of Bhawanigarh (Study Location)**



Google Maps (2017b)

The host school lies to the North of the map; approximately 10 minutes walk to the Gurudwara as denoted by  $\beta$  on the map. No data from the District is available on school enrolment but based on regional ASER (2016) estimates; school access continues to improve at the District level.

Contact was made with one of the low-cost private schools in the area, since both offer English-medium education from Class I, and meet the rigorous criteria for affiliation to CBSE. The proliferation of low-cost private schools has mostly been down to failing State provision and the desire for English (Tooley, 2009) like others around the world. Bhawanigarh's education climate reflects this position. However, the lack of schooling infrastructure requires parents to consider schools from outside the area.

In sum, understanding the local context was crucial in the development of the present work. Given the demands of educating local children, school leaders wanted to look at ways of involving the wider community regarding access and opportunities for learning English. With a maximum intake of 850 students from Nursery up to Pre College education, the school is at full capacity - limited in funds and location for new classrooms. The choice of location reflected the study parameters of a rural setting faced with difficult challenges in offering educational access and assuring quality in L2 English reading education.

## **1.6 Rationale for the Current Research**

The NCF (2005) recognises the crucial connections children make with letter-sound correspondences and their value in word construction and later, sentence building. It also identifies the importance of children reading for pleasure and without the predominant use of textbooks to fill-in gaps of teaching and learning experiences. Reading attitude research is severely redundant in India with a specific lacuna in understanding reading habits of children. The NCF (2005) enforces schools and teachers to shape reading instruction around child-driven interests. However, taken altogether, these claims are not substantiated by a clear rationale or agenda for teachers to follow. Moreover, implementing any changes is problematic based on English being introduced (and alongside other languages) at different stages in the Primary years, confusing the school and class level. Therefore, reading instruction becomes complex, as, transferal from one language system to English is not supported by a structured curriculum. The lasting effects of failing to read in English are then translated to the higher primary years where Board exams determine routes into higher education and the workforce.

One way to combat the curricula challenges that threaten reading instruction is through a synthetic phonics approach. There is substantial research on phonological awareness being linked to reading achievement (Adams, 1990). The synthetic procedure involves children



sounding out letters (phonemes) before blending them back into whole words. Children learn new sounds each day alongside actions to support recall and to give context to the sounds being mastered. Synthetic phonics methods have garnered much attention across the globe as a methodological choice for accelerating reading skills. However, only a few studies exist in international development contexts that use an experimental framework (see Eshiet, 2014 and Dixon, et al., 2011). Therefore, further research is required to understand more about its potential utility for improving literacy rates.

As detailed in the previous section, there remains a challenge for quality teaching in English, a problem associated with remoteness when considering rural school development and its impact on learning achievement (Mitra et al., 2008). Based on these issues, it's important to look at alternative approaches to promote L2 reading development. Peer learning is an umbrella term used for some methods and strategies commonly used in-group and one-to-one scenarios. It is deeply rooted in socio-cultural theory (Rogoff, 1994) as it includes a more knowledgeable other (Vygotsky, 1978, 1994) who is tasked with guiding the learning process. In particular, peer cooperative strategies might consist of a range of learning abilities but are governed by a specific goal and by those who are slightly advanced in achieving task related objectives. Research has found cooperative methods accelerate reading achievement (Lan et al., 2006) and motivation in (Liang, 2002) in L2 contexts (Mohanty & Roy, 2013) when the tasks have followed a specific structure. Furthermore, little research has utilised its method in educational backgrounds within India. Indeed, no research, using peer cooperative learning strategies embedding synthetic phonics strategies exist in the early primary years within the Indian education system.

### **1.7 Aims and Emerging Research Questions**

Based on the factors contributing to rural illiteracy, as previously discussed, the present study aimed to bridge peer learning techniques with a synthetic phonic programme. Both are well-researched areas, but little is known about its potential use in developing environments. Therefore, applying a sociocultural theoretical approach, utilising a peer cooperate framework, this study was guided by the following research question;

- What are the effects of a peer-mediated phonics intervention when considering the phonemic awareness, spelling ability, reading skills, and reading attitudes of children from a rural North Indian community?

In answering the above, the study operated under experimental conditions. To understand the effects of reading achievement and attitudes over an intervention period, the following research sub-questions guided the RCT to ask specifically;

- Is there a statistical difference between male and female Monitors' baseline reading, phonological processing, decoding and native language test scores?
- What are the effects of Monitors' phonological processing, reading and spelling abilities when guiding learners in a peer-mediated intervention?
- Is there a statistical difference between male and female learners' baseline reading, phonological processing and native language test scores?
- Are there statistical differences between learners' reading, phonological processing, and native post-test scores when controlling for pretest scores? If so, where do these differences lie?

The trial also investigated attitudes to reading using a blend of exploratory and experimental designs. This process was guided by the following sub research questions;

- Is there a statistically significant difference in baseline academic and recreational attitudes, when considering the average (neither satisfied nor dissatisfied) test scores?
- What are the effects of Monitors' recreational & academic reading attitudes, before and after a peer-mediated trial?
- Are there differences in Monitors post recreation, academic and total reading aspirations based on gender when controlling for pretest recreational, academic and total composite scores?
- What are the main factors concerning Monitors' reading attitudes in academic and recreational settings?
- Are there differences in control and intervention group's recreational, academic and total English reading satisfaction scores?

## **1.8 Significance of the Study**

Based on the research questions above, this research will redound to the International Development and Education community. More specifically, it will be of interest to Indian policymakers and educational planners who are engaged in developing and responding to the improvement of early reading instruction and attitudes to reading in a second language. The central tenet of this research is to understand the effects of a peer-mediated synthetic

phonics intervention, targeting children's early reading achievement and attitudes to L2 reading. Reading attitudes and their direct link to reading success is well established in the literature (Black, 2006, Guthrie & Greaney 1991, Mckenna, et al., 1995). The same can be said for peer learning pedagogies (see Johnson et al., 2000, Slavin, 2009) and synthetic phonic procedures that target struggling readers (Ehri et al., 2001, Hatcher, et al., 1994). However, all of these fields are underrepresented in the international development and education field of enquiry. Given the global movement of working toward an 'Education for All' paradigm, this research will be of particular interest to those working towards 'Goal 4' of the Sustainable Development Goals. It provides a direct focus on improving the quality of reading instruction and attitudes that contribute to accelerating reaching achievement while engineering lifelong literacy development.

### **1.9 Organisation of Thesis**

This thesis is presented in the following way. The first Chapter detailed some of the major challenges facing rural children in being able to access quality reading teaching. It also showed the aims, scope and emerging research questions of the current study. The rest of the thesis is organised into five Chapters, beginning with the literature review in Chapter two that draws upon the critical theories of language acquisition, models of reading and attitudes towards reading. Chapter three presents the methods used and describes the three phases constructed to search for, test and successfully implement a trial. Chapter four presents baseline and post-intervention results from both quantitative and qualitative data collections. These results are then discussed in Chapter five, where the implications for policy and practice are situated into the broader literature. Finally, Chapter six brings together the main findings and details their importance for future work and scholarly enquiry.

## Chapter Two - Literature Review

### 2.1 Introduction

This review considers the theoretical and practical work undertaken in understanding how children learn to read. It begins by acknowledging the importance of sociocultural theory as it interacts with emergent literacy development. From here, details drawing upon phonological models of reading instruction are explored and evaluated. Also, reading attitude research is referred to and identified as a contributory cause in accelerating reading development. Finally, the review investigates the scope of the peer learning method's effectiveness in creating interdependent readers via cooperative controls.

### 2.2 Sociocultural Theory & Emergent Literacy

For the present study, the significance of sociocultural theory (SCT) in second language acquisition (SLA) is crucial given the focus on cooperative learning structures outside the domain of formal learning. However, before investigating this further, it is worth visiting the foundations of SCT as it applies to culturally diverse groups.

Sociocultural theory is commonly associated with a mechanism that allows for higher-level learning that takes place via social mediation. In theoretical terms, it considers knowledge to be mediated through social experiences, individually and with intermediary assistance (Lantolf, 2000). In learning development situations, these interactions are best described through Vygotsky's (1978, p86) Zone of Proximal Development (ZPD) as;

*The distance between the actual development levels as determined by independent problem solving and the level of potential development as established through problem-solving under adult guidance or in collaboration with more capable peers*

Vygotsky acknowledges cognitive development as a process of enacting within the upper and lower ends of ZPD. Children operating at the top end will be exposed to complicated instructions and tasks requiring support for undeveloped skills. Through assistance from capable peers, children are instructed and supported before receiving less instruction. At this stage, children have created a level of mastery and become independent without the need for further guidance. The process is summarised by Vygotsky (1987, p211) as "*what a child is able to do in collaboration today he will be able to do independently tomorrow*"

(Vygotsky, 1987, p. 211). It assumes that children are not passively engaged in a filtering process (like in rote learning situations), instead, extend upon their imitation experiences leading to autonomous and independent learning (Lantolf, 2000). Independence is, thus, achieved when children are imitating others' and can complete tasks without assistance or guidance (Zuckerman, 2003).

Sociocultural approaches must consider the broader social circuit when individual development is concerned. The ZPD must reflect what Vygotsky relates to as whole activities, the requirement of social interaction and change in which the learner becomes independent (cited in Moll, 1990). The role of play in early childhood is an essential 'whole activity,' in which children can experiment with language and actions in unstructured situations. Indeed, imaginative play scenarios have symbolic references in the creation of experimental language; it is in use as a vehicle for transmitting cultural values (Sawyer, 1997). Children at play will be fully engaged in a task that is accessible, meaningful and unrestricted of formal instructions from adults or other authoritative figures. Social interaction is at the heart of these exchanges. Vygotsky (1981) refers to higher mental functions, such as learning to read, as a process that must go through an external stage where it is prepared (and accessible) in the social domain. In this regard, learning to read is a social construction that can only be obtained through interactions with other adults and children. After mastering a skill, the ZPD changes and children repeat the process in the acquisition of new skills. The sociocultural theory posits that ZPD is transient by nature and is constructed to create future development opportunities for knowledge and skill exchanges. This is because SCT is best explained through a mediated learning mechanism that has no established rules or linear processes (Barr, 2001), as with some of the earlier models described in this review. Moreover, as Gass and Selinker (2008) discuss, it can be seen in multidimensional form, but it primarily concerns how learners can develop new language systems (other than their own) at varying degrees of success.

As with the previously discussed theories, SLA through SCT has proponents in behavioural (Skinner, 1968) and cognitive (Roebuck, 1998) disciplines that have shaped our thinking about the topic. However, where cognitivism represents the mental organisation and, where behaviourism seeks stimulus for language development, SCT theories represent a kind of cognitive-social bridging. Lantolf & Pavlenko (1995) express this through the seminal work of Vygotsky and SCT that cites the biological function of language acquisition to be

triggered and mediated by sociocultural artifacts and symbols. What this suggests is the presumption that innate structures are intertwined with sociocultural events, such as, children playing in their cultural neighbourhoods (Counihan, 2015), engaging in meaningful activities and governed by their social cues.

We can attest to language and literacy development in social contexts as an emerging force. There is a plethora of research that indicates literacy as an emerging concept in children's later reading capabilities (Whitehurst & Lonigan, 1998, Lonigan et al., 2008). The next section reviews the process of emerging literacy through two hypotheses.

### **2.2.1 Relationships of L1 - L2 Reading and Emergent Readers**

Two hypotheses formulate the relationship between L1 and L2 reading development, the linguistic interdependence hypothesis (see Cummins, 1979, 1981, 1991) and the linguistic threshold hypothesis (Cummins, 1976). The former considers reading abilities to transfer from the L1 to the L2, and there's a fundamental similarity of properties or commonality between both processes (Bernhardt & Kamill, 1995). Reading in the target language will share the same skill properties in the L1. It suggests that if learners' are proficient readers in the primary language, such skills are in use on an experimental basis in the second (Moll, 1994). The theory posits that adequate exposure is central to how effective transfer is, whether in educative situations or the wider social context (Clark, 1980). Contrastingly, the linguistic threshold hypothesis argues that learners' must successfully gain advantages in using the target language before applying L1 reading strategies therein (Bernhardt & Kamill, 1995). In other words, there is a higher functioning, if not, proficient understanding of L1 language ability that is used to develop the L2 reading ability.

Both hypotheses link to various empirical studies examining their contribution to transfer. Indeed, as Grabe (2009) indicates, there is no debate on whether the transfer from L1 to L2 occurs, but an interest in when the transfer takes place. Both theories give some clues in explaining the complexities involved in understanding L2 reading as an emerging skill. However, it is 'emergent literacy' that goes further in explaining how both theories interact along a developmental reading timeline. There are various definitions of 'emergent literacy' in the literature, but it can be simply defined as a process of beginning to read from an early age – right through to proficiency and includes learning experiences in structured and unstructured situations (Lonigan, 2006). The unstructured element involves learning to read

in the home environment and includes skills, knowledge, and attitudes about reading ahead of formal instructions (Whitehurst and Lonigan, 1998). The social context provides a natural connection with printed media, books and stories for children to access. It facilitates a connection between the learner and what is to be learned. However, exposure is not enough to engineer reading and writing activities that have no formal instruction. Communities and environments that have complex language and limited instruction infrastructures (Bowman and Trieman, 2005). The relevancy and implication of such a situation is a major concern for the current study based on the factors as outlined in Chapter One. In the next section, the structured elements of reading are reviewed. This section considers the various contributions in understanding how reading is developed.

### **2.3 Theoretical Models & Reading Development**

This section identifies the various ideas and theoretical models involved in the reading process. It specifically targets ideas and contributions made in furthering understanding of how reading is acquired and developed. The first section concerns models (and nonmodels) of learning to read, both being popular during the 1980s (Snowling, 2007). The section considers contemporary ideas of reading development that rests on phonological pedagogies.

#### **2.3.1 Stages of Reading**

Broadly contested in the reading development literature are two opposing theories on the process of reading development- ‘stage models’ (Chall, 1983, Ehri, 1998, 1985) and ‘non-stage models’ (Goodman, 1986). The former considers reading as a developmental pathway, where one stage is completed before passing onto the other, whereas, the latter affirms that reading development is consistent throughout learning (Kurvers, 2006).

Stage model theorists have similar expectations in the progression and achievement of reading. Chall (1983), Frith (1985) and Ehri and Roberts (2006) have stages that are investigated further here given their relevance to emerging literacy, phonological processing, and developmental reading. For Frith (1985), pre-readers must go through three stages; logographic, alphabetic and orthographic. She suggests that children must visualise (logographic) and distinguish letters before proceeding to the next step. From here, pre-readers develop an awareness of the alphabet and also start to separate sound patterns from

building a richer understanding of the orthographic features through phoneme-grapheme recognition. Ehri (1995) also developed a similar model describing a process whereby pre-readers will go through pre, partial, full and consolidated alphabetic stages. The pre-alphabetic phase requires pre-readers to visualise letters and experiment with their use – but they're at the pre-reading stage, it is similar to the logographic stage as proposed by Frith (1985). From here, pre-readers' will make minor sound-letter connections within what Ehri (1995) names the partial-alphabetic stage. Only when children can differentiate words/sounds by decoding and segmenting, have they fully achieved the full-alphabetic stage. Finally, consolidate-alphabetic stage permits pre-readers to use words in different contexts through exposure of deeper orthographies inclusive of multi-visual morphological and syntactical arrangements. In summary, Ehri (1995) points to the process of sight word construction as an important advantage in developmental reading. After practice and exposure, words are confined to memory and recalled when required. Both Frith's and Ehri's models represent similar transitions in which children progress from visual to phonological cues. The only difference being the partial alphabetic stage, which, bridges to the visual stage, including the phonological processing skills pre-readers start to experiment. Moreover, both theories, although slightly different, agree that children must go through a visual, phonological and consolidated phase in word development to achieve reading achievement.

Chall (1996) offers a different perspective and model through her six stages. The first four stages are reviewed here given the significance of the current study's focus on developing pre-readers into fluent readers. The first phase identifies with pre-readers and is recognised by letters and book exposure. It rejects the idea of logographic experiences as it focuses on instructional, systematic exposure of letters and sounds, thus removing the guesswork required in word and latterly, sentence construction. Chall branded this stage as Stage 0 placing significance on the varying speeds, exposure and challenges children will face with printed words. It's a stage between 0-6 years and identifies children who move between the early years and the beginning of primary education. Chall (1996) identifies with various readers and different stages within Stage 0. In the context of this study, children who haven't had exposure to L1 books and reading opportunities - per se, may also find it difficult to transfer. The importance of letter recognition, alphabet development and some recognition of sounds are salient skills required at this Stage. The second stage focuses on decoding and sounding out words that are unfamiliar. It is a process whereby children experiment



more with sounds and start to make sound/letter correspondence ahead of word development. In the third stage children will bring together their decoding skills and use shift to building their vocabulary and have an emphasis on developing the fluency in their reading. One of the criticisms of Chall's model is the lack of fluidity between each stage. In other words, Chall's model is a sequential hierarchy that early readers pass through – one stage at a time. It's possible to understand some sound components without having a full understanding of the alphabet. Further, children may understand, at a fundamental level, a theme of a story, something that Chall argues happens in higher primary grades. Ehri (1999) acknowledges the weaknesses of the model to relate to the age specification it proposes. It's why she uses 'phases' in her developmental model.

### **2.3.2 Reading Pedagogy**

The best way to teach reading is an age-old debate with various methods and models emerging as the best way forward (Barr, 2001, Beard, 1990). In this section, the review is concerned with two methods given their significance to the current study; the whole language method and the phonics method. Both have been in vogue with policymakers at various stages of history. The first section details the historical and contemporary significance of whole language instruction. It argues for a holistic approach to learning to read, linking to social and cultural activities that facilitate learning. The second section concerns phonetic methods: analytic and synthetic approaches to reading. Both are reviewed in the context of developmental reading effectiveness and evaluated for use in the present study.

### **2.3.3 Whole Language Approach**

The whole language approach to reading isn't necessary a method, per se, it is an approach to viewing language holistically and not segmented, decoded or broken up into constituent parts (Dixon, 1996). It is closely related to constructivist learning and is seen as a *top-down* approach (Moats, 2007) in the development of reading. The method is closely linked to Goodman's (1986) work in that reading is a social practice that happens before formal schooling and rests on behavioural and maturation in learning contexts. The central argument against *bottom-up* approaches to reading development, where children rely on decoding, segmenting and phonetic techniques; is the mechanical structures they over-

emphasise, limiting previous reading skills obtained in non-formal environments (Donato, 1992). Whole language reading pedagogies rest on teaching reading as words appear in the text and not broken up into various arrangements and prepared separately, or sequentially, as previous methods in this review have demonstrated. Indeed, the whole language approach to reading isn't linear but allows access through speaking, listening, reading and writing – the 'whole approach' (Dixon, 1996). This approach considers the wider social aspect of reading, permitting users to discuss the text and arrange meaning. Such skills are interrelated and work in tandem in the promotion of learning new skills and fluency. For pre and developing readers, pedagogically, the approach can be used to promote the following; words that should be read naturally and part of a whole text, words to facilitate a discussion and have meaningful context and words that have the capability to grasp the learner's interest (Goodman, 1986, Gunderson & Shaprio, 1988). Although many definitions exist in the literature related to the various skills embedded into the whole language approach, Bergeron (1990, p6) offers an all-rounded view based on her extensive research by inferring it is;

*A concept which includes the use of real literature and writing in the context of meaningful, functional, and cooperative experiences to develop in students' motivation and interest in the process of learning*

The cooperative experiences allied with the functional and meaningful have special significance in the current study. Children may have already had access to printed books and have discussed reading materials in their home or school environment. While this might be conclusive with regards to L1 reading instruction, it isn't clear if it translates to L2 reading experiences. For it to be successful, readers must make sense of what they're reading in the creation of new knowledge, skills and language processing (Weaver, 1996). Some have argued that automaticity in word recognition is dependent on the number of books children read (Stanovich et al., 1985). However, in the context of this study, it is predicted that lower-income families may be at a disadvantage in the exposure of written text in the L1 and L2 reading contexts. Automaticity of word recognition is best described through structured instruction. One way in which the whole language approach translates to the classroom is teaching children using flashcards to enable a greater 'whole' word exposure to build vocabulary and later sentences (Watson, 1998). Flashcards are a useful tool in matching pictures to associated words and, in turn, permits children to interact with new words in their whole format (Brooks and Brooks, 2005). Children will use these words

in their natural learning environment to add meaning through discussion repetition. Whole word tasks are designed to promote recurrence and build new schema while engineering pre-reading development (Ashby and Rayner, 2006). In doing so, it rejects the notion that language should be broken up into sound bites and syllables that redirects the learner to concentrate on these functional tasks rather than contextual meaning. As Goodman (1986, p7) argues;

*We took apart the language and turned it into words, syllables, and isolated sounds. Unfortunately, we also postponed its natural purpose — the communication of meaning — and turned it into a set of abstractions, unrelated to the needs and experiences of the children we seek to help.*

However, some have argued that there is no isolation between whole language and phonic approaches. Stahl and Miller (1989) detail that both methods have similar elements. For example, they comment on the bridge between oral and written skills when children are producing language, there's a convergence of capabilities in recognising sound output and skills in the production of orthographic texts, in general. Second, instructions usually rest on a literature suitable for children to interact with. Both use literature accordingly; however, it's argued here that the whole language method will direct through semantic cues, whereas, the phonics method will use sound-letter cues for instruction.

#### **2.3.4 Analytic Phonics Approach**

There are complexities involved for L2 learners whose targets are to read in English. The English alphabetic system has a deep, complex orthography that has a limited phoneme-grapheme correspondence (Spencer, 2010). For transparent languages, the correspondence is mainly one-to-one where the sound is mapped to the grapheme. English is more complicated as some graphemes can be sounded through different phonemes and vice-versa. Concerns about the complexities of learning to read in opaque orthographies have led to studies investigating phonological approaches to early exposure to grapheme-phoneme correspondences (Dombey, 2006). The analytic approach is designed to 'break up' and 'analyse' whole words investigating sounds contained within, a process of segmenting words into phonemes (Torgerson et al., 2006). Goswami (1986) proposed an onset and rime method in word segmentation. For example, take the following words: *tap*, *tops*, *tin* and *ten*, children start by discussing each word, looking at its shape (grapheme)

and its relative likeness to the other words found in the series. For this example, the onset<sup>1</sup> Stress is the /t/ phoneme, and if successful, children will deduce that this onset sound is also found in the other corresponding words. Sometimes referred as the ‘whole-to-part’ method, analytic phonic methods instruct children to break down whole words into smaller parts and experiment with substituting consonant and consonant blends. If children can read the word: *car*, they recognise the onset as /c/ and the rime as /ar/ to complete the word. Moreover, they apply these techniques with similar rimes that might be studied in that sequence, such as *far*, *jar* and *bar*. Thus, if implemented correctly, children who can read *car* will also be able to read *far*, *jar* and *bar*. Success in the analytic approach is conducive to finding patterns and having some phonemic awareness (Ehri, 2005).

Over recent decades there has been substantial evidence in support of successful reading interventions that use phonological awareness teaching-instruction alongside correctly leveled texts in the promotion of early reading development (Hatcher et al., 1994, 2004, Iverson & Tunmer, 1993). One study in India reported benefits to English reading when printed text and picture cues in comics were used in the school and home environment (Kalia & Reese, 2009). Banerjee et al., (2007) hired young women to teach basic literacy skills using analytic dialog reading using flashcards and storybooks. Average literacy test scores of children as poor readers, and part of the intervention group, improved by 0.28 standard deviation – where the poorest readers at baseline, specifically, had significant gains. Similarly, there have been advantages to readers grouped by ability as part of the Read India programme, developed and endorsed by Pratham since 2003 (Chavan, 2003). Pratham's reached over 300,000 marginalised children (ASER, 2015) using a combination of analytic and remedial storytelling approaches to reading acceleration. Trained volunteers or local teachers identify struggling readers before classifying them into ‘paragraph’ or ‘word or letter’ child (Chavan, 2003). Through a ‘barahkhart chart,’ children are asked to segment words and reassemble them before moving onto sentence and then paragraph reading.

### **2.3.5 Synthetic Phonic Approaches**

In contrast to the top-down approach to reading through constructivist mechanisms, the phonics method advocates a bottom-up approach. Goswami (2002) acknowledges the

---

<sup>1</sup> An onset is the first part of the word – for example, /p/ is the first onset of the word <pit>

synthetic method as a process of identifying small units of sounds (phonemes) that are then blended into words. Therefore, phonemes are matched to particular graphemes and taught at a pace that enables children to recognise the formation of words through sounds that are later blended (Jolliffe et al., 2012). They do this by breaking down selected words (similar to the analytic method) by their units of sound and apply acoustic rhyming for word construction (Lewis and Ellis, 2006). For example, a teacher may display the word 'cap' in which children will respond by breaking the word down into its phonemic units, this being /c/ /a/ /p/. Children are then instructed to blend (synthesize) the sounds to reform the word 'cap' and read aloud. The method has gained currency throughout history as a multisensory and dynamic way to teaching early reading to children (Chall, 1967, Pressley, 2006, Snow, 1998, Stahl and Miller, 1989). Research has consistently found an association between being phonemically aware and developing more laborious tasks required to become an advanced reader (Ehri et al., 2001, Hatcher, et al., 1994). There is some research that indicates the necessity of having specific phonics training ahead formal reading instruction (Maclean et al., 1987, Goswami, 1999). However, whatever the angle on when phonics instruction takes place, there are various research syntheses and analytic reviews that all agree on pre- and early reading teaching should be geared towards phonological processing, decoding, vocabulary development and comprehension tasks (see Adams, 1990, Slavin, et al., 2009b, Snow, et al., 1998).

Most of the evidence of phonics instructions have taken place in western-styled classrooms with access to relevant materials and used within a broader rich literacy environment. For example, Stuart (1999) found that when comparing the Jolly Phonic (synthetic instructions) to the Big Book (analytic instructions), children's reading and spelling differences were effectively significant in favour of the synthetic method. Reading ages were more significant in the synthetic group compared to the analytic group and the effect sizes reported were large, reading  $+0.65$  and spelling  $.87$ .

Another study in the UK that re-energised the reading debate was a study by Johnson and Watson (2005) that tracked the development of readers over a period of 6 years. They compared three types of phonic methods (analytic, synthetic and a mixture of both). The final report concluded;

*After six years at school, children taught by the synthetic phonics approach read words, spelled words and had reading comprehension skills significantly in advance of those illustrated by the analytic phonics method (Johnson and Watson, 2005, p14)*

The results led the Rose Review (2006) to investigate the best practices of teaching with the results from the Watson and Johnson (2005) study leading to the adoption of synthetic approaches to be used in early reading development. However, the methodological rigour employed in the study is often criticised. Torgerson (2006) notes the crossover of pedagogical methods causing the analytic method not to be studied over the same period as the synthetic. Stannard (2006) argued that while there was good phonic knowledge and progress in reading there was little attention or progress on comprehension. This is often a criticism surrounding the synthetic approach. However, when considering the exposure of new words to promote confidence in reading, the synthetic method is perhaps geared towards early readers or those reading in L2. Cross-language transfer studies point to the mixed results in which phonological processing in the L1 affects L2 reading (Goswami, 2005). For languages sharing similar alphabetical structures, phonic instruction can enable successful transfer of target languages. Durgunoglu et al. (1993) in a Spanish class in the USA demonstrated that phonological awareness in Spanish is correlated to awareness in English that produced greater effects in word reading. However, there are some limitations when the language doesn't share a familiar script. In this regard, Wydell and Butterworth (1999) investigated this by offering the granularity and transparency hypotheses that view orthographic differences may impact learning to read. The transparency hypotheses acknowledge how well sounds map across to target languages, whether they share a similar one-to-one mapping. English is often considered an opaque language, and it's these complexities that make transfer difficult for L2 learners. For example, a study by Mishra and Stainthorpe (2007) investigated phonological awareness in English and Oriya (Indian language of Orissa). They found that phonological processing was consistent in Oriya word learning when children access it as their primary literacy language and not as a second language. However, English word learning was associated with phonological processing even when it was learned as a second language. The findings suggest the lack of transparency between both languages and in recognition of Oriya as an alphasyllabary script. This is a significant result given the Punjabi writing system (Gurumurki) will be used in the current study and may have similar implications for phonological processing and transfer. Another study completed in Canada on Punjabi speakers by Chiappe and Siegel (1999) found a relationship between English and Punjabi phonological processing

and reading. In a similar study, Stuart-Smith and Martin (1997) recognised the same properties located in English and Punjabi phonological processing to be the same. Both studies imply that there to be sharing some sound properties and infer there to be some transparency between both languages.

The literature points to a range of synthetic phonics studies and their effects in resources rich and developed classrooms. However, only a handful of studies have been completed in developing countries. One study by Dixon et al., (2011) conducted an intervention in 20 unaided low-cost private schools in Hyderabad, India improved reading ages by one year and one month against learners' chronological age. Using a battery of spelling, reading and dictation tests, the intervention group outperformed the control group, whose primary method was rote-based pedagogies. The intervention group received synthetic instructions over a one-hour time and followed the scheme of work similar to the Jolly Phonics package learning. A similar study was developed in Nigeria and reported benefits to learners' knowledge of spelling, decoding, and general reading when exposed to synthetic methods (Eshiet, 2014). In this study, further implications on the enjoyment of teaching using synthetic approaches was also reported, as the study used trained teachers from local schools in developing practice. It found that collaboration with teachers was an important factor when learning to sound and blend words in a L2. Using a quasi-experimental design with pre/post testing, the findings suggested a statistically significant difference between control and intervention groups' reading ability and synthetic knowledge. The intervention group had major gains when considering blending, decoding and creating new words.

#### **2.4 Reciprocity in L1 and L2 Reading Attitudes**

The literature renders reading motivation as a *sine qua non* for success in learning to read in a second language (Black, 2006, Guthrie & Greaney 1991, Mckenna, et al., 1995, Yamashita, 2007). Ever since seminal research undertaken by Gardner and Lambert (1959) on L2 motivation, there has been a gradual shift in understanding how motivation relates to reading aspirations and attitudes. As part of on-going research into this area, Gardner (1985, 2000, 2001) has redeveloped a theoretical model that is underscored by a sociocultural epistemology. Known as the socio-educational model of language learning, the main component of the model refers to the *integrative motivation* where learners want to learn the target language through positive attitudes to using it within a language

community (cited in Mori, 2002). It suggests that motivation to learn a L2 in this context will be driven by the linguistic community at large. Therefore, children aiming to learn a L2 for reading purposes require buy-in and openness for learning. It also applies to attitudes to learning in formal and non-formal situations. This refers to the exchanges the learner has with resources (including human resources, i.e., teachers, language instructors) and how they impact on the L2 learning context. It's a truism that children who're exposed to fun and exciting teaching will improve their motivation and desire for learning, compared to standards that are repetitive and unimaginative (Gardner, 1985).

While there has been some development in L2 attitudes, most of the research has been influenced by L1 models (Erton and Karakas, 2010), even when, and according to Day and Bamford (1998), learners' experience L1 exposure that informs their thoughts, feelings, and motivations in the L2. Indeed, most of the research has focused on cognitive processing (ibid, 1998), commonly agreed as essential in developmental aspects, however, it is argued in this thesis that more could be known in attitudinal properties, such as, essential factors with regards to L2 reading difficulties and home and schooling habits. Firstly, we consider the theoretical and scholarly work undertaken in this domain.

Reading attitudes is a multifaceted domain as it has different meanings in different contexts. Mori (2002) defines L2 reading attitudes within multiple domains where learners' are directed by intrinsic/extrinsic values while acknowledging its importance and efficacy when targeting achievement. Smith (1990) agrees that motivation is a factor due to its relevance in promoting reading experiences. While others consider the approach to reading that directs the learner to accept or avoid reading situations based on individual choices (Alexander and Filler, 1976). For children in India, it's important to develop a reading programme that enables access and development but also to promote reading itself. Much of the research confirms the importance of reading attitude and its successful development of children's ability to be literate (Baker & Wigfield, 1999, Wigfield and Guthrie, 1997). However, there is a lack of research into reading attitudes and literacy development concerning low-income countries. One study by Memon (2014) investigated the relationship between L1 and L2 reading attitudes amongst learners from Pakistan. The study also made comparisons across urban, rural areas and socio-economic backgrounds. Its main findings reported high correlations between L1 and L2 attitudes amongst rural learners when compared to their urban counterparts. With significance to the present study, it found background variables and accessibility in the social context to be important predictors of improved reading attitudes in the L2. Similar findings from India via the



National Survey on Youth Readership (NYRS, 2009) reported one-fifth of sampled literate children read non-academic books during their leisure time. Moreover, problems are reported with deficient levels of reading for pleasure opportunities in rural areas compared to urban centres. While both are considered low, rural settings are proportionally smaller. Overall, the survey presents a bleak view of reading in social and academic settings; it also suggests a disconnected and lack of engagement children have with reading materials.

Attitudes to reading must take a comprehensive appraisal of how pre-readers access reading and the conditions that present before, during and after reading events. As Mckenna, Kear, and Ellsworth (1995, p934) state; *“even for the fluent reader, poor attitude may occasion a choice not to read when other options exist, a condition now generally known as aliteracy.”* In summary, the conditions must be right for children to access reading options, whether in the L1 or the L2. In considering of the multifaceted domains of reading attitudes, some theorists have developed reading attitude models and are these reviewed here with their relevancy to the current study.

#### **2.4.1 Day and Bamford (1998) Model**

The first model offered by Day and Bamford (1998) is specific to L2 reading attitudes and at present, remains to be the only model to address L2 reading attitudes directly. They illustrate four domains of accessing L2 reading and subsequent attitudes. These include the following; L1 reading attitudes, previous experience of reading in a L2, the interaction of L2 environments (culture and native speakers) and the L2 learning environment. Central to how their model works is the onset development of L1 reading attitudes, they conclude;

*If students have had experiences with learning to read in the first language, these lessons will influence their attitudes toward reading in the second language (Day and Bamford, p 23)*

Therefore, according to this model, there is a transferal of experiences (positive/negative) from the L1 to the L2. Uniquely, there is no reference to proficiency in their model, which, might be seen as a contributing factor, where accelerated proficiency in the L2 may lead to more critical attitudes (Jimenez, 1995). However, this remains uncertain, as skills may not

be correlated to reading attitudes and not involved in L2 language processes (Yamashita, 2007).

The next models, although primarily developed with consideration to L1 reading attitudes, give further theoretical and practical justification for the current study. Indeed, already presented is the significance of L1 reading attitudes and the contribution it has on L2 reading attitudes.

#### **2.4.2 Ruddell and Speaker (1985) Model**

This model considered an interactional element comprising of four constructs embedded in understanding reading attitudes. They include the environment, knowledge utilisation and control, declarative and procedural knowledge and reader output. The theory was later developed (see Ruddell and Unrau, 1994) to place emphasis on teacher-learner partnership with the effective construct (belief/motivation) playing a crucial role in developing positive reading experiences. The knowledge utilization and control construct has a specific purpose for the current study. It bridges the affective and cognitive aspects of reading attitudes and shapes “*expectation for content, processing time, and product*” Ruddell and Speaker (1985, p757). In this regard, a meaningful text will receive higher rates of processing and output with regards to belief and future motivations, whereas, text that appears disinteresting will be the opposite leading to disengagement. Furthermore, the declarative and procedural knowledge component also has significance to this study, as it refers to the construction of language and prerequisite knowledge a reader possesses. It considers reading abilities through decoding, synthesizing and arranging words in conjunction with other constructs as predictors of reading attitudes. Children displaying difficulties in these constructs will invariably display negative attitudes towards reading.

#### **2.4.3 Mathewson (1985, 1994) Model**

The Mathewson model offers an understanding of reading attitude acquisition. It is made up of a tri-component approach to understand how reading attitude mediates reading development. The model consists of the following components; the cognitive aspect, this includes readers' previous thoughts and beliefs; the affective component, this taps into readers' feelings and motivations; lastly, the conative, which refers to the action readiness

(personal goals and behavioural outputs) readers' associated with reading. The theory has its routes in socio-cultural ideas as it involves influencing reading not just as a linguistic process but also as a social practice. Perhaps the most influential aspect of the theory is how it interacts with other actors and experiences in shaping reading attitudes. An earlier version of the theory discussed various constructs in the shaping process, Mathewson (1985, p847) stated;

*“The source of a book or magazine might be a friend, a teacher, or a club. Thus, the present model incorporates the notion that relevant attitudes toward reading may include attitudes toward the content, format, and form but may also include such attitudes as those toward reading as a universal good and attitudes toward the source of reading materials”*

As seen in the above quote, there's an acknowledgment of outside linguistic constructs, such as materials and experiences that influence reading attitudes. In light of the current study, the theory posits a useful framework when considering reading attitudes that apply linguistic traits alongside cognitive and affective constructs. Reeves (2002) completed an exhaustive review of reading attitudes and note this model to be wholly representable of an overarching understanding of reading attitude theory. However, Mckenna et al., (1995) criticise the lack of evidence in support of environmental and social influences. They argue that its reading incidence by way of feedback that is central to the original theory, and Mathewson's model doesn't represent reading attitudes over time.

#### **2.4.4 Mckenna (1994) Model**

In response to Mathewson's and Ruddell's models, Mckenna (1994, p30) combined these ideas into a single model which primarily focused on beliefs as the driver of reading attitudes, he stated,

*“The proposed model preserves Mathewson's notion of a decision to read, which is a consequence of subjective norms, intent, and attitude – all three of which are conditioned by contingencies”*

Central to these subjective criteria was how attitudes would change over time. What became his model included three constructs that reflected; beliefs about outcomes, assumptions about others' in regards to self when considering change, and the outcomes of specific incidents in reading (Mckenna, 1994, p938). The theory encompasses the social

environment where agents, actors/experiences, peers and other interrelated factors have a direct influence on how the reading experience is conditioned and ultimately played out. In turn, it has an impact on the individual as a reader, for example, how positively/negatively the reading experience has had and what it will lead to.

The theory offers a significant contribution to the current study due to its cultural and social outlook that is interlocked with reading attitudinal experiences. More specifically, peers and peer groups act as social actors and offer an invaluable avenue for knowledge, experience and motivation (whether positive/negative).

## 2.5 Peer Learning: Theoretical Dimensions

### 2.5.1 Introduction

This section highlights the theoretical footprints of peer learning dimensions and reflects the numerous designs and methods created over time. The first section gives an overview linking to pertinent research that locates its origins. Following this, the next section reviews social interdependence theory, which largely explains the theoretical blueprint for all peer-related methods

Peer learning is a well-researched area that has crossed many applications and streams in science, particularly in the social sciences (Johnson et al., 2000). Research has been verified in a myriad of experiments and observational studies that have led to theory generation and practice-based developments (ibid, 2000). Perhaps the reason why it is favoured methodologically in education is its ability to detail step-by-step instructions that various stakeholders can follow and benefit. The origins of peer learning are autochthonous within anthropological studies, which engineered its foundation and maturation into educational fields. Mead (1936) suggested primitive beings making collective decisions which, design sociological templates as a matter of conflict/resolution that is, in turn, behaviourally orientated (Coleman, 1961). For the latter, a whole body of research exists<sup>2</sup> that reinforces peer cooperation as much as a psychological construct and an apparatus for educational instruction. Indeed, with the behavioural movement in the 1960's, peer learning (and more pertinently, cooperative learning) became apparent in the social interdependence movement (Deutsch, 1962, Johnson & Johnson, 1989). It is attributed to seminal work

---

<sup>2</sup>The following researchers studied conflict resolution and distraction amongst peers in educational institutions Baron, et al., 1978; Coleman, 1961; Crandall et al., 1965; McClelland, 1961; Sanders et al., 1978.

conducted in reinforcement and context (Skinner, 1968, then later Bandura, 1977), cognitive-developmental (Piaget, 1950) and sociocultural studies (Vygotsky, 1978).

Given its widespread popularity, a meta-analysis conducted by Johnson and Johnson (2000) suggested that the breadth and dimension of peer learning are endless. At the time they found over 900 studies that produced more gains over individualistic and competitive methods of educational instruction. They suggest that peer cooperation levels are generalisable because these studies operate in different cultures relating to the economic status, age, race, ethnicity (see Johnson & Johnson, 1990, 1994, Sharan, 1980, Slavin, 1977, 1991). Further research suggests that children from diverse linguistic backgrounds, particularly with students who are not proficient in English, can benefit from peer learning structures (Cartledge & Kourea, 2008). Explanations in the literature cite relevancy to children's expectations, natural perspectives, aspirations and development levels (Diaz-Rico et al., 2010, Gollnick & Chinn, 2009). However, studies have warned of the laissez faire nature of under planning and off-task instructions mentioning that carefully prepared and monitored groups are far more likely to succeed than those who don't follow a structured plan.

### **2.5.2 Social Interdependence Theory**

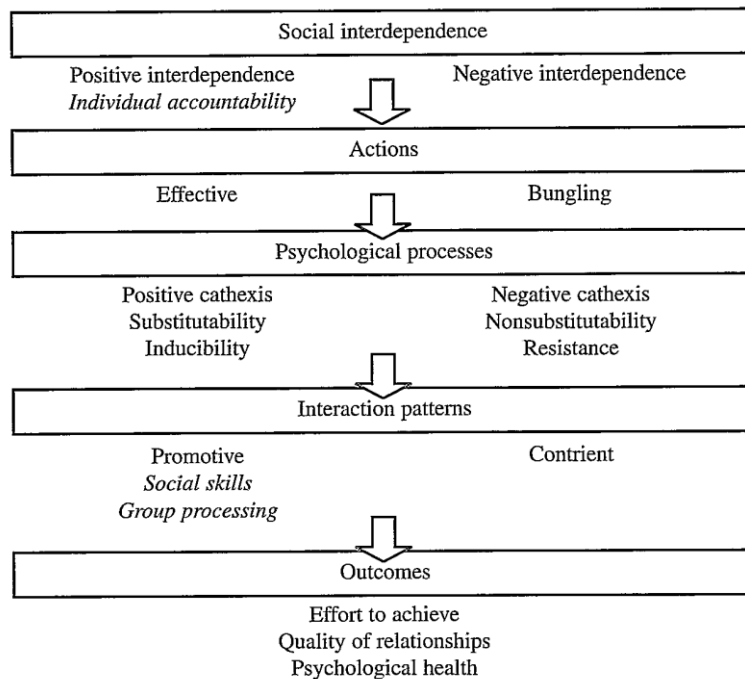
Historical roots of interdependence amongst learners began in the early 1900s emerging from Gestalt psychology that gave importance to *field* theories rather than mechanical instructions, thus denoting the shift from physics-orientated inquiry (Deutsch, 1962). The distinction of this change relates to the methodological emergence of *field* analysis in physics, to the gestalt (or whole), which became the preferred study method of perception and behavioural psychology during the 1900s. Gestalt psychologists like Koffka (1935) believed that the symmetry between psychological fields and group dynamics rested on the decisions (or recognition output) of wholes. The premise of this seminal work rested on human subjects primarily concerned with the organisation of perceived wholes and finalised meanings derived from their experienced realities, rather than a disorganised fractious property that may act as a summation of different parts (ibid, 1953, Deutsch, 1962, Deutsch & Krauss, 1965). The widely cited dictum of Gestalt psychology and other sciences where it has been studied and subsequently established is presented as “*the whole is greater than the sum of its parts*” (see Lewin, 1938, p211, Rock & Palmer, 1990, p84,

Van Boven & Epley, 2002, p263,). Thus, indicating that from a psychological perspective human subjects will perceive their realities as a whole and not broken down into parts.

Kurt Lewin (1935, 1948) and later Morten Deutsch (1949, 1962) both proposed that groups were like dynamic wholes that enabled interdependence to vary, thus positing the function of individuals with the need for common core goals (Deutsch and Krauss, 1965). Lewin (1935, 1948) was the first to develop Koffka's ideas of interdependence in a group. He proposed that humans would react to different situations based on a variety of perceptions. Thus a human's living space is moving fast and has dynamic properties that respond to the changes thrown in its path. Lewin (1935, 1948) also proposed that the level of "tension" amongst members was key to establishing the motivation trends amongst individuals.

Later, Deutsch (1949, 1962) expanded upon the behaviours and motivations of Lewin's group dynamic where he explored the interrelationships amongst individuals, in particular, highlighting the varying degrees of tensions and their significance on achieving interdependence. From this work, Deutsch produced two types of continua that relate to goals and actions of interdependence. Figure 1 illustrates these types as comparisons identifying the step process of positive and negative interdependence. The former, positive (cooperation) is formed as a correlation among individuals' goals and the chances of individuals attaining them are better if positive responses are applied by each person in the group. Johnson & Johnson (1974, p214) define this process as "cooperative goal structures" and suggest "... where the goals of the separate individuals are so linked together that there is a positive correlation between their goal attainments". Under the umbrella of positivity, therefore, individuals who are cooperatively linked must navigate and set reciprocal rules to achieve the goal of positive attainment. Deutsch's theory (1949, 1962) argues that for this to happen an *adequate* response is required amongst individuals' goals for a chance of attaining them. Conversely, negative (competition) interdependence is the polar opposite. It occurs when individual actions form negative correlations amongst groups members. Thus individuals perceive that they can only reach their goals by their cooperatively linked team members failing to reach theirs. An example of this may include negatively charged behavioural traits and possible tactics to put others off task, such as discouragement or obstruction. For Deutsch's (1949, 1962) theory and again highlighting the pathway found in Figure 1, it is observed that *bungling* is acknowledged as a proponent of negative interdependence.

Figure 2 Social Interdependence Theory



(cited in Johnson & Johnson, 2005, p289)

To briefly summarise, Deutsch's theory rests on two forces (positive and negative) that are found at the opposite end of the continuum. Both are exclusive to the paths they follow which lead to actions that are either effective or bungling (according to Deutsch's theory). These actions are again exclusive but lead to examine exactly how they react to three psychological processes (see Figure 1- substitutability, inducibility, and cathexis) and their interaction outcomes amongst individuals.

Substitutability, according to Deutsch (1949) relates to the actions of an individual that substitutes for the other. However, this will only work when the goals are evenly distributed and replaced for common purposes and decisive actions. A good example includes a group of relay racers who have different lengths to run. The goal is of course to win the race, but one runner may be slower in a particular section; thus other team members will need to run faster to substitute evenly for achieving the overall objective of the first position. Deutsch (1962) therefore argues that in cooperative group dynamics, individuals who are ineffective will not substitute for one's action or workload. Or to relate back to the relay example, racers cannot transfer their weakness to others racers (on the same team) without avoiding

a negative result. It should be noted that the tension levels in this dynamic are negatively charged and this falls into a resistant congruent. However, substitution can crossover and effect other cooperatives, such as if in the relay race example the competitor is running slower; thus a positive outcome for the other team as less effort is perceived by the group.

The investment in positive/negative and emotional attitude responses within a dynamic is related to what Deutsch calls a cathexis. This relates to the psychological energy expended on certain phenomena outside oneself (Deutsch, 1949). It is attitude formation and thus it can take many shapes and forms in different situations. The positive cathexis will help to support, encourage and develop those within a cooperative dynamic, whereas a negative will discourage and discharge values and opinions (ibid, 1949). People's goals and working together under the mantra of a 'one for all and all for one' dynamic will significantly increase the interdependence and strengthen common goal structures.

The final part of Deutsch's (1949, 1962) theory of social interdependence is the extent to which one person is open to influence and the practice of changing others. Like the other forces of cathexis and substitutability, the notion of inducibility is also positively and negatively charged. Respondents will induce each other for responses that are mutually positive for the benefit of the 'whole' task or not which will indicate negative inducibility. There is a growing literature on the subject of positive mutual influences in a cooperative dynamic against individualistic measures (Cronbag, 1966, Raven & Eachus, 1963, Johnson, et al., 1985, Johnson & Johnson, 2008). Further, linking studies have sought to find the effects of social relationships about behaviour amongst peers (Blumenfeld, 1992, Juvonen & Wentzel, 1996, Juvonen, et al., 2003) however, a known caveat of these studies reflects the when and why questions of intra- peer achievement. For example, we ask the questions- why is it that some positive peer relationships are established and achievement realised? The goal structures, cathexis, inducibility, and substitutability are positive, but when measured against a negative dynamic structure there are obvious differences. A better question will be, when does negativity enter the failing dynamic, and by what force? For educational interventions, how might we discourage or change it? Given the differences in group dynamics and external variables (such as personality traits), it is difficult to attribute one mechanism (as a failure amongst others) due to some disorderedly and unstoppable confusions amongst peers. Indeed, some studies offer possible answers related to behaviour and cite difficulties with multiple schooling environments. Particularly moving



between groups (Berndt et al., 2002, Berndt & Keefe, 1995, Ladd, 1990), anxiety, stress and depression (Brown, 1990, Wigfield, et al., 2006), and the changing nature of personality development and self-consciousness (Dryfoos, 1990). Some of these studies reflect the early adolescent years given the nature of physiological transitions and the notion of autonomy and self-independence at that time (Steinberg, 1990).

In summary, the success of interdependence rests on the outcomes of the whole group, so a change in a group member will effectively change the result of the entire sub-group (Johnson & Johnson, 1989). As discussed in this review the scope of negativity amongst peers is wide-open to interpretation based on physical, emotional and external factors. Therefore, should a change be made in a dynamic, this would affect the other members that were part of that group. Three structures exist in social interdependence theory- "positive," "negative" and "no interdependence" (Johnson & Johnson, 2008, p225). However, positive and negative interdependence must be distinguished from other similar phenomena. Take social dependence, where the association and links to interdependence may seem obvious. It is different as the goals and achievement factors of one person are affected by another's actions, but this process cannot be reversed (Johnson & Johnson, 2006). It might be said that this is a one-way realisation and one that cannot be reciprocated. No interdependence (individualistic goal structures) can be explained as when individuals attain or reach their goals regardless of others reaching or not reaching theirs (ibid, 2006). It is, therefore, a single process that is formed without others attention or actions for goal attainment. Personal factors and self-interest stimuli explain why this particular action is used. Its narrow form tends to lead individuals to *“seek outcomes that are personally beneficial without concern for others efforts to achieve their goals”* (Johnson & Johnson, 2008, p225). The foundation of social interdependence theory rests on how well individuals react with each other in cooperative structures; theoretical development has measured these outcomes as positive or negative intra-peer relations on achievement. The positive outcomes attributed to success from the studies mentioned seem to rest on altruistic and focused engagement within groups. However, external variables may impact overall outcomes. Conversely, negative forces tend to lean on competition and oppositional interaction. Finally, individualistic routes are measured as a force of indifference between groups.

The next section of this part of the review will look at the numerous studies that have contributed to the association of cooperative, individualistic and competition on

achievement factors. It strives to focus on the balanced context of cooperative interdependence that generates an understanding of what, in particular, works best for learning.

## **2.6 Peer Learning Methods**

### **2.6.1 Introduction**

This section begins with the identification of the various models of peer learning that have been identified within the literature. It bridges and thus furthers the conceptual frameworks that are involved with educational practice and pedagogy. Furthermore, it extends and builds on the previous section of theoretical development in the context of peer learning.

As previously stated, the theoretical blueprint for peer learning in relation to cognitive development and whole learning attainment are widely cited in the literature (Bull & McCalla, 2002, Buchs et al, 2004, Sharpley & Sharpley, 1981, Schunk & Zimmerman, 1994, Fitz-Gibbon, 1988, Nixon & Topping, 2000, Bruce, 1986, Mathes et al, 1988, Fuchs et al, 1987). Topping (2005) acknowledges the various trends and associations that have made peer learning one of the most researched areas in current educational practice. In particular, peer-cooperation, peer-tutoring, and peer-collaboration are suitably fashioned in the literature. Often advocating gains in the promotion of cognitive learning, scaffolding (Vygotsky, 1978, Van Zundert, 2010, Van Steendam, 2010, Gielen et al., 2010,) reciprocal talk and contextual meaning (Randi et al., 2010). All of which demonstrates mutually beneficial goals for small and large scale learning programmes (Bulotsky-Shearer et al., 2011, Siraj-Blatchford, 2009). Peer learning itself is used in many different contexts; it is essentially a generic term for a catalogue of methods that are appropriately chosen, given the tasks and goals of learning situations. However, there are some confusions and overlap regarding terminology in articles that advocate and use 'tutoring' and 'mentoring' interchangeably, with studies labeling incorrectly or changing study design half way through a learning programme (Topping, 2005, Dearden, 1998, Powell, 1997). The latter is best defined as a linear approach to higher learning, whereby a learner who purposely seeks skills or knowledge is guided by a person who knows capacity through a vis-a-vis design (Topping, 2005). Moreover, there is recent evidence to suggest peer mentoring is more effective as an intervention for behavioural and health strategies, custodial and pastoral programmes for primary/secondary school students rather than peer tutoring,

cooperation or collaboration (see Elledge, et al., 2010, Coppock, 2011, Mclean, 2004, Cowie, 2008). In summary, it can be accepted that the range and effectiveness of mentorship interventions are cross-age relevant and transient within a multidisciplinary network of agencies.

### **2.6.2 Peer Tutoring**

Peer tutoring, on the other hand, is somewhat different. Studies have reported benefits when applying a tutor-tutee scenario that enables both participants to swap roles, which is often mediated and controlled by a specific learning task (Greenwood, 1997, Miller et al., 1993, Fuchs et al., 1997). This design usually permits a teacher to expand on curricula content or provide remedial support for those who are behind. Topping (2005) provides guidance on this and suggests that content and learning goals must have clear objectives for them to work. A good example, regarding arrangement and goal setting, is provided in Hofstadter-Duke and Daly's (2011) study that examined the effects of peer-mediated reading fluency in the classroom. The class teacher trained three children who were above the reading frequency speed as per the test and study requirements. The three peer tutors would then coach a fellow student who was purposely identified as being below reading fluency frequency. Tests were based on reading word frequency speeds and examined for effects, which indicated a greater performance in reading while under the guidance of the peer tutors. However, the results are limited and are not generalizable, mainly because only one student took part who fitted a below average fluent reader. Regarding pedagogical arrangement, more students would need to be tested and examined to enhance validity and make accurate inferences for curricula implementation. The method showcases a typical peer-tutored study that fulfills the design of curricula matching, peer engagement with evaluation and diagnosis.

### **2.6.3 Peer Collaborative Learning**

Secondly, in support of the branches of peer learning, there is also a significant amount of literature that identifies collaborative learning as a relative extension to peer tutoring. Collaborative learning in its broadest terms is the coming together of large or small groups with a specific task and thus studied en masse or divided and shared amongst its members (Damon & Phelps, 1989, Mercer & Littlejohn, 2007). It is therefore dissimilar from peer tutoring, drawing upon the input and organisational structure that includes multiple voices and is not limited to various tasks. Also, as discussed earlier regarding constructivist and

cognitive approaches, peer collaboration is matched with Piagetian concepts with regards to interaction and progression, rather than Vygotsky's ZPD and social scaffolding. This is because children who form and work in collaborative peer environments are much more likely to be similarly placed regarding their natural understanding of the academic task at hand, and have similar competence levels. Indeed, for collaborative learning environments to be effective, it is premised that there is already a degree of mutual understanding and shared values towards a particular goal or learning objective (Teasley and Roschelle, 1993). Topping (2005) suggests that in some cases prior learning or indicative concepts must be grasped to excel in the chosen format. Furthermore, Totten et al. (1991) reported on the shared learning experience that if completed in the right circumstance it accelerates critical thinking and interdependence amongst its learners. Conversely, there are known associated problems and these are widely cited in the literature as showing mixed academic results (see Beers et al., 2005, Hallet & Cummings, 1997). One in particular relates to crowding or moving away from the associated learning goals without justification or reason (Kirschner, 2009). For example, Dillenbourg (2002) argues the identification of group roles and emphasises the power and commands of multiple voices that if not controlled or mediated may lead to chaos and confusion. This equalises the findings of Latané et al (1979) who reported observations in unstructured groups, where members revert to social relaxation or conflict when accessing learning tasks that are misguided or have limited structure. Others also say on behaviour as a caveat of learning attainment, which reduces the likelihood of goals being reached and in particular having an adverse effect on relationships and collaborative learning experiences (see Morgan et al., 2000, Sharan et al.1999). It can be therefore acknowledged that the success of peer collaboration is intrinsically linked to the organisation variables, particularly emphasising the direction of talk and communication. A summary that suits the narrative of successful peer collaboration is from Bruffee (1984, p652) who suggests “*organizing collaborative learning effectively requires doing more than throwing students together with their peers with little or no guidance or preparation*”. Peer collaboration must be mediated and rules should be applied in advance for learners to receive passive guidance to keep them engaged with content.

#### **2.6.4 Peer Cooperative Learning**

Working together to achieve a mutual set of goals is an age-old method. Its multi-usage in educational and non-educational environments has added to its popularity as a method for learning and attainment. It can be implicitly and explicitly controlled depending on task outcomes and the objectives that are arranged before tasks are completed. Therefore, for the benefit of this thesis we consider the pre-determined educational consequences, thus the explicit notion of a defined set of goals that are mutually agreeable and pursued. First, this section will review the practical elements of how cooperation is instructed and used for educational objectives. Second, the theoretical basis of how specific methods were developed, locating the research that engineered practice-based solutions. Thirdly, a look at the future and how technologies have been created utilising elements of peer cooperation methods.

Cooperative learning goes beyond usual group work that sometimes it gets mistaken for. In fact, they're many studies found in the literature that mix-up many different peer learning techniques, thus making it difficult to identify the links between their various methods and their success rates (Topping, 2005). Cooperation is studied at different levels and for different reasons (for example educational, behavioural and cognitive) therefore the critical difference against traditional group work formation in the classroom is the maturation of carefully planned ideas that are continually monitored in groups (Jacobs, 1996). In its broadest sense, the planning aspect would involve a teacher or learning mentor presenting an objective for learning to the whole class with detailed instructions. Following this, groups are formed and roles are identified readily for the task, with micro functions defined and in some cases unique assignments depending on the objective. But central to the idea of cooperation is the way individuals interact, thus placing stress on environmental factors and opportunities for learning formation reciprocity (Adams & Hamm, 1990, Kagan, 1994). Jacobs (1996) indicates that cooperation becomes more of a theme that reflects the engagement of social and academic outcomes, as participants are not thinking independently allowing for openness and guided transitions.

Thirdly, peer cooperation requires groups of children to work together and in most cases share responsibility for learning attainment, thus indicating that everyone learns together at the same speed with often the equal outcomes (Johnson & Johnson, 2002). Slavin (1995)

provides four theoretical perspectives that were broadly covered with relation to the theoretical development in the previous chapter. However, it is worth revisiting his third framework of cohesiveness that suggests children are best suited to working with each other because there is an operant mutual care system that reflects similarity in goal setting and learning achievement. Indeed, like peer collaboration, there is evidence in the literature that points to the sharing and development of ideas, which learners can construct and build upon (Webb & Mastergeorge, 2003, Sharan, 2010). Further evidence is measured through attitudes to learning while engaging in cooperative tasks. Research suggests that this improves when children are working cooperatively on shared responsibilities, and this is particularly significant when compared to individual or competitive methods of instruction (Johnson and Johnson, 2005). Other empirical studies report similar findings involving cooperative instruction that measure the improvement of individual self-esteem when shared tasks are completed (Kilic, 2008, Bertucci, et al., 2010). These results demonstrate the capacity cooperative learning has in advancing attitudes and engagement in learning content and accelerating individual achievement through team goals.

There is some evidence that children develop a critical discourse that goes beyond usual teaching transmissions and individualistic learning (Gillies, 2006). Of particular note, in cooperative groups children will negotiate with an informative discussion that is uniquely channelled between each conversation. Children devise their language and identify patterns that can form specialist reasoning, and create pedagogy to further a learning pattern or complete a learning task (ibid, 2006). However, one of the difficulties with cooperative learning is the various forms in which it takes, and the number of alternative programmes that have been designed from it. Peer cooperative learning has shown favourable results in teaching and learning situations and the literature acknowledges, through structured scenarios, it can be a method to accelerate learning experiences, self-esteem and achievement. However, while it is seen favourably in Western contexts, there is a limited literature that advocates its usage and availability in the developing world. Moreover, what is available suggests cooperative learning to have little effect on learning and achievement. One study conducted in Pakistan (Parveen et al, 2011) found no difference in learning performance in 8<sup>th</sup>-grade social science classes when comparing cooperative learning groups to regular instruction groups. Previously, Dean (1995) found similar results when using a cooperative (learning together) model as the primary method of instruction. The results indicated problems with roles and responsibilities in groups. Teachers and learners

reported the technique as being peculiar and unrelated to their learning context. For teachers, the method challenged the belief and attitudes of teaching itself. Some studies point to the cultural distinction in which knowledge is passively consumed through rote or one-way transmission mechanisms, where learners are consumers and teachers are the decision makers (Ghorbani, 2012). A meta-analysis by Thanh et al. (2009) reviewed cooperative methods in East Asia, and it reported over 50% of studies included in their review preferred specific teaching methods compared to cooperative instructions. Other factors include teacher authority and control on learning and assessment tasks (Tan et al, 2007), facilitating and mediating the cooperative process of group discussion (Eva, 2003, Sachs, et al., 2003) and methods that endorse memorisation (Messier, 2003). Although some of these studies were conducted in the international context, some mirror the Western education systems with resource-rich classrooms and available instructors. For some developing countries, this remains a challenge. To date, no research has aimed at using cooperative methods in non-formal learning 'development' contexts. It's perhaps unsurprising given the use of rote methods and difficulties in rural education planning.

#### **2.6.5 Peer Cooperative Learning and L2 Reading**

As we have seen in previous sections, there is an understanding of L2 reading failure when the L1 models have not been properly developed. One way in which to bridge the gap of struggling readers is to use cooperative learning strategies. Research in the L2 reading classroom has a major impact on the cognitive and linguistic processing of learners in this context (Kagan et al, 1995, McGroarty, 1989). Theoretical work has also been conducted which concentrates on the interaction element amongst learners. As Olsen and Kagan (1992) highlight, cooperative strategies present learners' with options in reshaping knowledge, expanding ideas and evaluating what is most meaningful in learning contexts. With specific relationship to reading; interaction can lead learners' to higher cognitive tasks and ultimately "*promote discussions of readings and to work with information from the readings, exploring different solutions for complex activities*" (Grabe, 2009, p396). As with the previously reviewed cooperative learning models, there is evidence of their suitability in improving reading attitudes, motivation and academic achievement. This is evaluated as being positively connected in the construction of activities, from a teacher-pedagogical perspective (Gwyn-Pacquette & Tochon, 2003, Horwitz, et al., 1997) and a learner perspective (Jacobs et al, 2002).

While there have been considerable gains in the promotion of cooperative learning models in education contexts, there remains an opportunity to understand more about how it works for improving L2 reading achievement. One study by Klinger and Vaughn (2000) investigated how cooperative reading in science can enable less-proficient peers to accelerate their knowledge and application of English. They report increases proficiency in vocabulary, understanding specific terms and generating meaningful conversations in the target language. Similarly, Jacob et al, (1996) reported benefits to reading where students were able to ask questions about the complex language found in their reading materials. More specifically for this study, Lan et al, (2006) found interaction benefits in peer-assisted learning techniques where students were operating in small reading groups. They reported an enhanced focus on tasks including phonological processing and general peer-assisted learning discussions. In their study, the use of a computer software programme was used for mediation, providing a place of reciprocal mentorship and learning. In an EFL study with Lebanese students, cooperative learning instructions increased reading achievement over a short period, however, in the same study measures between before and after attitudes to learning English reading remained the same (Ghaith, 2003). However, Liang (2002) reported benefits to attitudes and learning after reviewing five cooperative methods in Taiwanese junior school children. The study compared the various cooperative methods to the audio-lingual method and found significant differences (and gains) with the experimental cooperative groups. It reported that cooperative methods enabled communicative interactions that facilitated positive attitudes to learning English, compared to traditional methods.

Given the limited literature, the present study aims to develop a further understanding of how cooperative methods can be utilised in non-formal L2 reading situations. The next section reviews a variety of cooperative teaching strategies and methods.

## **2.7 Peer Cooperative Teaching & Learning Strategies**

### **2.7.1 Introduction**

In this section, a number of teaching and learning strategies have been selected for review by their value in advancing knowledge on cooperative structures. Where possible, procedures have been selected based on a methodological preference that can be used for reading strategies and to scaffold positive interdependence amongst learners.



### **2.7.2 Constructive Controversy**

The literature indicates many programmes that have evolved from the generic rules of peer cooperation. One of them is '*constructive controversy*' that enables children to formulate opinions and express differential views before making a judgement, and thus agree on an outcome (Johnson & Johnson, 2007, 2009). Topics are considered beforehand that induce sensitive responses or potential clashes of opinions to reach targeted objectives (ibid, 2009, 2007). The literature points to 'concurrency seeking' that helps teachers to mediate the direction of how much conflict can be explored, and the role of intervention at particular junctures to keep groups from moving away from the learning objective (see Smith et al., 1981, Johnson & Johnson, 2009). In the context of reading, children may have different views when working in pairs or groups. The constructive controversy enables groups of learners to have a healthy discussion on reading tasks or some other activity, so long as it's mediated through rational decision-making and inclusion of multiple views (ibid, 2007). On evaluating its worth in the present study, the process of using CI risks potential harm given the number of learners the study incorporates. It would require further human resources to oversee this successfully.

### **2.7.3 Teams Assisted Individualisation (TAI)**

Similar to CI, Teams Assisted Individualisation (TAI) was originally introduced to enable children with special educational needs a chance to develop mathematical concepts from their peers in mainstream settings (see Slavin et al., 1984). It has also featured in the higher primary years for all ability children with positive effects on motivation and task competence (Slavin et al., 1996). Teachers begin with setting a task or learning goal, specific to mathematics, and children are initially tested and then grouped based on their results. Once the groups form, all members are encouraged to problem solve before receiving another test, which is completed individually. Research on TAI, like other cooperative methods, has had positive effects on behaviour, perceptions, and productivity while on task (Slavin, 1985). The experimental design incorporated into the present study is fitting with the empirical nature that TAI follows. Involving data to arrange certain reading groups is one way in which to group heterogeneous learners' into manageable learning groups followed by further adjustments. However, the method would require numerous testing phases and group tracking to follow learners' from one group to the next.

#### **2.7.4 Student Teams Achievement Divisions (STAD)**

In bridging the inclusivity of cooperative peer methods, 'Student Teams Achievement Divisions (STAD) (Slavin, 1978) disregards grouping on academic performance and children are mixed when the task begins, usually 4-5 in a group. A lesson is first presented with objectives and puzzles are set. In some cases, there is a time limit (ibid, 1978, 1986) for tasks to be completed and these are often to induce the next task or are needed for progression. Similar to the monitorial system, all children must master the task and help others until they've reached the intended goals. Again, with relation to the previous chapter on underpinning theories, the premise of this arrangement lends itself to behaviourist methods, as it relies on reinforcement and praise from team members. However, the dynamic changes as the children are then tested individually based on what was discussed in each group. Further rewards are issued from the teacher who tallies individual scores and ranks them by all groups. For peer reading groups, this method has implications for minor rewards when achievement takes place and major rewards when tasks have been completed. One of the concerns in the present study is how this will be monitored and delivered equally. Achievement goals would need to be arranged before group assembly.

#### **2.7.5 Jigsaw**

Some methods require an indicative dynamic of ability before formation. The '*Jigsaw*' method introduced by (Aronson et al, 2008) is more specific on group allocation, allowing a maximum of six members per group who are tasked with understanding and reporting in layered academic material, in a specific core area. For example, a group might be tasked with understanding the history of the Roman Empire, with each group member designated to a particular section (i.e., history, main personalities, significant contributions, representations in modern culture). It is a two-step process that requires small groups acting in groups to accomplish the task set. Therefore, after individuals complete their section, the next step is to discuss their findings with a person who is also reading that section before reporting back to the group as a whole. It is from this meeting that all members are encouraged to listen to their peers and recognise the key details from their findings. A similar method, Group Investigation (Sharan & Sharan, 1992) builds on the Jigsaw by leaving group assembly up to the children with whom they form a group. One of its benefits is that it reaches all abilities. Further similarities are with the breaking up of sub-topics that are allocated within a formed group. However, this is not necessarily directed, as the children are free to choose which topics they want from a broader area of study that is often

projected from a school blackboard or whiteboard (Sharan, 2010). Following whole group discussions, the team is then instructed to give a presentation on what they have found. This method is quasi-self organised as the children are researching and thus interpreting what they believe to be a credible topic, worthy of fulfilling the wider objective. However, without the right support systems, there is a chance that members may not be researching or are possibly misguided in what they search for.

### **2.7.6 Complex Instruction (CI)**

Complex Instruction (Cohen, 1994) has the opportunity to develop this learning further as it is specifically targeted for all academic abilities, assuming the teacher or lead can identify the support required. Children form their own groups and are required to report on a particular theme, similar to what is found in the jigsaw model. Cohen & Lotan (1997) argue that successful transition in learning is maintained at group level before own understanding can solely take place, and this relates to the central theme that the group is investigating. Also, groups that may have children with low academic ability or who are failing to cooperate are targeted to participate, but this is reliant on the curriculum that is fostered to induce participation (ibid, 1994, 1997). Children with a low academic ability and who are not interdependent may find curricula topics challenging and unrecognisable. A parallel here is with the *constructive-controversy* method, as particular areas may lead to isolation. Teachers foster group reunification by introducing simpler tasks that are inclusive, to retrieve any member who is disengaged. From this, a teacher will '*assign competence*' to children who demonstrate an eagerness to participate by reinforcing the points they raise that might lead to a new conversation direction (ibid, 1994, p23). There are other support devices and interconnectedness that teachers and student themselves use to aid discourse and knowledge trajectories. However complex instruction is born out some assumptions, firstly that children who have low ability status in groups will contribute and secondly, that the teacher will recognise the individual contribution and move this forward with reinforcement that links to whole group objectives. The dynamic requires multiple groups studying various areas and without sufficient support systems, this may devise a competence hierarchy that is selective and that detaches or breaks away from individuals, thus leaving low ability learners in further isolation and off task. There are also issues with the classroom layout, often resulting in children who are situated in small groups but are not interacting because of the content and laissez-faire approach to constructing cooperative situations (Baines et al., 2003). In these situations children often result in switching back

to individualised learning or traditional patterns and for those who are in need of extra help, turn for adult guidance (ibid, 2003). Again, the literature cites organisation foci as the main caveat for the direct successful implementation of cooperation and learning success. It is therefore paramount that guidance and support is on hand to help with transmissions, something that is not necessarily achievable in a busy networking session.

### **2.7.7 Teams-Games-Tournaments (TGT)**

The literature also acknowledges competition amongst peer cooperation and the value of peripheral learning goals ad hoc to structured core areas. Indeed, 'Teams-Games-Tournaments' (TGT) extends the STAD method by replacing regular quizzes given to each group by introducing weekly tournaments, where each group is encouraged to play academic themed games for points for their team (see DeVries, et al, 1975). Students are paired against each other (from different groups) on ability, giving all children the chance to compete. In some cases children are able to prepare, but can only do this amongst the groups they are in to allow the competition to be fair. DeVries et al (1975) found positive group cooperation before task and an increase in mathematical ability for group members. Games were arranged based on numeracy tasks and a random control trial was set up, testing for individual-group effectiveness. In summary TGT is a suitable addition to create competition in the classroom and there is evidence from DeVries' findings that suggests weaker students show greater concentration and interest when on task. TGT is an incentivised method that invites groups of participants to challenge each other, something that Slavin (1987) notes as key to any peer cooperative network.

### **2.7.8 Peer-Led Team Learning (PLTL) & Learning Together (LT)**

Like some of the other cooperative methods, the origins of the Peer-Led Team Learning method has been mainly researched and developed in Higher Education. However, the key pedagogical arrangements allow it to be utilised in any teaching-learning situation (Gosser, 2001). It involves a trained mentor who has a particular skill that learners in h/er group can access through immersive activities (ibid, 2001). As Varna-Nelson and Coppola (2004) acknowledge, the PLTL method incorporates four areas of theoretical and instructional design; socio-cultural (Vygotskian approaches), group learning, reciprocal teaching/learning and holistic instruction. Typically, groups will be established with one trained or knowledgeable peer who provides direction and teaching for the group. The method has garnered favourable results in various studies; particularly partnership for

learning and motivation for learning itself. Rogers (2003) indicates the similar backgrounds both learners and mentors have in learning situations helps to unravel complex tasks. There is an unconscious admiration of those who have completed a specific learning task ahead of their peers. Similarly, the Learning Together method brings together the cooperative, individualistic and competitive aspects in-group learning (Johnson and Johnson, 1999). The significant factor here is to do with reinforcement and praise of tasks completed. Mentors can build on work and move learning forward by difficulty and groups can discuss new problems with relation to specific targets. Another advantage also relates to classroom management, which enables the teacher to ability monitor groups that are perhaps at different levels in the classroom.

### **2.7.9 Cooperative Integrated Reading and Composition (CIRC)**

As the name suggests, this cooperative method is used to teach reading and composition and has benefits for beginning writers (Durukan, 2011). There are three stages to CIRC where cooperative tasks are embedded at each step. The first stage involves story-related activities, where mentors introduce new vocabulary and read passages from a book (Stevens & Slavin, 1995). Learners are required to work in pairs on tasks associated with the story or new words given to them. During the next phase, learners are assigned comprehension tasks and challenged by their mentors on how much they know from selected passages of text (ibid, 1995). This is managed by slowly reducing the support each pair has on-task. Gradually, each pair is left to finish the task without help from their mentor or class teacher. The final stage incorporates knowledge within a broader language arts curriculum (Bridge & Hiebert, 1985) with the focus on writing development. Regarding reading development, the method allows for learners' the opportunity to read aloud to their co-learner and mentor. Slavin (1995) identifies the method to be reciprocally beneficial to learner and mentor, especially when constructing meaning. He states that this is what accelerates the comprehension of new text and fluency. The language arts stage is then used to explore writing abilities built from the previous discussions.

#### **2.7.9.1 Summary**

The methods outlined in this section have broadly put forward the significance of competition, cooperation and individualisation in furthering learning and achievement. Some have been developed to tackle specific problems while others have been constructed to translate broadly into teaching and learning situations. The next section discusses three major meta-analytical reviews conducted in the peer-cooperative learning field of inquiry.

## 2.8 Peer Cooperative Reading Meta-Analytic Reviews

### 2.8.1 Introduction

Major literature reviews and meta-analytical searches in the domain of cooperative learning have provided various contributions to knowledge on the effectiveness of this peer learning technique (Johnson & Johnson, 2000). However, only until recently have these analytic reviews considered effectiveness concerning specific cooperative methods. Historical analyses mainly focused on ability grouping learners (Slavin, 1987), special education (Elbaum et al, 1999) and literacy tasks (Kulik & Kulik, 1987) to further knowledge related to cooperative and collaborative learning effects. But, as documented in earlier sections, many peer methods have the capacity to improve general learning and more specifically reading instruction.

The analytic reviews included in this section are subject to inclusion through their impact in furthering understanding of peer cooperative learning methods to improve reading outcomes.

### 2.8.2 Johnson and Johnson (2000)

In this review, a total of 158 studies met the study inclusion criteria with a date range of 1970 to 1999. It was the first to understand 'cooperative effectiveness,' building on earlier reviews targeting individual cooperative models and ability grouping in the classroom (see Cohen & Lotan, 1997, Slavin, 1991).

The review included ten cooperative methods used across a range of educative levels – ranging from primary to adult education. In particular, cooperative learning methods were compared against individual and competition methods, these being typically used in classrooms. The primary aim of the review was to understand more about cooperative effectiveness, in particular, which method returned more significant achievement in teaching and learning situations. From this primary aim, the following criteria were used for investigation;

- Analyse the frequency of cooperative studies to evaluate specific cooperative methods
- Evaluation of previous cooperative methods
- To understand the particular techniques that accelerate achievement

- To understand the characteristics of the more useful cooperative method

Regarding characteristics, the review found most studies were conducted in primary schools, with over 50% of them lasting 60 mins or less between a 2-29 week period. Most studies were held in the West. However, four studies from Asia made the review. The review found significant effects for the Learning Together method when compared to competition (+0.82) and individualistic (+1.03) methods. Out of all of the cooperative methods, this had the most significant effect on achievement. The Constructive Controversy (CI) was ranked second and had large effect sizes when compared to competition (+0.67) and individualistic (+0.91) methods. The other methods also had significant effects, and these consisted of Teams-Games-Tournaments (+0.48, competition) and (+0.58, individualistic), Students-Teams-Achievement-Divisions (+0.51, competition) and (+0.62, individualistic). Lower effects (effects, nevertheless) were observed in Team Assisted Individualisation (+0.25, competition) and (+0.33, individualistic) and Cooperative Integrated Reading Composition (+0.18, competition) and (+0.18, individualistic). As suggested by these results, some had more substantial effects when comparing cooperative method type and cooperative strategy (i.e., competition, individualistic). Nevertheless, all had some impact on improving achievement.

If we take these scores alone, it's noted that cooperative methods have a more significant effect on achievement when compared to individual methods. As discussed throughout the earlier sections of this review and in Chapter 1, the prevalent pedagogies in Indian primary education rest on 'rote' and individual methods of knowledge transfer. The review contributes further to the idea that cooperative structures may have some role to play in the development of reading instruction. However, the review included various learning targets for different learning contexts, in other words, reading and comprehension were part of many different achievement conditions as set by the studies included. Therefore, it's mainly challenging to understand what effects, if any, there are on reading ability, proficiency, decoding and sound awareness alone. Moreover, the literature cites the need for more robust searches to highlight specific levels in education, since adults (operating in adult education) may have developed cooperative skills that primary children are yet to establish (Kutnick et al., 2006). This has significant implications for children in the Indian context who are mainly exposed to unique methods and memorisation tasks.

The next review, in this section, is conducted by Slavin, et al., (2009) who goes further by focusing on reading programmes that have utilised cooperative peer methods.

### **2.8.2 Slavin, et al., (2009b)**

In this review, the primary aim was to understand the effectiveness of reading programmes for children in the lower and middle school age range. It reviewed 63 studies that researched early reading programmes and 79 studies in the upper elementary year groups. They discussed various types of reading programmes including cooperative reading immersion programmes, like Peer Assisted Learning (PALS) and more specifically to the current study, Cooperative Integrated Reading and Composition (CIRC) methods. The strengths of the review were to inform teachers, reading instructors and mediators with efficient data to tell what works. It brought together some reading programmes and analysed their efficiency to enable educators to make better decisions when selecting a reading programme. However, given the range of reading programmes available to elementary practitioners, their use over time, practical issues, and diversity in skill approach (i.e., some studies targeted phonemic awareness, others vocabulary and word attack), it's difficult to accept the universal nature of the findings. In this regard, there are methodological limitations to the overall results, due to the lack of like-for-like comparisons the review couldn't control for. Nevertheless, the findings made similar projections in previous reviews and built on the positive effects of cooperative strategies. In this review, cooperative designs are considered instructional, as they're usually packaged for teachers to use and implement in class. Across the 17 studies (all used experimental designs) the results show effects on PAL and Peer Tutoring (+0.46). Further empirical evidence on the benefits of cooperative strategies is also reported in the upper elementary range. In particular, ten studies were included and found positive effects (+0.29) in approaches using the CIRC method. Interestingly, Slavin et al., (2009a) also reviewed cross-age and same-age peer tutoring, both resulting in the same positive effect on achievement (+0.29). They suggest that providing structure for student interaction mutually benefits metacognitive skills required for developing readers. These findings reinforce the nature of cooperative arrangements and their potential effects on learning to read. The next review by Puzio and Colby (2013) marks the latest analytical synthesis of cooperative literacy studies.

### **2.8.3 Puzio & Colby (2013)**

This review consisted of studies that incorporated both cooperative and collaborative strategies that specifically targeted achievement in reading. Out of the 30 studies included, 19 of them followed an experimental pathway and used standardised assessments on vocabulary and comprehension reading skills. These were analysed separately, however,



since it was challenging to separate collaborative and cooperative methods, general inferences were made across both methods. Indeed, as reported in other reviews here, it is difficult to control for practices in elementary education due to the range of activities different institutions and providers follow. However, the review adds further evidence to cooperative designs and reading research on the whole. It reported effect sizes on three reading outcomes; total reading (+0.16), vocabulary (+0.2) and comprehension (+22). Taken altogether, including other measures considering publication bias and weighed corrections, the review found that 90% of all effect sizes reported positively on cooperative and collaborative achievement in literacy interventions. Given the strength of these pedagogical techniques, they also positioned their findings to a similar review from Lou et al., (1996) who reviewed studies utilising cooperative and collaborative methods in mathematics.

Other findings relate to 'ability' grouping, where previous research has cited the procedure as harmful to lower 'ability' learners (Wheelock, 1994, Peterson, 1989) in group learning situations and promotes negative self-concepts of learning (Reutzel and Cooter, 1991). Cooperative methods, based on their structured nature, have fallen into this category where the focus is on ability rather than attainment. Cooperative groups require structure and planning before they are deployed in the classroom (Topping, 2005). Two reading programmes contributed to the overall effects of the review with the inclusion of Bramlett's (1994) study that utilised structured ability grouping through CIRC and another cooperative reading programme. Puzio & Colby (2013) note the contribution of Bramlett's (1994) study as an important finding in the current review. Bramlett (1994) specifically targeted structured cooperative methods to see what effect they would have on low, medium and high achievers in reading. The findings suggested that there were higher effects found in the lower ability group, thus, making the greatest gains in reading achievement. In summary, the review presents further evidence on the strengths cooperative and collaborative methods have in improving learning and achievement. This review is particularly useful to the main focus of this thesis given the gains made using these peer methods. It focused on reading attainment specifically with a focus on instructional methods with positive results. While this is encouraging, it didn't differentiate between cooperative or collaborative methods in their exhaustive review. This makes it difficult to see which method had the best effect and the authors acknowledge this as a limitation. However, their review found particular strengths with structured cooperative

ability matching in reading acceleration for lower ability children. Puzio and Colby (2013) note that ability grouping is intrinsically part of cooperative and collaborative learning groups since it's the diversity and abilities of learners' coming together that add to the learning process.

## **2.9 Theoretical Framework**

From the established theories explored in the cooperative and reading domains, this section provides details on the theoretical framework the present study employs. In this thesis, social interaction amongst peers is identified as a prerequisite for cognitive development in L2 reading (Donato and McCormick, 1994). This thesis incorporates the views closely associated with Donato (2000) who describes sociocultural processes involved in language learning as a socially mediated activity. It views the physical and mental tools (Wertsch, 1985) required for knowledge acquisition as being rooted in sociocultural learning experiences, highlighting the significant contribution of social actors operating in the wider social context.

This thesis considers the social domain and ZPD to be a cooperative process in which learning partnership and peer learning exchanges are enabled. Peer cooperative methods are rooted in social constructivism as they seek to provide opportunities for groups to work together to solve mutual goals but with interdependent gains. In this thesis, peer monitors are utilised as 'more knowledgeable others' trained in a synthetic phonic method before instructing their peers. The synthetic phonic method relies on systematic instruction to enable users quick access to sound proponents found in words. Phonemic awareness and pre-reading skills, through scaffolding from peers, will enable a cooperative practice to refine and develop individual development. In this regard, positive interdependence (Johnson, et al, 1998) is used as a framework to facilitate mutual goal setting of improving phonemic awareness, spelling and reading amongst developing L2 readers.

The theoretical justifications for peer cooperative learning and synthetic phonics instruction has given license to the development of a trial in the present study. The next section of this research presents how both methods were used in unison to target L2 developing readers in a rural Indian informal learning context.

## Chapter Three - Methodology

### 3.1 Introduction

This research is a timely response to the problems facing young children who are illiterate in rural India. The rationale for undertaking this work is focused on two areas. Firstly, it is my interest in exploring new pedagogies that have positive outcomes for children. Secondly, peer methods, as discussed, can improve learning opportunities for children who are not regularly exposed to English reading pedagogies. Alongside these aims, this thesis represents a piece of work that reflects the numerous calls from development and learned agencies for evidenced-based empirical research that aims to trial new initiatives that target illiterate groups (Banerjee and Duflo, 2011). Given these goals and the disappointing literature of child illiteracy and ineffective pedagogies, the adoption of a peer-mediated synthetic phonics intervention was selected with the aim of encouraging participation of hard to reach children from a rural town. In doing so, the readership is reminded of the study's primary research question followed by supporting questions;

- What are the effects of a peer-mediated phonics intervention when considering the phonemic awareness, spelling ability, reading skills, and reading attitudes of children from a rural North Indian community?

In support of the primary research question, this study employed an RCT format, targeting specific areas of reading achievement and attitudes that asked;

- Is there a statistical difference between male and female Monitors' baseline reading, phonological processing, decoding and native language test scores?
- What are the effects of Monitors' phonological processing, reading and spelling abilities when guiding learners in a peer-mediated intervention?
- Is there a statistical difference between male and female learners' baseline reading, phonological processing and native language test scores?
- Are there statistical differences between learners' reading, phonological processing, and native post-test scores when controlling for pretest scores? If so, where do these differences lie?
- Is there a statistically significant difference in baseline academic and recreational attitudes, when considering the average (neither satisfied nor dissatisfied) test scores?

- What are the effects of Monitors' recreational & academic reading attitudes, before and after a peer-mediated trial?
- Are there differences in Monitors post recreation, academic and total reading aspirations based on gender when controlling for pretest recreational, academic and overall composite scores?
- What are the main factors concerning Monitors' reading attitudes in academic and recreational settings?
- Are there differences in control and intervention group's recreational, academic and total English reading satisfaction scores?

This Chapter describes the rationale for choices of methodology and procedures involved in the collection of data from the research questions. It navigates through the preliminary choices for research design and discusses the steps taken in the development of the trial.

### **3.1.1 Philosophical Foundation**

For all research, it is essential to understand how different paradigms work as this assists the researcher in making methodological choices. Paradigms are sets of practices in which beliefs about something are guided. They are like lenses in which researchers use to make sense of the world (Greene & Caracelli 2003). Some define them as 'worldviews' to inform perspectives when underpinned by various philosophical assumptions (Teddlie & Tashakkori, 2009, Greene, 2007). Two major paradigms have been prevalent within the social sciences and educational research over recent decades. Both doctrines of positivism and constructivism (sometimes interpretivism) treat and diffuse knowledge differently, and it is within these domains that arguments have been vehemently debated. The 'paradigm wars' as they have come to be known, pursue epistemology, that is, 'what it means to know' (Guba & Lincoln, 1998, pg 108) through contrasting ontological viewpoints. Ontology itself asks 'what is the form and nature of reality?' (ibid, 1998, pg 108) Therefore, epistemological differences arise from holding different viewpoints about the significance of reality, or, what can be known. Researchers typically follow 'objective' or 'constructive' pathways in their pursuit of knowledge, these being navigated by a 'set of beliefs' to guide their actions (Guba 1990, in Creswell, 2009, pg 6). These beliefs are deeply rooted in the methodological approaches deployed in the interpretation of truth and knowledge of studied phenomena. There are multiple paradigmatic views for researchers' to follow based on their

ontological beliefs about how reality is constructed. A brief discussion of these views is presented here with the aim of highlighting the critical decisions made in the current study.

Positivism is realised through foundationalism, which developed from empiricist roots in the natural sciences (Robson, 2002). Social science research adopting this approach requires objectivity and knowledge creation to be independent of the researchers' experiences (ibid, 2002). Ontologically, the paradigmatic stance and following methods involve using quantitative procedures to explore relationships and effects of studied phenomena. The strength of such approaches follows strict procedural objectivity to allow for accurate inferences. Borrowed from the natural sciences, one way to interpret positivism is through hypothesis testing, using strict controls and procedures before offering accurate inferences. However, critics of positivism assert its limitations to the independent nature of how knowledge is constructed (Hustler, 2005). They agree that knowledge cannot be interdependent due, in part, to the complexities involved in controlling external forces (Silverman, 2000). Constructivist approaches directly oppose the positivist paradigm. Some argue that it goes beyond being a direct opponent to positivism by applying methods to understand the complexities involved in social behaviours at a personal, individual level (Guba and Lincoln, 1998). Methodically, the constructivist paradigm uses flexible qualitative approaches, such as interviews and observations in unravelling individual social differences (Patton, 2002). Its weakness is in its ability to generalise to a broader population and has problems in replication based on complex social situations, experiences, and events. Evidence-based research is particularly useful in gauging intervention effectiveness if controls are applied. It assumes members in a study may share similar experiences or traits, such as children grouped in a classroom or by examining socio-economic data (Shaw et al., 2006). The debate and subsequent paradigm wars are weakening (Mingers, 2003) with the emergence of alternative methods such as post-positivism and mix methods. This thesis acknowledges the differing approaches but follows an ontological pathway that suits the research questions and research design it employs. This is further discussed below.

### **3.1.2 Epistemological Framework**

This research has its roots within a pragmatic paradigm as it seeks to answer the 'what works?' questions. Broadly, pragmatism refers to practical approaches to solving specific problems and reflects numerous calls in International Development research for its inclusion (Banerjee & Duflo, 2011). It permits researchers to embrace certain methods

(whether qualitative or quantitative) and design instruments to answer targeted problems (Creswell, 2009). Theoretically, pragmatism combines subjectivism and positivism, but these are usually seen as competing paradigms. Thus, pragmatism is better understood as a methodological choice rather than an epistemological justification. It permits a merger of co-existence between methodological appropriateness and philosophical beliefs (Greene and Caracelli, 2003). This thesis did not initially start with a worldview or belief about how knowledge is constructed; instead, it was guided by the necessity to understand the best fit for intervention against the detailed research questions. The truths uncovered in this study, whether objective or subjective, manifest from a converging methodological paradigm that applies qualitative and quantitative tools. In this context, knowledge using multiple methods is seen as a meaningful process. Using a variety of methods, we can make further inferences about the collected data, something that, by function, unique methods would be limited in doing. Pragmatism is therefore flexible in its capacity to mix methods that are matched to relevant questions or inquiries. The combination of methods drives the methodological and epistemological direction of the study. Thus, ‘knowledge’ and ‘truths’ are driven by the available data where meaning can be explored concurrently, as Howe (1988, 1992) suggests, truths and knowledge are linked to outcomes rather than to paradigmatic debates. These choices to pursue meaning are therefore constructed by the knowledge that is based on what is working best. Specifically, evaluative research in education concentrates on the processes and outcomes of trials (Robson, 2002). In this study ‘processes’ are quantitative measures while ‘outcomes’ are treated as qualitative intrinsic values. Both measures are embedded within a mixed pragmatic framework. This is further discussed in the next section.

### **3.1.3 Mix Methods**

As it predictively suggests, ‘mix methods’ is the mixing of two or more methods that address a set of research questions in a field of inquiry. It supports the pragmatic paradigm in that research instruments can be utilised to best fit a specific problem (Johnson and Onwuegbuzie, 2007). In broader terms, it is the inclusion of quantitative and qualitative research methods that are applied to specific contexts aiding a greater understanding of a problem area. Education research has seen a rise in recent years with the number of studies incorporating mix methods into their study design (Creswell, 2002, Tashakkori and Teddlie, 2003). It is perhaps one of the reasons why some argue it to be the third

methodological choice (Cameron and Miller, 2007) when matched against traditional singular preferences. Moreover, some call for further adoption as it enables a better range of sources to give further meaning about the studied reality (Bryman, 2008, Johnson & Onwuebuze, 2007). Others support this and give specific criteria, such as, Green et al., (1989) who suggests the following benefits of exploring data through triangulation, complementarity, developing, initiating and expanding collections against relevant research questions. Bryman (2008) develops a similar approach but doesn't refer to it as 'mix methods' but instead combining methods for improving validity in all research endeavours. In efforts to find a universal definition, Johnson, et al., (2007, p123) set out to review multiple views and proclaimed the following;

*“Mixed-methods research is the type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for breadth and depth of understanding and corroboration.”*

One problem with Johnson et al., (2007) definition is its lack of positioning within a research paradigm and the controversies existing with mixing ontological and epistemological beliefs. Tashakkori and Teddlie (2003) address the difficulty in interpretation and hard to decipher the nomenclature of mix methods. They point to some factors, such as the little usage in research designs and the various mix method options available to researchers. To mitigate such commentaries, research methodologists have provided mix method typologies for study fields to follow, such as Morse (1991), Morse and Niehaus (2009), predominantly used in health research and Creswell (2002) for education and other social science disciplines. Such are the complexities involved in education studies; the application of multiple methods to answer research questions is becoming a suitable option, mainly, when we want to know why something works (Newby, 2010). This in-depth focus allows for a concentrated effort in research involving the evaluation and appraisal of new ideas from evaluation studies. As Creswell (1994, p177) states it is 'advantageous to a researcher to combine methods to understand better a concept being tested'. Moreover, there's a higher accuracy when involving multiple methods because whether data are analysed together or separately, consolidated results will take on a more meaningful appraisal and triangulated using the available methods akin to the research question(s) posed (Neuman, 2006).

### **3.1.4 Mix Method Research Design**

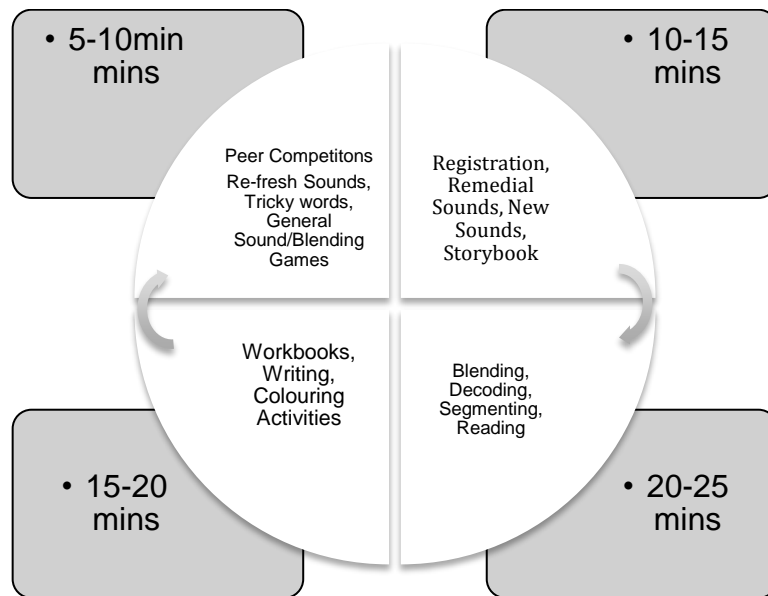
In the present study, a choice was made to apply an embedded design that uses a Qual – Quan pathway, that is, qualitative collections used in support of quantitative findings (Creswell, 2002, Creswell & Plano Clark, 2011). Qualitative methods are used to ‘assist, interpret’ and report on the internal workings of the RCT (Creswell et al., 2003, pg 178).

### **3.1.5 Mix Method Intervention**

Intervention research in education enables future programmes to be refined or further developed. In the case of this research, other interventions involving reading and peer cooperative frameworks encouraged the development of this study. Wright (2006) identifies the benefits of designing interventions in the language classroom. However, interventions designed to focus on EAL achievement are limited and present an opportunity for researchers and educational planners (Murphy, 2015). In response, this thesis considers a peer-cooperative intervention based on sociocultural principles. Over a period of four months, children acting as Monitors teach necessary phonic skills alongside their local peers. The teaching and learning framework in Figure 3 illustrates how each day of the intervention was constructed.



Figure 3 Intervention Teaching and Learning Structure



Multiple activities were set up to allow for the different skills required in developing sound and word knowledge. The framework gave Monitors structure in their approach and freedom to explore culturally created games to enhance experiences.

### 3.1.6 Why a Random Control Trial?

Random controlled trials (RCTs) are well established in the medical and health sciences due to their robustness and efficacy in intervention research (Boruch et al., 2002, Guyatt, et al., 2008, Rothwell, 2005). As a presence of authority in experimental designs, they are often referred to as the ‘gold standard’ for studies adopting evidence-based approaches (Maughan, 2013, Tymms, et al., 2008). However, extrapolation into the social sciences and across multiple disciplines therein, such as, healthcare practice (Semaan, et al, 2002), psychology (Stinson, et al, 2003) experimental criminology (Perry, et al, 2010), and social work (Naleppa & Cagle, 2010) often yield much skepticism due to their poor reporting. Education research also falls into this category, particularly when designs are not rigorously prepared (Torgeson & Torgeson, 2008), this being a contributor to the hostile commentaries it receives as a methodological option in the field (Ghate, 2001, Stewart-Brown, et al., 2011). Conversely, when designed correctly, it can be a powerful vehicle for programme scalability and refinement. Specific questions can be asked and research instruments deployed using a range of experimental designs. Only until recently have educational

researchers seen the value in RCT approaches, some calling for further implementation in programme evaluations (Goldacre, 2013). These calls have been echoed in international development studies targeting effectiveness, efficacy and exploratory designs (Banerjee & Duflo, 2011). In direct response to these calls, this study incorporated an exploratory cluster RCT (Torgerson & Torgerson, 2008) into its core design. The exploratory nature adds to the flexibility of evaluating intervention effectiveness alongside explanatory associations. Furthermore, in the protection of internal/external validity, ethical procedures, and reporting, this thesis followed a rigorous framework through the checklist provided by the Consolidated Standards of Reporting Trials CONSORT (see Altman, 2001). Primarily used in the health sciences, CONSORT has seen its criteria extended to other fields. Recently, extensions have been developed for psychological trials (Montgomery et al., 2013) using a modified formula. While there's yet to be an extension developed purposively for education trials, the criteria have been used in various educational interventions (Torgerson et al., 2013). Similar to other evaluation-based trials in education, this study reports the RCT method (exploratory pragmatic trial), randomisation procedures, statistical tests used for analysis (including steps taken to protect internal/external validity), trial attrition rates and overall effects, where applicable.

## **3.2 Quantitative Methods**

### **3.2.1 Introduction**

This section presents the quantitative collections used in the present study. Three test batteries are described alongside questionnaire and survey methods.

### **3.2.2 Family Background Questionnaire**

A family background questionnaire was developed to understand more about the home environment and socio-economic factors that may contribute to learning English. The final questionnaire (see Appendix A) was adapted from similar questionnaires completed in India (Tooley & Dixon, 2005) and other developing countries (Grosh & Glewwe, 2000) that have used background data to build vignettes to support the research picture. Building a profile of potential Monitors' and Learners' home-life and English experiences was

crucial in understanding, holistically, how English is accessed and used outside of formal learning situations. The final questionnaire was guided by the following access question;

- What specific socioeconomic and educational inputs contribute to learning to read English outside the schooling environment?

Specific themes were adapted to the research context of the investigation, such as incorporating questions to do with private tuition and study preferences. Questions were developed alongside the research team and school leaders with the remit of understanding the context of English usage at home for those children who were interested in taking part in the study. To enable a higher response, open and closed questions were used, thus, making it easier for children to answer.

### **3.2.3 Reading Aspiration Survey**

One of the aims of the research was to understand children's reading attitudes within and outside their primary learning environment. While the background questionnaires provided some clues to home life factors such as English exposure in the household and school type, the researchers wanted to elicit a more personal understanding from each student. Given the apparent linguistic difficulties in translating and transcribing multiple Punjabi interview transcripts, a decision was made to develop a reading aspiration survey that focused on reading in school and during recreational time. The final survey was adapted from the Elementary Grade Reading Assessment Survey (ERAS) (Mckenna & Kear, 1990) to reflect Punjabi recreation and academic environments. In this survey, the same procedures, as suggested by Mckenna & Kear, (1990) were followed and included children responding to twenty items linked to their recreational (10 items) and academic (10 items) reading settings. It is possible to calculate a total score based on both constructs, thus, reflecting total attitudes towards reading. Children complete the survey by responding to a 4-point (1 being extremely satisfied and 4 reflecting extremely unsatisfied) scale that is usually depicted by an animated character to help with translating relevant emotions. Total happiness in either the recreation or academic construct would be equal to 40 (4 x 10 items = total satisfaction). Therefore, total satisfaction would be equal to 80, where indifference is observed at scores close to 40 (or 20 per construct – recreational/academic). An example of a recreational reading item includes "How do you feel about reading English books away

from school?” The academic alternative asks the same question but relates to the academic environment. The full survey developed for this study can be found in Appendix B. For this study, questions were redeveloped from the ERAS (1990) survey to reflect Punjabi academic and recreational culture. The reason to modify the original survey was simply down to items that wouldn't translate in the Punjabi climate. Most of the ERAS (1990) survey questions ask how children feel when they “receive books for presents” and when they “read stories in class.” Further questions were adapted to reflect typical situations Punjabi children were found in on a daily basis. One example of this was to include purchasing English reading books from the local market and reading aloud in class on an individual basis. Indian classrooms tend to be rote focused. Therefore, each question had to reflect the norms found in a typical Punjabi learning environment. All questions developed relate back to the academic or recreational constructs for the analysis to offer accurate inferences.

#### **3.2.4 Phonological and Character Test Battery Scoring**

A diagnostic battery that includes three separate tests (Nfer A, B, and C) requiring children to recognise English character letters, phonemes and blend 2/3 sounds to form basic CVC words (Ames, 1980). All three tests were administered altogether in one sitting by a RTM on a one-to-one basis. Many studies infer letter characters play an essential role in developmental spelling (see, 1982, Pollo, et al., 2005) mainly as children attempt spelling words, such as RUDF? (are you deaf?) Bissex (1980, p3) and start to build a partial alphabet creating a semi-phonetic code (Bowman & Trieman, 2005 & Kessler, 2003), where only a portion of the sounds have been decoded in a given word. The Nfer A test can be found in Appendix C and displays ten rows of progressively tricky matches of letters. Children are asked to match each letter (and sequence of letters) to the same corresponding letter or sequence of letters from the same row. To be successful, they would need to recognise each letter character and then match the letter(s) along the same row. Each correct match is recorded by the research team member on a separate piece of paper – away from the child.

The Nfer B test (Appendix D) followed the same principles as above with regards to pace, allowing children enough opportunities and encouragement. The test aims to accurately sound out all phonemes contained rather than name each letter. Children were instructed to attempt all phonemes received one correct mark if all sounds (36 in total) were successfully

read aloud. The test is not prepared alphabetically, but some children will recite letter names. In these situations, the instructor must remind the child before marking incorrectly.

The Nfer C test requires children to identify sounds and blend them into whole words. Therefore, skills needed to complete this test successfully are twofold: accurately identify the phoneme(s) found in CV/CVC words and secondly, blend them to form a whole word. A total of 34 words are represented in this test and available in Appendix D. One mark is given for each successfully blended word, and children are instructed to attempt all words. This test aims to understand more about children's blending abilities and requires prior skills in the identification of constituent sounds found in words.

Lastly, the Miskin Nonsense word test (2006) was selected to further strengthen an understanding of children's abilities and application of their phonological processing skills. In this test, a series of phonemes and pseudo-words (the nonsense element) require segmenting and blending. The rationale behind test inclusion is that children who've acquired phonological processing abilities will be able to recognise the sound patterns through segmentation successfully. Moreover, it is especially useful in understanding whether or not children recognise words in parts (analytic method of recognition) or through memorisation (rote method of recognition) compared directly to the synthetic techniques. It is for these reasons that similar tests form parts of diagnostic assessment benchmarks in UK schools to understand more about children's decoding, blending and sound awareness (DFE, 2015). The list begins with phonemes and then moves on to irregular words for which children who possess decoding abilities can segment. Letters and words are batched together in order of difficulty, and similar to the Nfer tests, each correctly sounded pseudo-word blend represented one correct mark. The raw scores were finalised and inputted into the master spreadsheet ahead of subsequent analysis.

Following the Phonological Awareness and Processing Battery, to test the reading skills set three tests were selected and are discussed in the section below.

### **3.2.5 Reading Test and Dictation Test Battery Scoring**

The choice to include a reading test battery was necessary for understanding reading habits and current abilities. Following a component skills approach (Carr & Levy, 1990), where

various underlying cognitive and functional skills are assessed by their relative contribution to reading development; this study tested for word recognition, word construction, and passage reading skills. Tests were selected on their contextual, historical and reliable influences on reading test performance. For the latter, there is substantial evidence that by including such a battery can help understand the various skill subsets required in the reading process (Koda, 2004). Moreover, research continues to find highly correlated associations between spelling and reading testing, mainly, in the early years of literacy development (see Bishop & Leonard, 2000 and Frith, 1985). Having only one component, such as word recognition, would reduce the findings to word perceptiveness. This study wanted to go further by including alternative tests to investigate reading ability. Fluidity in the recital of words positioned in a sentence is a vital subset of numerous components in learning to read, in this regard. It was, therefore, useful to start with a word test to investigate word knowledge before moving onto spelling and reading passages. The Burt Reading test (1974) (cited in Thorpe, 1976) was selected as the first test in the battery based on the suggestions of previous studies in development contexts of Dixon, et al., (2011) and Eshiet (2014). The Burt Reading test (1974) tests for children's ability to read 110 progressively tricky English words. It begins with 2-3 letter words and with each word that is read correctly, one mark is awarded.

The Burt Reading Test (1974) has good construct validity - test-to-re-test - through its various standardised norms based on various countries, such as Gilmore et al., (1981) in New Zealand. However, these were not used in the primary analysis, as there is no Indian scale for converted scores. Therefore, the primary analysis used raw scores at both the baseline and post-test stages to ascertain individual and group differences. Final totals were tallied as raw scores, which formed part of the primary analysis and primary results.

The second component of the Reading and Dictation Test Battery included the Schonell Spelling Test (Schonell & Goldacre, 1971). It is a famous test commonly used in UK classrooms and has also been used in international education contexts (see Dixon et al., 2011, Eshiet, 2014) to understand more about spelling abilities. The test was first standardised by Bookbinder (1976) to reflect the modern usages in English vocabulary. This was a significant factor when researching the vast number of spelling tests available from the literature and the practicalities involved with large samples.

The following factors were considered. First, it was vital to select a test that reflects modern, cultural and linguistic capabilities. Indeed, in early spelling development, children will perform better if they are familiar with the words commonly used in educative and social situations (Joshi & Aaron, 1991, 2003). Unfamiliar words, or those, as such, not part of a natural setting may skew results and also render the test incompatible for what is being tested. In this regard, words that were found on spelling tests in the host school mirrored those found on the Schonell test. Second, the investment of timing naturally required attention. One of the advantages of running the Schonell is it can be administered to large groups. Other tests in this study followed a one-to-one arrangement, thus, taking more time to run and administer. To successfully run the Schonell test, one test master (selected research team member) and support for children, such as offering supply pencils, paper and responses to general enquiries (from peer monitors) form the minimal requirements. Lastly, consideration of the predictive reliability and word knowledge involved in spelling words were sought after. Tests that have strong concurrent validity were researched alongside the parameters mentioned above. Research conducted into the modalities of spelling tests is not evident in the literature due to the range of tests deployed within various contexts and preferences (Calleia & Howard, 2014). For example, some tests apply proofreading skills, while others use multiple choice questions, dictation or error correction involving the deletion of incorrect words (Moats, 2005). These are all modalities in the delivery of a spelling test, and all have varying results. This has led to much debate on the universality and validity of standardised scores as an accurate measure of spelling ability and relative inferences made after that. Central to this argument is the spectrum of tests and their power at accurately predicting spelling ability holistically. In theory, if the dependent variable (spelling ability) were the focus of each test then this would imply that all modalities would show a high correlation in overall test performance. Studies conducted to address this issue found that adding ‘dictation tasks’ are far superior to any of the other test modalities (Willet & Gardiner, 2009, Calleia & Howard, 2014). Both studies found that dictation tests can significantly improve the relationship a learner has with words. This strengthened the choice to include a dictated spelling test but placed restrictions on what could be developed in the short amount of time before study commencement.

One of the caveats of spelling tests, from an Indian perspective, is the severe lack of spelling and general dictation guidance to inform practice. Schools tend to follow instructions from established, sometimes foreign textbooks that are bought independently. NCERT (2012)

offers some guidance of the necessity to undertake spelling tests alongside other tests for spelling, reading, and writing. However, there is no standardised data on the state of spelling, nor any pedagogical guidance on successful implementation. Since writing this, no such research exists targeting spelling modalities in India, and this leaves a significant gap for further research. While this presents a unique opportunity for the current study, it also posed some risks. Firstly, like the Burt Test, the Schonell test have no Indian norms, even with calls from some for its apparent ease of introduction within English-medium schooling environments (Singh, 2004). The test itself comprises a total of 100 words typically reflecting a spelling test of progressive difficulty. Venkatesen & Holla (2011) demonstrate the capability to construct a hierarchical graded word-reading test for children with learning difficulties. They report excellent internal/external validity for their test, comparing composite scores across grade ranges (Kindergarten up to Primary 5). Similar strategies were adopted in the present study, such as using words from a range of Kindergarten-Primary V Standard textbooks related to the CBSE/ICBSE syllabuses. Venkatesen & Holla's (2011) work clarified the test selection method in an Indian context follow a delicate process. A decision was made not to replicate the test only because Schonell's listed words mapped to the CBSE curricula and textbooks typically found in the grade ranges involved in the experiment. Moreover, the analysis provided by Newton & Thomson (1978) illustrates the Schonell test has a high (historical) concurrent validity, the coefficient being .84 and a good predictive validity at .64 when compared to other tests used in language diagnostics. When compared to Venkatesen & Holla's (2011) self-created test, the Schonell has a historical legacy enabling it to be more powerful for analysing raw spelling ability in the primary analysis and predicting spelling ages in the secondary.

Children are very creative when learning to spell, the pioneering work of Read (1975, 1986) detailed this through exaggerated spellings of words and challenged previous set beliefs that children memorised words solely, or by rote. This research documented these as a developmental process, as Read and others since (Ehri, 1995) have documented in that spelling perceptions are as just as an essential stage than correctly spelling achievements themselves. However, spelling errors are not the primary focus of this research and only correctly spelled whole words were counted in the final analysis.



### **3.2.6 Dictation Test Delivery and Scoring**

The final test in the battery included a passage reading (or dictation) test. This required all children taking the test to read aloud progressively difficult sentences. As a central theme throughout this thesis, most Indian children are severely at risk of not developing basic reading skills, which has a profound impact on their educational journeys. In the current context, the latest research from rural Punjab points to 60% of children in Standard III is unable to read a Standard II textbook (ASER, 2015). This improves slightly as children progress into higher Standards, but the numbers are still very high. Reading fluency has connections to phonological (Stanovich, 1990), word recognition (Torgerson et al., 1999) and accelerated reading comprehension after mastery of requisite skills (Fuchs et al., 2001, Jenkins, et al., 2003). Therefore, the inclusion of a reading test following the phonological, word recognition and spelling tests enabled a greater insight into how children learn to read.

Mirroring the Pratham method (Banerji & Bobde, 2013) of selecting an appropriate test for children to interact with, a mini-review of the reading test literature was also conducted. Reading tests are in abundance throughout the research literature. They aim to understand recognition, fluency and pace best to enable readers the chance to develop an unconscious connection with written sentences (Logan, 1997). The salient technique and ability to be able to read, therefore, reflects the mastery of translating the written sentence into the spoken language (Rasinski et al., 2011, Adams, 1990). These theoretical assumptions along with the pedagogical advice from the various Pratham Reports informed the type of test required for children in the present study. Given the sporadic nature of English teaching and its introduction at various grades, Pratham undertook rigorous pilots across India to investigate how best to deliver a reading assessment that is universally accepted (Banerji & Bobde, 2013). They found combining National curricula learning goals with State-wide textbooks the best approach to understanding reading practices in schooling environments. Whether English (as a subject) was introduced at the Primary grade 1 or 3, their reading dictation tests were found to represent words/sentences familiar to children.

Textbooks and reading materials from the host school were provided, and various conversations on key words were discussed with the Principal, teachers, and children from grades 1 -3. The finalised sentences were mapped to textbooks, syllabuses to reflect the test literature on word familiarity. In the final test, to be successful, each child was asked to

read aloud each of the sentences and would receive one mark for each sentence correctly read aloud. The final test can be found in Appendix F, where it is noted the progressive difficulty (and length) of each sentence. A total of 10 marks are awarded and used in the subsequent analysis for the primary research strand.

### **3.2.7 Native Reading Test Battery Scoring**

Two native tests were constructed to understand more about reading in the home languages. As Punjabi (State language) is predominately used as the primary L1 and Hindi (National language) is mixed as a L1/2. Thus, there was an interest to see how well the children grasped basic reading skills in their home languages. It was discussed at the pre-pilot stage that children tend to learn Hindi from Primary 1 (1st grade) with Punjabi used in the early years and social situations, due to it being the State-adopted language. However, this proved not to be universal when considering various non-mainstream education situations, such as private tutoring or cultural activities in the local area. Advice from teachers and the Principal of the host school recommended investigating reading abilities due to a lack of localised understanding of language divergence and its practical implications for school readiness. To classify this, Hindi is often used as a bridging language that links to English, whereas, Punjabi borrows from Hindi (and vice-versa) in many spoken/written contexts due to the similarity in both scripts (Gurmukhi – Punjabi, Devanagari – Hindi) and because of its alphasyllabic properties. Understanding how these work and whether there is core knowledge of one or both is vital to the reading process. Various research reports the benefits of cross-linguistic transfer when English is the L2, and the alphabetic system is different, word recognition abilities increase (Hamada & Koda, 2008, & Carr, 1990).

Similar to the other reading battery, the same process was followed in test selection. A Punjabi dictation test was selected influenced by Standard 3 textbooks. The final ten sentences were developed in unison with teachers, parents, children and school leaders following the initial pilot stage of the research. The translated sentences included in the final test can be found in Appendix G. Each sentence followed the same procedure to that of the English reading dictation. Children are asked to recite each sentence in turn and were given time between transitions. One mark is recorded for each correctly read sentence to

mirror the same conditions present in the English dictation test. Marks are then recorded onto a master copy against each research participation identification number.

The Hindi word test followed the same processes of all the other tests with regards to selecting appropriate age words, level and links to future study. The main problems with word selection for the test were down to the confusion of when Hindi is taught (and used), whether in learning or social contexts. From the host school's curricula point of view, the language is not formally introduced until the first grade or Primary school. Here, the alphabet (Devanagari) is introduced, and knowledge of speaking and listening skills is developed. However, parents asserted that it is primarily understood from a very young age due to the mainstream media, such as cartoons and movies. It was advised by school leaders that 'passage reading' would prove to be too difficult, this relates back to language exposure, that is, the inclusion of reading, writing and listening to the language. Moreover, and following advice from local Hindi teachers operating private tuition or who were registered in nearby schools, it is estimated that the 'reading' component of the language is not taught until the 3rd Standard. It was advised that whole sentences would probably be too complicated and not map across to the sample, instead, either letters/words in their basic form should be considered.

A selection of tests was investigated from Central and State Boards, however, most were targeting middle to higher primary ages ranges; thus the difficulty raised. However, Pratham has developed and run numerous tests used in rural contexts as part of their annual reports (See ASER, 2016). The tests include a paragraph that children are requested to recite and words that they must identify in Hindi. Given the targeted age range being lower – middle primary school children, the Pratham tests would need reconfiguring to suit local requirements. Instead, a decision was taken to include a word list, similar to the Burt reading test with progressively harder words that are suggested given the age banding. Hindi schoolteachers from the host school supplied further guidance on suggested words suitable for lower to middle ranges. A decision was made to exclude a passage reading test due to the parameters as discussed with trained local teachers, parents and the host school. It was determined that it could take up valuable time and possibly cause confusion with the Punjabi dictation test. Moreover, as the primary aim of the research was to understand how pre-reading phonetic skills could be used to accelerate English learning, it was logical to

include knowledge of other languages but not analysed within a cross-linguistic format. One mark is awarded for reading each word correctly (30 in total) and transferred to SPSS for analysis. The translated words used in the test can be found in Appendix H.

### **3.3 Qualitative Methods**

#### **3.3.1 Introduction**

This section presents the qualitative collections used in the present study. A pre-trial focus group was used as a supportive extension of understanding reading habits amongst Monitors in academic and recreational settings. The justification for its inclusion is outlined below.

#### **3.3.2 Focus Groups**

The placement of the focus group discussion used in this study followed an embedded mixed explanatory design in support of the initial quantitative findings. A focus group interview is a useful qualitative research method in facilitating a synergistic discussion among all participants (including the researcher) on topics of interest (Patton, 1990). In this fashion, multiple focus groups are used to enhance quantitative collections that initially surveyed reading aspirations. As discussed in Chapter 2, understanding-reading attitudes can have an everlasting impact on the quality associated with being able to read. Therefore, it was crucial to understand how Monitors interact with reading experiences within academic and recreational environments. The design of the focus group discussion was governed by the factors and subsequent questions derived from quantitative collections. Typically, focus group methodologists suggest groups be between 6 and 12 members (Patton, 2002, Denscombe, 2007) to enable a broader rich discussion. However, group discussions in this study aimed to look at specific content and permit the greater depth of reading experiences (Anderson, 1990). Recruitment was via purposive sampling, applying maximum variation sampling techniques (Patton, 2002). Related to these principles, the underlying criteria for the maximum variation sampling in this study included the following;

- Registered as a Monitor in the experimental arm of the study

- Participation ceiling of N = 16 Monitors (n = 8 male, n = 8 female) to form the focus group
- Selection stratified against positive and negative attitudes of academic and recreational reading (using baseline scores)

Due to time constraints, it was not possible to include all Monitors. Instead, the selection was organised to include Monitors' who gave a mix of positive/negative responses from the English reading attitude survey. The 16 participants were divided into four groups of four and given time slots for interviews. Focus group interviews were allocated to take place over four afternoons at the host English medium school. The focus group interviews were planned in English, inclusive of questions from the lead researcher and answers from participating children. All participating children were happy to converse in English, however, to help with remove any ambiguity during interviews, a language teacher was present for support. A total of 12 questions (including probes) formed the interview guide (see Appendix N), with minor deviations to support children's answers and encourage responses. Each focus group was given a time range of 15 – 20 minutes to allow for responses from children in keeping within the interview framework (Bogdan & Biklin, 1998). The interview procedure included either a warm-up activity or starter question to settle the children. To begin with, a general question related to 'first thoughts' of English reading experiences was planned, the idea being children would familiarise themselves with the interview process and generate basic discussions. Following brief discussions and introductions, the next stage of the interview included questions related to the positive and negative aspects of reading English in recreational and academic settings. Children were required to give as much detail about their experiences and expand on examples wherever possible. Data were captured via dictaphone alongside researchers notes. Two research team members were also present for each meeting and took notes to use for later comparisons. Data analysis followed all completed interviews and is described in the next sections.

### **3.4 Data Analysis**

#### **3.4.1 Quantitative Data**

All test data; questionnaire and survey responses were entered into SPSS for cleaning and analysis. Various statistical tests including t-tests, ANCOVA, and Principal Component

Analysis were used in a variety of ways to understand trends and data patterns. Following experimental procedures, data were collected before and after the intervention period for all participants. Questionnaire data were collected before randomisation took place to stratify participants into equally weighted groups. These procedures, including their reliability, validity (internal, construct and external) are discussed during the various research phases and results section. However, universally, all parametric tests must follow a normal distribution (Field, 2013), that there is uniformity in the data set. Analysis using graphical outputs and Levene's (1960) homogeneity of variance test statistic is used in this thesis for understanding group parity before proceeding to interpret whether any statistical differences exist. Tests of significance are meaningless if there is no accompanying discussion on why there is (or perhaps not) a statistically significant difference between groups. Since the conventional strategy to interpret pre/post data includes: Type errors, level of power in the control group (responses and standard deviation) and expected treatment effect (observed differences), it is essential to recognise these procedures before inferring effects.

For education research, the basis of any statistically significant differences ( $p = < 0.05$ ) infers that 1 in 20 (or 5 in 100, 5%) is related to chance or some other factor followed by rejecting the null hypothesis and accepting an alternative one (Peers, 1996). The acceptance of setting the probability value at  $p = < (0.05)$  is universally accepted in the social science research arena (Field, 2013). However, some studies adopt more stringent criteria by selecting the values to be  $< p = 0.01$  or lower. This is more to do with the outcomes that are to be measured and how confident researchers want to be that the null hypothesis is to be rejected. Hypothesis testing can often be misleading if we rely solely on their findings alone. As Kirk (1996, p754) suggests, "what we want to know is the size of the difference between A and B and the error associated with our estimate; knowing that A is greater than B is not enough." Therefore, it is perhaps unsurprising that there has been much criticism of studies that don't report the full story of what is happening to the analysed datum (Cohen, 1994, Schmidt, 1996). In summary, 'the primary product of a research inquiry is one or more measure of effect size, not p values' Cohen (1990, p1310). This is a hot topic in hypothesis testing studies. We might report a statistically significant difference between two groups at the 95% confidence level, but this can be somewhat arbitrary if there is no indication of what it means. In essence, a statistically significant difference between two test subjects could infer that the sample employed was indeed large enough (Lang et al.,

1998). Thus, the need to locate, in detail, how significant an effect is and what it means to a group is what is typically missed and in some cases misinterpreted. What should follow is an interpretation that is guided by questions about the sample. Such questions like how significant are any detected effects? Moreover, to what end can these be replicated in future studies? To facilitate this, some procedures can be, and arguably should be, adopted when reporting statistical significance.

Conducting a sample size calculation is one procedure that is helpful to determine the size of effects. Educational evaluation research has a redundancy with not reporting a priori sample calculations; indeed, some studies give no clear rationale on how samples were formed (Torgerson, et al., 2005). To discover small effects, group allocations tend to be large, and this can result in over-prescription and thus wasting valuable resources for what is required. Indeed, if large effect sizes are required, then, the reverse is to be true; a smaller sample is required.

This study conducted a power analysis using the G\*Power programme (Faul et al., 2009) to establish an adequate sample size for the intervention. Power analysis revealed that a total sample of 651 participants was required to 80% power to detect a small effect size of 0.11. Given the variability of peer-cooperative and reading effectiveness studies as reported in Chapter 2, it became clear that a larger sample would be required to avoid Type I and Type II errors.

External validity is threatened if studies reporting statistical information is uninformed or not explained to the untrained eye. Brooks et al., (2014) studied claims of ‘understandability’ from a range of standard reporting (i.e.,  $r^2$ ,  $d$ ) measures compared to non-traditional methods. They found non-traditional methods, in particular, the ‘common language effect size’ indicator (McGraw and Wong, 1992) a better vehicle for communicating meaningful results. The common language effect size represents an index that considers possible differences between matched pairs from two groups. In its simplest form, an effect size is a test statistic that quantifies differences between two (or more) groups, the magnitude and size of difference as measured are often reported following significant results, whether positive or negative (Kline, 2004, & Sun, et al., 2010). Reported effect sizes vary across studies, mainly due to the correlations/differences a studies report. To that end, effect sizes fall into two main categories: the measure of total mean difference

and strength of association (Henson, 2006). For this study, an effect size is simply the magnitude of the observed mean differences when comparing a group that receives a peer-cooperative intervention to a control group that receives traditional teaching. Cohen (1960, 1988) provides an interpretation of effect sizes that can be used as a guide to interpreting magnitude effects. The effect sizes used range from 0.2 (small), 0.5 (medium) and 0.8 (large) and are used in the interpretation of results in this study. Including a sample of N = 651 or more improves the chances of detecting small, medium or large effects (Cohen, 1988) are reported with 80% accuracy.

### **3.4.2 Qualitative Data**

Qualitative data from focus group participant responses were analysed using an interpretive approach to establish how Monitors' perceive reading experiences in recreational and academic settings. Focus group questions were created following a data reduction process using quantitative instruments to extract meaningful patterns that could be enhanced at interview. Factor analysis identified major categories that informed the interview schedule and initiated a triangulation process.

The analysis was completed using the constant comparison method (Strauss & Corbin, 1998). A procedure of open, axial and selective coding was used in searching for significant themes (ibid, 1998). Firstly, all interviews were tape-recorded (using a Dictaphone) and then transcribed in preparation for the first stage of coding. Krueger and Casey (2009) suggest various methods of transcription and recording, but a decision was made to use MS word rather than specific qualitative software. This made data extraction more natural and more manageable when considering the challenging climate for undertaking this type of research inquiry. During the open coding phase, one RTM assisted with the examination of preliminary data, with the aim of identifying and interpreting participants' feelings, ideas, thoughts, and aspirations associated with reading experiences. Every attempt was made to remain faithful to the emerging codes by identifying participants' voices in the context of their original thinking. Following this, a process of categorisation was incorporated and developed to link any sub-categories from the interviews. Sub-categories refer to additional themes that emerged as a supporting function to one (or more) of the major categories. These categories were compared with the RTM's coding and further developed before



finalising. An example of the steps taken in interpreting a range of open codes and relationship patterns - leading to sub and major categories can be found in Appendix H(i).

### **3.5 Ethical Considerations**

In all research, there is an ethical challenge to recognise, and that must be embedded into the design, implementation, and follow-up of the research cycle (Payne & Payne, 2004). For this research, the guidance provided by the British Research Educational Research Association (BERA, 2011) was adopted.

#### **3.5.1 School Clearance**

Initial contact with the host school came through an invitation from the school's Chairman to investigate ways of improving English teaching and learning opportunities in the early years. The first phase visit culminated with an agreed plan to investigate the effectiveness of a synthetic phonics programme mediated by peer Monitors. A letter of introduction (See Appendix I) was created to formalise the programme and sent along with a draft research proposal to the host school requesting permission to undertake the programme. After initial approval, regular Skype meetings were arranged to discuss strategies and how best to implement the study framework, taking account of known cultural barriers, such as, language, cultural, social-emotional awareness and working with vulnerable children and adults. Guidance from these meetings proved to be valuable in the creation of new plans and provided a more precise focus on how best to implement the work. One of the main risks in gaining access to subjects, participants, and other related groups is how the researcher may influence positive, and in some cases, unintentional negative responses from gatekeepers (Yates, 2004). The reality of having agreed to an agenda with support from the host school enabled a natural flow in all of the preparations.

Following school management and the agreed research plan, two further school visits were arranged to meet with teachers, parents and search for possible locations to conduct the study. It was identified that the researcher creates a school team that was able to assist with the project before, during and after the study. On arrival, the researcher delivered a presentation to school management and teachers on the nature and purpose of the proposed research. The presentation focused on the possible merits of providing such an intervention

counter-balanced with potential risks as identified through previous discussions with management. It also demonstrated the types of tests and conditions required to deliver a study of this type. Teachers who expressed an interest in assisting with the project were invited to other sessions that discussed roles and responsibilities required for the development of the work. This was conducted through voluntary informed consent concerning points raised by social researchers like Bryman (2015) who states that all participants must be aware of the primary purpose of the research goals and has the right to withdraw. Moreover, participants were informed that the choice to leave the programme had no bearing on their employment status at the school and this was supported both by the management committee and school Principal.

### **3.5.2 Request for Study Location**

One of the significant challenges at the preparation stage was to identify and seek permissions for a study location that was safe for children to attend. It was suggested that the Gurudwara was suitable because of its scale to house large numbers and because of its proximity to the main town centre. Meetings with Gurudwara management were arranged and later agreed to the plans of the study but recommended that there should be no interference with those wishing to pray and for any children to enter religious ceremonies. Entry into the Gurudwara itself was off limits to all children and management agreed to help with the transit of local children from the outside into the designated areas. A small hall and various large rooms were allocated as these were seen as not interfering with the main traffic into the temple. Further to this, agreed dates were outlined in the calendar to avoid confrontations with religious festivals and public gatherings. Plans were adjusted to meet the dates by either cancelling the teaching for that day or by starting earlier or later.

Given the significance of the temple to the local village and the school's reputation, Gurudwara management placed a vast degree of trust in the proposed study. Ryen (2011) notes the importance of being trustworthy when carrying research in the field, in this case, the management granted access based on the school's reputation and the impact it has on local society. To enhance relations and in line with the guidance provided by BERA (2011, p6) on ethical standards, a letter describing the trial scope and procedures was presented for consideration. After this consultation period, the Gurudwara management permitted

access to their facilities based on the points raised through discussion. This letter and approval can be seen in Appendix J.

### **3.5.3 Recruitment of Children**

To fulfil the study requirements, two sets of children were sought, those who had basic English skills and those who either did not attend school or were registered to a government alternative. As it was recognised by school management that access would require alternative approaches, two separate procedures were followed. In educational research, it is particularly important to receive consent from parents, children, and teachers even if the school are happy to act as gatekeepers (Hutchinson & Styles, 2010). It was thought that informed consent through written letters and attachments would not be enough to convey study goals with possible key information lost through translation or be deemed ambiguous. Indeed, Homan (1991) argues that informed consent in which all participants can fully understand the proposed research before the study may lead to further confusion, leaving what happens in reality unknown. BERA (2011, p6) provide a universal description in through their guidelines acknowledging alternative measures that seek better approaches in articulating the proposed research, they state;

*“In the case of participants whose age, intellectual capability or other vulnerable circumstance may limit the extent to which they can be expected to understand or agree voluntarily to undertake their role, researchers must fully explore alternative ways in which they can be enabled to make authentic responses.”*

In line with this and given the complexities of language, religious and social factors, informed written consent was one way in which to articulate the research agenda. Therefore, two approaches were employed in recruiting children into the study and these are explained below in more depth.

### **3.5.4 Monitor Recruitment**

For the Monitors, a school-wide campaign over each of the three visits was conducted to gain interest in participating in the research. This led to the researcher conducting presentations in school and inviting children to attend and discuss their potential involvement. Children were told about the tests and shown some examples of questions they might expect to see, this allowed for transparency and a level of trust between the

researcher and participants. Children were also informed that they had the right to leave the programme and of the relevant teaching personnel if they wanted to speak in confidence about problems or potentially leaving the programme. Confidentiality was, therefore, a crucial proxy for developing trust with the children, built-in to understand their realities and to find common ground that would be enacted as a safeguard for their protection. Guidance from Berg (2004) ensures children's choices, responses, and other related personal information would not be recognised or discussed publically with others or identified in final reports. Registered teachers, the researcher, parents and school management had access to personal records to those in the study, and great care was discussed to ensure that no information could be linked. Photo's and media used in this thesis and during the study was permitted by parents from consent letters and debriefing meetings.

Following the presentations, those children who expressed an interest were given a letter of invitation (See Appendix K) that required parents to attend meetings at the school to discuss the project. Two meetings were set up to introduce the project and allow questions about the logistics, timing, impact, and purpose of the study. Teachers recruited as assistants for the programme provided language support as parents in attendance chose to converse in Punjabi, Hindi, and English. Final registration into the programme only permitted signed consent forms following a period for parents to consider options.

### **3.5.5 Learner Recruitment**

Recruitment for learners was managed in two ways, firstly by visiting local government schools, meetings with regional educational leaders and second by placing advertising posters around town. Meetings with government school Principals and leaders helped to gain further access to children under their care. Three school Principals gave access to talk to the children about an after-school English learning programme. Two schools declined by telephone and were therefore not subjected to further inquiries. Brief demonstrations at the schools that accepted were given providing an insight into the programme. Smaller posters were handed to the children (see Appendix L) to take home to their parents accompanied with a translated letter of invitation (Appendix K). It was crucial to keep the wording to a limit at this stage as it was suggested that most parents are illiterate in their mother tongue, therefore, a standard letter requesting consent be impractical.

Subsequent meetings then took place with interested parents who wanted to know more about the study. Once again, language support from recruited teachers improved the chances of the study being understood fully. Some parents provided written consent while others gave it verbally. RTMs provided support on the meeting days and requested to complete the formal registration process at the earliest. The letter of consent and study debriefing guide used at the meetings can be found in Appendix M. Parents were informed to register their child for the programme and were permitted the same rights as discussed above to Monitors' parents.

## **3.6 Phase One –Ground Work**

### **3.6.1 Introduction**

This section presents the three research phases in the current study. The first phase included the first trip to India to understand the opportunity and scope for developing literacy practice. Following remote conversations and dialogue with school leaders, the second phase details steps taken in the creation of a peer-mediated cooperative phonics intervention. A pilot study and initial training for all participants is also explained. Lastly, the final research phase documents the transition from programme registration right through to intervention post-tests.

### **3.6.2 Setup**

During the early stages of the PhD, possible projects were evaluated against the researcher's interest in early year's literacy and in particular, phonological pedagogies that can be used in a development context. Initial ideas considered setting up teacher trials in Ghana and Sierra Leone to measure phonic effects on learning to read. However, this proved to be costly and presented further difficulties in setting up trials.

Through connections from my Doctoral supervisor, a rural school in India wanted to trial a new phonics literacy strategy that looked to improve the quality of early year's English within the school and potentially as an outreach project for local, non-registered children. This opportunity proved to fit well within a sociological framework the research initially wanted to follow. Conversations with the Chairman and Principal of the school followed to understand specific problems and challenges the school and community were facing.

From these initial conversations, three themes emerged that helped shape the direction of thinking;

- English literacy is a significant problem for the school. Children who progress to the higher Primary years struggle to pass the Board examination. Other children who struggle leave the school for alternative Boards, namely, Punjab Education State Board (Punjabi medium)
- The number of school places (and options for English medium schools) is limited throughout the area. Limited enrolment (850 school-wide) places much strain on finding adequate provision for children from the village. The school mentioned the difficulty in offering places as a consequence.
- Parents have high aspirations for their child to learn to read and converse in English. It is unclear what the children's attitudes are to reading English.

Based on these central themes, some ideas were proposed, but it was difficult to understand without context. Three research trips were subsequently planned to entail three phases of development. The first research trip had the following objectives;

- Understand the context of English teaching and learning for reading purposes in the early primary years at the host school
- Analyse the local problem of aspirations to learn English and opportunities that can facilitate this option

The first trip lasted a month with various meetings with school leaders, teachers, parents and local education advisers from the area. It became evident that teaching in the host school followed a rote-style delivery. Textbooks and chapters were covered in the middle to higher primary years whereas lower first classes focused on alphabet delivery and letter writing. Although the school is an English-medium enterprise, children conversed in Punjabi socially and early years' teachers had no choice but to instruct in the home language.

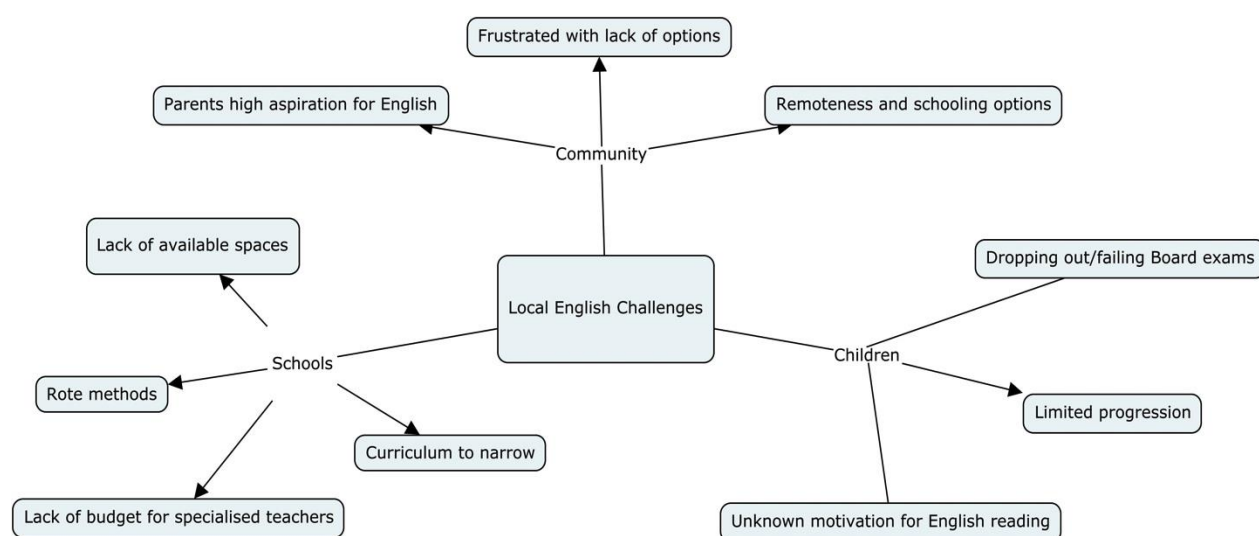
Outside the school, the local area is served by two government schools (primary/secondary), one PSEB school, one Khalsa school (religious denomination) and one other CBSE affiliated low-cost private school. Visits to these schools only triangulated

the same issues as initially communicated. Lessons were very much rote-led, and teachers were brought from outside the town to cater for the growing demand for English in the area. Parents from the area would make it explicitly clear that English was the most sought-after subject as it enables children to progress to College/University and also work in skilled labour jobs such as in farming. School places were filled with some parents opting to send their children outside the village for schooling. Whereas others decided against schooling on a regular basis because it was either unattainable or the quality offered was not acceptable. Children who were not registered to any provider or in government managed schools were identified as being particularly vulnerable to falling behind key learning concepts in all subjects.

### 3.6.3 Project Searchlight

Further meetings with the school Principal and school team about how best to tackle the situation were arranged. The following concept map in Figure 4 was crafted to visualise the common problems faced by a child, school and community level.

Figure 4 Project Searchlight Concept Map



One major theme that ran throughout the meetings and from the concept map is the lack of skilled teachers. Original proposals considered recruiting specialised teachers and testing for overall effectiveness, however, this conflicted with budgetary limitations and was previously explored by the host school. Other ideas included specialised training events and master classes shared by government-private-religious denominated schools as this would alleviate financial pressures and share expertise. However, there were issues related

to the different curriculums schools follow and scheduling of events. Timetable clashes would be inevitable as each school runs to specific schedules and Boards of education require strict adherence to their policies. This proved to be too complicated and would mean disruption to children who attended school.

#### **3.6.4 Towards a Peer-Mediated Trial**

Throughout the time spent at the school and working with parents in the town and connected villages, it became apparent that some children had a basic grasp of English. Frequent conversations with children about their favourite cartoons, movies, and jokes were often shared during house visits. Some children, as part of their home life after school, were required to teach their younger siblings what they had learned during their school day. Visiting some pupils' houses from the host school confirmed this. Issues to do with schooling availability, affordability and remoteness were given again as reasons why younger children were not able to attend provision. Nevertheless, some children were observed reading their schoolbooks and sharing stories with their brothers and sisters before completing homework. This spring boarded the idea of developing a trial that included children as co-creators of knowledge rather than sole recipients.

Further discussions with the School Principal confirmed the practical nature of running a trial that enabled children from the host school to cooperatively learn alongside their village/town peers. Subsequent meetings followed with teachers and a selection of parents who also took to the idea and recognised the useful nature of English language skills utilised in this way. Moreover, a child teaching other children cooperatively wasn't uncommon. Children from the host school were already part of a mentorship programme that was created to develop paired/guided reading, math games, and ICT support. These programmes would run every second Saturday to free-up teachers to allow for planning and book marking. Children from the higher year groups would be assigned to a specific class and follow instructions left by class teachers. Paired reading included children reading while mentors would listen and guide them through difficult words and offer general advice on responses. Math games were also created to encourage counting, addition, and subtraction. The arrangement would include groups with two or more mentors playing alongside their younger peers. Finally, two computer rooms were made accessible for middle to higher classes wishing to develop their ICT skills including basic word



processing and drawing. Five to ten mentors would walk around offering assistance and demonstrating basic skills while children implemented their instructions.

The flexible approach to learning through play and discovery was well received by teachers and Heads of School. All senior teachers reported the social and cognitive benefits of the peer-mentoring system in place at the school. Behavioural problems were virtually non-existent and accounted for a small number of cases. A learning log – a kind of handover book was completed by each mentor and given to the section in-charge (senior head teacher) to document learning tasks and any relevant information to feed-forward to the next session.

One encouraging benefit of the peer mentorship programme the school adopts is the differentiation of pedagogical techniques used to target skill enhancement. For example, the focus on reading development with a one-to-one scenario compared to a more open dialogue with multiple mentors for ICT instruction. The system, therefore, demonstrated capacity and flexibility for the adoption of different peer learning designs. Following further meetings about the proposal, a mini-literature search was completed that confirmed the traditional and practical benefits of sociological approaches (Johnson & Johnson, 1991) to peer-mediated learning in an Indian context (Counihan, 2015). However, at this stage, it was unsure which peer method would be adopted as the principle pedagogy – this being a requirement for lesson structure and transitions.

### **3.6.5 Establishing a Research Team**

To be successful in the development and delivery of the trial and study, there needed to be a significant support network with clearly defined roles. Both trained research assistants and monitors were required to enable the planned trial to operate. The recruitment of each group is detailed below.

### **3.6.6 Research Team Members (RTMs)**

Discussions with the Principal from the host school outlined the local challenges and some of the difficulties in enabling the procedures of the proposed trial. Themes included recruiting teachers with necessary key skills, teacher availability, as the trial runs parallel

to the school year, and costing of teacher time or recruitment of specialised skills. In considering these factors, there was a joint decision from the school and researcher to allocate teachers from the host school to various roles. Training was identified as a significant factor and was built in to accommodate the various phases of the study, from pre – to – post. Principal suggested the following core conditions a teacher should possess in the successful delivery of the study;

- A willingness to participate and explore new knowledge
- Have a BA in English, Punjabi, and Hindi
- Have a postgraduate qualification in English or be teaching in the higher (Standard 10+) year groups
- Have an awareness of English phonetic properties combined with a linguistic background
- Ensure that classwork and general planning are not disturbed

The conditions were met by six teachers working in the higher year groups and had an interest in the pedagogical design of the trial. All of them have a linguistic background at the postgraduate level (minimum qualifications to teach in the higher years) combined with teaching qualifications awarded from various Universities. At a later stage, four Hindi teachers were recruited to help co-construct in-school Hindi tests to enable a fairer and more balanced approach for children who were asked to undertake them.

All teachers were entered into a weekly training programme to gain insights and scope of the trial and more specifically the roles and responsibilities that each teacher would have. Teachers were seconded to the study every Saturday to coincide with general staff training and school-wide planning. Each training session had a specific focus that included the following;

- Test training and trailing events
- Location visits
- Data entry and management
- Translation and transcription

Training events and notifications of procedures were routinely discussed to enable each team member a chance to understand what was expected of them. Teachers from the host

school played a significant role in developing the study to make it viable and pertinent to the local and educative conditions it was to be set within.

### **3.6.7 Searching for a Location**

Alongside lesson observations and meetings with parents/school leaders, an important factor was to scout locations and delineate study boundaries. The final location needed to be accessible and open to accepting children for an hour each day. One option was to use the grounds in the host school either before or after each school day. However, this proved to be problematic due to the proposed numbers and safety concerns with school extra-curricular activities running at various times.

For the location searchlight, the following conditions were agreed and the following criteria formed;

- The venue would be ideally located in the centre of the market centre of Bhawanigarh to ensure the safe passage of children between places of residence and location
- Central locations were seen as advantageous in promoting the trial and the necessary attendance required
- Children should be located from the same community to fulfil the conditions of randomisation required in experimental designs

### **3.6.8 Gurudwara**

A Gurudwara is a temple and focal place of worship for people who follow the Sikh religion. It is also open to people of other faiths and beliefs. Bhawanigarh's Gurudwara is centrally located and serves the local people through daily prayers, ceremonies, National/Regional events and social functions. Interest in the Gurudwara was mainly down to its central location and halls that can accommodate large numbers. It has numerous annexed rooms primarily used for religious readings and individual prayers. For this reason, the temple was not initially considered, as it was thought prayers might be disturbed. However, Gurudwara temples are welcoming places with various open social events happening throughout the year. Local and National festivals attract large crowds, and it was these reasons why the location was further explored.

Initial meetings with the management team were encouraging. The suggested trial times (2 pm – 3 pm) worked parallel to the daily routines at the temple and considered to be minimal risk to prayer-goers due to the trial running outside of busy periods. Large crowds would usually form early in the morning and then later in the evening. Guidance was shared indicating known busy days due to National/Regional festivals, and it was encouraged trials be postponed on these days. The initial trial plans were outlined and discussed receiving supportive feedback on potential risks and opportunities. Overall, the trial was looked upon favourably and mirrored the values and outreach focus tailored towards local needs. A provisional acceptance was made for the trial to run. However, regional leaders were required to confirm before implementation. Final confirmation was granted during phase two of the research.

### **3.6.9 Searching for Monitors**

To be able to run the programme there needed to be interest from children, whether acting as a Monitor from the host school or participating as a learner from the community. During the first school visit and following the searchlight phase – it was seen as an opportunity for children from the host school to feature in a monitorial role during the intervention. As the host school is an English-medium school, convenience-sampling techniques were adopted in the recruitment of Monitors. Convenience sampling is, as its names suggest, a sampling technique that is used by way of convenience (Bryman, 2015). Therefore, for this phase of the study, children were selected who are registered to the host school operating as an English medium School. The primary aim of the study is to understand if children from an English medium school can cooperatively peer teach using synthetic phonics pedagogies.

Alongside the research team, Principal and school leaders, Monitor criteria was discussed and finalised into a framework for trial entry and included the following mandatory points;

- Consent from parents or guardian
- Studying in Classes fifth -tenth Standard
- Interested in developing their English language skills and using them to help others with limited English ability
- Achieved above C1 grade in English speaking and listening formative assessments
- Native Punjabi speaker or has achieved above B1 in speaking and listening formative assessments

- Above 90% school attendance

Following meetings related to the above Monitor criteria, a request was made to visit classes to understand how children use English. Unsurprisingly, the higher the year groups, the better proficiency, and grasp in learning and informal learning tasks. Teachers would predominately instruct in Punjabi in the lower primary years and then a blend of mixed instruction in the middle and higher years.

### **3.6.9.1 Parent Communication**

Some presentations were set up by the principal researcher with assistance from the research team to demonstrate the roles of monitors. Children from classes fifth – tenth Standard were invited to attend in free periods and after school. A letter outlining initial plans for a trial involving children teaching and learning alongside other children was created and translated into Punjabi. Children showing interest were given a copy of the letter to take home to their parents. The letter detailed parent meetings ahead of the trial and further information about the proposed timings and location. Parents who were interested in their child participating in the trial were instructed to sign a slip acknowledging which meeting time they wanted to attend as part of a trial briefing scheduled at the school. Given the complexities of running a trial and intervention, it was thought that a letter might not fulfil enough detail. Meetings were thought to curb any doubts and were seen as an ideal opportunity to meet parents and answer any questions that were overseen or required further explanation. It also enabled a more fluid discussion related to the intervention – as some parents might not be able to read Punjabi or English. Developing appropriate materials and widening access is vital to the overall success of gathering research data successfully (Grosh & Glewwe, 2000). Enabling a forum for questions and answers with subjects was therefore crucial in all exchanges and partnership development.

### **3.6.9.2 Family Background Questionnaire**

The questionnaire was sent home and required parents to record their levels of income, education and language preferences to help understand the socioeconomic status of the family environment. These demographics are vital in understanding the context of those

concerned in studies that include background factors (Grosh & Glewwe, 2000). Children who returned completed slips were instantly enrolled into the Monitor programme. Some parents required additional support with completion; RTMs assisted with language support at the school site.

### **3.6.9.3 Searching for Learners**

Since the school opened in 1995, it has always been the ambition to enrol as many students as possible. However, restrictions due to land, regional and national education policies, the scope for inclusion has taken time to grow. Records indicate that each year, more children take an interest in registering with the school, but most are unsuccessful in gaining a place. During meetings with school leaders, the proposal of running a trial was appealing and played a significant part in the development of an educative outreach initiative to reconnect with the community, particularly identifying those children not registered with a provider.

As there was an initial contact with host school parents from the initial visits, plans were made to set up new meetings to understand more about the habits of non-registered children. Meeting parents from the locality always garnered interest from neighbours, thus, parents who had children enrolled in other schools, unregistered to a provider, or interested in our plans, would be easy to find and talk to. Most parents seemed receptive to the idea of children mentoring their peers and had no objection including their children into the trial. Some talked about the limited development of school choice in the area. Others focused on the difficulties in meeting transportation costs between school and home as another factor. Most parents have their children in government free schools but only attend during set times of the day due to work patterns or other family duties. This was taken into account when considering the timing of the trial. To maximise regular attendance, it was vital to understand work and school attendance patterns. Most parent meetings also explored further opportunities for I.T teaching and possible links to further education (equivalent). It was made clear that the programme was a trial and held no certification for transitions to educational providers. Nor did it substitute regular lessons or hold any value towards formative or summative assessments commonly used to assess children's abilities in core subjects before progression. Details of interested parents were recorded and stored on a password protected excel spreadsheet at the host school. Information about future meetings and final decisions about trial enrolment would be communicated via a text message or

phone call. Interested parents were also encouraged to call or visit the host school to discuss any further points about the trial.

Additional visits were made to the primary government and Khalsa school in Bhawanigarh. Both were receptive to the ideas about the trial and permitted access to speak to children and at later dates their parents. It was essential to build bridges to promote the ideas behind the trial and to generate interest alongside regular providers. The school times ran parallel with the host school and weren't seen as a potential clash. During the summer months, all providers start their first period at 7.30 and finish at 1.30pm. The trial was planned to run for an hour from 2 pm – 3 pm, permitting enough time to travel between schools and places of work before returning. It also enabled the trial to run outside of school hours and during extended breaks from work due to the heat of the day.

Alongside the numerous visits to the local areas, it was decided by the research team, school leaders and the principal researcher to advertise in areas not visited during the first phase. Posters and leaflets were designed and sent to print before the end of the visit. All parties decided not to distribute until the end of the second visit given the time difference leading up to the trial.

#### **3.6.9.4 Phase One Summary**

The searchlight phase was completed after all areas of enquiry were addressed. The feasibility of conducting a trial that identified children from the host school as peer-monitors, cooperatively teaching in their neighbourhoods, was seen as an opportunity for children who do not access mainstream education or who have limited access due to various economic and geographic constraints. Positive meetings with local parents strengthened the ideas of running an English trial. All were very accommodating and receptive to the ideas discussed. However, it was noted that there needed to be more clarity on the unique nature of running a trial and that it had no connection to current formative or summative credit within the mainstream system.

Alongside initial concepts of a peer-mediated trial, the location was seen as central to the success or failure of the planned intervention. Gaining permission to access the areas of the Gurudwara enabled the trial to take a maximum of N = 800 children without disturbing

local prayers. Given its size, multiple unused rooms, large hall and central location, the temple offered a familiarity to all residents and access.

The next section details phase two implementation and account for the planning and organisation ahead of the final trial.

### **3.7 Phase Two –Follow-Up and Pilot Work**

#### **3.7.1 Introduction**

Returning to India to start Phase 2 of the research plan came three months before the live trial. Regular Skype meetings were arranged in between visits to mediate any communications related to recruitment and other information. The sessions were also used to plan the second phase visit with the aim of arranging key targets to save time. The following action areas were established pre-visit and used to guide this phase of the research, and they included the following steps;

- Location
  - Followed-up with Gurudwara management regarding access
- Recruitment (Learners)
  - Repeated visits to local communities, schools and marketplaces to promote participation
  - Gave two presentations with parent groups at the host school
  - Issued posters, leaflets and other marketing material around the local area
- Research Team Pre-Baseline Reliability Checks
  - Conducted preliminary reliability testing with Research Team Members on oral baseline measures
  - Completed sample size calculation for final trial
- Recruitment (Monitors)
  - Finalised Monitor recruitment and ran baseline tests
  - Following baseline tests and questionnaire analysis, entered Monitors into a peer teaching synthetics phonics training programme
  - Completed a mini-pilot with Monitors and children from the host school
  - Issued an English reading attitude survey



### **3.7.2 Location Follow-up**

The Gurudwara management team acknowledged the scope of the study and its importance for children's access to learning programmes. The second visit confirmed the start and end date of the programme and detailed the plan to cater for up to 1000 children. This number was a ceiling limit due to the potential conflict with prayers and general health and safety. Further inspections of rooms and halls were observed with estimates of how groups will form in certain areas. A brief walkthrough, from entering the Gurudwara, to leaving it was trialled – tagging areas where children could disappear. From the tagged areas, zones were created where research team members and Gurudwara custodians would stand and direct children and worshipers. All children were asked not to enter the main prayer areas without adult supervision. These arrangements were finalised through a consent letter for which the Gurudwara team had time to reflect on. It was approved during the second visit and can be viewed in Appendix J.

### **3.7.3 Recruitment by Media**

As agreed at the end of the first visit, the school management team agreed to cost a small marketing campaign to promote the trial. Posters and leaflets were identified as an appropriate medium for information as they're relatively inexpensive and can be distributed in surrounding areas to attract participants. The posters and leaflets included the following information in keeping with the trial aims and scope of the research;

- Trial aims, procedures, and expectations
- Location including dates and times
- Trial demographic sample (min/max age, government/out-of-school children)
- Host school address, phone number, and email contact information
- Registration information including prior meetings and Q&A

An example of the media used in promoting the trial can be seen in Appendix L. The plan was to reach out to surrounding communities proximal to the Gurudwara and host school locations. Posters were distributed within a 5km radius of the market, which acts as a central forum for residents from the surrounding areas. We selected specific areas known to where large crowds would form, in particular, children and their families. The market was one of the primary locations as it houses numerous sweet shops, play areas and the bus and cycle station which are very populous throughout the day. Further to this, we asked shop owners

and other local proprietors if we could display in their businesses. Lastly, we recognised that most of the sample would come from either of the two large government schools at opposite ends of the town. We visited both, and each Principle agreed to take a batch of leaflets and posters to market the event.

Permission was always sought before pasting posters around town and within businesses. We faced no trouble or issue with the marketing materials. Indeed, one request from the Gurudwara management was upheld as no posters/leaflets, or related marketing is strictly prohibited on religious artefacts and physical structures, in this case, external walls. This was confirmed, and no materials were used to promote in the immediate surrounding area of the Gurudwara.

#### **3.7.4 Pilot Study Development**

The pilot study was considered to be the first practical step in initiating the final trial that investigated the effects of a peer-mediated phonics approach to learning basic English reading.

In its most straightforward context, pilot studies enable study methods, and data collection plans to be examined before going live with an initial study (Henk, 1987). Most are identified as a small-scale study of test instruments to test any flaws (Welman & Kruger, 1999) and to understand the limits and potential risks in study design (Van Teijlingen & Hundley, 2001). The checking of these processes and their implications for final live studies is necessarily a “reassessment without the tears” (Blaxter et al., 1996, p121). For this study, it was essential to understand more about the test instruments and how they might be used in the context of a trial. It was also important to gauge whether children from the host school were able to peer teach using a semi-structured teaching plan. These factors were first demonstrated through a pilot and in this section details how the final trial was structured and implemented.

#### **Pilot Goals**

The pilot stage involved two areas in the development of the research. Firstly, it considered the role of Research Team Members (RTMs) in managing tests conducted by children. RTMs would be required to perform each of the tests and record responses from each child

participant. Some tests required RTMs to record spoken sounds and words. Thus, recording responses accurately were considered a significant factor. The following questions were set up to guide this process;

- To what degree are RTMs can accurately record spoken elements of tests from child participants?
- Is there a high level of concordance amongst all RTMs on spoken test elements?

The second part of the pilot was to understand more about how the Monitors would take to mirroring the conditions of teaching letter sounds through a peer-cooperative approach. This phase included two steps following finalised recruitment as outlined in the previous section. The first phase involved a 2 and a half-week training programme delivered by the principal researcher. The Initial Mentor Training (IMT) focused on techniques for teaching commonly found sound properties in CVC and CVCC words, with the aim promoting phonemic awareness amongst the group. It also enabled Monitors to play around with sounds and discover different combinations found in words. Most of all, the IMT programme enabled each Monitor to understand how peer-cooperation can utilised for the study. The pilot was guided by the following question;

- What are the main threats to Monitors learning to share English sounds and teaching techniques in small cooperative groups?

### **3.7.5 Testing the Tests**

Following the recruitment of research team members and initial training, all of the test batteries were introduced and explained to each member with practical guidance and application. Ahead of Monitor recruitment, children based on Primary Classes 1-6 from the host school initially expressed an interest in the experiment. Unfortunately, due to children being outside stipulated age range, this wasn't possible. Instead, they were invited to take part in a mini-pilot of tests. These children (n =15) were chosen at random as a matter of convenience sampling due to the difficulty in trying to pilot test in other schools, and identified as a threat to validity.

Following approval from parents, all children were entered into the programme and pre-tested against a range of phonological, reading and native tests as the principal aim of the study is to understand more about children's pre-reading skills. To ensure that each test was recording children's responses accurately, each research team member individually assessed each child alongside the principal researcher on each of the tests. This enabled the principal researcher to investigate the strength of reliability of each research team on responses from children for each analysis. This was considered a significant step in the development of the trial, as research team members would be scoring in the final trial.

At the pre-test pilot, each child was given a chance to complete all of the tests from each battery. This was run over two days to allow for the time within free periods during school. The tests took place in the library due to its quiet environment. Dictaphones were used to record each response for every child on each of the tests. The one-to-one arrangement had reciprocal benefits, where research team members could experience giving each test for the first time and the child in giving responses to them. Each recording after each test was played to each researcher individually to limit any discussion or debate for scoring. Following completion of all of the tests, each score was tabulated into SPSS and prepared for analysis.

### **Reliability of Tests**

The Intra-class correlation coefficient (ICC) examines the amount of variance between research members recording of responses. There are six different ways in which ICC can be applied, and this is determined by study design and arrangement of data (Shrout & Fleiss, 1979). ICC is best suited for the current data set because of the employment of multiple random research team members who were assigned to record children's responses from associated tests (Hallgren, 2012). Recording test scores accurately are of paramount importance. Thus, any differences between research team members' scoring would seriously threaten the internal consistency of each test instrument and any external generalisations made after that. In light of this, it was important to investigate both individual and group reliability of test scores awarded.

For the final tests, six research team members were tasked with recording responses (correct/incorrect) from children on both reading and native tests. Cohen's Kappa (Cohen, 1960) is usually performed for pairwise analysis, that is when two raters (or in this case

examiners) observe responses and are analysed for their agreeability. There are other versions of Kappa (for example, weighted) but these data were recorded on the interval scale; thus, ICC was selected. Moreover, ICC output illustrates single and average measures of agreement. Single measures will answer the question; how accurate is a single assessor on grading student performance for reading and native test batteries? Whereas, average measures report on the same question but for every assessor within the group. It is usually acceptable to report average scores when there are multiple raters, however, in this case, the estimates of the agreement were investigated at the individual level to enable a more in-depth appraisal of whether assessors agreed and the differences therein.

For the pilot study, the ICC method adopted a two-way mixed model while applying conditions for an absolute agreement. A two-way mixed model was used as it assumes that research team members are drawn from the population and are only of interest. The absolute agreement was, therefore, a more stringent estimation of any variance from single and averaged coefficients as a consequence of scores awarded from individual members and total group. Both methods are reported in Table 1 where ICC1 represents individual units of analysis (i.e., the ICC estimates for an individual research member), and ICC2 represents the average unit of analysis (i.e., ICC estimates for the mean of all research team members).

Only tests requiring verbal responses were analysed based on the sensitivity of gauging responses. Table 1 reports the ICC results for tests from specific test batteries alongside confidence intervals to indicate where the actual mean lies.

**Table 1 Intraclass Coefficients for Tests**

<b>Battery</b>	<b>Test</b>	<b>ICC<sub>1</sub></b>	<b>Confidence Interval (95%)</b>	<b>ICC<sub>2</sub></b>	<b>Confidence Interval (95%)</b>
Reading	Burt	.98	.97 - .99	.99	.99 - .99
	Dictation	.91	.84 - .96	.98	.97 - .99
Native	Punjabi	.75	.58 - .89	.94	.89 - .98
	Hindi	.98	.95 - .92	.99	.99 - .99

Table 1 illustrates the final results of individual research team members ICC scores and including the grouped total for each test observed. Scores are rated on a scale from 0 – 1, where 0 indicates no agreement and one total agreement (Fleiss, 1975).

We consider the first column ICC1 as it reports single measures for each research team member on each of the tests. Both reading tests (Burt and Dictation) show a high level of inter-reliability when considering the estimates by Cicchetti (1994) and Landis & Koch (1977). The confidence intervals show some variance for Dictation scores, but these are marginal across each member. This can be further examined in ICC2 that illustrates the average measures of each research member on each of the tests. It shows a high level of agreement; thus, inter-reliability is excellent, indicating almost 100% agreement for the Reading tests.

The Native tests report an acceptable and high inter-agreeability on each of the Punjabi and Hindi tests, based on the estimates by Cicchetti (1994) and Landis & Koch (1977). There is a wider variance for the Punjabi test as reported by the confidence intervals in ICC1; this is detailed in the section below. However, the agreement was considered acceptable. Across all assessors, as reported in ICC2 there is a high agreement when factoring group associations. It shows that with marginal differences the group as a whole demonstrate a high inter-reliability across all tests. The Hindi test was highly correlated across both individual and group scores.

### **Follow-up and Other Tests**

Following the analysis of agreement on test measures, it was evaluated that all six-research team members had good to high agreeability when measuring children's responses on test batteries. One assessor, on the Punjabi test, showed disagreement for two student responses contributing to individual scores being low. The assessor reported a dialectal difference, which possibly contributed to the scores being lower than the other research team members. Such variations will occur in this type of testing as responses rest on interpretation. With that said, and based on the summary of results presented in Table 1, there is a high consistency amongst all research team members and assessment tasks, and any errors are taken as a limitation in this research.

### **Phonological Battery**

Children scored zero across all of the sound and irregular word (Miskin) tests. Ahead of the tests, two English teachers (with linguistic backgrounds) verified each test with the following recommendations;

- Children should complete all the tests within the same battery and not independently at alternative times
- All digraphs, trigraphs and irregular words should be attempted
- Children to be shown a brief demonstration on what constitutes a ‘sound’ in starter words ahead of testing

The conditions were adhered to and replicated at the baseline for the final trial. The advice confirmed that tests were not beyond the abilities of children even though the pilot test found zero scores across the phonological battery. Both teachers explained that phonic approaches are not taught in primary schools and are not used to accelerate reading skills.

### **Letter Matching and Schonell Spelling Test**

The Nfer test required children to find letter matches in sequences for lists of varying difficulty. These are marked out of ten, and there were no issues with scoring consistency across all research team members. Ahead of the test, lessons were observed in the early years and primary year groups. All English lessons targeted alphabet learning where children were asked to recite letter names and attempt writing tasks. Matching tasks (similar to the Nfer A) were set up in some KG and Primary classes. Instructions mirrored those of those proposed for the final test with no reported difficulties. Children were active in their responses and demonstrated a capacity to recognise English letters.

The Schonell Spelling test was administered to one Primary 3 class with no reported issues other than some questions from teachers about word meaning. One observation made during the piloting of the test was potential for children to copy answers. It was noted that there would need to be specific arrangements made for children to be placed away from their peers.

Research team members each demonstrated their capability in administering the assessment to small groups of children. Each child presented a list of words on their sheet, which enabled each assessor to grade each correctly spelled word. Spelling tests and their procedures are frequently used in the school as a method of evaluation for new words and

literacy development ahead of summative assessments. The school also runs spelling bee competitions against to promote spelling within the school. Children are very used to spelling tests, and as a result, the Schonell spelling test was no different, by design, in the way spellings are achieved.

### **3.7.6 Reading Aspiration Survey Administration**

As detailed in the previous Chapter, the survey aimed to understand children's attitudes to L2 reading. The survey considers questions linked to reading in academic and recreational settings. To make the survey more appealing to children, a cartoon character illustrated four poses to reflect the different emotions against each statement. The worksheet used both Punjabi and Hindi sentences on different sides, but during the pilot test stage, it was easier for the research team to sit with each child individually and read the sentences aloud. In preparation of this, each of the researchers was given guidance on how to administer the survey following similar instructions to their training in test batteries. RTMs were asked to read each statement aloud and time was given for the children to respond as accurately as they can to each construct. Researchers were given full autonomy to translate as accurately as they can but without deviating from the agreed questions. Upon receiving the revised survey, no researcher found any of the questions to pose a risk to multiple interpretations or deviation from meaning. Some researchers reported they repeated most questions but after a brief demonstration most reported that the children made fluid responses to each question. Monitors completed their survey independently, not requiring input or support from RTMs.

### **3.7.7 Monitor Baseline Tests and Focus Group**

Following the pilot, all Monitors were subjected to all of the test batteries before being formally enrolled in the Initial Monitor Training (IMT) programme. All eligible Monitors completed the tests during their free periods and with RMTs. The process allowed for RTMs to build on their experiences from the pilot and improve further. Monitors also completed the reading aspiration survey during school time. After analysing the baseline data,  $n = 16$  Monitors were called to take part in a focus group interview to learn more about their responses surrounding reading aspirations in recreational and academic settings.



### 3.7.8 Developing Peer Monitors Through Initial Mentor Training (IMT)

After the initial screening phase, 55 mentors were registered into the IMT programme to be prepared and trained in basic phonics training and peer cooperative techniques. The programme ran for an hour and thirty minutes over 14 days at the host school to enable each child enough time to understand the minimum expectations for experimental conditions. All children enrolled into the programme had to receive at least ten days of IMT to take part in the experiment, this was judged to be enough time for each mentor to grasp experiment rules and techniques required for successful completion.

#### Phonics Instruction

The IMT aimed to provide each mentor a foundation in phonics teaching; this required each mentor to learn new sounds daily and then successfully cooperatively learn these with other mentors to replicate last experiment conditions. The first day after induction introduced the first six sounds and actions required to blend basic English words and familiarise each mentor with how words are constructed.

The picture in Figure 5 shows a small cooperative group of Monitors during the IMT using a big book to introduce a story and first sound /s/. Peer Teaching IMT: Big Book and Initial Sounds.

Figure 5 Peer Teaching IMT: Big Book and Initial Sounds



## **Peer Cooperative Method**

The delivery of the phonics instruction followed an adaptive peer-led team method (Gosser, 2001) alongside the theoretical framework of cooperative learning (Johnson & Johnson, 1991). In doing so, Monitors are arranged into small groups that are evenly balanced, that is, all roughly the same stage in their understanding of phonics. Each Monitor learns one of the different sounds; therefore, on the first day of sound training, six students teach six different sounds and actions and learn alternative sounds/actions from their peers. The IMT schedule can be viewed in Table 2 following the procedure each Monitor participated in;

### **Step One:**

Monitors are assigned to groups. At this stage, all had no phonic knowledge and are therefore equal in their ability across each group. Each group has between 6-8 members depending on attendance on each day.

### **Step Two:**

The principal researcher introduced six sounds and actions to 6 different Monitors and asked them to share their new knowledge with their fellow group Mentors. To start with, each Monitor takes it, in turn, to orally teach the sound with accompanying actions. From here, another Monitor offers a new sound and action, and this process continues until all sounds and actions have been taught within each group. Over the 14 day programme, every Monitor had the chance to take the lead disseminating sounds and actions to each member of their group.

### **Step Three:**

Sounds and actions are exchanged and learned within the group using a variety of techniques and support learning material. Monitors familiarise themselves with flashcards for both phonemes and blending words. Time was allocated to recognise phoneme-grapheme associations. This coding became trickier as digraphs and trigraphs were introduced. They then incorporated the stories of each sound from the Jolly Phonics 'Big Books', each taking turns to read the story aloud (a maximum of 3-4 sentences) before introducing the new sound. This would eventually become the first part of the process in the introduction of new sounds on a daily basis.

**Step Four:**

Small games and techniques such as phonic bingo, hunt for the missing sound and speed rounds of 'show and tell' using the flashcards were used to develop Monitor's phonemic abilities and confidence further. Monitors were exposed to these games and encouraged to develop their ideas for sound testing in each of their groups.

**Step Five:**

After each story, sound and action had been shared; blending techniques were taught as a whole class method from the principal researcher. The blending flashcards encouraged Monitors to physically blend each of the sounds, to showcase how words are constructed. Oral segmenting was used in this regard to breaking down each word into phonemes before blending it back together. Techniques such as dragging a pen under each of the sounds, fill-up boards and physically bringing phoneme flashcards together to synthesize each sound for each word displayed. Monitors were invited to engage with these methods for remediation and then to trial back in their groups.

**Step Six:**

Monitors were then assigned to a different group to practice blending sounds and recognise the structure of sounds in each word. As the introduction of digraphs and trigraphs were implemented, enough time was given for each group to experiment with counting the sounds, using the flashcard blends to help recognise the patterns and then practice with group members. Each Monitor practiced writing the digraphs and trigraphs on the magic whiteboard sheets to mimic how they were presented on the flashcards. Again, enough time was allocated to cultivate the necessary skills.

A supported study guide (see Appendix O) was produced and illustrated the phonics approach, exposing the children to some of the techniques discussed during the IMT training programme. The booklet held key information on dates and contact information they would need to produce when taking part in the experiment. Each study guide contained information about the nature of the experiment, decoding puzzles, practice sounds and word blends for each mentor to complete in their own time. It also had a breakdown of each day to show roughly where each peer group should be up to. The guide along with the IMT

sessions gave an insight into how the experiment would evolve and the expectations of working in partnership with the learners.

**Table 2 Synopsis of Initial Monitor Training and Sound Awareness**

<b>Day</b>	<b>Objective</b>	<b>Action</b>
1	Welcome and introduce whole programme objectives	Brief students on programme overview and answer questions
2	Introduce and explore the 42 sounds of English Introduce experiment materials and tests Introduction to sound training and blending techniques	Children to acknowledge the 42 sounds and experiment with the first 6 sounds - /s/a/t/i/p/n/
3	Working in ‘peers’ workshop Sounds & Blending Training- Set Sounds 2	Children to identify - /c/k/e//r/m/d/ sounds
4	Working in “peers” and further sound training Sound & Blending Training- Set Sounds 3	Experiment with Set Sounds 3 - /g/o/u/l/f/b/
5	Sounds & Blending Training- Set Sounds 4	Experiment with Set Sounds 4 - /ai/j/oa/ie/ ee/or
6	Visit to Gurudwara and Sound/Blending Training Sounds & Blending Training- Set Sounds 5	Experiment with Set Sounds 5 - /z/w/ng/v/ oo/oo/ Familiarise with the study location
7	Sounds & Blending Training- Set Sounds 6	Experiment with Set Sounds 6 - /y/x/ch/sh/ th/th
8	Sounds & Blending Training- Set Sounds 7	Experiment with Set Sounds 7 - /qu/ou/oi/ue/ er/ar
9	Tricky words introduction	Experiment with a selection of tricky words
10-14	Mini-Teaching Events	Experiment with all sounds/peer techniques (1 hour)

### **Time Management and Pilot Learning**

As part of the initial development of the study, there was an original focus to not embed a scheme of work into the learning together method. One idea was to allow for a self-organised approach with the possibility of children joining different groups based on their interest levels and learning desires. Therefore, teaching materials, methods of delivery and timing of what needed to be covered (i.e., the number of sounds taught) would be variable for each group. However, this didn’t work during the pilot as Monitors had previously decided learning content for the last quarter of the hour. New Monitors to the group disturbed what was already established, and this led to confusion. It was a similar situation

might arise during the live trial, where late arrivals or poor attendees may join different groups and unsettle group cohesiveness. Much of the difference in learning achievement using cooperative methods is about scaffolding learning and allowing time for positive interdependence to form. Given the short timeframe in the experiment, it was decided that children should remain in their registered learning groups and not freely move around the Gurudwara halls until the delivery of core learning objectives had been covered.

### **Teaching and Learning Materials**

There are many synthetic phonic packages on the market for schools to purchase. Indeed, the DfE has a list of verified materials that have been widely used in the early years and primary classes in the UK. The list proved to be very helpful when considering the pros and cons of each product and gave clues into how each could be adapted for the study. The primary goal was to use a package that is visually enriching and considered the L2 user. Initially, it was considered to involve interactive software that would be presented on affordable tablets, phones and mini projectors for visual stimulation. There are some free online resources commercially available. However, these are mostly incomplete and would require adaption to fit the research context. Furthermore, after scouting and research planning visits to Bhawanigarh, it became clear that there were some issues to do with power availability, potential risks to disruption of prayer/worship and the general cost of shipping hardware to India.

After consultation with colleagues, it was decided the Jolly Phonics learning package would be selected based on its pace and visual representations of each day of learning. Moreover, the learning materials available appeared to be child centred with a focus on visual apparatus that children would like to experiment with. One other highlight of the package was the element of pace and the ease of adapting the scheme of work to reflect the experiment plan. The pace is key in learning to understand basic sounds, and the Jolly Phonics scheme of work enables children to start to blending from day three after learning the first three letter sounds s/a/t. This study used the following materials as developed by Jolly Learning Inc based on their practicality for children delivering the programme;

- 25x letter sound flashcards used to teach each of the 42 sounds by quickly showing each student before moving on

- 25x Blending Flashcards – cards that display words ranging from basic CVC blends to more significant and more challenging words
- 25x Alternative word cards with varying degrees of difficulty.
- 30x Big Books – large colourful books that contain large letters and sentences about each of the 42 sounds to support the development of physical movement (children are required to mimic the actions of the characters in the plot).

Other supporting materials included;

- 150x Magic Whiteboard A4 Sheets – the product is designed to be reusable with a nonpermanent marker. Each group had two sheets they used to display letters, words, and drawings created by their peer group.
- 650x Phonics Workbook – each child in the learning group was presented with a workbook that was locally produced after design. Each page reflected the associated Jolly Phonics storyline but with keywords/sentences to blend, a space to form letters and words and an area to colour in pictures.
- 1500x crayons and pencils that went with each learner workbook
- 100X actions worksheet to remind each student of the sound-action correspondence. The worksheet listed each action against each sound illustrating how to create the action. For example /S/ required each mentor to move their hands like a slithering snake.

### **3.7.9 Phase Two Summary**

The main aims and objectives of Phase two were to prepare for the intervention in Phase three. This included testing the reliability of test instruments and the continuation of recruiting children into the trial. In addition to this, Monitors completed baseline tests, following reliability checks, and were inducted into the peer-synthetic phonics programme. Monitors found first sounds and associated materials a challenge in practice; however, this was improved through confidence and familiarity. Selected Monitors also completed a pre-trial focus group based on a reading attitude survey. This provided invaluable information about reading preferences in academic and recreational settings. Data from the interviews were considered for the final trial design. Recruitment was accelerated via media and various visits to schools and residential areas. However, its final numbers were still

uncertain. This posed a major threat to maximising the sample required for detecting effects. Before the end of this phase, it was agreed a further push for recruitment would be conducted via ‘callouts’ and messages to prospective parents, acting as a reminder ahead of final registration during Phase three.

## **3.8 Phase Three – Registration, Baseline Tests, and Intervention**

### **3.8.1 Programme Registration**

A total of N = 842 children initially registered into the programme during this phase of the research. Due to the strict timing of the intervention period, it was important to get all consenting participants enrolled into the programme. The research team initially planned for two days to accommodate this process. However, based on the visited numbers this proved to be problematic. Enrolment lasted seven days but permitted late participants to join. Given the timeframe, no participant was enrolled after this deadline due to threats to validity and trial controls. However, as Watson and Torgerson (2006) suggest, successful trials require a degree of flexibility in retaining participants and note that the recruitment stage is crucial in mediating smooth transitions. Moreover, at the time, there was no way of knowing exactly how many people would turn up to finalise their place in the programme, this being typical of researching in development climates (Jament and Feng, 2007). Parents were informed not all children would be accepted due to the required study numbers, health, safety and agreement with Gurudwara management on the ethical considerations given to places of worship. Every effort was made to provide answers to questions and signpost with relevant information. One of the external benefits of the enrolment process was the ‘word of mouth’ and ‘curiosity’ it created around town. This unexpected outcome was welcomed alongside established methods from other studies conducting similar trials.

The enrolment process involved collecting basic data from parents/primary carers that included the following;

- Number of children to enrol in the programme (including names)
- Age and Gender
- Year (Standard) of education (Standard/year group)

- Language spoken at home
- Residential address
- Contact telephone numbers
- Child to be collected by (enter name)
- Packed lunch or Langar (please tick)

Parents were asked to provide proof of school registration that included enrolment cards or letters of correspondence. One of the most significant challenges in the enrolment process was age verification. Some parents were unsure of their child's age due to missing birth records. It isn't uncommon for parents not to know their child's age in India; records are usually kept at district head offices with associated fees for release. Most parents were able to give a rough estimate of the year of birth, and this was correlated with the standard of education from school reports and letters. Based on similar assumptions made in other studies, the transferal of programmes from one context to the other can present many problems. The universality of designing trials should reflect onset criteria, inter alia, by recognising potential barriers to its access, delivery, and summary. The lesson learned in this regard is recognition of evidence for enrolment with the assumption that parents would have birth records. Debriefing discussions with the research team suggested future programmes should communicate enrolment criteria clearer on any promotional materials and communications to interested parties.

All information was entered into SPSS and kept securely ahead of randomisation and following baseline testing. In return, parents were given a slip and unique identification number to act as a receipt and entry into the programme. If there were multiple children from one family, these were recorded under a single-family identification number. Parents were informed to keep the number until the end of the programme and that it wasn't redeemable or held any financial reward. One identified risk included parents selling their registration number to interested third parties. All parents were informed that this would be regarded as gross misconduct and result in an instant disqualification from the programme. One way to combat was recording addresses and telephone numbers during enrolment week. Parents were informed that there would be planned correspondence in case of any disruption or cancelled teaching days. Peer monitors would also police this, ensuring all registered children in their groups were recorded when attended. Monitors were asked to



report any new attendees with myself or research team members. Details were checked with collected enrolment data for verification.

On receiving enrolment numbers, parents were given four choices each day from the following;

1. Join an information video presentation (every 15 mins) and Q&A with a research team member
2. Parents receive assistance (if required) questionnaire with English support from research team members or completed at home and returned
3. Child to complete peer reading satisfaction survey
4. Child to complete designated test

For successful enrolment, all parents were required to complete all of the above for their child to enter the programme. There was flexibility with the household questionnaires as some parents decided to take home and complete. Language support and translations were offered to help mediate this process. Each enrolment receipt was stamped and initialled by a research team member after completing one of the above. Once all tests and questionnaires were completed (and stamped), parents would return their test log slips to finalise enrolment ahead of the final day randomisation.

### **3.8.2 Baseline Testing**

Baseline tests were completed during enrolment week and lasted for 14 complete days running from Monday to Saturday. In total, there were eight tests to complete that included the Reading and Phonological Reading tests batteries, Native reading tests, one reading attitude survey (for children) and one household survey completed by parents. All tests were mandatory and formed the requirements ahead of the peer-teaching programme itself. Parents were informed throughout the process that failure to complete all tests would result in automatic disqualification. Due to study population numbers being high, some tests ran over stipulated times and therefore carried over to the next day.

The tests themselves were collectively known and delivered in 'batteries.' Test batteries refer to a group of tests collectively testing one or more skills. They have been prevalently used in psychological testing (Silverman, 1990) due to the accuracy required when

prescribing interventions in clinical trials. Administering one test and then making assumptions based on performance will be greatly improved if other data were available that tested a similar characteristic. Having only administered one test in isolation before and making generalised claims is limited in power. It's perhaps obvious and insufficient to have one test that can accurately answer a single complex research question. Indeed, having multiple tests improves this as researchers aim to triangulate results, confirm overall effectiveness and suitability. Moreover, tests that maximise accuracy and consistency in understanding children's skills, abilities and mastery are unequivocally the primary aims of any educational intervention's architecture.

### 3.8.3 Test Scheduling and Management

A total of n = 842 children completed all of the tests ahead of the randomisation phase. Testing lasted 14 consecutive days apart from Sundays. The final test schedule, including the overlap of specific tests, can be found in Table 3 below. Up to 6 rooms and one hall were utilised (on the same day) within the Gurudwara surroundings. Each room could accommodate up to 40 adults, however, a maximum of 30 children per room and 50 in the hall were allocated at each time and depending on the type of test administered.

Table 3 Test Schedule

Task	1	2	3	4	5	6	7	8	9	10	11	12	13	14
<b>Nfer &amp; Miskin Tests</b>														
<b>Schonnel Spelling Test</b>														
<b>Burt Reading Test</b>														
<b>Hindi &amp; Punjabi Tests</b>														

Separate rooms were allocated to ensure all children had an equal chance of sitting a test with limited disruptions. One initial concern with test scheduling was to do with potential contamination (Keogh-Brown et al., 2007) and cross over effects amongst children. Most tests ran concurrently throughout each day, mixing children who've completed a test with those who were yet to start one. Furthermore, due to time constraints with getting through

all of the tests and reacting to unscheduled (and unannounced) regional holidays, religious gatherings at the Gurudwara – time was limited to separate into morning and afternoon slots. Moreover, two tests (Schonell and Burt Tests) were ran at the same time to ensure children had adequate time in completing each of the tests to the best of their ability. The potential for children who sat a test the previous day meeting up with children who were about to complete a test was identified as a minor threat to validity. Even though the tests were prepared in a way to limit answer sharing, children could relay information about test experiences or what to expect, thus, leading to potential small advantages for children about to sit tests. To mitigate potential contamination threats, plans were made ahead of the baseline phase to locate and direct children to different areas of the Gurudwara. As seen in Table 3 the crossover of the spelling and reading tests, running between days 5-7, resulted in test locations being at opposite ends of the Gurudwara. Parents and children were asked to visit different locations where the tests took place and this offered a logistical challenge to the team. Further, the Gurudwara is a place of worship and so it was essential that the inner prayer and worshiping areas were not disturbed. This being an agreement and documented in the ethical planning of the study. In the planning of transitions and potential routes parents/children took to testing rooms, signpost stickers and help from Monitors were on hand to help this procedure. Moreover, parents received a room number specifying location and approximate times tests would last. The research team directed parents/children to test locations after registration and answered questions related to test locations.

To begin with, the phonics and character test battery was run over the first four days from the beginning of the week. Also, the Miskin Nonsense Word blending test was administered as the last of a four tests completed over four days. Nfer A (letter matching) and B (phonemes) were completed in three days with Nfer C (blending 2/3 words) and Miskin Nonsense Word blending finishing completed on the fourth day. Before testing, a brief demonstration was given to each child with instructions given in Punjabi on a draft copy and not used on the final test paper. RTMs were instructed to be open, encouraging and friendly to each child when conducting the tests. They were asked to be discrete when marking (no ticks, crosses) final papers with scores tallied at the end after the student had left. It was natural that children wanted to know how they performed on each test but scores were not disseminated until study completion. However, supportive feedback was given to encourage holistic participation with the entire programme.

Out of the six research team members, 2 members were allocated one of the Nfer tests on the first three opening days. This enabled a quicker transition between each test in the battery. It made sense to include the Miskin test to follow the Nfer C due to the sounding element, this also provided children with enough time to recognise any sound patterns in words, should they have missed them in the first testing of sound. Four research team members each were tasked in delivering it, while the other two members finished the Nfer C on the last day of testing in this battery.

The Schonell Spelling test was part of the reading battery that required children to spell words from a list read by research team members. To maximise time, 3 rooms were isolated to accommodate up to 30 students in one sitting, thus taking 90 students per hour. The research team members considered this manageable. One and quarter-hours was allocated for each test to allow for instructions, questions about the test and time completing it. At baseline, most tests lasted between 45 minutes to an hour, following the instructions as presented earlier. Following each test, volunteer students and Monitors would assist the children to their bags and parents.

The Burt reading test ran parallel to the spelling test in an effort to ensure each child had enough time to complete it. At the planning stage, it was identified as one of the harder tests due to its length and language difficulty. The test involved a one-to-one scenario with each child who was requested to read a progressively difficult word list. This required planning to ensure each child had a fair chance at the test. However, due to the cross over with the Schonell spelling test, half of the research team members were otherwise engaged. The test ran for 6 complete days with the remaining team members joining on the 4th day. Post baseline discussions indicated that future testing should consider preparation time as a collective instruction, rather than 5-10 minutes before actual testing. In this case, given the opportunistic dimension of the RCT, it was perhaps overseen due to uncertainty of numbers. However, future studies should factor in the inner structuring of tests before subjects are tested.

The Native tests were selected last due to their perceived ease based on the pilot test results and conversations with teaching professionals from the locality. Four days were planned to complete the tests, with allocation of two days for each test. The first two days involved

the Hindi Word test, which included a list of progressively difficult words, similar to that of the Burt test. As the words list was predetermined and agreed with Hindi teachers locally, there was no set guidance on test procedure. It was agreed that test administration would follow the same principles of the Burt equivalent. That is, allowing for each word to be read in turn before moving onto the next. Batches of ten were scored and when children had more than 6 consecutive wrong answers the test would stop. This decision proved to work well in the pilot testing and initial development of both native tests. Unlike the Burt test, the Hindi word test was able to run quicker. This was partly down to test performance and the structure of allowing all of the research team to take part in the testing procedure. It was originally planned to follow-up with the Punjabi dictation test, since it was seven sentences long. However, the pilot test results suggested that some children found this confusing, due to similarity of both Hindi and Punjabi scripts. To reiterate the design and function of the RCT in respect of test selection, the primary outcome of all tests was to ensure that each child had a fair chance at completing each one, without confusion of instruction or other misleading factors.

The Punjabi dictation test was the last test battery and concluded the baseline phase. It was introduced on the 13th day where all research team members were allocated to its functioning. Similar to the English dictation test, the Punjabi test was administered on a one-to-one basis. One issue that arose was the number of children waiting to be tested, this being particularly popular in the morning sessions when large groups formed. Future testing should consider placing accurate times for each child and parent to subscribe to. This was considered in the early development phase, but it was uncertain exactly how many children would continue to take part in the testing phase and the external events happening at the Gurudwara that sometimes cut short testing times. It was perceived that the Punjabi dictation test would be one of the most straightforward tests based on the pilot results and the socio-linguistic preferences in the local area, thus, fortunately, there was limited disruption to testing procedures with on-hand support and guidance from Monitors, research team members, and Gurudwara management.

#### **3.8.4 Post - Baseline**

The process involved in the planning, development, and application of the baseline tests were run in accordance with the original plan and agendas as discussed. To finish the

baseline-testing phase, all registration cards and test cards required collecting and verified by research team members. This finalised the phase for participants and automatically entered their verification number ahead of the randomisation phase.

As stated at the pre-enrolment stage, children who didn't complete all of the tests were unable to participate in the programme. This was rearticulated on a number of occasions throughout the two weeks. A total of  $n = 82$  children were excluded at this stage of the programme due to a number of factors. The majority of these children had failed to complete all of the tests by not showing up on the allocated test days, while others had trouble meeting the study inclusion criteria as originally stipulated on various media and communications. Upon reflection, the attrition rates were considered to be low as based on the overall participation rate of over 800 children who attended the testing period regularly. The research team looked to accommodate as many children as possible but the standards were set before testing, and it was difficult to remove or be flexible with all criterion.

A debriefing session concluded the testing phase with indication on next steps and a question/answer forum. This process is detailed below at the randomisation stage of determining groups for the trial.

### **3.8.5 Randomisation**

Once all of the enrolment questionnaires had been completed and verified, a new spreadsheet was created ahead of the randomisation phase. Parents were reminded at the debriefing session that groups would be determined by random allocation. This was managed by doing a concealed randomisation to protect the internal validity and reduce bias of selection whether intentionally or by chance. Concealed randomisation is a process of minimising bias and improving upon simple randomisation techniques where children (in this case) do not know which group, whether control/intervention, they will be allocated to (Torgerson & Roberts, 1999, Mayo-Wilson, et al., 2013). A previous review of RCT randomisation trials estimated that those that didn't report concealment procedures from subjects/primary investigators accelerated the treatment effect by 41% when comparing to trials that reported concealment procedures (Schultz & Grimes, 2002). It was, therefore, important to conceal the processes of how groups were selected to everyone associated with the study, including the principal investigator.

The block factors were predetermined based on preliminary questionnaires conducted at enrolment. Parents and children were required to supply types of identification that demonstrated they met pre-trial eligibility criteria (language, gender, home address, age, the standard of education) as discussed previously. The blocked stratification technique ensures there is equal balance represented in each group as different factors could bias the statistical analyses completed at post-test and comparisons between intervention and control groups. To limit this, an individual stratification technique was employed (Kerry et al., 2005). The study used three factors (gender, the standard of education and locality) to determine control and intervention groups. Stratified randomisation was selected over simple randomisation techniques purely because the representation of this trial was varied and unequal groups could lead to biased outcomes. Education Standard reached was significant because it identifies those potentially as completing a level of education. Having more children from the lower Standard in the Control group compared to children represented at the higher levels could lead to a bias representation of how well each group did over the trial period. This assumption builds on the thread that those children who've reached higher levels of education will have had more time studying English. This is an assumption as children in India who progress into the middle primary year classifications take no tests. Thus, it could be seen as a process of maturity rather than academic ability. Nevertheless, this was one factor that was discussed and controlled for. Second, gender was included. Most trials will factor gender (unless the intended outcome doesn't relate to gender differences) to ensure parity across groups. In this trial, gender was a factor in effects of the trial and was considered a vital block to include. Finally, the locality was identified as a factor based on the two central locations the respondents came from. This wasn't factored initially to be a block and only came to light once all of the registration forms were collected. It was noted from the registration questionnaires that there were two areas parents recorded as their addresses. Bhawanigarh is a large densely populated area with many sub-areas and co-linked villages connected by various roads and alleys. Most of the children were recorded as living within Bhawanigarh/Kakra (Kakra being the adjacent village) or Gandhi Nagar, which is located to the South of the main highway road from Bhawanigarh to Sangrur. Using each unique person code, all participants were recorded into SPSS and cross-tabulated against each stratum: gender and current primary standard (year of education) and locality. No further information was included such as baseline test scores or factors that could lead to bias selection. It is often advised that a third party who

is not part of the study conduct the randomisation process due to the non-interest in group selection and study goals (Sullivan, 2011, Torgerson & Torgerson, 2008, Torgerson, et al., 2013). Furthermore, no parent was told, or information shared based on the blocking factors. The host school's computer teacher acted as a non-invested stakeholder in the study and was thus identified as someone who had no interest in the allocation of particular subjects to specific groups. He was supplied a new separate spreadsheet that detailed registration numbers and block factors required for successful distribution. No names, date of birth, test scores or other factors were shared with him or anybody else to help limit any selection biases that could have a profound effect on the trial.

The randomisation process was conducted by the computer teacher who applied the =Rand() function using excel based on the data we sent him. This is completed by setting up rows of columns based on the following stratum: standard of education (coded lower primary/upper primary), gender type (coded male/female), and locality (coded Bhawanigarh/Gandhi Nagar). The techniques allow for arbitrary numbers to be assigned before sorting them and then randomly selecting into control and intervention groups. As the trial required equal groups based on the stratum, a blocking technique was used. The host school's computer teacher used the blocking technique BAAB that has six permutations for both groups (A = allocated to intervention group, B = allocated to control group), which used the following sequence for determining group selection;

AABB, BBAA, ABBA, BABA, BAAB, ABAB

Table 4 illustrates the stratification descriptive statistics based on the Standard of education, Gender and Location. Within excel, using the ABBA method, each stratum was selected and thus represented in either a control or intervention group. The final groups were weighted by approximation due to the stratified nature and blocking technique used. The final groups included n = 378, intervention group and n = 382 control group, with n = 82 participants excluded from the final randomisation (see Figure 6 below). Torgerson et al. (2013) acknowledge stratified random assignment outcomes will almost always produce unequal numbers per grouping, this is simply based on the blocking techniques and conditions studies put in place. In the present study, there were six blocks used which required careful consideration for group allocation. The small difference in final group sizes was thought not to be large enough for unequal differences that could lead to bias results.



Instead, the rigorous method employed in determining groups was considered to make the overall inferences more reliable and thus generalizable.

**Table 4 Stratified Conditions for Random Allocation**

<i>Stratum</i>	<i>N</i>
<i>Standards 1<sup>st</sup> – 4<sup>th</sup></i>	
Males (Gandhi Nagar )	93
Males (Bhawanigarh/Kakra)	89
Females (Gandhi Nagar)	84
Females (Bhawanigrh/Kakra)	92
<i>Standards 5<sup>th</sup> – 9<sup>th</sup></i>	
Males (Gandhi Nagar)	115
Males (Bhawanigarh/Kakra)	124
Females (Gandhi Nagar)	127
Females (Bhawanigarh/Kakra)	118
Total	842

### **3.8.6 Group Announcements and Feedback**

Following the group allocation randomisation, the computer teacher met with the lead researcher and research team with the final group allocations. These were grouped by registration number and illustrated in list format. No details of children’s age, registered address, baseline tests or other identifiable data were shared. One finalised list was provided to the lead researcher that only contained a list of registration numbers that were listed under control and intervention group columns. The school Principal checked all blocking stratifications independently to ensure that each stratum followed the blocking sequence and that there was a genuine attempt to equalise each group. Once the final checks were completed, the last lists were used on the orientation day, which, for study purposes is considered to be ‘Year 0’ reflecting the first-day children were allocated into their groups. The orientation schedule is detailed in the next section.

### **3.8.7 Orientation – Year 0**

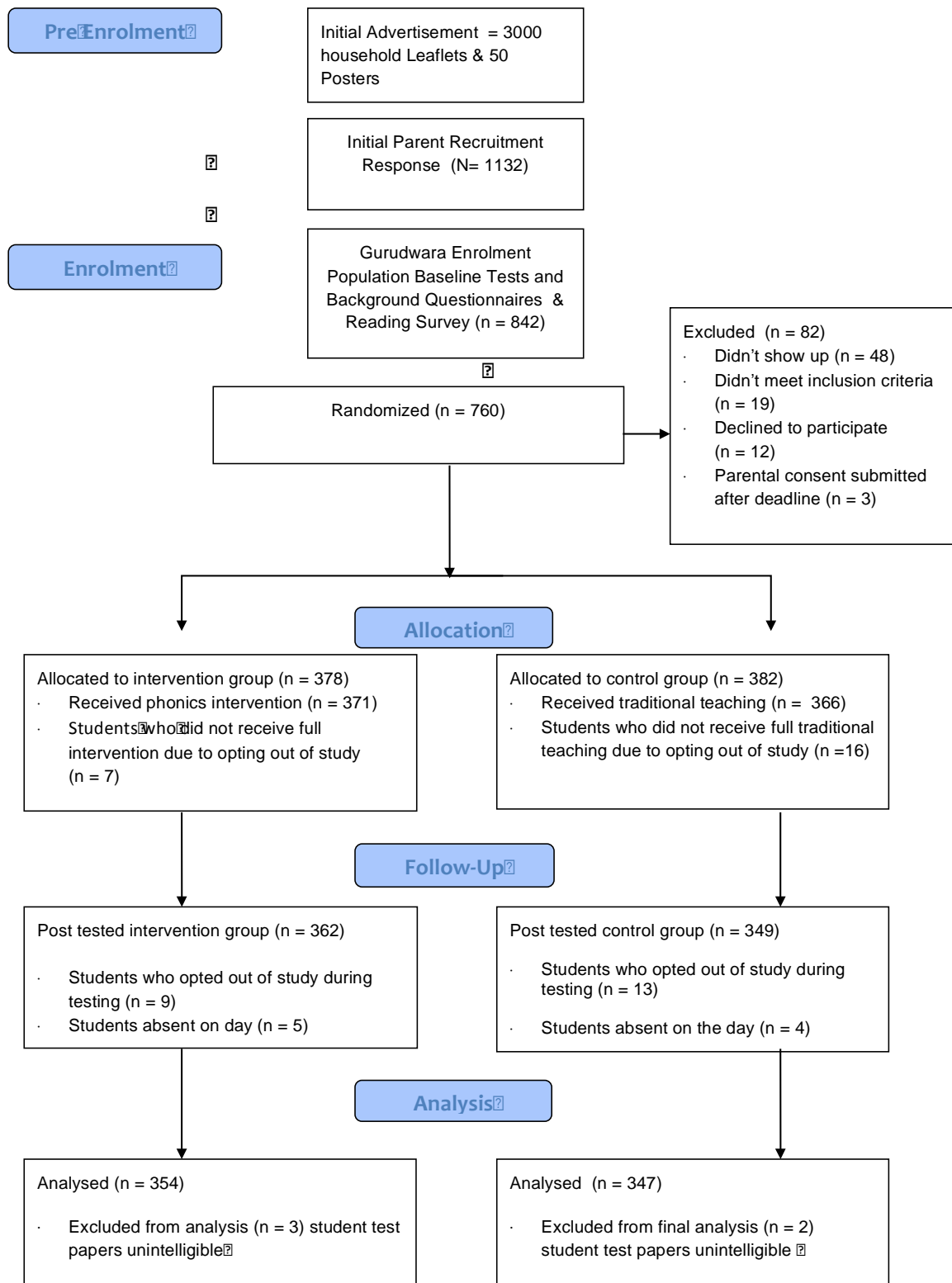
Following two weeks of baseline testing and two days of randomisation and planning of study locations, the first day of the trial was an orientation day which enabled study

participants to be allocated into groups and move to their essential study locations. The two study locations included the Gurdwara hall and two government school sites in each of the sample population areas (Kakra and Gandhi Nagar). This enabled all registered children to access to their usual schooling environment (control group participants) or the Gurdwara (intervention participants).

Orientation day ran from morning into the late evening due to parents and registered children's work and duties. Again, to reflect the difficult circumstances of running a trial of this size in a rural area that has basic facilities, the trial needed to be as flexible as possible to enable full participation and minimise attrition. Parents arrived at various times throughout the day and evening to receive information on which group their child was assigned to but most arrived in the morning. Parental feedback included questions of study commitment and duration each day, although this was agreed at preliminary meetings and enrolment, some parents did not have the chance to attend. Further guidance was given about the study duration, finishing times and information about the core aims and scope. There were some requests for children to be grouped due to the proximity of their addresses and moving between school/workplace and home. These requests came from parents who were travelling across from the Gandhi Nagar area, because of the busy double dual carriageway that acts as the central access route across the Punjab. This was identified as a potential risk to children travelling to and from the venue. Thus, it was decided to use two of the host schools' buses to move between locations. Pick up - drop off points and times were arranged with those who required it. Given the efforts expended on providing guidance throughout the registration period, baseline testing phase and trial initiation, it was crucial not to lose numbers to logistical problems.

Some parents didn't attend the orientation day but were contacted by telephone in the evening for instructions. Other parents decided not to join, or there were complications with providing enough evidence for entry. Figure 6 details the study flowchart as required by CONSORT standards in RCT reporting, it details sampling, recruitment, randomisation, allocation and finally analysed samples in each study arm. The next section further explores the attrition rates and general participant information at each stage of the study.

Figure 6 CONSORT Flowchart of RCT Profile

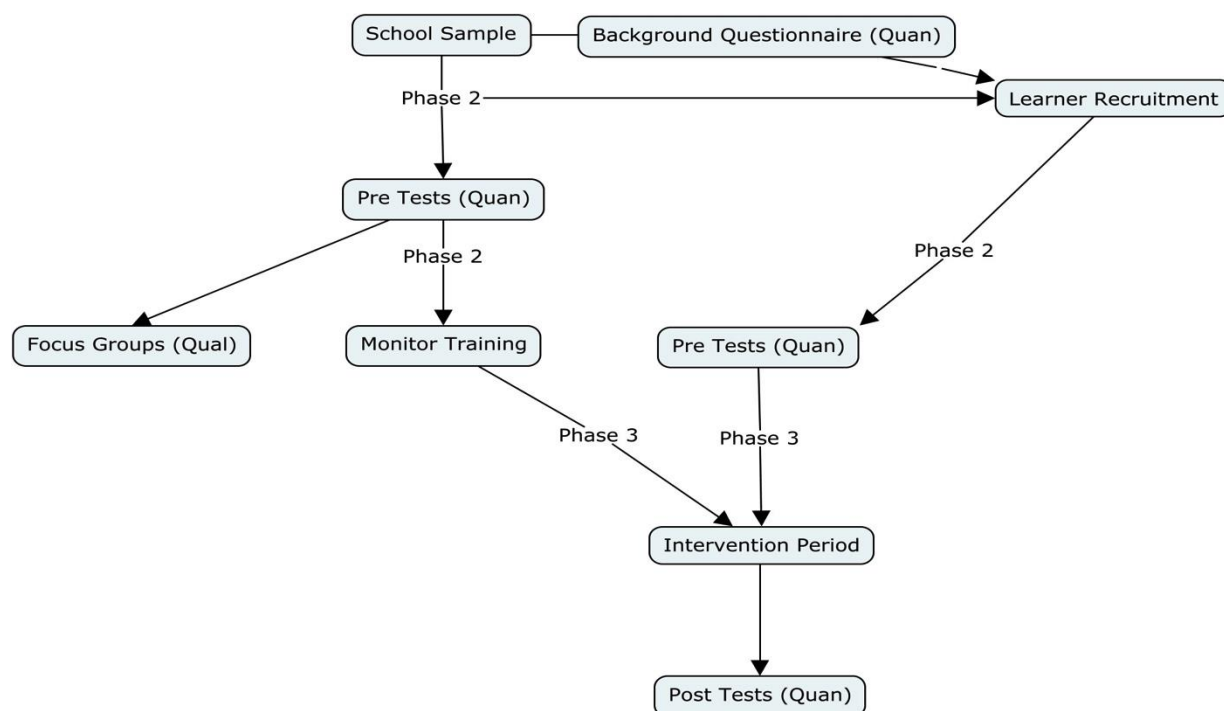


### **3.8.8 Losses and Exclusions**

The attrition rates were quite high at the beginning of the study due to the difficulties encountered in enrolling children into the programme. This smoothed out leading up to randomisation and final allocation. A total of  $n = 82$  parents were not randomised, out of these,  $n = 12$  parents decided to withdraw from the process; this was communicated via text message, mobile or face-to-face contact, or via a member of the research team. Further,  $n = 19$  participants couldn't produce evidence that met the study criteria as outlined previously. In some cases it was challenging to make an informed guess about birthdates or evidence supplied regarding registered address didn't conform to the geographic parameters detailed in first communications and media. One lesson learned from this process was that other trial designs might have been selected to allow for a wider catchment. However, to enable a balanced approach ahead of the randomisation process, these decisions were vital in the protection of reducing bias. Finally,  $n = 3$  parents didn't meet the final deadline and were omitted from the study altogether. The decision to omit these participants was a methodological one and based on threats to validity. If participants joined the intervention/control groups late, it could lead to a positive/negative bias. Further losses to control ( $n = 16$ ) and intervention ( $n = 7$ ) during the initial trial phase were down to many factors including, family relocation, commitment to work duties and general withdrawal. At post-test some children were absent (intervention,  $n = 5$ , control,  $n = 4$ ) on the day of testing. Efforts were made to find a suitable day during the post-test phase, but children couldn't take part due to other commitments. Matched pre-test data wasn't included in the final analysis. Finally, due to torrential rains over a period of about three days, some test papers were ruined (intervention,  $n = 3$  batches, control,  $n = 2$  batches). Fortunately, the pre/post test batches of lost papers were the only identified as unintelligible. After all responses and follow-up calls were made, the next stage included the intervention period which consisted of peer-cooperative phonics teaching for the intervention group at the Gurudwara and regular teaching (native) instructions from the schools found in Kakra and Gandhi Nagar areas of Bhawanigarh. The next section details the results from the intervention period and supporting evidence of reading attitudes at pre-trial.

Figure 7 illustrates the various arms of the study and trajectory of multiple data collection points.

Figure 7 Conceptual Schematic of Study Trajectory



The choice to use this design permits the study to explore exploratory factors associated with pre-trial L2 reading attitudes. The embedded design (Creswell, Plano Clark, 2011) used focus group data (Qual) derived from pre-test data (Quan) as secondary support ahead of the intervention period. This allowed for a more in-depth understanding of pre-trial aspirations while enhancing meaning behind pre-test surveys.

### 3.8.9 Phase Three Summary

Phase three was the final stage of the research plan, which culminated in the completion of pre-tests (for learners), the intervention period and post-tests. One of the challenges in successful delivery included enrolment and retention of participants, this mirroring other reading interventions in developing climates (Davidson and Hobbs, 2013). Every effort was made to include participants who met the study criteria and to provide regular communications, where applicable. Baseline testing took longer than expected, a knock on effect from the complications at registration. Baseline measures, randomisation and group

allocation were followed by the start of the intervention. Teaching was conducted for one hour every day except Sunday's and holidays. Some pictures from the intervention can be viewed in Appendix P. Attendance was recorded at 75.4% for the scheduled peer teaching days, and this included Monitors and Learners. Following the intervention, all children were post-tested with support from RTMs. Based on the learning from the baseline tests; 14 days were allocated to allow for completion. All data were then analysed using statistical procedures, the final results of which are presented in the next Chapter.

## Chapter Four Results

### 4.1 Introduction

To reflect, the purpose of this research was to study the effects of a peer-mediated phonics reading intervention and understand L2 reading attitudes with children from a rural village in North India. Based on the scholarly work in the area of peer cooperative pedagogy and the impact synthetic phonics has on early stage readers, it is hypothesized that this type of learning can potentially lead to improved outcomes for children's holistic reading development, compared to, traditional methods of learning commonly found in India. Therefore, utilising a random control trial framework, children were given a battery of English word, spelling, blending and reading tests before and after a four-month intervention period. Children from a low-cost private English medium school were selected to deliver the intervention, which consisted of teaching letter sounds and blending techniques for word construction. In guiding the study, the following research questions asked;

- Is there a statistical difference between male and female Monitors' baseline reading, phonological processing, decoding and native language test scores?
- What are the effects of Monitors' phonological processing, reading and spelling abilities when guiding learners in a peer-mediated intervention?
- Is there a statistical difference between male and female learners' baseline reading, phonological processing and native language test scores?
- Are there statistical differences between learners' reading, phonological processing and native post test scores when controlling for pre test scores? If so, where do these differences lie?

Specific to understanding more about English reading attitudes, the following research questions asked;

- Is there a statistically significant difference in baseline academic and recreational attitudes, when considering the average (neither satisfied, nor dissatisfied) test scores?
- What are the effects of Monitors' recreational & academic reading attitudes, before and after a peer-mediated trial?
- Are there differences in Monitors post recreation, academic and total reading aspirations based on gender when controlling for pre test recreational, academic and total composite scores?

- What are the main factors concerning Monitors' reading attitudes in academic and recreational settings?
- Are there differences in control and intervention group's recreational, academic and total English reading satisfaction scores?

The study employed three groups within a total population of  $N = 745$ . The first group were selected from a low-cost private English medium school in the host village. These children formed the Monitor group ( $N = 44$ ) and given basic phonological training ahead of the intervention. Further, local children  $N = 701$  were pre/post tested and were randomly allocated into control ( $n = 347$ ) and intervention (354) groups. Ahead of the intervention period, family background questionnaires were submitted to all groups with the aim of understanding the socio-economic and educational climate. These results are presented first before the trial findings.

## 4.2 Monitors Socio-Economic Background Vignette (N =44)

### 4.2.1 Introduction

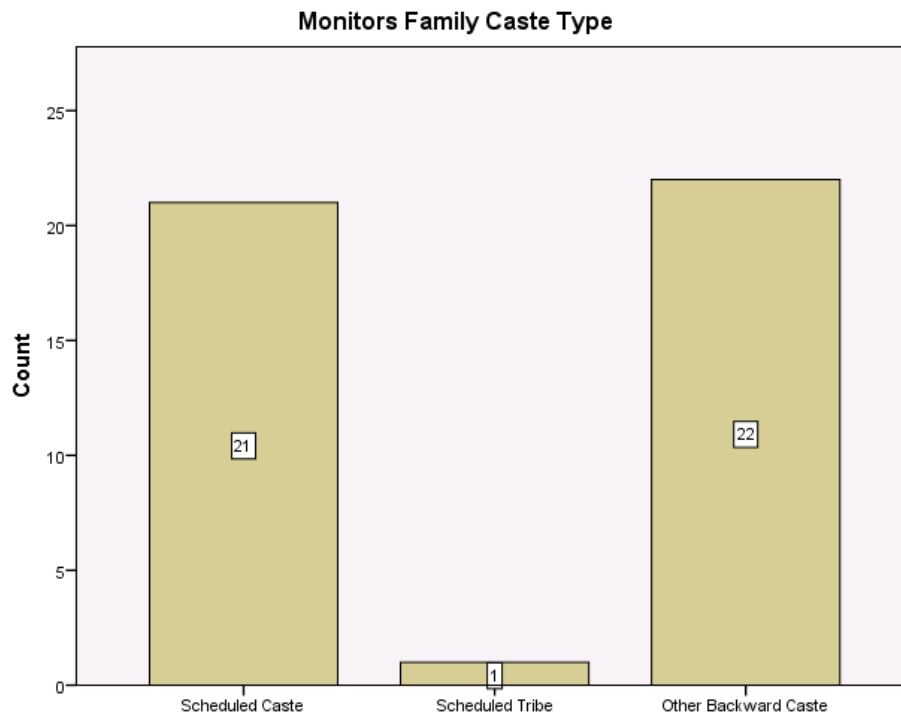
In this section information about the socio-economic backgrounds of the Monitor and Intervention/Control groups are reported. Given the disparities in sample sizes, two groups were formed in the analyses. In this regard, the Monitor group was analysed separately from those who acted as learners (and controlled learners) in this research project. The first section reports descriptive statistics on the socio-economic data gathered from the Monitor group. It considers background information related to household income, education and preferences related to learning English outside of the mainstream provider.

### 4.2.2 Socio-Economic and Family Findings

Figure 8 reports the percentage of monitor's caste type as identified from the family questionnaire. There is almost an equal split between Scheduled and Other Backward Caste type.



Figure 8 Family Caste Type



Both these categories recognise that these children are from poorer backgrounds and that 1 monitor is part of a tribal community. These caste categories are represented in the Indian Constitution and are often referred to as a collective understanding on the social economic status of groups of people throughout the country. The average income was reported at being RS 11,198 each month from the Monitor sample. This roughly equates to around RS 2800 per week or RS 400 (\$6 a day). It suggests that income levels are well above the poverty indicators as set by the World Bank (2003). One potential clue in understanding more about income distribution is home ownership. Figure 9 shows that just over half of the responses (55%) reported to owning their own (ancestral) home. One assumption is that these types of homes may be owned outright and with no rents or loans to be repaid. Ancestral homes are often passed down to the next generation.

Figure 9 Home Ownership

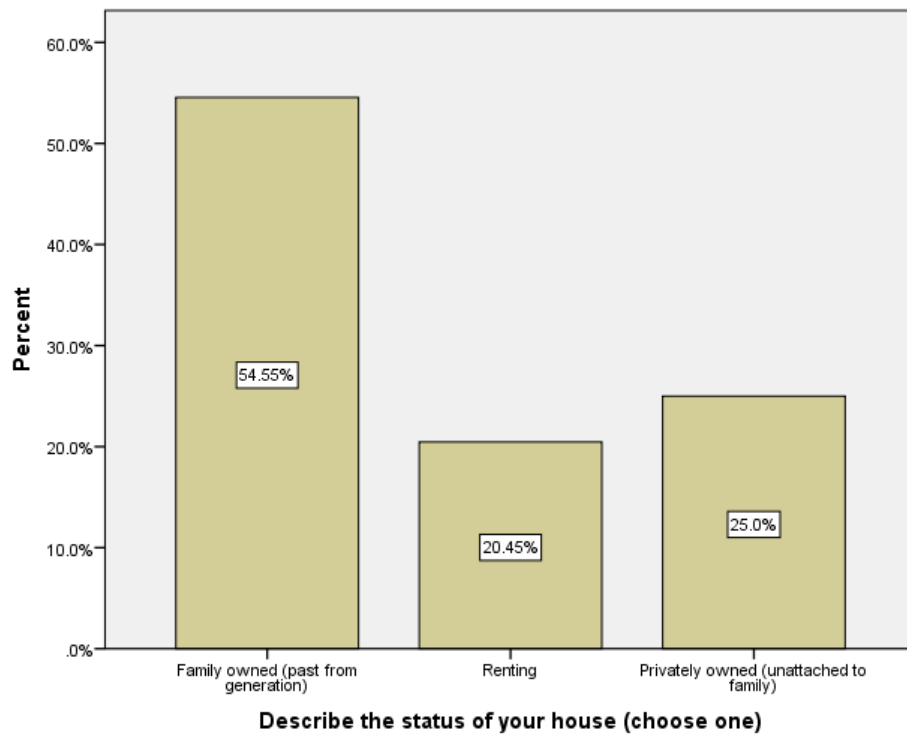


Figure 10 illustrates the job type of the primary carer for each household. Male carer's reported being employed with 36% reporting working for a private company, 34% as being employed in a government job and 29% as being self-employed. No male response was recorded as being unemployed from the sample.

Figure 10 Father's Job Type

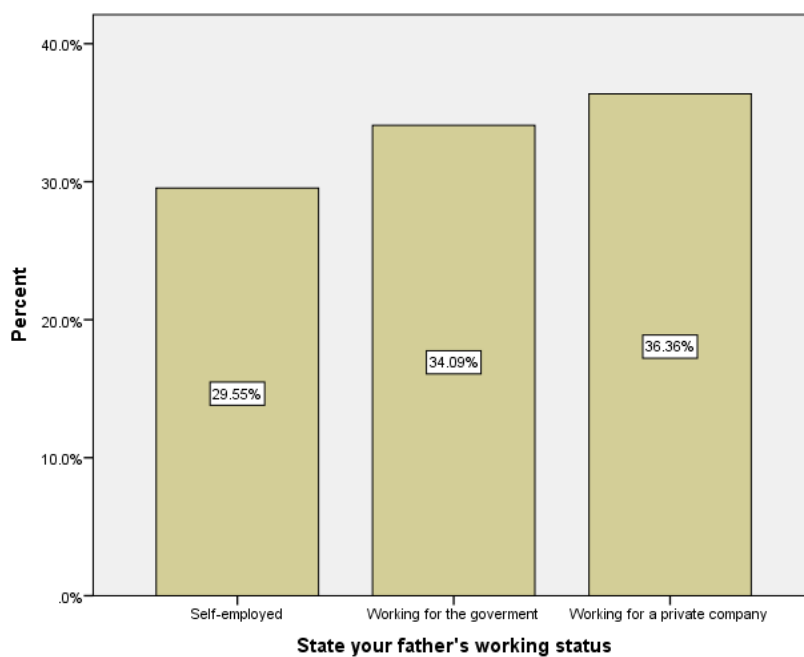
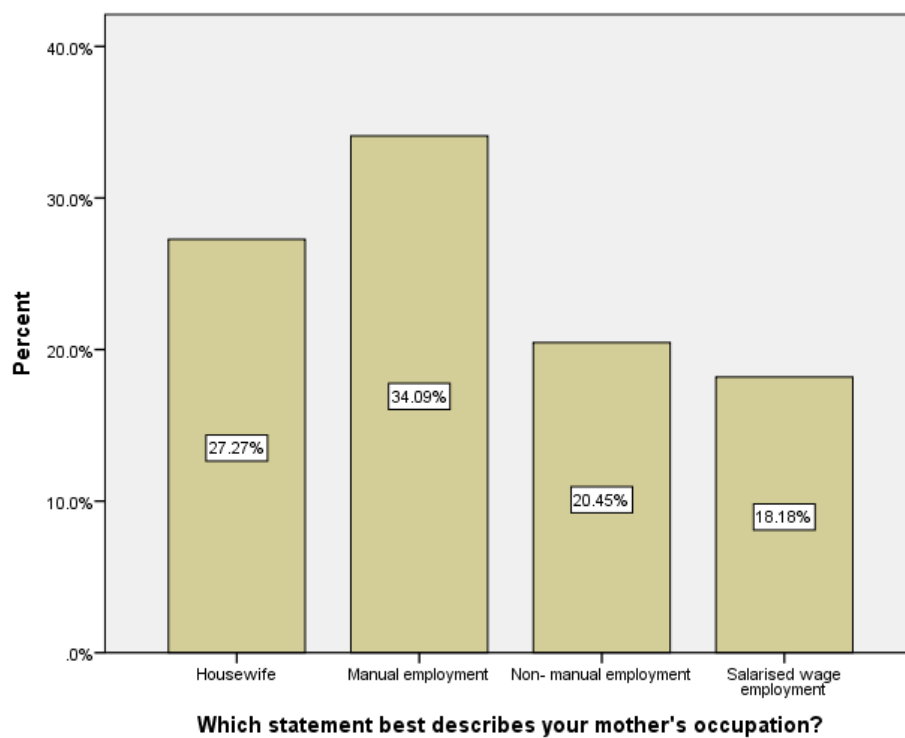


Figure 11 illustrates the job type of the primary (female) carer for each household. It shows that 34% of female carers reported being in manual employment, 20% in non-manual jobs and 18% in salaried jobs. However, unlike male carers, 27% of female carers reported being a housewife and not in employment of any kind. This isn't too uncommon for females in a typical Indian family setup, however, as data suggests, low – income families will typically work in ad – hoc roles to supplement household incomes

Figure 11 Mother's Occupation



All responses indicated that Punjabi is the main language used in the household with Hindi spoken as an alternative. Nevertheless, 100% of responses reported Punjabi as the primary language used in the home. In addition to both native languages, it was reported that at least one other person in the family could read, write and converse in the English language. Similar to the native languages, 100% of the sample reported this from the questionnaire. Further, over half of the responses (82%) identified as having English speaking family members in the household who read to Monitors in English. Data reveals that there is an English presence in the majority of households, however, it isn't clear how much time is spent speaking in English or time spent reading in English to Monitors.

Figure 12 illustrates the highest level of qualifications gained by the primary male carer. It suggests that at least 52% of male carers achieved a University level standard (undergraduate) qualification. All responses from the questionnaire reported Primary education was achieved with the lowest level of education recorded as a vocational or diploma (16%). The majority of Universities in India and higher education institutions are English medium, suggesting that male carers requiring a good level of English language ability in order to progress at the highest level.

**Figure 12 Father's Education Level Achieved**

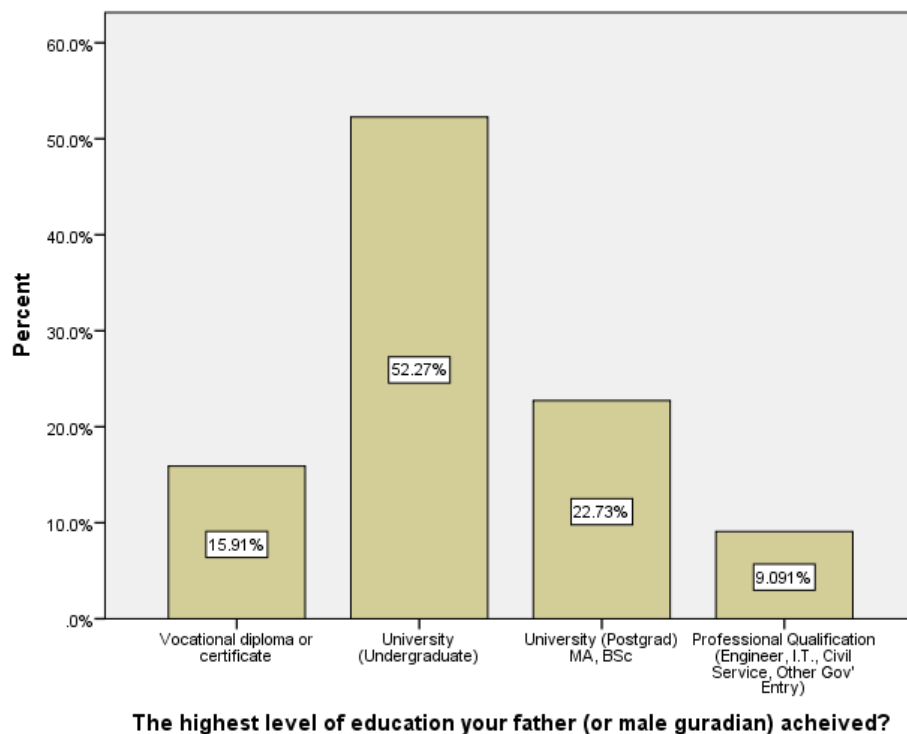
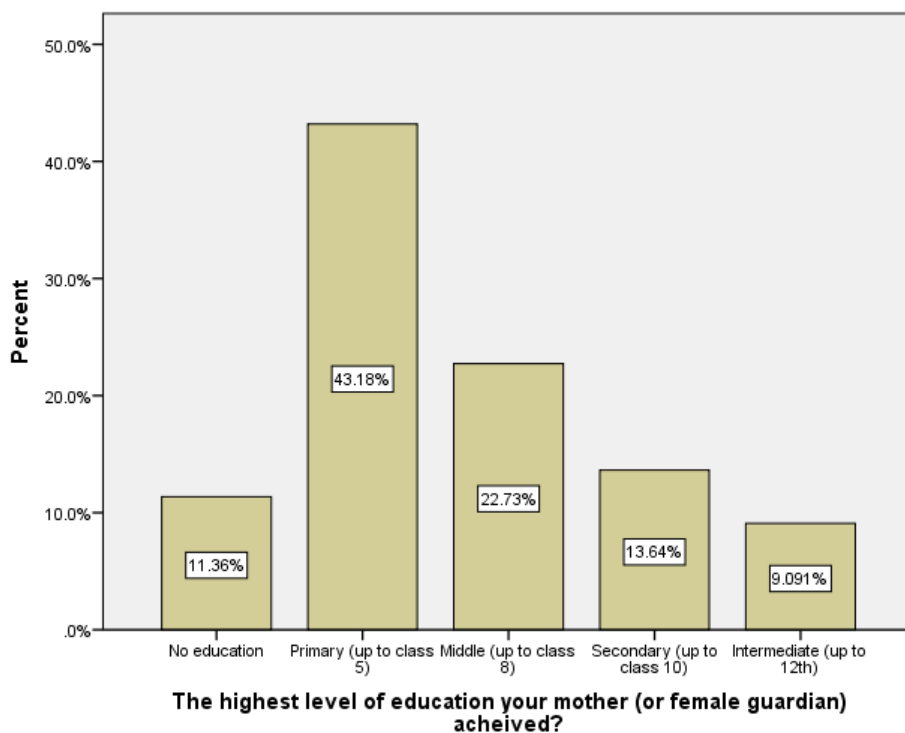


Figure 13 illustrates responses from female primary carers on their highest level of education. It reports that 89% completed some form of education with 11% of the respondents reporting no education. Deeper inspection reveals disparities at different levels. For example, it shows that 43% of female carers achieved Primary Class 5, whereas, 9% of respondents reported they only reached 12<sup>th</sup> Standard. To give this context, 12<sup>th</sup> standard is similar to completing an A-level and is often completed in registered schools or at centres that have affiliated status to educational Boards. While the majority of female carers reported receiving an education, it cannot be ignored that the levels are very low in comparison to their male counterparts.

Figure 13 Mother's Education Level Achieved

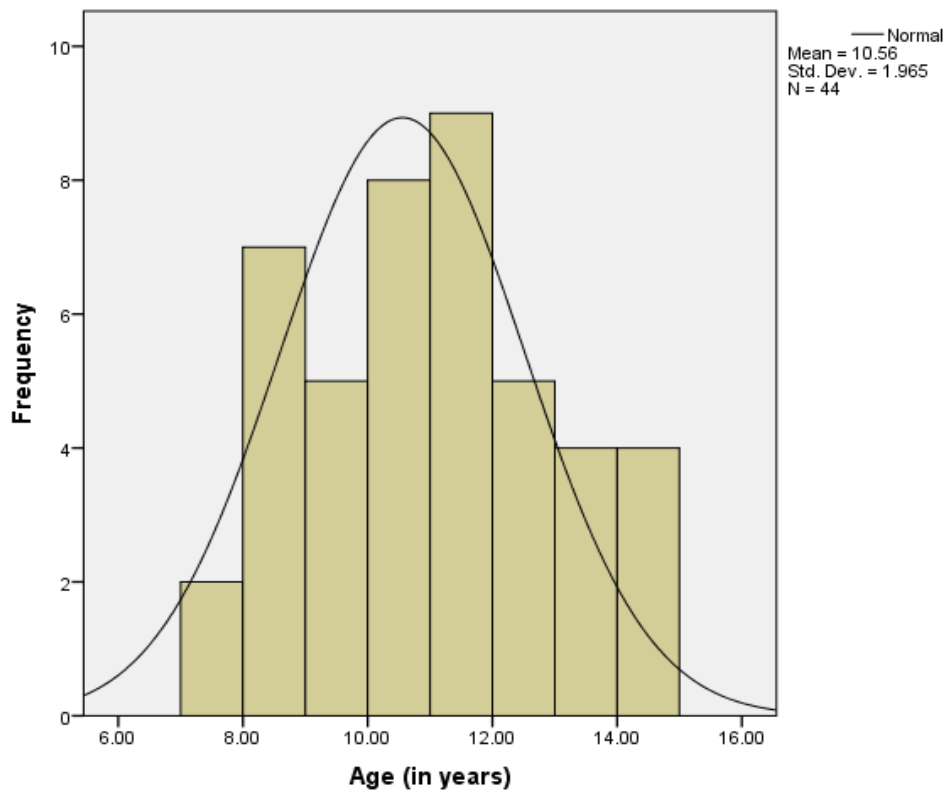


#### 4.2.3 Monitors' Characteristics and Education Inputs

This section considers the education inputs from a Monitor perspective. It reports the findings of Monitors' educational experiences outside of their mainstream provider.

The final group consisted of 23 males and 22 females, an almost equal split. All students were at the time registered as being full time students at the host School. Further to this, all students reported living in and around the areas connected to Bhawanigarh. Figure 14 illustrates the age frequency of the finalised monitor group with the majority of students being aged 10 and 11 at the time of study commencement.

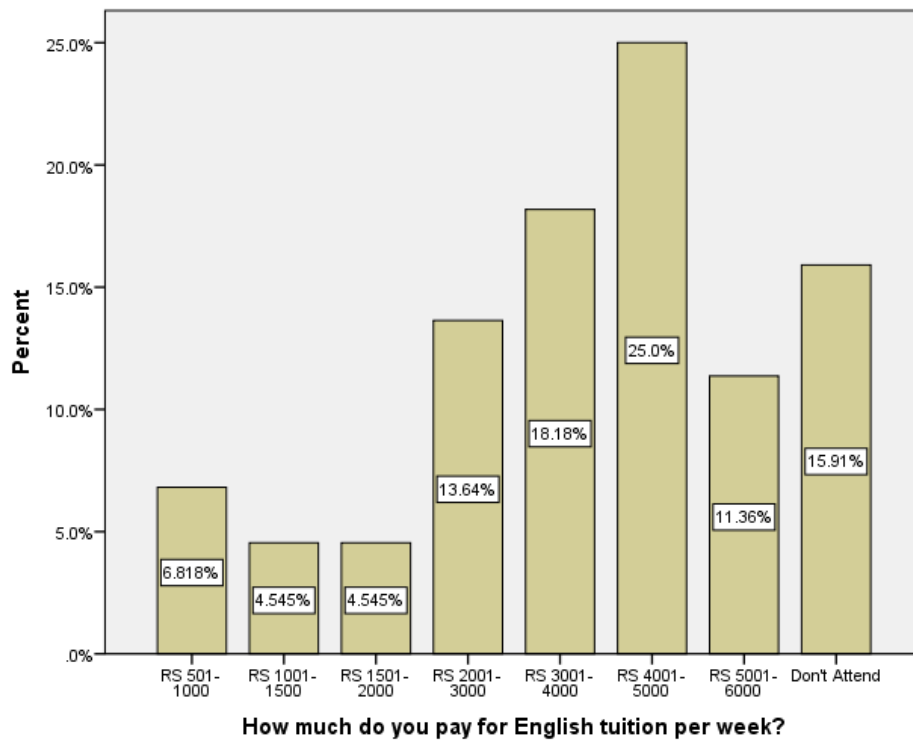
Figure 14 Distribution of Monitors' Age (in years)



It is a habitual for most CBSE registered students to receive fee-paying tuition outside of their main provider. These are usually locally accessed at teacher's private residencies or at specific academies or training centres. All students reported that they are in some way engaged in some form of tuition. This is represented in various modes of delivery such as teachers providing paid support in their home environment or students travelling to an academy to receive support and guidance. Based on the responses 84% of Monitors recognise with fee-paying tuition, which based on this sample means that 37 out of 44 students are attending after school tuition.

Tuitions are regarded as an extra household cost and not part of the regular provider of education. None of the Monitors reported receiving free or concessionary places. Figure 15 shows the range of fees paid on a weekly basis and illustrates that 55% of the monitors reported paying above RS. 3000 per week. At the time of writing, this roughly translates into £30 per week.

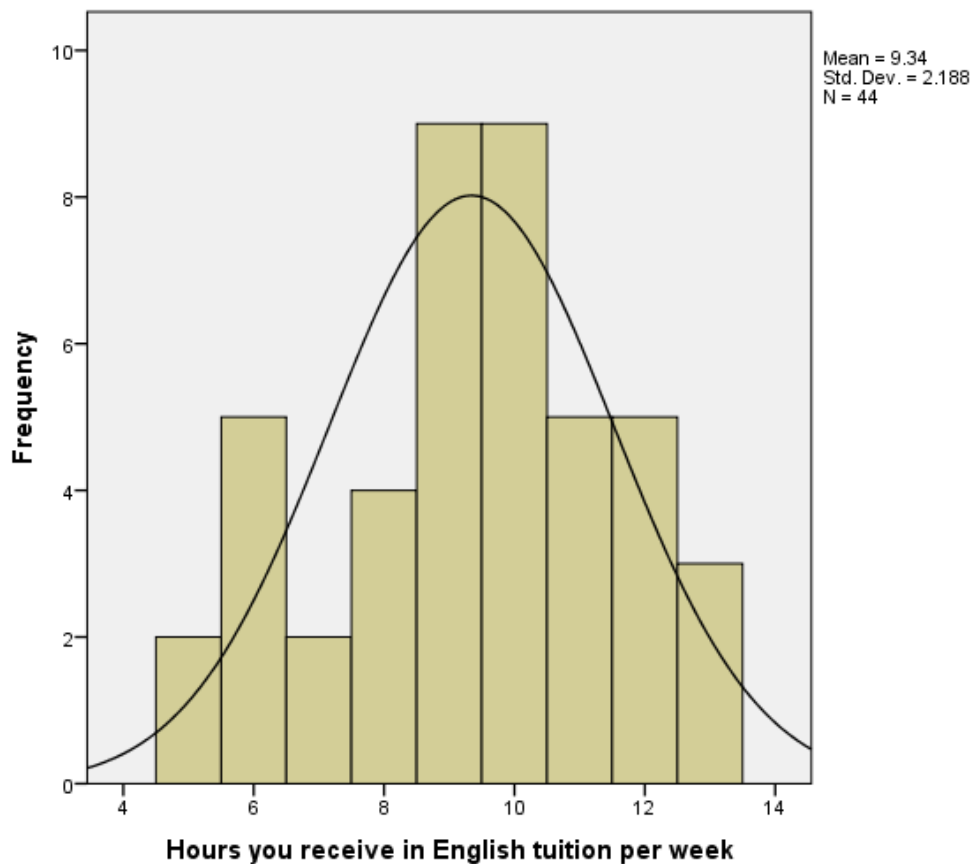
Figure 15 Percentages of English Tuition Fees



Sixteen Monitors reported paying between RS 4001 – RS. 6000 (£40 - £60) per week, which is way above the fees paid to their regular school provider. To give this perspective, at the time of writing, fees paid for higher year students was no more than RS. 5000 per term. Of those who attend English tuitions, 25% pay between 4001 – 5000 RS per week. The cheapest tuition cost for English was reported as being RS 501 – 1000 per week. However, as Figure 14 suggests, there are disparities in terms of cost to the student.

The average amount of time spent at tuition is 9.34 hours per week. As reported in Figure 16, 31 Monitors (70%) are spending above 9 hours at English tuition. If we factor in the usual time spent at school which is roughly 6 ½ hours and include the 6 day school week, then for the 84% of students who attend tuition will be adding 1 ½ hours per day on average to their learning week.

Figure 16 Number of English Tuition (in hours)



### 4.3 Learner Socio-Economic Background Vignette (N = 701)

#### 4.3.1 Introduction

The learner sample represented N= 701 of the targeted questionnaires. All questionnaires were completed at the registration phase with support from the RTMs and the lead researcher. Data were analysed to understand more about learners' socio-economic and educational inputs.

#### 4.3.2 Social Economic and Family Findings

The average monthly income reported RS 4768 for the learner sample. This equates to on average RS 1192 per week, or RS 170 (\$2.63) per day. These data suggests that the average daily earnings are just over the poverty indicator threshold as recognized by the World Bank (2003). Furthermore, the sample responded on homeownership, indicating 88% of families living in rented accommodation compared to 12% who own their own ancestral home.



In terms of employment, data suggests that 99% of respondents are employed. Of these, 59% recorded themselves' as being self-employed with 40% as working for a private company and 1% recognized themselves' as being unemployed. The majority of primary male carers (99%) work in manual labour jobs. Although it wasn't identified from the questionnaire responses, low – skilled manual labour jobs are vast in the Punjab, typically involving agricultural work and engineering assistants. From the female responses, it was reported that 100% of female carers recognize themselves as being a housewife and not receiving any form of income.

#### **4.2.3 Leaners' Educational Inputs**

Responses on the question of the primary language used in households found that 97% of households use the State language of Punjabi, with others recognising Hindi and Urdu (1%) as their primary language. Some missing data (incomplete forms) represented the other 1%.

English wasn't recognized as a primary language in the household, only 6% of the sample population reported reading in English to children in their households.

The results from Figure 17 suggest that half of the learner sample's male carer's didn't complete basic education (50%). Further to this, only 24% completed up to class 5 and 18% up to class 8. The results suggest attainment levels are very low in terms of receiving a basic education. No respondents mentioned that they attended higher education with the highest attainment being a vocational diploma with 6 responses (less than 1% of the total sample). The graphic displays 3 respondents as achieving 12<sup>th</sup> standard (less than half a per cent) and compared to vocational training certificates this would seem strange and not a natural progression. However, online education and accreditation from work programmes are becoming more prevalent for a range of jobs.

Figure 17 Highest Level of Education (male carer)

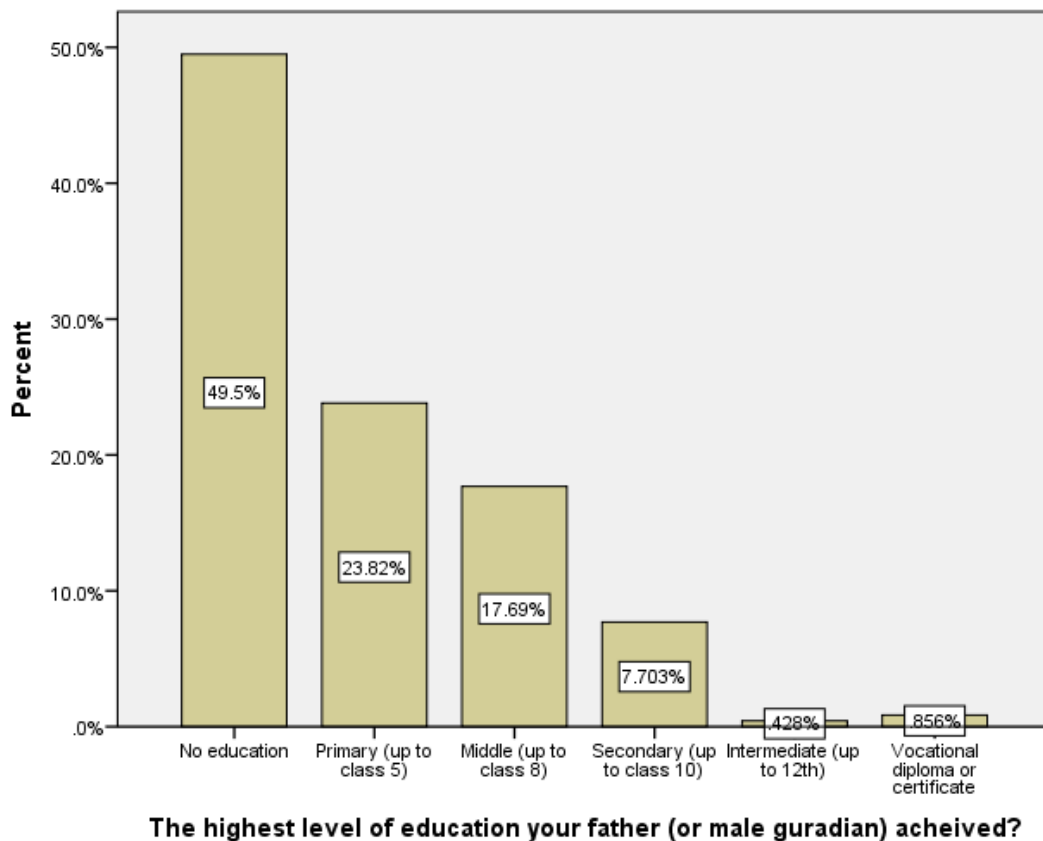
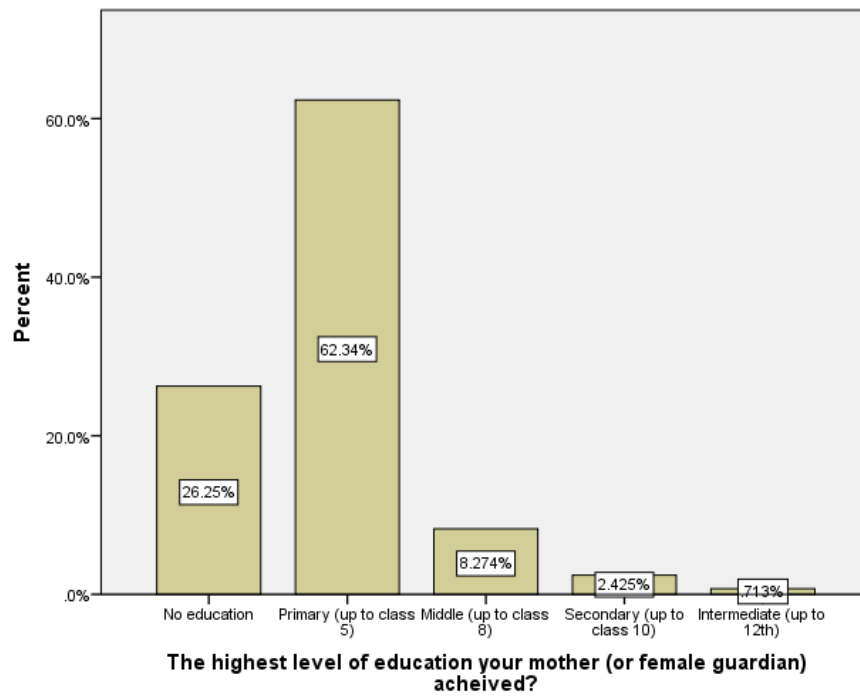


Figure 18 illustrates the responses from female primary carers in relation to their highest level of education. It reports that 26% of female primary carer's did not achieve a basic education, whereas, the average was achieving up to primary class 5 (62%). Other responses included 8% for middle school with less than 5% of the total sample of female carer's achieving secondary schooling (2% up to class 10, less than 1% up to class 12). Similar to the primary male carer's responses from the same household, no female carer reported going to University or receiving accreditation for work.

Figure 18 Highest Level of Education (female carer)



#### 4.3.4 Summary

The results from the family background questionnaires found clear differences between samples. The level of education, job type and obvious differences in tuition indicate disparities between groups and their access to English. However, these results must be taken at caution, as they don't directly compare both groups. Analysis was conducted on each group separately, given the obvious differences in size. Instead, what can be learnt from the background questionnaire is; Monitors have regular contact with English compared to the Learner sample that has minimal exposure. Furthermore, English is identified in Monitors' households and is accessed during after school tuition. While this cannot be generalised to the whole population, the descriptive statistics presented here give a basic understanding of how different groups access English.

## 4.4 Monitors Attitudes to Reading English in Recreational and Academic Settings

### 4.4.1 Introduction

Ten questions formed the recreational construct and asked children about their reading aspirations in non-academic situations, such as being at home, with family/friends. In general, these questions were adapted from the ERAS survey that asked about how much children liked to read English books for pleasure. The following research sub-question was used to guide the study;

- Is there a statistically significant difference in baseline academic and recreational attitudes when considering the average (neither satisfied, nor dissatisfied) TEST scores?

### 4.4.2 Monitor Pre/Post Recreational & Academic Scores

#### Baseline Reading Attitude Results

Firstly, reliability checks were conducted and had a good level of internal consistency as determined by an alpha of .76 for the pre recreational construct and an acceptable level of internal consistency (.63) for the pre academic construct. Table 5 displays the Monitor groups' pre/post-test scores for reading aspiration in academic and recreational constructs.

Table 5 Monitor Pre/Post Recreational and Academic Scores

Variable	N	M	SD	Minimum	Maximum
Pre Recreational	44	32.09	4.43	26	40
Pre Academic	44	22.34	5.20	14	40
Post Recreational	44	35.43	3.55	20	40
Post Academic	44	33.18	4.54	10	38

The pre recreational mean ( $M = 32.09$ ,  $SD = 4.43$ ) suggests a good level of satisfaction amongst the Monitor group and was higher than the pre academic ( $M = 22.43$ ,  $SD = 5.20$ )

mean. A one- sample t-test was run to test if satisfaction scores were statistically significantly different from the average satisfaction responses (neither satisfied, nor dissatisfied), this reflecting (and defined) as 20, the middle value in both construct scores. The test value of 'indifferent' was used in the final results. Both pre (academic/recreational) satisfaction scores were normally distributed as assessed by graphical boxplots and illustrated no outliers greater than 1.5 box lengths. Results found that the pre recreational satisfaction scores ( $M = 32.09$ ,  $SD = 4.42$ ) were higher than the indifferent score of 20.00, a statistically significant positive mean difference of 12.09, 95% CI (10.74, 13.43),  $t(43) = 18.085$ ,  $p = .00$ ,  $d = 2.7$ . Taking the very large effect size and overall positive difference, the main inference from the Monitor group would suggest that pre recreational satisfaction to be positive, that the children enjoy reading English books for fun and unrelated to academic study. To understand more about the pre academic satisfaction scores, another one-way sample t-test was run to test for statistical differences between Monitor's pre academic satisfaction levels and the test value 20 (indifferent satisfaction). The results suggest that the pre academic satisfaction mean score ( $M = 22.34$ ,  $SD = 5.20$ ) to be slightly higher than the indifferent test value, a statistically significant mean difference of 1.27, 95% CI (.0524, 2.4930),  $t(43) = 2.103$ ,  $p = .041$ ,  $d = .01$ . The low effect size, while statistically significant, suggests that satisfaction levels are not practically significant from the indifferent test value used in this analysis. It infers that satisfaction levels of reading academic books before the intervention are neither towards being satisfied or unsatisfied.

Using the same data, a paired samples t-test wanted to investigate if attitudes changed over the intervention period. The following sub –research question asked;

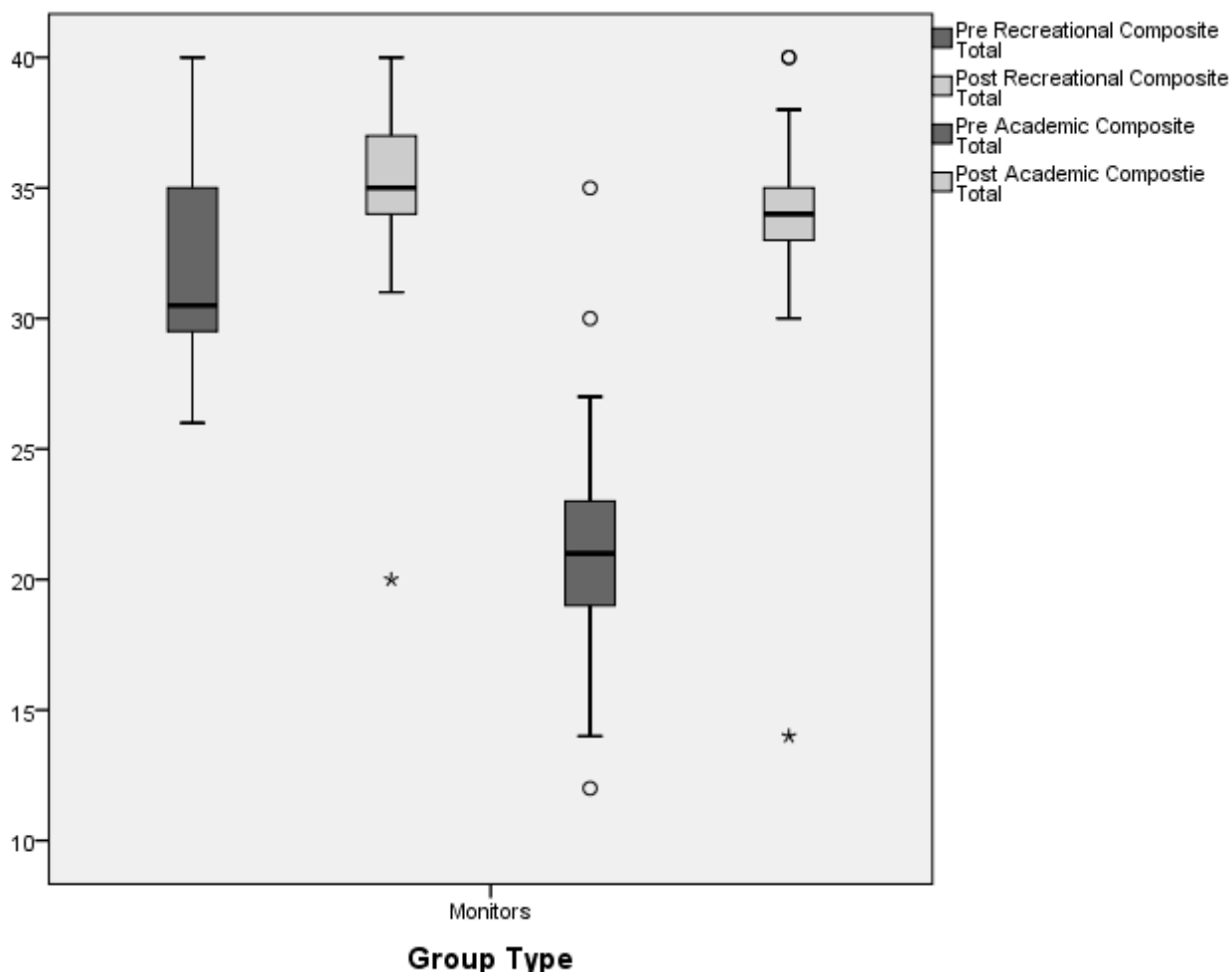
- What are the effects of Monitors' recreational & academic reading attitudes before and after a peer mediated phonics trial?

### **Pre/Post Reading Attitude Results**

Firstly, internal consistency tests were run on the post-test constructs of recreational and academic satisfaction scores. Both constructs had good internal consistency as determined by Cronbach's alpha of .73 (post recreational) and .77 (post academic). Inspection of distributions of normality were investigated and found data to be normally distributed by approximation. Calculated differences (post tests – pre tests) were used in the analysis to look for any plotted points that didn't follow a normal distribution. One outlier in the

academic constructed difference scores was revealed. Usually, outliers can pose a threat to the overall spread of data, especially if they are frequently distributed outside the main cluster of data. In this case, however, they were not considered to be extreme and was therefore kept in the final analysis. The paired samples t-test results suggested the Monitor group improved their pre recreational English reading satisfaction ( $M = 32.09$ ,  $SD = 4.43$ ) when compared to post recreational English reading satisfaction ( $M = 35.43$ ,  $SD = 3.55$ ), a statistically significant mean increase of, 3.34, 95% CI (1.471, 5.210),  $t(43) = 3.604$ ,  $p = .001$ ,  $d = 0.8$ . The large effect size suggests that during the intervention period, the Monitor group improved their recreational reading satisfaction scores and the results are statistically significant over the two time points. The second paired samples t-test was run to investigate if differences occurred between the pre and post academic satisfaction scores. Originally, the pre academic mean scores were indifferent, thus, indicating responses are neither satisfied nor unsatisfied when responding to questions about reading English books in academic situations. The paired samples t-test results suggested that the pre academic satisfaction scores ( $M = 21.27$ ,  $SD = 4.01$ ) when compared to scores reported on the post academic satisfaction survey ( $M = 33.79$ ,  $SD = 3.81$ ), there was a statistically significant mean increase of 12.52, 95% CI (11.11, 13.92),  $t(43) = 17.988$ ,  $p = 0.00$ ,  $d = 2.7$ . The extremely large effect size indicates that Monitors improved their satisfaction levels of reading English academic reading following a peer mediated English phonic intervention. The mean improved over the time period suggesting there to be a more positive outlook on academic English reading.

Figure 19 Boxplot of Pre/Post Composite Scores

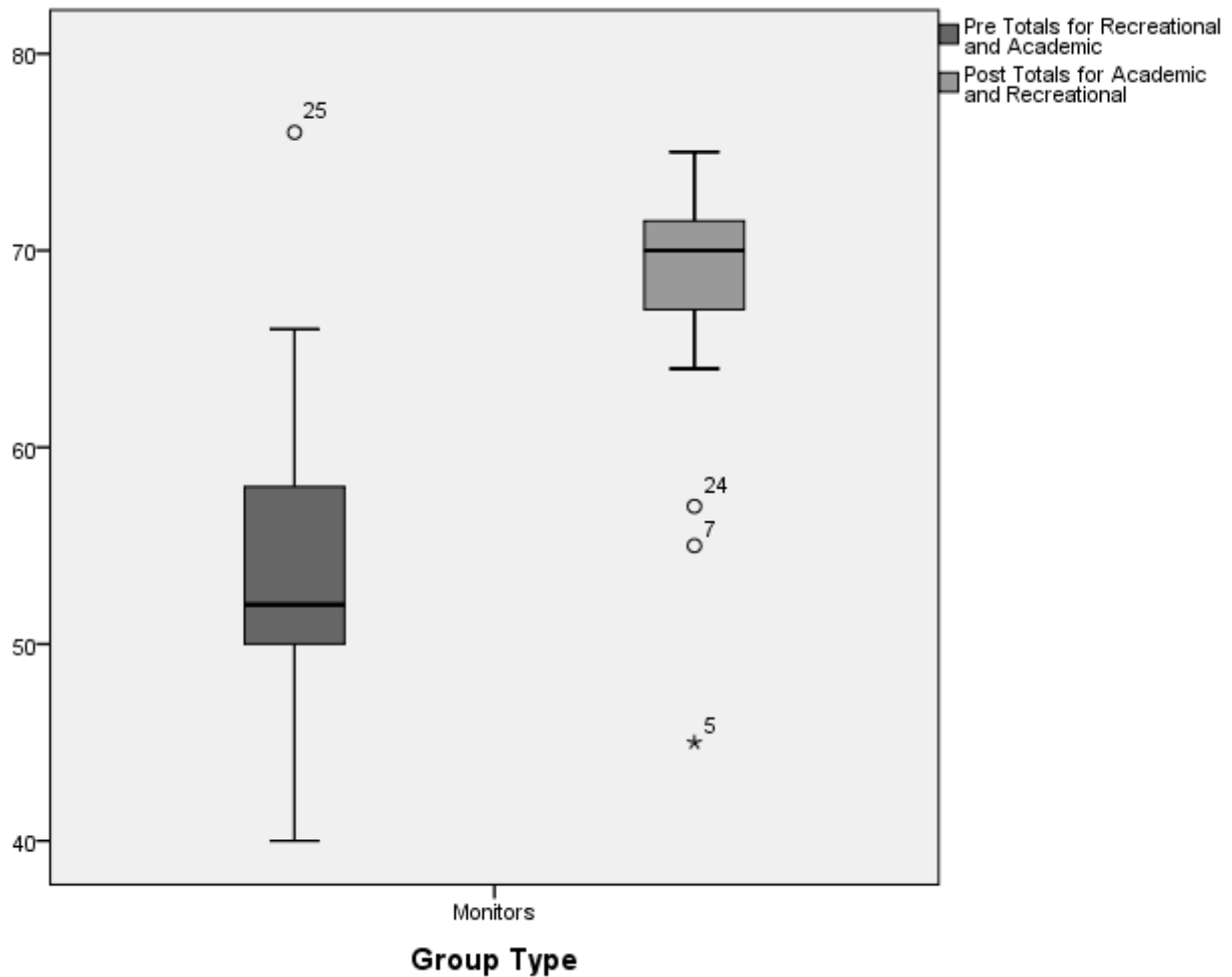


The boxplot illustrates the gains made in the mean scores for both recreational and academic satisfaction levels pre/post intervention period. It shows that there was a large gain in academic satisfaction at post-test and the spread of data to be fairly uniform towards being very satisfied with regards to reading English books in academic situations. The median also improved for recreational satisfaction at post-test.

To further understand the magnitude of these gains, a total score from both composite scores was calculated by summing the total pre recreational scores with the pre academic scores. The same process was also completed for the post-tests and a new variable created that measured the difference. Tests of normality were once again performed and found data to be normally distributed. Figure 20 is a boxplot that illustrates the total summed differences of pre recreational and pre academic total satisfaction scores compared to post

recreational and post academic total scores. The observed outliers are not considered a threat to the internal validity based on limited number falling outside clustered boxplots.

Figure 20 Boxplot of Recreation and Academic Attitude Scores



It shows that there is a noticeable increase from combined pre totals ( $M = 53.72, SD = 6.90$ ) compared to combined post totals ( $M = 68.61, SD = 5.46$ ), a statistically significant mean increase of 14.88, 95% CI (12.332, 17.440),  $t(43) = 11.755, p = 0.00, d = 1.7$ . The large effect size suggests that total satisfaction from pre intervention to post intervention improved considerably. Furthermore, if we consider the maximum total of 80 to indicate total satisfaction (pre & post composite scores summed) then total satisfaction of reading English books in recreational and academic situations is deemed to be high.



#### 4.4.3 Monitors' Gender Effects on Academic and Recreational Attitudes

This section reports the findings of the academic and recreational English reading attitudes when considering gender over time. Specifically, the following research sub question guides the study by asking;

- Are there differences in Monitors post recreation, academic and total reading aspirations based on gender when controlling for pre test recreational, academic and total composite scores?

To test for any gender differences between each construct over the two time points, an ANCOVA was selected. The fixed effect variable for the model included gender type (coded 0 = male, 1 = female) and the dependent variable post-test scores (for both recreational and academic constructs) with the covariate (control factor) being the pre test satisfaction scores (again for both the recreational and academic constructs). Normality test results found there to be no extreme outliers from inspection of the pre recreational data, therefore, both gender types were normally distributed with linear association, this, being supported by similar regression lines and thus fulfilling homogeneity of slopes with a statistically significant result,  $F(1, 40) = 1.197, p = > 0.05$ . Inspection of scatterplots revealed there to be homoscedasticity and homogeneity of variances meeting further assumptions that data follow a normal distribution. After performing an adjustment for pre test recreational satisfaction scores, the main effect of gender was not statistically significantly different to post recreation satisfaction scores,  $F(1, 41) = .60, p = .808$ . The final model can be seen in Table 6.

Table 6 ANCOVA of Recreational Post Attitude Results Controlling for Baseline Measures

<i>Source</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Partial Eta Squared</i>
Recreation Pre Test	17.048	1	17.048	1.330	.031
Gender	.768	1	.768	.060	.001
Error	525.358	41	12.814		

\*\*= $p = 0.01$

Table 7 illustrates the unadjusted and adjusted mean scores of the post recreational satisfaction scores when controlling for pre test scores. It shows that there is little difference between male and female monitors over the time period.

**Table 7 Unadjusted and Adjusted Recreational Attitude Means (by gender)**

Group	<i>N</i>	<i>Unadjusted</i>		<i>Adjusted</i>	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Male	23	35.52	4.176	35.55	.747
Female	21	35.33	2.816	35.29	.782

The inference made from the results suggests that there was no difference made when adjusting for pre recreational attitudes between male and female Monitors. Using the data from Table 7, it is clear to see that there was no difference over time based on these factors with the similar mean scores reported.

Pre academic satisfaction scores using the same ANCOVA procedures were carried out to test for gender differences. Normality test results found there to be no extreme outliers with gender types normally distributed by approximation. Homogeneity of slopes was not violated as the interaction term was not statistically significant,  $F(1,40) = 72.653, p = > 0.05$ . Graphical outputs from scatterplots illustrated that data had homoscedasticity and homogeneity of variances fulfilling the assumptions that data follow a normal distribution. The ANCOVA results from the adjustment for pre academic satisfaction scores found the main effect of gender to be not statistically significantly different from post academic satisfaction scores,  $F(1, 41) = 20.654, p = .220$ . The final model is presented below in Table 8 The covariate, pre academic scores was statistically significant against post test scores but there were no difference between male and female responses.

Table 8 ANCOVA of Academic Post Attitude Results Controlling for Baseline Measures

<i>Source</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Partial Eta Squared</i>
Academic Pre Test	72.391	1	72.391	5.431	.025
Gender	20.654	1	20.654	1.550	.220
Error	546.499	41	13.329		

\*\*= $p = 0.01$

Table 9 displays the adjusted and unadjusted mean scores when taking into consideration of the covariate (pre academic satisfaction scores). It shows that the mean satisfaction scores for both male and female are very similar.

Table 9 Unadjusted and Adjusted Academic Attitude Means (by gender)

<i>Group</i>	<i>Unadjusted</i>			<i>Adjusted</i>	
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Male	23	33.43	4.69	33.11	.773
Female	21	34.19	2.58	34.53	.810

The main inferences based on these results suggest that gender wasn't a factor of academic satisfaction over the intervention period. When controlling for pre intervention satisfaction levels, the scores were not statistically significantly different between male and female Monitors.

Using the total satisfaction scores of the recreational and academic constructs, a new variable was created to understand if any differences occurred in overall satisfaction and whether these differences occurred between gender types. An ANCOVA procedure was carried out to test these assumptions. Results from tests of normality found the data to fit a normal distribution by approximation. No extreme outliers were present in the graphical

outputs that could lead to Type errors and distributional error. Homogeneity of slopes was fulfilled with the interaction term resulting in non statistical significance,  $F(1, 40) = 1.269$ ,  $p = > 0.05$ . Further inspection of graphical outputs found data to have homoscedasticity and homogeneity of variances, thus fulfilling normality assumptions about data and its covariance in the final model. ANCOVA results when performing an adjustment for pre satisfaction total scores found the main effect of gender to not be statistically significantly different to post total satisfaction scores,  $F(1, 41) = .012$ ,  $p = .913$ . The final model is presented in Table 10 below and details that adding the covariate (pre total satisfaction scores) to the model results in no statistically significant difference,  $F(1, 41) = .363$ ,  $p = .550$  to the post test totals and thus makes no difference to the overall model.

**Table 10 ANCOVA of Total Post Attitude Results Controlling for Baseline Measures**

<i>Source</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Partial Eta Squared</i>
Total Pre Test	11.282	1	11.282	.363	.009
Gender	.372	1	.372	.012	.000
Error	1273.035	41	31.050		

\*\*= $p = 0.01$

Perusal of Table 11 informs how the mean changed when considering the covariate (pre total satisfaction scores) and its effect on gender at the post-test stage. It shows that there is not much difference.

**Table 11 Unadjusted and Adjusted Total Attitude Means (by gender)**

<i>Group</i>	<i>Unadjusted</i>			<i>Adjusted</i>	
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Male	23	68.56	6.59	68.52	1.16
Female	21	68.66	4.04	68.71	1.21

The main inferences based on these results suggest that gender wasn't a factor of academic satisfaction over the intervention period. When controlling for pre intervention satisfaction levels, the scores were not statistically significantly different between male and female Monitors.

## 4.5 PCA on Monitors' L2 Baseline English Reading Attitudes

### 4.5.1 Introduction

From the aspirational survey, the attention shifted from differences to what factors are important in Monitors' feelings towards accessing and using English. Using an exploratory design, a principal component analysis (PCM) was run on the 20 likert styled questions representing aspirations of reading English in academic and recreational environments. In support of these factors, focus group data are also used to triangulate responses about interactions with English. Specifically, the following research sub question asked;

- What are the main factors concerning Monitors' reading attitudes in academic and recreational settings?

Data presented here represents all Monitors (n =44) responses to the pre recreational aspiration survey. The survey was developed to represent localized and culturally applicable responses conducive to the Monitors' reading environment. Principal components analysis was used to determine survey responses would load onto a particular factor, using the composite scores to generate new factors via an exploratory design. Highly loaded factors are explored against their real-world meaning (Field, 2013) through explanations provided at Monitors' pre-trial focus group. Principal components derived from analysed data provided the conceptual framework for focused questions related to recreational and academic attitudes towards English reading respectively.

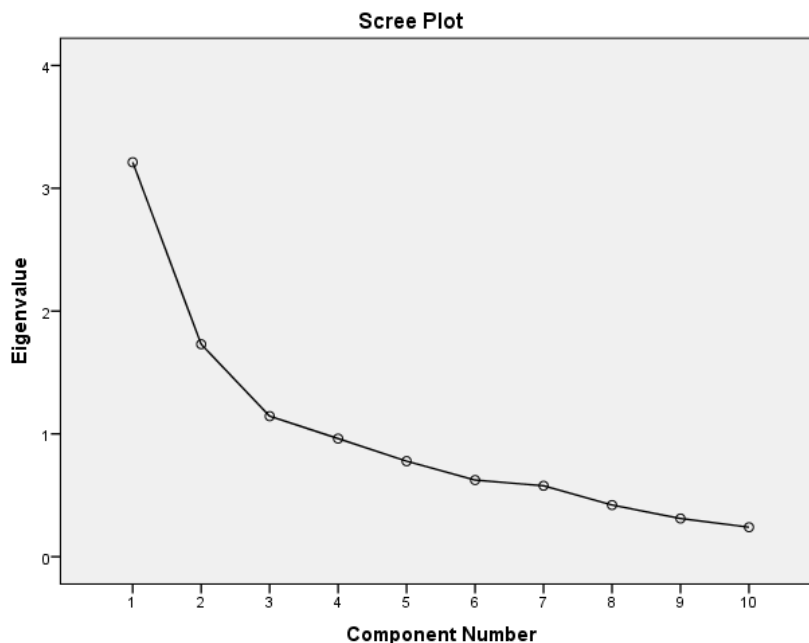
Responses from the recreational attitudes are reported first. Ten questions were adapted from the ERAS survey (Mckenna & Kear, 1990) to yield responses akin to the localized environment. At writing, no similar survey measured these constructs and as such, there are no standardized norms in fitting with the Indian context. The recreational survey asked ten questions about Monitors' attitudes to reading English in a social environment. Each question instructed the Monitors to think about responses in a non-education environment, thus, being at home or with friends at the park. Some questions provided context (such as being at home) but others didn't, therefore, it was important that Monitor responses were instructed to think about giving responses accurately linked to their recreational attitudes to each question.

#### 4.5.2 English Reading Attitudes Within a Recreational Setting

The suitability of PCA was assessed prior to the analysis and reporting of final data. Inspection of the correlation matrix met the conditions set by having at least one coefficient greater than 0.3. The final Kaiser Meyer Olkin (KMO) measure was inspected at 0.64, the minimum requirement being 0.6 for conducting PCA according to Kasier's (1974) advice for linear relationships. Bartlett's test for Sphericity was statistically significant,  $p = 0.00$ , thus, indicating the factorizable nature of the analyzed data.

PCA revealed there to be 4 components that had eigenvalues greater than one explaining 32.1%, 17.3%, 11.4% and 9.6%, respectively of the total variance. Visual inspection from the scree plot in Figure 21 shows the inflection point and cut off for each of the four components. Only 3 components were kept based on their interpretability and general guidance provided by Cattell (1966) for factor loading and simple structuring (Thurstone, 1947). Finally, Table 12 displays a communality column within the final factor model, it shows that all items reported being over .3, leading to interpret the findings as each item having some common variance with other items in the final model.

Figure 21 Scree Plot of Recreational Reading Attitude Eigenvalues >1



A varimax orthogonal was used to determine interpretability based on the 3 new factor loadings. An oblimin rotation was also used but there was no difference in either approach based on the 3 accepted factors. Total variance of the 3 components explained 61% of the

total cumulative variance. The final matrix of factor loadings are presented in Table 12, it shows that all 10 items are retained and had a primary factor loading  $> .4$ , with items 1, 2 and 3 loading onto factor one, items 8, 5, 9 and 6 onto factor two and items 4, 10 and 7 onto factor three. Interpretations of the loadings and their composite scores are also presented, here the factors have been interpreted as attitudes to reading for fun (factor 1), reading books in different locations (factor 2) and reading with peers (factor 3).

**Table 12 Factor Loadings & Commonalities for English Recreational Reading (N = 44)**

	Reading for Fun (factor 1)	Reading Books in Different Locations (factor 2)	Reading with Peers (factor 3)	Communality
How do you feel about reading English books for fun?	.89			.79
How do you feel about reading English books with family at home?	.69			.63
How do you feel about reading English books away from school?	.67			.65
How do you feel about reading English books instead of playing after school?		.76		.45
How do you feel about learning to read new words from an English reading book outside of school?		.68		.50
How do you feel about seeing new English books at the market or somewhere outside of school?		.65		.46
How do you feel when you read English textbooks/stories to your friends outside of school?		.55		.48
How do you feel when a friend gives you an English reading book to share?			.80	.66
How do you feel about reading English books not related to school work with friends?			.73	.59
How do you feel about listening to your friends read English textbooks/stories outside of school?			.60	.48

Note:  $< .4$  are suppressed

Based on the loadings for each factor, all have high positive loadings, with the exception for item 6 (*how do you feel when you read English textbooks/stories to your friends outside of school*), with a correlation of .55 in the second loading. This item was retained as eigenvalues lower than .4 were identified as either showing cross-loadings (loading onto multiple factors), which invariably led to data not following a simple structure (Thurstone, 1947).

Internal consistency for each of the new factors was examined using Cronbach's alpha. The alpha levels reported an acceptable level .72 for the three items within the 'reading for fun' factor, low alphas .67 for the four items within the 'reading books in different locations' factor and the three items related to 'peer reading' .63, respectively. Removing items did not contribute to alpha scores improving (i.e. > .7). However, guidance from Everitt & Dunn (2001) suggests that PCA is primarily used for reduction purposes and applies a different examination akin to the theoretical bases found in factor analysis, thus, whilst the internal consistency is moderately low for two of the factors (reading stories and peer reading), PCA results in the reported factors are considered to be valid, and useful enough, in terms of themes to explore at focus group interviews.

Following the PCA analysis, composite scores were calculated for each of the three new factors based on the primary loadings using the mean. Perusal of histogram plots for each construct reported a normal distribution by approximation based on the sample size and theoretical basis for data reduction analysis. Table 13 reports the composite mean scores for each factor indicating that the reading stories composite total to be higher than the other factors. Higher composite scores indicate a greater happiness within each construct.



	No. of items	<i>M</i> ( <i>SD</i> )	Skewness	Kurtosis	Alpha
Reading for fun (factor 1)	3	9.44 (1.9)	-.22	-.85	.72
Reading books in different locations (factor 2)	4	12.7 (2.3)	-.04	-.65	.67
Reading with peers (factor 3)	3	9.86 (1.6)	-.25	-.52	.63

Table 13 Descriptive Statistics for Three Recreational Scale Factors (N = 44)

With reference back to the coding (1 = very dissatisfied, 4 = very satisfied) the composite total for complete satisfaction is 40, however, this is based on each respondent reporting ‘very satisfied’ (4) on each of the ten items from the original survey. Therefore, lower responses ‘very dissatisfied (1) would yield a lower total. For these specific factors, we use the same parameters but in addition calculate a percentage of the overall satisfaction contained within. As with the first factor, and based on recreational pursuits, Monitors reported a ‘Reading for Fun’ factor ( $M = 9.44$ ,  $SD = 1.9$ ), with 79% of them satisfied with ‘Reading for Fun’ in recreational environments. The second factor was interpreted as ‘Reading Stories’ and reported a higher average ( $M = 12.7$ ,  $SD = 2.3$ ) compared to the other factors and greater variability when considering the standard deviation. However, it included more items and with more scope for deviation. Based on the total happiness within these items, 79% of Monitors reported satisfaction with reading stories in their recreational environment. This equalises with the first factor, which also reported the same percentage after rounding. Finally, the last factor was interpreted as ‘Reading with Peers’, reporting similar totals ( $M = 9.86$ ,  $SD = 1.6$ ) to the first factor and with the same number of item correlations. However, this factor was slightly higher with 81% of Monitors reporting satisfaction with ‘Reading with Peers’ in a recreational environment.

The PCA on recreational reading attitudes revealed three underlying factors related to fun, stories and peers. This was explored through further justification at focus groups conducted at the pre-pilot stage of the research.

### 4.5.3 Recreational Focus Group Findings

Each of the factors was explored in-depth to gain a deeper understanding of Monitors' attitudes to reading English in a recreational (or non academic) environment. A total of N = 16 Monitors were selected for focus group questioning based on the factors revealed by PCA. Table 14 illustrates the descriptive statistics of Monitors on each of the three factors. It shows the raw scores obtained on each factor. Monitors were selected to a focus group based on gender (to give equal weighting) and their factor scores.

Table 14 Descriptive Statistics of Monitors' Focus Group Information

Monitor ID	Focus Group	Age	Gender	Reading for Fun	Reading Books in different Locations	Reading with Peers
M1	A	14	Female	9.00	12.00	9.00
M17	A	11	Male	11.00	11.00	10.00
M7	A	9	Male	10.00	15.00	9.00
M23	A	13	Female	8.00	16.00	9.00
M3	B	8	Male	9.00	12.00	9.00
M21	B	8	Female	6.00	12.00	12.00
M16	B	12	Female	9.00	12.00	9.00
M35	B	12	Male	9.00	16.00	12.00
M42	B	11	Female	8.00	14.00	9.00
M29	C	8	Female	11.00	11.00	8.00
M14	C	13	Male	12.00	16.00	12.00
M18	C	9	Male	11.00	11.00	10.00
M27	C	11	Male	9.00	12.00	9.00
M39	D	10	Female	8.00	14.00	9.00
M43	D	8	Male	10.00	10.00	8.00
M31	D	7	Female	10.00	12.00	9.00

Each focus group was completed using predetermined questions linked to each factor. The first factor considered 'reading for fun' with very high loadings on the initial survey

indicating an above average level of satisfaction. Findings from the four focus groups revealed reading for fun to be associated with the types of books read, appealing stories and interesting characters and practising new words. In response to the question; *‘Thinking about reading in the home environment, what makes reading fun for you? What do you read and why is it fun?’* The following highlight the emerging themes from a selection of Monitor quotes conducive to each emergent theme;

#### *Types of Books*

M17: *“I enjoy the comics, they good because colourful and big to read. They make me laugh because they are not like textbooks because pictures are there and the funny faces make me laugh so many times”*

M3: *“I love the space books from the market, so many are there. I sit in my house and read about the space and how it takes too much time to travel there”*

M27: *“Four stories about the Himalaya are fun, this is very exciting reading because one day I want to travel there to see the snow.”*

Similar responses associated with reading material and in the quote below, specific reference to characters that fuel the enjoyment of reading included the following responses;

#### *Appealing Stories & Interesting Characters*

M35: *“I read Doramon (cartoon character) story books because it makes me laugh. He is cartoon person and has a friend who helps him. Problem for him sometimes when he faces daily trouble with school activities. I buy all books every month when I visit Banala market stand and take them home to read.*

M43: *“All types of books I read for fun are there. I like investigator books because it shows mystery of major event. These books are fun because you try to find clues for problem”.*

M21: *“ One story was about a boy who painted his cat by mistake, this made me laugh because he thought he was washing it (unintelligible) very funny story”.*

Whereas some responses also reported the difficulty in reading English even though there was some understanding of text;

#### *Practising New Words*

M29: *“Sometimes I read a word I don’t understand so use dictionary to find what it means for story. It can be confusing for me but it is fun to find what it means.*

M16: *“It is fun when I read a new word but difficult for me to write using it my school work. Sometimes I look on computer so I can use the joke on my friends if it is funny.*

M42: *“Even reading my favourite book is difficult because so many words I don’t know. It helps me read more and remember which I like to do by myself and when (name deleted) visits house”*

Reading for fun was also thought to be an activity limited by the academic focus of schoolwork and preparing homework and/or revising for weekly tests. Other causal examples limiting reading for fun included domestic duties, lack of enjoyable reading material and problems with electricity.

The second factor revealed high loadings on reading in different locations. Perusal of Table 14 indicates high raw scores from the selected Monitor sample. With a maximum score of 16, indicating the highest level of satisfaction, it was important to explore what ‘reading in different locations’ meant to each Monitor. As such, the following question was asked;

*Describe your experiences of reading outside of school, what experiences do you have and where do you go?*

Most of the responses were linked to reading in the home environment on an individual basis such as those mentioned here;

M7: *“At home after duties have finished I take the books to read”*

M18: *“Sometimes in my bedroom and sometimes I like to read my storybook outside the house by the petrol pump stand. It is much quieter there because my brother won’t find me.*

M1: *“With my mother in house. She reads the English and helps me with words. Then after some time I read to my sister in kitchen because the light is much more there.*

Other responses included going to the market and reading in bookstores, the following locations were discussed;

M16: *“Bookshop at market is good one, Mr (name removed) gives permission for the childrens to read book we want to read.*

M43: *“Reading by myself, I go to the market to pick up a book I want to read. I think most books are in English. Hindi and Punjabi books are also there but for study. I like to read the computer magazine there.*

M31: *“The market has many books but sometimes too much cost, Mam brings new books to tuition so I can borrow which ones I like”*

In terms of outdoor reading experiences, there were varied responses, which related to positive and negative experiences when reading in related to location;

M23: *“Problem is with time, sometimes I go outside house to read – it is difficult when so much noise in house”*

M42: *“Weather is difficult situation here in Punjab, too much hot and then rains. Reading outside is difficult when power is cut because no light”*

Whereas others reported positively;

M3: *“I like to be outside because much more space to sit and move away from busy area. Reading in quiet zones is much more beneficial, we go there to leave trouble from daily problems like being at school and homework”*

M14: *“From outside to inside I don’t mind but I think I prefer outside much more. I go to the farm near (unintelligible) hill and read and sometimes late”*.

Reading in the home environment was a major factor when considering reading recreationally. Initially, it was thought that the home environment would be heavily referenced, however, this was extended to other locations, such as, market places, farms and quiet places outside. Therefore, the descriptions detailed here extend the meaning of reading recreationally; it isn't singularly associated with reading in the home environment, instead, it extends to other locations.

The final factor had high loadings and was interpreted as ‘reading with peers’. This presented an interesting inquiry summary as plans were being made to develop a peer-led research trial. In this regard, responses to the question;

*Tell me about your experiences of reading with your peers (friends) outside of school. What does this look like and describe what you do?*

The responses gathered around themes to do with sharing stories, study clubs, homework and peer formations. Monitors reported heavily on sharing stories, explaining how this is completed in peer groups of varying dimensions, some comments included;

M39: *“Meeting (name) after tuition we go to (name) house and read the comic and play games. Sometime if power is on, (name)’s dad let we play on internet. We go to this place where you make up a story and change ending. Sometime we argue and can’t decide what to do next. (name) is really good, he knows what to do with character and then move to next stage”*.

M3: *“I read with my friend (name) in book shop near Patiala junction. He goes to (School name) English school and tells me about his favourite story books. He likes same as me. We have some time to discuss and then I have to meet father because tuition.*

M35: *“My sister friend is at (name) College and she has friends who visit the house for studying. They read me stories that makes me laugh and I enjoy.*

M29: *“Reading stories with my friends is happy time because no study or duty. We read in my house and sometime I go to (name)’s house near bus stand. We read story from school library in our houses, then swap if time, we make short quiz because sometime Mam asks us about the story we took from library”.*

The children talked about the positive experiences of reading with peers in a number of situations and contexts. At this stage of the research, it was encouraging to see the different peer relationships children formed and how reading activities were set up and arranged.

Following the recreational reading focus groups, further analysis was conducted on the academic responses of the survey and investigated for contextualised themes explored through group interviews.

#### **4.5.4 English Reading Attitudes Within an Academic Setting**

The same PCA processes were performed on the academic part of the English reading attitude survey. Monitors completed the second part of the survey on a different day and within school, an obvious environment to complete the academic element of the survey. The instructions were once again provided by the research team and the principle researcher, all Monitors completed the survey and reported their level of happiness to virtually the same questions but with an academic outlook. The academic component of the survey can be found in Appendix B.

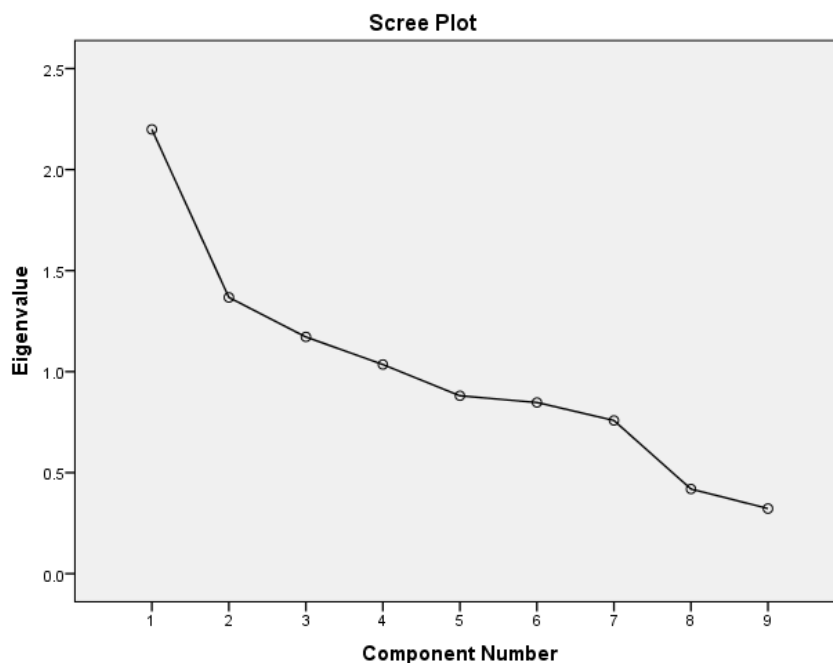
Inspection of the correlation matrix reported that the conditions for one coefficient greater than 0.3 were met. The Kaiser Meyer Olkin (KMO) measure was reported at 0.78, the minimum expected being 0.6 for conducting PCA according to Kaiser’s (1974) guidance for linear relationships. Finally, the Bartlett’s test for Sphericity was statistically significant,  $p = 0.02$ , thus, data were representable in accordance to the basic assumptions required for factorizable designs.

The analysis revealed there to be 4 components that had eigenvalues greater than one, which explained 24%, 15%, 13% and 12%, respectively of the total variance. In keeping with this, visual inspection of the scree plot in Figure 22 below shows the inflection point

and cut off for each of the components. It is noted here that the point of inflection starts around factor 4 with the leveling off of items thereafter. Therefore, a four factor model was retained at this stage of the analysis based on eigenvalues greater than 1.

The scree illustration from Figure 22 notes the 5<sup>th</sup> component falls below the eigenvalue of 1, rendering it below the level of acceptance and total variance explained in this structure. Further inspection of the four-component model revealed the item, “*Q6. How do you feel when you have to read English textbooks/stories in class during school time?*” was excluded from final models, as it didn’t load onto any of the extracted factors (Matsunaga, 2010, Thompson, 2004).

Figure 22 Scree Plot of Academic Reading Attitude Eigenvalues >1



The nine-item scale was adopted based on factor loadings showing greater than .4. An initial four-factor component model represented 65.1% of the total variance, however, it didn't represent a simple structure due to cross-loadings and subsequent low loadings within each factor. Given the close proximity of the fourth factor close to having an eigenvalue less than 1, the same analysis was run with a three factor model using both varimax and oblimin methods. Both methods produced the same results with the three factor model, each explaining 24.4%, 15.1% and 13% of the total variance, similar to the four factor model but with no cross-loadings as illustrated in Table 15 PCA revealed the final three

component model, following a simple structure and with eigenvalues of  $> .4$ , to explain 53% of the total cumulative variance. A varimax orthogonal rotation was used in generating the final interpretability of the model. The factors can be seen in Table 15 with items 2, 9 and 8 loading highly on to factor 1, items 4, 10 and 7 onto factor 2 and items 1, 5 and 3 onto factor three, all incorporating high correlations. In this scale, the factors are interpreted as attitudes to reading books to teachers and play (factor 1), reading with peers (factor 2) and reading for academic purposes (factor 3). All loadings are presented in Table 15.

**Table 15 Factor Loadings & Commonalities for English Academic Reading (N = 44)**

	Reading Books to Teachers & Play (Factor 1)	Reading with Peers (Factor 2)	Reading for Academic Purposes (Factor 3)	Communality
Q2. How do you feel about reading English books with teachers at school?	.87			.79
Q9 How do you feel about new textbooks/storybooks when at school?	.66			.46
Q8 How do you feel about reading English books to teachers instead of playing at break time?	.52			.39
Q4. How do you feel when a friend shares an English book during school time?		.70		.51
Q10. How do you feel about reading English books related to school work with friends?		.68		.65
Q7. How do you feel about listening to your friends read English textbooks/stories in a lesson at school?		.62		.47
Q1. How do you feel about reading English books in school?			.72	.64
Q5. How do you feel about learning to read new words from an English reading book selected in school?			.60	.46
Q3. How do you feel about reading English books for work in school?			.56	.33

Note:  $< .4$  are suppressed

Measures of internal consistency were then completed on each of the new factors using Cronbach's alpha. The first factor, which indicated 'Reading Academic English books to Teachers and Play' reported an acceptable alpha level of .64. The second factor indicating 'Reading with Peers' also had a moderate alpha level of .61. Lastly, the third factor



“Reading for Academic Purposes’ reported a good level of consistency with an alpha of .72. Based on the acceptable (but low) alpha levels on the first two factors, items deleted did not improve alpha in either of the first two factors. Both were deemed acceptable against the theoretical basis of capturing a smaller number of variables within a larger scale measuring the same latent construct (Bland & Altman, 1997, DeVellis, 2003, Everitt & Dunn, 2001, Kline, 2005).

To better understand the directional pathways of attitudes to reading English in an academic environment, based on the extracted factors, composite scores were calculated from the primary loadings using the mean. Assumptions to do with normal distributions were met by approximation, some skewness was present but perfectly normal for a small sample size and the theoretical basis, as previously discussed. Table 16 illustrates the final factors and their calculated mean score, including standard deviation and reported alpha levels.

**Table 16 Descriptive Statistics for Three Academic Scale Factors (N = 44)**

	No. of items	<i>M</i> ( <i>SD</i> )	Skewness	Kurtosis	Alpha
Reading books to Teachers and Play (factor 1)	3	6.39 (2.1)	.57	-.47	.64
Reading with Peers (factor 2)	3	6.23 (1.8)	.60	1.0	.61
Reading for Academic Purposes (factor 3)	3	6.22 (1.6)	.29	.33	.63

As before, the descriptive statistics found in Table 16 are composite scores found within the original scale of ten items. Each item was coded 1 (very unsatisfied) to 4 (very satisfied), thus, a maximum composite score on the first factor (reading books to teachers and play) indicating total happiness would be 12, as there are only 3 items.

Reading books to teachers and play had high loadings in the first extracted factor. Monitors reported their happiness within these items ( $M = 6.39$ ,  $SD = 2.1$ ) representing just over half (53%) of the total sample  $N = 44$ . It demonstrates that Monitors are neither satisfied nor dissatisfied when reading books to teachers and playing. The second factor associated reading with peers with similar findings ( $M = 6.23$ ,  $SD = 1.8$ ) to the first factor. Findings indicate that 51% of Monitors had indifferent views about reading with peers in an

academic setting. This was equalised with the last factor – reading for academic purposes ( $M = 6.22$ ,  $SD = 1.6$ ) with half (51%) of the Monitors sharing indifferent attitudes.

In summary, on each of the factors related to academic reading aspirations, just over half of the Monitors had a positive outlook. Further, when considering reading factors associated with academic environments and total happiness, it is unclear why just under half of the sample was either indifferent or leaning towards unhappiness. These data alone cannot answer this question. Instead, focus group interviews were utilised as they go beyond quantitative collections by trying to understand the ‘why’ questions behind, in this case, pre academic English aspirational scores to gain some clarity on why choices were made.

In particular, three factors were interpreted as having high loadings that were clustered around particular items on the likert-styled survey. These items represented common themes, or factors based on Monitors’ response and thus warranted further investigation. Similar to the recreational survey, each factor was used to develop a series of questions to pursue Monitors’ reasoning.

#### **4.5.5 Academic Focus Group**

Using the descriptive statistics from each of the raw scores summed on each factor, Table 17 presents a detailed snapshot of Monitors’ responses, including how focus groups were assigned. The sample widened to include Monitors who didn’t feature in the recreational focus group. A total sample of  $N = 16$  Monitors were recorded and analysed based on the summaries found in Table 17.

The first factor was interpreted as ‘reading books to teachers and play’ a total of three items within. A score closer to 12 indicates total happiness for each of the created factors. As with the findings based on the full Monitor sample, the mean score ( $M = 6.39$ ,  $SD = 2.1$ ) indicted Monitors as being indifferent. It demonstrates that just over half of the sample was happier when reading books to teachers. This was explored in further detail at focus group interviews.

Table 17 Descriptive Statistics of Monitors' Reading Aspirations

Monitor ID	Focus Group	Age	Gender	Reading Books to Teachers	Reading with Peers	Reading for Academic Purposes
M9	E	12	Female	10.00	3.00	8.00
M25	E	12	Male	4.00	3.00	3.00
M20	E	10	Male	6.00	7.00	5.00
M4	E	10	Female	3.00	3.00	3.00
M44	F	10	Male	9.00	12.00	9.00
M28	F	10	Male	9.00	6.00	3.00
M36	F	7	Female	6.00	6.00	3.00
M26	F	13	Female	8.00	6.00	5.00
M8	G	13	Male	4.00	7.00	6.00
M19	G	14	Male	10.00	9.00	8.00
M6	G	8	Female	8.00	5.00	7.00
M32	G	14	Female	10.00	5.00	8.00
M22	H	11	Female	8.00	6.00	9.00
M34	H	14	Female	10.00	5.00	8.00
M5	H	11	Male	4.00	8.00	7.00
M40	H	11	Male	8.00	7.00	6.00

Findings from the four groups revealed reading books to teachers and play to be associated with, teacher continuity, free periods and reading pedagogy. In response to the following question;

*Thinking about reading English books to teachers at school, what does this look like and how do you feel about doing it? How do you feel about reading at break time*

Each Monitor reported various experiences, whether positive/negative, when presenting their views on reading to different teachers. The following quotes exemplified this;

#### *Teacher Continuity and Free Periods*

M4: *“I don’t like reading to mam because problem is that different teacher. They give different book and I read but sometimes I read wrong book. Teacher shout if I read wrong book but it is difficult because I don’t know teacher who I read to on that day. So many times I miss lunga (lunch) period because teacher is discussing with me.*

M19: *“Teacher asks to read the book but sometime I get from library and teacher becomes impressed with me. I like reading to all English teacher, we sit outside with lunch when no rain and enjoy. She help me with words that I don’t see before. I enjoy Mr (name) and mam (name), they help me too much with each chapter – it is hard”*

M25: Mam (name) is good and wishes me well every time I give difficult words but Mr (name) has challenging nature. He comes every time Mam is on tea leave or not in class. We read difficult books and he makes me finish before break period. I wish I see Mam because she help me and makes me happy.

M34: “We sit in class reading the books with teacher and my friends too. It is easy task because I practice at tutorial with Mr (name) and then same book in class. I see Mam (name) on Wednesday and Mr (name) on Monday reading different English books. I like story one, and reading newspaper article with Mr (name). I missed break too much because I don’t understand”.

The Monitors reported varying responses depending on the teacher they were assigned to and how this sometimes affected free periods. Firstly, all Monitors seeing different teachers had a major effect on their happiness. That is, seeing multiple teachers for reading affected their ability to perform, lose focus or cause confusion. Teacher continuity is therefore a major sub-theme that emerged from the discussions. Secondly, most Monitors reported either being happy or unsatisfied with missing breaks when matched to positive and negative reading experiences with teachers, respectively. It demonstrates that Monitors with positive reading experiences are mostly satisfied, whereas those who struggle are set new objectives and are required to catch up, this being in their free time.

To highlight this further, Monitors also reported the types of pedagogy each teacher employed when responding to the original reading with teachers question; “*what does it look like and how do you feel about doing it?*”

#### *Reading Pedagogy*

M36: “Mr (name) let me read and then stops after some time. He ask questions about what I read to him, about story. It’s ok, I can answer question about story but sometimes I don’t understand. He help me to find answer”.

M8: “After reading in class, I stay behind to read book with Mam (name) and I go first. She reads the first sentence and then I take over. The problem I have is when she ask the big question. This is after a chapter I read, she ask so many questions, I forget and look for answer in book. It makes me angry when I can’t remember on big question”.

M6: “Look at the pictures and she (teacher) ask question then read for some time. I like picture book, Mam help me to read story using picture, it help me too much but I don’t like words, too long”

M5: “We sit outside but classroom it too noisy. Mam shout at the childrens through window and me also. I read chapter and she help me but stop after small time. I don’t like to stop – lose my place and forget teacher question”

All Monitors described various techniques teachers used when listening to them read. Most recalled teachers pointing to pictures and then asking questions, stopping and asking general questions or reading small passages before handing over. Based on the responses, there was a clear divide in how this was managed, with some Monitors showing preference for one or more techniques while others explaining a disregard for the processes involved in guided reading.

The next factor reported high loadings that were interpreted as ‘reading with peers’, again, in response to questions developed to reflect Monitors’ thoughts on reading aspirations within an academic setting. The final factor aggregated three survey items and PCA revealed there to be indifference when considering reading with peers ( $M = 6.23$ ,  $SD = 1.8$ ) in an academic setting. Focus group interviews were arranged to explore the indifference, with the primary aim of trying to understand underlying causes. Monitor focus group interviews provided their views on the following question based on the interpreted factor;

*“Do you read with your friends in school? What do you do and can you tell me what this looks like?”*

Monitor responses were again varied for working with their peers. The following themes emerged from data that provided further clarification on why there is indifference between Monitors’ academic reading attitudes. The majority of responses were clustered around peer arrangements, time on task and reading academic subjects. Comments on the first theme included the following;

#### *Peer Arrangements*

M44: *“We sit with other childrens from other class and read the book. Enjoy it on second Saturday with older childrens, they help complete homework and understand better”.*

M20: *“Too many fights because we become over 60 in class. I sit and read my book and my friend play instead helping me. Mam get angry because we are not with partner reading”.*

M9: *“(Name) helps me a lot, she is my friend in Mam (name) class so makes me sad when she go. We work on reading task and help each other to do well”.*

M26: *“Too much reading in class, I on my own. Childrens come from higher standard and help us but noisy and I get the pains in head, it makes me angry because I read too many childrens”.*

Monitors had various views on reading with peers when reflecting on academic tasks. Positive comments tended to cluster around intimate reading arrangements, where Monitors either were paired with their friends or worked proactively on learning tasks set by other children or teachers. Minor categories such as overcrowding in the class and general classroom management were causal claims made by the sample. It suggests that environment factors and planning have a major effect on Monitors' aspirations to read academic content in pre-arranged peer groups. This was further detailed through the 'time-on-task' category with supporting comments that included;

#### *Time-on-Task*

M:28 *"Last time I waited for (name) but he didn't come. Then Sir asked me what I was doing waiting outside class. He say I work with (name) but I wasted too much time before and didn't want to do anything but (name) make me laugh telling the joke from library book:.*

M36: *"Problem is we have time every Wednesday but Mam say we need to complete Chapter to complete syllabus. She then give me and (name) before break time for reading task which I like to do. Time is difficult because class test and syllabus.*

M4: *"I read my science book with (name), he help me with difficult words. Reading time is ok, yes, morning time is better, after assembly and small snack, I like reading with snack – it help me think.*

It was clear from the majority of comments that Monitors would prefer more time with their peers on academic reading tasks. Some commented on in-class activities, while others mentioned the peer teaching days which involves children from higher standards supervising English reading, mathematics and science tasks.

Finally, the last theme that emerged from the analysis was interpreted as 'reading for academic purposes'. PCA found similar responses to the other factors ( $M = 6.22$ ,  $SD = 1.6$ ), thus, the three items within the 'reading for academic purposes' factor implied an indifferent response. Focus group interviews were targeted to understand more about Monitors' perspectives by answering the following question;

*Can you give me some details of how you feel when asked to read an English book in school? Remember to think about what you do, how you respond to reading tasks and any experiences you have when reading in school*

The major themes to emerge from the focus group interviews included achievement on tests, reading subjects and ability to read. Achievement on tests was regarded as a major theme

with responses from all of the Monitors. Mixed opinions about testing revealed sub-categories of worry, confusion, expectations and empowerment. The following identifies with these themes through the following responses;

#### *Achievement on Tests*

M40: *“Reading the Science book is difficult, words I don’t see before and hard. Too many, I get lost before homework, monthly quiz, it is too hard. SA one (summative assessment) I score low, so many words on test I didn’t see before”.*

M28: *“Mr (name) Sir give the readings daily. Sometimes panic from my side, yes, the Science and Math readings is difficult because of class quiz and SA1 and 2 are there, also. You can’t read difficult words then I take worry on test”.*

M22: *“I feel shaking because I don’t get the words correct in class quiz and then low house points. Problem is that I can read basic textbook but after some time I confuse words and get them wrong”.*

M34: *“Too much to remember but I like learning new words. Mam put new word on board every day and ask us to read each Chapter. She help me for test but I don’t understand and see her during break or for tuition”.*

Although there were concerns around testing, some Monitors reported their enjoyment reading textbooks in particular subjects;

#### *Reading Subjects*

M19: *“I like reading my school books, I like SST (social studies) Chapters because the history parts are very interesting”.*

M22: *“To get the knowledge, I read my textbook for everything I need to know. GS (general studies) is easy for me and I enjoy. India past is interesting and learning about Lok Sabha I am reading for class quiz tomorrow about this, lots of facts about Indian government I remember”.*

M25: *“I don’t like reading the textbook for Math, difficult, hard sentences to follow, I can’t do most time but I try”.*

M20: *“Reading Chapters with fill-ups and questions in English book, I struggle, too much for me, yes, the problem is that to remember all words”.*

Finally, when considering reading for academic purposes, the majority of Monitors commented on their reading ability, in some cases, setting targets on what they’d like to achieve but in other cases providing details about their thoughts on the academic reading process. Some of the comments included;

## *Reading Ability*

M8: “Mam (name) makes me feel good when she help me read difficult Chapter, she answer question at end or break time in staff office, I good reader, if, when, I read easy Chapter, like, newspaper article in English, Mam always praise me”.

M40: “I like reading the story book in English class, I feel good inside when Sir ask me to read in class. The childrens laugh when I get word wrong, the longer words, some easy, some not”.

M6: “Confidence is there but I worry, take tension, and get wrong. Reading the English and Hindi cause problem, reading take the time away from Mam, she help me. Can’t read the fast pace, (name) is first position in class, he’s too quick reading the English book”.

M26: “Problem for me, I get angry and lose my place on lines. English book too hard for me, yes, I get angry for reading after long time, reading is major problem when no time is there”

The varying comments on ability to read provided a much more in depth understanding on the struggles each Monitor faces when trying to read English for academic purposes. Encouraging were their efforts expended on trying to improve fluency, however, comprehension and general understanding of texts came through the interviews as a major impeders of cognitive literacy functions.

## **4.6 Descriptive Findings of Monitors’ Pre/ Post Test Scores**

### **4.6.1 Introduction**

Following the intervention period, the Monitors were tested again on the same reading, phonological and native tests administered at the pre – test stage. This section first reports the descriptive statistics of each test battery before exploring the following inferential research questions that asks;

- Is there a statistical difference between male and female Monitor’s baseline reading, phonological processing, decoding and native language test scores?
- What are the effects of Monitors’ phonological processing, reading and spelling abilities when guiding learners in a peer-mediated intervention?

Pre and post test scores from the Monitor sample are presented in the next section. This is then followed by tests that measured effects over the two time points.



#### 4.6.2 Monitor Reading & Dictation Test Battery

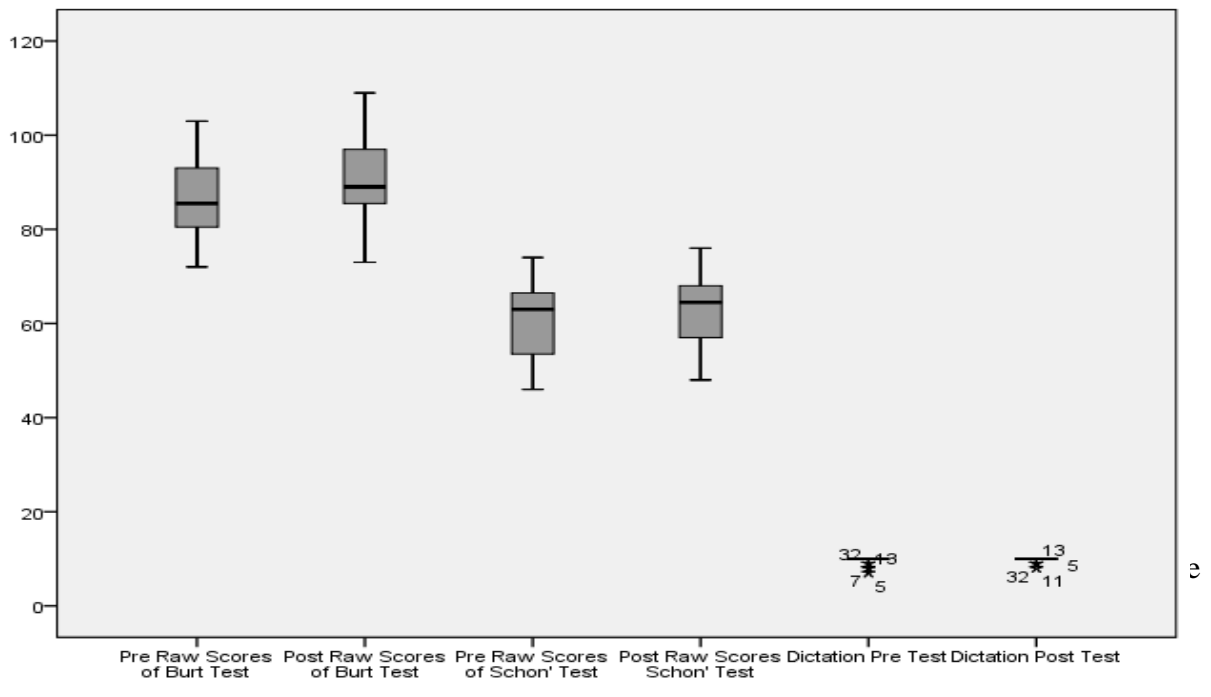
The descriptive statistics from Table 18 group the test results of reading, spelling and dictation at the pre and post stages. The results from the Burt test suggest that there was an improvement in words read correctly at post test ( $M = 90.68$ ,  $SD = 9.38$ ) when compared to the pre test ( $M = 85.55$ ,  $SD = 8.37$ ). The median improved but there was greater variability from the mean at the post test stage when considering the standard deviation pre/post scores. The Schonell spelling test mean also increased at post test ( $M = 63.23$ ,  $SD = 6.58$ ) when compared to the pre test mean ( $M = 60.27$ ,  $SD = 8.30$ ) which indicates that the Monitors sample improved their spelling over the study period. However, there was less variability away from the mean after the intervention. Finally, the Monitors' reported a similar Dictation mean at the pre test stage ( $M = 9.61$ ,  $SD = .841$ ) as they did at post test stage ( $M = 9.82$ ,  $SD = .495$ ) suggesting that there was marginal improvement over the two time points. The median confirms as the majority of students scored the maximum ten marks for this assessment.

**Table 18 Descriptive Statistics of Monitors' Baseline Reading Test Scores**

Burt Reading		<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>Mdn</i>
	Pre	44	86.55	8.37	85.50
	Post	44	90.68	9.38	89.00
Schonell					
	Pre	44	60.27	8.30	63.00
	Post	44	63.23	6.58	64.00
Dictation					
	Pre	44	9.61	.841	10.00
	Post	44	9.82	.495	10.00

The Boxplot graphic in Figure 23 illustrates the activity over the pre and post time points as observed within the reported descriptive statistics from Table 18. Visually, each Boxplot shows a similar spread of data for each of the reading tests administered at the pre and post test stage. Furthermore, the median for each test appears to show no obvious disparity when considering each time point. For the Burt test, there is a noticeable difference in the spread of data but the median line remains almost at par when comparing with the pre-test result.

Figure 23 Boxplots of Monitors' Pre/Post Reading Test Scores



### 4.6.3 Monitor Character and Phonological Processing Test Battery

Descriptive statistics from Table 19 reports on the additional tests used for letter, sound recognition and blending techniques. The results for the letter matching exercise show a slight increase from the pre test ( $M= 9.82, SD= .582$ ) against the post test ( $M = 9.93, SD = .254$ ), however the median remains the same, thus, indicating a minor change over the two time points. For Blending, the Monitors' results report a much higher average score from pre test ( $M = .50, SD = .976$ ) to post test ( $M = 33.07, SD = 1.30$ ). Similarly, the Miskin sound and blending nonsense test reported gains from pre test ( $M = 2.52, SD = 2.16$ ) to post test ( $M = 84.38, SD = 3.34$ ). Given the training all monitors received it is perhaps unsurprising to see greater gains in sound knowledge and awareness of blending techniques.

**Table 19 Descriptive Statistics of Monitors' Pre/Post Letter Matching and Blending Scores**

Letter Matching		<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>Mdn</i>
	Pre	44	9.82	.582	10.00
	Post	44	9.93	.254	10.00
Blending					
	Pre	44	.50	.976	.00
	Post	44	33.07	1.30	34.00
Miskin Nonsense					
	Pre	44	2.52	2.16	2.00
	Post	44	84.38	3.34	84.50

### 4.6.4 Monitor Native Reading Test Battery

Further tests were administered to explore if differences existed for native reading tests. Table 20 reports that for the Punjabi dictation test there were slight increases in the mean from pre test ( $M = 5.59, SD = 1.40$ ) to post test ( $M = 5.61, SD = 1.36$ ). For the Hindi word test, a similar pattern, with the mean increasing from pre test ( $M = 42.47, SD = 10.28$ ) to post test ( $M = 42.70, SD = 10.52$ ). Both medians remain constant over time and there is similar variability leading to the conclusion that no major differences occurred in native tests over time.

**Table 20 Descriptive Statistics of Monitors' Pre/Post Native Reading Test Scores**

Punjabi		<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>Mdn</i>
	Pre	44	5.59	1.40	6.00
	Post	44	5.61	1.36	6.00
Hindi					
	Pre	44	42.47	10.28	47.50
	Post	44	42.70	10.52	48.00

The next section reports findings from baseline measures when considering gender as the dependent variable.

## **4.7 Monitors' Gender and Test Battery Inferential Findings**

### **4.7.1 Introduction**

This section reports the findings of the Monitors' baseline reading, phonological processing, decoding and native language test scores. Firstly, baseline gender effects are reported. The following research sub question was used to guide the study;

- Is there a statistical difference between male and female Monitor's baseline reading, phonological processing, decoding and native language test scores?

The stated question is of particular interest as there are a number of factors that occur in Indian middle to high school education. One factor is that boys and girls don't tend to sit with each other, this is observed in many Indian education institutions. It would be interesting to know whether girls or boys performed better on the tests in order to utilise potential strengths with the peer phonics intervention.

Using the findings from the focus group and from the fieldwork notes there wasn't an obvious link to whether girls or boys showed more enthusiasm to learn basic reading through a synthetic method. It was decided to explore for differences using statistical tests of significance using 'gender' as the dependant variable. The following sections are the results from the reading, letter matching and phonological reading test batteries that were used to determine if any differences occurred when gender was the dependent variable.

Firstly, preparation for these data must be conducted so that it follows a Gaussian distribution. There are a number of ways in which to interpret whether data are normally

distributed, such as, histograms, test statistics from normality tests and Q-Q plots that illustrate linearity. As the sample population for the Monitors' group is  $< 50$  and considered to be small, the use Shapiro-Wilk's test statistic has given its power for small samples and sensitivity to minor deviations from the mean. For data to be considered as normal the results must have a greater significance of  $p = .05$  as it tests the null hypothesis, comparing the studied data to a normal distribution. Graphical outputs such as histograms and the statistics literature refer to data being shaped as a 'bell curve' or by approximation for data not to be violated.

Each test is grouped by its significance to what it is testing. The results from the reading battery include tests that test Monitors' ability to read, spell and dictate English words and sentences. Results are reported using the mean and standard deviation score with tests for statistical significance for the pre test level for each gender type. The first results reported are the Monitor's pre reading battery that tested for Monitors' pre reading, spelling and dictation abilities.

#### **4.7.2 Gender Effects: Reading & Dictation Test Battery**

Table 21 indicates there is a marginal difference in the mean when comparing male and female pre Burt test scores. An independent-samples t-test was selected to test for any differences in pre Burt Test scores between male and female Monitors. For this analysis, a between subjects method was applied where the independent variable 'gender' was fixed against each of the pre test scores. The pre Burt test scores between gender revealed there to be no outliers in the data, as assessed by inspection of a boxplot. Pre Burt test scores for each level of gender were therefore normally distributed, as assessed by Shapiro-Wilk's test ( $p > .05$ ), and there was homogeneity of variances, as assessed by Levene's test for equality of variances ( $p = .783$ ). Analysis revealed the pre Burt test scores not to be statistically different for male Monitors ( $M = 87.61$ ,  $SD = 8.66$ ) compared to female Monitors ( $M = 85.38$ ,  $SD = 8.10$ ),  $t(42) = .879$ ,  $p = .385$ . The mean difference in test scores was  $M = 2.23$ , 95% CI (3.00, 7.00), demonstrating a small effect size according to Cohen (1988),  $d = 0.26$ .

Table 21 Descriptive Statistics of Monitors' Pre/Post Reading Test Battery Scores (by gender)

Burt Reading		<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
	Pre Male	23	87.61	8.66	72.00	102.0
	Pre Female	21	85.38	8.10	74.00	103.0
Schonell						
	Pre Male	23	62.83	7.48	47.00	74.00
	Pre Female	21	57.48	8.43	46.00	71.00
Dictation						
	Pre Male	23	9.61	0.83	7.00	10.00
	Pre Female	21	9.62	0.65	7.00	10.00

The pre Schonell test was subjected to the same tests for distribution bias and revealed there to be no outliers as inspected visually using a boxplot. Pre Schonell test scores for each level of gender were normally distributed as assessed by the Shapiro-Wilk's test ( $p = > .05$ ), with there being homogeneity of variances, as assessed by Levene's test for equality of variances ( $p = .245$ ). Analysis revealed that there is a statistical difference for male Monitors ( $M = 62.83$ ,  $SD = 7.48$ ) when compared to female Monitors ( $M = 57.48$ ,  $SD = 8.43$ ),  $t(42) = 2.23$ ,  $p = .031$ , a mean difference of  $M = 5.53$ , 95% CI (.50, 10) between males and females but a moderate to high effect size according to (1988),  $d = 0.67$ .

Finally, from the reading battery, the pre Dictation test revealed there to be a number of outliers as inspected visually using Boxplot... for reference. Moreover, the Shapiro-Wilk statistic test reported ( $p = < .05$ ) thus violating tests for normality. Instead, a non-parametric method was considered based on the previous violations and used to test for differences. Table 22 illustrates the mean ranks for male and female pre Dictation test scores with a range between 7 and 10 marks for each level of gender. A Mann Whitney U analysis was run to test for differences on the Dictation test exist between male and female Monitors.

Table 22 Monitors' Pre Dictation Mean Rank Test Scores (by gender)

<i>Dictation</i>	<i>N</i>	<i>Mean Rank</i>	<i>Min</i>	<i>Max</i>
Pre Male	23	22.28	7	10
Pre Female	21	22.74	7	10

Distributions for both the male and female pre Dictation test scores were assessed to be normally distributed by approximation. Pre Dictation test scores for females (mean rank = 22.74) and males (mean rank = 22.28) were not statistically significant,  $U = 246$ ,  $z = .167$ ,  $p = .867$ . Indicating that the differences in the mean rank from Table 22 is marginal and that both male and female Monitors are not different at the pre test for reading basic sentences.

#### **4.7.3 Gender Effects: Phonological and Character Awareness Test Battery**

The descriptive statistics for this test battery can be found in Table 23 below. The Nfer A test required each Monitor to recognise an English character and match it correctly to the same character in amongst a line of other English characters. It's a test of letter recognition that gets progressively harder with more letters being added after each line with a total of 10 matches being tested. Table 23 illustrates the results of the Nfer A test and suggests that females ( $M = 9.95$ ,  $SD = .218$ ) were slightly ahead of their male ( $M = 9.70$ ,  $SD = .765$ ) counterparts in letter matching at the baseline level. The standard deviation suggests there's a greater variability for males given the spread of data from ranges 7 to 10 correctly matched letters. An independent t-test was run to investigate if any differences were present in the gender pre scores. No outliers were present as assessed by graphical boxplots in the SPSS output and Nfer A pre test scores for each level of gender were visualised inspected and found to be normally distributed. However, the assumption of homogeneity of variances was violated, as assessed by the Levene's test for equality of variances ( $p=.003$ ). As equal variances were not assumed it was decided to report the unequal assumed test statistic following guidance from Welch (1947). The results found there to be no statistically significant difference in mean pre male phonological test scores and pre female phonological test scores even with a mean difference of  $M = 0.25$ , 95% CI (-5.9, 0.85),  $t(25.872) = -1.542$ ,  $p = .135$ . The results infer that there was no difference at the gender level for identifying basic letter sounds and that letter recognition mean scores are constant across both groups.

For the Nfer B test, each Monitor was asked to sound out each letter presented on the test sheet, only sounds were recorded as being correct and not letter names. The results suggest that both male and female Monitors have limited sound awareness with both means being very low based on the scores achieved. Males were slightly higher ( $M = 2.00$ ,  $SD = 1.56$ ) than females ( $M = 1.67$ ,  $SD = 2.10$ ) with there being greater variability for female scores. An independent samples t-test was conducted to test for differences. No outliers were

present by visual inspection of graphical plots and homogeneity of variances were considered to be equal as tested by the Levene’s test of equal variance ( $p = .173$ ). There was no statistically significant difference in male pre Nfer test B scores and female Nfer Test B pre test scores when considering the mean difference,  $M = .33$ , 95% CI (-.789, 1.456)  $t(36.786) = .599$ ,  $p = .552$ . The results infer that there are no differences at the gender level when considering basic phoneme awareness.

**Table 23 Monitors' Phonological and Character Awareness Test Battery (by gender)**

Nfer A (Letter Matching)		<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
	Pre Male	23	9.70	.765	7	10
	Pre Female	21	9.95	.218	9	10
Nfer B (Sounding)						
	Pre Male	23	2.00	1.56	0	5
	Pre Female	21	1.67	2.10	0	5
Nfer C (Blending 2/3 Sounds)						
	Pre Male	23	.30	.635	0	2
	Pre Female	21	.71	1.23	0	5

The Nfer C test required each Monitor to correctly decode the sounds found in common 2/3 CVC words. Similar to the Nfer B, only correctly sounded blends would be recorded as being correct, even if the word was read correctly. The results suggest that females ( $M = .71$ ,  $SD = 1.23$ ) did better than males ( $M = .30$ ,  $SD = .635$ ) but the difference was once again marginal. Further, the results suggest that both male and female Monitors found this task difficult. Independent samples t-test was ran to test for differences in gender at the pre test stage. Data were visualised to be normally distributed by approximation through the use of Q-Q plots and on the basis that there were no outliers present. However, the Levene’s test for equality of variances was violated ( $p = .042$ ), therefore, the equal variances not assumed test statics are reported. Results found there to be no statistically significant difference in male pre Nfer C and female pre Nfer C scores taking into account the mean difference of  $M = .41$ , 95% CI (-1.02, .202),  $t(29.33) = 1.369$ ,  $p = 1.81$ . The results suggest that there are no differences at the gender level when considering pre – blending skills.



The results from Table 24 relate to the Miskin nonsense word tests. For this test Monitors were asked to decode the words by sound unit (phoneme) and asked to re-build the word. It tests for phoneme awareness ability. The test is usually administered as a whole with word items. The test gets progressively difficult asking children to correctly identify the different sound patterns found in the words (phonemes, digraph, trigraphs, for example). One way could have been to analyse the total amount of words correctly sounded and blended and compared these at the pre test level. However, the following results were analysed separately grouping each set of words by difficulty starting with simple phoneme sounds and then moving on to harder blends.

Monitors were exposed to tests that involved decoding irregular words, or words that have no meaning in the English language.. Presented in Table 24 below are the Nonsense words groupings that identify, by gender, the number of sounds for each group. The test starts with 'Phonemes and Digraphs', which are Group 1 letters and asks each student to reproduce their corresponding sound. Table 24 illustrates that the pre mean for the phonemes and digraphs pre test was higher ( $M = 1.70$ ) for males than for female Monitors ( $M = 1.52$ ). Perusal of the scores indicates the standard deviations of female scores were more spread compared to males. An independent samples t-test was performed to test for statistical significance. Visual inspection of graphical outputs generated by SPSS found there to be no outliers from the sample. Further inspection of distribution biases found there to be homogeneity of variances as reported from the Levene's test for equality of variances,  $p = .189$ . The results found there to be no statistically significant difference between pre male Miskin Phoneme and Digraph scores and pre female Miskin Phoneme and Digraph scores when considering a mean difference of  $M = .172$ , CI 95% (-.875, 1.21),  $t(42) = .331$ ,  $p = .742$ . The results suggest that at the pre – test level for the Miskin phonemes and digraph test to be no difference when considering gender.

Table 24 Descriptive Statistics of Monitors' Miskin Nonsense Words (by gender)

Miskin Nonsense Words		<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Phonemes and Digraphs	Pre Male	23	1.70	1.55	0	5
	Pre Female	21	1.52	1.88	0	5
3-4 Letter Sounds	Pre Male	23	.61	.891	0	3
	Pre Female	21	.76	.768	0	3
4-6 Letter Sounds	Pre Male	23	.17	.388	0	1
	Pre Female	21	.14	.359	0	1
4-13 Letter Sounds	Pre Male	23	.09	.288	0	1
	Pre Female	21	.05	.218	0	1

Analysis of the 3-4 letter sound words for both pre male and pre female tests yielded very low mean scores. Perusal of Table 24 illustrates that females did slightly better ( $M = .76$ ) than males ( $M = .61$ ) at the pre test stage. The standard deviation of male scores were spread out more compared to the female standard deviation. An independent t-test was ran to explore for differences at the gender level for pre test 3-4 letter sounding irregular words. No distribution biases were detected and inspected boxplots revealed no outliers greater than 1.5 of a boxplot. There was homogeneity of variances as inspected by the Levene's test of equality of variances,  $p = .272$ . The results of the t-test found there to be no statistically significant difference between pre male Miskin grouped 3-4 letter sounding irregular words and pre female Miskin grouped 3-4 letter sounding irregular words when considering the mean difference between each group,  $M = .153$ , 95% CI (-.662, .355),  $t(42) = -.608$ ,  $p = .547$ . The results suggest that there is no difference in blending irregular sounding 3-4 letter based on the Monitor's pre – test gender scores.

Analysis of the 4 -6 letter-sounding irregular words reported low mean scores for both male ( $M = .17$ ) and female ( $M = .14$ ) pre test scores. Results from Table 24 suggest there were a minimum of 0 and a maximum of 1 scored and this explains the low means for both groups.

An independent t-test was run to determine if any statistical differences were present at the gender level. A visual inspection found there to be no outliers greater than 1.5 from the edge of the top of the boxplot. Homogeneity of variances was met as reported by the Levene's equality of variances test statistic,  $p = .583$ . The t-test results found there to be no statistically significant difference between pre male Miskin 4-6 letter sounding irregular words and pre female Miskin 4-6 letter sounding irregular words when considering the mean difference between each group,  $M = .031$ , 95% CI,  $(-.197, .259)$ ,  $t(42) = .275$ ,  $p = .785$ . The results suggest that there were no differences found at the pre test of 4-6 letter sounding irregular words based on gender.

Analysis of the 4-13 letter sounding irregular words reported low mean scores and were predictably similar to the 4 -6 letter sounding irregular words. Taking the descriptive statistics from Table 24 it suggests that the pre male 4 -13 letter sounding irregular words were slightly higher ( $M = .09$ ) when compared to pre female 4 -13 letter sounding irregular words ( $M = .05$ ). The results show a similar pattern to that of the 4-6 letter sounding irregular words in that the range of correct scores were recoded as being between zero and one. An independent t-test was run to test for any statistical differences between gender and pre 4-13 letter sounding irregular words. No outliers were present when testing for any distribution biases and a visual inspection of the generated boxplot illustrated this. Homogeneity of variances was met as reported by the Levene's test of equality of variances,  $p = .310$ . The t-test results found there to be no statistically significant difference in pre male 4-13 letter sounding irregular words when compared to pre female 4 -13 letter sounding irregular words taking into account the mean difference of each group,  $M = .039$ , 95% CI  $(-.117, .196)$ ,  $t(42) = .507$ ,  $p = .615$ . Similar to the 4-6 letter sounding irregular words, the results suggest that there is no difference between Monitor's gender and sound awareness when considering difficult blends.

#### **4.7.4 Gender Effects: of Native Reading Battery**

Monitors were asked to complete two types of tests related to their native languages – Punjabi and Hindi. From the questionnaire results as reported earlier 99% of households speak Punjabi as their first language. It was assumed that this was to be the case for the children involved in this study as well. As such, a Punjabi dictation test was set up to see if children were able to read in their first language. Furthermore, given the other native language is Hindi and with guidance provided from the pilot test, some children reported that they have just started learning Hindi formally and it wasn't suited to testing using a

dictation reading test. Instead, a word test was selected as being appropriate with the results of both reported in Table 25.

Based on the independent variable: gender, the mean for the pre Punjabi dictation test is slightly higher for females ( $M = 5.71$ ) than it is for males ( $M = 5.47$ ). Variability for males is higher suggesting that scores are more spread based on the standard deviation. An independent t-test was run to test for any differences between gender and pre Punjabi dictation reading scores. No outliers were present in graphical outputs from SPSS and the homogeneity of variances was not violated as inspected by the Levene's test for equality of variances,  $p = .253$ . The results of the t-test suggest that there to be no statistically significant difference between pre male Punjabi Dictation reading scores and pre female Punjabi Dictation scores when considering the mean difference between both groups,  $M = -.23$ , 95% CI (-1.09, .625),  $t(42) = -.553$ ,  $p = .583$ . The findings infer that there is no difference in gender type for pre test Punjabi dictation scores.

Table 25 Native Language Reading and Word Tests (by gender)

Punjabi Dictation Test	Grouping Variable	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
	Pre Male	23	5.47	1.53	3	7
	Pre Female	21	5.71	1.27	3	7
Hindi Word Test						
	Pre Male	23	41.73	11.20	7	50
	Pre Female	21	43.28	9.37	17	50

The analyses of pre Hindi Word Test scores from Table 25 indicate that female Monitors have a slightly higher mean score ( $M = 43.28$ ) when compared to male Monitors ( $M = 41.73$ ), with a mean difference of 1.54. Variability is more spread for male Monitors when compared to female standard deviation scores. An independent t-test was run to test for statistical differences between male and female pre test scores. Visual inspection of boxplots generated from SPSS found there to be no outliers greater than 1.5 box lengths, therefore a normal distribution was assumed based on there being no violations. Homogeneity of variances was tested and no violations were found, as inspected by the Levene's test of equality of variance,  $p = .502$ . The results of the t-test found there to be no statistically significant difference between the pre male Hindi word test and the pre female

Hindi word test when considering the difference between each group,  $M = -1.54$ , 95% CI (-7.86, 4.71),  $t(42) = -494$ ,  $p = .624$ . The results infer that there was no difference between pre male and pre female Hindi word knowledge.

The next section presents the findings of Monitors' reading, phonological awareness, spelling and native language scores following the intervention period. Specifically, the following research sub question asks;

- What are the effects of Monitors' phonological processing, reading and spelling abilities when guiding learners in a peer-mediated intervention?

#### 4.7.5 Monitor Reading and Dictation Test Battery

The reading and dictation battery included tests that required the Monitors to read a word list (Burt reading), complete a written spelling test (Schonell) and read aloud 10 sentences with varying difficulty as they progressed (Dictation). Table 26 reports the normality test results using the Shapiro-Wilk's statistic  $W$  for the Reading and Dictation battery. It shows that the Burt Reading ( $p > .05$ ) and Schonell Spelling ( $p > .05$ ) were not statistically significant and would suit parametric testing, for both samples, a paired samples t-test was selected to make inferences about the data over the two time points. However, the Dictation test reported a statistical difference, thus violating a normal distribution and would be more suited to non-parametric testing. Details related to these tests and their results are presented below in Table 26.

Table 26 Reading and Dictation Tests for Normality

	<i>N</i>	<i>W</i>	<i>P</i>
Burt Reading	44	.972	.365
Schonell Spelling	44	.977	.503
Dictation	44	.454	0.00

As the Burt Reading and Schonell Spelling test were normally distributed and fell within the parameters of the central limit theorem, a paired t-test was chosen to determine if the mean difference between the paired observations are significantly different from zero. The pairings are pre-post for each test with a dependant variable of time. Firstly, it is determined that the pre test scores for each test and whether they're any differences for each group at

this time point. The pre tests were administered at registration before randomisation took place, thus a significant difference here could lead to potential contamination issues. However, it is important to note significant differences don't give the size of the difference, therefore, in addition to this, effect sizes using Cohen's (1988) norms are used for interpretation.

In the present study, we wanted to know whether there are any significant differences between time 1 (pre test) and time 2 (post test) of Monitors Burt Reading and Schonell Spelling test scores. The paired samples t-test revealed Burt reading scores improved at time 2 ( $M = 90.68$ ,  $SD = 9.38$ ) compared to time one ( $M = 86.55$ ,  $SD = 8.37$ ), resulting in a statistically significant mean increase of 4.13 when considering the 95% CI (3.00, 5.27),  $t(43) = 7.338$ ,  $p < 0.01$ ,  $d = 0.4$ . A moderate effect size, according to the norms set by Cohen (1988) but observed as practically insignificant. Tests of significance were also performed for the Schonell Spelling test with improvements observed from pre test ( $M = 60.27$ ,  $SD = 8.30$ ) to post test ( $M = 63.23$ ,  $SD = 6.58$ ), a statistically significant difference including a mean increase of 2.95 when considering the 95% CI (1.73, 4.17),  $t(43) = 4.900$ ,  $p < 0.01$ ,  $d = 0.3$ . The effect size can be interpreted as low, suggesting that while spelling did improve over time and these results were only slight over the two time points.

As presented in Table 26 the results from the Dictation test violated the assumptions required for a sample that is normally distributed, therefore, parametric tests like the paired samples t-test are not suited for making accurate and reliable inferences. Instead, a Wilcoxon signed-rank test, which is a non-parametric test, was used as it doesn't require the sample to be normally distributed. The Wilcoxon-signed rank, however, does have an assumption that requires the median of the paired samples (in this case pre – post-tests) to be symmetrical. For this study, the median was negatively skewed and thus violated the assumptions posited by the Wilcoxon signed-rank test. Therefore, a paired sample signed test, as the name suggests, analyses the positive and negative differences, from the sample median, are equal or different to zero. As in usual hypothesis testing, if the median (like the mean) is unequal to zero then we reject the null for an alternative conclusion. Table 27 illustrates the positive, negative and ties for the Monitor sample. It reports that there were  $N = 7$  positive differences against  $N = 0$  negative differences over the two time points. From the frequencies illustrated in Table 27, we can infer that none of the Monitors did worse in

their Dictation test at time 2, instead  $N = 7$  representing 16% (when rounded) improved at time point 2 when median differences are considered.

**Table 27 Sign Test Frequencies of Monitors' Dictation Pre/Post Tests**

Dictation Post Test – Dictation Post Test	<i>N</i>
Positive Differences	7
Negative Differences	0
Ties	37
Total	44

Table 27 also reports the ties that occurred over the two time points. Once again, as the name suggests, ties are data that report neither a positive nor negative difference which in practical terms that  $N = 37$  (84%) of the Monitors did not increase their level of dictation at time point 2 when measuring the median occurrences at time point 1. As there are fewer than 25 positive and negative responses, in this case 7, an exact test statistic is computed based on a binominal distribution. The exact test revealed that dictation scores at the post test stage improved over the two time points, a statistically significant result,  $p = .016$ . However, the medians of both pre and post-tests reported the same score ( $Mdn = 10.00$ ), this referring to the maximum scores each Monitor could reach on each of the tests. It is important to note that the sign test determines whether the sample medians in both pre and post tests are statistically significant to zero, not whether they are significantly statistically different from each trial. Further, perusal of Table 27 indicates that 84% of the sample data are tied.

#### **4.7.6 Phonological and Letter Character Test Battery**

These tests were conducted to fulfil the sound and letter awareness children are required to learn when forming basic reading skills. All tests failed tests of normality required for parametric testing and are therefore subjected to non-parametric methods. Firstly, a Wilcoxon signed-ranked test was used to search for distributions that had similar medians, these are presented in Table 28. The Nfer A letter matching test illustrated that the distribution of the paired differences are symmetrical by approximation, confirming a signed-ranked test to be suitable in the interpreting the results over the two time points.

However, all tests reported differences in medians illustrating graphically that they were not symmetrical in shape and therefore a paired sample signed test was used for hypothesis testing for analysing these tests.

**Table 28 Wilcoxon Signed Test on Phonological and Character Awareness Test**

	<i>N</i>	<i>W</i>	<i>P</i>
Nfer A (Letter Matching)	44	.630	.000
Nfer B (Sound values)	44	.901	.001
Nfer C (Blending 3-4 letter sounds)	44	.788	.000

### **Results of the Phonological and Character Awareness Test Battery**

Table 28 reports descriptive statistics for each of the pre and post Nfer tests. It illustrates the positive and negative ties that were observed over the two time points. Data are medians unless otherwise stated.

#### **Nfer A – Letter Matching Test**

Descriptive statistics from Table 28 show that there were 40 ties made over the two time points with  $n = 4$  students improving. This suggests that only 4 students improved their letter matching tests scores over time with the remaining 40 Monitors making no improvement. Given that there are less than 25 ties, the exact test was applied and reported no statistically significant difference of Monitors’ pre Letter Matching (Mdn = 10) and post Letter Matching (Mdn = 10),  $p = .125$ . Taken together with the descriptive statistics from Table 28 the results infer no improvement of letter matching over the two time points. This suggests that Monitors’ had already achieved basic letter recognition and these are reflected in the findings presented here.

#### **Nfer B – Phonemic Awareness Test**

The Nfer B tested for phonemic awareness and like the other tests was administered at both the pre and post time points. The descriptive statistics from Table 29 shows that there was a 100% improvement from pre/post-tests with all Monitors improving their phonemic sound awareness. A sign test, allowing for continuity correction was used to compare the differences over the two time points, given the number of ties are  $25 >$  a  $z$  score is reported



alongside the asymptotic  $p$  value which reports a statistically significant difference over the two time points. Taken together, data reveal that all Monitors' improved their sound knowledge over the two time periods.

**Table 29 Wilcoxon Signed Tests of Monitors' Phonological and Letter Character Differences**

Tests & Variable Calculation	<i>N</i>	Positive Differences	Negative Differences	Ties	Total
Nfer A Post – Nfer A Pre Test	44	4	0	40	44
Nfer B Post – Nfer B Pre Test	44	44	0	0	44
Nfer C Post –Nfer C Pre Test	44	7	0	37	44

### **Nfer C – Blending Test**

The Nfer C tested for the blending of two and three letter sounds in phonetic words. Only words that were correctly 'sounded out' are considered a correct answer and all Monitors were made aware of these instructions at both the pre - post test stage. The analysis revealed that there were 7 positive differences over the two time periods, indicating that some students made progress over time. Given that there are fewer than 25 ties, the exact test was used and reported a statistically significant difference of Monitors' pre phonemes sounds (Mdn = 0.00) and post phonemic sounds (Mdn = 34),  $z = -6.482$ ,  $p = 0.00$ .

### **Nonsense Phonemes and Decoding**

The Miskin Nonsense tests were first subjected to tests for normal distributions based on the conditions set out in conducting parametric testing. Table 30 reports the Shaprio-Wilk  $W$  test statistics and  $P$  values for each of the different types of Miskin Nonsense tests.

**Table 30 Monitors' Miskin Nonsense Tests (by type)**

Miskin Nonsense Words	<i>N</i>	<i>W</i>	<i>P</i>
Phonemes and Digraphs	44	.750	.000
3-4 Letter Sounds	44	.872	.000
4-6 Letter Sounds	44	.662	.000
4-13 Letter Sounds	44	.753	.000

Analysis revealed each of the tests is presented and report that all tests failed distribution assumptions with all being statically significant. Each of the tests were analysed individually over the two time points and with each test becoming progressively harder based on the different types of sound combinations within each word. Non parametric tests were therefore selected, firstly, a Wilcoxon signed ranked test was applied and found that medians were not symmetrical by approximation. A paired sample test was used for hypothesis testing and given that all tests had fewer than 25 ties an exact test reported Monitor's pre phonemes and digraphs (Mdn = 1.00) and post phonemes and digraphs (Mdn = 28.50) to be statistically significant,  $z = -6.482, p = 0.00$ . This suggests that Monitors improved their decoding ability of small units of sounds found in these word-types over the two time points. Furthermore, the analysis revealed that there to be a statistically significant difference between pre 3-4 letter words (Mdn = .50) and post 3-4 letter words (Mdn = 12.00),  $z = -6.482, p = 0.00$  over the two time points. Similar statistical differences reported for the pre 4-6 letter words (Mdn = 0.00) and post 4-6 letter words (Mdn = 12.00),  $z = -6.482, p = 0.00$ , pre 4-13 letter words (Mdn = 0.00) and post 4-13 letter words (Mdn = 38.00),  $z = -6.482, p = 0.00$ , again, over the two time points. These results demonstrate that the identification of sounds in unrecognizable words improved over time.

#### **4.7.7 Native Reading Test Battery**

The native reading test battery included a Punjabi dictation test and Hindi word test. Both tests fulfilled normality tests and were analysed using a paired samples t-test. The results of the Punjabi dictation test found no statistical significant differences from pre-test ( $M = 5.59, SD = 1.63$ ) results compared to post test ( $M = 5.61, SD = 2.23$ ), 95%, CI (-.023, .068),  $t(43) = 1.00, p = .323$ . The results suggest no significant improvement of Punjabi dictation scores over the period of the intervention. The same observation was made for Hindi word test scores with no improvement from pre-test ( $M = 42.47, SD = 4.23$ ) to post test scores ( $M = 42.70, SD = 3.42$ ), 95% CI (-351, .805),  $t(43), .792, p = 4.33$ . The inference made based on these results suggests native languages didn't improve as a consequence of the intervention period.

## 4.8 Descriptive & Inferential Findings of Learners' Pre/Post Test Scores

### 4.8.1 Introduction

A total of  $N = 701$  children were registered into the study and then randomly selected into control and intervention groups after the pre –test stage. Results in this section analyses the pre test data from all of the test batteries and reports any statistical differences based on gender, this being similar to the Monitor group results reported in the previous section. As such, the following research sub question guided the study by asking;

- Is there a statistical difference between male and female learners' baseline reading, phonological processing and native language test scores?

### 4.8.2 Gender Effects: Reading and Dictation Test Battery

Table 31 below reports the pre test reading battery of the registered children by gender before randomisation took place. It shows that for the pre Burt Reading test that males ( $M = 4.54$ ,  $SD = 4.32$ ) performed slightly lower than females ( $M = 4.59$ ,  $SD = 4.55$ ) although the variability for females was greater, the mean difference between each group was .531 and considered small. An independent sample t-test was ran to test for statistical differences between pre male and female Burt reading tests. No outliers were found when inspecting graphical plots from SPSS output. However, the Levene's test of equal variances reported a statistical difference,  $p = .049$ . Using the guidance from Welch (1974) based on the statistics used when considering equal variances are not assumed, the t-test results found there to be no statistically significant difference between male and female pre Burt reading test scores,  $M = .531$ , 95%, CI (-1.191, .129),  $t(699) = -1.581$ ,  $p = .114$ . These results confirm that male and female learners that made up the study population were equal in terms of basic reading ability as measure by the Burt reading test.

The pre schonnel spelling test results revealed similar findings with females ( $M = 4.25$ ,  $SD = 3.10$ ) performing slightly better at spelling than males ( $M = 4.06$ ,  $SD = 2.92$ ). A t-test was run to test for differences between both male and female pre spelling scores. No violations were found in the distribution tests, in particular, no outliers were found to deviate from the generated box plots. However, the Levene's test of equal variances was violated so the equal variances not assumed statistics were used to report the findings. The results found there to be no statistically significant difference in male and female pre

Schonell spelling scores when considering the mean difference from both groups,  $M = .194$ , 95%, CI (-.641, .252),  $t(699) = -.854$ ,  $p = .393$ . The results infer that there was no difference in spelling ability when considering gender from the study population. Finally, from the reading test battery the trend of female scores being slightly higher than male scores are mirrored in the dictation test, reporting the results of the pre female dictation ( $M = 1.21$ ,  $SD = 2.43$ ) displaying improved scores compared to pre male dictation ( $M = 1.00$ ,  $SD = 2.32$ ), both groups reporting very low average scores in total. A t-test was run to determine if there were any statistical differences between genders. No violations were reported with outliers being 1.5 box lengths greater than the top of the box plot, thus, data were assumed to be normally distributed by approximation. The Levene's test for equality of variances reported a homogenous test statistic,  $p = .748$ . The results found there to be no statistically significant difference between pre male and pre female dictation scores when considering the differences between the mean scores,  $M = .53$ , 95%, CI (-1.163, .099),  $t(699) = -1.655$ ,  $p = .098$ .

**Table 31** Learners' Pre Reading Test Battery (by gender)

Burt Reading	Grouping Variable	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
	Pre Male	345	4.54	4.32	1	16
	Pre Female	356	4.59	4.55	1	16
Schonell						
	Pre Male	345	4.06	2.92	0	13
	Pre Female	356	4.25	3.10	1	12
Dictation						
	Pre Male	345	1.00	2.32	0	10
	Pre Female	356	1.21	2.43	0	10

The reading battery found that there was no statistical differences in the gender mean scores at the pre –test stage. This finding confirms the original hypothesis that reading, spelling and oral English scores would be low for children who don't attend English medium schools and may not attend school on a regular basis. The null hypothesis is therefore retained on the basis that there are no gender differences in reading, spelling and oral dictation scores.

### 4.8.3 Gender Effects: Phonological & Character Awareness Battery

The study group were also given the same tests as the Monitors to understand more about their phonological and letter character awareness. Three tests were found in this battery and are reported here. Firstly, the Nfer A (letter matching) test was administered, requiring children to match English letter characters to gain an insight into their ability to recognise the English alphabet in its printed form. The maximum score each child could achieve was 10 marks, one mark was awarded for each time the child correctly matched the letters to the corresponding letters. The pre test for reported male mean scores ( $M = 4.78$ ,  $SD = 4.26$ ) to be slightly lower than female scores ( $M = 5.31$ ,  $SD = 4.24$ ), demonstrating that females correctly matched just over 50% of the letters on the test in comparison to males who were slightly below 50% of total correct answers. The spread of data is similar for both groups as detailed by the standard deviation from the sample. An independent t-test was run to test for statistical differences between both pre male and pre female test Nfer A test scores. No violations were reported when testing for distributional biases and no outliers greater than 1.5 box lengths were reported in the output. Data were homogenous as reported by the Levene's test for equality of variances,  $p = .748$ . The results from the t-test found there to be no statistically significant difference in pre male and pre female Nfer A (letter matching) test scores when considering the mean difference from the group  $M = .532$ , 95% CI, (-1.16, .099),  $t(699) = -1.65$ ,  $p = .098$ . In summary, the results suggest that while females did slightly better at the letter matching, there were no differences when considering gender and letter matching.

The second test, Nfer B, tested children's phonemic awareness by asking them to sound out each English letter character shown. The pre test scores reported low means for both male ( $M = .34$ ,  $SD = 1.07$ ) and female ( $M = .43$ ,  $SD = 1.04$ ). A t-test was run to test for differences between both male and female pre Nfer B test scores. No violations were found in tests for distribution biases, no outliers were reported from the generated boxplots. Data were homogenous as reported by the Levene's test for equality of variances,  $p = .06$ . The results from the t-test found there to be no statistically significant difference in pre male and pre female Nfer B test scores when considering the mean difference from both groups,  $M = .09$ , 95%, CI (-.280, .093),  $t(699) = -.986$ ,  $p = .324$ . Given the low mean scores, the inferences made from the results suggest there to be a very low average of phonemic awareness amongst the children and there were no differences in these scores when taking into account gender at the pre-test stage.

The last test within the Phonological and Character Awareness Battery asked children to blend 2/3 letter sounds. The Nfer C test carried a total score of 34 with one mark being awarded for every correctly blended word. The results from the test suggested that both means for male ( $M = .04$ ,  $SD = .293$ ) and female ( $M = .11$ ,  $SD = .836$ ) pre test scores recorded as being very low with greater variability for female scores based on the standard deviation from the mean. An independent t-test was run to test for differences between male and female pre test scores. No violations were found in tests for distribution biases and there were no outliers that could skew the distribution. However, data were heterogeneous as reported by Levene's test for equality of variances that returned a statistically significant result,  $p = .006$ . A decision was made to use the test results from the unequal variance t-test statistics based on the guidance as set by Welch (1947). The results from the test found there to be no statistically significant difference in pre male and pre female Nfer C (blending) test scores when considering the mean difference of each group,  $M = (.06)$ , 95% CI,  $(-.159, .026)$ ,  $t(433) = -1.407$ ,  $p = .160$ . The results infer that blending skills of the sample were seen to be very low and that there were no differences in male and female blending scores at the baseline test.

In the next test, the study group children were exposed to phonemes and blending of irregular words on the test sheet. The test is administered altogether but the analysis has been broken down into 4 different tests to reflecting the difficulty presented to each child. For example, the first part of the analysis looks at the results from the children's phoneme and digraph awareness. From here, three and four letter nonsense words are presented and children are asked to 'sound out' each letter sound. As the test progresses, longer words appear with different types of blends asked of the children. Only correctly blended words are sounded out and these are recorded as being correct in the final analysis.

The first test asked children to recognise phonemes and digraphs and made up the Miskin phonemes and digraph pre test scores. The reported mean for both male ( $M = .51$ ,  $SD = 1.90$ ) and female ( $M = .42$ ,  $SD = 1.32$ ) pre test scores to be very low. The variability based was greater for male pre test scores but both scores were similarly low in phoneme and digraph awareness. An independent t-test was run to test for statistical differences between baseline male and female scores. No violations were reported in distributional biases, all data were therefore normally distributed by approximation. The Levene's test for equality of variances reported a statistically significant test statistic,  $p = .212$ , thus indicating data having met homogeneity of variances. The t-test reported there to be no statistical

difference in pre male Miskin phonemes and digraphs test scores and pre female Miskin phonemes and digraphs test scores when taking into consideration the mean difference of both groups,  $M = .086$ , 95% CI ( -1.49, .322),  $t(743) = .720$ ,  $p = .472$ . The results suggest that there are no differences based on gender and that both groups of children are equally low in their phoneme and digraph awareness.

The second Miskin word test asked children from the sample to decode 3/4 sounds contained in the nonsense words. The means by gender type reported males ( $M = .12$ ,  $SD = .768$ ) as being slightly greater than females ( $M = .10$ ,  $SD = .707$ ) with similar variability based on the standard deviation from the mean. An independent t-test was run to test for statistical differences between each gender type. When observing the data, no distributional violations were reported with outliers, thus a normal distribution by approximation was concluded. Data were also homogenous after consulting the Levene's test for equality of variances with no statistical differences found,  $p = .480$ . The results of the t-test reported no statistically significant difference in pre male Miskin 3/4 sound awareness and pre female Miskin 3/4 sound awareness when considering the mean differences between each group,  $M = .019$ , 95% CI (-.087, .125),  $t(743) = .347$ ,  $p = .480$ . The results show that children from both male and female groups have a poor grasp of blending 3/4 sounds in irregular words and there are no differences between each group at the pre – test stage.

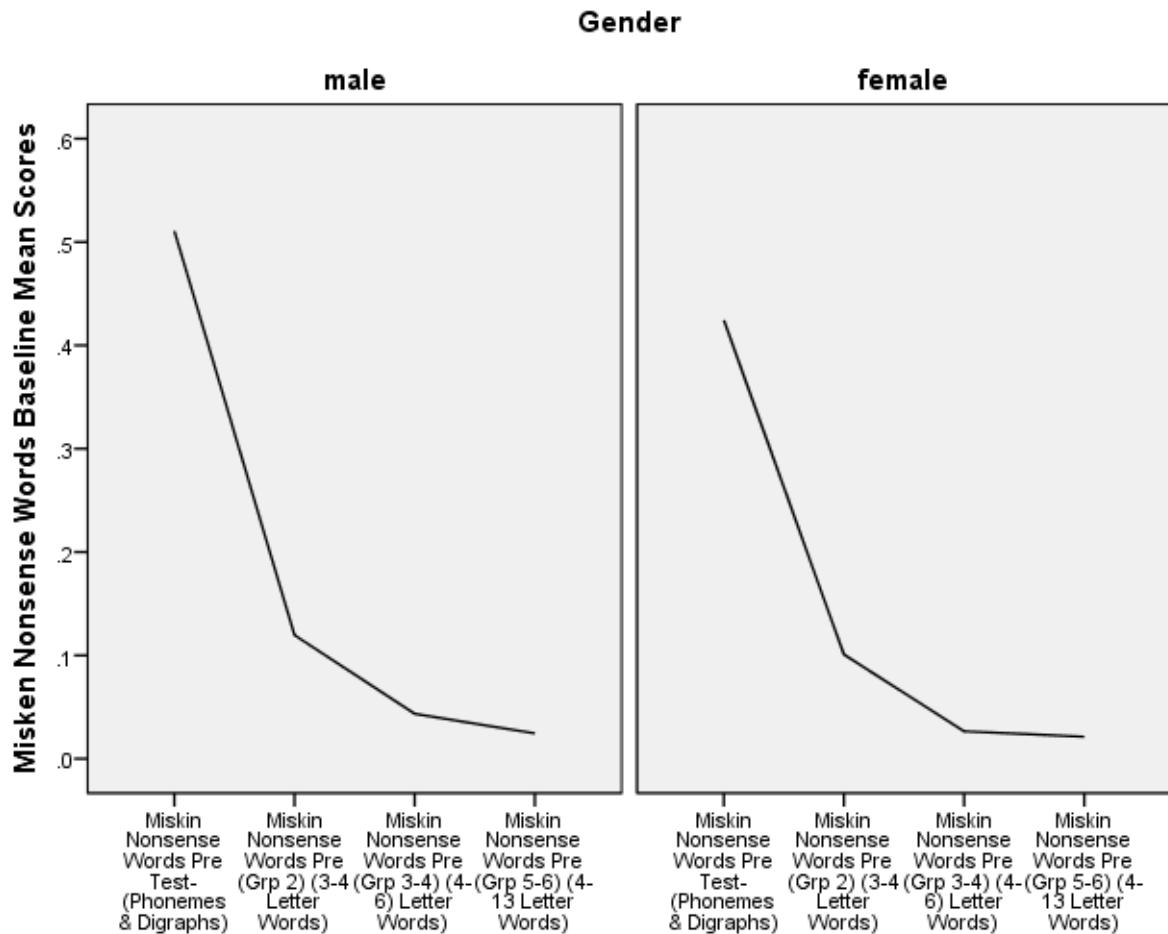
The third Miskin nonsense word test grouped words by 4-6 letter sounds. Analysis of data indicated male mean scores ( $M = .04$ ,  $SD = .422$ ) were slightly higher than pre female scores ( $M = .03$ ,  $SD = .191$ ) but with greater variability when taking into consideration the standard deviation from the mean. An independent samples t-test was run to test for any statistical differences. No violations were reported in distribution bias, no outliers contributed to any positive or negative skews in the preparation of the data. Homogeneity was met as reported by the statistical difference statistic found in the Levene's test for equality of variances that reported,  $p = 1.53$ . The results infer there to be no statistical significant difference in pre male Miskin nonsense Group 4 letters and pre female Miskin nonsense Group 4 letters when considering the mean difference between each group,  $M = .017$ , CI 95%(-.30, .064),  $t(743) = .709$ ,  $p = .478$ . Similar to the other results from the Miskin test, the low means for each group suggest that there is no basic awareness of sounds in nonsense words containing 4-6 letter sounds and that there is no difference based on gender type.

Finally, the results from the pre male and female Miskin Group 5- 13 letter sounds reported pre male means ( $M = .02$ ,  $SD = .270$ ) and pre female means ( $M = .02$ ,  $SD = .241$ ) to be the same but the variability being more spread when considering the male standard deviation score. A t-test was run to test for any statistical differences between the pre test scores of both gender groups. No violations were reported in the distribution of the sample data, in particular, no outliers were greater than 1.5 box lengths from the generated output. Using the Q-Q plots to test for distribution biases, the data were considered normal by approximation. Further to this, data were considered to be homogenous based on the Levene's test for equality of variances which reported a statistically significant result,  $p = .729$ . The results from the t-test found there to be no statistically significant difference in pre male Miskin Group 5-13 letter sounds and pre female Miskin Group 5-13 letter sounds when considering the mean difference,  $M = .003$ , 95% CI (-.034, .040)  $t(743) = .173$ ,  $p = .863$ . The results confirm that both male and female means scores are very low at the baseline level and that there are no differences when considering gender on this test type.

The Miskin Nonsense Words test battery aimed to test children's decoding ability. All of the test results reported low mean scores for the study sample indicating that children had a limited phonemic awareness and skills in decoding non-sight words. Figure 24 illustrates the trend of the mean score by gender type against each of the Miskin tests. Both male and female line graphs show a similar decreases in mean scores as words become harder and with complicated blends and sounds. The inference made from this test battery suggests that children from the sample to have a low level of phonemic awareness and knowledge of basic – difficult blends.



Figure 24 Miskin Nonsense Words Baseline Test (by gender)



Parallel to the low mean scores from each group, no statistical differences were found between male and female children, this being mainly reported through similar deviations from the mean score and the low mean overall. The main inference to come out of this test battery is both male and female children would suit a synthetic phonics trial based on the limited sound awareness and skill in blending complicated English words.

#### 4.8.4 Gender Effects: Native Language Test Battery

Two native language tests were administered to the study population, the same tests that were given to the Monitor sample with each test exposed to words from both languages. Firstly, the Punjabi test was administered where the male mean scores ( $M = 3.54$ ,  $SD = 2.13$ ) were lower compared to the female scores ( $M = 3.71$ ,  $SD = 2.02$ ), with both groups having similar variability from the mean. An independent t-test was run to test for statistical significance between gender and the Punjabi pre dictation test. No distribution violations

were observed and data were normally distributed by approximation. Homogeneity of variances was also achieved through inspection of the Levene's test of equality of variances that reported,  $p = .06$ . The results from the t-test found there to be no statistically significant difference in pre male and female Punjabi dictation scores, when considering the mean difference of each group,  $M = .162$ , 95% CI (-.4713, .145),  $t(699) = -1.03$ ,  $p = .30$ . The results suggest there to be no difference in reading Punjabi when considering gender as a factor. The same procedures were used in measuring the Hindi word test, using Gender as dependent variable. No violations to distributions were observed. The t-test found there to be no statistically significant differences between male and female Hindi word test scores, when considering the mean difference of each group,  $M = .455$ , 95% CI (-.991, .0794),  $t(699) = -1.67$ ,  $p = .117$ . Similar to the Punjabi dictation scores, the Hindi word test results indicate there to be no differences when considering gender as a factor. Indeed, both scores are observed to be extremely low, when observing the overall mean.

The next section presents the findings of the learners' reading, phonological awareness, spelling and native language scores following the intervention period. Specifically, the following research sub question asks;

- Are there statistical differences between learners' phonological processing, spelling, reading post test scores when controlling for pre test scores? If so, where do these differences lie?

Results from the Reading and Dictation Test Battery are presented first followed by Phonological Processing and Native Reading Test Batteries.

#### **4.8.5 Learners' Reading and Dictation Baseline Comparisons**

$N = 701$  children were part of the final analysis and were subjected to numerous test batteries. The results from Table 32 illustrate the baseline test results from the reading battery. The reported statistics indicate that the control group performed slightly better on the Burt reading test ( $M = 5.44$ ,  $SD = 4.44$ ) compared to the intervention group ( $M = 4.20$ ,  $SD = 4.38$ ). An independent samples t-test was run to test for statistical significance between each group against the null hypothesis that there is no difference at the Burt test baseline. Tests for distribution biases were fulfilled and were considered to be normally distributed, both baseline scores reported a statistical significance as assessed by the

Shapiro Wilks test,  $p = > .05$ . Homogeneity of variances as inspected by the Levene's test of equality of variances were assumed, also, reporting statistical significance,  $p = .603$ . The analysis revealed that there was a statistically significant difference of 1.24, (95% CI, .580, 1.88),  $t(699) = 3.705$ ,  $p = 0.00$ ,  $d = 0.2$  at the baseline level of Burt scores between each group. However, the result, while statistically significant, reported a low effect size based on Cohen's (1988) suggestions and was considered to be practically non-significant. Similarly, compared to the results of the Burt reading test, the Schonell spelling test represented the same outcome between both groups. The control group had a greater mean ( $M = 4.81$ ,  $SD = 3.26$ ) compared to the intervention group ( $M = 3.52$ ,  $SD = 2.59$ ). It implies at this stage of the analysis, the control group were marginally better at reading and the basic spelling of English words. The maximum words read correctly included a parity of 16 total words with almost parity of words that were correctly spelt. Following on from the descriptive statistics, inferential measures were undertaken to test for statistical significance differences between both groups spelling scores. Distribution biases were ruled out through inspection of the statistically significant Shapiro –Wilks test statistic,  $p = > .05$ , however, equality of variances were violated as inspected using the Levene's test,  $p = .025$ , unequal variances are reported based on the advice from Welch (1947). The final analysis revealed there to be a statistically significant difference of 1.29, (95% CI, .849,1.72),  $t(659) = 5.766$ ,  $p = .004$ ,  $d = .04$ . The results are similar to the Burt reading test in that they offer a low to moderate effect size and are considered to be practically non-significant when considering the study design and overall aims of the research.

**Table 32 Learners' Baseline Reading Test Scores**

<i>Intervention Group</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Burt Reading	354	4.20	4.38	1	16
Schonell Spelling	354	3.52	2.59	0	11
Dictation	354	1.49	2.82	0	10
<i>Control Group</i>					
Burt Reading	347	5.44	4.44	1	16
Schonell Spelling	347	4.81	3.26	1	13
Dictation	347	.71	1.73	0	10

Finally, descriptive statistics from Table 32 indicate that the intervention group were slightly better at reading sentences in the dictation test ( $M = 1.49$ ,  $SD = 2.82$ ) when compared to the control group ( $M = 0.71$ ,  $SD = 1.73$ ). The results indicate that the intervention group had a marginal advantage on the number of sentences read correctly when compared to the control group. To test the hypothesis that there are no differences between each groups' dictation scores and independent samples t-test was run. Data were normally distributed by approximation and were statistically significance when considering the assumptions normality using the Shapiro-Wilks,  $p = >0.05$ , however, equality of variances were not assumed as reported by statistically significant Levene's test,  $p = 0.00$ . Based on Welch's (1947) guidance, the reported results considered using the equality of variances not assumed calculations. It found there to be a statistically significant difference of .78, (95%, CI, -1.126, -.433),  $t(588) = -4.418$ ,  $p = 0.00$ ,  $d = .03$  baseline reading scores between both control and intervention groups. Again, similar to the other tests within this test battery, the reported effect size is interpreted as being practically low, therefore, the inference made about these results suggest there to be no practical difference at the baseline level for each group, although statistically significant.

#### **4.8.6 Learners' Reading & Dictation Effects**

A one-way analysis of covariance (ANCOVA) was run to test for post-test differences between the intervention and control group when controlling for the pre test. The independent variable, control and intervention groups have two levels. The dependent variable was the post Burt test scores and the covariate was the pre Burt test scores. The assumptions of normality and linearity were met. In particular, there was a linear relationship between pre and post raw Burt scores for each group type as assessed by visual inspection of a generated scatterplot. There was homogeneity of regression slopes as the interaction term reported a non-statistically significant result,  $F(1, 697)$ , .411,  $p = >.05$ . Standardised residuals for the groups and for the overall model was normally distributed by approximation – there were some deviations from normality but these were considered to be normal when considering the distribution of both groups and the total sample. Further, visual inspections of plotted data confirmed there to be homoscedasticity and homogeneity of variances. No outliers greater than 3 standard deviations were visualised concluding data to be a best fit for analytic ANCOVA procedures. The final results of the ANCOVA suggest

a statistically significant effect of the covariate (pre Burt test) on the dependent variable (post Burt test),  $F(1, 698), 242.584, p = 0.00$ . Furthermore, there is a statistically significant effect for group type,  $F(1, 698) = 1450.368, p = 0.00$ , with a large effect size and strong power (partial  $\eta^2 = .675$ , observed power = 1.00). With relation to Table 33, results suggest a high effect that explains 68% of the total variance at the Burt post test stage that can be accounted for by group type (control/intervention) when controlling for the pre Burt test score results.

**Table 33 ANCOVA of Burt Reading Test Controlling for Baseline Measures**

<i>Source</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Partial Eta Squared</i>
Burt pre test	7082.92	1	7082.92	242.58**	.258
Group Type	4234.76	1	4224.76	1450.56**	.675
Error	2038.00	698	29.198		

\*\* $p < 0.01$

Post hoc analysis was performed with Bonferroni adjustment to control for Type 1 error across the pairwise comparisons. Pairwise comparisons of the estimated marginal means revealed there to be a statistically significant difference between the control and intervention groups ( $p = 0.00$ ). Further analysis of the mean differences can be found in Table 34 that represents the unadjusted and adjusted mean scores for each group.

**Table 34 Unadjusted and Adjusted Burt Reading Test Means (by group)**

<i>Group</i>	<i>Unadjusted</i>			<i>Adjusted</i>	
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Control	347	5.44	4.27	4.82	.292
Intervention	354	20.14	7.73	20.57	.289

The unadjusted mean score (i.e. prior to controlling for Pre Burt test scores) for the control group ( $M = 5.44, SD = 4.27$ ) was smaller than that of the intervention group ( $M = 20.14, SD = 7.73$ ), a difference of 14.28. However, when considering the adjusted mean score (i.e. when controlling for the pre Burt test score) the difference is greater, where the mean difference of the intervention group ( $M = 20.57, SE = .289$ ) is greater than the control group ( $M = 4.82, SE = .292$ ), an increase of 15.75. Thus, the inference drawn from these results

is when the covariate (pre Burt test scores) was added to the model, a large statistical significant difference occurred between the intervention and control groups. In summary, and based on data from Table 34 children who were exposed to the intervention treatment made better progress at the post Burt reading test than those children in the control group (traditional teaching method) when controlling for the baseline Burt reading tests.

### **Schonell Spelling Test Results**

The same procedures in answering the above question were run using an ANCOVA to answer the research sub question. The independent variables remained the same, control and intervention groups were used in this regard. The dependent variable used was the post test Schonell spelling test results with the covariate being the pre test Schonell spelling test results. The assumptions of normality and linearity were met indicating a linear relationship between the pre and post spelling test scores, the results were inspected by visual inspection of generated scatterplots. Homogeneity of regression slopes was also not violated as results reported a non-statistically significant result,  $F(1, 697), 2.407, p = >.05$ . Standardised residuals for the groups and the overall model were considered to be normally distributed, these inspections confirmed there to be homoscedasticity and homogeneity of variances. Further inspection of boxplots confirmed there to be no outliers greater than 3 standard deviations concluding that all assumptions were met for analysis with ANCOVA design. The results suggested a statistically significant effect of the covariate (pre Schonell spelling test scores) on the dependent variable (post Schonell spelling test scores),  $F(1, 698), 248.447, p = 0.00$ . At the group level (control and intervention) the results indicated a statistically significant effect,  $F(1, 698) = 605.108, p = 0.00$ , with a moderate to large effect size and strong power (partial  $\eta^2 = .464$ , observed power = 1.00). The effect size suggests that 46% of the variance in post Schonell spelling test scores can be accounted for by group type, when controlling for pre Schonell spelling test scores. Table 35 breaks down the statistics in more detail.

Table 35 ANCOVA of Schonell Spelling Test Controlling for Baseline

<i>Source</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Partial Eta Squared</i>
Schonell pre test	4906.83	1	4906.83	248.44**	.263
Group Type	11950.90	1	11950.90	605.10**	.464
Error	13785.53	698	19.750		

\*\* $p < 0.01$

Post hoc analysis using the Bonferroni procedure was run with pairwise comparisons of the estimated marginal means, these data are presented in Table 36. There was a statistically significant difference between the control and intervention groups ( $p = 0.00$ ), with group type contributing to 46% of the total variances for post test Schonell test scores.

Table 36 Adjusted and Unadjusted Means for Schonell Spelling Test Means (by group)

<i>Group</i>	<i>Unadjusted</i>			<i>Adjusted</i>	
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Control	347	5.14	3.39	4.55	.241
Intervention	354	12.44	6.45	13.00	.239

Table 36 also reports the unadjusted and adjusted mean scores. Prior to controlling for the pre Schonell spelling test scores, where the intervention group ( $M = 12.44$ ,  $SE = 6.45$ ) was greater than the control group ( $M = 5.14$ ,  $SD = 3.39$ ), a difference of 7.3. Further, when considering the adjusted mean scores that included the pre Schonell spelling test results, the difference is greater where the mean difference of the intervention group ( $M = 13.00$ ,  $SE = .239$ ) was greater than the control group ( $M = 4.55$ ,  $SE = .241$ ), a difference of 8.54. Of particular note, when the pre test scores are added to the model, the control group's mean score decreases, indicating a negative movement over the two test points. The main inference drawn from the results when the covariate (pre Schonell spelling test) was added to the final model, a large statistical significant difference occurred between the control and intervention groups. Based on the adjusted mean scores the main difference occurred in the intervention group, who outperformed children who took the same test from the control group when controlling for pre test scores. To summarise, those children who were exposed to the peer-mediated phonics approach did significantly better at spelling when taking into

account the changes made between pre and post tests compared to traditional methods of teaching.

### Dictation Test Scores

To answer the above question, another ANCOVA was run to test for differences in the dictation scores of both groups that acted as the independent variables. The dependent variable used in this analysis was the post – test Dictation scores, with the covariate being the pre-test Dictation scores. Normality tests were carried out and inspected which found a linear relationship between the pre and post Dictation test scores, these results were considered normal although there were some distortion, graphical outputs pointed to a normal distribution by approximation given the large sample size and robustness of the ANCOVA procedure to non-normal data. Homogeneity of regression slopes was assumed as results reported a non-statistically significant output,  $F(1, 697), 198.83, p = >.05$ . Standardised residuals for the groups and tests of homoscedasticity and homogeneity of variances were met by approximation. No outliers greater than 3 standard deviations were reported, thus, qualifying the ANCOVA procedure and limiting any Type errors that may impede later inferences. From the output, a statistically significant effect of the covariate (pre Dictation scores) on the dependent variable (post Dictation scores) was reported,  $F(1, 698), 2164, 31, p = 0.00, r = .52$ . There was a large effect size and strong power (partial  $\eta^2 = .303$ , observed power = 1.00). The large effect size suggested 30% of the variance in post Dictation scores can be accounted for by group type when controlling for pre Dictation test scores. These statistics are reported in Table 37 and show in more detail where the variances occur in the model.

Table 37 ANCOVA of Dictation Test Scores Controlling for Baseline Measures

<i>Source</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Partial Eta Squared</i>
Dictation pre test	2471.89	1	2471.89	345.81**	.211
Group Type	11950.90	1	11950.90	302.78**	.303
Error	4989.25	698	7.148		

\*\* $p < 0.01$



Given the statistical differences in the adjusted means pairwise comparison, the Bonferroni procedure on the estimated marginal means was run to detect where the differences lie. This detailed analysis is contained in Table 38.

**Table 38 Adjusted and Unadjusted Dictation Test Means (by group)**

Group	N	Unadjusted		Adjusted	
		M	SD	M	SE
Control	347	1.09	2.21	1.40	.145
Intervention	354	5.28	4.04	4.96	.143

Table 38 details the adjusted and unadjusted mean scores for the Dictation test. The unadjusted mean scores for the intervention group ( $M = 5.28$ ,  $SD = 4.04$ ) were reported as being higher than the control group ( $M = 1.09$ ,  $SD = 2.21$ ), a difference of 4.19. However, when controlling for pre test scores, the adjusted mean score of the intervention group ( $M = 4.96$ ,  $SE = .143$ ) is statistically significantly ( $p = 0.00$ ) higher than the control groups ( $M = 1.40$ ,  $SE = 1.45$ ), a difference of 3.56. The results suggest that the difference in mean scores when controlling for the Dictation pre test scores is significantly higher for the intervention group at post test. Although the adjusted mean score for the intervention group decreases when pre Dictation scores are added to the model, the results are still significant. The moderate to large effect size from the results suggests that children who attended the intervention were more likely to succeed at the Dictation post-test (when considering the pre test scores) against children who were exposed to traditional methods.

### **Summary of Reading Battery**

The reading battery consisted of three tests that aimed to understand more about children’s basic reading and spelling ability. All tests were administered over two time points with children completing them at each interval. The analysis revealed that children from the intervention group improved their basic reading and spelling skills when exposed to a peer-cooperative phonics pedagogy. The next test battery follows this section that reports the results of phonological awareness and letter character matching.

#### **4.8.7 Learners’ Phonological and Character Awareness Baseline Comparisons**

Like in the Monitors sample, three tests were used to test for phonological awareness and letter character for the complete sample. Table 39 presents the descriptive statistics for each

of the tests and used for subsequent inferential reporting. The first test (Nfer A) tested children’s ability to recognise and match a series of English letter characters to corresponding letters on the same test sheet. Like before, children received one mark for every correct association found. Using the descriptive statistics from Table 39 it reports the intervention groups mean score ( $M = 4.31$ ,  $SD = 4.08$ ) is slightly lower when compared to the control group ( $M = 5.82$ ,  $SD = 4.30$ ). Both standard deviations show similar patterns of skewness from the mean for each group. To test for statistical differences at the baseline level for group differences, an independent samples t-test was run. No distributional biases were reported as inspected by the Shapiro – Wilks test statistic  $p = >0.05$  and supporting Q-Q plots, the results suggested a normal distribution by approximation. Homogeneity of variances was assumed by the Levene’s test for equality of variances that reported statistical significance,  $p = .234$ . The final analysis revealed there to be no statistically significant difference of .68, (95% CI, -.119, .254),  $t(699) = .712$ ,  $p = .477$  between the baseline letter matching scores of each group. The results suggest that both groups are practically equal at the baseline tests for letter matching and character identification.

Table 39 Baseline Phonological and Letter Matching (by group)

<i>Intervention Group</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Letter Matching (Nfer A)	354	4.31	4.08	0	10
Phonemes (Nfer B)	354	.35	1.34	0	16
Blending (Nfer C)	354	.15	.882	0	12
<i>Control Group</i>					
<i>Control Group</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Letter Matching (Nfer A)	347	5.82	4.30	7.00	16
Phonemes (Nfer B)	347	.42	1.15	.00	6
Blending (Nfer C)	347	.00	.00	0	10

The second test (Nfer B) tested for children’s understanding of phonemes where each child was given one mark for each correctly sounded letter. The results indicate that both groups have very low mean scores, thus inferring that there is a lack of phonemic awareness at the

basic level for each group. An independent samples t-test was run to test for statistical significant differences after preliminary analysis found there to be no distributional violations as inspected as reported by the Shapiro –Wilks statistically significant result,  $p = >.05$ , however, data were assumed to be unequal as reported by Levene’s test for equality of variances,  $p = 0.00$ , therefore, unequal measures were used in the final calculation, following the guidance provided by Welch (1947). Final results from the independent samples t-test indicated that scores were not statistically significantly higher for the control group ( $M = .42$ ,  $SD = 1.15$ ) than for the intervention group ( $M = .35$ ,  $SD = 1.34$ ),  $t(699) = .712$ ,  $p = .477$  (95% CI, -2.40, -.054). The results suggest that both groups are equal when considering phonemic awareness at the baseline level. The inferences made from the results suggest two important features, firstly, that both mean scores are low indicating that the sample has limited awareness of phonemes and individual units of sound. Secondly, that there are no differences when analysing both groups independently.

The third test (Nfer C) required the children to blend 2/3 sounds found in CVC words. Children were awarded one mark for every word sounded out phonetically correct. Based on the descriptive data, it is clear that the mean scores from the sample population are very low. An independent samples t-test was run to test for statistical differences at the baseline level for each group. Normality was assumed to be normally distributed by inspecting graphical outputs, however, equality of variances was not assumed and therefore calculations using Welch’s (1947) guidance were followed. The final results suggested that the intervention group had a statistically significant higher blending test score ( $M = .15$ ,  $SD = .882$ ) than the control group ( $M = 0.00$ ,  $SD = 0.00$ ) given the control group scored zero marks across the test,  $t(353) = -3.135$ ,  $p = .002$ ,  $d = .02$ . The inference made from the result indicates that while both groups are statistically significant, the effect size reported is very low. Therefore, when considering practical significance, the low means of both groups are not considered to be different. Given the low mean scores on the blending tests for the total sample, it didn’t come as a surprise to see similar low scores in blending. As a general rule of thumb, children who demonstrate basic phoneme awareness may exhibit the ability to bring two or more sounds together to form basic words. However, from the baseline results of the phoneme and blending tests reported that the sample population were yet to develop these skills. The analysis then investigated if there was a correlation between the low phoneme and blending test scores. To test this, a Pearson’s product moment correlation was run to assess the relationship of pre phoneme (Nfer A) and pre blending (Nfer B) test scores

across the sample population. Preliminary analysis revealed that both variables were normally distributed by approximation given the large data set (N= 701). The results reported a moderate correlation between pre phonemes scores and pre blending scores,  $r(699) = .525, p = 0.00$  with pre phoneme test scores explaining 28% of the variation in pre blending test scores. The results suggest that children had no phonemic awareness and blending knowledge and that these two tests scores are positively correlated. The next test battery reports the results of the Miskin Nonsense words test.

### Baseline Miskin Nonsense Words Test

Table 40 presents the descriptive statistics for the Miskin Nonsense word tests used in the study. Like in the Monitor sample, only correctly sounded words would elicit one mark for each of the tests. Each test got progressively harder by adding different sound combinations and words became longer.

Table 40 Baseline Miskin Nonsense Word (by group)

<b>Intervention Group</b>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Phonemes and Digraphs	354	.62	2.12	0	26
3-4 Letter Words	354	.15	1.00	0	12
4-6 Letter Words	354	.05	.452	0	7
4-13 Letter Words	354	.04	.359	0	4
<b>Control Group</b>					
<b>Control Group</b>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Phonemes and Digraphs	347	.16	4.30	0	5
3-4 Letter Words	347	.00	.00	0	0
4-6 Letter Words	347	.00	.00	0	0
4-13 Letter Sounds	347	.00	.00	0	0

The analysis concentrated on the first groups of words included phonemes and digraphs for children to locate. The low mean scores for each group indicate that children from the sample population had severely limited phoneme and digraph awareness in nonsense words

at the baseline level. To test for statistical differences, an independent samples t-test was run with no violations in normality but with adjustments to the calculations where equal variances were not assumed,  $p = 0.00$ , guidance from Welch (1947) was followed when reporting equality of variances not assumed statistics. The final results indicate that the intervention group ( $M = .62$ ,  $SD = 2.12$ ) had a statistically significant higher score than the control group ( $M = .16$ ,  $SD = .430$ ),  $t(431) = 3.800$ ,  $p = 0.00$ ,  $d = 0.1$ . The low effect size confirms that while there is a statistical difference between each group on the baseline scores, there is no practical significance when considering study assumptions and random assignment. The low mean scores right across the board confirm this inference.

The second test required children to blend 3-4 letter-sounding words. The low mean scores from the 3-4 letter test scores indicate that children struggled to find the associated sounds contained in the words. To test for group differences at the baseline level an independent samples t-test was run, data were normally distributed by inspection of graphical outputs from SPSS, however, Levene's test of equality of variances were statistically significant,  $p = 0.00$ . Based on Welch's (1947) guidance, calculations and reported statistics followed the equality of variances not assumed results. The final results concluded that the intervention group ( $M = .15$ ,  $SD = 1.00$ ) had a statistically significant higher score at the baseline level compared to the control group ( $M = 0.00$ ,  $SD = 0.00$ ), 95% CI (-.252, -.042),  $t(353) = 2.751$ ,  $p = .006$ ,  $d = .02$ . The results indicate a statistical difference but when considering the low effect size, the practical significance in relation to the study goals and randomisation process revealed there to be no effects at the baseline level for 2-3 sounding nonsense words.

Analysis on Miskin 4-6 sounding nonsense words also recorded low mean scores at the baseline level. With reference to Table 40 the results suggest that the intervention group performed slightly better than the control group when considering the mean and range of scores. To test for statistical differences, an independent samples t-test was run, results from the assumption of normality were met through inspection of graphical outputs. However, Levene's test of equality of variances were violated ( $p = .000$ ), therefore, equality of variances not assumed, calculations and statistics were used in the final output as suggested by Welch (1947). The results of the t-test indicate the intervention group ( $M = .05$ ,  $SD = .452$ ) had a statistically significant higher baseline score compared to the control group ( $M = 0.00$ ,  $SD = 0.00$ ), 95% CI (-.252, -.042),  $t(353) = -2.571$ ,  $p = .006$ ,  $d = 0.1$ . The results

suggest differences between each group's baseline scores but when considering the low effect size, this is considered to be practically non-significant.

Lastly, analysis of the Miskin 4-13 sounding nonsense words was also predictably low. Table 40 outlines the descriptive statistics between both groups and highlights the low mean scores achieved. Analysis of an independent samples t-test after testing for normality assumptions and following Welch's (1947) guidance on using equality of variances not assumed statistics, the final results reported the intervention group ( $M = .04$ ,  $SD = .359$ ) had a statistically significant higher baseline score compared to the control group ( $M = 0.00$ ,  $SD = 0.00$ ), 95% CI (-.077, -.002),  $t(353) = -2.074$ ,  $p = .039$ ,  $d = 0.1$ . Similar to the other analysis from the other test scores, the low reported effect size limits the strength of a statistically significant result. Instead, this result, like the others, indicates a practically non-significant difference suggesting that although there are differences between each group, these differences are small and are left to chance.

#### **4.8.8 Learners' Phonological and Character Awareness Effects**

Another ANCOVA was run to test if the difference between the post test scores of the letter matching test differed when controlling for pre tests, and where these differences were found. Thus, the dependent variable (post test Letter Matching scores) was selected along with the covariate (pre-test Letter Matching scores). Before the analysis was finalised, several normality test were run which confirmed a linear relationship by approximation. The scatterplots revealed some deviations but these were regarded as normal based on the large sample size and the robustness of ANCOVA procedures. The assumption of homogeneity of regression slopes reported a non-statistically significant result,  $F(1, 697)$ , 53.31,  $p > .05$ . Further inspection of the results found the standardised residual plots to suggest a normal distribution with homoscedasticity and homogeneity of variances also meeting the required assumptions. However, 3 cases were removed from the final analysis as these data points affected the overall plots of fulfilling normal distribution. The cases were checked for possible coding errors but remained consistent to original test papers and backup data files. The analysis was run with inclusion of and exclusion of all three outliers with the final result being the same. A choice was therefore to leave out the 3 outliers to fulfil the assumptions of normality and to make accurate inferences. The final results suggest a statistically significant effect of the covariate (pre Letter Matching scores) on the dependent variable (post Letter Matching scores),  $F(1, 695)$ , 59.309,  $p = 0.00$ ,  $r = .24$  with both variables (group type and pre test Letter matching scores) contributing 24% of the

variability of post test Letter Matching scores. This was considered a large effect size when considering the contribution each variable makes to the final model. Table 41 breaks down the model further by looking at how each variable contributed. After adjustment of pre test Letter Matching scores, the results inferred a statistically significant difference between post letter matching test scores between group type  $\eta^2 = .08$ . Furthermore, the covariate (pre test) also provides a statistically significant result  $\eta^2 = .22$ , suggesting pre test scores is a higher contributor to the overall model, in this case, greater than group type.

**Table 41 ANCOVA of Letter Matching Test Scores Controlling for Baseline Measures**

<i>Source</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Partial Eta Squared</i>
Letter Matching pre test	4138.83	1	4138.83	198.84**	.222
Group Type	1271.65	1	1271.65	61.096**	.080
Error	1452.81	698	20.815		

\*\* $p = <0.01$

Table 42 reports the different mean scores calculations based on the unadjusted and adjusted means for the post Letter Matching test. If we take the unadjusted mean scores for the intervention group ( $M = 8.74$ ,  $SD = 5.95$ ) and compare them to the control group ( $M = 6.87$ ,  $SD = 4.21$ ) it is assumed that that the intervention group performed better at the post test stage, a mean difference of 1.87. Post hoc analysis to look for pairwise comparisons was performed using the Bonferroni procedure. It reported that with the inclusion of the pre test covariate scores, the adjusted mean scores for the intervention group ( $M = 9.17$ ,  $SE = .244$ ) compared to the control group ( $M = 6.43$ ,  $SE = .247$ ) has a statistically significantly ( $p = 0.00$ ) higher mean difference of 2.74.

**Table 42 Adjusted and Unadjusted Means for Letter Matching (by group)**

<i>Group</i>	<i>Unadjusted</i>			<i>Adjusted</i>	
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Control	347	6.87	4.21	6.43	.247
Intervention	354	8.74	5.95	9.17	.244

The overall findings suggest that while there is a statistically significant difference between control and intervention groups, children who performed better on the pre-test were more likely to do better overall. The intervention group had a statistically significant higher mean score that increased when including the pre-test scores. The large effect sizes reported conclude that children who attended the intervention programme were more likely to improve their letter matching scores when compared to traditional methods.

Phonological awareness test scores were then subjected to analysis and ANCOVA procedures to test for differences between the control and intervention groups. More specifically, by running the ANCOVA were there statistical differences between the groups mean scores when controlling for both pre-test scores. The dependent variable (post test Phonological test scores) was used alongside the fixed effect (group type) while controlling for the covariate (pre Phonological test score). Normality results suggested a linear relationship after the removal of seven outliers that were considered to be extreme. Parallel testing was completed to compare whether omitting the outliers affected the final model. Apart from the means and adjusted mean scores, the same result was reported and a decision to remove the outliers to fulfil normality assumptions was made. From this, a linear association by approximation was inspected and fulfilled via scatterplots and other graphical output. The assumption of homogeneity of regression slopes reported a non statistical significant results,  $F(1, 690), 14.422, p >.05$ . Data were considered to have homoscedasticity and homogeneity of variances meeting normality assumptions. Following the normality test results, there was a statistically significant effect of the covariate (pre Phonological test scores) on the dependent variable (post Phonological test scores),  $F(1, 691), 1145.12, p = 0.00, r = .657$  with group type (control/intervention) and pre Phonological test scores contributing 66% of the variability of post test scores, a large effect size. Table 43 displays the results of how much each variable contributes to the final model.

**Table 43 ANCOVA of Phonological Test Scores Controlling for Baseline Measures**

<i>Source</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Partial Eta Squared</i>
Phonological pre test	842.780	1	842.780	9.765	.014**
Group Type	1145.12	1	1145.12	1326.78	.658**
Error	5963.34	691	86.309		

\*\*= $p = 0.01$



Following the adjusted Phonological pre test scores, there was a statistically significant difference between group type and post test scores,  $F(1, 691), 1326.78, p = 0.00, \eta^2 = .658$ . The covariate (pre test) scores when controlling for post test scores also reported a statistically significant result,  $F(1, 691), 9.765, p = 0.00, \eta^2 = .014$  but contributed a lesser effect to the overall model when compared to group type.

The unadjusted and adjusted mean scores for the Phonological tests are found in Table 44. Taking the unadjusted mean score prior to controlling for pre-test scores, the control group results ( $M = 0.01, SD = .162$ ) are extremely low when compared to the intervention group ( $M = 25.61, SD = 13.163$ ), suggesting the intervention group had a positive mean difference of 25.60.

Table 44 Adjusted and Unadjusted Means for Phonological Test Scores (by group)

Group	Unadjusted			Adjusted	
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Control	347	0.01	.162	-0.64	.501
Intervention	354	25.61	13.163	25.68	.497

Post hoc analysis using the Bonferroni procedure found the difference increases when controlling for pre test (covariate) test scores, where the intervention group improves ( $M = 25.68, SE = .497$ ) with a statistically significant ( $p = 0.00$ ) mean increase of 26.32 compared to the control group ( $M = -.64, SE .501$ ). While pre test scores contributed to the final model, a large effect size was reported based on group membership. The intervention group had a higher mean score at post test and when controlling for pre test scores. This suggests that the likelihood of becoming phonemically aware was down to which group children participated in and not entirely down to your previous knowledge.

The next analysis focused on blending skills of the two samples. Children were pre and post tested and scores were recorded and prepared for analysis. An ANCOVA was run to test if the differences between the post test Blending scores were statistically different amongst group type when controlling for pre-test Blending scores. Investigating the results

of normality found there to be a linear relationship of the data by approximation. Scatterplots revealed some deviations but these were acceptable given the large sample sizes for both groups and the trends on display. The assumption of homogeneity of slopes reported a non statistically significant result  $F(1, 698) 693.92, p > 0.05$ . Tests of homoscedasticity and homogeneity were met and no other violations from the data being non normal were observed from these results. Final analysis revealed after adjustment for Blending pre test scores, there was a statistically significant difference in Blending post-test scores between the group types,  $F(1, 698), 693.92, p = 0.00, r = .499$ . With the inclusion of the covariate in the model, there was a large effect size that explained 49% of the total variance. Further analysis from Table 45 revealed statistically significant effects for Blending pre test scores,  $F(1, 698), 14.592, p = 0.00$ , partial  $\eta^2 = 0.02$  and group type,  $F(1, 698), 693.92, p = 0.00, \eta^2 = .499$ , with the later contributing more to the final model.

**Table 45 ANCOVA of Blending Test Scores Controlling for Baseline Measures**

<i>Source</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Partial Eta Squared</i>
Blending Pre Test	1421.11	1	1421.11	14.592	.020**
Group Type	6757.98	1	6757.98	693.92	.499**
Error	6797.68	698	97.388		

\*\*= $p = 0.01$

Table 46 reports the unadjusted and adjusted mean scores for the Blending tests. It shows that the unadjusted scores (before controlling for pre test covariate scores) for the intervention group ( $M = 20.11, SD = 14.02$ ) were higher than the control group ( $M = 0.00, SD 0.00$ ) due to this group not scoring at the post-test stage. The result suggests that without controlling for pre test scores, the control group had no basic awareness of blending skills based on the post-test scores. Therefore, the post-test scores of the intervention group are statistically significantly higher.

**Table 46 Adjusted and Unadjusted Means for Blending Test Scores (by group)**

Group	<i>Unadjusted</i>			<i>Adjusted</i>	
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Control	347	0.00	0.00	.169	.532
Intervention	354	20.11	14.02	19.94	.526

The adjusted mean scores for both groups are also reported with the standard error in Table 46. Means have been adjusted following a Bonferroni procedure to estimate marginal means. The results include the covariate (Blending pre test) to predict if any differences occur at the post test stage between both groups. When adding the covariate to the model, the intervention group ( $M = 19.94$ ,  $SE = .526$ ) produced a higher mean compared to the control group ( $M = .169$ ,  $SE = .532$ ), a positive difference of 19.41. When pre –test scores were added to the model the control group appeared to perform slightly better at pre test stage, however, the small increase is due to the conservative calculations performed by Bonferroni methods of estimating marginal means. Similarly, the intervention group appear to perform slightly worse when pre test scores are added, the same principle of adjustment applies here too. Nevertheless, the results infer that the intervention group had a statistically significant higher mean score than the control group at post test. Based on these results, the likelihood of children being able to blend basic 2/3 letter words together would be suited to peer-mediated pedagogies compared to traditional methods of teaching and learning.

Three tests formed the Phonological and Letter Character battery that aimed to test children's letter character, phonological and blending skills. Both groups performed reasonably well on identifying letters and matching them to the corresponding characters, it infers that children have a basic knowledge of a foreign alphabetical system.

#### **4.8.9 Miskin Tests Over Time**

Tests were conducted on the decoding and segmenting abilities of both groups using the Miskin Nonsense word test. The control group reported a zero score on the pre and post-tests, thus the difference between the groups would indicate that even with a marginal gain from the intervention group the results would be significant. To investigate this further, a paired-samples t-test was performed that aimed to test for statistical differences between the two time points (pre-post). Each test was subjected to tests of normality and equal variances. Preliminary results suggested that each of the 4 Miskin grouped words were normally distributed by approximation. Group 4 words had 3 outliers but these were considered not to be extreme based on the large sample size used in the analysis and the robustness of a paired t-test in this regard. Each of the Miskin Nonsense word tests for the intervention group underwent this procedure, the results of which are detailed in the following sections.

The number of letter and sound combinations found in each word were grouped by difficulty. Therefore, the Miskin Nonsense word Group 1 tests asked children to sound out phoneme and digraph sounds. Group 2 – 4 asked to find different sounds (trigraphs) in each of the longer words where the tests got progressively harder.

To test the whether there were statistical differences between the pre and post phonemes and digraphs (Group 1 Sounds), analysis using a paired samples test are reported. Table 47 indicates that the Group 1 sounds reported low mean scores at the pre-test ( $M = .62, SD = 2.12$ ) but improved at post test ( $M = 21.32, SD = 10.84$ ), a statistically significant difference of 20.6, 95% CI (19.578, 21.812),  $t(353) = 36.42, p = 0.00, d = 2.6$ . The results suggest a very large effect size over the two time points, thus, indicating children from the intervention group have a better grasp of phonemes and digraph sounds at the post test stage and these were statistically significant from pre –test scores. Group 2 sounds, made of 3-4 letter words required children to segment and decode each sound they came in contact with. The pre test scores reported a low mean ( $M = .15, SD = 1.00$ ) compared to post test scores ( $M = 6.56, SD = 5.15$ ) a statistically significant mean difference of 6.41, 95% CI (5.878, 6.947),  $t(353) = 25.58, p = 0.00, d = 1.7$ . The large effect size between the two time points suggests that children were able to segment and decode 3-4 words, an improved statistically significant difference before the intervention period.

Table 47 Descriptive Statistics and T-Test Results for Miskin Nonsense Word Groups (1-4)

Miskin Test	N	Pretest		Posttest		95% CI for Mean Difference	t	df
		M	SD	M	SD			
Group 1 Sounds	354	.62	2.12	21.32	10.84	19.578, 21.812	36.42**	353
Group 2 Sounds	354	.15	1.00	6.56	5.15	5.878, 6.947	23.58**	353
Group 3 Sounds	354	.05	.452	3.88	5.00	3.304, 4.340	14.51**	353
Group 4 Sounds	354	.04	.359	5.71	10.30	4.600, 6.759	10.38**	353

\*\* p < 0.01

Group Sounds 3 required children to decode and segment 4-6 letter nonsense words. Pre-test scores indicated a low mean score ( $M = .05, SD = .452$ ) which was found to be

statistically significantly different to post-test scores ( $M = 3.88$ ,  $SD = 5.00$ ), a mean difference of 3.8, 95% CI (3.304, 4.340),  $t(353) = 14.51$ ,  $p = 0.00$ ,  $d = 0.9$ . The large effect size infers that there was a sizeable difference in children's abilities to decode and segment 4-6 letters at post-test compared to pre test scores. Finally, Group 4 Sounds required children to decode and segment 4-13 letter nonsense words. The extremely low mean at pre test ( $M = .04$ ,  $SD = .359$ ) was statistically significantly different at post-test ( $M = 5.71$ ,  $SD = 10.30$ ), a mean difference of 5.67, 95% CI (4.600, 6.759),  $t(353) = 10.38$ ,  $p = 0.00$ ,  $d = 0.5$ . The moderate effect size infers that while there was a statistically significant difference between pre and post-tests of 4-13 nonsense letters, the variability was more spread when considering Group 2 and 3 sounds. The results show Group 3 sounds to have a lower mean than Group 4 sounds. Given the variability for these tests, it's reasonable to make assumptions about children having difficulty with decoding and segmenting longer, more challenging words. The next section reports the native reading tests.

#### 4.8.9.1 Learners' Native Reading Test Baseline Comparison

Two tests were also conducted in two other languages, the official State language, Punjabi, and one of the other National languages, Hindi. Instructions and responses were given and received in both of the languages with the featured descriptive statistical results reported in Table 48. The Punjabi dictation test included 7 sentences that would be commonly found in Standard 3 class text books. One mark was given for every correct sentence read.

Table 48 Descriptive Baseline Statistics for Native Language Tests (by group)

<i>Intervention Group</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Punjabi Dictation Test	354	3.53	2.13	0	7
Hindi Word Test	354	9.20	3.23	0	15
<i>Control Group</i>					
<i>Control Group</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Punjabi Dictation Test	347	3.72	2.01	0	7
Hindi Word Test	347	9.74	3.94	0	50

At baseline on both tests the control group had a slight advantage compared to the intervention group. Independent samples t-test was run that first examined distribution biases with no violations, data were normal. However, Levene's test for equality of variances reported a statistically significant difference,  $p = .039$ , therefore, guidance from Welch (1947) was followed when reporting the final results. The Punjabi dictation test elicited a low-mid mean raw score with the control group ( $M = 3.72$ ,  $SD = 2.01$ ) but still achieving a higher baseline score compared to the intervention group, ( $M = 3.53$ ,  $SD = 2.13$ ), however, the result was not statistically significant 95% CI (-.124, .4921),  $t(698) = 1.171$ ,  $p = .242$ . The results suggest that there was no difference in Punjabi language dictation between both groups at the baseline level.

A similar trend was noted for the Hindi word test that required children to read a total of 50 words with progressive difficulty. After passing the assumptions of normality tests and testing for equality of variances,  $p = .886$ , the final results indicated that the control group ( $M = 9.74$ ,  $SD = 3.94$ ) reported a higher baseline score that was statistically significant compared to the intervention group ( $M = 9.20$ ,  $SD = 3.23$ ), 95% CI (.008, 1.07),  $t(699) = 1.99$ ,  $p = .047$ ,  $d = 0.1$ . However, given the very low effect size, the main inference from these results suggests that there is no practical significance when considering the study aims and any advantages that may have occurred before randomisation took place. The Hindi word test results were very low for both groups but both tests reported that there was some knowledge of basic literacy in the home languages at the baseline level and these were similar within the analysis of both groups.

#### **4.8.9.2 Learner' Native Reading Effects**

The next battery of tests wanted to see if there were any differences made to native languages between each group and method of teaching instruction. An ANCOVA was run to test for these differences between both tests for the effects on Punjabi dictation scores. The results of normality testing revealed normally distributed groups with linear relationship based on visual inspection of generated scatterplots and supported statistics. Homogeneity of slopes reported a statistically significant result,  $F(1, 697)$ , 2.056,  $p = > 0.05$  indicating regression slopes are equal and would fit an ANCOVA regression. Further test results found homoscedasticity and homogeneity of variances to meet criteria of equal residuals and confirm data to be normally distributed. No significant outliers were visualised from all of the generated test results, including graphical plots. After adjustment to pre Punjabi test scores, the main effect of group type was not statistically significantly different in post Punjabi test scores,  $F(1, 698)$ , 1.620,  $p = .203$ ,  $r = .96$  Pre test scores (covariate) were statistically significant,  $F(1, 698)$ , 1675.64,  $p = 0.00$ ,  $\eta^2 = .96$ , contributing 96% of the total variance in the final model, but these differences only have a statistically

significant relationship with post test scores and not the final model as reported above and in Table 49.

**Table 49 ANCOVA of Punjabi Post Test Scores Controlling for Baseline Measures**

<i>Source</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Partial Eta Squared</i>
Punjabi Pre Test	2835.28	1	2835.28	1675.64	.000**
Group Type	.274	1	.274	1.620	.203
Error	118.106	698	.169		

\*\*= $p = 0.01$

There was no need to run post hoc tests given the final model wasn't statistically significant. Instead, the unadjusted mean scores are reported in Table 50 that shows at post-test there was very little difference in the overall mean score. To test these assumptions further, a paired samples t-test was run on each group to investigate whether post test results were statistically different from pre-test scores. The difference to ANCOVA procedures here is based on analyzing the groups separately, mainly because the paired samples t-test can only handle one group at a time. The results of the ANCOVA confirm there are no statistically significant mean differences between group type but how did each group perform (taken separately) over time?

**Table 50 Adjusted and Unadjusted Means for Punjabi Dictation Test Scores**

<i>Unadjusted Mean Scores</i>			
<i>Group</i>	<i>N</i>	<i>M</i>	<i>SD</i>
Control	347	3.85	2.01
Intervention	354	3.72	2.09

Two separate paired samples t-tests were run to investigate if each group improved over the two time points. The intervention group was subjected to normality tests before confirming the validity of the final results. A new variable (difference) was calculated that subtracted the post-test scores from the pre test scores. Normality graphical test outputs

detected 7 outliers; inspection of these revealed them not to be too extreme and they were kept in the final analysis. Further visual inspection of Q-Q plots confirmed that the difference in pre and post Punjabi dictation results to follow a normal distribution by approximation. Children in the intervention group performed better on the post-test ( $M = 3.70$ ,  $SD = 2.09$ ) as opposed to the pre-test scores ( $M = 3.53$ ,  $SD = 2.13$ ), a statistically significant mean increase of .180, 95% CI (.1347, .2268),  $t(353) = 7.726$ ,  $p = 0.00$ ,  $d = 0.08$ . The small effect size based on Cohen's (1988) classifications suggest there to be a minor change from pre-test to post-test but this is not practically significant.

The control group were also subjected to paired samples procedures based on their Punjabi dictation results. The same measure of difference was computed and used for normality testing. The results suggested that 4 outliers were considered to be extreme and lead to distribution problems. These were removed from the final analysis after running the tests that reported the same significance levels when included. The choice to remove these values was made due to one of them having a negative result at the post test, leading to a skew in the data set. It was concluded that this particular data point was an error in the input stage and was thus removed. The other 3 outliers were removed due to the same issues with deviation from normality. Visual inspection from generated Q-Q plots confirmed the adjusted data to be normally distributed by approximation. Children from the control group performed better on the post test ( $M = 3.85$ ,  $SD = 2.02$ ) compared to the pre test ( $M = 3.72$ ,  $SD = 2.03$ ), a statistically significant mean difference of .125, 95% CI (.0901, .1605),  $t(342) = 7.001$ ,  $p = 0.00$ ,  $d = 0.06$ . The small effect size demonstrates no practical significance when considering the two tests over the two time points.

The next test in the native language battery (Hindi word test) wanted to investigate if any differences occurred between control and intervention group mean scores, controlling for pre test scores as a covariate. An ANCOVA was selected to addresses these parameters and to test for statistical differences. Data were setup as follows; the fixed effect was group type, with the dependent variable post-test scores and the covariate (control factor) pre-test scores. Normality tests uncovered 1 extreme outlier that didn't fit the regression line from the generated scatterplot. A decision was made to keep the outlier in after running the final analysis that reported the same statistical results when the outlier was removed. Taking this into consideration, the output revealed there to be normally distributed groups with linear relationship based on visual inspections of the available scatterplots. Homogeneity of



slopes reported a statistically significant result, indicating equal measures on the regression line,  $F(1, 697), .073, p = > 0.05$ . Homoscedasticity and homogeneity of variances met the criteria of equal residuals and confirm data to follow a normal distribution. After performing an adjustment from pre test Hindi scores, the main effect of group type was not statistically significantly different to post Hindi test scores,  $F(1, 698), 1.198, p = .274, r = .81$ . The covariate (pre Hindi test scores) reported a statistically significant effect,  $F(1, 698), 2978.05, p = 0.00, \eta^2 = .81$ , contributing 81% of the total variance in the final model, however, these differences only have statistical significance with post test scores. The final statistics are reported in Table 51.

**Table 51 ANCOVA of Hindi Post Test Scores Controlling for Baseline Measures**

<i>Source</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Partial Eta Squared</i>
Hindi Pre Test	9425.19	1	9425.19	2978.05	.810**
Group Type	3.719	1	3.719	1.198	.002
Error	2209.08	698	3.165		

\*\*= $p = 0.01$

No post hoc analysis was required due to the final model not being statistically significant. Therefore, the unadjusted mean scores are reported in Table 52. It shows the control group having a slightly higher mean score and greater variability compared to the intervention group.

**Table 52 Adjusted and Unadjusted Means for Hindi Test Scores (by group)**

<i>Unadjusted Mean Scores</i>			
<i>Group</i>	<i>N</i>	<i>M</i>	<i>SD</i>
Control	347	9.50	4.40
Intervention	354	9.09	3.72

In order to test for differences made over two time points, a paired samples t-test was run for each group. Firstly, the intervention group was subjected to test to see if data were normally distributed. A new variable was calculated (difference) from subtracting the post

test scores from the pre test scores and used in the analysis. Normality tests were conducted and reported that data were normally distributed by approximation. No significant outliers were visualized from the graphical outputs. Tests of homogeneity were conclusive in that data followed a normal distribution and would be suited based on the assumptions required to conduct a paired samples t-test. The final analysis found that children in the intervention group performed worse at the post test stage ( $M = 9.09$ ,  $SD = 3.72$ ) compared to the pre test stage ( $M = 9.20$ ,  $SD = 3.23$ ) and these results were not statistically significantly different with the mean decreasing between over two time points,  $.107$ , 95% CI  $(-.2937, .07901)$ ,  $t(353) = -1.133$ ,  $p = .258$ . The results suggest that even when considering the small decrease in the mean at time two (post test), there was no difference in the test scores for the intervention group over the intervention period. In summary, the peer-mediated phonics programme had no effect on the intervention group's basic Hindi word knowledge, with the scores remaining constant.

A paired samples t-test was also performed for the control group to investigate what differences (if any) were made over time from pre test to post test. Similar to the intervention group, a new variable (difference) was created and used to test for normality. The results of tests of normality found the data to be normally distributed by approximation. Using Q-Q plots and other graphical outputs, there was no significant outliers that would impede the levels of distribution. Furthermore, tests of homogeneity of variances was met and concluded that data would suit a paired samples procedure based on the large sample size and that data fit regression lines by approximation. Results of the paired t-test reported a statistically significant decrease in the difference from pre-test scores ( $M = 9.74$ ,  $SD = 3.94$ ) compared to post test scores ( $M = 9.50$ ,  $SD = 4.40$ ), a difference of  $1.77$ , 95% CI  $(-.4322, -.05754)$ ,  $t(346) = -2.571$ ,  $p = .011$ ,  $d = .05$ . The results suggest that the mean score for the control group over the two time points decreased and that these results are statistically significant. However, the small effect size gives a practical response that infers that the overall difference is marginal and that there are no differences over time. In summary, the control group didn't improve their Hindi word knowledge when being exposed to traditional teaching methods. Furthermore, there was no practical significance in the decrease of the overall mean score at time two. In other words, there is little evidence in the present results to suggest that children did practically worse overall on their Hindi word test.

## 4.9 Learners' Attitudes to Reading English in Academic and Recreational Settings

### 4.9.1 Introduction

In this section baseline reading satisfaction scores are reported from control (N = 347) and intervention (N = 354) groups before randomisation took place. The following sub research question was set up to understand;

- Are there differences in control and intervention group's recreational, academic and total English reading satisfaction scores?

### 4.9.2 Control Group Pre/Post Academic and Recreational Findings

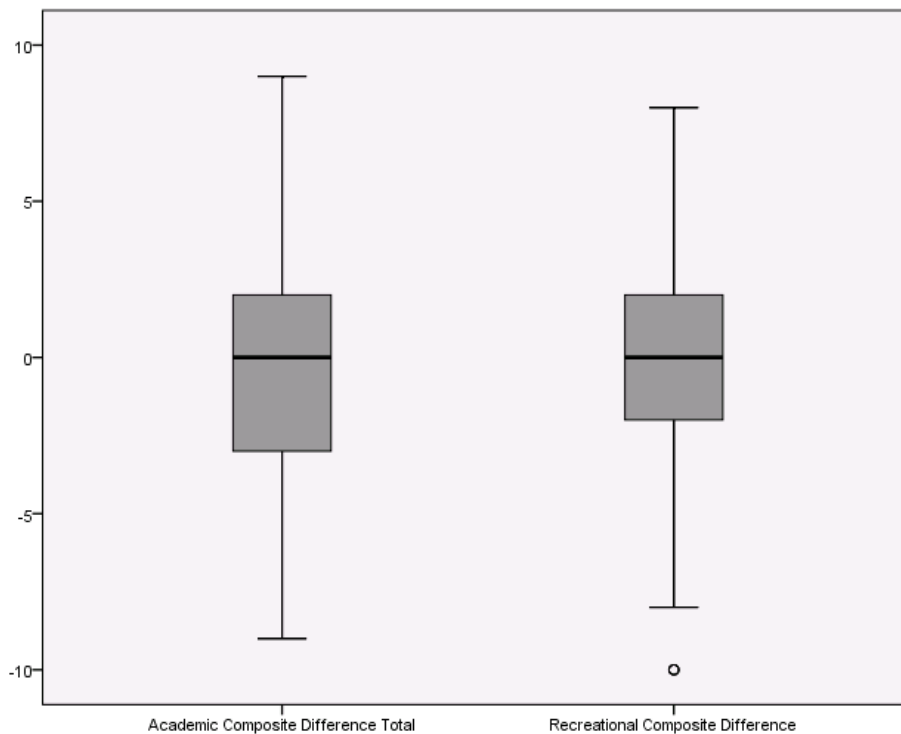
Table 53 presents the control group's mean, standard deviation, minimum and maximum satisfaction for pre recreational and academic composite scores. It shows there to be extremely low means for both recreational and academic attitude reading. This is further reflected in the minimum and maximum columns showing low scores over the two time points.

Table 53 Baseline Recreational, Academic and Total English Reading Attitude Scores

<i>Construct</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Minimum</i>	<i>Maximum</i>
Pre Recreational	347	12.97	2.83	10.00	20.00
Pre Academic	347	14.96	1.64	11.00	20.00
Totals (rec + aca)	347	27.94	3.26	21.00	37.00
Post Recreational	347	13.19	2.73	10.00	19.00
Post Academic	347	14.61	2.80	10.00	20.00
Totals (rec + aca)	347	27.81	4.17	20.00	38.00

These measures can also be examined through the difference calculated between pre/post variables. Figure 25 illustrates a similar spread of the differences with both for recreational and academic parallel at 0 median. The spread of data indicates that for the control group there was no movement in attitude in either recreational or academic construct.

Figure 25 Boxplot of Composite Differences of Academic and Recreational Reading Attitudes (Control group)



To test for statistical differences over each of the two time points, for each construct, a paired samples t-test was run. Firstly, recreational satisfaction scores were analysed for differences before and after the intervention. Tests of normality were conducted to assess whether data would fit a normal distribution and confer with parametric testing. The Shapiro-Wilks test reported a statistical significant difference ( $p=0.00$ ) implies that these data violated the assumption of normality and that parametric testing would not be suited. However, and with guidance from Chesher and Spady (1991) we are able to make better judgements about data that are larger in kind with graphical outputs, in this case, a Q-Q plot was used to test assumptions of normality. Perusal of the Q-Q plot distribution output showed that these data illustrate an approximate asymptotic distribution as data are linear and close enough to the 45 degree line (Chesher & Spady, 1991, Davidson & Mackinnon, 1998). One reason, as discussed by the authors cited, is that larger samples may deviate from normality as they continue infinitely such as the recreational difference scores illustrated observed. A decision was made to accept normality based on these conditions thus permitting the use of paired samples t-test. A paired samples t-test found that there was no significant statistical difference of pre recreational ( $M = 12.97$ ,  $SD = 2.83$ ) and post recreational ( $M = 13.49$ ,  $SD = 2.73$ ) reading aspirations  $t(346) = 1.18$ ,  $p = .236$ . The inference made about these results suggests that the control group didn't improve their

recreational satisfaction scores over the intervention period, a non-statistically significant result. Furthermore, both mean scores indicate that the group were not satisfied reading English books in recreational conditions.

### **4.9.3 Academic Scores**

The academic construct satisfaction scores were also assessed to follow a normal distribution by approximation and based on the guidance for large samples, asymptotic distributions were clearly evident of having linearity based on the guidance of (Chesher & Spady, 1991) and (Davidson & Mackinnon, 1998). A paired samples t-test was selected to investigate if there were statistically significant differences between pre academic satisfaction scores and post academic satisfaction scores. The results found that there was a negative decrease over the two time points, where the pre academic mean score ( $M = 14.96$   $SD = 1.64$ ) decreased slightly compared to the post academic mean score ( $M = 14.61$ ,  $SD = 2.82$ ), a mean difference of  $-.348$ , 95% CI ( $-.69671$ ,  $-.00069$ ),  $t(346) = -1.971$ ,  $p = 0.06$ . The results suggest that academic satisfaction levels didn't improve over the intervention period and that the means for both conditions were extremely low. It indicates that the control group's general satisfaction and aspirations of reading English books for pleasure in an academic situation are generally quite low.

To test these assumptions further, an overall total was calculated to understand total satisfaction over the two time points. Normality tests were run and found data to be normally distributed. The total satisfaction score from pre test ( $M = 27.94$ ,  $SD = 3.26$ ) was not statistically significantly different to the total satisfaction score at post test ( $M = 27.81$ ,  $SD = 4.17$ ), with a negative decrease of  $-.129$ , 95% CI ( $-.6419$ ,  $.3825$ )  $t(364) = -.498$ ,  $p = .619$ . The results indicate that over the intervention period, total satisfaction of reading English books slightly decreased and was not statistically different to satisfaction levels before the intervention took place.

### **4.9.3 Intervention Group Pre/Post Academic and Recreational Findings**

The descriptive statistics reported in Table 54 below suggest satisfaction levels for English reading was low, indicating unsatisfied responses to pre recreational ( $M = 13.73$ ,  $SD = 13.73$ ) and pre academic ( $M = 14.26$ ,  $SD = 2.62$ ) constructs. However, the mean increases for both recreational ( $M = 39.92$ ,  $SD = 2.55$ ) and academic ( $M = 14.61$ ,  $SD = 2.80$ ) construct scores at the post test stage. Of note, the descriptive statistics show the minimum

and maximum scores for recreational reading improving but the minimum score for the post academic construct remaining the same.

**Table 54 Baseline Recreational, Academic and Total English Reading Attitude Scores**

<i>Construct</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Minimum</i>	<i>Maximum</i>
Pre Recreational	354	13.73	2.68	10	20
Pre Academic	354	14.26	2.62	10	20
Pre Totals (aca & rec)	354	28.05	3.70	20	38
Post Recreational	354	39.92	2.55	31	40
Post Academic	354	17.37	7.86	15	42
Post Totals (aca & rec)	354	65.82	7.78	44	80

A paired samples t-test was run to test for statistically significant differences between pre recreational and post recreational composite scores. The results found that there was a statistically significant difference from pre recreational composite satisfaction scores ( $M = 13.73$ ,  $SD = 2.68$ ) and post recreational composite satisfaction scores ( $M = 39.92$ ,  $SD = 2.55$ ), a positive mean increase of 23.12, 95% CI (22.74, 23.51),  $t(353) = 118.09$ ,  $p = 0.00$ ,  $d = 1.0$ . The effect size indicates that there is a very large difference between both time points. Based on the higher mean at post intervention, the intervention group’s mean has improved indicating a highly satisfied response of reading English books in recreational conditions following a peer-mediated phonics programme.

A paired samples t-test was run to test for statistically significant differences between pre academic and post academic satisfaction scores. The descriptive statistics from Table 54 suggest the mean improved over the two time points. Results from the paired t-test found there to be a statistically significant difference between pre academic attitude composite scores ( $M = 14.26$ ,  $SD = 2.62$ ) and post academic attitude composite scores ( $M = 39.92$ ,  $SD = 2.55$ ), a positive mean difference of 3.10, 95% CI (2.210, 3.993),  $t(353) = 6.843$ ,  $p = 0.00$ ,  $d = 9.9$ . The large effect size offers a statistically significant difference over the two time points when considering academic attitudes towards reading. In this regard, the inference

made about this result is to accept practical differences over the intervention programme that can be examined by improved attitudes to reading English in academic situations.

Finally, both recreational and academic constructs were totalled to give an overall satisfaction score for reading English. Table 54 displays the descriptive statistics indicating that the range of scores to be higher at post intervention compared to the baseline measures. To test whether there was a large statistically significant difference between the scores, a paired samples t-test was run. Normality results indicated that data were approximately normally distributed based on no significant outliers and the differences between the tests were judged to be normal using graphical Q-Q plots based on large sample sizes. The final results indicated that there was a statistically significant different mean increase from total pre satisfaction scores ( $M = 28.05$ ,  $SD = 3.70$ ) compared to total post satisfaction scores ( $M = 65.82$ ,  $SD = 7.78$ ), an overall increase of 37.76, 95% CI (36.868, 38.662),  $t(353) = 82.810$ ,  $p = 0.00$ ,  $d = 6.5$ . The very large effect size suggests that total attitudes to reading English improved over the intervention period when compared to attitudes levels before the intervention took place. The low pre test mean 28.05, indicated dissatisfaction in both recreational and academic environments. However, this improved significantly post intervention, with a mean of 65.82 suggesting moderate to high satisfaction levels.

#### **4.94 Chapter Conclusion**

The aim of the intervention was to understand whether children could peer teach basic reading skills using a cooperative-synthetic approach. Monitors from the host school were trained over a two-week period before teaching children from their local village. Pre-trial focus group interviews with Monitors found a mix of positive and negative attitudes towards reading. Both academic and recreational aspirations of Monitors improved as a result of the peer-teaching experiment. Children who formed the learning group were considered to be at threat of not obtaining English literacy skills. Following the intervention, the results suggest children who're exposed to peer-cooperative pedagogies gain in phonemic awareness; spelling and reading. Furthermore, reading aspirations improved over the time period. All results and their implications are discussed in Chapter five.

## Chapter Five - Discussion

### 5.1 Introduction

This Chapter presents a discussion of the findings as they relate theory and wider literature. It also includes information that can be useful for future research and policy design specific to progression in early reading skills for education in developing countries. In particular, it reviews the contribution of synthetic and cooperative methods on learning development and achievement. It also makes the case for further acknowledgment of how L2 reading attitudes in rural Indian environments ultimately shape holistic reading experiences. Together, these results are then discussed with acknowledgements to their limitations and providing information to direct future work.

The aim of the current study was consider the effects of a peer-mediated phonics programme on the phonemic awareness, spelling ability, reading skills, and reading attitudes of children from a rural North Indian community. Firstly, a summary of the research is presented with reference to specific research questions. Second, the implications of the findings for practice development and research is discussed. Thirdly, limitations are presented with the aim of offering guidance on where future trials incorporating similar methods should concentrate their efforts.

### 5.2 Commentary on Findings

Guided by a peer cooperative framework while operating within a sociocultural theoretical lens, the purpose of this mixed method random control trial study was to understand the effects of a peer mediated phonics strategy on L2 English readers from a rural Northern town in India. Ahead of the trial, it also aimed to understand more about peer monitors' reading habits, as this is severely underrepresented in the research literature that concentrates on education in developing countries.

The study enrolled (N = 745) participants acting as either learners (n = 701) or Monitors (n = 44). Monitors were exposed to a systematic phonic approach utilising basic sounds, blending and segmenting techniques through a peer cooperative method. Learners were recruited from surrounding villages and stratified into groups before being randomly allocated to control and intervention groups. The intervention group received systematic phonics teaching, mediated by Monitors over a four-month period, while the control group received their usual teaching over the same period.



The main findings of the study indicate gains in phonemic awareness, phonological processing, spelling and reading skills, when children were exposed to a peer-cooperative phonics intervention. Moreover, attitudes to reading improved over the intervention period with gains in recreational reading pursuits akin to the games and fun activities the synthetic phonics approach applies. All results are further described in this Chapter against their relevant research sub questions and consider the wider literature. The first section in this Chapter concerns the attitudes to reading based on the thoughts of the Monitors who mediated the learning processes involved in the trial. The second section discusses the findings of the intervention with implications for further study and identifies limitations and opportunities for improvement.

### **5.2.1 Pro-Social Recreational Reading Attitudes**

At the pre-trial stage, Monitors were asked to complete a reading attitude survey before (and then after) the trial with the aim of understanding more about their reading habits in academic and recreational settings. Reading attitudes are considered to be an important motivational factor in participating in reading activities. A myriad of research recognises the important correlation of reading motivation and higher reading achievement (Black, 2006, Mckenna & Kear, 1990) but less is known about L2 reading experiences in international development contexts. This presented a challenge to develop an understanding of the underlying reading habits of Monitors who were later challenged with guiding learners in an intervention. To guide this process, the following research sub questions focused on Monitors' reading attitudes and asked;

- What are the effects of Monitors' reading attitudes before and after a peer mediated phonics trial?
- What are the main factors concerning Monitors' reading attitudes in academic and recreational settings?

Using a mix method approach at the pre-trial stage, quantitative findings from the survey were then used to inform a focus group interview schedule. Here, Monitors were asked to expand on their attitudes to reading English books in recreational and academic settings. Indeed, McGuire (1989) acknowledges the importance to unify ideas of attitudes to represent social and symbolic references, where responders are able to offer open responses

to questions related to their attitudes. This approach is deeply rooted in the symbolic and cultural understanding of how language is formulated within socio-cultural contexts (Lantolf & Pavelenko, 2000). Attitudes to reading are therefore forged as a social practice representing sociocultural values and used to determine the preference of one experience compared to another.

Findings from the pre-trial survey found Monitors reported higher satisfaction scores in recreational compared to academic settings when the difference was statistically higher than the average (raw score) of indifferent attitudes. This essentially infers, at pre trial, the Monitors were more satisfied reading in their home environments, non-formally, when compared to academic, formal settings. Recreational reading reported a higher satisfaction level through obtaining a large effect size, indicating 99% of the sample as being strongly positive to reading in recreational settings compared to being indifferent or negative. This was observed as being much higher than the academic responses, which also reported an effect size but was considered practically insignificant. At this stage of the research, it was reassuring to understand children's positivity towards reading. The findings agree with what Smith (1990, p215) suggests as reading choices that are shaped by our 'state of mind' on whether reading happens or not. For Monitors in this study, the satisfaction level of reading suggests reading happens in both academic and recreational settings. Further analysis was completed to understand reading preferences within recreational and academic domains. PCA conducted on both survey constructs with the aim of looking for factorized loadings based on each survey item. Final loadings from the Monitor survey on recreational attitudes found highly correlated responses on *reading for fun*, *reading books in different locations* and *reading with peers*. These factors were then extrapolated and used for further probing at focus group interviews with selected Monitors. Each of the factors and their responses are discussed below with relevant links to the literature.

### *Reading for Fun*

Monitors gave multiple answers as to how they view reading for fun. Secondary themes emerged such as; the *types of books* each Monitor had access to and drew negative comparisons to the rigid instructional manuals academic books provide. The *types of books* theme built on the aspirations of future travel and exploration via the immediate characters readers interact with. Interestingly, *practicing new words* was also discussed as being associated with recreational attitudes. Monitors explained their focus of practicing new

words and the guesswork involved when trying to interpret new words they came across. These findings indicate a L2 'input' where new words presented to L2 readers are just ahead of the readers' ability to understand them (Krashen, 1982, Gregg, 1984). Within this context, reading becomes a social practice whereby a lack of formal instruction (authority) is replaced by a natural curiosity to locate new word experiences via recreational pursuits. This thesis recognizes learning new words within a 'fun' framework, where children are actively engaged in a narrative that is purposefully designed to arouse pleasurable reading experiences (Mckenna, et al, 1995). Indeed, this requires further scrutiny as to whether 'fun' is solely responsible for the linguistic 'guesswork' in new word retention, or, whether it is correlated with other factors. This goes beyond the findings of this research. Instead, it's acknowledged that reading for fun and the guesswork involved in reading new words are connected through discovery, exploration and a desire to read in recreational spaces.

#### *Reading Books in Different Locations*

Monitors reported their desire to read in different locations in their immediate surrounding areas. Some reported reading in their house alone whilst others mentioned their preferences for quieter places and areas setup to promote reading activities. The findings suggested *reading books in different location* contributed positively to the recreational model. Social beliefs play an important role in the decision making process of where to read. These interact with culture, home environment and wider social actors in enhancing the readers' experience. It equalizes Mckenna's (1994) reading attitude acquisition model, with particular emphasis on environmental factors that promote reading spaces and interact with personal beliefs and reading experiences. Taken as a whole, the findings indicate the importance of the social environment in promoting reading opportunities that has a positive effect on individual reading pleasures.

#### *Reading with Peers (Recreationally)*

The last theme to emerge from the PCA analysis within the recreational construct was linked to Monitors' immediate peer group. Factor loadings on *reading with peers* are well correlated from the survey analysis to suggest Monitors having various reading interests with peers in different contexts. Reciprocal storytelling and listening to others' story experiences were the main themes to emerge from the Monitor transcripts. Monitors particularly enjoyed explaining stories to their peers and informing them of what they think will happen to specific characters they were reading about. Taken altogether, recreational

reading pursuits promoted positive reading experiences for children when their immediate peer group, environment and types of reading material were accessed. These positive experiences were vital to the development of the intervention in terms of pedagogical arrangement and the promotion of pro-social reading exchanges. Monitors demonstrated their recreational reading experiences as a self-organized cooperative experience. The recreational attitudes rest on experiences outside formal situations, in this context, children's experiences with schooling. To balance this, virtually the same survey questions were asked but related reading attitudes to the academic school environment. As detailed in Chapter 1, academic environments in India present different challenges for children who're developing their reading skills. It was important to locate the reading attitudes of L2 readers in the academic context.

### **5.2.2 Academic Reading Attitudes**

As highlighted above, there was a marginal difference of Monitors' attitudes to reading English in an academic setting. Whilst these differences were statistically significant, they were considered practically insignificant based on the extremely low effect size. The majority of Monitors felt indifferent about reading in academic settings and this was explored at the focus group level. Before this, the same PCA processes were run to look for particular relationships and factorized loading on items from the academic survey. The final themes were interpreted as *reading academic English books to teachers and play*, *reading with peers* and *reading for academic purposes*. These are discussed below in further detail.

#### *Reading Academic English Books to Teachers and Play*

At focus group, Monitors described their reasons and the following sub-themes emerged as being closely related to reading to teachers in academic settings and play. Monitors reported having mixed opinions with regards to reading English for academic purposes in formal academic settings. Some indicated the lack of continuity with teachers and reading practice was a problem. Seeing multiple teachers with varying reading expectations was a major concern in this regard. Certainly, Monitors who detailed pleasurable reading experiences expressed their satisfaction when seeing the same teachers repeatedly. The continuity of reaching out to known teachers who continually show interest and support in a range of learning and remedial situations seems to improve satisfaction. The findings suggest *instrumental motivation* (Gardner, 1985, 2001) as a precursor for guiding Monitors as there

is an aspect of performing reading tasks for rewards or for teacher appreciation. It is difficult to see any intrinsic motivational attitudes in this context.

The next theme to emerge from the focus group interviews was interpreted as reading with peers.

#### *Reading with Peers (Academically)*

Similar to the recreational construct, reading with peers emerged as major theme when considering reading in academic settings. The Monitors provided a range of peer reading situations to include; peer tutoring (Topping, 2005), cross-age peer reading pairs (Johnson and Johnson, 2005) and peer-ability (Slavin, 1987) that shaped reading attitudes in the classroom. However, the unstructured and often complex descriptions cited were difficult to see if each method was properly implemented. With that said, the positive aspects grouped around paired reading situations, instruction and communication playing a major part in knowledge transfer. One sub-theme related to these exchanges indicated confusion as being associated with variable schedules. It's concluded that the disorganization of planning is a problem when designing peer-reading events. What is interesting is the peer-partnership theme emerging as a matter of importance and similarly equalizing the findings from the recreational construct. To revisit these briefly, Monitors reported being happiest when reading together, sharing stories and following different cues. These principle exchanges were underdeveloped in the academic domain and thus reduce L2 readers' attitudes as a consequence.

The final theme to emerge under the academic construct at focus group interviews was interpreted as reading for academic purposes and is discussed below.

#### *Reading for Academic Purposes*

This theme had various emerging sub-themes about reading experiences in the school environment. Overall, there was an indifferent attitude to reading attitudes as expressed through a focus on class tests and assessments, the types of reading subjects and ability to read. These findings relate to a low expectancy for success (Mori, 2002) because of the lack of attention to developing L2 reader skills via student centered teaching and learning. It relates to the lack of designing in facilitating peer learning and investigating other pedagogical approaches for enhancing practice in the classroom. Reading tasks are found to be similar in each subject, whether Science or History. The similarity of book reading

chapters are synonymous with Indian primary schools, where syllabi are created from book chapters and read out during class. As discussed in Chapter 1 and throughout this thesis, rote learning methods and one-way transmitted lessons are threats to learner achievement when used as the primary medium of instruction. The findings in this section also indicate the negative connotations ‘one teaching’ style has on reading attitudes for academic purposes.

In summary, the focus on setting (Gardner, 1985, 2001) plays a significant role in attitudes towards reading. The findings at the pre-trial found Monitors with greater satisfaction when *reading for fun*, *reading books in different locations* and *reading with peers* in recreational environments. Academic settings had an indifferent outlook largely due to limited time *reading with peers*, being exposed to rote (or singular) pedagogies when *reading for academic purposes* and inefficient planning including structure when *reading books to teachers and play*. Positive aspects in academic reading situations related to more time with teachers or peers who targeted ability (Gardner, 1985, 2001) as a way of improving attitudes and learning achievement in L2 reading acquisition.

### **5.2.3 Pre – Post Trial Attitudes**

One of the main aims of the study was to understand the effects of English reading attitudes (recreational and academic), comparing a peer-cooperative synthetic phonic method to traditional English teaching, typically conducted in Indian schools. It wanted to know if attitudes towards reading would change in each domain. Synthetic phonics instruction is identified as being fun, creative and user friendly in L2 reading international development contexts (Ekpo, 2007, Eshiet, 2014, Dixon, 2011). In this thesis, it was developed through a sociocultural framework outside of formal teaching and learning settings. Therefore, an interest in whether reading attitudes, compared to the usual ‘rote’ methods deployed in academic settings, improved over the trial were investigated. More specifically, the following research sub question asked;

- Are there differences in Monitors post recreation, academic and total reading aspirations based on gender when controlling for pre test recreational, academic and total composite scores?

Firstly, the Monitor results are discussed noting the progress made in reading attitudes as a consequence of the intervention period. Monitors played an active role in mediating and co-constructing reading practice alongside their peers. As discussed in Chapter 2, the role of social interdependence rests on positive relationships in peer groups (Buchs, et al, 2004, Johnson and Johnson, 2009), that individuals are able to cooperate on learning tasks leading to success and achievement. Children actively engaged via cooperative methods have repeatedly demonstrated a capacity to achieve where there has been limited instruction from qualified teachers. A plethora of research, using experimental designs, finds success in cooperative methods and creating knowledge partnerships (Johnson and Johnson, 2009). Engagement has led to motivation on-task and this has similarly improved attitudes from a range of diverse linguistic backgrounds (Cartledge & Kourea, 2008). The findings in this thesis are consistent with these studies and their theoretical dispositions. Findings from the posttest attitudes when controlling for pre –test as a covariate indicate total attitudes improved over the intervention, and these were statistically significant. The pooled recreation and academic scores show a mean increase of 15 percentage points across the time period. The large effect size indicates 95% of pre test scores had a score below the average subject at the post-test stage. Inferences from these results indicate the important role peer-cooperative synthetic phonics teaching has on total L2 reading attitudes. This is an important finding when considering the limited scientific work undertaken in developing countries that target reading attitudes as a feature of reading enhancement.

Further analysis revealed there to be gains in both recreational and academic constructs. The biggest gain when considering pre-test scores as a covariate was in academic reading attitudes. Pre-test scores indicated an indifferent satisfaction towards English in academic settings. However, this changed following the intervention period with a mean increase of 12 points, a statistically significant difference to pre test scores. The large effect size indicates 99% of the pre test responses to have a lower score compared to the average score at post-test. Posttest recreational attitudes also improved with a mean difference of 3 points over the intervention period. It indicates that 79% of the pre-test scores were below the average subject at post-test. The inferences drawn from both results suggest the impact of teaching and learning peer-instructions on academic attitudes. There is a convergence of attitudes to reading English in recreational and academic settings. It implies that Monitors don't favour one construct over the other. Indeed, there is a pluralistic tendency, geared towards reading English in either setting and these are mutually inclusive to their

aspirations. Further analysis investigated if gender could explain differences over the intervention period. Research highlighting gender differences in reading attitudes and motivation continues to gain interest in policy and programme design. Scholarly work completed in key areas of reading motivation (Senn, 2012), ability (Marinak and Gambrell, 2010) and engagement (Guthrie, et al, 2006) being salient themes in efforts to understand what works for reading achievement. The main findings from this experiment indicate no gender differences in English reading aspiration. This was observed on each of the recreational and academic constructs for Monitors where the mean remained constant over the two time points. The results are consistent with Guthrie's, et al (2006) assertion that 'engagement' is key to reducing differences in attitudes and incite motivation of reading tasks. When considering 'gender' as a factor of reading promotion, the results suggest attitudes to be intrinsically homogenous, that is, not one specific gender promotes reading attitude itself. Therefore, it rejects gender as an extrinsic factor in the development of reading attitudes amongst peer Monitors. This finding promotes the use of male and female peer-monitors in a co-constructive process of maintaining and developing reading aspirations amongst L2 readers. It is an important finding for programme development and consistent with Senn's (2012) ideas of using role models in reading promotion.

#### **5.2.4 Learners' Pre/Post Trial Attitudes**

The same attitude survey was given to learners' ahead of randomisation and trial. All learners' (N = 701) completed the survey before and after the trial with the aim of understanding if attitudes to reading English changed, and which domain was preferred. In guiding this enquiry, the following research sub-question asked;

- Are there differences in control and intervention groups' recreational, academic and total English reading satisfaction scores?

At baseline there was no difference between control and intervention group's attitude scores in both recreational and academic reading attitudes. This was analyzed, as one homogenous group ahead of randomisation and the slight differences in each construct had no bearing on group selection at the pre-trial phase. To help us understand the patterns of transition between pre-post tests, the results in Chapter 8 were presented as pre control, pre intervention, post control and post intervention for each construct. On both constructs, if we consider total satisfaction is equal to 40, control and intervention reported low mean



scores at baseline. It indicates that children reported being unsatisfied with English reading in either recreational or academic settings. The findings correlate with Memon (2014) who suggested family background variables, such as, access to books in the home environment and a lack of English instruction contribute to L2 reading attitudes. It found low-aspirations as a consequence of redundant measures to introduce L2 reading activities and discussions in both formal and informal environments. The same is reported in this study, family background information reveals that 94% of families don't read to children in English, for fun or academic purposes. Indeed, half of the families in this study have not completed any form of education, the majority only completing up to Standard 8. It's fair to assume that attitudes to L2 (English) reading will be low if the provision and chances to use and apply it are limited. This is further argued in the total scores that were analyzed over the trial period. It shows statistically significant differences between control and intervention groups' recreational and academic attitudes scores. The mean increase of 23 points over the trial period for the intervention group suggests the impact of the peer-mediated reading programme to have an impact on English reading attitudes in recreational settings. The large effect size indicates 84% of the pre-test scores to be lower than the average subject from the posttest scores. The majority of learners completing the intervention improved their attitudes to reading English over the time, nearing total satisfaction with a high mean score of 39 out of a possible 40. Similar findings were also reported on the academic construct. Total reading attitude improved as a consequence of the intervention. These findings are relative to the Monitor findings where there seems to be a convergence of academic and recreational aspirations towards reading English. Learners in the intervention group relate to books, peer engagement and wider reading positively as a result of the intervention.

Overall, attitudes to English reading improved over time and at a higher rate compared to traditional methods in Indian schools. The findings are consistent with the literature that finds programme engagement and resources as key drivers in reading aspirations. It also adds to a weak literature on the subject of L2 reading experiences in the developing world. More about how this can be scaled is reviewed in the implications section that follows this section. The next section discusses the effectiveness of the synthetic phonics peer-mediated trial.

### **5.2.5 Trial Impact**

The principle aim of this thesis was to understand whether a peer-mediated synthetic phonics trial had an effect on phonological processing, spelling and reading of children from a marginalized rural town in Northern India. As discussed in Chapter 1, providing quality literacy education to rural parts of India is a major challenge for policy and programme development in schools and places of education. Numerous calls from development agencies and educational planners seek to provide answers on how best to tackle illiteracy (Banerjee & Duflo, 2011). This study aimed to address such calls by acknowledging foundational reading skills and trialing a method that places the learners at the co-constructive process of learning to read. Chapter 2 detailed the achievement levels peer cooperative designs can have on learning. Moreover, synthetic phonic methods have also gained currency in developing struggling readers but the literature in development contexts is limited. Therefore, utilizing a RCT design, this study was specifically guided by the following research question;

- What are the effects of a peer-mediated phonics programme on the phonemic awareness, spelling ability, reading skills of children from a marginalised North Indian community?

### **5.2.6 Assessing Monitors' Phonological, Spelling and Reading Abilities**

As with the attitude survey, baseline tests were conducted to investigate current abilities in phonemic processing, spelling, reading and native reading. The following research-sub question guided the study by asking;

- Is there a statistical difference between male and female Monitor's baseline reading, phonological processing, spelling, reading and native reading test scores?

The aim of understanding Monitors' abilities at the baseline was to acknowledge any differences that might be seen to reduce/enhance the learning experience of learners' at the trial. Moreover, it also wanted to know the effect gender has on abilities as outlined in the research sub question above. Typically, Indian classrooms separate boys and girls on learning tasks, and there is research on Indian census data to suggest a disparity between girls and boys reading ability (White & Ruther, 2016). Boys tend to outperform girls based on a number of factors to do with girls dropping out of school, family income, availability of schools and culture (ibid, 2016).

Whilst this trial was not designed to investigate the effects of Monitors ability to peer teach L2 reading, it did consider the possible impact of skills difference and if gender differences exist. Before baseline, it wasn't understood if Monitors had previous experiences with English sounds or were previously exposed to phonic methods. Therefore, it was important to acknowledge any existence of any such skills.

Baseline measures revealed there to be high levels of English letter character and word reading between both male and female Monitors. Word reading did generate a small effect, however, this was seen as practically insignificant given the similarity of the mean. However, spelling was statistically significant with a moderate effect size explaining 73% of female Monitors to be lower than the average male Monitor at baseline. The spelling effects remained practically insignificant because the dictation test revealed no differences. The results indicated that apart from the moderate spelling effects, Monitors had similar reading abilities and these were consistent across letter character awareness, spelling and reading as a whole. Further baseline measure found low awareness of letter sounds including irregular sounding nonsense words. These findings were consistent across male and female results and suggested there to be no phonemic awareness amongst the Monitor group. More evidence from the extremely low blending scores support this. The findings from the phonological battery suggested there to be no formal phonics instruction or relative pedagogies in school. The low mean supports the findings from Chapter 1 that learning to read is by rote or through memorization. Words read on the Burt and spelt on the Schonell tests suggests the prevalence of rote/memory methods as core practice. Finally, no gender difference were reported on the native tests but the high mean score indicate a good foundational knowledge of the home languages.

### **5.2.7 Monitors as Peer Instructors**

Monitors were given two weeks synthetic phonic training ahead of the intervention period. This time was used to expose 44 sounds and basic CVC, CVCC blending techniques through cooperative design. Workbooks and basic games were utilized in the training period along with instructions of how the programme would run during the trial. Instructions were issued that enabled core features of story development, introduction of new sounds, revisions of old sounds, blending techniques, writing development and intra-peer exchanges. This was interpreted by each Monitor group and developed in accordance with their peers' learning requirements. The cooperative method followed a 'learning together' approach (Johnson and Johnson, 1991) underscored by a sociocultural theoretical

framework (Vygotsky, 1978). Monitors took on the role of a ‘more knowledgeable other’ in the experiment (ibid, 1978) based on their current exposure and learning within a low-cost English medium private school. This initial immersion with English enabled them to refine their skills through training of letter sounds and blending techniques to accelerate their peers reading achievements. Before this is discussed, Monitors’ phonological processing, spelling and reading abilities were reassessed at post-test alongside their peers (learners). This was to control for any differences in reading ability as a process of being a peer monitor during the experiment. As such, the following research sub-question asked;

- What are the effects of Monitors’ phonological processing, reading and spelling abilities when guiding learners in a peer-mediated intervention?

The question considers reading as a developmental process and informed by sociocultural influences. Reading engagement can foster reading achievement (Marinak and Gambrell, 2010) when learners are on-task with meaningful content. Monitors actively engaged in teaching and learning situations are also considered to benefit even when skills have been mastered. Peer research demonstrates the cognitive gains made in reciprocal learning partnership situations (Johnson and Johnson, 1991, Topping, 2005) in refining and maintaining core skills. Monitors also developed new skills and applied them in learning contexts. Baseline tests revealed limited phonemic awareness and thus indicators of previous reading achievement were solely based on word and sentence memorization. Progress at post-test revealed Monitors improved their phonemic and blending skills and these were statistically significant from pre-test measures. It suggests Monitors developed training experiences into real teaching and learning experiences, using new knowledge through application and refinement. It was hypothesized achievement in generating phonemic awareness amongst Monitors would improve over the intervention; this being based on training notes and observations during the trial period. However, unknown was how much could Monitors achieve in their own learning based on a relatively short intervention period. The answer is consistent with most phonics research that indicates the rapid acquisition of sounds as a precursor to reading text (Johnson and Watson, 2005). Monitors demonstrated a capacity to recognize all phonemes at posttest and utilize blending techniques in teaching and learning situations. Blending skills were predictably low at pre-test as there was virtually no phonemic awareness. This changed over the training event and intervention period with gains being observed from the non-parametric analysis. Post

tests also revealed gains in reading and spelling although these were observed as moderate and minor effects. Word identification improved explaining 66% of pre-test scores to below the average subject at posttest. Similarly, spelling reported a minor effect explaining 54% of pre-test scores to be below subjects at posttest. However, reading scores dipped when compared to pre-tests measures. It's noted here that we should take caution when inferring gains in reading, the results are positive in respect of Monitors improving their reading skills over the intervention, however, these are moderate. If we look at the dictation (reading) scores over the period of the intervention, these were slightly lower compared to baseline measures. It suggests Monitors might be selecting various methods for word recall and using linguistic guesswork (Krashen, 1982) for word. These findings relate to Bulotsky-Shearer (1986) cautionary guidance on switching between phonic instruction and non-phonetic instruction for established readers, even when there is some evidence that decoding non-phonetic words may happen long after initial phonic exposure (Johnson and Watson, 2005).

#### **5.2.8 Learning to Read with Peer Instruction**

Learning to read presents many challenges to native readers in situations that are well resourced and instructions are guided and developmental. Second language learners experience problems when faced with learning a language that doesn't follow similar rules and principles as their mother tongue. Moreover, learning to read in these contexts, without tailored pedagogies targeting L2 reading skills is virtually impossible. These learning dynamics are consistent with children in India who are on the periphery of quality reading education. Given the linguistic diversity and challenges presented in learning English, this thesis aimed to investigate the effects of slightly advanced peers on the L2 reading achievement of learners from a rural Northern Indian village. Using an experimental design, the intervention group consisted of children receiving a peer-mediated synthetic phonic instruction while control group members followed their usual traditional teaching. The results of this intervention are discussed here in relation to inferences made at the baseline and post intervention level.

Peer cooperative methods were designed and implemented into the final trial. Peer Monitors from the host English medium school were sampled as being slightly advanced in their skills of English reading. Cooperative 'learning together' (Johnson and Johnson, 1991) procedures were developed for the trial and Monitors were trained in the core competencies related to its structure and delivery. Monitors were also trained in synthetic

phonic methods and these were incorporated into cooperative practice. Before the trial all learners were pre-tested in reading, phonological processing and spelling test batteries. These being the same tests the Monitors completed. The guiding sub-research question wanted to know if the learner groups were homogenous, as any differences found at baseline could limit the inferences made post intervention. Although, minor deviations and differences are almost certain to occur, any major differences require careful observation. It asked;

- Is there a statistical difference between male and female learners' baseline reading, phonological processing and native language test scores?

The phonological battery found no differences in male and female test scores. Of wider interest was confirmation of the original hypothesis that learners would have low mean scores and these are not different when considering gender. It mirrors Monitors' scores on the same test, although learners were observed to be somewhat lower. It reaffirms other findings that rural Indian children have limited phonemic awareness and these skills are untapped in advancing their usage in L2 reading acquisition. Further evidence of low mean scores is found in the reading, spelling and native tests, all with no effects on gender but parity with low performance. All children generally have troubles with reading in native and second languages. This is consistently found across all of the test batteries and in similar studies that report reading achievement using synthetic approaches based in international development contexts (Eshiet, 2014, Dixon, 2011). The low mean scores also relate to poor performances in English and native tests conducted within rural climates and conducive to an emerging trend of stagnation in reading achievement within the middle to higher year groups (ASER, 2016). In particular, recent findings indicate alarming volumes of children from Standard VIII are still unable to read in their mother tongue (ibid, 2016). When factoring English into this equation, rural data suggest just over a quarter of Standard III children could read simple English words (ibid, 2016). The baseline findings corroborated with the National picture of reading difficulty in rural settings and implied learners rest on singular pedagogies for reading achievement.

In an effort to find a new approach to reading teaching, cooperative synthetic phonic methods were targeted based on children's prior ability in English reading. In guiding this process, the main research question asked;

- Are there statistical differences between learners' reading, phonological processing and native posttest scores when controlling for pre test scores? If so, where do these differences lie?

The intervention period included learners as co-constructors in the learning to read process, sharing their knowledge and expectations with Monitors. These skills were taught using a fun and friendly approach using sounds of English words through a systematic approach. The learning process involved children being exposed to letter sounds (phonemes) and experimenting with harder sound combinations (digraphs, trigraphs) before using both skills in sentence reading and writing activities. The systematic approach allowed for a pace that suited each learner's ability to imitate Monitors in task-orientated activities (Lantolf & Pavlenko 1995). Reading as a social constructivist concept permitted learners to acquire knowledge (Vygotsky, 1978) via cultural artifacts and symbols, such as, phonic stories, physical actions and imaginative play. The symbolic nature of this approach, alongside the promotion of positive interdependence (Johnson and Johnson, 1991), fosters memory retention of new concepts and skills. At the beginning, learners were on the periphery of learning new sounds before applying blending techniques while experimenting with these techniques to form words. This process changed when learners met Monitors within the ZPD, where Monitors acted as the more knowledgeable others (Vygotsky, 1994) based on their previous training in phonic pedagogies. Monitors permitted exchanges within the subgroup, this being particularly useful if learners had missed previous lessons allowing them to catch-up. The repetition of sounds was essential for memory tasks and consolidation of newly acquired sound skills. These exchanges are fixed in social learning, the intersubjectivity stage (Wertsch, 1985) where learners and Monitors are twinned with common goals and co-develop new learning schemas ahead of individual internalization. These dynamics are consistent with the results from the trial. The results found large effects of phonological awareness and processing when peers guided learners, compared to traditional methods. Whilst this finding may not come as a surprise from a constructivist view, that is; children exposed to sounds over a four month period are bound to make progress in sound awareness. It demonstrates capacity for children who're slightly advanced in phonic skills are capable to mediate learners who're developing them. Results suggest that children are able to cooperatively peer teach basic letter sounds and have a positive effect of phonemic awareness/processing in learner children, and compared to

traditional teaching methods. However, there is evidence explaining that complex decoding of irregular sounds is still being developed. The Miskin test results showed improvement in decoding and processing ability, however, overall mean transition scores were low. It suggests learners operating in the ZPD are in a fluid process of skill development (Vygotsky, 1998). In this experiment, learning development is a process measured through continuous dynamic assessment (Lidz, 1991). The test-intervention-retest format allows for the detection of skill modification and subsequent pedagogical re-shaping. This thesis argues phonemic/spelling/reading modification occurred as a result of peer-mediated controls. However, it recognises these skills within a developmental continuum, where irregular words and sounds are functioning but have not yet been internalised or mastered. It identifies peer-mediation as a cultural tool to facilitate mastery of new skills compared to traditional methods (Vygotsky, 1998, Johnson & Johnson, 1991). Reading and spelling batteries also resulted in large effects and follow a similar pattern of increased achievement at post-test. Spelling improvement found in this study is substantiated by Johnson & Watson's (2005) views in that children exposed to synthetic phonic drills are better able to decode unfamiliar words using their detection abilities.

Native reading and spelling tests were not affected by the intervention and remained relatively unchanged over the two time points. It has to be pointed out again that Hindi isn't taught until Standard III in some schools and this could be a contributing factor to the minor reversal of scores at posttest.

The intervention period enabled groups of learners to be supported in the development of their phonemic, spelling and reading. Each day, new letter sounds would be introduced and eventually used for the blending of words. This study supports the mnemonic and physical/oral nature of learning sounds for beginning L2 readers. It argues against singular methods or visual whole word cues used in isolation, and supports the engineering process in sound-word detection. Words that are unfamiliar may not be attempted and thus unused in vocabulary or reading development contexts (Bowey, 2007). In recognition of similar studies utilizing phonics methods, it's important to note that these findings support the development of reading programmes that enable letter characters, initial sounds, blending techniques and physical actions as a vehicle for promoting reading achievement (Adams, 2001).



### **5.2.9 The Significance of Peer Reading Instruction and Attitudes**

The findings of this thesis offer a significant contribution in furthering innovative pedagogies that aim to tackle illiteracy. In direct response to calls from development agencies/educational planners for more experimental trials in improving literacy rates in rural areas, this thesis offers some promising results. Peer-cooperative learning in conjunction with a systematic phonic approach accelerates phonemic awareness/processing, spelling and early reading in L2 learners. It also provides information related to specific factors that are useful in fostering positive reading attitudes constructed around peer exchanges in recreational and academic environments. Attitudes to reading is an underdeveloped area in Indian education policy, it is not referred to in any policy documents or curricula. The significance of peer reading promotion in academic and recreational settings is crucial to accelerating reading achievement. Currently, Indian classrooms are filled with singular teaching methods that depend too much on children memorizing words. These results add to the vast literature that supports synthetic phonic/cooperative learning strategies as an important factor in reading success.

## Chapter Six - Conclusion

### 6.1 Introduction

Peer cooperative learning and synthetic phonic strategies are dominant in the research literature for advocating successful learning and achievement. This study used both methods in the delivery of an intervention targeting developing L2 readers. Chapter one presented the difficulties involved in imparting quality reading instructions in rural education environments, with specific reference to rural Punjab. Chapter two identified the relevant literature in support of peer methods and reading strategies known to work for beginning readers. Chapter three and four followed by demonstrating the processes involved in establishing the trial, utilising methods and piloting the tests and procedures involved in running a successful RCT. Chapter five reported the findings from the mix method experiment before being discussed in Chapter six. In this final Chapter, the main findings are addressed and state overall conclusions of the study. As with all research, there are limitations and implications for future research based on the current study's findings. These are discussed with particular attention to the methodological, pedagogical and policy implications of developing this work.

#### 6.1.2 Contribution to Knowledge

Firstly, it's worth restating the study's research aim was designed to understand phonemic awareness/processing, spelling, reading achievement and attitudes amongst L2 English beginner readers from a rural India location. It treated reading achievement through a sociocultural lens and instructions were imparted via peer cooperation. The methods used in the trial incorporated a peer cooperative approach in the delivery of a synthetic phonic programme.

This study found significant effects of the peer cooperative-synthetic phonic intervention on generating phonemic awareness/processing, spelling and some gains in reading. These findings suggest children are able to co-create and mediate their reading development with minimalistic teacher instruction. It found, when compared to traditional – rote teaching methods, the intervention group significantly outperformed the control group and with large effects. Attitudes to reading also improved over the trial period where all children reported being highly satisfied. Pre-trial focus groups were run to understand reading motivation and attitudes towards reading, it found peer exchanges and teaching structure to be salient factors in promoting positive experiences of reading as a practice. The study was

undertaken in a small – marginalized community in the North of India where children have limited access to English reading books and associated learning materials. Taken altogether, one of the major contributions of this work is to the field of informal learning and its function outside of mainstream education. In theoretical terms, this thesis contributes to the sociocultural movement that recognizes learning to happen within social ‘peer’ exchanges, through concomitant symbolic and cultural partnerships. More specifically, the findings make a contribution to the peer learning and L2 English reading attitudes literature. It expands upon research into peer cooperative learning effects with further theoretical insights into its use in international development education contexts. In sum, this is an arena with limited scholarly attention but with scope to enhance further enquiries. This thesis also contributes to the limited literature on L2 English reading attitudes in development contexts. In this field, it provides a dynamic contribution to understanding reading attitudes from children operating in rural, marginalized and unsupported contexts. Indeed, the findings are unique in that scholarly work undertaken in the reading attitude domain is comprehensively redundant in India. Therefore, this study presents a gateway to understand more about L2 English reading attitudes and its interaction with reading access and overall achievement.

### **6.1.3 Future Research and Implications**

Future research will build on a number of the findings and contributions as outlined above. This section reviews the more immediate opportunities for research and scholarly activity. Firstly, the trial did not build-in or test for reading comprehension mainly as a response to understand more about the potential in combining peer/synthetic methods. As of writing, no research combining these approaches in international development education exists. Only recently has there been progress employing experimental/action research designs using synthetic phonic approaches in development contexts (Dixon, 2011, Ekpo, 2007, Eshiet, 2014). This presented an opportunity to further research in both domains using an evidence-based approach in programme evaluation. Therefore, this thesis aimed to explore these areas but was limited in what it could achieve. Future research should incorporate comprehension models using a longitudinal format to allow more time for learners’ to develop their comprehension skills. As discussed in Chapter two, there is an argument that synthetic phonic methods appeal to developing pre-reading activities and provides a blueprint for later comprehension (Adams, 1990, Johnson and Watson, 2005). But it cannot be assumed synthetic methods alone can improve reading achievement. It should be structured and operate within a broader curriculum while being promoted through positive

exchanges – as found in this thesis. Future studies looking to understand more about the implications of phonic strategies and reading comprehension would benefit from longitudinal trials similar to experiments run by Johnson and Watson's (2005). As detailed in Chapter five, whilst there were minor gains in reading achievement, it wasn't understood if Learners or Monitors made a semantic connection with words/sentences. Future research is required to determine if any effects occur. Similar calls to improve L2 reading comprehension effects via synthetic deliveries are made by Eshiet (2014), where the findings in this thesis match those found in Nigeria.

Under the banner of longitudinal designs, the implications of learning to read were understood within a dynamic assessment (Vygotsky, 1994) of individual ability, thus, providing a snapshot of what all children could/could not achieve during the short intervention period. Reading is a lifelong process and operates at the intersection of sociocultural and cognitive exchanges. The findings in this thesis don't recognise learners as being fully accomplished readers, nor does it recognize their mastery in decoding, segmenting and/or working with irregular words. Instead, the short intervention provided evidence that children, who previously found learning to read a challenge following rote methods, modified their learning through a peer-cooperative synthetic approach. In order to make accurate inferences on the proficiency and mastery of reading achievement, a longer intervention should be designed.

This study employed a pre/post design with limited focus on the transitions made by learners. Some results (native reading/word tests) had negative effects at post –test and these couldn't be fully explained. Future studies may incorporate multiple testing times to enable the ability to track specific trends over time. Furthermore, if future studies developed longer trials, this would create an opportunity to test L2 reading effects using different peer learning designs (i.e. peer tutoring, collaboration). This study reports effects of a cooperative strategy for instruction and is limited by this approach. Future studies should incorporate a complex multi-method RCT that tests for overall effectiveness. In other words, although this research has found positive effects in cooperative methods, it might mean other methods (i.e. tutoring) might have a greater impact.

The arrangement of the trial, including criteria as set by the CONSORT statement (Altman, et al, 2001) incorporating rigorous procedures undertaken in this study's initiation, development, analytic and subsequent interpretation. This study made every use of the

experimental conditions that were uniquely available in using a RCT format (Torgerson, et al, 2013) in an informal education setting. As documented in Chapter three, applying the strict parameters of random assignment often presents challenges to educational researchers, where quasi-experimental methods are often preferred. This study calculated a large sample size in order to detect effects (with strong power) given the lack of clear guidance on peer/synthetic effects found in the literature. Final guidance on the strength of effects was provided by Cohen (1988). With all of these careful measures, some tests violated assumptions and non-parametric tests were used. The choice to use non-parametric tests reduces overall power, as it doesn't require distributions to be normal. However, non-parametric tests require independent random sampling where the median is used instead of the mean. Therefore, future research should consider any confounding factors when reporting effects. This is something that isn't always transparent in some papers; thus, a call for better reporting in social science research should be acknowledged.

Internal validity was improved by stratifying the sample population to give equal weighting into control and intervention groups by restricting known characteristics (i.e. gender, location, age) to have an advantage in one group over the other. Every effort was made in separating control/intervention groups as per the experimental conditions when the intervention was running. However, these were only applied to the duration of the intervention period – an hour per day. It didn't apply to any condition before or after each intervention/control-teaching day. This left potential for children to meet and discuss methods externally. Any such engagements were identified as being beyond the control of the experiment. Based on the results, this appeared to be minimal or certainly not regular.

Another threat to validity included the use of props and learning materials. This study used colorful books and accessibility to writing instruments to aid learning tasks. It could leave room for a novelty effect based on the input of apparatus to arouse increased levels of interest. A qualitative component may have enhanced the interpretation of any effect as a direct result of using learning props/materials. Future research should target overall usage, or, time-on-task with materials to understand interaction elements. Moreover, along with the other factors mentioned above, RCT designs should consider the practical elements to mitigate against threats to validity wherever possible.

The qualitative research strand embedded a pre-trial focus group interview framework to understand specific factors in Monitor reading attitudes. Analysis by way of factorized

loadings from the pre-attitude questionnaire guided the types of questions asked. This type of reduction method strengthened the naturalistic positive/negative responses from Monitors by recognizing the complexities involved in understanding human responses (Lincoln & Guba, 1985). Whilst these steps prompted a deeper understanding of reading attitudes related to each setting, some data were deemed unintelligible or missed due to complex responses. The organization of groups was structured by analyzing clustered data from a quantitative survey. Monitors were organized in this way as a matter of convenience sampling and ease of access. However, highly structured groups limit potential social interactions (Lewis, 1992) from other actors involved in the study, further limiting any exchanges that added to the emerging themes. Future research should consider focus group interviews that rest on reading exchanges in peer groups' themselves. A pre-questionnaire could identify peer relationships and match each member before grouping them for interview.

Finally, during the initial registration, there was some confusion of children's ages. This is a typical problem found in development environments, as some records are either misplaced or not kept. Therefore, a decision was made not to use chronological ages when analyzing results. The majority of children were registered to a school provider (primary/middle equivalent), thus, even if children were older/younger than documented, their schooling environment was mapped to their current ability. Future research will need to consider reading ages if there is enough scope for standardization processes.

#### **6.1.4 Implications for Pedagogy**

Synthetic phonics and cooperative methods continue to gain currency in developing achievement in reading and other learning tasks. This thesis advocates the usage of such methods in the classroom. It demonstrates the scope for using children as peer mentors in traditional learning situations when reading pedagogy is fun and engaging. Focus group interviews revealed positive attitudes to learning to read when there is structure and scope for peer exchanges. This must be the sole focus when developing reading programmes and policies should reflect the playful nature when designing curricula. Peer cooperative methods require flexibility to permit groups of learners' autonomy in cognitive tasks. Positive interdependence is realized when peer groups homogeneously work on solving problems and individually benefit from whole group output. Teaching targeting reading and associated skills must be designed to stimulate learners' enthusiasm for developing new knowledge and be demonstrable through actions or active discussion.

The ease and practical use of employing a synthetic phonic package makes for an encouraging opportunity for classroom-based practice. Materials, including flashcards and storybooks are relatively inexpensive and accessible. Teachers and school leaders also have the option to create their own materials and incorporate their own activities following the guidance set out in Chapter three. Policy/practice makers should consider investing in time to allow children to experience holding flashcards and physically blending words. This was an assumption that was taken for granted at the pilot stage. Some children had difficulty in using the materials because of their weight or shape; other children presented materials upside down or regularly dropped props when trying to explain key concepts. Future practice using children as mentors should encourage these mistakes in a fun, collaborative and non-judgmental way.

#### **6.1.5 Implications for Policy**

The scope of these findings has wider implications for policy development at the National and International level. In addressing International agreements on advancing quality education and lifelong learning (Goal 4 of the SDGs), this thesis calls upon stakeholders to adopt peer/synthetic approaches as a quality benchmark in improving children's reading chances. Peer mentorship and synthetic phonics offers a vehicle for accelerating opportunities for children's reading development. At the National level, there are major gaps within Indian curriculum bodies when considering pedagogical guidance. Of the two largest bodies in the Punjab - CBSE and PSEB, both offer a limited focus on speaking, listening and reading. Information is limited to broad statements teachers interpret and then develop into practice. In response to the limited information, this thesis offers practical guidance that can be easily transmitted into the classroom. The methods will be particularly useful for English-medium schools at entry points lower-upper KG or Primary Standard One. For most schools this will depend on when English is introduced as either a topic or medium of instruction. The early years focus of developing play-orientated activities marries particularly well with phonic/peer exchanges based on the physical actions, stories and games it employs. Children learning outdoors using recreational spaces also add to the variety of developing pre-reading activities. Findings from this study indicate children to regard structure and peer exchanges as salient factors in their positive attitudes towards reading. Future practice must consider allowing these freedoms to co-exist alongside traditional rote mechanisms. In this regard, children acting as co-constructors of knowledge

and working in partnership on structured tasks is seen as vitally important in scaffolding positive attitudes in pursuit of reading achievement.

#### **6.1.6 Concluding Thoughts**

Evidence-based research is key to unlocking the potential in the discovery of unique programmes that can have positive effects on learning and achievement. The benefits of using a cooperative approach in the teaching of sounds to beginner L2 readers presents an interesting journey in programme refinement and re-implementation. Future work will aim to address some of the shortfalls in this thesis (notably building in comprehension models), but the theoretical and practical aspects remain constant. This is because peer-synthetic methods have an important role to play in creating interdependent literate learners, whilst generating positive attitudes and accelerating reading achievement.



## References

- Adams, D.N. & Hamm, M.E. (1990) *Cooperative learning - Critical thinking and collaboration across the curriculum*, Springfield, IL: Charles C Thomas
- Adams, M. J. (1990). *Beginning to read. Thinking and learning about print*. Cambridge, MA: MIT Press.
- Adams, M. J. (2001). Alphabetic anxiety and explicit systematic phonics instruction: a cognitive science perspective. *Handbook of early literacy research*. S. B. Neuman and D. K. Dickinson. New York. The Guilford Press: 66-80.
- Alexander, J. E., & Filler, R. C. (1976). *Attitudes and reading*. Newark, DE: International Reading Association.
- Altman, D.G., Schulz, K.F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Gotzsche, P.C. and Lang, T. (2001) The revised CONSORT statement for reporting randomized trials: Explanation and elaboration, *Annals of Internal Medicine*, 134(8), 663-694
- Ames, T (1980) *Nfer Nelson – Diagnostic Reading Programme*, Berkshire, Britain.
- Anderson, G. (1990). *Fundamentals of educational research*. London: The Falmer Press.
- Aronson, E. (2008). *The Social Animal* (10th ed.). New York: Worth/Freeman.
- ASER (2013) *Annual Status of Education Report (ASER)*. New Delhi, India: Pratham
- ASER (2016) *Annual Status of Education Report (ASER)*, New Delhi, India: Pratham
- ASER (2015) *National Read India Report 2014–15. Measurement, Monitoring & Evaluation Unit*, ASER, New Delhi, India, Pratham.
- Ashby, J. and K. Rayner (2006). Literacy development: insights from research on skilled reading. *Handbook of Early Literacy Research*. D. K. Dickinson and S. B. Neuman. New York. The Guilford Press: 52-63.
- Azam, M, Kingdon, G. (2011). Are girls the fairer sex in India? Revisiting intra-household allocation of education expenditure. Discussion Paper 5706
- Baines, E. Blatchford, P. Kutnick, P (2003) Changes in grouping practices over primary and secondary school *International Journal of Educational Research*, 39 (2003), 9–34
- Baker, L., & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly*, 34(4), 452-476.
- Bandura A. (1977) Self-efficacy: Toward a unifying theory of behavioural change, *Psycho/. Rev.* 84, 191-215

- Banerjee, A, V., and Duflo, E. 2011. *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: Public Affairs.
- Banerjee, A, V R, Banerji, M, Chavan, Esther Duflo, and M Walton. (2007). *Impact Evaluation of Pratham's "Learning to Read" and "Reading to Learn" Interventions*. J-PAL South Asia at IFMR and Pratham Resource Center
- Banerji, R, Chavan, M & Rane, U (2005) 'Learning to Read' *Changing English*. Vol.12, No.2, 177-190
- Banerji, R., Bobde, S, (2013) Evolution of the Aser English tool. In, Berry, V. *Investigating English Language Learning Outcomes as the Primary School Level in Rural India*, London: British Council
- Baron, R., Moore, D., & Sanders, G. (1978). Distraction as a source of drive in social facilitation research. *Journal of Personality and Social Psychology*, 36, 816–824
- Barr, R. (2001). Research on the teaching of reading. *Handbook of Research on Teaching*. Richardson. V. (ed.) Washington, American Educational Research Association: 390-415
- Barro, R. J. (2001) Human Capital and Economic Growth *American Economic Review*, 91 (2) 2001, 12-17
- BERA (2011) Ethical Guidelines for Educational Research. Available from <http://www.bera.ac.uk/wp-content/uploads/2014/02/BERA-EthicalGuidelines-2011.pdf>. Accessed 20th March 2012
- Beard, R. (1990). *Developing Reading 3-13*. London, Hodder and Stoughton
- Beers, P. J. (2005). *Negotiating common ground: Tools for multidisciplinary teams*. Unpublished doctoral dissertation, Open University of The Netherlands, Heerlen, The Netherlands.
- Berg, B. (2004) *Qualitative Research Methods for the Social Sciences*. (5th ed.) Pearson, Boston.
- Bergeron, B. (1990). What does the term whole language mean? Constructing a definition from the literature. *Journal of Reading Behavior*, 12, 300-330.
- Berndt, T. J. (2002). Friendship quality and social development. *Current Directions in Psychological Science*, 11, 7–10.
- Berndt, T. J., & Keefe, K. (1995). Friends influence on adolescents adjustment to school *Child Development*, 66 (13), 121-329
- Bernhardt, E., & Kamill, M. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics*, 16, 15-34.
- Bertucci, A., Conte, S., Johnson, D. W., & Johnson, R. T. (2010). The impact of size of cooperative group on achievement, social support, and self-esteem. *The Journal of General Psychology*, 137(3), 256-271.

- Bhattacharya, U (2016) 'Colonization and English ideologies in India: a language policy perspective', *Language Policy*, 1-21
- Bishop, D. V. M. & Leonard, L. B. (Eds.). (2000). *Speech and language impairments in children: Causes, characteristics, intervention and outcome*. London: *Psychology Press*, 27-32
- Bissex, G L. (1980). *Gnys at Wrk*. Cambridge, MA: Harvard University Press
- Black, R. W. (2006). Language, culture, and identity in online fanfiction. *E-Learning*, 3(2), 170–184
- Blaxter, L., Hughes, C. and Tight, M. (1996). *How to research*, Open University Press, Buckingham.
- Blumenfeld, P. C. (1992). Classroom learning and motivation: Clarifying and expanding goal theory. *Journal of Educational Psychology*, 84, 272–281.
- Bonell, C., Oakley, A., Hargreaves, J., Strange, V., & Rees, R. (2006). Assessment of generalisability in trials of health interventions: suggested framework and systematic review. *BMJ*, 333, 346–349.
- Bookbinder, G.E. (1976). *Salford sentence reading test: Manual for the Salford sentence reading test*. London, UK: Hodder and Stoughton.
- Boruch, R., De Moya, D. & Snyder, B. (2002). The Importance of Randomized Field Trials in Education and Related Areas. In: Mosteller, F. & Boruch, R. (eds.) *Evidence Matters: Randomized Trials in Education Research*. Washington: Brookings Institution Press.
- Bowey, J. A. (2007). "Need for systematic synthetic phonics teaching within the early reading curriculum." *Australian Psychologist* 41(4): 79-84.
- Bowman, M. and Treiman (2005). "Stepping stones to reading." *Theory into Practice: Developmental Psychology* 43(4): 295-303.
- Bramlett, R. (1994). Implementing cooperative learning: A field study evaluating issues for school based consultants. *Journal of School Psychology*, 32, 67–84.
- Bridge, C., & Hiebert, E. (1985). A comparison of classroom writing practices, teachers' perceptions of their instruction, and textbook recommendations on writing practices. *Elementary School Journal*, 86: 155-172.
- Brooks, M., & Brooks, J. (2005). Whole language or phonics: Improving instructional language through general semantics. *ETC. A Review of General Semantics*, 29, 271-280.
- Brooks, ME, Dalal, DK, Nolan, KP, (2014) Are common language effect sizes easier to understand than traditional effect sizes?, *Journal of Applied Psychology*, 99(2): 332-40
- Brown, B. B. (1990). Peer groups and peer cultures. In S. S. Feldman & G. R. Elliott (Eds.), *At the threshold: The developing adolescent*, Cambridge MA
- Bruce, P (1986) A peer tutoring project with a class of 9 and 10 year olds. *Paired Reading Bulletin*, 2,

- Bruffee, K. A. (1984). *Collaborative learning and the “conversation of mankind”*. *College English*, 46, 635–652
- Bryman, A (2008). *Social Research Methods Oxford*, UK, Oxford University Press
- Bryman, A (2015) *Social Research Methods*, 5<sup>th</sup> Edition, Oxford University Press, London
- Buchs, C., Butera, F., & Mugny, G. (2004). Resource interdependence, student interactions and performance in cooperative learning. *Educational Psychology*, 24, 291–314
- Bull, S., & McCalla, G. (2002). Modelling cognitive style in a peer help network. *Instructional Science*, 30, 497–528.
- Bulotsky-Shearer, R. J., Manz, P. H., Mendez, J. L., McWayne, C. M., Sekino, Y. and Byrne, D. (1986) *Teaching Oral English*, Longman
- Calleia, M, A. Howard, J, S. (2014). Assessing what students know: Effects of assessment type on spelling. *Journal of Student Engagement: Education matters*, 4 (1),14–24
- Cameron, R & Miller, P (2007). Mixed methods research: phoenix of the paradigm wars, in Proceedings of the 21st ANZAM Conference, Australian & New Zealand Academy of Management, Sydney, December 2007.
- Carr, T., & Levy, B. (1990). *Reading and its development*. San Diego, CA: Academic Press.
- Cartledge, G., & Kourea, L. (2008). Culturally responsive classrooms for culturally diverse students with and at risk for disabilities. *Exceptional Children*, 74(3), 351-371.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276.
- Census of India (2011) *Punjabi: District Census Handbook Sangrur, Directorate of Census Operations*, New Delhi, India
- Chakraborty, T. & Kapur, S. (2008). *English language premium: Evidence from a policy experiment in India*. St. Louis: Washington University in St. Louis.
- Chall, J. S. (1967). *Learning to read: The great debate*. New York: McGraw-Hill.
- Chall, J. S. (1983). *Stages of Reading Development*. New York: McGraw Hill, 10-24
- Chall, J. S. (1996). *Stages of reading development*. Fort Worth, Harcourt Brace College Publishers.
- Chavan, M. (2003). *The Pilot Phase of Pratham’s Accelerated Reading Program Jan 2003- Jun 2003*. India. Read India-Pratham Resource Centre
- Chesher, A., and R. Spady (1991). “Asymptotic Expansions of the Information Matrix Test Statistic,” *Econometrica*, Vol. 59, No. 3., 787–815

- Chiappe, P and Siegel, L.S (1999) Phonological awareness and reading acquisition in English and Punjabi Speaking Canadian Children, *Journal of Educational Psychology*, 91(1), 20-28
- Cicchetti DV. (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*. ;6(4):284–290.
- Clark, M. (1980). The short circuit hypothesis of ESL reading—or when language competence interferes with reading performance. *The Modern Language Journal*, 64, 203-209
- Cohen J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*.; 20(1):37–46.
- Cohen, E., & Lotan, R., (Eds.). (1997). *Working for equity in heterogeneous classrooms: Sociological theory in practice*. New York: Teachers College Press.
- Cohen, E.G. (1994) Restructuring the classroom: Conditions for productive small groups, *Review of Educational Research* 64: 1-35.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates
- Cohen, J. (1990), “Things I have learned (so far),” *American Psychologist*, 45(12): 1304–1312
- Cohen, J., (1994). The earth is round ( $p < 0.05$ ). *American Psychologist* 49, 997–1003
- Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A metaanalysis of findings. *American Educational Research Journal*, 19, 237–248.
- Coleman, J. S. (1961). *The adolescent society: The social life of teenagers and its impact on education* New York: Free Press
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation. Design and analysis issues for field settings*. Chicago: Rand McNally.
- Coppock, V. (2011) Children as Peer Researchers: Reflections on a Journey of Mutual Discovery *Children and Society* Vol, 25, 6 435-466
- Counihan, C. (2015) ‘Endogenous Education in India and the Implications of Universal Peer Teaching in the 19th Century’. In *Handbook of International Development and Education*, 1st edn. Ed. Dixon, P., Humble, S, Counihan, C. MA: Edward Elgar, 291-304
- Cowie, H. Hutson, N, Oztug, O. Myers, C. (2008) The impact of peer support schemes on pupils’ perceptions of bullying, aggression and safety at school *Emotional and Behavioural Difficulties*, Vol 13, 1, 63-71
- Crandall, V. C., Crandall, V. J., & Katkovsky, W. (1965). A children’s social desirability questionnaire. *Journal of Consulting Psychology*, 29, 27–36.
- Creswell, I.W. (1994) *Research Design: Qualitative and Quantitative Approaches*, London: Sage Publications.

- Creswell, J. W. (2002). *Educational research: Planning conducting, and evaluation quantitative and qualitative research*. Upper Saddle River: Merrill Prentice Hall.
- Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. (3rd, Ed.) London: Sage Publications Ltd.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research*, Thousand Oaks, CA: Sage.
- Cronbag, H. (1966) Cooperation and competition in means-interdependent triads: A replication *Journal of Personality and Social Psychology*, 4, 692-695
- Cummins J. (1976). *The influence of bilingualism on cognitive growth: a synthesis of research findings and explanatory hypotheses*. Work. Pap. Biling. 9, 1-43
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49, 222-251.
- Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* Los Angeles: California State University, Evaluation, Dissemination and Assessment Center.
- Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language Processing in Bilingual Children*. Cambridge, England: Cambridge University Press.
- Damon, W., & Phelps, E. (1989). Critical distinctions among three approaches to peer education. *International Journal of Educational Research*, 13, 9-19
- Davidson, M, Hobbs, J (2013) Delivering reading intervention to the poorest children: The case of Liberia and EGRA-Plus, a primary grade reading assessment and intervention, *Journal of Educational Development*, (33), 3, 283 -293
- Davidson, R. and MacKinnon, J. G. (1998), *Graphical Methods for Investigating the Size and Power of Hypothesis Tests*. The Manchester School, 66: 1-26.
- Day, R. R., & Bamford, J. (1998). *Extensive Reading in the Second Language Classroom*. Cambridge: Cambridge University Press
- De, A., Khera, R., Samson, M., and Kumar, S. (2011). *PROBE revisited: A report on elementary education in India*. New Delhi: Oxford University Press
- Dean, B. (1995). *Implementing cooperative learning in classrooms in Pakistan: Problems and possibilities* (Unpublished master's dissertation). Aga Khan University, Karachi, Pakistan.

- Dearden, J (1998) 'Cross Age Peer Tutoring' *Educational Psychology in Practice* Vol 13, 4, 250-257
- Denscombe, M. (2007). *The Good Research Guide for Small-Scale Social Research Projects*. (3rd ed.). New York: McGraw-Hill.
- Department of Education (2015). Phonics Screening Check: Evaluation Report, *Department for Education*, Available here:  
[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/434821/R418A\\_Phonics\\_screening\\_check\\_evaluation.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/434821/R418A_Phonics_screening_check_evaluation.pdf) (Accessed July 2015)
- Deutsch, M. (1949). A theory of cooperation and competition. *Human Relations*, 2, 129–152.
- Deutsch, M. (1962). Cooperation and trust: Some theoretical notes. In M. R. Jones (Ed.), *Nebraska symposium on motivation*, 275-319. Lincoln, NE: University of Nebraska Press
- Deutsch, M, & Krauss, R. M. (1965). *Theories in Social Psychology*. New York: Basic Books
- DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- DeVries D L, Mescon IT, Shackman SL (1975). *Teams-Games-Tournament in the Elementary Classroom: A Replication*, Baltimore: Johns Hopkins University, Centre for Social Organization of Schools
- Department for Education (2015) Phonics Screening Check: 2015 Materials, London, DfE
- Diaz-Rico, L, T. & Weed, K. Z. (2010). *Crosscultural, Language, and Academic Development Handbook: The Complete Reference Guide* (4th ed.). Boston, MA: Allyn & Bacon
- Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. *Three worlds of CSCL. Can we support CSCL?*, 61-91.
- Dixon, J. (1996). *Whole Language: An Integrated approach to Reading and writing. A Guide for Literacy Practitioners*. Amherst, MA : The Centre for International Education, Massachusetts
- Dixon, P (2013) *International Aid and Private Schools for the Poor: Smiles, miracles and markets*. Edwar Elgar, UK.
- Dixon, P., Schagen, I. and Seedhouse, P. (2011). The impact of an intervention on children's reading and spelling ability in low-income schools in India. *School Effectiveness and School Improvement* 2011, 22(4), 461-482.
- Dombey, H. (2006). Phonics and English orthography. In M. Lewis & S. Ellis (Eds.), *Phonics: Practice, research and policy*. London, England: Paul Chapman Publishing.
- Donato, R (1992). Discourse perspectives on formal instruction. *Language Awareness*, (1), 73-90
- Donato, R. and McCormick, D. (1994) A sociocultural perspective on language learning strategies: The role of mediation. *The Modern Language Journal* 78(iv), 453-464

- Donato, R., 2000. Sociocultural contributions to understanding the foreign and second language classroom. In: Lantolf, J. P., ed. *Sociocultural theory and second language learning*. Oxford University Press, 27-50
- Dornyei, Z. (2009). *The psychology of second language acquisition*. Oxford, Oxford University Press.
- Dreze, J. and Sen, A. (2002), *India: Development and Participation*, Oxford and Delhi, Clarendon Press.
- Drozdek, A (2007), *Greek Philosophers as Theologians: The Divine Arche*, Ashgate publishing
- Dryfoos, J. (1990). *Adolescents at risk: Prevalence and prevention*. New York: Oxford University Press
- Durukan, E. (2011). Effect of Cooperative Integrated Reading and Composition (CIRC) technique on reading – writing skills. *Educational Research and Reviews*, 6(1): 102-109
- Durgunoglu, A., W. Nagy, and B. J. Hancin-Bhatt. (1993). Cross-Language Transfer of Phonological Awareness. *Journal of Educational Psychology*, Vol. 85, 453–465.
- Ehri, L. C. (1995). Phases of development in learning to read words. *Journal of Research in Reading*, 18(2), 116-125
- Ehri, L. C. (1998). "Research on learning to read and spell: a personal-historical perspective". *Scientific Studies of Reading*, 2(2): 97-114.
- Ehri, L. C. (2005). Development of sight word reading: Phases and findings. In M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook*, Blackwell: Oxford.
- Ehri, L., Nunes, S., Willows, D., Schuster, B., Yaghouh Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's metaanalysis. *Reading Research Quarterly*, 36, 250–287
- Ehri, L.C. (1999). Phases of development in learning to read words. In J. Oakhill & R. Beard (Eds.), *Reading development and the teaching of reading: A psychological perspective*. Oxford: Blackwell Science
- Erhi, L. C. and T. Roberts (2006). The roots of learning to read and write: acquisition of letters and phonemic awareness. *Handbook of early literacy research*. D. K. Dickinson and S. B. Neuman. New York, Guilford Press 2: 113-134.
- Ekpo, C. M., et al. (2007). Jolly Phonics Strategy and the ESL Pupils' Reading Development: a Preliminary Study. 1st Mid Term Conference University of Ibadan
- Elbaum B., Vaughn, S., Hughes, M., & Moody, S. W. (1999). Grouping practices and reading outcomes for students with disabilities. *Exceptional Children*, 65, 399–415
- Elledge, L. C. Cavell, T.A. Ogle, N.T. Newgent, R. A. (2010) School Based Mentoring as Selective Prevention for Bullied Children: A Preliminary Test *The Journal of Primary Prevention* Vol. 31, 3, 171-187



- Erton, I., Topcaya, Z., & Karakas, M. (2010). Exploring motivational constructs in foreign language reading. *Journal of Education*, 39, 185-196.
- Eshiet, O. (2014) *Synthetic phonics as a tool for improving the skills of Nigerian pupils*, PhD Thesis, Newcastle University
- Eva, C.E.W. (2003). *The application of cooperative learning in a remedial classroom in Hong Kong – A case study*. Unpublished dissertation, Hong Kong Baptist University, Hong Kong.
- Everitt B S, Dunn G. 2001 *Applied Multivariate Data Analysis*. 2<sup>nd</sup> Edition, London, Arnold
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149-1160.
- Field, A (2013) *Discovering Statistics Using IBM SPSS Statistics*, 4<sup>th</sup> Edition, London, SAGE
- Fitz-Gibbon, C. T. (1988). Peer tutoring as a teaching strategy. *Educational Management, Administration and Leadership*, (16), 3, 217-29
- Fleiss JL (1975) Measuring agreement between judges on the presence or absence of a trait. *Biometrics* 31: 651-659.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, J. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia, Neuropsychological and Cognitive Studies of Phonological Reading*. London: Erlbaum, 301-330
- Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34, 174–206.
- Fuchs, L. S., Fuchs, D., Hosp, M, K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 5(3), 239-256.
- Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold Publishers.
- Gardner, R. C. (2000). Correlation, causation, motivation and second language acquisition. *Canadian Psychology*, 41, 1-24.
- Gardner, R. C. (2001). Integrative motivation and second language acquisition. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* Honolulu: University of Hawaii, Second Language Teaching & Curriculum Center, 1-19
- Gardner, R. C. & Lambert, W. E. (1959). Motivational variables in second language acquisition. *Canadian Journal of Psychology*, 13, 266-272.
- Gass, S. M., & Selinker, L. (2008) *Second Language Acquisition*, Taylor & Francis, Routledge, Oxon, 30-34.

- Ghaith, M. G. (2003). The relationship between cooperative learning, perception of social support, and academic achievement. *System*, 30, 263-273.
- Ghate, D. (2001). Community-Based Evaluations in the UK: Scientific Concerns and Practical Constraints. *Children and Society* 15 (1): 23–32.
- Gholamin, M & Geva, E. (1998) Orthographic and Cognitive Factors in the Concurrent Development of Basic Reading Skills in English and Persian, *Language Learning*, 49(2), 183-217
- Ghorbani, M, R. (2012) Cooperative learning boasts EFL students grammar achievement, *Theory and Practice in Language Studies*, Vol. 2, No. 7, 1465-1471
- Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction*, 20(4), 304e315
- Gillies, R. (2006) Teachers' and students' verbal behaviours during cooperative and small-group learning *British Journal of Educational Psychology*, 76, 271–287
- Gilmore, A., Croft, C., and Reid, N. (1981). *Burt Word Reading Test: New Zealand Revision*. Wellington: New Zealand Council for Educational Research.
- Google Maps (2016b) Map of Central Punjab (online). Google. Available from: <https://www.google.co.uk/maps/place/Bhawanigarh,+Punjab+148026,+India/@30.3039537,76.1278209,48978m/data=!3m1!1e3!4m5!3m4!1s0x391049088e35c4f9:0x81a7bac10f6b1ddf!8m2!3d30.2640601!4d76.0370947> (Accessed 2<sup>nd</sup> October 2017)
- Google Maps (2017b) Map of Bhawanigarh (online). Google. Available from: <https://www.google.co.uk/maps/place/Bhawanigarh,+Punjab+148026,+India/@30.2715306,76.0403755,766m/data=!3m1!1e3!4m5!3m4!1s0x391049088e35c4f9:0x81a7bac10f6b1ddf!8m2!3d30.2640601!4d76.0370947> (Accessed 2<sup>nd</sup> October 2017)
- Goldacre, B. (2013) *Building Evidence into Education*, Department of Education, London, UK
- Gollnick, D. M.& Chinn, P. C. (2009). *Multicultural education in a pluralistic society* (8<sup>th</sup> ed.). Upper Saddle River, NJ: Merrill/Pearson Education
- Goodman, K. A. (1986). *What's whole in whole language?* Richmond Hill, Ontario, Canada: Scholastic. (Distributed in the United States by Heinemann).
- Gosser, D, Mark S, Cracolice, J.A. Kampmeier, Strozak, V and Varma-Nelson, P (2001) *Peer-led Team Learning: A Guidebook*, Prentice-Hall, Upper Saddle River, New Jersey
- Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology*, 42(1): 73-83.
- Goswami, U. (1999). Causal connections in beginning reading: The importance of rhyme. *Journal of Research in Reading*, 22, 217–240

- Goswami, U. (2002) 'Rimes, phonemes and learning to read: Interpreting recent research' in M. Cook (ed). *Perspectives on the Teaching and Learning of Phonics*. Roystone: UK Literacy Association
- Goswami, U. (2005). Synthetic phonics and learning to read: a cross-language perspective. *Educational Psychology in Practice*, 21, 273e282
- Govinda, R and Bandyopadhyay, Madhumita (2008) *Access to Elementary Education in India. Project Report. Consortium for Research on Educational Access, Transitions and Equity (CREATE)*, Falmer, UK.
- Government of India. (2011). Socio economic & caste census 2011 in rural India. Ministry of Rural Development,. Available from: <http://rural.nic.in/sites/BPL-census-2011.asp> Accessed on November, 2016
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York, NY: Cambridge University Press.
- Graddol, D. 2010. *English Next India: The future of English in India*. New Delhi: British Council.
- Graddol, D. (2006). English next: Why global English may mean the end of 'English as a Foreign Language.' London: The British Council
- Greene, J. C. (2007). *Mixed Methods in Social Inquiry*. San Francisco, CA: John Wiley & Sons.
- Greene J. C., Caracelli V. J. (2003). Making paradigmatic sense of mixed methods practice. In Tashakorri A., Teddlie C. (Eds), *Handbook of mixed methods in social & behavioral research* Thousand Oaks: Sage, 91-110
- Greene, J.C., Caracelli, V.J. & Graham, W.F., (1989). Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *Educational Evaluation and Policy Analysis*, 11(3), 255–274.
- Greenwood, C. R. (1997). Classwide peer tutoring. *Behavior and Social Issues*, 7(1), 53-57
- Gregg, K. (1984). Krashen's monitor and Occam's razor. *Applied Linguistics*, 5: 79-100.
- Grosh, M. E. and P. Glewwe, (2000). *Designing household survey questionnaires for developing countries : lessons from 15 years of the living standards measurement*, Paris
- Guba, E. G., & Lincoln, Y. S. (1998). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.). *The landscape of qualitative research: Theories and issues*. Thousand Oaks, CA: Sage, 195-220
- Guba, E. G., & Lincoln, Y. S. (1984). *Do inquiry paradigms imply inquiry methodologies?* Copyrighted manuscript.
- Gunderson, L. & Shapiro, J. (1988). Whole language instruction: Writing in first grade. *Reading Teacher*, 41(4), 430-37.

- Guthrie, J. T. & Greaney, V., (1991), 'Literacy acts', in R. Barr, M.L. Kamil, P. Mosenthal & P.D. Pearson (eds.), *Handbook of Reading Research*, Longman, New York, 68-96
- Guthrie, J. T., Wigfield, A., Humenick, N. M., Perencevich, K. C., Taboada, A., & Barbosa, P. (2006). Influences of stimulating tasks on reading motivation and comprehension. *Journal of Educational Research*, 99(4), 232-245
- Guyatt, G., Oxman, A., Kunz, R., Vist, G., Falyk-ytter, Y. & Schuneman, H. (2008). What is "quality of evidence" and why is it important to clinicians?, *British Medical Journal* 995–8
- Gwyn-Paquette, C., & Tochon, F.V. (2003). The role of reflective conversations and feedback in helping preservice teachers learn to use cooperative activities in their second language classrooms. *The Canadian Modern Language Review*, 59, 503-545
- Hallet, K., & Cummings, J. (1997). The virtual classroom as authentic experience: Collaborative, problem-based learning in a WWW environment. In Proceedings of the annual conference on Distance teaching and learning: *Competition-connection-collaboration*. Madison, WI: University of Wisconsin-Madison, 103-107
- Hallgren, A, K. (2012) Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial, *Tutor Quantitative Methods Psychology*, 8(1), 23-34
- Hamada, M., & Koda, K. (2008). Influence of first language orthographic experience on second language decoding and word learning. *Language Learning*, 58, 1–31.
- Hanushek, E,A, Woessmann, L (2008) The role of cognitive skills in economic development, *Journal of Economic Literature*, 43, (3), p607-668
- Hanushek, EA, Woessmann, L (2012) Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation, *Journal of Economic Growth*, 17 (4), 267–321
- Hardman, F., Abd-Kadir, J. and Tibuhinda, A., (2012). Reforming Teacher Education in Tanzania. *International Journals of Educational Development*, Vol. 32, pp. 826 – 834.
- Hardman, F. (2012). *Review: Teacher Support and Development Interventions*. London, Save the Children.
- Hatcher, P.J., Hulme, C., & Ellis, A.W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development*, 65, 41–57
- Hatcher, P.J., Hulme, C., & Snowling, M.J. (2004). Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure. *Journal of Child Psychology and Psychiatry*, 45, 338–358.
- He, F., Linden, L., & MacLeod, M. (2007) *Helping Teach What Teachers Don't Know: An Assessment of the Pratham English Language Program*. Mimeo Press, USA.
- Henk, W.A. (1987). Methodological concerns in experimental reading research: All that glitters *Reading Psychology: An International Quarterly*, 8(2), 61-71

- Henson, R. K. (2006). Effect size measures and meta-analytic thinking in counselling psychology research. *The Counselling Psychologist*, 34, 601– 629.
- Heyneman, S & Stern, J. (2015) ‘Development and Education’. In, Dixon, P., Humble, S., Counihan., C (Eds). *Handbook of Education and International Development*, London: Edgar Elgar Publishing
- Horwitz, E. K., Bresslau, B., Dryden, M., McLendon, M. E., & Yu, J. (1997). A graduate course focusing on the second language learner. *The Modern Language Journal*, 81, 518-526.
- Howe, K. R. (1988). Against the quantitative–qualitative incompatibility thesis or dogmas die hard. *Educational Researcher*, 17, 10–16.
- Howe, K. R. (1992). Getting over the quantitative–qualitative debate. *American Journal of Education*, 100, 236–256.
- Homan, R. (1991) *The Ethics of Social Research*, Longman: London
- Hulten BH, Devries DL 1976. Team Competition and Group Practice: *Effects on Student Achievement and Attitudes*, Johns Hopkins University, Centre for Social Organization of Schools.
- Hustler, D. (2005) in Goldbart, J. and Hustler, D., ‘Ethnography’. In Somekh, B. and Lewin, C. (eds). *Research methods in the social sciences*. London: Sage Publications.
- Hutchinson D and Styles B (2010) *A Guide to Running Randomised Controlled Trials for Educational Researchers*. Slough: NFER
- Iverson, S., & Tunmer, W.E. (1993). Phonological processing skills and the reading recovery programme. *Journal of Educational Psychology*, 85, 112–126.
- Jacobs, D. L., Watson, T. G., & Sutton, J. P. (1996). Effects of a cooperative learning method on mathematics achievement and affective outcomes of students in a private elementary school. *Journal of Research and Development in Education*, 29, 195–202.
- Jacobs, G. M., Power, M. A., & Inn, L. W. (2002). *The teacher’s sourcebook for cooperative learning*. Thousand Oaks, CA: Corwin Press, Inc.
- Jain, C, (2016) Dakar, are We There Yet! Analysing Gender Gaps in Learning in Rural India. NUECON-WP2016-002. Available at SSRN: <https://ssrn.com/abstract=2875590> (Accessed 12<sup>th</sup> October 2017)
- Jament, J. and Feng, Y. (2007) *Doing educational research in developing countries: challenges and opportunities*. Paper presented to: Research Students Conference 2007, University of Birmingham, Birmingham, 30 June 2007.
- Jamison, EA, Jamison, DT, Hanushek, EA (2007). The effects of education quality on income growth and mortality decline *Economic Education Review.*, 26 (6), 771–788

- Jenkins, J. R., Fuchs, L. S., Espin, C., van den Broek, P., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, 95(4), 719-729.
- Jhingran, D. (2005): *Language Disadvantage*. Delhi: ADP
- Jhingran, D. (2009): 'Hundreds of Home Languages in the Country and Many in Most Classrooms: Coping with Diversity in Primary Education in India' in Skuttnabb-Kangas, T. and Phillipson, R. (Eds.) *Social Justice Through Multilingual Education*. Clevedon: Multilingual Matters. Chapter 4, 263–282.
- Jiménez, R. T., Garcia, G. E., & Pearson, P. D. (1995). Three children, two languages, and strategic reading: Case studies in bilingual/monolingual reading. *American Educational Research Journal*, 32, 67–97.
- Johnson, D. & Johnson, R. (2009): Energizing learning: The instructional power of conflict, *Educational Researcher*, vol. 38, no. 1, 37-51
- Johnson, D. W, & Johnson, F. (2008) Promoting Early Adolescents' Achievement and Peer Relationships: The Effects of Cooperative, Competitive, and Individualistic Goal Structures *Psychological Bulletin*. Vol. 134, No. 2, 223–246
- Johnson, D. W. & Johnson, F, Roy, P, & Zaidman, B. (1985) Oral interaction in cooperative learning groups: Speaking, listening and the nature of statements made by high-medium, and low achieving students, *Journal of Psychology*, Vol: 119, No 4, 303-321
- Johnson, D. W., & Johnson, F. (2006). *Joining together: Group theory and group skills* (9th ed.). Boston: Allyn & Bacon.
- Johnson, D. W., & Johnson, R. (1999). *Learning together and alone: Cooperative, competitive, and individualistic learning*. Boston: Allyn & Bacon
- Johnson, D. W., & Johnson, R. T. (2013). The impact of cooperative, competitive, and individualistic learning environments on achievement. In J. Hattie & E. Anderman (Eds.), *International handbook of student achievement* (372-374). New York: Routledge
- Johnson, D. W., & Johnson, R. T. (2014). Cooperation in the 21st century. In J. M. Serrano & R. M. Pons (Eds.), *Anales de Psicología* (Special Issue on Cooperative Learning), 30(3, Octubre).
- Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and Research*, Edina, MN: Interaction Book Company.
- Johnson, D. W., & Johnson, R. T. (2005). New developments in social interdependence theory, *Psychological Monographs*, 131(4), 285-358.
- Johnson, D. W., & Johnson, R. T. (2007). *Creative controversy: Intellectual challenge in the classroom* (4th ed.). Edina, MN: Interaction
- Johnson, D. W., and Johnson, R. T. (2009). An educational psychology success story: social interdependence theory and cooperative learning. *Education Research*. 38, 365–379

- Johnson, D.W. & Johnson, R.T. (1990). Cooperative learning and achievement. In S. Sharan (ed.), *Cooperative learning: Theory and research*, 23-37. New York: Praeger
- Johnson, D.W. & Johnson, R.T. (1994) *Learning together and alone* (4th ed.), Needham Heights, MA: Allyn and Bacon.
- Johnson, D.W., & Johnson, R.T. (1991). *Cooperation in the classroom*. Edina, MN: Interaction Book Company.
- Johnson, D.W., Johnson, R.T., and Stanne, M.B.,(2000) "Cooperative Learning Methods: A Meta- Analysis," Available at: <http://www.cooperation.org/pages/cl-methods.com> Accessed September 2013
- Johnson, R & Watson, J. (2005) *The Effects of Synthetic Phonics Teaching on Reading and Spelling Attainment: A Seven Year Longitudinal Study*. Retrieved from: <http://www.scotland.gov.uk/library5/education/sptrs.pdf>
- Johnson, R. B., Onwuegbuzie, A. J. & Turner, L. A. 2007. Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1, 112-133.
- Johnson, R.T., Johnson, D.W., and Holubec, E.J. (1998). *Cooperation in the Classroom*. Boston: Allyn and Bacon.
- Johnson, D. & Johnson, R. (2002) Learning together and alone: overview and meta-analysis *Asia Pacific Journal of Education*, 22 (2002), 95–105  
Psychological Monographs, 131(4), 285-358.
- Johnston, R. S. and J. E. Watson (2005). The effects of synthetic phonics teaching on reading and spelling development: a seven year longitudinal study. *S. E. E. Department*. Edinburgh
- Johnstone, R. (2009). An early start: What are the key conditions for generalized success? In J. Enever, J. Moon & U. Raman (Eds), *Young learner English language policy and implementation: International perspectives*. Reading: Garnet Education, 31-41
- Jolliffe, W., et al. (2012). *Teaching systematic synthetic phonics in primary schools*. London, Sage.
- Joshi RM, Aaron PG (1991). Specific spelling disability: Factual or artifactual? *Reading and Writing: An Interdisciplinary Journal*, 2(1): 107-125.
- Joshi RM, Aaron PG (2003). A new way of assessing spelling and its classroom applications. In: RM Joshi, B Kaczmanek, CK Leong (Eds.): *Literacy Acquisition Assessment and Instruction: The Role of Phonology, Orthography and Morphology*. Amsterdam/Holland: IOS Press, 153-161.
- Jun, S.W., Ramirez, G., & Cumming, A. (2010). Tutoring Adolescents in Literacy: A Meta-Analysis. *Journal of Education*, 45 (2), 219-238
- Juvonen, J. J. (1996). Self-presentation tactics promoting teacher and peer approval: The function of excuses and other clever explanations. In J. Juvonen & K. Wentzel (Eds.), *Social motivation: Understanding children's school adjustment*. New York: Cambridge University Press, 43-65

- Juvonen, J. J., Graham, S., & Schuster, M. (2003). Bullying among young adolescents: The strong, weak, and troubled. *Pediatrics*, 112, 1231– 1237.
- Kagan, M., Robertson, L. & Kagan, S. (1995) *Cooperative learning structures for class building*, San Clemente, CA: Kagan Cooperative Learning.
- Kagan, S. (1994) *Cooperative learning*. San Juan Capistrano, CA: Kagan Cooperative Learning
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 32-36.
- Kalia, V & Reese, E (2009) 'Relations between Indian Children's Home Literacy Environment and Their English Oral Language and Literacy Skills', *Scientific Studies of Reading*, 13: 2, 122 — 145
- Kalof, L., Dan, A., and Dietz, T. (2008) *Essentials of Social Research* Maidenhead, England: Open University Press
- Kanu, Y. (1996) Educating teachers for the improvement of the quality of basic education in developing countries, *International Journal of Educational Development*, 16(2), 173-184
- Keogh-Brown, K, Bachmann M, Shepstone L, Hewitt C, Howe A, Ramsay CR, et al. (2007) Contamination in trials of educational interventions. *Health Technology Assessment*;11(43).
- Kerry SM, Cappuccio FP, Emmett L, Eastwood JB, Plange-Rhule J. (2005) Reducing selection bias in a cluster randomised trial in 12 West African villages. *Clinical Trials*, 2: 125-9
- Kilic, D. (2008). The Effects of the Jigsaw Technique on Learning the Concepts of the Principles and Methods of Teaching. *World Applied Sciences Journal*, 4(1), 109-114
- Kingdom, G. and Muzammil, M. (2009). *A political economy of education in India: The case of Uttar Pradesh*. *Oxford Development Studies*, 37(2): 123–144
- Kirk, R.E., (1996). Practical significance: a concept whose time has come. *Educational and Psychological Measurement*, 56, 746–759
- Kirschner, F., Paas, F., & Kirschner, P. A. (2009). A cognitive-load approach to collaborative learning: United brains for complex tasks. *Educational Psychology Review*, 21, 31-42.
- Kline, R. B. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guildford.
- Klingner, J.K., & Vaughn, S. (2000). The Helping Behaviors of Fifth Graders While Using Collaborative Strategic Reading during ESL Content Classes. *TESOL Quarterly*, 34, 69-98



- Koda, K. (2004) *Insights into Second Language Reading: A Cross-Linguistic Approach*, New York, Cambridge University Press
- Koffka, K. (1935). *Principles of Gestalt psychology*, New York: Harcourt, Brace, & World.
- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Oxford. Pergamon
- Krueger, R. A. and M. A. Casey (2009). *Focus groups: a practical guide for applied research*. Los Angeles, Sage Publications.
- Kulik, J. A. and Kulik, C. -L. C. (1987). Effects of ability grouping on student achievement. *Equity and Excellence in Education*, 23: 22–30.
- Kutnick, P., Ota, C., & Berdondini, L. (2006). Improving the effects of group working in classrooms with young school-aged children: Facilitating attainment, interaction and classroom activity. *Learning and Instruction*, 18, 83-95
- Ladd, G. W. (1990). Having friends, keeping friends, making friends, and being liked by peers in the classroom: Predictors of children's early school adjustment *Child Development* (61), 1081-1100
- Lan, Y. J., Chang, K. En., & Sung, Y. T. (2006). Cooperative learning in small EFL early reading groups. *Pan-Pacific Association of Applied Linguistics*, 9, 319-341.
- Landis JR and Koch GC (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159-174.
- Lang, J.M., K.J. Rothman and C.I. Cann (1998), "That confounded p-value," *Epidemiology*, 9(1): 7-8.
- Lantolf, J.P., 2000. Introducing sociocultural theory. In: Lantolf, J. P., ed. *Sociocultural theory and second language learning*. Oxford University Press, 1-26
- Lantolf, J. P., & Pavlenko, A. (1995). Sociocultural theory and second language acquisition. *Annual Review of Applied Linguistics*, 15, 108–124.
- Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, 37, 822-32.
- Lesaux, N, K and Siegel, L, S, (2003) The Development of Reading in Children who speak English as a Second Language, *Developmental Psychology*, 39(6), 1005-1019
- Levene, H. (1960). Robust tests for equality of variance. In I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow & H. B. Mann (Eds.), *Contributions to probability and statistics: Essays in honor of Harold Hotelling* (Stanford: CA: Stanford University Press, 278, 292
- Lewin, K. (1935). *A dynamic theory of personality*. New York: McGraw-Hill.
- Lewin, K. (1948). *Resolving social conflicts*. New York: Harper.

- Lewin, K., 2007. The Limits to Growth of Non-government Private Schooling in Sub-Saharan Africa. In: Srivastava, P., Walford, G. (Eds.), *Private Schooling in Less Economically Developed Countries: Asian and African Perspectives*. Symposium, Oxford, UK, 41-65.
- Lewis, A. (1992). Group child interviews as a research tool. *British Educational Research Journal*, 18(4), 413-421.
- Lewis, M. and Ellis, S. (eds) (2006) *Phonics: Practice, Research and Policy*. London: Sage
- Liang, T. (2002). *Implementing cooperative learning in EFL teaching: Process and effects*. Doctoral dissertation, National Taiwan Normal University
- Lidz, C. S. (1991). The extent of incorporation of dynamic assessment into cognitive assessment courses: A national survey of school psychology trainers. *The Journal of Special Education*, 26, 325-331.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications.
- Lindberg, D. C. (2007) *The Beginnings of Western Philosophy: The European Scientific Tradition in Philosophical, Religions and Institutions Context, Prehistory to A.D. 1450* 2<sup>nd</sup> Edition, USA, Chicago University Press
- Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading and Writing Quarterly*, 13, 123–146.
- Lonigan CJ, Schatschneider C, Westberg L. (2008) National Early Literacy Panel, Developing early literacy: Report of the National Early Literacy Panel. Washington, DC: National Institute for Literacy; Impact of code-focused interventions on young children's early literacy skills; 107–151
- Lonigan, C. J. (2006). Conceptualizing phonological processing skills in prereaders. *Handbook of Early Childhood Literacy Research*. D. K. Dickinson and S. B. Neuman. New York, The Guilford Press. 2: 77-89
- Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within class grouping: A meta-analysis. *Review of Educational Research*, 66, 423–458
- Macleane, M., Bryant, P.E., & Bradley, L. (1987). Rhymes, nursery rhymes and reading in early childhood. *Merrill Palmer Quarterly*, 33, 255-282.
- Mallikarjun, B. (2010) Patterns of Indian Multilingualism *Language in India*. Vol 10, 1-18.
- Marinak, B. A., & Gambrell, L. B. (2010). Reading motivation: Exploring the elementary gender gap. *Literacy Research and Instruction*, 49(2), 129-141.
- Mathes, P. G., Howard, J. K., Allen, S. H., & Fuchs, D. (1998). Peer-assisted learning strategies for first-grade readers: Responding to the needs of diverse learners. *Reading Research Quarterly*, 33, 62–94.
- Mathewson, G. C. (1985). Toward a comprehensive model of affect in the reading process. In: R. B. Ruddel, M. R. Ruddell, & H. Singer (Eds.) *Theoretical Models and Processes of Reading*. Newark, DE: International reading association, 841-856

- Mathewson, G. C. (1994). Model of attitude influence upon reading and learning to read. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical Models and Processes of Reading* (4th ed) Newark, DE: International Reading Association, 1131-1161
- Matsunaga, M (2010). How to factor-analyse your data right: Do's, Don't's and How To's, *Journal of International Psychological Research*, 3, (1), 97-110
- Maughan, B. (2013). Editorial: "Better by design" - why randomized controlled trials are the building blocks of evidence-based practice. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 54, 225–6.
- McGuire, W. J. (1989) The structure of individual attitudes. In A. Pratkanis, S. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 129-151). Hillsdale, New Jersey, Erlbaum.
- Mayo-Wilson E, Grant S, Hopewell S, Macdonald G, Moher D, Montgomery P. (2013) Developing a reporting guideline for social and psychological intervention trials. *Trials*.14:242.
- McCarty, T. L. (2015). Ethnography in educational linguistics. In M. Bigelow & J. Enns-Kananen (eds.), *The Routledge handbook of educational linguistics* (pp. 23–37). New York, NY: Routledge
- McClelland, D. C. (1961). *The Achieving Society*. New York: Free Press.
- McGraw, K. O., & Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin*, 111, 361–365.
- McGroarty, M. (1989). The benefits of cooperative learning arrangements in second language instruction. *NABE Journal*, 13, 127-143.
- McKenna, M. C. (1994). Toward a model of reading attitude acquisition. In: E. H. Castle, & M. Castle. (Eds.) *Fostering the life-long love of reading: The affective domain in reading education*. Newark, DE: International reading association, 18-40
- McKenna, M.C., Kear, D.J. & Ellsworth, R.A., (1995), 'Children's attitudes toward reading: A national survey', *Reading Research Quarterly* 30 (4), 934–956.
- McKenna, M.e., & Kear, D.J. (1990). Measuring attitude toward reading: A new tool for teachers. *The Reading Teacher*. 43. 626- 639.
- McLean, M (2004) Does the curriculum matter in peer mentoring? From mentee to mentor in problem-based learning: a unique case study *Mentoring and Tutoring: Partnership in Learning* Vol. 12, 2,173-186
- Mead, G.H. (1936). *Mind, self and society*. Chicago: University of Chicago Press
- Meganathan, R. (2011) "Language Policy in Education and the Role of English in India: From Library Language to Language of Empowerment,". In: Coleman Hywel, editor. *Dreams and Realities: Developing Countries and the English Language*. London: British Council

- Memon, S. (2014) *Reading attitudes in L1 and L2 among rural and urban learners in a Pakistani context*, PhD Thesis, Bedfordshire University, Bedfordshire
- Mercer, N., and K. Littlejohn. 2007. *Dialogue and the Development of Children's Thinking: A Sociocultural Approach*, London: Routledge.
- Messier, W. (2003). Traditional teaching strategies versus cooperative teaching strategies: Which can improve achievement scores in Chinese middle schools? *Journal of Student Centred Learning*, 2 (3), 231-238.
- Miller, L. J, & Kohler, F. W. (1993). Winning with peer tutoring. *Preventing School Failure*, Vol 37,(3) 14-18
- Mingers, J. (2003) "The paucity of multimethod research: a review of the information systems literature", *Information Systems Journal*, Vol. 13, 233-249.
- Miskin, R. (2006). *Read-Writ Inc: Phonics Handbook*. Oxford University Press.
- Mishra, P. K. (2000). English language, postcolonial subjectivity, and globalization in India. *A Review of International English Journal*, 31(2), 383–410
- Mishra, R. & Stainthorpe, R. (2007) \_The relationship between phonological awareness and word reading accuracy in Oriya and English: A study of Oriya-speaking fifth- graders. *Journal of Research Reading*, Vol 30 (1) 23-37
- Mitchell, L. (2009). *Language, emotion, and politics in South India: The making of a mother tongue*. Bloomington, IN: Indiana University Press
- Mitra S, Dangwal R, Thadani L.(2008). Effects of remoteness on the quality of education: A case study from North Indian schools. *Australasian Journal of Educational Technology*, 24(2), 168-180
- Moats, L. (2007). *Whole language high jinks: How to Tell When Scientifically-Based Reading Instruction*. Washington, DC: Thomas B. Fordham Institute
- Moats, L.C. (2005). How spelling supports reading: And why it is more regular and predictable than you may think. *American Educator*, 29 (4), 12–22
- Mohanty, D, Roy, B. (2013) *Need of Cooperative Learning in Indian Language Classroom: A Study*, *International Multidisciplinary Research Journal*, Vol 2, 1-4
- Moll, L. C. (1994). Literacy research in community and classrooms: A sociocultural approach. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading*. Newark, DE: International Reading Association, 179-207
- Moll.C. (1990). *Vygotsky and Education: Instructional implications and applications of socio-historical psychology*. Melbourne: The Cambridge Pres.
- Montgomery, P., Grant, S. P., Hopewell, S., Macdonald, G., Moher, D., Michie, S., & Mayo-Wilson, E. (2013). Protocol for CONSORT SPI: An extension for social and psychological interventions. *Implementation Science*, 8, 99.

- Morgan, R. L., Whorton, J. E., & Gunsalus, C. (2000). A comparison of short-term and long-term retention: Lecture combined with discussion versus cooperative learning. *Journal of Instructional Psychology*, 27, 53-58.
- Mori, S. (2002). Redefining motivation to read in a foreign language. *Reading in a Foreign Language*, 14, 91–110.
- Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research*, 40, 120–123
- Morse, J. M., & Niehaus, L. (2009). *Mixed method design: Principles and procedures*. Walnut Creek, CA: Left Coast Press, Inc.
- Mukherjee, D. (2011). Reducing out-of-school children in India: lessons from a micro study. National University of Educational Planning and Administration, Delhi, India
- Mullainathan, S (2005). “Development Economics Through the Lens of Psychology” in *Annual World Bank Conference in Development Economics 2005: Lessons of Experience*, edited by Francois Bourguignon and Boris Pleskovic. Oxford, UK and Washington, DC: Oxford University Press and World Bank
- Muralidharan, K and Das, J, Holla, A & Mohpal, A, (2016) The Fiscal Cost of Weak Governance: Evidence from Teacher Absence in India. World Bank Policy Research Working Paper No. 7579. Available here: <https://ssrn.com/abstract=2738335> (Accessed September 2016)
- Muralidharan, K & Sundararaman, V (2015). The aggregate effect of school choice: evidence from a two-stage experiment in India, *The Quarterly Journal of Economics*, 130 (3), 1011-1066
- Murphy, V. (2015). *A systematic review of intervention research examining English language and literacy development in children with English as an Additional Language*, Education Endowment Fund, Oxford, London
- Nair, P. (2004). *India: Desk Study Of Non- State Providers of Basic Services*; International Development Department, University Of Birmingham
- Naleppa, M. J., & Cagle, J. G. (2010). Treatment fidelity in social work intervention research: a review of published studies. *Research on Social Work Practice*, 20, 674–681
- Nambissan, G. B. (2003) Educational deprivation and primary school provision: a study of providers in the city of Calcutta, IDS Working Paper 187 (Brighton, Institute of Development Studies).
- Nayar, U. (2002). “Education of girls in India: An assessment.” Pp. 35-46 in *India Education Report*, edited by Govinda, R. New Delhi: National Institute of Educational Planning and Administration and Oxford University Press.
- NCERT (2005) *National Curriculum Framework*, New Delhi,
- NCERT (2006), *Syllabus for classes at the Elementary Level*; National Council for Educational Research and Training, New Delhi.

- NCERT (2012) *Teaching of English at Primary Level in Government Schools*, New Delhi,
- Neuman, W.L. (2006) *Social Research Methods: Qualitative and Quantitative Approaches* (6th Ed), Boston, New York: U.S.A.
- Newby, P. (2010) *Research Methods for Education*, England: Pearson Education Limited
- Newton, M.J., & Thomson, M.E. (1978). *The Aston Index: A classroom test for screening and diagnosis of language difficulties (Revised)*. Wisbech, Cambridge: Learning Development Aids.
- Ng, M. & Lee, C. (1996) What's different about cooperative learning? - and its significance in social studies teaching, *Teaching and Learning* 17:1, 15-23.
- Nixon, J., & Topping, K. J. (2000). Emergent writing: The impact of structured peer interaction. *Educational Psychology*, 21, 41–58
- NYRS (2009). *National Youth Readership Survey, Basis of Indian Youth: Demographics and Readership*, NBT-NCAER, New Delhi
- Oakley, A., V. Strange, T. Toroyan, M. Wiggins, I. Roberts, and J. Stephenson. (2003) Using Random Allocation to Evaluate Social Interventions: Three Recent U.K. Examples.” *The Annals of the American Academy of Political and Social Science*, 589 (1): 143-167
- Parveen, Q., Mahmood, S. T., Mahmood, A. & Arif, M. (2011). Effect of cooperative learning on academic achievement of 8th grade students in the subject of social studies. *International Journal of Academic Research*, 3(1), 950-954.
- Patton, M.Q. (1990). *Qualitative evaluation and research methods*. Newbury Park, California: Sage.
- Patton, M.Q. (2002). *Qualitative evaluation and research methods*. (3rd ed.). Thousand Oaks, CA: Sage.
- Payne, G & Payne, J. (2004) *Key Concepts in Social Research*, Singapore, SAGE Publications, Asia Pacific
- Peers, I (1996) *Statistical Analysis for Education and Psychology Researchers: Tools for researchers in education and psychology*, London, Falmer Press
- Pellegrino, W, J. Assessment as a positive influence on 21st century teaching and learning: A systems approach to progress, *Psicología Educativa*, Volume 20, Issue 2, 2014, 65-77
- Perry, A. E., Weisburd, D., & Hewitt, C. (2010). Are criminologists describing randomized controlled trials in ways that allow us to assess them? findings from a sample of crime and justice trials. *Journal of Experimental Criminology*, 6, 245–262.
- Peterson, J. M. (1989). Tracking students by their supposed abilities can derail learning. *American School Board Journal*, 176(5), 38–46.

- Piaget, J. (1950). *The Psychology of Intelligence*. M. Piercy, & D. E. Berlyne, Trans. London: Kegan Paul, Trench, & Trubner.
- Pocock Stuart J., M. D. Hughes, and R. J. Lee. 1987. Statistical problems in the reporting of clinical trials. *New England Journal of Medicine* 317:426-32
- Pollo, T C. Kessler, B., & Treiman, R. (2005). Vowels, syllables, and letter names: Differences between young children's spelling in English and Portuguese. *Journal of Experimental Child Psychology*, 92 (2): 161-181
- Pontefract, C & Hardman, F (2005) The discourse of classroom interaction in Kenyan primary schools, *Comparative Education*, 41(1), 87-106
- Powell, M, A. (1997) *Academic Tutoring and Mentoring: A Literature Review* USA, California Research Bureau
- Prabhu, N. S. (1987). *Second language pedagogy* (Vol. 20). Oxford: Oxford University Press.
- Pressley, M. (2006) *Reading instruction that works, the case for balanced teaching* (3rd ed.). New York: The Guilford Press.
- Probe Team. (1999). *Public report on basic education in India*. New Delhi: Oxford University Press.
- Puzio, K. & Colby, G. T. (2013). Cooperative learning and literacy: A meta-analytic review. *Journal of Research on Educational Effectiveness*, 6(4), 339-360
- Ramanathan, H. (2008). Testing of English in India: A developing concept. *Language Testing*, 25(1), 111–126.
- Randi, J. Newman, T. Grigorenko, E. (2010) Teaching Children with Autism to Read for Meaning: Challenges and Possibilities *Journal of Autism and Developmental Disorders* Vol 40: 7, 890-902
- Rao, N, Kai-Ming C, and Narain, K. (2003). “Primary schooling in China and India: Understanding how socio-contextual factors moderate the role of the state.” *International Review of Education* Vol 49:153-176.
- Rasinski, T. V., Reutzel, R., Chard, D. & Linan-Thompson, S. (2011). Reading fluency. In M. L., Kamil, P. D. Pearson, E. B. Moje, & P. Afflerbach (Eds.), *Handbook of Reading Research* (Vol., 4), 296-319.
- Raven, B, & Eachus, H (1963) Cooperation and competition in means-interdependent triads, *Journal of Abnormal and Social Psychology*, 67, 307-309.
- Read, C. (1975). *Children's categorization of speech sounds in English*. (NCTE Research Report No. 17). Urbana, IL: National Council of Teachers of English.
- Read, C. (1986). *Children's Creative Spelling*. London: Routledge & Kegan Paul.

- Reese, E., et al. (2010) A Review of parent interventions for preschool children's language and emergent literacy. *Journal of Early Childhood Literacy* 10, 1, 97-117
- Reeves, C. (2002), July–August). *Literacy attitudes: Theoretical perspectives*. Paper presented at the 19th World Congress on Reading, Edinburgh, Scotland.
- Reutzel, D. R., & Cooter, R. B. (1991). Organizing for effective instruction: The reading workshop. *The Reading Teacher*, 44, 548–554
- Robson, C. (2002). Real world research: A resource for social scientist and practitioners-researchers, 2nd edition, Cornwall, UK T.J. International Ltd.
- Rock, I, Palmer, S (1990) *The Legacy of Gestalt Psychology*, Scientific American, USA.
- Roebuck, R. (1998) *Reading and recall in L1 and L2: a sociocultural approach*, Ablex Publishing Corporation, Stamford, 8,55-60.
- Rogers, E. M. 2003. *Diffusion of Innovations*. New York: Free Press.
- Rogoff, B. (1994). Developing understanding of the idea of communities of learners. *Mind, Culture, and Activity*, 1, 209-229.
- Rose, J. (2006). *Independent review of the teaching of early reading*. Nottingham, DfES Publications.
- Rothwell P. M. (2005) External validity of randomised controlled trials: “to whom do the results of this trial apply?” *Lancet*; 365: 82–93
- Ruddell, R. B., & Speaker, R. (1985). The interactive reading process: A model. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (3rd ed). Newark, DE: International Reading Association, 751-793
- Ruddell, R. B., & Unrau, N. J. (1994). *Reading as a meaning-construction process: The reader, the text, and the teacher*. *Theoretical models and processes of reading* (4th ed), 996-1056
- Ryen, A. (2011). *Ethical issues. Qualitative Research Practice*. C. Seale, G. Gobo, G. J. F. and D. Silverman. London, Sage Publications Inc.: 230-247.
- Sachs, G. T., Candlin, C., Rose, K. R. & Shum, S. (2003). Learner behaviour and language acquisition project: Developing cooperative learning in the EFL/ESL secondary classroom. *RELC Journal*, 34 (3), 338-369.
- Sanders, G., Baron, R., & Moore, D. (1978). Distraction and social comparison as mediators of social facilitation effects. *Journal of Experimental Social Psychology*, 14, 291–303
- Sawyer, K. (1997). *Pretend Play as Improvisation*. Mahwah, NJ.: LEA.
- Schmidt, F., 1996. Statistical significance testing and cumulative knowledge in psychology: Implications for the training of researchers. *Psychological Methods* 1, 115–129
- Schonell, F., & Goodacre, E. (1971). *The psychology and teaching of reading*. London: Oliver & Boyd.



- Schulz, K.F., Grimes, D.A. (2002) Allocation concealment in randomised trials: defending against deciphering. *Lancet*. ;359:614–618
- Schunk, D. H., & Zimmermann, B. J. (Eds.). (1994). *Self-regulation of learning and performance*. New York: Lawrence Erlbaum.
- Semaan, S., Kay, L., Strouse, D., Sogolow, E., Mullen, P. D., Neumann, M. S., et al. (2002). A profile of U.S.-based trials of behavioral and social interventions for HIV risk reduction. *Journal of Acquired Immune Deficiency Syndromes*, 30, S30–S50.
- Senn, N. (2012). Effective approaches to motivate and engage reluctant boys in literacy. *Reading Teacher*, 66(3), 211-220.
- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143–174
- Sharan, S. (1980) Cooperative learning in small groups: Recent methods and effects on achievement, attitudes and ethnic relations, *Review of Educational Research*, 50(2), 241-271.
- Sharan, S. (2010) Cooperative learning: a diversified pedagogy for a diverse classroom, *Intercultural Education*, 21:3,195-203
- Sharan, S., Shachar, H., & Levine, T. (1999). *The innovative school: Organization and instruction*. Westport, CT: Bergin & Garvey.
- Sharan, Y. (2010) Cooperative Learning for Academic and Social Gains: valued pedagogy, problematic practice *European Journal of Education, Research, Development and Policy* Vol 45, 2, 300-313
- Sharan, Y., & Sharan, S. (1992). *Expanding Cooperative Learning Through Group Investigation*. New York, NY: Teachers College Press
- Sharpley, A. M., & Sharpley, C. F. (1981). Peer tutoring: A review of the literature. *Collected Original Resources in Education*, 5(3), 7–C11.
- Shaw, I. F., Greene, J. C., & Mark, M. M. (2006). *The SAGE Handbook of Evaluation*. London: SAGE Publications.
- Shrout, P. E. and Fleiss, J. L. 1979. Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86: 420–428
- Silverman W. (1990) Frequently Performed Psychological Tests. In: Walker HK, Hall WD, Hurst JW, editors. *Clinical Methods: The History, Physical, and Laboratory Examinations*. 3rd edition. Boston: Butterworths; Chapter 208. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK321/> (Accessed September 2011)
- Silverman, D. (2000). *Doing Qualitative Research: A Practical Handbook*. Thousand Oaks, CA: Sage
- Singh, P, V. (2004) *Problems with Educational Backwardness*, Sarup and Sons, New Delhi

- Siraj-Blatchford, I. (2009) Learning in the home and at school: how working class children 'succeed against the odds' *British Educational Research Journal* Vol, 36:10, 463-482
- Skinner, B. F. (1968) *The Technology of Teaching*. Englewood Cliffs, NJ: Prentice-Hall
- Slavin, R. (1978) Student Teams and Achievement Divisions. *Journal of Research and Development in Education*, 12, 39-49
- Slavin, R. (1991). Group rewards make groupwork work. *Educational Leadership*, 5, 89-91.
- Slavin, R. E. (1977). How student learning teams can integrate the desegregated classroom, *Integrated Education*, 15, 5-58
- Slavin, R. E. (1991). Synthesis of research on cooperative learning. *Educational Leadership*, 48, 71-82.
- Slavin, R. E. (1995). *Cooperative Learning: Theory, Research, and Practice*. Englewood Cliffs, New Jersey: Prentice-Hall
- Slavin, R. E. 1987. Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57: 293–336.
- Slavin, R. Leavey, M. Madden. N (1986) *Team Accelerated Instruction- Mathematics*, Watertown Mass: Mastery Education Company
- Slavin, R.E. (1985) Cooperative Learning: Applying Contact Theory in Designated Schools *Journal of Social Issues* 41, 3: 42-62
- Slavin, R.E. (1987). Developmental and motivational perspectives on cooperative learning: A reconciliation. *Child Development*, 58, 1161-1167.
- Slavin, R.E., Lake, C., & Groff, C. (2009a), Effective programs in middle and high school mathematics: A best evidence synthesis. *Review of Educational Research*, 79 (2), 839 – 911
- Slavin, R.E., Lake, C., Davis, S., & Madden, N. (2009b), *Effective Programs for Struggling Readers: A Best Evidence Synthesis*, Baltimore MD: Johns Hopkins University, Centre for Research and Reform in Education (CRRE)
- Slavin, R, Leavey, M, Madden, N (1984) Combining cooperative learning and individualized instruction: Effects on student mathematics achievement, attitudes, and behaviors. *Elementary School Journal*, 84:409–422.
- Smail, A (2013) Rediscovering the teacher within Indian child-centred pedagogy: implications for the global Child-Centred Approach, *Compare: A Journal of Comparative and International Education*, Vol 44 (4), pp 613-633
- Smith, K., Johnson, D. & Johnson R. (1981): Can conflict be constructive? Controversy versus concurrence seeking in learning groups, *Journal of Educational Psychology*, vol. 73, no. 5, 651-663.

- Smith, M. C. (1990). A longitudinal investigation of reading attitude development from childhood to adulthood. *The Journal of Educational Research*, 83, 215–21
- Snow, C.E., Burns, M. S. Griffin, P. Ed. (1998). Preventing reading difficulties in young children. Prepared by the committee on the Preservation of Reading Difficulties in Young Children. Washington, DC: Department of Education.
- Snowling, M. J., et al. (2007). Children at family risk of dyslexia: A follow-up in adolescence. *Journal of Child Psychology and Psychiatry*, 48: 609-618.
- Sonalde, D. & Vanneman, R. (2005) *National Council of Applied Economic Research. India Human Development Survey*, Inter-university Consortium for Political and Social Research
- Spencer, K.A. (2007). Predicting children’s word-spelling difficulty for common English words from measures of orthographic transparency, phonemic and graphemic length and word frequency. *British Journal of Psychology*, 98, 305-338
- Sriprakash, A. 2010. Child-Centred Education and the Promise of Democratic Learning: Pedagogic Messages in Rural Indian Primary Schools. *International Journal of Educational Development* 30 (3): 297–304.
- Stahl, S. A., & Miller, P. D. (1989). Whole language and language experience approaches for beginning reading: A quantitative research synthesis. *Review of Educational Research*, 59, 87-116.
- Stannard, J. (2006). Keeping Phonics in Perspective. In M. Lewis & S. Ellis (Eds.), *Phonics. Practice, research, policy*, London, England: Paul Chapman Publishing.
- Stanovich, K.E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: The Guilford Press.
- Stanovich, K. E., Nathan, R., West, R. F. & Vala-Rossi, M. (1985) Children’s word recognition in context: Spreading activation, expectancy, and modularity. *Child Development*, 56:1418–28
- Strauss, A. L., & Corbin, J. (1998). *Basics of qualitative research: techniques and Procedures for Developing Grounded Theory*. London: Sage Publications.
- Steinberg, L. (1990). Autonomy, conflict, and harmony in the family relationship. In S. S. Feldman & G. R. Elliott (Eds.), *At the threshold: The developing adolescent*. Cambridge, MA: Harvard University Press.
- Stevens, R. J., & Slavin, R. (1995). Effects of a cooperative learning approach in reading and writing on academically handicapped and non-handicapped students. *Elementary School Journal*, 95: 241-262.
- Stewart-Brown, S., R. Anthony, L. Wilson, S. Wintanley, N. Stallard, H. Snooks, and D. Simkiss. (2011). “Should Randomised Controlled Trials be the ‘Gold Standard’ for Research on Preventive Interventions for Children?” *Journal of Children's Services*, 6 (4): 228–235

- Stinson, J. N., McGrath, P. J., & Yamada, J. T. (2003). Clinical trials in the Journal of Pediatric Psychology: applying the CONSORT statement. *Journal of Pediatric Psychology*, 28, 159–167.
- Stuart-Smith, J. & Martin, D. (1999). Developing assessment procedures for phonological awareness for use with Punjabi/English bilingual children. *The International Journal of Bilingualism*, 3, 55–80.
- Stuart, M (2004) Getting ready for reading: A follow up study of inner-city second language learners at the end of Key Stage 1. *British Journal of Educational Psychology*, 74(1)15-36
- Stuart, M. (1999). Getting ready for reading: Early phoneme and phonics teaching improves reading and spelling in inner-city second language learners, *British Journal of Educational Psychology*, 69(4), 587-605
- Stringer, D. (2015), ‘Embedded wh-questions in L2 English in India’, *Studies in Second Language Acquisition* pp. 1–33.
- Sullivan, M, G (2011) Getting Off the “Gold Standard”: Randomized Controlled Trials and Education Research. *Journal of Graduate Medical Education: Vol. 3, No. 3*, 285-289.
- Sun, S., Pan, W., & Wang, L. L. (2010). A Comprehensive Review of Effect Size Reporting and Interpreting Practices in Academic Journals in Education and Psychology. *Journal of Educational Psychology*, 102 (4), 989-1004
- Tan, I. G. C., Lee, C.K.E. & Sharan, S. (2007). Group investigation effects on achievement, motivation, and perceptions of students in Singapore. *The Journal of Educational Research*, 100(3), 142-154.
- Tashakkori, A., & C. Teddlie (Eds.), (2003) *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: Sage Publications.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage
- Teasley, S. D., & Roschelle, J. (1993). Constructing a joint problem space: The computer as a tool for sharing knowledge. In S. P. Lajoie, S. J. Derry (Eds.), *Computers as cognitive tools*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Teddlie, C. & A. Tashakkori (2009). *Foundations of Mixed Methods Research*. Thousand Oaks, CA: Sage Publications.
- Thanh-Pham, TH, Gillies, R and Renshaw, P. (2009). ‘Cooperative learning (CL) an academic achievement of Asian students: A true story’, *International Education Studies*, 1(3), 82-88.
- Thompson, B. (2004). *Explanatory and confirmatory factor analysis*. Washington DC: American Psychological Service.
- Thorpe, W, G. (1976) *The Burt Reading Test 1974 Revised*, SCRE, Scotland
- Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago, IL: University of Chicago Press

- Tooley, J. (2009), *The Beautiful Tree: A personal journey into how the world's poorest people are educating themselves*, Penguin, New Delhi.
- Tooley, J., et al. (2005). "Private and public schooling in low-income areas of Lagos State, Nigeria: A census and comparative survey." *International Journal of Educational Research*, 43: 125-146.
- Topping, K (2005) Trends in Peer Learning, *Educational Psychology: An International Journal of Experimental Educational Psychology*, (25), 6, 631-645
- Torgerson C. J., Torgerson D. J., Birks Y. F., Porthouse J. (2005). A comparison of RCTs in health and education. *British Educational Research Journal*, 31, 761–785
- Torgerson D. J, Roberts C. (1999) Randomisation methods: concealment, *BMJ*, 319 :375
- Torgerson, C. J. (2009). Randomised controlled trials in education research: A case study of an individually randomised pragmatic trial. *Education 3-13: International Journal of Primary, Elementary and Early Years Education*, 37(4), 313-321.
- Torgerson, C. J. and Torgerson, D. J. (2001), The Need for Randomised Controlled Trials in Educational Research. *British Journal of Educational Studies*, 49: 316–328.
- Torgerson, C. J., Wiggins, A., Torgerson, D., Ainsworth, H., & Hewitt, C. (2013). Every Child Counts: Testing policy effectiveness using a randomised controlled trial, designed, conducted and reported to CONSORT standards. *Research in Mathematics Education*, 15, 141–153.
- Torgerson, C., et al. (2006). *A Systematic review of the research literature on the use of phonics in the teaching of reading and spelling*. DfES. University of Sheffield.
- Torgerson, D. J., and C. Torgerson. (2008). *Designing Randomised Trials in Health, Education, and the Social Sciences: An introduction*. New York: Palgrave Macmillan.
- Torgesen, J.K., Wagner, R. K., Rashotte, C.A., Rose, E., Lindamood, P., Conway, T. , & Garvin, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology*, 91, 579-593.
- Totten, S., Sills, T., Digby, A., & Russ, P. (1991). *Cooperative learning: A guide to research*. New York: Garland
- Treiman, R. (1994). Use of consonant letter names in beginning spelling. *Developmental Psychology*, 30, 567–580.
- Treiman, R., & Kessler, B. (2003). The role of letter names in the acquisition of literacy. In R. Kail (Ed.), *Advances in Child Development and Behavior*, Vol. 31. San Diego: Academic Press.
- Tymms, P., Merrell, C., & Coe, R. (2008). Educational policies and randomised controlled trials. *The Psychology of Education Review*, 32, 3–7 & 26–29.

- UN (2015) *Transforming Our World: The 2030 Agenda for Sustainable Development*, UN, New York
- UNDP (2003) *Human Development Report 2003*, New York, United Nations Development Programme.
- UNESCO (2009). *Education for All global monitoring report, Overcoming inequality: why governance matters*. Oxford University Press, Oxford, UK
- UNESCO (2011) *Financing Education in SubSaharan Africa Meeting the Challenges of Expansion, Equity and Quality*. UNESCO Institute of Statistics, Montreal
- Vaidyanathan, A. and G. Nair (eds) (2001). *Elementary Education in Rural India: A Grassroots View*. New Delhi: Sage Publications.
- Van Boven, L, Epslet, N. (2003) The unpacking effect in evaluative judgements: When the whole is less than the sum of its parts, *Journal of Experimental Social Psychology*, 39: 263-269
- Van Steendam, E., Rijlaarsdam, G., Sercu, L., & Van den Bergh, H. (2010). The effect of instruction type and dyadic or individual emulation on the quality of higher-order peer feedback in EFL. *Learning and Instruction*, 20(4), 316-327.
- Van Teijlingen E et al (2001) The importance of conducting and reporting pilot studies: the example of the Scottish Births Survey. *Journal of Advanced Nursing*. 34, 3, 289-295.
- Van Zundert, M., Sluijsmans, D., & Van Merriënboer, J. (2010). Effective peer assessment processes: research findings and future directions. *Learning and Instruction*, 20(4), 270-279
- Varma-Nelson, P., and B. P. Coppola. 2004. Team Learning. In *The Chemists' Guide to Effective Teaching*, eds. N. Pienta, M. M. Cooper, and T. Greenbowe, 155-169. Upper Saddle River: Prentice Hall
- Venkatanarayana, M. (2004) Educational deprivation of children in Andhra Pradesh: levels and trends, disparities and associate factors, Working Paper 362, Centre for Development Studies. Available on line at: [www.cds.edu](http://www.cds.edu) (accessed August 2012).
- Venkatesan S, Holla MJ (2011). Validation of graded word list for testing spelling competencies in children with learning disabilities. *Indian Journal of Psychology*, 2(2): 73-82.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Cambridge, MA: MIT Press
- Vygotsky, L. S. (1987). Thinking and speech (N. Minick, Trans.). In R. W. Rieber & A. S. Carton (Eds.), *The Collected Works of L. S. Vygotsky: Vol. 1. Problems of general psychology*. New York: Plenum Press. (Original work published 1934)
- Vygotsky, L. S. (1994). *The problem of the cultural development of the child. The Vygotsky reader*. R. V. D. Veer and J. Valsiner. Cambridge, Blackwell: 57-72.
- Vygotsky, L. S. (1998). The Problem of Age. In *The Collected Works of L. S.*, Cambridge, Blackwell.

- Watson JM, Torgerson DJ. (2006) Increasing recruitment to randomized trials: a review of randomized controlled trials. *BMC Medical Research Methodology*; Vol 6, 34.
- Watson, J. (1998). *An investigation of the effects of phonics teaching on children's progress in reading and spelling*, University of St Andrews. PhD.
- Weaver, C. (1996). *Creating support for effective literacy education*. Portsmouth, NH: Heinemann.
- Welch, B. L. (1947). "The generalization of "Student's" problem when several different population variances are involved". *Biometrika*. 34 (1–2): 28–35
- Webb, N & Mastergeorge, A. (2003) Promoting effective helping in peer-directed groups *International Journal of Educational Research*, 39 (2003), 73–97
- Welman, J.C. & Kruger, S.J. (1999). *Research methodology for the business and administrative sciences*. Johannesburg: International Thompson.
- Wertsch, J. (1985). *Vygotsky and the Social Formation of Mind*. Cambridge, MA: Harvard University Press.
- Wheelock, A. (1994). *Alternatives to tracking and ability grouping*. New York, NY: R&L Education.
- White, G. & Ruther, M, Khan, J. (2016) *Educational Inequality in India: An analysis of gender differences in reading and mathematics*, New Delhi, IHDS
- White, H., & S. Sabarwal (2014). *Quasi-experimental Design and Methods, Methodological Briefs: Impact Evaluation 8*, UNICEF Office of Research, Florence.
- Whitehurst GJ, Lonigan CJ. (1998) Child development and emergent literacy. *Child Development*, 69(3): 848-72.
- Wigfield, A., & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology*, 89, 420–432
- Wigfield, A., Byrnes, J. P., & Eccles, J. S. (2006). Development during early and middle adolescence. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology*. Mahwah, NJ: Erlbaum.
- Willet, L. & Gardiner, A. (2009). Testing spelling: Exploring NAPLAN. Paper presented at Australian Literacy Educators Association Conference, Hobart, 9–12 July. URL: [http://www.englishliteracyconference.com.au/files/documents/hobart/conferencePapers/n\\_onRefereed/Willet-Gardiner.pdf](http://www.englishliteracyconference.com.au/files/documents/hobart/conferencePapers/n_onRefereed/Willet-Gardiner.pdf) (accessed 05 August 2016).
- World Bank (2003) *Making Services Work for Poor People: World Development Report 2004*, New York, World Bank/Oxford University Press.
- Wydell, T. & Butterworth, B. (1999). A case study of English–Japanese bilingual with monolingual dyslexia. *Cognition*, 70, 273–305.

Yamashita, J. (2007). The Relationship of Reading Attitudes Between L1 And L2: An Investigation of Adult EFL Learners in Japan, *TESOL Quarterly*, Vol. 41, No. 1, 81-105

Yates, L. (2004) *What does good education research look like?*, Berkshire: Open University Press.



## Appendices

### Family Background Questionnaire

#### Appendix A Family Background Questionnaire

Q1. Which language do you use in your household?

- i. Punjabi
- ii. Hindi
- iii. Urdu
- iv. English
- v. Other (please specify)

Q2. Please state which Caste you belong to

Q3. Please state the age of the child you're enrolling into the Phonics Programme  
(yrs/months, where possible)

Q4. Please state your household income per month

Q5. Please select the ownership of your house

- i. Family Owned
- ii. Renting
- iii. Private

Q6. What is the job role of the primary male carer in the household?

- i. Unemployed
- ii. Government Job
- iii. Private Company
- iv. Labour/Manual Work (i.e. agriculture)

Q7. What is the job role of the primary female carer in the household?

- i. Unemployed
- ii. Government Job
- iii. Private Company
- iv. Labour/Manual Work (i.e. agriculture)
- v. Housewife

Q8. Please tick up to the highest level of education the primary male carer has achieved

- i. No Education
- ii. Nursery
- iii. KG (lower/upper)
- iv. Up to Primary 5
- v. Up to Primary 8
- vi. College (vocational/academic qualification)
- vii. University (undergraduate)
- viii. University (postgraduate)
- ix. Professional Qualification (work based/centre based)

Q9. Please tick up to the highest level of education the primary male carer has achieved

- i. No Education
- ii. Nursery
- iii. KG (lower/upper)
- iv. Up to Primary 5
- v. Up to Primary 8
- vi. College (vocational/academic qualification)
- vii. University (undergraduate)
- viii. University (postgraduate)
- ix. Professional Qualification (work based/centre based)

Q10. Do you pay for after-school tuition?

Q11. Please state how much you pay for after-school tuition (total cost)

Q12. How many hours do you send you child to after-school tuition each week?

Q13. How many hours does your child spend at after-school English tuition?

## English Reading Aspiration Survey

### Appendix B – English Reading Aspiration Survey

#### Recreational Construct Questions

1. How do you feel about reading English books away from school?
2. How do you feel about reading English books with family at home?
3. How do you feel about reading English books for fun?
4. How do you feel when a friend gives you an English reading book?
5. How do you feel when you read new words from an English book?
6. How do you feel when you read English stories to your friends?
7. How do you feel about listening to your friends read English stories?
8. How do you feel about reading English books instead of playing outside?
9. How do you feel about seeing new English books at the market?
10. How do you feel about reading English books not related to school work?

#### Academic Construct Questions

1. How do you feel about reading English books in School?
2. How do you feel about reading English books with teachers at school?
3. How do you feel about reading English books for schoolwork during school time?
4. How do you feel when a teacher gives you a reading book during school time?
5. How do you feel about learning to read new words from an English book?
6. How do you feel when you have to read English books during class time?
7. How do you feel about listening to your friends read English books during school time?
8. How do you feel about reading English books instead of playing at break time?
9. How do you feel about seeing new English books when you're at school?
10. How do you feel about reading English books related to school work?

## Nfer A Test (Letter Matching)

### Appendix C –Nfer A Test (Letter Matching)

1	<b>a</b>	b	a	o	e	a	d	o	a
2	<b>s</b>	n	e	s	a	o	s	t	s
3	<b>on</b>	no	on	oh	ho	an	na	ha	ra
4	<b>da</b>	ab	ob	ba	da	da	ad	ab	da
5	<b>wau</b>	uam	wau	man	wan	wua	uwa	mau	
6	<b>dpb</b>	bpd	dpd	bpd	dpb	bpb	dqb	dpb	
7	<b>chaf</b>	cuaf	cuat	fano	canf	chfa	chaf		
8	<b>dogl</b>	bopl	dogl	bogl	dagr	dogl	dagl		
9	<b>swzc</b>	szwc	czsw	swzc	scwz	czws			
10	<b>laos</b>	soal	loos	laso	loas	laas	laos		

## Nfer Test B (Phoneme Identification) and Nfer Test C (2/3 Blending)

### Appendix D – Nfer Test B (Phoneme Identification) and Nfer Test C (2/3 Blending)

Nfer-Nelson Test B Sound value of letters:

m	u	d	l	o	g	j	o	b	v	e	t
d	i	g	r	o	d	c	o	t	p	a	f
h	o	p	w	a	g	k	i	t	y	e	:

Total .....

Nfer-Nelson Test C Blending 2/3 letter words:

in	at	up	on	an
dim	van	god	ham	sip
jet	nip	mop	led	dab
pen	cup	lit	ban	mix
pan	nap	fix	bud	sum
got	hid	ram	fit	tub
jug	tax	rug	web	

Total .....

## Miskin Nonsense Words

### Appendix E –Nfer Test C (Blending/Segmenting)

Ruth Miskin's Nonsense Word Test

1

m	a	s	d	t	i
n	p	g	o	c	k
u	b	f	e	h	sh
l	r	j	v	y	w
th	z	ch	ng	qu	x

2

feg	jes	vok
gop	ruch	dez
thob	cag	shug
wiss	miff	sleck

3

sprell	creld	splind	flond
blim	flut	smez	shrid
sprund	shrong	brost	flamp

4

spow	clirt	thronn	scight
droy	scray	troud	drair
weeg	grort	ploon	frarp

5

cleab	scrule	gart	hoimt
splafe	scry	chie	floke
grooring	shrowed	scurnly	clared

6

phantrite	strawble	frubehabe	doilble
snoiggal	wacespink	disclorping	threption
illarptacture	naightenfance	stobesourous	feanlissable

Letters should be read as sounds, not names. (Note: When the test is read as an oral essay assessment, seconds a word. Stop when 20 errors are made in one box.) Complete test as far as is possible to ascertain a child's knowledge and blending ability. Teacher has separate copy to record teacher's responses for detailed word-level test should be given in conjunction with a real word and comprehension test to ascertain reader's full skills at a context.

## Dictation Tests

### Appendix F – Dictation Test

1. A cat in a hat
2. A dog in a box
3. Bob is top of the pops
4. The shop is open
5. It is windy by the sea
6. The ball is red and green
7. How can it rain in the morning
8. The sun is high in the sky
9. Top cat is happy and so are his friends
10. The train is running late today and will take ten minutes extra

### Appendix G – Punjabi Dictation Test (translated)

1. Rani goes to school
2. Rani likes school
3. Rani has lots of friends
4. Rani and Deepika are best friends
5. She is in class eight and her class is big
6. She is the second tallest in her class after Deepika
7. Her school has over one thousand children
8. Rani's school is one of the largest in the district
9. Rani enjoys reading books because it makes her happy
10. Rani and her best friend Deepika like to play at the park after school

## Hindi Word Test

### Appendix H – Hindi Word Test (translated)

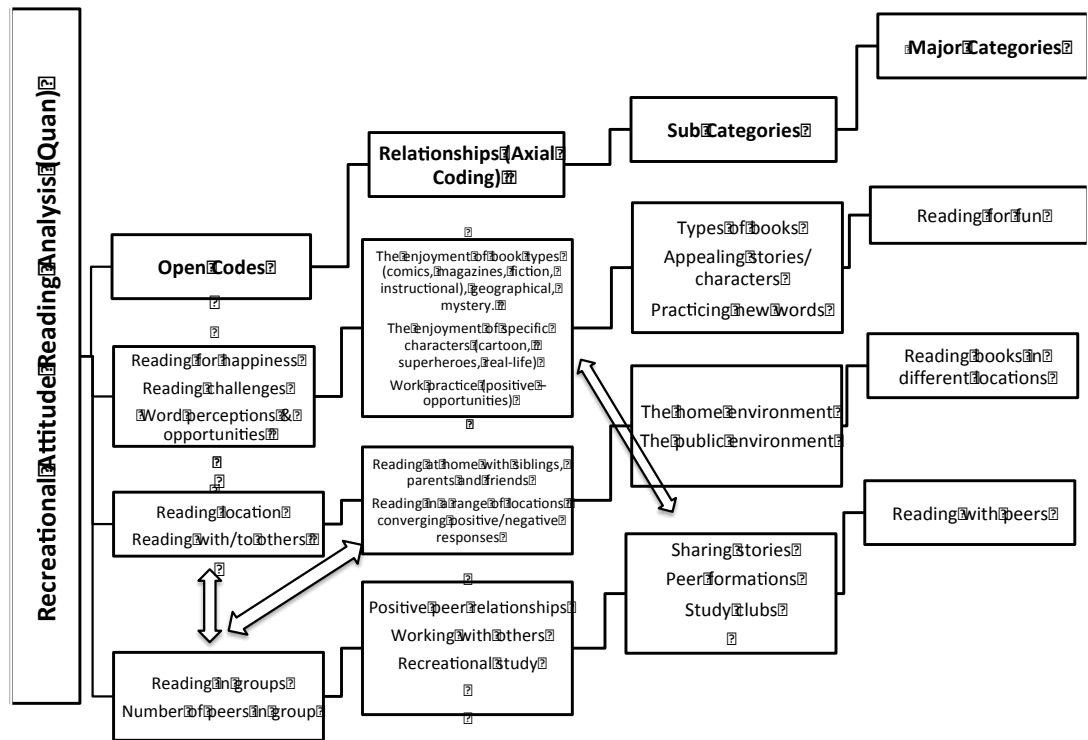
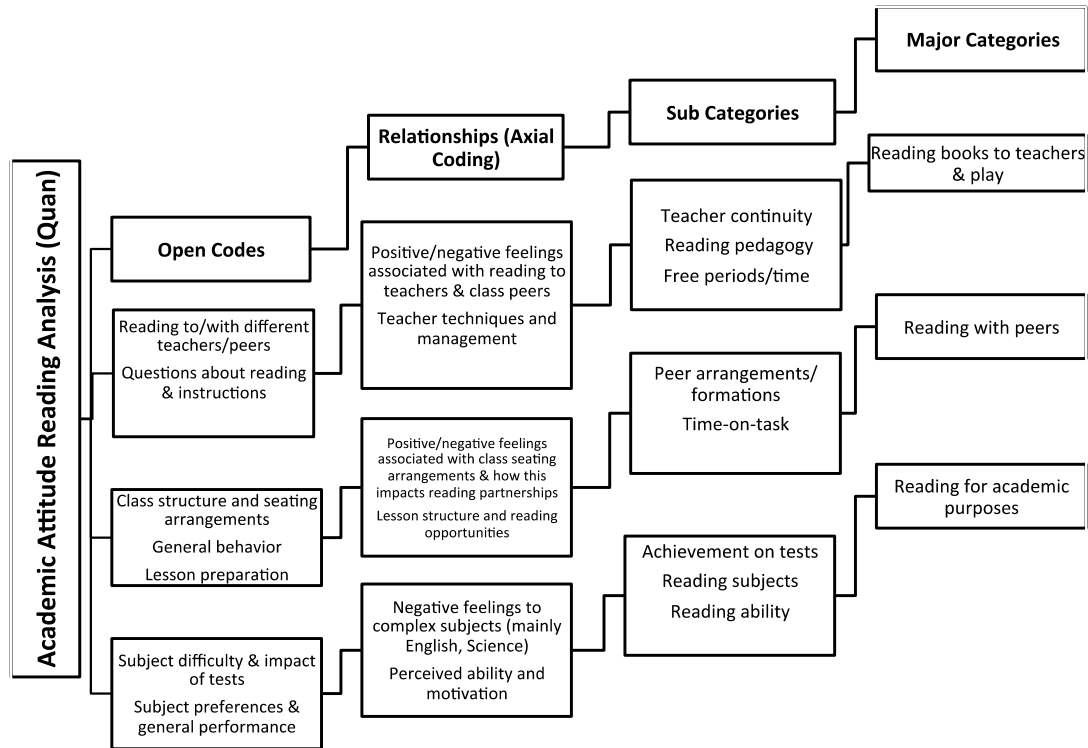
1. Boy
2. Bus
3. Hat
4. Girl
5. Day
6. Top
7. Car
8. Old
9. Cat
10. Hand
11. Star
12. Wind
13. Class
14. Book
15. Bike
16. Few
17. School
18. House
19. Tree
20. Shop
21. Park
22. Water
23. Garden
24. Market
25. Temple
26. Friend
27. Uncle
28. Station
29. Flower
30. Winter





# Recreational and Academic Reading Attitude Focus Group Analysis

Appendix H(i)



Key ⇔ = Intra-peer & cross-over coding relationships

## Letter of Enquiry

### Appendix I – Letter of Enquiry



20<sup>th</sup> September 2011

EG West Centre, Newcastle University  
ECLS , King George VI Building  
Newcastle Upon Tyne  
NE1 7RU  
United Kingdom

Dear Principal

I am interested in conducting a study into peer learning strategies to induce reading attainment in a second language (English). This research will, in part, fulfil the demands of Doctor of Philosophy (PhD) from the host institution, Newcastle University. My aim is to study the effects of a synthetic phonics programme as managed by pupils who study at your school. Your school has been identified as a host school because of its links to a number of local communities and socio-economical groups.

In general terms, the Alpine students will spend one hour each day for 4 months teaching children who don't access mainstream Indian government/private education. These children will typically not receive education, or who are enrolled in a local government school. Alpine students will collect data including whole word and individual letter sound tests before and after their teaching. Through portfolios, diaries and presentations Alpine students will gain academic credit from school formative assessments and CBSE examinations.

Along with Alpine Public School teachers, it is my aim to help prepare the peer-teachers with regular teaching workshops, seminars and practical assignments to help them mature into the role of peer coaching. The synthetic phonic method allows for games, songs and mixed media that develop pre-English reading through creative approaches to second language acquisition.

It is proposed that students from 8<sup>th</sup>- 10th Standard are suitably identified for the peer roles based on their foundational knowledge of English and likelihood of participation.

Please see below the proposed breakdown and key milestones of the intervention under the following headings:

#### Duration:

It is proposed the intervention last for 4 months. Your school will receive regular support and contact regarding the whole process throughout the year.

#### Procedure:

As highlighted above, three visits to your school and local area-

- Phase 2                      To train students and further generate interest from local families/children based on the above criteria. Conduct a pilot test.
  
- Phase 3                      To refresh student knowledge and begin intervention including pre/post tests for children and disseminating/de-briefing seminar for parents, children, teachers and peer-monitors on exit.

**Participation:**

Students from your school who attend 8<sup>th</sup>-10<sup>th</sup> Standard year groups are identified at this stage as appropriately aged for this type of study. Separate consent forms will be produced and transcribed locally for parents of the students involved.

Children from local areas and in particular those who don't receive education and those who are enrolled at a local government school will participate in the learning side of the study. Parents of these children will also need to provide consent allowing their children to participate in the study.

All consent forms will be translated into Punjabi and where necessary extra help will be needed to complete them. It is requested that extra provision be made to help parents/guardians who are unable to read in English or Punjabi so they are able to understand the implications of the study. No child, whether peer teacher/child learner will be selected for the programme of study until full consent is granted from the school and parent/guardian. All participation is voluntary and no names will be used in the disseminating process.

**Location:**

The teaching location will be dependent on the geographical accessibility of children who will be peer coaching/learning the method. At this stage it is proposed to ensure safety and security Gurudwara temples from each village will be targeted as possible locations where teaching can be delivered. A letter and potential meetings with Gurudwara temple management to allow passage of entry must be agreed before study commencement.

Using the information overleaf please can you read and sign this form, as this will be needed for ethical/project approval from Newcastle University. You may scan and send this version as a digital copy to the address provided below.

I'm extremely excited about this project and look forward to hearing from you regarding possible pathways. Please feel free to contact me anytime with any questions and suggestions you have before committing to the study.

Yours Sincerely

Christopher-John Counihan (PhD Candidate)

---

**Important Information: Please read**

I am willing to allow the proposed study to take part at Alpine Public School in conjunction with the following:

1. As mentioned in this letter, all registered Alpine Public School students will only feature with consent from parents/guardians.
2. Participating students will receive training from the researcher and Alpine teachers including pastoral care where necessary.
3. Study locations will be arranged and finalised before students of Alpine School undertake peer-teaching outside of school grounds.
4. Extra help and support will be provided from selected gatekeepers (nominated teachers/parents) to ensure safety and security at all times when teaching/learning.
5. All students will benefit educationally and will be free from psychological and physical harm.
6. The school will continually receive updates and support through seminars, email exchanges and telephone calls regarding any problems to do with the study.
7. The school and parents have a right to remove their children from the study during any stage of the research.
8. I have had the chance to ask questions about the study and have received information, support and guidance that is relevant to Alpine Public School.
9. No names of teachers, parents and children will be used in the disseminating process and all collected data will not breach any Data Protection Laws. All data will be secured in a safe and password protected file.

To the best of my knowledge, I have read this entire document and agree to the following proposal.

Signature  
Researcher

Signature  
Principal

\_\_\_\_\_

Printed Name  
Name

Printed

## Gurudwara Letter

### Appendix J – Gurudwara Letter

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

Dear Sir or Madam:

I am writing to you for your help with an educational study that will commence during the summer. My name is Christopher Counihan and I'm currently pursuing an PhD with the University of Newcastle, England. I'm extremely interested in developing new methods of teaching and learning for children who don't typically access mainstream school. Therefore, I have decided to develop a study that focuses on improving English reading. The aim of the study is to research the effects of a peer learning experiment on English reading attainment. The location of the study must be in a rural setting and involve local children. Therefore, the Gurudwara is seen as an ideal location.

The study will involve approximately 800 children and will last for one hour each day but not including Sundays and festival holidays.

I respectfully request your permission to use the Gurudwara as the primary location for the study for duration of 4 months between July - October. This request is based on the access opportunities for local children. No cost from either involved children or myself will be levied, as this is an experiment that aims to fulfill the Doctorate criteria as set by my University. I would be grateful if we could discuss my ambitious project and to hear your views.

I/We have had the opportunity to discuss the study in accordance with Gurudwara rules and Regional Boards. I/We permit access for the use of the Gurudwara for the aforementioned study as detailed above.

**Researcher**

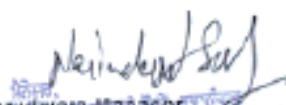
Christopher Counihan

Email [REDACTED]

Newcastle University

United Kingdom

**Gurudwara Senior Manager**

  
Gurudwara Manager

## Recruitment Letter

### Appendix K – Recruitment Letter



Dear Parent:

Allow me to introduce myself, my name is Chris Counihan and I am a Doctoral student at Newcastle University, UK. As a research student my main interests are to do with reading in a second language and in particular international education and development. As a result, I am writing to inform you of a new and exciting academic programme that will run for approximately 4 months next year in partnership with Alpine Public School (APS) and the Alpine School New Literacy Programme. The proposed programme will also fulfill the requirements of Doctor of Philosophy (PhD) through the host University.

We aim to investigate and measure the effects of peer teaching that will involve students from 8<sup>th</sup>-10<sup>th</sup> classes teaching local children using English letter sounds. We believe this method of teaching can significantly help those children who don't access mainstream education, or those who don't attend a low cost private school. We also believe that this programme is mutually beneficial for children beginning to read and those with acquired skills. Teaching will take place outside of school in the Gurudwara grounds, Bhawanigarh. Games, songs and puzzles will be used to induce pre-reading skills and maximize every child's potential whether through teaching or learning. There will be no formal teaching instruction from teachers or professionals; the aim of this study is to understand more about how children can teach.

A three week training programme will be arranged next year (dates to be confirmed) for those parents/students who are interested. Students from APS will learn basic English letter sounds and will be encouraged to keep a reflective diary to help them develop their own English and academic study skills. This project file can also be used to gain credit for those developing formative portfolios and evidence of continual learning. Ultimately, the entire programme will be managed by the research team who are teachers at the school and me. Please see the table below that illustrates the key dates and milestones for the proposed programme-

Month	Action Required
December 2011	1 <sup>st</sup> meeting with parents outlining project objectives
March 2012	2 <sup>nd</sup> meeting with parents outlining project objectives
June/July 2012-	2 week training of Alpine students using the peer and phonic methods
July-October 2012-	Testing Period and then teaching begins in communities with support from myself and APS teachers (outside school). 45mins-1hour per day after school.

At this stage we would like to explain more about the proposed programme. We are also conscious that some parents are unable to read English or Punjabi, therefore we would like to invite you to two presentations at Alpine School on the following dates and times to find out more and for you to ask questions-

- 20<sup>th</sup> December 2011 at 2.45pm
- 27<sup>th</sup> March 2012 at 2.45pm

Please choose a suitable day and let us know by completing the slip below and returning it to school. We welcome any questions you have and encourage you to participate in at least one of the presentations if you are interested for your child to take part. Both meetings will be conducted in Punjabi medium to allow for any questions or queries you may have.

We very much look forward to hearing from you and hope that you will consider this fantastic opportunity

Yours Faithfully

Christopher-John Counihan M. Ed

At this stage please can you indicate whether by placing a tick (✓) next to the relevant box indicating your response.

X- - - - -

**PLEASE TICK (✓) ONLY 1 BOX INDICATING YOUR RESPONSE**

- I can confirm my attendance on the 20<sup>th</sup> December 2011 at 2.45pm
- I can confirm my attendance on the 27<sup>th</sup> March 2012 at 2.45pm
- I would like to attend both meetings
- I am not interested in this programme and would like no further communication.

Parent Signature..... Date.....

Your Child's name:..... Class.....

X- - - - -



## Media Recruitment Poster

### Appendix L - Media Recruitment Poster

**ਕੀ ਤੁਸੀਂ ਮੁਫਤ ਅੰਗਰੇਜ਼ੀ ਸਿੱਖਣਾ ਚਾਹੁੰਦੇ ਹੋ**  
**ਕੀ ਤੁਸੀਂ ਵਿਦੇਸ਼ (ਇੰਗਲੈਂਡ) ਦੀ ਯੂਨੀਵਰਸਿਟੀ ਤੋਂ ਸਰਟੀਫਿਕੇਟ ਲੈਣਾ ਚਾਹੁੰਦੇ ਹੋ** 

**ਅਸੀਂ ਅੰਗਰੇਜ਼ੀ ਭਾਸ਼ਾ ਨੂੰ ਪ੍ਰਮਾਣਿਤ ਕਰਨ ਲਈ ਤੁਹਾਨੂੰ ਤਲਵਾਰ ਕਰ ਰਹੇ ਹਾਂ।**  
**ਇਹ ਪ੍ਰੋਗਰਾਮ ਇੰਝੇ ਕਾਲਜ ਯੂਨੀਵਰਸਿਟੀ (ਇੰਗਲੈਂਡ) ਵਾਲੇ ਅੰਗਰੇਜ਼ੀ ਭਾਸ਼ਾ**  
**ਨੂੰ ਪੜ੍ਹਨ ਅਤੇ ਲਿਖਣ ਲਈ ਬਿਲਕੁਲ ਮੁਫਤ ਕਰ ਦਿੱਤਾ ਗਿਆ ਹੈ।**  
**ਇਸ ਖੇਤਰ ਲਈ ਕਿਸੇ ਬਿਲਕੁਲ ਵੀ ਕੋਸ਼ੀ ਵੀ ਫੀਸ ਨਹੀਂ ਹੈ।**

ਕੀ ਤੁਹਾਡਾ ਬੱਚਾ ਕਿਸੇ ਵੀ ਸਕੂਲ ਵਿੱਚ ਨਹੀਂ ਪੜ੍ਹ ਰਿਹਾ ਜਾਂ ਸਰਕਾਰੀ ਸਕੂਲ ਵਿੱਚ ਪੜ੍ਹ ਰਿਹਾ ਹੈ ਅਤੇ ਕੀ ਉਸ ਦੀ ਉਮਰ 5 ਤੋਂ 14 ਸਾਲ ਦੇ ਵਿਚਕਾਰ ਹੈ?  
**ਜੇਕਰ ਹਾਂ ਤਾਂ ਤੁਹਾਡਾ ਬੱਚਾ ਇਹ ਕਲਾਸਾਂ ਬਿਲਕੁਲ ਮੁਫਤ ਨਗਾ ਸਕਦਾ ਹੈ।**  
 ਇਹ ਪ੍ਰੋਗਰਾਮ 23 ਜੁਲਾਈ 2012 ਤੋਂ ਸ਼ੁਰੂ ਹੋ ਕੇ 29 ਸਤੰਬਰ 2012 ਤੱਕ ਚੱਲੇਗਾ।  
 ਹਰੇਕ ਖੇਤਰਵਾਰ ਅਤੇ ਅਠੀਨ ਹੇ ਫੁਲਰੇ ਸ਼ਹੀਦਾਬ ਨੂੰ ਛੁੱਟੀ ਹੋਵੇਗੀ।  
 ਜਿਹੜੇ ਬੱਚੇ ਇਸ ਵਿੱਚ ਭਾਗ ਲੈਣਗੇ ਉਹਨਾਂ ਨੂੰ ਇੰਝੇ ਕਾਲਜ ਯੂਨੀਵਰਸਿਟੀ (ਇੰਗਲੈਂਡ) ਵੱਲੋਂ ਸਰਟੀਫਿਕੇਟ ਜਾਰੀ ਕੀਤਾ ਜਾਵੇਗਾ।

**ਮੁਫਤ ਕਿਤਾਬਾਂ ਅਤੇ ਟੈਪੋਗਰਾਮ ਲਾਭ ਵਜੋਂ ਹੀ ਵਿੱਤਾ ਜਾਵੇਗਾ ਅਤੇ ਕਿਸੇ ਵੀ ਕਲੀ ਵੀ ਭੇਜ ਨਹੀਂ ਹੈ।**  
 ਸਮਾਂ ਅਤੇ ਸਥਾ: ਹੁਦਦੁਆਰਾ ਲਾਇਫ ਸਾਫਟਵੇਅਰ ਨੈੱਟਵਰਕ ਵਿਖੇ ਦੁਬਿੰਗ 2:45 ਤੋਂ 3:45 ਤੱਕ

**ਇਹ ਪੀਰ ਕਲਾਸਾਂ ਅਤੇ ਰਜਿਸਟਰੇਸ਼ਨ ਬਿਲਕੁਲ ਮੁਫਤ ਹਨ**  
**ਆਪਣੇ ਬੱਚੇ ਦੀ ਰਜਿਸਟਰੇਸ਼ਨ ਕਰਵਾਉਣ ਲਈ ਹੇਠ ਲਿਖੇ ਪਤੇ ਜਾਂ ਫੋਨ ਨੰਬਰਾਂ 'ਤੇ ਸੰਪਰਕ ਕਰੋ:**

**ਐਲਪਾਈਨ ਪਬਲਿਕ ਸਕੂਲ**  
 ਕਾਕਰਾ ਰੋਡ, ਬਵਾਨੀਗੜ੍ਹ  
 ਫੋਨ: 01672-272470, 272670



## Letter of Consent and Debriefing Guide

### Appendix M

Dear Parent

I'm writing to you following our debriefing meetings held over the past couple of months. I would also like to take this opportunity to thank you for your attendance, asking relevant questions and making the whole experience enjoyable. Of course, I apologise for my limited Punjabi, a language I'm trying to pick up and learn, but I was very impressed and somewhat surprised with the level of English, which certainly helped me to understand more of your queries.

I wish to remind you that we've been extremely fortunate in gaining access to the Gurdwara, thus, we must respect the daily routine and take care not to disturb prayer and other religious activities. Any child who is displaying poor behavior will be removed from the programme and will be disqualified from any certificates.

The structure of each day will be as follows;

1.30pm	School Class finishes
2.00pm	Teaching, learning and educational games at Bhawanigarh Gurdwara
3.00pm	Back to school for tuition and more games - a bus service will be provided for those who require it

As we mentioned at the debriefing sessions we will need your consent and approval in line with both Alpine School and Newcastle University's ethical frameworks.

Please indicate with a tick (✓) in the boxes provided which applies to you:

- I have had the chance to ask questions regarding the study                      Yes                       No
- English is not my main language but I have been given support with all documents requiring my attention                      Yes                       No
- I am unable to read in English or Punjabi, but I have been given support with all documents requiring my attention                      Yes                       No
- I have had attended at least one of the debriefing sessions                      Yes                       No

You can confirm your child does not attend a private school	Yes <input type="checkbox"/>	No <input type="checkbox"/>
I require my child to be dropped home by the school bus service	Yes <input type="checkbox"/>	No <input type="checkbox"/>
I have had the opportunity to ask questions related to the study	Yes <input type="checkbox"/>	No <input type="checkbox"/>
I am also aware that this study will omit any names, links and associations in the final report.	Yes <input type="checkbox"/>	No <input type="checkbox"/>
I am aware of my rights to withdraw my child during any period of the programme	Yes <input type="checkbox"/>	No <input type="checkbox"/>

---

I am happy for my son/daughter to take part in this programme      Yes       No

Parent/Guardian Name.....

Signature.....

Child's Name.....

Child's Age (if known).....

Home Address (including village name and telephone contact details)

.....  
 .....  
 .....  
 .....

Emergency Contact Number (if different, please state)

.....  
 .....

Work Address (if applicable)

.....  
 .....  
 .....

## Study Debriefing Guide and Participation Information Sheet

20<sup>th</sup> December 2011 & 27<sup>th</sup> March 2012

---

### Notice:

This guide was used to give an oral explanation of the project outline and was particularly designed for direct responses from parents and interested parties. The school Principal requested that all verbal communications were to be made through Punjabi medium to allow for greater accuracy and understanding. A further request was made limiting the amount of written literature based on those parents who cannot read English or Punjabi. Therefore, consent forms and other literature were verbally explained by trained Punjabi teachers. Parents were told that they were welcomed into the school to ask further questions following each presentation. However, given the nature of the project and to ensure the protection of children during the study, it was outlined at each of these meetings that no child would be able to take part without parental/guardian consent.

---

### Focus:

To investigate the effects of peer learning and reading attainment utilising a synthetic phonic pedagogy as the medium of instruction.

### Why Peer Learning?

Peer methods have the capability of reaching audiences who are not set within the traditional educational parameters. Reasoning, partnership and in particular minimally invasive educative goals are generated through self-organisation. Thus, children will begin to demonstrate key academic competencies with no formal instructions. These include- problem solving, sharing their experiences, language and communication and actively taking part in their own learning through the guidance and support from their fellow peers. Because instruction is administered through a peer-teacher, learning becomes less formal and places the emphasis on group discussion and discovery. The synthetic phonic model advocates child creativity using actions and sounds to prepare children for reading. Children who teach in an informal way can help learning using child-like communication and specific instructions to articulate a pathway for learning.

### General Information

#### So who will be the peer teachers?

Peer monitors will be registered APS students and come from 8<sup>th</sup> - 10<sup>th</sup> classes and they will teach children English using a systematic synthetic phonic programme aimed at pre-readers.

#### What children will they teach?

They will teach children who either follow a basic education covered in a government school, or who don't receive any formal education. This will therefore take place outside of school and feature in the community.

**Where will the peer monitors teach?**

They will teach children in the local Gurudwara, where it is safe and in the public eye. No adult will participate with regards to teaching or offer learning guidance. However, there will be an adult (research on duty teacher) present to ensure safety and security.

**Some students have tuition after school, will they have enough time?**

Each lesson will only last for a maximum of 1 hour and will not go beyond this time limit.

**How long will the programme last for?**

Teaching will last for three months in accordance with the first stage of learning to read using the synthetic model. However there will be two weeks before and after the programme for tests.

**What are the tests? Do I need to buy revision books or prepare somehow?**

The tests are a series of spelling, reading and sound tests designed to help the research understand the effects of the peer intervention. These will be completed at the beginning and end of the teaching programme in a quiet area of the temple. It is important that no child feels any stress when conducting the tests, therefore if a child is finding a section particularly difficult then the test will stop.

(Show each test and demonstrate what needs to be completed for each)

**Some children from the rural villages can't read or write, how will this help?**

The programme is designed to target these children specifically. Children should be encouraged to participate even if they have had no prior education, or find it difficult. This programme is most suited for these children as it engineers the linguistic areas of their brains. So it is much better if children cannot read or write in English.

**How old must the children be?**

Any child participating will be typically between the ages of 6-42 years. However, as stated before we are willing to take children who are interested in learning English for perhaps the first time and is not younger than 5 on the first day of the programme (July 23<sup>rd</sup>). Sadly we cannot accept children as young as five due to health and safety and the rules as set out by the Gurudwara management.

#### **Information: Peer Teachers**

**I'm interested in becoming a peer teacher, but what exactly will I need to do?**

After receiving consent from parents you will participate in a number of exciting training workshops to help familiarise yourself with the teaching.

**What happens next?**

You will be trained how to teach English reading using the Jolly Phonics materials based in school. A dedicated team of APS teachers will prepare you before you start to teach.

**What is Jolly Phonics and how will I be taught this method?**

Jolly Phonics is an English learning programme aimed at young children who are beginning to learn to read. Recent research, including research at APS has shown that children are significantly better readers when exposed to this programme. The method is very simple as it highlights all of the English sounds typically used to form basic words, for example;

(Make time for brief demonstrations with the parents)

Say the following letter sounds really fast and see if you can hear the word-

/S/A/T/.....

The word is SAT, as in 'she SAT on the cat'

You will be taught all of the English letter sounds and techniques used to bring letter sounds together to form basic words like the example above.

(if time, demonstrate the actions and include tricky words- non-phonetic).

**Where will I be teaching and how much time will I need?**

In the Gurudwara grounds. The programme lasts for no longer than 1 hour per day over a 3 month period.

**When will I start teaching?**

The first teaching day will begin on July 23<sup>rd</sup> 2012. However, you will be shown the various techniques 2 weeks before and during the intervention.

**I've never taught before, why don't normal teachers go instead?**

Chances are you have taught before, but you haven't realised it. Think about the last time you corrected a friend in class, demonstrated a football skill on the playing field or talked about a favourite movie- this is teaching! Demonstration, discussion, group reflection and administration are very much what good teachers practice. Therefore, it is about creating opportunities for these skills to flourish, we believe peer-learning can help bridge the learning gap.

**Kids in my area don't like going to school, some of them don't like teachers- how can this help?**

Remember this isn't like being in a school environment, simply because you won't be wearing uniform (although this is your choice) and it will take place in the Gurudwara. There won't be any adults teaching and participation is down to the children alone who wish to come and have fun. The Jolly Phonics method relies on playing games, group work and singing songs, therefore it isn't like learning English at school.

**Will this affect my school work?**

In short, no! This will contribute to your continual professional development across all the academic disciplines. For instance, you might consider developing your school portfolios of evidence. APS school curriculum managers have agreed that through a diary exercise you are able to gain credit towards end of year grades. This is something you might consider if you want to develop your school portfolios, but it is not mandatory. Also, the research team are willing to accommodate any other ideas that you may have that can help towards your formative assessments. (See Heads for more discussion on this).

**What is in the portfolio of evidence?**

This you can keep and update throughout your academic life, not just at APS. Here contains everything you have achieved within your current education journey. So for this particular project, you will manage class lists, develop lesson plans using your own ideas, keep written diaries/blogs/personal web-spaces to document each day. As stated above, this can be used as evidence towards your end of year grades.

**Will I need to do an exam?**

There are no formal examinations; only the reading, spelling and sound tests as outlined previously.

**Will I get credit for this?**

After the three months you will receive a certificate of participation and teaching the first stage of the Jolly Phonics programme from Alpine Public School.

**General Study Setup and Design Brief**

**How many peer monitors are needed?**

As many as possible as we want to maximise this unique learning opportunity. We believe that at least 50 students should be considered.

## Focus Group Interview Guide

### Appendix N – Focus Group Questions

1. **Small activity icebreaker**
2. **General warm-up questions about school day and lessons attended**  
Which lessons have you attended today?
3. **Focused questions to do with reading**  
Describe the last time you read to make you feel happy – what did you read?  
Why?
4. **Recreational Directed Questions**
  1. Thinking about reading in the home environment, what makes reading fun for you? What do you read and why is it fun?
    - a. Probe – why do you think that? What specifically makes it fun?
  2. Describe your experiences of reading outside of school, what experiences do you have and where do you go?
    - a. Probe – describe more about the setting and how it makes you feel
  3. Tell me about your experiences of reading with your peers (friends) outside of school. What does this look like and describe what you do?
5. **Academic Directed Questions**
  1. Thinking about reading English books to teachers at school, what does this look like and how do you feel about doing it? How do you feel about reading at break time?
    - a. Probe – more on positive and negative feelings
  2. Do you read with your friends in school? What do you do and can you tell me what this looks like?
    - a. Probe – more on positive and negative feelings



## Extracts from Supported Study Guide

### Appendix 0 - Extracts from Supported Study Guide

#### What is Jolly Phonics?

Jolly Phonics is a system that was produced in the late 1980s to help young children recognise the letter sounds in English words. For example, let's look at the word CAT.

Now, let's look at each LETTER SOUND in the word C-A-T, it looks like this,

1<sup>st</sup> Letter sound- C

2<sup>nd</sup> Letter sound- A

3<sup>rd</sup> Letter sound- T

Using the Jolly Phonics method, we can find the sound of each letter and once we have them all we can blend them to make the word. Each individual sound is called a **phoneme**, we need to know how to sound each phoneme to sound the letters found in words. **Blending**, or to blend basically means putting all the sounds (**phonemes**) together to make a word. Children who might see an English word for the first time will be able to use this technique more comfortably.

Here is a table of all of the phoneme sounds we will use-

S	A	T	P	I	N	C	K	E	R
M	D	G	O	U	L	F	B		



#### Assignment 1

Can you identify how many sounds in each of these words? The first one is done for you as an example;

WORD	LETTER SOUNDS	HOW MANY SOUNDS
Sun	S-U-N	3
Top		
Pin		
Tap		
Pip		
Nap		
Pot		

## Other Sounds

Now that we know the individual letter sounds (**phonemes**) we can progress onto more complex sounds. Not all sounds found in English words are **singular**, or by themselves. For example, take the word **SHOP**. We might split the sounds in **singular** form like before, so, S-H-O-P. This will mean that there are 4 sounds needed to make the word. However, when we pronounce the word it will be incorrect. Therefore, we use what we call **DIGRAPHS** to help bring together **TWO** sounds to make one overall sound.

Let's look at the word **SHOP** again using this method.

1<sup>st</sup> Letter sound SH

2<sup>nd</sup> Letter Sound O

3<sup>rd</sup> Letter sound P

ai	Oa	ie	ee	or	ng	oo	oo	ch	sh
th	Th	qu	ou	oi	ue	er	ar		



### Assignment 2

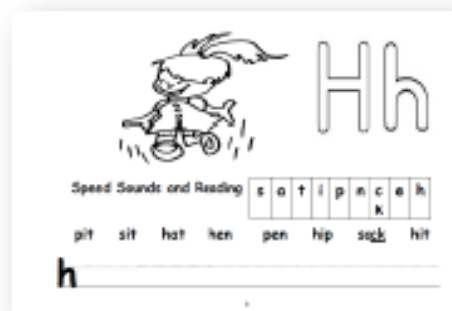
With these sounds we can make so many English words. In the space provided below, see how many words you can make, also how many sounds are there in each of the words you make. The first two lines are completed as an example. You can make any English words you like.

WORD	LETTER SOUNDS	HOW MANY SOUNDS?
Rain	r-ai-n	3
Flour	f-l-ou-r	4

## Why Peer Learning?

Peer learning techniques have been around a very long time. In Ancient Greece the famous philosopher Aristotle would get his students to teach each other to progress scientific thought.

In these days, peer learning techniques are used in schools around the world to help with many different academic subjects and puzzles.



Children will work in pairs or small groups to share their experiences and this has found to improve their overall academic experience. Peer learning for Jolly Phonics will mean that we will work in pairs and teach a maximum of 10 students. No teachers will be present and it is up to you as a student to help your fellow peers. We will have great fun teaching and playing loads of games too!

## Materials

We will use a range of materials to help us with our teaching. See the main materials we'll use to help make our lessons fun and rewarding-

## Finger Phonics Books

Each day we will turn the page to read a story (only a maximum of 3 sentences) to introduce a new letter sound. The stories help the children to recognise characters and future letter sounds.

## Flashcards

We use the flashcards in two ways, firstly to introduce the new letter sound at the beginning of the lesson and again when we repeat all other sounds.

## Perfect Partners Workbook

The perfect partner workbook is a chance for the children we're teaching to work in partners and practice the sounds themselves. Also, they will learn how to write the letter and practice it by writing it across the bottom dotted line. Finally, they'll use colour pencils to colour the picture and letters that they write!

## Jolly Phonics Blending Cards

The blending cards we will use in small groups to help the children to make small words. In the beginning we will use only 3 letter words like, s-a-t and p-i-n. But as we learn more sounds we can make **BIGGER** words like, s-t-a-m-p and c-l-o-ck.

## Magic Whiteboard

The Magic Whiteboard is a recent technological development that we are very fortunate to use for our teaching. It is a board that we can use time again like a normal blackboard. We will write using our magic pens and when we have finished we will wipe ready for next use. The Magic Whiteboard will be used when we get children to come up and practice writing letters and words, before they write them in their workbooks.

## Photos from the Intervention

