



USAID
FROM THE AMERICAN PEOPLE



Sindh Basic Education Program



Research Study on Early Grade Reading Innovations for Quality Education: Prospects for Scaling Up

SINDH CAPACITY DEVELOPMENT PROJECT

Research Study on
Early Grade Reading Innovations
for Quality Education: Prospects for Scaling Up

This research report has been prepared by a research team from the Aga Khan University Institute for Educational Development for USAID's Sindh Capacity Development Project.

The views expressed in the report are those of the research team, and do not necessarily represent those of USAID or AKU.

Research Team

Dr. Sajid Ali

Dr. Sadia Muzaffar Bhutta

Research Assistants & Associates

Ahmed Mughni, Sohail Ahmed, Nazeer Ahmed, Mazari Khan,
Ghulam Bhutto, Imran Qambrani, Nida Ijaz, Nasira Parveen,
Barkat Ali, Shahnaz Bhutto, Kiran Qasim Ali and Aniq Shah

Cover photographs: USAID-CMP/Amjad Jamal

Graphic Designing: Rizwanullah Khan



Contents

Page #s

LIST OF ABBREVIATIONS	VII
EXECUTIVE SUMMARY	IX
1. INTRODUCTION	1
Background	1
Current Study	1
Organization of the Report	2
2. LITERATURE REVIEW	3
Interventions for Reading Enhancement in Early Years	3
Reading Interventions and Student Learning Outcomes	3
Factors for Reading Enhancements	5
<i>Design of teacher training to improve practice</i>	5
<i>Home literacy environment</i>	6
<i>Textbook/reading material development as a goal of the intervention</i>	7
Intervention Projects: Issues of Implementation and Sustainability	9
Conclusion	10
3. METHODOLOGY	11
Research Design	11
Quantitative Approach	12
<i>Sample and sampling</i>	12
<i>Data collection process</i>	14
<i>Data collection tools</i>	14
<i>Administration of the data collection tool</i>	17
<i>Data analysis</i>	19
Qualitative Approach	20
<i>Research participants</i>	21
<i>Data collection methods</i>	21
<i>Data collection tools</i>	22
<i>Data collection procedures</i>	22
<i>Data analysis</i>	22
Limitations of the Study	23

4. FINDINGS: QUANTITATIVE ANALYSIS	25
PART-A: COHORT-1 (TLA MODEL)	25
Overall EGRA Results: Non-SRP vs SRP – Cohort-1	25
<i>EGRA scores: overall comparisons</i>	25
<i>EGRA scores: comparing performance on reading domains</i>	25
<i>EGRA scores: zero vs non-zero readers on reading domains</i>	28
Comparison by Language: Non-SRP vs SRP – Cohort-1	31
<i>EGRA-Sindhi – overall comparisons</i>	31
<i>EGRA-Sindhi – comparing performance on reading domains</i>	32
<i>EGRA-Sindhi – zero vs non-zero readers on reading domains</i>	34
<i>EGRA-Urdu – overall comparisons</i>	36
<i>EGRA-Urdu – comparing performance on reading domains</i>	36
<i>EGRA-Urdu – zero vs non-zero readers on reading domains</i>	38
Conclusion: Cohort-1	40
PART-B: COHORT-2 (GT MODEL).....	41
Overall EGRA Results: Cohort-2	41
<i>EGRA scores: overall comparisons</i>	41
<i>EGRA scores: comparing performance on reading domains</i>	41
<i>EGRA scores: zero vs non-zero readers on reading domains</i>	42
Comparison by Language: Non-SRP vs SRP – Cohort-2	44
<i>EGRA-Sindhi – overall comparisons</i>	44
<i>EGRA-Sindhi: comparing performance on reading domains</i>	44
<i>EGRA-Sindhi: zero vs non-zero readers on reading domains</i>	45
<i>EGRA-Urdu – overall comparisons</i>	46
<i>EGRA-Urdu: comparing performance on reading domains</i>	46
<i>EGRA-Urdu: zero vs non-zero readers on reading domains</i>	46
Classroom Practice: Non-SRP vs SRP – Cohort-2	47
<i>Overall quality of classroom practice</i>	47
<i>Quality of classroom practice by ERCOP items</i>	48
Home Literacy Environment: Non-SRP vs SRP – Cohort-2	51
Conclusion: Cohort-2	51
5. FINDINGS: QUALITATIVE ANALYSIS.....	53
EGRA Scores and Factors Explaining Students’ Performance	53
<i>Within-case analysis</i>	53
<i>Cross-case analysis</i>	54
TLA Vs GT Model: Motivational / Structural Bottlenecks	55
<i>Within-case analysis</i>	55
<i>Cross-case analysis</i>	59
Implementation of SRP: Role of Stakeholders	60
<i>Within-case analysis</i>	60
<i>Cross-case analysis</i>	63
Technology-based Assessment	64
<i>Within-case analysis</i>	64
<i>Cross-case analysis</i>	65

Sustainability and scaling up of SRP	65
<i>Within-case analysis</i>	65
<i>Cross-case analysis</i>	68
6. CONCLUSION AND RECOMMENDATIONS	71
Students Performance on EGRA	71
Factors Influencing EGRA Scores	72
<i>Teachers' classroom practices</i>	72
<i>Home literacy environment</i>	72
SRP Intervention and Factors Influencing EGRA Scores	72
<i>Content, SRP material and teachers' professional development</i>	72
<i>Physical infrastructure</i>	73
<i>Parental engagement</i>	73
TLA vs. GT Model Comparison	73
Motivational/Structural Bottlenecks	74
Role of Stakeholders in Implementing SRP Intervention	75
Technology-enabled Assessment	75
Prospects of Sustainability for the GT Model	75
<i>Alignment between SRP and STBB material</i>	75
<i>Timings for intervention</i>	76
<i>District / school level coordination</i>	76
<i>Duration and follow-up training</i>	76
<i>Parental involvement</i>	76
<i>Sensitization of senior political and bureaucratic leadership</i>	76
References	77
Appendix: EGRA - Detailed Results of Statistical Analysis	83

List of Tables

Table 1: Data (% correct) from the large scale study of reading skills at the end of grade 1 in 14 European languages.....	4
Table 2: Parents' qualification - comparison across group	14
Table 3: A brief overview of EGRA	15
Table 4: Reliability of ERCOP.....	17
Table 5: Qualitative approach: research participants	21
Table 6: Zero readers - comparison across group and progression over time	28
Table 7: Phonemic awareness –distribution of readers across group for pre- and post-test	30
Table 8: Non-word fluency – pre-test distribution of readers across group	30
Table 9: Expressive vocabulary – distribution of readers across group for pre- and post-test	30
Table 10: Oral reading fluency – distribution of readers across group for pre- and post-test	30
Table 11: Reading comprehension – distribution of readers across group for pre- and post-test	30
Table 12: Listening comprehension – distribution of readers across group for pre-and post-test	31
Table 13: Zero readers – comparison across group and progression over time on EGRA-Sindhi.....	35
Table 14: Zero readers – comparison across group and progression over time	39
Table 15: Students' performance in the various domains of EGRA	43
Table 16: Percentage of lessons by ranges of ERCOP score across group	47

Table 17: Quality of classroom practice – mean score on ERCOP pedagogy items	48
Table 18: Quality of classroom practice – mean score on ERCOP reading items	49
Table 19: Variety of teaching strategies	50
Table 20: Classroom management strategies	50
Table 21: Home literacy environment – comparison across group	51
Table 22: TLA vs. GT Model – comparing children’s performance across groups	74
Table 23: TLA vs GT model: quality, sustainability and impact	74

List of Figures

Figure 1: Overall research design	11
Figure 2: Quantitative approach: sample and sampling for Cohort-1	12
Figure 3: Quantitative approach: sample and sampling for Cohort-2	13
Figure 4: ERCOP pedagogy: example of an item	16
Figure 5: ERCOP reading: example of an item	16
Figure 6: Students’ performance on EGRA in cohort - 1: Difference across and progression within group	25
Figure 7: Students’ performance on EGRA in pre-test across group	26
Figure 8: Students’ performance on EGRA in post-test across group	26
Figure 9: Non-SRP group: progression from pre- to post-test	27
Figure 10: SRP group: progression from pre- to post-test.....	27
Figure 11: Overall comparisons on EGRA-Sindhi: Differences across and progression within group.....	32
Figure 12: Children’s performance on EGRA-Sindhi in pre-test across group	32
Figure 13: Students’ performance on EGRA-Sindhi in post-test across group.....	33
Figure 14: Non-SRP group: progression from pre- to post-test on EGRA-Sindhi.....	34
Figure 15: SRP group: progression from pre- to post-test on EGRA-Sindhi	34
Figure 16: Overall comparisons on EGRA-Urdu: Differences across and progression within group.....	36
Figure 17: Students’ performance on EGRA-Urdu in pre-test across group	36
Figure 18: Students’ performance on EGRA-Urdu in post-test across group	37
Figure 19: Non-SRP group: progression from pre- to post-test on EGRA-Urdu.....	37
Figure 20: The SRP group: progression from pre- to post-test on EGRA-Urdu.....	38
Figure 21: Students performance on EGRA in cohort-2 across group.....	41
Figure 22: Students’ performance on EGRA reading domains across group	41
Figure 23: Zero readers on EGRA in cohort-2 across group	42
Figure 24: Overall comparisons on EGRA-Sindhi on cohort-2.....	44
Figure 25: Students' performance on EGRA-Sindhi in cohort 2 across reading domains	45
Figure 26: Zero readers on EGRA-Sindhi in cohort2 across group	45
Figure 27: Overall comparisons on EGRA-Urdu in cohort 2	46
Figure 28: Students' performance on EGRA-Urdu in cohort-2 across reading domains	46
Figure 29: Zero readers on EGRA-Urdu in cohort-2 across group	47

LIST OF ABBREVIATIONS

ASER	Annual Status of Education Report
BoC	Bureau of Curriculum
CMP	Community Mobilization Program
ED-LINKS	Links to Learning Programme: Educational Support to Pakistan
EGRA	Early Grade Reading Assessment
ERCOP	Early Reading Classroom Observation Profile
ESRA	Education Sector Reform Assistance
FGD	Focus Group Discussion
GT	Guide Teachers
HM	Head Master
HT	Head Teacher
NCHD	National Commission for Human Development
PEAC	Provincial Education Assessment Center
PMIU	Project Management and Implementation Unit
RFP	Request for Proposal
SAP	Social Action Programme
SAT	School Achievement Test
SBEP	Sindh Basic Education Program
SCDP	Sindh Capacity Development Project
SELD	School Education and Literacy Department
SMC	School Management Committee
SPDC	Social Policy and Development Centre
SPSS	Statistical Package for Social Sciences
SRP	Sindh Reading Program
STBB	Sindh Text Book Board
STEP	Strengthening Teacher Education Pakistan
TLA	Teaching Learning Associates
TORs	Terms of Reference



Executive Summary

BACKGROUND

The Sindh Reading Programme (SRP), as part of Sindh Basic Education Programme (SBEP), was launched to develop the capacity of grade 1 and 2 teachers in government schools of Sindh to enhance students' reading skills. The research aims to study the efficacy of SRP intervention for improving early grade reading. It also explores the contribution of teaching practices and the home literacy environment towards students' reading skills. Furthermore, it aims to undertake an in-depth analysis of implementation of SRP interventions and to extract lessons for sustainability.

METHODOLOGY

A mixed-method research methodology was employed for a comprehensive analysis of both the learning outcomes and the processes of intervention. The study investigates the two training approaches employed by the SRP (TLA - Teaching Learning Associates and GT - Guide Teachers) to determine their relative advantages, which are referred to as Cohort-1 and Cohort-2, respectively. In order to enhance the robustness of findings, the SRP schools were compared against the non-SRP schools. Of the eight target districts, five were strategically selected to represent a variety of geographical locations, spread and medium of instruction. The results are based on 1,722 grade-3 children's EGRA tests from 104 schools. The EGRA test encompasses six reading domains: *phonemic awareness, non-word fluency, expressive vocabulary, oral reading fluency, reading comprehension, and listening comprehension*. Teaching practices of 104 grade 1 and 2 teachers were observed and fifty-six parents were surveyed for the home literacy environment. Focus Group Discussions and Interviews

from 107 stakeholders representing a broad spectrum of implementers were also carried out.

FINDINGS

EGRA Scores

1. *The SRP schools showed better performance on EGRA tests as compared to non-SRP schools across cohorts. On average SRP children scored 6% higher than non-SRP in Cohort-1; and 5% higher in Cohort-2. Despite small magnitude, the difference was significant.*
2. *The SRP intervention has managed to significantly reduce the number of zero-readers as compared to non-SRP across cohorts. However, only a fraction have qualified to be considered fluent readers in most of the domains. In Cohort 1, SRP intervention has influenced a significant reduction in the number of zero-readers in phonemic awareness (17%), non-word fluency (14%), oral reading fluency (14%), and reading comprehension (16%). While in Cohort 2, the number of zero-readers was reduced in phonemic awareness (8%), expressive vocabulary (7%) and oral reading fluency (7%).*
3. *The students have performed better in oral reading as compared to reading comprehension for both SRP and non-SRP schools across cohorts. This means that students are able to read the given text but found it difficult to comprehend.*
4. *The students' performance on phonemic awareness stayed at the bottom end for both SRP and non-SRP schools across cohorts, indicating that students found it challenging to identify phonemes in the given words.*

5. *In Sindhi medium schools, the SRP schools have shown significantly better performance on EGRA tests as compared to tests as compared to non-SRP schools across cohorts.* On average, SRP children have scored 9% higher than non-SRP both in Cohort-1 and Cohort-2. *The SRP intervention has managed to significantly reduce the number of zero-readers as compared to non-SRP across cohorts.* In Cohort 1, SRP intervention has influenced a significant reduction in the number of zero-readers in phonemic awareness (21%), non-word fluency (17%), oral reading fluency (19%), and reading comprehension (22%). In Cohort-2, the number of zero-readers was reduced in all 6 reading domains – phonemic awareness (8%); expressive vocabulary (6%); oral reading (13%); reading comprehension (13%) and listening comprehension (9%).
6. *In Urdu medium schools, there was no difference between the performance of SRP and Non-SRP schools in the EGRA tests. The SRP intervention has managed to significantly reduce the number of zero-readers in phonemic awareness (14%) in Cohort-1.* While in Cohort-2, there was no significant difference in any reading domain. Exploring the factors which explains difference in students' performance across language was beyond the scope of this study. However, some speculations can be made to explain this difference. Perhaps, the sample drawn for Urdu EGRA was more heterogeneous in terms of languages. In other words, majority of children in Urdu-medium schools have mother-tongue which is different from Urdu. Therefore, they may only get an opportunity to practice Urdu in schools. On the other hand, Sample drawn for Sindhi EGRA comes from homogeneous group in terms of language. Majority of them speak the same language they were tested in.

Factors Influencing EGRA Scores

7. *SRP teachers have shown significantly better teaching practices than their non-SRP counterparts.*

The SRP teachers have shown better skills in generic teaching (e.g. questioning, active involvement, teacher-child interactions) and reading instructions (e.g. phonemic awareness, reading aloud, silent reading). However, these teaching practices have remained 'mediocre' (1.51-2.50 on 3-pt scale) for SRP schools and 'weak' (<1.5 on 3-pt scale) for non-SRP.

8. *The parents of SRP children tend to provide more literacy enhancement opportunities for their children at home.* However, the difference was not big enough to be significant and needs further investigation with a bigger sample.
9. *The teachers found the SRP material attractive, however it was not considered complementary to their regular textbooks.* Therefore introducing this material was considered additional burden by the teachers.
10. *Teachers found the phonemic method helpful in improving students' reading, but were reluctant to continue using this method* as it did not align with their regular pedagogical practice (hijje method). Teachers feel that it needs more time, effort and academic support to implement this method to improve learning outcomes.

Implementation Bottlenecks

11. *There was no consensus in terms of efficacy of the two models - TLA and GT.* The SRP officials preferred TLA as more intensive and of higher quality; on the contrary the government officials preferred GT for its better prospects of sustainability due to the involvement of government teachers.
12. Common lessons learnt from the two models highlighted: *the weak coordination amongst SRP and government at district level; the shorter training duration; the lack of synchronization between SRP training and academic sessions; the insensitivity of the SRP material for multi-grade and large classes; the frail monitoring by the government officials; and lack of motivation and ownership at district and taluka levels.*

Role of Stakeholders

13. *Government stakeholders could not effectively play their respective roles in implementation of SRP at classroom, school and taluka levels in a coherent manner.* At classroom level, teachers felt improvement in their teaching skills through SRP training; however, only the self-motivated ones showed willingness to continue teaching through phonemic method after the project ends. At school level, head teachers were not engaged centrally in the schools and therefore they remained at the margins of intervention. At taluka level, the TEO did not feel involved and responsible in implementing and monitoring of the process.

Technology Based Assessment

14. *The technology was used in assessment, but not in enhancing teaching skills.* Guide teachers learnt to use tablets for assessment, though some of them struggled with this technology. They also complained about the lack of tablets. Overall, the government officials felt positive about technology-based assessment.

RECOMMENDATIONS TO ENHANCE PROSPECTS FOR SUSTAINABILITY

The EGRA scores revealed that the students of SRP schools have shown better reading performance and SRP teachers have exhibited better classroom practices as compared to non-SRP schools, though the magnitude of the difference is small. This shows that SRP interventions have potential for enhancing students' reading scores. However, for these interventions to have greater chances of sustainability, they need to:

1. *Align* SRP material with Sindhi and Urdu textbooks of STBB and also include writing in addition to reading focus only.
2. *Increase* the duration of the intervention to at least 4 to 5 years where the intervention schools should remain connected and followed up with classroom support and continuous professional development support.

3. *Initiate* the training of teachers at the beginning of the academic session of the government schools.
4. *Improve* district and school level coordination and accountability mechanisms to get all the relevant stakeholders including head teachers involved.
5. *Involve* parents in their children's reading enhancement.
6. *Sensitize* senior political and bureaucratic leadership on importance of specialized reading interventions.

CONCLUSION

Arguably, SRP intervention focuses heavily on improving reading through phonemic understanding, yet it remains one of the weakest areas of EGRA performance. Interestingly, students have performed better in the other five domains regardless of groups, cohorts and medium of instruction. The further questions that this research indicates are: *what is the role of phonemic awareness in improving students' reading skills? Can students read an unknown text and comprehend it without having adequate skills in decoding phonemes? What sort of language teaching approaches should be used by the teachers and what type of material should accompany them?*

1

Introduction

BACKGROUND

The Sindh Basic Education Program (SBEP) through its interventions aims to improve different aspects of primary and secondary education in the selected districts of Sindh. The areas of intervention include: school construction, policy support, improvements in early grade reading, community mobilization and capacity building. The Project Management Implementation Unit (PMIU) supports school construction, consolidation and up-gradation. The Sindh Reading Program (SRP) is intended to develop the capacity of primary teachers for enhancement of reading skills at grade 1 and 2. Training and resources were provided by SRP for improvement in early grade reading in the selected government schools in eight target districts of the province. The program aims to impact 400,000 students and provides professional development opportunities for up to 15,000 government teachers in order to improve their capacities to teach reading effectively.

The Sindh Community Mobilization Program (CMP) aims to enhance community mobilization for improving girls' enrolment and retention. The Sindh Capacity Development Project (SCDP) works to build overall capacity of the PMIU and the School Education and Literacy Department (SELD) of the Government of Sindh.

So far, the SRP has implemented two models of reading interventions including the Teaching Learning Associate (TLA) model and the Guide Teachers training (GT) model. The former is referred to as cohort 1 while the latter is identified as cohort 2. For the TLA model private staff were hired to train and support reading

teachers. On the other hand, the GT model recruited government teachers as facilitators for the project. For both projects selected teachers from eight districts were trained for three to six days in using strategies to improve reading in the early grades. They were then provided with classroom support for almost six months during which time the schools were visited by the facilitators (TLA or GT) once or twice a month. Informal assessment of children was part of the follow-up visits.

One of the mandates of the SCDP is to generate knowledge around various interventions being introduced under SBEP to guide PMIU and SELD towards sustaining these interventions, learning lessons through best practices and institutionalize them. Thus, a research study was commissioned by the SCDP to study the interventions by the SRP for improving early grade reading abilities of grade 1 and 2 students of government schools in selected districts of Sindh.

CURRENT STUDY

The Terms of Reference (TORs) advertized with the Request for Proposals (RFP) suggested that the main aim of this research is to ascertain if the interventions introduced by the SRP made any difference in teachers' pedagogies and students' reading abilities, and if so, what factors and processes explain that difference (or non-difference). Furthermore, it aimed to undertake an in-depth analysis of various aspects of SRP interventions towards enhancing reading skills and to extract lessons that can improve the intervention and gather policy lessons for scaling up and institutionalization of the interventions.

The RFP outlined the following specific questions that this research should explore:

1. What explains the variation among the students learning outcomes (zero scorers vs. readers)?

Explanation: Several factors could affect the learning outcomes such as pedagogic materials, teaching content, quality of instruction, physical infrastructure, parental engagement, etc.

2. How has the intervention addressed the factors mentioned above?

3. What are the lessons learned from the TLA model?

4. To what extent are the objectives of the new cluster-based GT model being met? What are the key motivational or structural bottlenecks and how can they be addressed?

5. What are the roles of various actors, including head teachers, teachers of target grades/classes, parents (SMCs)¹, government, etc. in implementing the interventions of the SRP?

6. How successful and sustainable is technology-enabled assessment in target schools?²

7. What are the prospects for the sustainability of the cluster-based GT model for replication and scaling up?

A cursory analysis of these questions indicated that the current research had certain predispositions. For example, it was hoped that the intervention schools ought to show positive change in reading abilities of children. Thus, the factors that play a more effective role needed further exploration to learn lessons. Also, the viability of the two models of intervention (TLA and GT) was to be explored to suggest a better model to the government for sustaining and scaling-up. The role of various actors in this regard required further investigation.

However, the history of various projects in Sindh (e.g. SAP, EDLINKS, ESRA, STEP) shows that the projects may not achieve significant effects in terms of student learning achievements. And even if an intervention shows positive outcomes during the life of the project the effect fizzles out quickly as soon as the project is phased out. Various achievement surveys (e.g. ASER, PEAC, SAT) have shown very weak achievements of students indicating very little effect, whatever the

interventions by whichever project in Sindh in general. With that backdrop a more comprehensive approach was adopted for the current study to not only explore the differences in student learning outcomes across non-SRP and SRP groups, but also to investigate various factors which may have contributed to these outcomes. The results would inform reading related practices and policies to bring improvement in implementing relevant models in future. The current project focused only on SRP intervention (as suggested by the RFP) which limits the researchers to explore any other good practice (if that exists at all) in improving early grade reading in Sindh. That said, a component of the study (experimental design) did help to compare the outcome of interventions against standard practice, which was further complemented by the qualitative data to explore processes.

ORGANIZATION OF THE REPORT

The report is organized in six chapters including: introduction to discuss the background of the SRP intervention and the rationale of the current study; literature review to highlight existing literacy and reading interventions from around the globe and factors which contribute to the student learning outcome; methodology in order to bring to light research methods which have been used to collect and analyze data; findings of quantitative analysis; findings of qualitative analysis; and conclusion and recommendations.

2

Literature Review

INTERVENTIONS FOR READING ENHANCEMENT IN EARLY YEARS

It is essential to learn how to read to be literate. In order to achieve high proficiency in literacy, children must effectively learn to read around the age of five or six years (Abadzi, 2017). Neuroscience research indicates that the ideal time to develop basic reading skills is in early childhood (Abadzi, 2017). Moreover, research shows that children who do not do well in reading and fail to improve by the end of their first grade are more likely to attain lower competency in other subjects throughout their academics life (McIntosh, Horner, Chard, Boland & Good, 2006). According to Mahar & Richdale (2008) almost half the children who find reading difficult at the age of seven will have difficulty with reading at the age of thirteen too. Therefore, if a child does not start learning to read around 1st grade, ‘the negative affects cascade’ limiting the child’s future progress in education (Abadzi, 2017, p. 8).

Early grade reading interventions are a solution to countering the problems that may arise as a result of failure to acquire reading skills at an appropriate age. An early start in reading is not only crucial in predicting a lifetime of literacy experience (Cunningham, 1997), but a persistent difficulty in reading can also negatively impact a child’s self-esteem and reduces their later employment prospects (Snowling, Muter & Carroll, 2007). Apprehending the potential problems that the failure to acquire reading skills can cause, numerous reading intervention programs have been designed and implemented around the world. These programs play a critical role in the efforts to eradicate reading and school failure (Slavin, Karweit, & Wasik, 2001), enhance children’s

scholastic development and contribute to reducing the need for later remedial actions (Reynolds, 1998).

Whilst most of the interventions (e.g. Reading Recovery, Success for All, READ) in developed countries aim at assisting at-risk students and struggling readers, the interventions (e.g. EGR) in developing countries aim to develop basic reading skills to increase the literacy rate in the long run. However, the effects of these interventions in terms of their alignment with the needs and demands of the education system to deliver a long term change is yet to be seen.

READING INTERVENTIONS AND STUDENT LEARNING OUTCOMES

The central emphasis of reading interventions ever since the National Reading Panel’s report has been on developing phonological awareness and using systematic phonics instructions to deliver reading lessons.

The effectiveness of phonological awareness in the acquisition of reading skills has been an ongoing debate since the publication of Bruce’s (1964) seminal publication ‘The analysis of word sounds by young children’. However, the insistence to provide phonological awareness to children to improve their reading skills is a recent phenomenon. There are numerous studies (for example; Bus & Van Ijzendoorn, 1999; Ehri et al., 2001; National Institute of Child Health and Human Development [NICHD], 2000a, 2000b; Troia, 1999; 2004) that established the role of phonological awareness in improved word decoding and enhanced reading skills confirming that the knowledge of grapheme-phoneme correspondence is

directly related to reading acquisition (Backman, Bruck, Hebert & Seidenberg, 1984; Manis & Morrison, 1985). The most comprehensive of these studies has been the report presented by the National Reading Panel (NRP) convened by the US Congress after the promulgation of “No Child Left Behind Act” in 2001.

The NRP conducted the most rigorous review of studies relevant to the process of reading acquisition and teaching ever undertaken, providing unequivocal evidence that the majority of children could learn to read if teachers were given necessary training to implement scientifically validated and effective instructions. The NRP analyzed research in the categories of: alphabets (phonemic awareness and phonics), fluency, comprehension, teacher education, technology and methodology (NICHD, 2000b).

In an article, Kimberly Kolba (2005) argues that reports produced by the National Academy of Sciences and the National Reading Panel evidently suggest that a comprehensive, scientifically based approach to reading instruction is necessary if all children are to learn to read efficiently. Based on the findings, it is essential that all reading programs must include systematic and direct instructions in phonemic awareness, phonics, reading fluency, vocabulary development, and comprehension strategies.

It is important to note that reading pseudo words are also considered as reading as these words assess the ability to pronounce unknown written words (Ehri, 2001). Therefore, all research initiatives that aim to gauge the efficacy of a reading intervention program should include the ability of children to read pseudo words (or non-words) as part of any assessment.

Synthetic phonics programs are also useful in teaching children systematically and sequentially the correspondences between graphemes and phonemes of the language to decode and blend unfamiliar words by sounding out the letters (Ehri et al., 2001). The use of systematic phonics instructions for reading proficiency is not limited to English. The research indicates that to be confident readers, children need to be able to decode print fluently which depends on their

awareness of phoneme and knowledge of letters (Catts, Fey, Zhang & Tomblin, 1999; Muter, Hulme, Snowling & Stevenson, 2004). Since in alphabetical languages, letters in print denote phonemes in spoken words, an awareness of phonemes in speech is essential to learn to read an alphabetic script (Lieberman, Shankweiler, Fischer & Carter, 1974; Savin, 1972).

Usha Goswami (2005) conducted a large-scale and carefully controlled cross-language reading comparison. For this study, only those schools were chosen where all children were experiencing phoneme-level ‘phonics’ teaching. Goswami (2005) notes that the children who were acquiring reading in languages with consistent spelling systems (Greek, Finnish, German, Italian, Spanish) were close to the ceiling in both word and non-word reading by the middle of first grade, irrespective of age. The results are presented in table 1. Paradoxically, English speaking children performed considerably poorly because of the language’s phonological complexity.

Table 1: Data (% correct) from the large scale study of reading skills at the end of grade 1 in 14 European languages

Language	Familiar real words	Pseudo-words
Greek	98	92
Finnish	98	95
German	98	94
Austrian German	97	92
Italian	95	89
Spanish	95	89
Swedish	95	88
Dutch	95	82
Icelandic	94	86
Norwegian	92	91
French	79	85
Portuguese	73	77
Danish	71	54
Scottish English	34	29

Source: Adapted from Seymour et al.(2003), cited in Goswami (2005)

Further support for the phonics instructions in a language other than English has been provided by Raynor et al. (2002) who carried out an experiment in which the investigators trained two groups of English-speaking college students to read Arabic. One group learnt the phonemes associated with individual Arabic letters (the phonics approach), while the other group learnt entire words associated with certain strings of Arabic letters (whole-word). Both groups were then made to read new set of words constructed from the original letters. In general, the students from the group who were trained to read phonetically performed better.

The results from these studies are of particular interest to us as Urdu and Sindhi are highly phonetic languages and are closely related to Arabic in their construction. Furthermore, Ziegler and Goswami (2005, cited in Tibi & Kirby, 2018) concluded that reading is acquired more readily in transparent orthographies (in which grapheme-phoneme correspondence is simple, e.g., German or Finnish) than in more opaque orthographies (e.g., English or French).

FACTORS FOR READING ENHANCEMENTS

This section discusses relevant factors that may impact students' reading skills.

Design of Teacher Training to Improve Practice

Since teachers are at the heart of all teaching contexts, it is imperative that their preparation, in terms of the resources and training available to them, must be taken into account before the outcomes of an intervention program could be measured.

Successful interventions largely depend on the knowledge and aptitudes of classroom teachers and school leaders (Meiers, 2013). In his review of five effective intervention programs, Pikulski (1994) acknowledged that professionally prepared accomplished teachers are the mainstay of successful early intervention programs. He suggested that in addition to provide initial training to the teachers, continuous professional support should also be available to the teachers particularly in the first year of intervention.

Teachers' preparation and training in teaching pupils and helping with their reading skills is essential and one of the most important determining factors in the success or failure of any intervention. Research into educational practice offers ample evidence on the positive influence of a well-trained teacher on student reading achievements (Darling-Hammond, 2004; 2006; Rowan, Correnti & Miller, 2002; Sanders & Horn, 1994), and recognizes teacher experience to be a statistically significant predictor of students' achievements (Rowan et al., 2002). Fielding-Barnsley (2010) believes that teachers ought to have an explicit knowledge and skills involved in the process of word decoding if they are to assist beginner readers to gain these skills. Furthermore, Mahar and Richdale (2008) assert that teachers must possess proficient knowledge of the structure of oral and written language to be able to deliver phonics-based instructions.

One of the most effective reading interventions for at-risk students, Reading Recovery – that provides one-to-one tutoring to struggling readers – owes its success to the quality of teaching. To begin with, it employs highly skilled, specially trained teachers to provide the intervention (Shanahan, 1995). "Selection criteria require that teachers have at least three years of teaching experience, including teaching in primary grades" (Groom et al., 1991, p. 21). Additionally, these teachers undergo training for approximately 75 hours spread over one year. Despite such stringent training protocols, Reading Recovery teacher training model is not believed to be powerful enough or appropriate to influence class room practices.³

This raises the question that if all interventions have research based design of training; do all the teachers trained for specific interventions actually carry out teaching the way they are expected or not?

Kennedy (1999) presents an explanation for the discrepancy between what teachers are trained and what they deliver by arguing that teachers, through their own childhood experience of teaching, are likely to teach in the way they themselves were taught as

children. Moreover, Clark-Chiarelli and Louge (2016) maintain that typically teachers struggle to retain and put into practice new knowledge gained from trainings. It is also common for teachers to only partially follow instructional guides (RTI International 2011b). Given these facts, one of the most crucial questions that need to be addressed before any intervention is executed is: what kind of training could help teachers re-conceptualize and reorient their teaching practice and how to ensure that this re-conceptualization and re-orientation lasts long enough to produce significant positive impact?

There are no simple answers on the training of teachers as literature does not provide unanimous parameters for the training design of teachers selected to be a part of any intervention. However, it is understood that the training programs that stretch over several hours or days with limited follow-up activities are most likely to succeed only with those teachers whose beliefs match with the assumptions inherent in the innovation (Richardson & Placier, 2001). Stallings and Krasavage (1986) assert that innovative practices can only be maintained until teachers and students remain interested and excited about their own learning. Consequently, a good staff development program can create excitement about learning to learn. Stallings and Krasavage (1986) support their claim through an in-depth analysis of a four-year study of a popular staff development program, which trained teachers in a structured approach to instruction. They found that in the third year teachers implemented the intended behaviors much less often than they had in the first two years (Richardson & Placier, 2001).

Teachers' habits also act as a barrier and prevent sticking to innovations as rather than working to develop new skills/strategies, it is simply easier for teachers to continue teaching in the same fashion (Greenberg & Baron, 2000). Generally, people derive a sense of security from doing things in familiar ways; disrupting teachers' well-established professional and instructional patterns could result in a fear of the unknown (Fullan, 2001; Greenberg & Baron, 2000).

Nevertheless, this can be countered by effective leadership (Tschannen-Moran et al., 1998) and coaching as well as continual feedback (Graham & Kelly, 2018).

Garet et al. (2001), in a survey of 1027 teachers, found that those teachers who attended professional development courses in line with their other professional experiences, aligns with standards and assessments and fosters professional communication—are more likely to change their practices. Their results further indicated that sustained and intensive professional development programs are more likely to have an impact as compared to the ones last for shorter duration. Several studies (for example; Shields, Marsh & Adelman, 1998; Weiss, Montgomery, Ridgway & Bond 1998) confirm the role of intensity and duration of professional development in determining the extent of change among teachers.

Home Literacy Environment

Well executed reading instructions alone cannot guarantee improved reading skills. There are several other factors that should be considered. One of the factors is parental engagement.

Parents and caregivers play a crucial role in their children's literacy development and social advancement (Adams, Frampton, Gilmore & Morris, 2010). The premise that language development occurs long before a child utters his or her first words (Bruner, 1978) underlines the role that parents or home environment plays in developing a child's language and literacy acquisition. Children take their reading seriously when they are motivated and praised for reading well at home (Cook-Cottone, 2004; Dodici & Pertson, 2003; Rashid, 2005); exposure to books at home is likely to contribute to children's better reading performance (Sénéchal & LeFevre, 2002).

Parental reports obtained by Burns and Collins (1987) indicate that gifted kindergartners, who were early readers, had more exposure at home to discussions of letter-sound correspondences, letter names and word identification experiences than did those children who were not early readers (Baker et al., 1998). Literacy

skills like alphabet knowledge, print concepts and early writing develop through socio-cultural experiences from birth – and are strong predictors of conventional reading and writing ability (Adams, 1990; Cohen & Cowen, 2007; Teale & Sulzby, 1986; Whitehurst & Lonigan, 1998).

The most effective interventions are those where parents and school personnel work together to implement a two-way exchange of information (e.g. parent-teacher action research teams), and those involving communication between school and home (e.g., daily report cards, school-to-home notes) (Cox, 2005). Literature has established the role of parent involvement in increasing the likelihood of long term effects of early childhood programs (Reynolds, Maurogenes, Bezruczko & Hagemann, 1996; Seitz, 1990). Moreover, parent involvement at home and in school also appears to be a good predictor of school achievement and grade promotion (Clarke-Stewart, 1988; Reynolds, 1991; Reynolds & Bezruczko, 1993). Early childhood programs that involve parents significantly through home visitation or through parent education or center involvement show longer-term effects (Campbell & Ramey, 1995; Schweinhart et al., 1993; Yoshikawa, 1994). Direct parent involvement in such programs is expected to enhance parent-child interactions as well as attachment to school, thus promoting school readiness and social adjustment (Reynolds, 1998).

Pikulski (1994), in his analysis of five effective reading interventions, mentions home involvement as one of the key contributing factors to the programs' success. Similar findings have been reported by Morrow & Young (1997) who used a combination of school-based literacy program and family literacy program. Children reported to have been involved in reading more often as well as enjoyed working with their parents.

In another study, focusing on transitioning from a trainee to a teacher in Pakistan, Westbrook et al. (2009) found that newly qualified teachers believed that the students with more literate parents and siblings made more progress with their learning. The

teachers also pointed out that home background and community are additional barriers to their preferred pedagogical approach (pg. 5).

Research establishes the importance of involving parents in early literacy interventions. However, the majority of parents of participating students in these programs in developing countries are illiterate. Therefore, the role that these parents play in facilitating their children's learning and achievements, other than providing motivation, is unclear. However, it emerges that the supportive home environment does make a difference in children's reading skills, so even illiterate parents can play a supportive role.

Textbook/Reading Material Development as a Goal of the Intervention

Textbooks are vehicles for delivering content knowledge and determining largely what happens in a class (Hummel, 1998. as cited in Lebrun et al., 2002). Access to and availability of textbooks is a particularly significant factor in predicting academic achievements (Heyneman et al, 1978, cited in Oakes & Saunders, 2004). Altbach (cited in Lockheed et al., 1986) UNESCO (2014) declared textbooks and supplementary reading material to be essential for building foundational literacy and developing lifelong reading habits.

The emphasis on textbooks and reading materials as part of literacy interventions is based on the idea that textbooks can provide students an exposure to written materials that is otherwise unavailable in the environment (Heyneman et al., 1989 as cited in Lockheed et al., 1986). The textbooks and reading materials help students to learn independently, particularly through completing their homework (Rohlen, 1983 as cited in Lockheed et al.; 1986).

Moreover, in order to be able to become fluent readers, children need individual practice along with other important skills (Abadzi, 2013). This sometimes becomes an added challenge for children in low-income countries due to poor reading habits as compared to children in high-income countries due to better literacy standards and good reading habits (Abadzi, 2013). Low-income countries face 'print poverty', often leaving

the students with limited or poor reading practice from the blackboard (Abadzi, 2013). Furthermore, it is common for the teachers to 'focus on few best students' in developing countries and 'nonexistent materials' is one of the causes. This situation has created a dire need for the development of textbooks that could aid children in developing nations to master reading in their language. These countries need to spend on textbooks because 'next to an engaged and prepared teacher, well-designed textbooks in sufficient quantities are the most effective way to improve instruction and learning' (UNESCO, 2016).

The analysis of data from regional assessments in 22 sub-Saharan African countries by Fehrer et al. (2009) further strengthens the notion that pedagogical resources especially textbooks are effective ways of improving learning. They found that provision of textbooks to each student could increase literacy scores by 5-20%.

However, as pressing the need for developing books and supplemental materials is – it is essential to explore the kind of books and materials best suited to the learning needs of the local students.

A review of early grade reading in Latin America and the Caribbean revealed that children recognize the type of written material (for example; traffic signs, medicine labels and story books) that are associated with their experiences (USAID, 2016). This leads to the idea that reading materials and pedagogies should include live experiences (USAID, 2016) and thus give a direction that could contribute to the production of effective reading materials.

Nevertheless, it must be made clear that textbooks cannot function alone (Mohammad & Kumari, 2007) and therefore textbook writers and editors have a responsibility of clarifying every detail for teachers so that they know what are they teaching and why are they teaching (Mohammad & Kumari, 2007).

Mohammad and Kumari (2007) highlight an important phenomenon in Pakistan where teachers leave out the sections from textbooks that they are not familiar with. This revelation has serious implications for the

agencies and organizations seeking to develop textbooks and other materials for Pakistani students. Will they be able to orient teachers to an extent that they become well versed in the content and strategies introduced? If not, this will pose serious questions on time and resources spent on attempts to develop new and better materials.

Another major issue is the perception of what textbooks can achieve in developing countries. Lockheed et al. (1986) carried out an analysis of the longitudinal data drawn from national sample of eighth-grade mathematics classrooms; they learned that textbooks are seen as a substitution for postsecondary teacher education in developing countries and are believed to have the potential to deliver a more comprehensive curriculum. Such a perception is problematic since it separates teachers and textbooks as two independent components of education.

In 2005, USAID launched 'The Textbooks and Learning Materials Program (TLMP)' in Ethiopia, Ghana, Malawi, Senegal, South Africa and Tanzania. Tarnoff (2005) believes that this program exemplifies USAID's efforts to improve education quality through the development and distribution of textbooks and learning materials. However, an analysis carried out by Allan and Horn (2013) of the same program revealed numerous weaknesses (that they define as 'lessons learned'). They noted that for introducing teaching and learning materials, extensive research is needed to assess the readiness of teachers and the adequacy of their training for making the best use of the new textbooks and materials. Similarly, the aptitude of the students must also be gauged to avoid any assumptions about their potential. Most importantly, teachers must be allowed to try out the newly-developed materials in classrooms over a considerable span to determine the challenges that may arise; this should lead to further research and revision before final printing.

In a country like Pakistan, where the mechanism for implementing educational innovations is severely deficient, carrying out such a rigorous process may pose extraordinary challenges affecting the overall success of any initiative.

INTERVENTION PROJECTS: ISSUES OF IMPLEMENTATION AND SUSTAINABILITY

Successful programs for educational change have a different profile than the less successful programs (Verspoor, 1989). The three most critical factors for the implementation of any educational reform that Verspoor (1989) identified are (i) administrative development (ii) in-service teacher training (iii) strategies to gain commitment of all stakeholders.

Warwick et al. (1992) carried out an analysis of the implementation of five educational innovations in Pakistan. They note that for successful implementation of an intervention, it is imperative that both kinds of intelligence; Initial (works out program planning and design) and Ongoing (evaluates various aspects during implementation) perform efficiently. Furthermore, any implementation that is seen as an imposition by outsiders is likely to face resistance. Therefore, it is compulsory that persons with local influence participate in the planning and execution, field implementers are convinced about the purpose of the innovation and parents are taken on board, so much that they do not see the innovation as threatening their culture.

It is equally important to realize that no educational change can be 'teacher-proof' (Warwick et. al, 1992); therefore any innovations that fail to help the teachers to understand their purpose and convince them about being useful will not succeed. Only a few studies indicate that non-certified teachers can increase student reading skills effectively after going through sufficient training (Miller, 2003). But it is unclear as to what this training should look like. Findings from various studies outline autonomy and support as two factors that contribute and make an intervention self-sustaining (Pelletier, Fortier, Vallerand & Briere, 2001). The teachers implementing an intervention should be made part of decision making and allowed to control some aspects of it (Fired, 2012). They must also be provided with opportunities for self-reflection and feedback and their work should be recognized (Couston-Theoharis et al., 2007).

Warwick et al., (1992) cite that one of the reasons for the failure of teachers in using teaching kits effectively is its conflict with their own classroom style, which emphasized on lecturing and rote memory and the lack of energy in trying anything new. This implies that sufficient reorientation of thoughts is needed before the participant teachers in any intervention could be expected to perform well.

It is challenging to incorporate all characteristics mentioned above in a teacher training program in Pakistan, because of a generally low level of education among practicing teachers (Kanu, 2005). Usually teachers recruited neither have adequate subject knowledge, nor the attitudes and pedagogical and classroom management skills required to inculcate conditions for effective student learning (Kanu, 2005). The international consensus on quality teaching posits that learning should be learner centered (UNESCO, 2004).

Poor infrastructure of public schools in Pakistan also poses additional challenges for the successful implementation of interventions. The poor state of furniture, including boards is alarming (Westbrook et al., 2009). The classes are overcrowded, causing additional difficulties for teachers (Little, 2006). The lack of resources for the preparation and storage of audiovisual aids also discourage their use (Westbrook et al., 2009).

It is often observed that the approaches to teaching literacy in the developing countries have their basis in curricula devised in developed countries giving way to learning objectives that are not based on the reality of students' abilities (Abadzi, 2017, 2-3). Similarly, provision of reading materials is essential, but books must be aligned with the local context, culture and reading competency of the students (Malik et al., 2015, 2). Simply providing books is not enough (Graham & Kelly, 2018).

In case if the challenges outlined above are overcome, what is the possibility that the intervention is bound to produce the intended results?

Shanahan (1995), while endorsing what Durkin (1974-75) had underlined, provides a convincing answer to our question. She states that children fail to maintain their initial achievement advantage over their peers, because subsequent instructions fail to capitalize on the advantage. In order for the effects of an intervention to sustain in the long-run, subsequent instructions must be responsive to the higher achievement shown by children.

CONCLUSION

In the face of so many challenges, what can be done to be able to predict some degree of success and sustainability of an intervention? Literature outlines certain principles, factors and characteristics that can make an intervention effective.

Reynolds (1998) elucidates eight essential principles of effective early childhood programs. One of the principles which is particularly relevant to the context of Pakistan is keeping the class size small enough to promote individualized learning experiences which in turn will lead to greater school achievement in later years. He further maintains that teachers with bachelor's degrees or relevant certification are more likely to offer developmentally appropriate practices in the classroom. Graham and Kelly (2018) assert that in order for an intervention to sustain, continued trainings (not conference-style) and support on how to use the materials is essential. This emphasis on using the correct teaching methods has been bolstered by Slavin et al. (2009) as they compared schools using two different reading interventions employing phonics instructions and note:

“... imply that whereas the importance of phonics and phonemic awareness in reading instruction are well established, the addition of phonics to traditional basal instruction is not sufficient to bring about widespread improvement in children's reading. Other factors, especially relating to teaching methods, are also consequential”(p.4).

Building the capacity of teachers to participate fully is of immense importance. Professional development should be accompanied with practical courses that not only improve teacher's knowledge, but also train them in the real sense of its practice (Villegas & Reimers, 2000). Furthermore, school leadership must realize that teachers' confidence may decrease initially as they try out new strategies and hence they must respond with necessary feedback and reassurance (Tschannen-Moran et al., 1998). A supportive environment will ensure that teachers do not revert back to their old instructional strategies and methods due to their stress (Goleman et al., 2002).

Interventions programs can be effective when they are not merely short-term initiatives, but rather as long-term programs that are taken over by host-country governments to sustain (Graham & Kelly, 2018). An innovation is only likely to survive if it is cohesive both internally (in terms of theory, training, program design, evaluation) and with the host system (i.e., it must be workable, can effectively contribute, cost effective and a winner with the stakeholders) (Clay, 1987).

3

Methodology

This chapter discusses the research methodology employed to explore the efficacy of the Sindh Reading Program (SRP) intervention in enhancing reading skills in the early years. In particular, this chapter presents an insight into research methodology organized in two main sections, i.e. a quantitative approach and a qualitative approach. These main sections are further categorized into three sub-sections. The first sub-section describes the research design along with the sample and sampling procedure. The second sub-section discusses the data collection methods, data collection tools, and tool administration process. The third sub-section discusses the strategies employed for data analysis. The chapter concludes with a brief discussion on the limitations of the study.

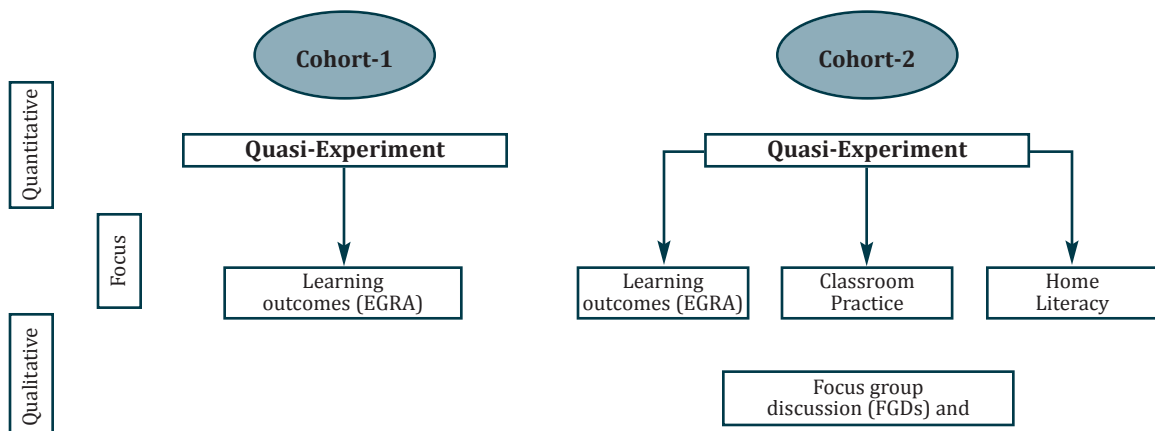
RESEARCH DESIGN

For studying the impact of intervention, mixed method research methodology was employed. Use of both quantitative and qualitative methods provided researchers with a greater scope to investigate using both numbers and words; moreover, it allowed for

more comprehensive analysis (Almalki, 2016). It is important to note that both quantitative and qualitative data were collected simultaneously, followed by analysis.

As presented in Figure 1, quasi-experimental design was employed as a quantitative design to explore efficacy of the SRP intervention for student learning outcomes for both Cohort-1 (TLA model) and Cohort-2 (GT model). Secondary data were analyzed for Cohort-1 where the research team coded and recoded the data to make comparison across groups (non-SRP and SRP) and overtime (pre- and post-tests). Primary data for the quantitative part were collected for Cohort-2 at three levels: children’s learning outcomes, classroom practice and home literacy environment. In addition, qualitative data were collected from various stakeholders through semi-structured interview and Focus Group Discussions (FGDs). The next section discusses in detail the quantitative and qualitative approaches used in this study.

Figure 1: Overall research design



QUANTITATIVE APPROACH

Considering the nature of the research questions, the quantitative paradigm was chosen to answer research question 1: “*What explains the variation among the students learning outcomes (zero scorers vs. readers)?*” The first research question was addressed using a quasi-experiment with two variants for Cohort-1 (non-intervention pre- and post-test design) and Cohort-2 (non-intervention post-test only) design. The design was carefully crafted in coordination with implementers. In order to explore the efficacy of two models of interventions employed in Cohort-1 and Cohort-2, respectively, two sets of children’s assessment data were analyzed quantitatively. Secondary data were used to analyze the effects of Cohort-1, while primary data were collected for Cohort-2.

Sample and Sampling

The section presents a description of the sample and sampling procedure used in this study. The sampling was carried out separately for Cohort 1 and 2.

Cohort-1 (TLA model)

This section presents the sampling strategy used for Cohort-1. The secondary data, which comprised of only children’s scores on the EGRA baseline (pre) and midline (post) were collected and entered by the SRP team as part of their regular monitoring system. The SPSS file was forwarded to the AKU-IED research team for analysis along with the coding scheme. Multiple meetings were held with the SRP team to select a sample for Cohort-1 (as well as Cohort-2) to align with the proposed research design ‘non-intervention pre- and post- tests experiment’ design. The sampling was carried out at four levels (i.e. district, school, classroom, and children).

In the original file forwarded to the research team, the data collected at two points were coded as baseline and midline. The team defined a new variable to identify non-SRP and SRP schools. Henceforth, baseline and midline will be referred to as pre- and post-test, respectively. Figure 2 presents a brief overview of the sample and sampling strategies employed for Cohort-1.

Figure 2: Quantitative approach: sample and sampling for Cohort-1

Levels	Non-SRP	SRP
Districts (n = 5)	n = 5	
Schools (n = 50)	n = 20 (40%)	n = 30 (60%)
Children (n = 1113)	n = 436 (42%)	n = 648 (58%)
Baseline (pre)	232	333
Midline (post)	232	316

The next section discusses the sampling procedure at each level.

Districts: The SRP intervention has been carried out in 12 different jurisdictions including districts from interior Sindh and some areas of Karachi. Of the eight SRP target districts, five were strategically selected for four main reasons: (a) to represent a variety of geographical locations i.e., southern (Karachi), central (Dadu, Larkana, Sukkur) and northern (Kashmore) districts of Sindh, (b) to cater for the geographical spread across urban (i.e. Karachi) and rural (i.e. Kashmore, Larkana, Sukkur, Dadu) settings, (c) to showcase a range of practices i.e., good (Kashmore), average (Dadu, Sukkur, Larkana) and below-average (Karachi) as per the general assessment of the SRP team, and medium of instruction (Sindhi and Urdu). This strategy also helped to align the sample for Cohort-1 (secondary data) and Cohort-2 (primary data).

Schools: From within 5 districts, 50 schools were selected for secondary analysis of Cohort-1 to represent non-SRP (n=20, 40%) and SRP (n=30, 60%) groups. Both SRP and non-SRP groups represented single-sex and co-education system. The schools were selected in close consultation with the SRP field and the research team. It is important to note that schools remained the same for pre- and post-tests; however, different cohorts of Grade 3 children were tested at two points. All non-intervention schools which were available in the secondary file included a non-SRP

group for analysis. As expected, the rural sample was bigger (n=776, 70%) than the urban (n=339, 33%). The rural districts were over-represented in the SRP population and the proportion is reflected in the sample of the study. This ratio was also comparable with the Cohort-2 sample.

Children: Altogether the EGRA score of 1113 children was analyzed which represents non-SRP (463, 42%) and SRP (648, 58%) groups. Of the 1113 children 565 (51%) were tested before the intervention and 548 (49%) completed the EGRA test after the intervention. The number of children tested for baseline and midline was not significantly different across groups [$\chi^2 (1) = 0.186$; ns]. The sample size for boys (n=611, 55%) was bigger than for girls (n=502, 45%); however, the gender ratio across groups was significantly different [$\chi^2 (1) = 10.735$; $p < 0.01$]. Particularly, the SRP group overrepresented boys (n=382, 59%).

Cohort-2 (GT model)

This section presents the sampling strategy used for Cohort-2. For Cohort-2, the sampling was carried out at five different levels (i.e. district, school, classroom, children, and parents). Figure 3 presents a brief overview of the sampling strategies employed for Cohort-2.

District: Of the eight SRP target districts, five districts were strategically selected for the same reasons cited in the Cohort-1 sample section (i.e. geographical

spread, settings, and range of EGRA performance). This strategy also helped to align the sample for Cohort-1 (secondary data) and Cohort-2 (primary data).

Schools: The participants of the Cohort-2 were recruited from 54 randomly selected schools to represent the non-SRP (n=26, 48%) and SRP (n=28, 52%) groups. Two-thirds of the schools (n=36, 67%) were recruited from the rural districts while the rest (n=18, 32%) represented urban districts. As expected, the ratio of schools in two different settings remains the same across groups. The rural districts were over-represented in the SRP population and the proportion is reflected in the sample of the study. Since this study employed quasi-experimental design which seeks to compare the scores of the SRP and non-SRP groups, it was important that both groups had similar characteristics (e.g. districts, school system, class size) in order to minimize the effect of extraneous variables.

Classrooms: Grades 1 and 2 were chosen on purpose as the SRP intervened in these two classes to represent early years. Altogether, 104 classrooms were recruited to represent the non-SRP (n=50, 48%) and SRP (n=54, 52%) groups. The mean class size for the whole sample was 26 (SD=12.6) children and the class size varied from 6 to 61. The class size was slightly bigger in the SRP (M=27, SD=13.8) than the non-SRP (M=25, SD=11.3). However, the difference across groups was not significant [Mann Whitney U=39540.500; ns]. Overall, there were more co-education (n=78, 75%) than single sex (n=26, 25%) classes and the ratio was similar across groups [$\chi^2 (1) = 0.188$; ns]. On average, the teachers in this sample spent 21 minutes (SD=9.1) in teaching a language lesson either Urdu or Sindhi with a range of 5 to 45 minutes. The SRP teachers were observed to have spent more time on teaching a target language (M=22.1 minutes, SD=9.4) than their non-SRP (M= 19.6 minutes, SD= 8.5) counterparts and the difference was found to be significant [Mann Whitney U=33955.500; $p < 0.01$]. It is important to note that ‘prescribed’ time for language teaching reported by the school ranged from 30 to 40 minutes. The observer expected to spend this duration in the classroom with

Figure 3: Quantitative approach: sample and sampling for Cohort-2

Levels	Non-SRP	SRP
Districts (n = 5)	n = 5	
Schools (n = 54)	n = 26 (48%)	n = 28 (52%)
Classrooms (n = 104)	n = 50 (48%)	n = 54 (52%)
Children (n = 609)	n = 287 (47%)	n = 322 (53%)
Parents (n = 56)	n = 23 (41%)	n = 33 (59%)

prior consent from the teachers. However, in some cases teachers switched to the other subjects with a clear cue to the observer that *the language lesson is over*. In this case the observer had to cease the observation after getting a 'goodbye cue' from the teacher.

Children: On average, 12-15 students were randomly selected from Grade 3. In schools where the class size ranged from 6-15, the whole class was recruited as a sample for the study. Of the 609 children who participated in the study, 311 (51%) were girls and 298 (49%) were boys. However, there were more girls in the SRP (n=190, 61%) than the non-SRP (n=121, 39%) and the gender ratio across group was significantly different [$\chi^2 (1) = 17.233$; $p < 0.001$]. The mean age of the children for the whole sample was 9.70 years (SD=1.41) with slightly younger children in SRP (M=9.68, SD=1.32) than non-SRP (M=9.72, SD=1.50). However, the difference between the mean age of the two groups was not significant [Mann Whitney U=45831.000; ns].

Parents: In this study, 56 parents participated to represent the non-SRP (n=23; 41%) and the SRP (n=33; 59%) groups. Of the 56 parents who participated in the face-to-face structured interview, 30 (53%) were fathers and 26 (46%) were mothers. However, the gender ratio was equal in the non-SRP and the SRP groups [$\chi^2 (1) = 2.128$, $p > 0.05$]. Table 2 presents a comparative overview of the parents' qualification based on the information gathered from the participant parent. Regardless of group, a majority of the parents in this sample did not attain any academic qualification with a relatively higher percentage in the non-SRP group than their SRP counterparts. The remaining sample parents showed

variability with some having completed primary/ elementary school, others secondary school or a bachelor's degree⁴. Evidently, parents of SRP children reported having an edge over non-SRP in all categories of education with a significant difference only for fathers [U=235.000; $p < 0.05$].

Data Collection Process

This section presents the process of data collection along with the details of data collection tools and its administration to the study sample. The data were collected from October 2017 to January 2018; however, planning for data collection and field visits was initiated in July 2017. In rural districts, the data were collected over a period of three weeks during October 2017; whereas, in urban districts, the data were collected during October 2017 and January 2018. Data collection timelines were developed in coordination with the SRP team to schedule the field work. For this study, the data were collected at three levels i.e., children, classroom, and parents.

Data Collection Tools

This section discusses in detail the assessment tools which were developed and adopted/adapted as part of this study. The description of the assessment tools is organized under three subsections: (i) children's assessment tool; (ii) classrooms' assessment tool; and (iii) parents' assessment tool.

Children's assessment: EGRA

A children's assessment tool was employed to gauge the efficacy of the intervention for enhancing target students' reading skills. The Sindh Reading Program (SRP) shared with the research team a contextually relevant and theme-specific assessment tool entitled 'Early Grade Reading Assessment (EGRA)' to assess

Table 2: Parents' qualification - comparison across group

Parents / Groups	Mother		Father	
	Non-SRP	SRP	Non-SRP	SRP
No-schooling	16(70%)	17(53%)	14(61%)	09(28%)
Primary / elementary	04(17%)	07(22%)	05(22%)	10(31%)
Secondary	02(9%)	08(25%)	03(13%)	10(31%)
Graduation	01(4%)	0%	01(4%)	03(9%)

Table 3: A brief overview of EGRA

Constructs	Structure	Instructions/Activity	Time
Phonemic	10 words	Identify last sounds of the given words	Non-timed
Non-word fluency	40 non- words (pseudo words)	Read 40 words accurately in one-minute	Non-timed
Expressive vocabulary	10 pictures	Identify object, subject, and action to describe a given picture	Non-timed
Oral reading fluency	60 words story	Read a story in one minute	Timed
Reading comprehension	5 questions to comprehend a sixty-words story	Respond to five questions one-by-one	Non-timed
Listening comprehension	3 questions	Listen to a brief story and respond to three questions	Non-timed

children’s reading skills. The SRP has been using this tool for their project’s assessments; therefore, this tool was employed to maintain consistency.

Overview of the EGRA: The EGRA tool comprised six main constructs i.e., phonemic, non-word fluency, expressive vocabulary, oral reading fluency, reading comprehension, and listening comprehension. Each construct is comprised of a specific structure along with the set of instructions to assess children’s reading skills. A brief overview of EGRA is shown in Table 3.

As shown in Table 3, five types of skills have been tested as part of the pre-test of Cohorts-1 and Cohort-2 and the post-test of Cohort-1. All EGRA-based target skills (i.e. phonetics, non-word recognition, vocabulary, passage reading followed by a couple of questions to test comprehension; and listening skills-short story telling followed by questions) were tested as part of the post-test of Cohort-2. It is important to note that each construct had a specific structure and a set of questions which were particularly designed to measure an aspect of reading skill. Clear instructions were also mentioned alongside each construct to (a) guide the researcher for the process of data collection; (b) ensure proper administration of the tool, and (c) maintain consistency throughout the assessment process.

Establishing the reliability of EGRA: In order to establish the internal consistency, the Cronbach’s Alpha on six constructs of the EGRA was calculated. The calculated value of the Cronbach’s Alpha was found to be (0.79) for Cohort-1 and (0.76) for Cohort-

2. This value is greater than the standard alpha value recommended for such tests i.e. (0.70). Hence, the test was considered as consistent and reliable for the study.

Classroom assessment: ERCOP

Structured observations of SRP-trained teachers and non-SRP teachers were carried out in Grades 1 and 2 of all schools in order to assess the efficacy of the intervention program. The classroom observations provided information on the context of teaching and learning at classroom level (e.g. classroom interactions, reading-oriented activities, availability and accessibility of resources, classroom physical environment). Classroom Observation Scale (COS) was adapted and renamed Early Reading Classroom Observation Profile (ERCOP) for this study to gather data on classroom practice.

Overview of the ERCOP: ERCOP is a rating scale which consists of 20 items. In this tool, the items fall under two major categories i.e., ERCOP-pedagogy and ERCOP-reading.

ERCOP-pedagogy: In the category of generic pedagogy, there are 10 items which mainly deal with the generic pedagogical aspects of classroom practice (e.g. questioning, active involvement of students, teacher-child interaction). Often items, eight items are on a three-point rating scale; whereas, the other two items deal with teaching strategies and classroom management strategies in the form of a checklist. Figure 4 presents an example of an item from ERCOP pedagogy.

Figure 4: ERCOP pedagogy: example of an item

Guidelines	Items	Rating		
		1	2	3
<p>Does the teacher interact with children in a supportive manner? ‘Supportive’ includes support given by the teacher in words as well as through gesture and body language, for example, listen to children’s talk/questions attentively, exhibit calm tone, do not discriminate.</p> <p>No credit for unpleasant interaction (overly controlling teachers, rejects children’s ideas/questions, not involved with children); partial credit for some pleasant interactions, full credit when interaction are pleasant throughout the observation (teacher listens attentively, treats fairly and children listen to teacher when she/he speaks).</p>	<p>2. Teacher-Child Interaction</p> <p>Teacher usually responds to children in a supportive manner (e.g. teacher and most of the children seem relaxed, voices cheerful).</p>	<p>Comments:</p>		

Figure 5: ERCOP reading: example of an item

Guidelines	Items	Rating		
		1	2	3
<p>This refers to the awareness of the sounds within spoken words and activities employed to promote this. It includes the teaching of letter sounds, blending, words sounds, and rhymes. It may also include decoding sounds and words.</p>	<p>12. Phonological Awareness Activities</p> <p>Teacher uses phonological awareness activities</p>	<p>Comments:</p>		

ERCOP-reading: In the category of ERCOP reading, there are ten items, all on a three-point rating scale. The items in this category mainly deal with aspects related to reading instruction (e.g. phonemic awareness activities, loud reading, silent reading, group reading, vocabulary teaching). Figure 5 displays an example of an item from ERCOP reading.

Tool adaptation: Early Reading Classroom Observation Profile (ERCOP) was adapted from Classroom Observation Scale (COS) (Bhutta, Anwar & Chauhan, 2011). COS is a rubric which consists of 15 items. The descriptors are anchored at three points. In addition, specific guidelines are provided with each item for the observers. COS encompasses various aspects of classroom quality including *physical set-up, classroom interaction, teaching learning methods and planning and monitoring*. Of the 15 COS items two items are

defined as checklists in order to capture a variety of active methods (e.g. questioning, role-plays) and classroom management strategies (e.g. whole class, group/pair work) used during the lesson which was observed. COS has been validated in more than 500 classrooms across grade levels (primary to secondary); geographical settings (urban to rural); and observation of various subjects (e.g. English, mathematics, science, social studies) classrooms in a variety of classrooms across the country.

For this study, the COS was adapted at three levels. First, the number of items was reduced from 15 to 10. Only those items which had direct relation with the research question were included in the tool. Second, new items were added to serve the need of the project. Ten new items were adopted from the tool used in a study where ‘Target Child Observation’ was employed

as a strategy to observe literacy practices in classrooms (Sylva, Hurry, Mirelman, Burrell, & Riley, 1999). Third, the marking scheme was modified. In the original scale, the items were defined for target classroom observation where coding of classroom activities was carried out using minute-by-minute observation; however, for this particular study, the items were taken from the original tool and were defined on a three-point scale. This step was carried out to ensure a similar marking scheme throughout the tool. Hence, the adapted tool-ERCOP comprised of 20 items which dealt with two major themes/categories i.e., ERCOP-pedagogy and ERCOP reading.

Establishing the reliability of ERCOP: To determine the internal consistency of the ERCOP, the Cronbach's Alpha was calculated for ERCOP-overall (18 items), as well as for ERCOP pedagogy (8 items) and ERCOP-reading (10 items). Table 4 shows the results of the reliability analysis. Reliability values are greater than the cut-off point (0.70) which suggests that the classroom observation tool was considered reliable for the study.

Table 4: Reliability of ERCOP

Total	Cronbach's Alpha	Items
ERCOP - Overall	0.87	18
ERCOP - Pedagogy	0.89	08
ERCOP - Reading	0.76	10

Parents' assessment tool

While there is no evidence of 'direct' engagement of parents in reading intervention, data were collected to gather quantifiable evidence, if any, of parents' engagement in enhancing students' reading skills. A questionnaire titled 'Home Literacy Environment' was adopted from PISA (2015) which has been extensively used in all PISA studies.

Overview of Home Literacy Environment: The Home Literacy Environment questionnaire comprised of ten items. The items were scored on a four-point frequency rating scale. The questionnaire was

administered face-to-face; however, it was translated into Urdu – lingua franca of Pakistan – to not only make it user-friendly, but also to ensure consistent administration. Despite this, the statements had to be translated into the local language during administration especially in rural settings. It was easier for the field researchers to translate statements from Urdu than from English.

Establishing the reliability of the Home Literacy Environment: In order to establish the internal consistency of the parental questionnaire, the Cronbach's Alpha on ten items was computed. The calculated value of the Cronbach's Alpha was found to be 0.88, indicating that the tool was reliable.

Administration of the Data Collection Tool

This section presents the details of the administration process including the composition of the research team, the training of the researchers and the data collection process.

The data for this study was collected by an AKU-IED research team which comprised of Principal Investigator (PI), Co-Principal Investigator (Co-PI), a research coordinator and two MPhil students from AKU-IED. Additionally, three local research assistants were hired for EGRA administration in rural districts. All three were oriented in the process of EGRA administration prior to field work. Concentrated field work was carried out in rural Sindh over three weeks where eight team members (including PI and Co-PI) were engaged in data collection simultaneously. It is important to note that PI and co-PI accompanied the research coordinator and assistants in the field throughout the data collection process not only to maintain the quality of the data collection process, but also to participate actively in the field work.

Before the data collection process began, the research coordinator corresponded with district focal persons in target districts. The SRP team helped tremendously during this process. These district focal representatives helped in locating the sampled schools in each district. The process was followed by the

negotiation for data collection by the research team. The research team coordinated with the head teacher and the administration staff of each school to setup schedules for assessments, select a proper venue within the school for test administration, and coordinate with selected parents for data collection. The core-team, research coordinator and field researchers were oriented in administration of EGRA by the SRP representative. The main team then oriented field researchers in rural areas in administration of EGRA. At the outset, the data were mainly collected by PI and Co-PI whereas other team members were engaged in learning the process of data collection. Upon attaining a certain level of familiarity with the research field, and consistency in administration of tools (i.e. EGRA, classroom observation, parents' questionnaire) the research team started collecting data for this study. The data were collected from children, classroom, and parents using EGRA, ERCOP, and the parent questionnaire, respectively. The next section presents the process of the tool administration.

Children assessment: EGRA

For the test administration, a specific protocol was followed which was kept consistent for all schools. According to the protocol, the test was administered to randomly selected Grade 3 children in a one-to-one setup where a sufficient distance was maintained between the researcher and the sample child. Prior to the administration of the EGRA, the researcher built rapport with the child by asking general questions. Next, the researcher began to administer the tool construct-by-construct following the protocol. It is important to note that the instruction phrases, language, and tone were kept similar during administration of EGRA in order to avoid possible effects on children's responses. On average, each EGRA administration took around 15-20 minutes. Upon completion, the researcher thanked the child and requested to call the next child. At the end, the assessment sheets were kept safely in an envelope. The

data collector recorded necessary information (name of the school, number of children, date of data collection) on each envelope. The envelopes from all field teams were handed over to the research coordinator after the school visit. A similar procedure was followed in all schools across districts.

Classroom assessment: ERCOP

A 20-item assessment tool, the Early Reading Classroom Observation Profile (ERCOP) was employed to observe teachers' classroom practices in both SRP and non-SRP classroom across districts. Grades 1 and 2 were the target classrooms for observation. In the SRP schools, those classrooms were observed where the lessons were taught by SRP-trained teachers; whereas, in non-SRP schools, teachers of Grade 1 and 2 were the part of classroom observations. The next section discusses in detail the procedure of classroom observation. Prior to the classroom observation, the research team coordinated with the school staff to schedule observation for the selected classrooms.

Initially, the research team was oriented by Co-PI on observing classrooms using ERCOP. Two research team members independently conducted classroom observations in pairs. After classroom observation, the team members discussed the observations to ensure the inter-rater agreement. Upon achieving 70% -80% of the inter-rater agreement, the research team members started observing classrooms independently. It is important to note that observations were undertaken by PI, Co-PI, two field researchers (students of MPhil in Education and AKU-IED) and research coordinator. The local field researchers were restricted to EGRA administration.

During classroom observation, the researcher sat at the back of the classroom and remained silent throughout the lesson. In cases where activities or group discussion were part of the lesson, the researcher first sought the permission and then observed the activities and group discussion from a distance to avoid eye contact with children. The

observation sheets were kept in an envelope after completion. The envelopes from all field teams were handed over to the research coordinator after the school visit. The data collectors recorded the necessary information (name of the school, date of data collection, observation of grades – Grade 1 /Grade 2) on each envelope. A similar procedure was followed in all schools across districts for classroom observations.

Parents' questionnaire

The 'Home Literacy Environment' questionnaire was administered to 56 parents who represented both SRP and non-SRP groups. In both groups, only those parents were called for data collection, whose children have been part of the test (EGRA). For the administration of the parental questionnaire, face-to-face structured interview strategy was employed. The research team coordinated with the school staff to schedule parents' visits for the selected children. The tool administration took approximately 10-15 minutes with each participant. For this assessment only, the researcher was accompanied by an interpreter who was well-versed in the local language (in cases where translation was required) as well as with the items in the questionnaire.

Prior to the tool administration, the researcher introduced the purpose of the research study to the parents and obtained their verbal consent. Afterwards, the researcher began to administer the tool item-by-item following the protocol. The researcher first read aloud the statement and then gave four options for the statement. The researcher provided some time for the participant to think and respond. Upon receiving the response, the researcher marked the response on the questionnaire. In cases, where language was a barrier, the interpreters translated the statement immediately. The completed questionnaires were kept in an envelope and handed over to the research coordinator. A similar procedure was followed in all schools across districts for parental questionnaires

Data Analysis

The aim of the study was to carry out an in-depth analysis of various aspects of SRP interventions towards enhancing reading skills. The quantitative arm of this study sought to explore whether SRP introduced interventions made any difference in teachers' pedagogies and children's reading abilities in the target districts of the province of Sindh, and if so, what factors and processes explain that difference. The components of the quantitative framework include children's learning outcomes, teachers' classroom practice and parents' views about the intervention. The study was conceptualised by the research team at AKU-IED along with SRP representatives. A battery of assessment tools were employed to collect data from children, classrooms, and parents. The quantitative data collected were analyzed using the Statistical Package of Social Science (SPSS version 20) which is considered as reliable software for analyzing quantitative data.

For Cohort-1, a SPSS file of test scores was obtained from the SRP team. For Cohort-2, data which were gathered from children (EGRA scores), classrooms (observation ratings), and parents (questionnaire ratings) were entered by the research team at item level for individual cases. After cleaning, average scores were computed for variables measured at interval level (e.g. children's test scores). Categorical variables were transformed into dummy variables by coding (e.g. non-SRP =1; SRP =2). The characteristics of all measures were then explored using objective tests of normality (e.g. comparing skewness and standard error of skewness) and homogeneity (e.g. Levene's test). In order to reduce the potential bias due to skewed data, bootstrapping was employed to generate confidence intervals for the mean differences. It helped to re-estimate the standard error of the mean difference (Field, 2013).

Students' EGRA scores were analyzed in five distinct steps. *Firstly*, overall percentage mean scores were computed and compared across groups for Cohort-2.

For Cohort-1, analyses were carried out to compare differences in non-SRP and SRP groups as well as progression in reading skills over time. *Secondly*, a detailed analysis was carried out to explore children's performance in six EGRA domains across groups. Percentage scores were computed to put all domains on the same scale. It is important to note that the raw data on 'timed-items' (oral non-word fluency, oral reading fluency) were treated to compute fluency before generating percentage scores for main analysis. In order to make the above mentioned comparisons t-test was used with bootstrapping. The effect sizes were computed for all comparisons in order to provide the reader with a sense of the magnitude of the statistically significant differences reported in results. The effect size is denoted as 'r' and its value varies from 0 to 1. The range was used to define small (less than 2); medium (starts with 3); and large (starts with 5) effect size. The large effect size indicates a bigger magnitude of difference. *Thirdly*, new variables were computed to present a comparative overview of zero readers (those who cannot read a single word) and non-zero readers (who can at least read one word) for both Cohort-1 and 2. In order to compare the ratio of zero readers across groups, chi-square was used along with effect size (i.e. Phi) to gauge the magnitude of the difference. *Fourthly*, benchmarks were defined to understand how the readers' scores (non-zero readers) were distributed across four levels. After analysis of mean scores and zero readers ratio four performance categories were defined to demonstrate distinguished levels of 'readers': level 1 (1% to 25.9% score); level 2 (26% to 50.9% score); level 3 (51% to 75.9% score); and level 4 (76%-100% score). Performance of those children who 'qualified' to be readers (non-zero readers) was compared against these four levels for overall scores in six EGRA domains. A detailed analysis was undertaken to determine students' performance in the six domains defined for EGRA. *Finally*, the EGRA scores were analyzed by language (i.e. Sindhi and Urdu) to explore distinct patterns based on the medium of instruction.

In order to make the report more 'reader-friendly', only alpha value (p), effect size (r) and percentage differences (for significant differences only) are presented in the text. Detailed results of statistical analysis for EGRA (both Cohort-1 and 2) are appended for readers' reference.

Classroom observation data were analyzed by computing overall ERCOP scores as well as scores on ERCOP pedagogy and ERCOP reading. These scores were compared across groups using t-test with bootstrapping to highlight the differences between classroom practices in terms of overall teaching practice, generic pedagogy as well as reading-specific instructions. A detailed analysis was carried out at item level to compare various aspects of classroom practices across group. The effect size is denoted as 'r' and its value varies from 0 to 1. The range was used to define small (less than 2); medium (starts with 3); and large (starts with 5) effect size. The large effect size indicates a bigger magnitude of difference. Additional analyses were carried out to examine the characteristic of classroom practices by defining quality as three broader categories. Any score less than or equal to 1.5 was defined as 'weak practice', a score between 1.51 and 2.5 was defined as 'mediocre practice' while a score of more than 2.5 was categorized as 'good practice'.

The parents' questionnaire was analyzed to represent 'home literacy environment' in the sampled children's home. The mean score was computed to present an overall home literacy environment across groups. Additionally, comparisons were made at item level to present a micro picture of home literacy environment in non-SRP and SRP groups. In order to make comparison t-test was employed. However, the sample size in this category was too small (n=56) to make any viable inferences.

QUALITATIVE APPROACH

In addition to quantitative analysis, the study also aimed to conduct an in-depth analysis of various

Table 5: Qualitative approach: research participants

Research Participants	Number
Teachers	60
HM	25
SRP (field)	4
SRP (Central)	4
GTs	4
Government District Official	7
Government Central office	3 (BoC, Curriculum Wing, STBB)

aspects of SRP interventions towards enhancing reading skills in early years. The quantitative arm of this study attempted to answer research question 1 which intended to explore the efficacy of SRP introduced interventions in enhancing children’s reading abilities in the target districts of the province of Sindh. Research questions 2-7 are addressed through qualitative design which tried to explore in-depth the deeper factors contributing towards the success (or non-success) of intervention. The qualitative part explored such aspects as design of innovation, the implementation processes, the role of various actors at various stages of the project, structural issues and the possibility of sustainability and scalability of the innovation. The qualitative elements of this research study, therefore, bears significant importance as it does not only explores the reasons for success, but for non-success as well. It is a major aim of the research to investigate the possibility of sustainability and scalability of the innovation, therefore, emphasis was laid on exploring the in-project factors as well as the out-project factors e.g. structural and policy issues.

Research Participants

The section presents a description of the research participants selected for the qualitative arm of this study. For qualitative research, there were three layers of respondents: school (headteacher and teacher), taluka/district (guide teachers, SRP district officials, TEOs and DEOs), central (SRP central office

representatives and government officials at central level). Table 5 presents details of participants who were recruited for qualitative data collection.

Data Collection Methods

This section presents the process of qualitative data collection along with the details of data collection methods, tools and their administration to the research participants.

Focused group discussions (FGDs)

At school level, the HTs and teachers were interviewed through FGDs to explore the out-of-classroom factors complementing or hindering the in-classroom factors and intervention process. Most importantly the teachers of Grades 1 and 2 who have been given training were the key persons to be interviewed to know the effect of the intervention. Their views about what works and what does not work was crucial to draw inferences on how and why certain parts of intervention work or do not work as envisaged.

Semi-structured interviews

The quantitative information gathered in the first stage was complemented by the detailed qualitative interviews with district government officials and district SRP officials. The interviews helped in gathering the data regarding the nature of training imparted to the teachers, the monitoring mechanisms and classroom supports along with other project related interventions.

Interviews were carried out with the SRP project staff at central and district level to understand the details of design and implementation of TLA model and GT model. Since the TLA model was implemented a year before and the GT model was just finished at the time of data collection, there was more recent information available about the GT model’s effect.

Both models of SRP intervention (GT and TLA) relied heavily on the support of the regular government structure of the education department; therefore, the government officials (district and provincial levels)

were also interviewed to explore the facilitating and hindering factors emanating from the structure and official processes, which affected the success of the intervention. In addition, government officials were also interviewed to explore the sustainability and scalability factors for example the possibility or impossibility of regularizing the 'Guide Teachers or Guide Schools' were explored.

Data Collection Tools

In order to develop tools, detailed discussions with the SRP team were carried out to understand program modalities, program intervention, stakeholders and other activities. Specific information about program intervention was also gathered to enhance researchers' understanding. Additionally, a reconnaissance visit to some SRP schools was also carried out to better understand the context. The information was helpful in developing appropriate tools before the start of the field work.

In order to develop tools specific to the research project, the research team had several discussions to extract themes from the research questions. Based on the extracted themes, first hand data collection tools were developed by the research team in order to capture various stakeholders' data. For this study, interview guides were developed to explore research participants' understanding regarding the research questions. A guide for FGD was developed to understand teachers', HTs' and guide teachers' viewpoints on various aspects related to the SRP intervention. Initially, the research team planned to conduct semi-structured interviews with head teachers and FGDs with teachers separately; however, initial field interactions revealed that it would be better if head teachers and teachers were interviewed together in a focus group situation. Further, two semi-structured interview guides were developed to conduct interviews with government officials (at district and central level) and SRP officials (at district and central level).

Data Collection Procedures

The fieldwork was carried out between September, 2017 and January, 2018 across five districts of Sindh. Data collection timelines were developed in coordination with the SRP team to schedule the field work.

At school level, all team members were actively engaged in data collection process for both quantitative and qualitative parts; however, it is important to note that the PI and Co-PI accompanied research team members during the data collection process and shuttled between the team to have maximum first-hand experience and guide the team. In order to overcome language barriers in rural districts, a good number of research team members were chosen from amongst native Sindhi speakers who were well trained in the data collection tools.

At central and district levels, semi-structured interviews were conducted with the government and SRP officials. These interviews were mainly conducted by the PI in order to capture maximum information at the strategic level. The close supervision of the PI and Co-PI ensured the proper collection of data and its quality.

Data Analysis

The analysis of qualitative data began with the transcription of interviews. During the data collection process, the interview sessions were recorded with the permission of interviewees and were later transcribed for the purpose of analysis. The majority of the interviews were transcribed by research assistants who were well-aware of the research project as well as of the transcription process. Interview sessions were conducted in Urdu or Sindhi. All interviews were transcribed in the original language to prevent potential harm to credibility of the data and data loss during translation.

After interviews' transcription, the research team began the process of coding. In this stage, transcripts were analyzed in detail and sections of text units (words, phrases or sentences) which referred to a thematic idea were extracted and assigned different

codes so that they could easily be retrieved at a later stage for further comparison and analysis. The codes provided the basis to move forward with thematic analysis.

The codes that emerged during the second stage were then clustered based on similar themes. The codes were further analyzed to look for their relationship with the research questions and the connection among codes in order to come up with different themes. As a result, different themes and sub-themes emerged corresponding to the research questions which were further examined at two stages: a) within-case analysis and b) cross-case analysis, as suggested by Miles and Huberman (1994).

For ‘within-case analysis’, the matrices which were developed according to the participants were looked at cumulatively across themes under each of the research participants and write-ups were developed for each group of research participants. For ‘cross-case analysis’, the write-ups were further looked at to extract summative sub-themes under each of the major research themes. Inferences were made to enable reflective comments to be written under each of the sub-themes.

LIMITATIONS OF THE STUDY

This section addresses the four main limitations of the study which include: (i) different cohorts of children for pre- and post-tests for Cohort-1, (ii) EGRA vs. observational data, (iii) parents’ sample size, and (iv) qualitative data limited to SRP schools.

Different cohort of children for pre- and post-tests: As mentioned earlier a pre-post comparison design was employed for analyzing Cohort-1 (TLA model) data. EGRA-based data which have already been collected by the SRP team were analyzed as part of the study. While comparison schools were sampled from the ‘matched’ neighbourhood, different cohorts of Grade 3 children participated in pre- (baseline) and post- (midline) tests. In this case, the claim about ‘progression overtime’ needs to be interpreted with caution. The

results are indicative for exploring the efficacy of intervention in improving children’s reading skills in sample schools.

EGRA vs. observational data: EGRA has been developed for and administered in Grade 3 by the SRP team. The research team did not deviate from the norm while using the assessment for Cohort-2 (GT model). On the other hand, for pragmatic reasons, teaching practices were observed in Grades 1 and 2 – as teachers of these grades were trained for reading instructions. Observation of Grade 3 would not be helpful in making ‘inferences’ about the efficacy of intervention in enhancing classroom practices in target schools. Arguably, ERCOP ratings are used as a proxy of the teaching and learning environment – both generic and reading instruction which children in the two groups (SRP and non-SRP) would have experienced before they progressed to Grade 3. The background needs to be kept in consideration for interpreting observational results.

Parents’ sample size: The sample size for home literacy environment assessments (parents) was an obvious limitation for the study. A larger group of parents would have further strengthened the quantitative analysis. It is important to note that about one-tenth of the sample children’s parents participated in the study. There were different reasons for the low participation rate of parents: *firstly*, the data were collected during school timings which clashed with fathers’ ‘working time’ and mothers’ ‘prime home-chores time’. *Secondly*, parents were invited based on children’s random selection for the test on the day of data collection. This ‘short-notice’ contributed to the limited participation of parents yet it was unavoidable. *Thirdly*, despite repeated dialogue between the researchers and the contact person (usually the head teacher) in school for inviting ‘relevant’ parents for interview those were called whose children were not sampled for the study. This small sample did not allow developing any sophisticated statistical models to demonstrate, if any, contribution of home-literacy practices in students’ reading skills as a predictor or

mediator. That said, SRP has not directly involved parents in reading intervention. They were recruited to explore any 'unplanned' contribution SRP parents (or non-SRP parents) may have made in providing the literacy environment at home.

Qualitative data limited to SRP schools: Qualitative data were confined to the SRP group. It would have been valuable to have gathered information from non-SRP stakeholders in order to explore their ideas and

experiences of teaching languages in the early years. However, this comparison at qualitative level was beyond the scope of qualitative analysis which set out to explore differences in SRP implementation models, bottlenecks in implementation of intervention, efficacy of technology-based assessment, and prospects for sustainability and scalability of the intervention.

NOTES:

1. Engagement with parents/ community is not part of the SRP intervention. However, it is an important factor to investigate. CMP's interventions regarding community mobilization could provide a good link in this regard.
2. Technology enabled assessment is not a standalone intervention. It is an integral part of the SRP intervention but needs to be spotlighted.
3. Shanahan & Barr suggest this on the basis of the incongruence between classroom teaching methods and those of the intervention. For better understanding refer to their article.
4. Only one participant (i.e. a father) from the SRP group reported having attained a Master's degree. The case was merged with Bachelor's degree.

4

Findings: Quantitative Analysis

This chapter presents findings generated through quantitative and qualitative data. The results are organised to respond to the major questions developed for the study and are presented in two parts – Cohort-1 and Cohort-2.

PART-A: COHORT-1 (TLA MODEL)

OVERALL EGRA RESULTS: NON-SRP VS SRP – COHORT-1

What explains the variation among the students learning outcomes (zero scorers vs. readers)? -*There could be several factors affecting the learning outcomes such as pedagogic materials, teaching content, quality of instruction, physical infrastructure, parental engagement, etc.*

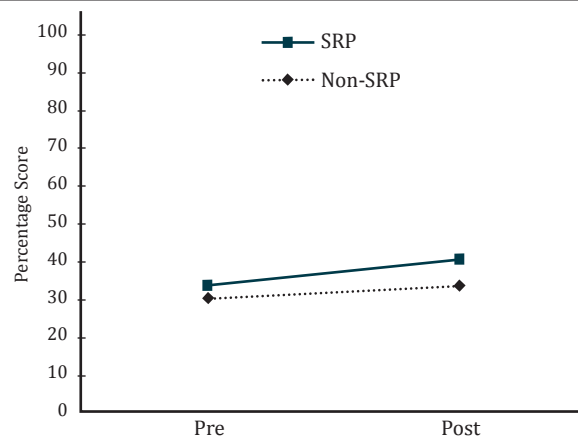
This section presents results of EGRA assessment for cohort-1. The section is organized into 3 main subsections including (i) overall comparison (ii); comparing performance on reading domains; and (iii) zero vs. non-zero readers on reading domains.

In order make the report more ‘reader-friendly’ only alpha value (p), effect size (r) and percentage differences (for significant differences only) are presented in the text. Detailed results of statistical analysis for EGRA (both cohort - 1 and cohort - 2) are appended for readers’ reference (Appendix).

EGRA Scores: Overall Comparisons

Figure 6 depicts an overall comparison between (SRP vs. Non-SRP) and within group (progression over time). Results revealed that on average the two groups exhibited similar performance at the outset ($p>0.05$). After the intervention, the two groups show significant difference in reading performance which was in favor of SRP ($p<0.001$, $r=0.15$). In terms of percentage

Figure 6: Students’ performance on EGRA in cohort-1: Difference across and progression within group



difference, SRP scored 6% higher than their non-SRP counterparts. While SRP intervention seems to have contributed significantly in improving reading skills of participant children, the magnitude of the difference is rather small.

Evidently, both groups have made progression over time in their reading skills. However, SRP has shown statistically significant improvement ($p<0.001$, $r=0.16$) with a small effect size. On the other hand, the progression made by the non-SRP is negligible ($p>0.05$). SRP has made a progression of 7% in average EGRA scores from pre- to post-test while non-SRP only made a progression of 3%.

EGRA Scores: Comparing Performance on Reading Domains

The total EGRA score does not provide details about students’ performance on various aspects of reading skills (i.e. reading domains) assessed in the test. Therefore, mean scores were computed for six domains to compare performance across groups and progression within each group over time.

EGRA scores: comparison by groups – non-SRP vs. SRP

This section presents comparative overview of students’ performance on EGRA before and after the intervention.

Comparison of groups on pre-test: Figure 7 presents a comparative overview of students’ performance on pre-test across group. Of the six comparisons made at pre-test the two groups evenly matched on four with no significant difference ($p>0.05$). However, the SRP group started with better performance in two reading domains including - non-word fluency ($p<0.05$; $r=0.12$) and expressive vocabulary ($p<0.05$; $r=0.10$). Nevertheless, the magnitude of the difference was small for both domains. In terms of percentage differences, SRP started with 6% higher score in reading non-words in the given time (*non-word fluency*) and 5% in describing the given picture by identifying object, subject and action (*expressive vocabulary*).

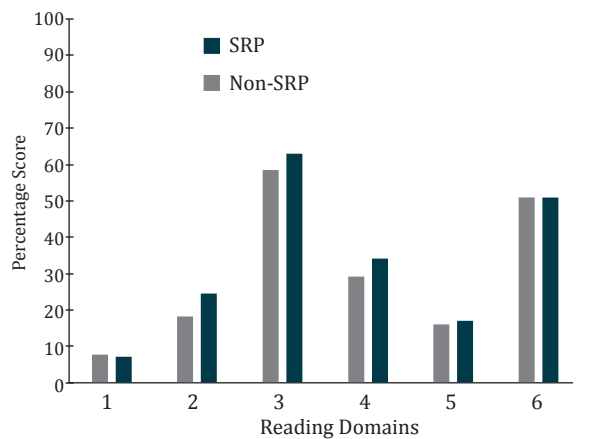
The pattern of performance across groups on six domains revealed that Grade 3 children exhibited better skills in expressive vocabulary by describing the action, subject and object presented to them in picture form. They also seem to have better listening skills as compared to the other domains (save *expressive*

vocabulary). On the other hand, performance on phonemic awareness stays at the bottom end indicating that children found it challenging to identify phonemes in the given words - in this case the last sound. Non-word fluency which required children to read ‘non-words’ exhibits children’s difficulty in reading these words accurately in the given time (1 minute). Though SRP has an edge over non-SRP in this domain, none of them managed to read more than 25% of the words accurately in a minute. Interestingly, children in both groups show better performance in oral reading than reading comprehension. In other words, children were able to read the given paragraph without necessarily comprehending the message given in the text.

Comparison of groups on post-test: Figure 8 presents a comparative overview of students’ performance on post-test across group. It is evident that, on average, SRP shows a consistent edge on non-SRP across all six domains; however, the differences were found to be significant in four domains (*phonemic awareness, non-word fluency, oral reading and reading comprehension*).

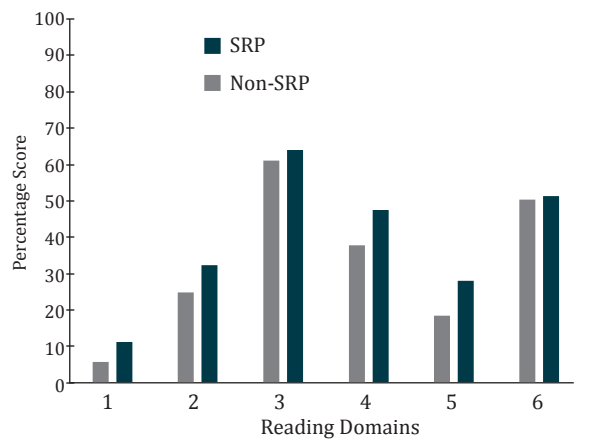
Interestingly, phonemic awareness still stayed at the bottom end, but the SRP group shows a significantly better performance in identifying the last sound in given words as compared to the non-SRP counterparts

Figure 7: Students’ performance on EGRA in pre-test across group



1 = Phonemic awareness | 2 = Non-word Fluency
 3 = Expressive Vocabulary | 4 = Oral Reading
 5 = Reading Comprehension | 6 = Listening Comprehension

Figure 8: Students’ performance on EGRA in post-test across group



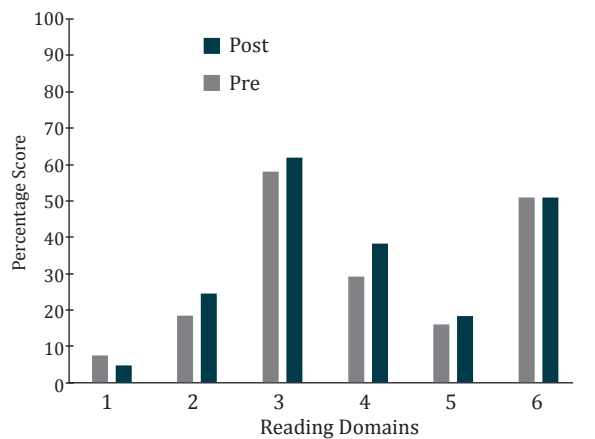
1 = Phonemic awareness | 2 = Non-word Fluency
 3 = Expressive Vocabulary | 4 = Oral Reading
 5 = Reading Comprehension | 6 = Listening Comprehension

($p < 0.001$, $r = 0.18$). Moreover, the SRP group shows a significantly better performance in accurately reading 'non-words' in the given time as compared to the non-SRP group ($p < 0.01$, $r = 0.14$). Similarly, the SRP group exhibited better skills in reading the given text ($p < 0.01$, $r = 0.16$) accurately in one minute followed by better comprehension ($p < 0.001$, $r = 0.20$) as well. That said, the magnitude of the differences were small for all comparisons. Putting it differently, children in SRP schools scored 6% (*phonemic awareness*), 8% (*non-word fluency*), and 10% (*oral reading fluency, reading comprehension*) higher than the non-SRP group. It is important to note that reading skills are better than comprehension regardless of group. Interestingly, listening comprehension skills do not seem to be affected by the SRP intervention and neither does the expressive vocabulary. A scan of the performance across domains revealed a similar pattern as for the pre-test where phonemic awareness stayed at the bottom followed by non-word fluency, reading comprehension, oral reading, listening comprehension and expressive vocabulary.

Progression in EGRA scores over time

This section presents progression over time in EGRA mean score for six reading domains from pre- to post-test in the SRP and non SRP group separately.

Figure 9: Non-SRP group: progression from pre- to post-test



1 = Phonemic awareness | 2 = Non-word Fluency
 3 = Expressive Vocabulary | 4 = Oral Reading
 5 = Reading Comprehension | 6 = Listening Comprehension

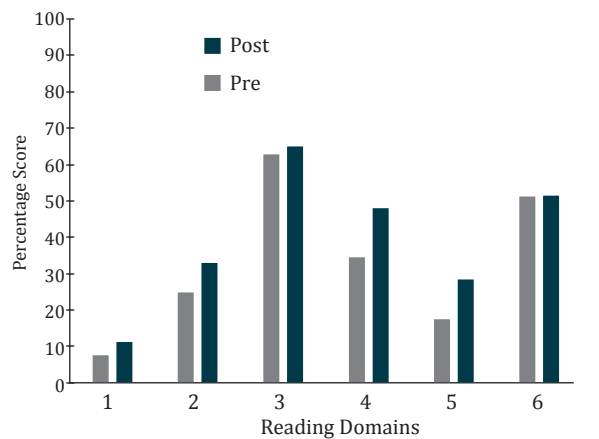
Progression over time – Non-SRP group: Figure 9 presents progression over time for the non-SRP group. Of the six comparisons made to gauge progression in the non-SRP group three were found to be significant (i.e. *non-word fluency, oral reading fluency*).

In the non-SRP group, children show a significant improvement in reading non-words accurately in the given time ($p < 0.01$, $r = 0.12$) and reading the given text in a specified time ($p < 0.01$, $r = 0.12$). However, the magnitude was rather small with a percentage difference of 7% (*non-word fluency*) and 10% (*oral reading fluency*). No significant progress was observed in comprehending the text they have read (*reading comprehension*) and text read to them (*listening comprehension*).

Progression over time – SRP group: The SRP group shows significant improvements over time in all reading domains except expressive vocabulary and listening comprehension as shown in Figure 10.

A significant progression has been observed in the SRP group in decoding phonemes in given words ($p < 0.05$, $r = 0.10$); reading non-words accurately in the given time ($p < 0.01$, $r = 0.14$), reading the given text in a specified time ($p < 0.01$, $r = 0.18$); and, comprehending that text ($p < 0.001$, $r = 0.19$). In terms of percentage

Figure 10: SRP group: progression from pre- to post-test



1 = Phonemic awareness | 2 = Non-word Fluency
 3 = Expressive Vocabulary | 4 = Oral Reading
 5 = Reading Comprehension | 6 = Listening Comprehension

differences progression made in various reading domains ranged from 4% (*phonemic awareness*), 8% (*non-word fluency*), 11% (*reading comprehension*) and 14% (*oral reading fluency*).

EGRA Scores: Zero vs Non-Zero Readers on Reading Domains

Zero readers – comparison across group

The data were reanalyzed to compare ratio of zero readers (those who cannot read a single word accurately) across group and overtime. Table 6 presents a summary of zero readers’ comparison within and across groups.

Zero vs Non-zero: comparison of pre-test across group:

Of the six comparisons made at pre-test the two groups evenly matched on five reading domains with no significant difference ($p > 0.05$). However, the SRP group started with better performance in oral reading fluency. In other words, at the outset, the percentage of SRP children who were not able to read a single word accurately in the given text was 15%, significantly less as compared to their non-SRP counterparts [$p < 0.001$, $j = 0.14$]. Nevertheless, the magnitude of the difference was small. Evidently, an overwhelming majority (more than 80%) of children fell in the category of zero readers in phonemic awareness regardless of groups. Moreover, two-thirds of the children across groups did not manage to comprehend the given text. In non-word fluency, half of the children (non-SRP= 54%; SRP= 46%) scored zero across two groups. On the other hand, both groups exhibited the lowest ratio of zero readers on expressive vocabulary (less than 10%) followed by listening comprehension (less than 20%).

Zero vs Non-zero: comparison of post-test across group:

Of the six comparisons made at post-test, the differences were found to be significant on four reading domains. It is evident that the ratio of zero readers and readers remain the same across groups for *expressive vocabulary* and *listening comprehension* ($p > 0.05$ for both domains). The intervention seems to have influenced *phonemic awareness* significantly [$p < 0.001$, $j = 0.20$] with small magnitude. Similarly, the differences were found to be significantly in favour of SRP for *non-word fluency* [$p < 0.01$, $j = 0.14$]; *oral reading fluency* [$p < 0.01$, $j = 0.15$]; and *reading comprehension* [$p < 0.001$, $j = 0.16$]. In terms of percentage difference, the SRP intervention managed to reduce the frequency of zero readers by 17% in decoding the phonemes in the given words (*phonemic awareness*), 14% in reading the non-word accurately in the given time (*non-word fluency*) 14% in reading the given text accurately in given time (*oral reading fluency*), and 16% in comprehending that text (*reading comprehension*).

Zero readers – progression over time

This section presents progression over time in terms reducing the frequency of zero readers in the SRP and non SRP group separately.

Zero vs Non-zero in the non-SRP group – progression over time:

Of the six comparisons made to gauge progression in the non-SRP group, only one was found to be significant (i.e. *oral reading fluency*). In the non-SRP group, frequency of zero-readers significantly reduced from pre- to post-test [$p < 0.01$, $j = 0.16$] though the magnitude of the difference was small. In terms of percentage difference, the numbers of zero readers reduced by 16% in the domain of oral reading fluency

Table 6: Zero readers - comparison across group and progression over time

Reading Domains	Non-SRP		SRP	
	Pre %(n)	Post %(n)	Pre %(n)	Post %(n)
Phonemic Awareness	82% (190)	87% (202)	83% (275)	70% (220)
Non-Word Fluency	54% (125)	47% (108)	46% (155)	33% (105)
Expressive Vocabulary	08% (18)	08% (18)	04% (13)	07% (21)
Oral Reading Fluency	57% (132)	41% (95)	42% (140)	27% (85)
Reading Comprehension	68% (158)	65% (151)	66% (220)	49% (154)
Listening Comprehension	15% (35)	18% (42)	16% (54)	17% (52)

overtime. No significant decrease in the number of zero readers has been observed in decoding phonemes, fluency in reading non-words, stating action depicted in a picture along with identification of subject and object, comprehension of the given text and comprehending the text read to them.

Zero vs Non-zero in the SRP group – progression over time: Of the six comparisons made to gauge progression in the SRP group, four reading domains exhibited a significant reduction in the number of zero readers over time. In the SRP group, frequency of zero-readers has significantly reduced from pre- to post-test in decoding phonemes in the given words [$p < 0.001$, $j = 0.15$]; reading non-words accurately in the given time [$p < 0.01$, $j = 0.14$]; reading the given text accurately in a specified time [$p < 0.001$, $j = 0.16$]; and, comprehending that text [$p < 0.001$, $j = 0.18$]. That said, the magnitude of the differences was small for all comparisons. In terms of percentage difference, the SRP intervention has managed to reduce the frequency of zero readers by 13% in decoding the phonemes in the given words (*phonemic awareness*), 13% in reading the non-word accurately in the given time (*non-word fluency*) 15% in reading the given text accurately in the given time (*oral reading fluency*), and 17% in comprehending that text (*reading comprehension*). On the other hand, SRP intervention does not seem to have contributed in reducing zero readers in expressive vocabulary (describing the given picture by identifying action, subject and object) and listening comprehension (comprehending the text read to them).

EGRA scores: comparing the distribution of readers across groups and over time

After analysis of mean scores and zero readers ratio, four performance categories were defined to demonstrate distinguished levels of 'readers': level 1 (1% to 25.9% score); level 2 (26% to 50.9% score); level 3 (51% to 75.9% score); and level 4 (76%-100% score). Performance of those children who 'qualified' to be readers (non-zero readers) was compared against these four levels for overall scores in six EGRA domains. A detailed analysis was undertaken to

determine students' performance in the six domains defined for EGRA. This section presents the results of readers' scores against the four levels of performance across group.

Phonemic awareness: Table 7 demonstrates the distribution of readers for pre- and post-tests across two groups for decoding the last sound of ten words. Before intervention, both non-SRP and SRP followed the same trend except minor discrepancies in levels 2, 3 and 4. However, a visible trend can be observed in readers' distribution for post-test which is in favor of SRP. That said, even after the intervention a majority of readers remained at the bottom end – who managed to decode phonemes only in one to two words. Only a fraction qualified for the highest level and identified phonemes in eight to ten words (level 4).

Non-word fluency: Table 8 illustrates the distribution of readers before and after the intervention across two groups for reading forty non-words. As discussed earlier, SRP has started with better performance and the trend is visible in pre-test distribution especially at the highest ends. A similar trend is observed in post-test readers' distribution for non-word fluency. A little more than one-tenth of children fell into the highest level in SRP while less than one-third reached level 4 in non-SRP (thirty one to forty words accurately in one minute).

Expressive vocabulary: Table 9 displays the distribution of readers for pre- and post-test across two groups for describing the action, subject and object presented in ten pictures. Evidently, this is one of the highest scoring domains across group and the trend is maintained over time where the ratio of readers is higher in the highest level (quite a few managed to describe eight to ten pictures). In other words, this was one of the easiest reading domains of EGRA for the sample children. Visibly, both non-SRP and SRP followed a similar trend in pre- and post-test except level 3 in post-test where the ratio of SRP children is considerably higher in level 3 (can describe six to seven pictures) than their non-SRP counterparts.

Table 7: Phonemic awareness -distribution of readers across group for pre- and post-test

	Non-READERS		READERS			
	Levels of Reading	Zero Reader	Level 1 (1-25%)	Level 2 (26-50%)	Level 3 (51-75%)	Level 4 (76-100%)
Pre-test	SRP Children (%)	82	6	6	4	2
	Non-SRP Children (%)	82	6	7	2	3
Post-test	SRP Children (%)	70	14	9	3	4
	Non-SRP Children (%)	87	5	4	1	3

Table 8: Non-word fluency – pre-test distribution of readers across group

	Non-READERS		READERS			
	Levels of Reading	Zero Reader	Level 1 (1-25%)	Level 2 (26-50%)	Level 3 (51-75%)	Level 4 (76-100%)
Pre-test	SRP Children (%)	47	10	22	13	8
	Non-SRP Children (%)	54	12	21	12	1
Post-test	SRP Children (%)	33	15	20	18	14
	Non-SRP Children (%)	47	11	19	15	8

Table 9: Expressive vocabulary – distribution of readers across group for pre- and post-test

	Non-READERS		READERS			
	Levels of Reading	Zero Reader	Level 1 (1-25%)	Level 2 (26-50%)	Level 3 (51-75%)	Level 4 (76-100%)
Pre-test	SRP Children (%)	4	2	28	34	32
	Non-SRP Children (%)	8	4	28	31	29
Post-test	SRP Children (%)	7	4	17	31	41
	Non-SRP Children (%)	8	6	20	23	43

Table 10: Oral reading fluency – distribution of readers across group for pre- and post-test

	Non-READERS		READERS			
	Levels of Reading	Zero Reader	Level 1 (1-25%)	Level 2 (26-50%)	Level 3 (51-75%)	Level 4 (76-100%)
Pre-test	SRP Children (%)	42	14	12	9	23
	Non-SRP Children (%)	57	5	8	11	19
Post-test	SRP Children (%)	27	12	12	12	37
	Non-SRP Children (%)	41	8	13	13	25

Table 11: Reading comprehension – distribution of readers across group for pre- and post-test

	Non-READERS		READERS			
	Levels of Reading	Zero Reader	Level 1 (1-25%)	Level 2 (26-50%)	Level 3 (51-75%)	Level 4 (76-100%)
Pre-test	SRP Children (%)	66	10	8	8	8
	Non-SRP Children (%)	68	10	8	6	8
Post-test	SRP Children (%)	49	12	12	12	16
	Non-SRP Children (%)	65	10	7	7	10

Table 12: Listening comprehension – distribution of readers across group for pre-and post-test

	Non-READERS		READERS			
	Levels of Reading	Zero Reader	Level 1 (1-25%)	Level 2 (26-50%)	Level 3 (51-75%)	Level 4 (76-100%)
Pre-test	SRP Children (%)	16	0	28	43	13
	Non-SRP Children (%)	15	0	32	38	15
Post-test	SRP Children (%)	16	0	28	40	16
	Non-SRP Children (%)	18	0	31	32	19

Oral reading fluency: Table 10 shows the distribution of readers for pre- and post-test across two groups for reading a given paragraph of sixty words in the given time. SRP started with a better performance than non-SRP and the trend is visible in most of the levels. In pre-test almost a quarter of the sample children read forty-two to sixty words accurately in a minute while less than one-fifth reached that level for non-SRP. A visible gain was observed in SRP after the intervention and it maintained its significant edge on non-SRP. Expressly, after the intervention more than one-third of the SRP children managed to reach forty-two to sixty words accurately in a minute. On the other hand, only a quarter of non-SRP children qualified for the highest level.

Reading comprehension: Table 11 explains distribution of readers for pre- and post-test scores across two groups for responding to five comprehension questions asked to evaluate their skills in comprehending the text they read. Before intervention, both non-SRP and SRP followed the same trend except for minor discrepancy in levels 3. However, a visible trend can be observed in readers' distribution for post-test which is in favor of SRP and this gap has gradually increased across four levels. That said, only one-tenth of the non-SRP and a little more than one-tenth of the SRP sample children managed to comprehend the text well by responding to four to five questions accurately.

Listening comprehension: Table 12 shows the distribution of readers for pre- and post-test across two groups for responding to three comprehension questions posed to children after a short paragraph was read to them. Regardless of group and time, the

majority of the children responded to two questions accurately. Interestingly, the two questions which were responded to accurately by the majority were pitched at 'knowledge' level while the third question required children to make 'inferences'.

In summary, before intervention both non-SRP and SRP followed almost the same trend. Nevertheless, the distribution of readers on six domains of EGRA highlights that the ratio of zero-readers has gone down over time in both groups. However, the ratio of fluent readers varied across six domains whereby expressive vocabulary stayed at the top end while phonemic awareness remained at the bottom end. While it is important to 'monitor' the drop in zero-readers it is imperative to scrutinize the distribution of readers across four levels. The latter may help to make targeted amendments in various components of the intervention (i.e. teachers' training, follow-up, adequacy and quality of material, validity of assessments).

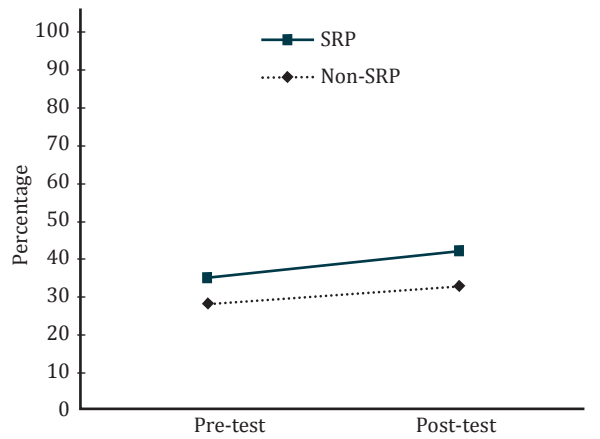
COMPARISON BY LANGUAGE: NON-SRP VS SRP – COHORT-1

This section presents results of a comparison of EGRA assessment for cohort-1 by language i.e., EGRA-Sindhi and EGRA-Urdu. These sections are further divided into 3 main subsections: (i) overall comparison; (ii) comparing performance on reading domains over time; and (iii) zero vs. non-zero readers on reading domains.

EGRA-Sindhi – Overall Comparisons

Figure 11 depicts an overall comparison between (SRP vs. Non-SRP) and within (progression over time) group. Results revealed that SRP started off with a better performance (6%) than non-SRP with a

Figure 11: Overall comparisons on EGRA-Sindhi: Differences across and progression within group



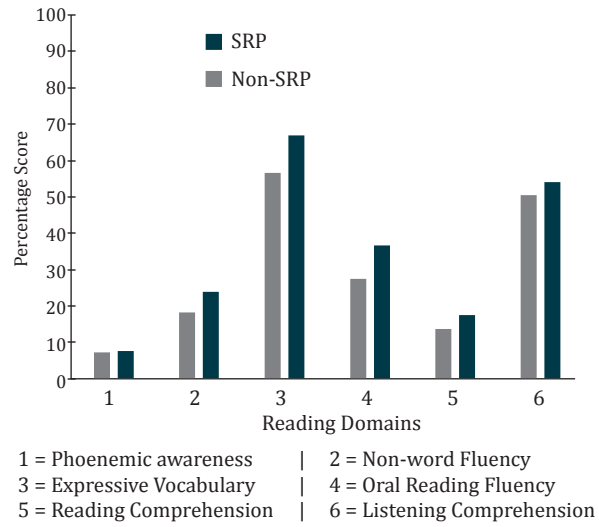
significant difference ($p < 0.01$, $r = 0.14$). After the intervention, both groups show improvement; however, SRP maintains its significant edge ($p < 0.001$, $r = 0.19$) with 9% higher score than their non-SRP counterparts. While SRP intervention seems to have contributed significantly in improving reading skills of participant children, the magnitude of the difference is rather small.

Evidently, both groups have made significant progress over time in their reading skills with significant differences for the non-SRP ($p < 0.05$, $r = 0.12$) as well as for the SRP ($p < 0.001$, $r = 0.18$) group. However, in terms of percentage difference, SRP shows better progression with the difference of 7% overtime as compared to their non-SRP counterparts (4%). While both groups made progress over time, the magnitude of the difference fell in the category of small effect size for both.

EGRA Sindhi – Comparing Performance on Reading Domains

The total EGRA-Sindhi score does not provide details about students’ performance on various aspects of reading skills (i.e. reading domains) assessed in the test. Therefore, mean scores were computed for six domains to compare performance across groups and progression within each group over time.

Figure 12: Children’s performance on EGRA-Sindhi in pre-test across group



Comparisons between groups: Non-SRP vs SRP

This section presents comparative overview of students’ performance on EGRA-Sindhi before and after the intervention.

Comparison of groups on pre-test: Figure 12 presents a comparative overview of students’ performance on pre-test across group.

Of the six comparisons made at pre-test, both groups have shown comparable performance only on the domain of *phonemic awareness*. For the rest of the five domains, SRP shows an edge over the non-SRP group in pre-test; however, the difference was found to be significant only for three domains including- *non-word fluency* ($p < 0.05$, $r = 0.11$); *expressive vocabulary* ($p < 0.001$, $r = 0.22$), and *oral reading fluency* ($p < 0.05$, $r = 0.12$). Nevertheless, the magnitude of the difference was small for all three domains. In terms of percentage difference at the outset, SRP scored 6% higher in *non-word fluency*, 11% higher in *expressive vocabulary* and 9% higher in *oral reading fluency*, as compared to their non-SRP counterparts.

Before intervention, the pattern of performance across groups on six domains revealed that Grade 3 children in both groups have exhibited better skills in *expressive vocabulary* by describing the action, subject and object presented to them in picture form. They also seem to

have better *listening* skills as compared to the other domains. On the other hand, performance on *phonemic awareness* received the lowest scores indicating that children found it challenging to identify phonemes in the given words – in this case the last sound. Furthermore, in both groups, children scored relatively lower in *non-word fluency*, *oral reading fluency* and *reading comprehension* indicating that it was difficult for children to read the given words accurately in the given time, read the given text fluently in the given time and respond to the questions to comprehend the text after reading it.

Comparison of groups on post-test: Figure 13 presents comparative overview of students’ performance on post-test across group. It is evident that on average SRP shows a consistent edge on non-SRP across all six domains; however, the differences were found to be significant in four domains (*phonemic awareness*, *non-word fluency*, *oral reading fluency* and *reading comprehension*).

Interestingly, *phonemic awareness* still stayed at the bottom end, but the SRP group shows a significantly better performance in identifying the last sound in given words as compared to their non-SRP counterparts ($p < 0.001$, $r = 0.18$) with a 7% difference in average EGRA scores. Furthermore, EGRA score of

the SRP group was 9% better than the non-SRP group in accurately reading ‘non-words’ in the given time as compared to their counterparts ($p < 0.01$, $r = 0.14$). Similarly, the SRP group exhibited better skills in reading the given text accurately in one minute- *oral reading fluency* ($p < 0.01$, $r = 0.16$) followed by better *reading comprehension* ($p < 0.001$, $r = 0.20$). In terms of percentage differences, SRP scored 13% higher than their non-SRP counterparts in both *oral reading fluency* and *reading comprehension*. That said, the magnitude of the difference remained small for all comparisons which generated significant differences.

Interestingly, *listening comprehension* skills do not seem to be affected by the SRP intervention and neither do the *expressive vocabulary*. A scan of the performance revealed a similar pattern as for the pre-test where *phonemic awareness* received the lowest score followed by *reading comprehension*, *non-word fluency*, *listening comprehension* (oral reading fluency for non-SRP), *oral reading fluency* (listening comprehension for non-SRP), and *expressive vocabulary*.

Progression over time: from pre- to post-test

This section presents progression over time in EGRA-Sindhi mean score for six reading domains from pre- to post-test in the non-SRP and the SRP group separately.

Progression over time – non-SRP group: Figure 14 presents progression overtime for the non-SRP group. Of the six comparisons made to gauge progression in the non-SRP group, significant differences were found only on three domains (i.e. non-word fluency, oral reading fluency and reading comprehension).

In the non-SRP group, children show a significant improvement in reading non-words accurately in the given time ($p < 0.05$, $r = 0.12$), reading the given text accurately in the given time ($p < 0.01$, $r = 0.16$), and comprehending the given text ($p < 0.05$, $r = 0.11$); however, the magnitude is rather small in all three domains. In terms of percentage difference, students gained 6% in *non-word fluency*, 14% in *oral reading*

Figure 13: Students’ performance on EGRA-Sindhi in post-test across group

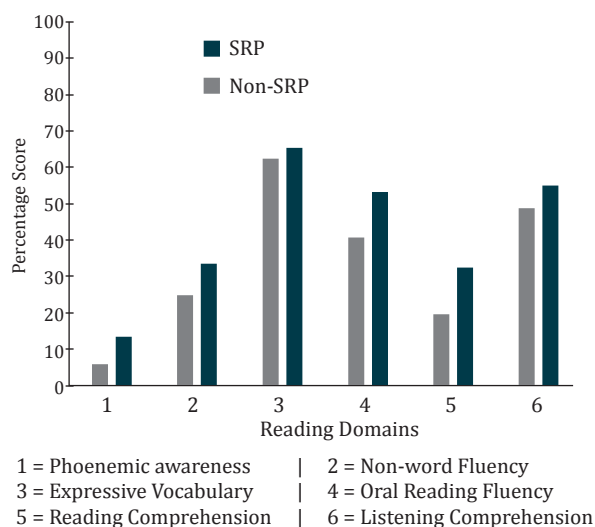
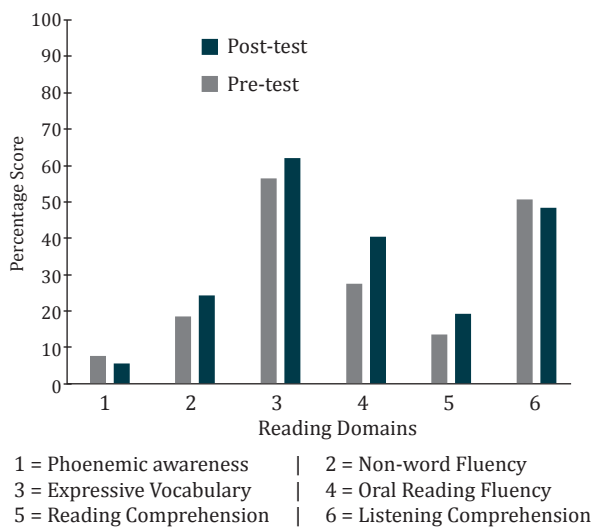


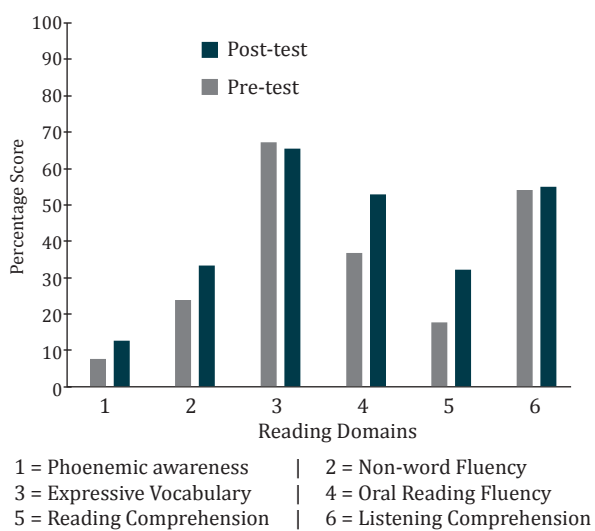
Figure 14: Non-SRP group: progression from pre- to post-test on EGRA-Sindhi



fluency, and 6% in reading comprehension over the period. Surprisingly, a marginal decline was observed children’s performance in identifying the last sound of the given words (*phonemic awareness*) and comprehending the text read to them (*listening comprehension*).

Progression over time – SRP group: The SRP group shows significant improvements over time in four domains including *phonemic awareness*, *non-word fluency*, *oral reading fluency* and *reading comprehension* as shown in Figure 15.

Figure 15: SRP group: progression from pre- to post-test on EGRA-Sindhi



From pre-to post-test, a significant progress has been observed in the SRP group in decoding phonemes in the given words – *phonemic awareness* ($p < 0.05$, $r = 0.12$), reading non-words accurately in the given time – *non-word fluency* ($p < 0.001$, $r = 0.16$), reading the given text accurately in the specified time – *oral reading fluency* ($p < 0.001$, $r = 0.20$), and comprehending that text – *reading comprehension* ($p < 0.001$, $r = 0.24$); however, the magnitude of the differences remained small for all four domains. In terms of percentage difference over time, students in the SRP group have gained 5% higher scores in *phonemic awareness*, 10% higher scores in *non-word fluency*, 16% higher scores in *oral reading fluency*, and 15% higher scores in *reading comprehension*. Performance of students seems to remain unaffected in identifying subject, object and action in the given picture (*expressive vocabulary*) as well as in comprehending the text read to them (*listening comprehension*).

EGRA Sindhi – Zero vs Non-Zero Readers on Reading Domains

Zero readers – comparison across group

The data were reanalyzed to compare the ratio of zero readers (those who cannot read a single word accurately) across group and over time. Table 13 presents a summary of zero readers’ comparison within and across groups.

Zero vs non-zero – comparison of pre-test across group:

Of the six comparisons made at pre-test the two groups evenly matched in two domains (*phonemic awareness* and *listening comprehension*). SRP started with a relatively low number of zero scorers on the rest of the four domains; however, the difference was found to be significant only for *oral reading fluency* [$p < 0.001$, $\phi = 0.24$]. To put it differently, at the outset, SRP had 24% less zero-readers as compared to the non-SRP group, which indicates that the percentage of children who were not able to read a single word accurately in the given text was significantly less in the SRP group as compared to their non-SRP counterparts. Nonetheless, the magnitude of the difference was found to be small. Evidently, a majority of children (more than 80%) fell in the category of zero readers in *phonemic awareness*

Table 13: Zero readers – comparison across group and progression over time on EGRA-Sindhi

Reading Domains	Non-SRP		SRP	
	Pre %(n)	Post %(n)	Pre %(n)	Post %(n)
Phonemic Awareness	81% (151)	86% (149)	82% (183)	65% (125)
Non-Word Fluency	57% (105)	48% (83)	50% (110)	31% (59)
Expressive Vocabulary	9% (16)	8% (14)	4% (9)	7% (14)
Oral Reading Fluency	59% (110)	40% (69)	35% (78)	21% (41)
Reading Comprehension	71% (132)	64% (111)	64% (141)	42% (80)
Listening Comprehension	15% (28)	20% (35)	15% (33)	16% (30)

regardless of groups. Moreover, almost two-thirds (SRP) and nearly three-quarters (non-SRP) of the children across groups did not manage to comprehend the given text in the specified time (*reading comprehension*). In *non-word fluency*, half (SRP) and more than half (non-SRP) of the children scored zero across the two groups. On the other hand, both groups show the lowest ratio of zero readers on *expressive vocabulary* (less than 10%) followed by *listening comprehension* (less than 20%).

Zero vs non-zero – comparison of post-test across group: Of the six comparisons made at post-test, SRP intervention seems to have contributed in reducing the number of zero readers with a significant difference in four reading domains. Evidently, the ratio of zero readers remained the same across groups for *expressive vocabulary* and *listening comprehension* ($p > 0.05$ for both domains). The intervention seems to have worked in reducing zero-readers significantly for *phonemic awareness* [$p < 0.001$, $\phi = 0.24$]; *non-word fluency* [$p < 0.01$, $\phi = 0.17$]; *oral reading fluency* [$p < 0.001$, $\phi = 0.20$]; and *reading comprehension* [$p < 0.001$, $\phi = 0.22$]; however, the magnitude of the difference was found to be small in all cases. In terms of percentage differences, SRP intervention had the most influence on reading comprehension (22%), followed by phonemic awareness (21%), oral reading fluency (19%) and non-word fluency (17%) in reducing the number of zero readers in SRP as compared to non-SRP.

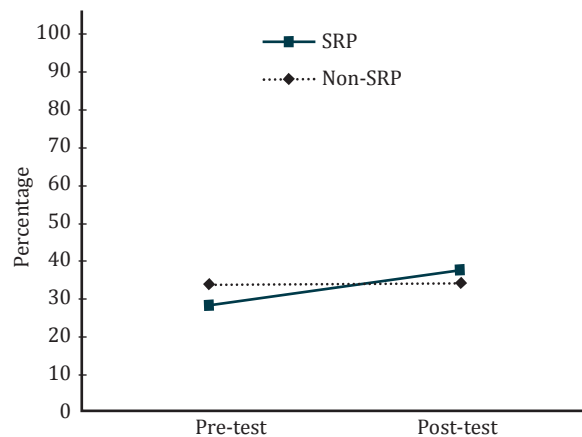
Zero readers – progression over time

This section presents progression over time in terms reducing the frequency of zero readers in the SRP and non-SRP group separately on EGRA-Sindhi.

Zero vs. non-zero in the non-SRP group – progression over time: Of the six comparisons made to gauge progression in the non-SRP group only one was found to be significant (i.e. *oral reading fluency*). In the non-SRP group, frequency of zero-readers significantly reduced from pre- to post-test [$p < 0.001$, $j = 0.20$]. The magnitude of the difference was found to be small. In terms of percentage differences, non-SRP managed to reduce zero readers in oral reading fluency by 19%. No significant progression was observed in decoding phonemes, fluency in reading non-words, stating action depicted in a picture along with identification of subject and object, comprehension of the text and comprehending the text read to them.

Zero vs non-zero in the SRP group – progression over time: Of the six comparisons made to gauge progression in the SRP group, four reading domains exhibited a significant gain over time. In the SRP group, the frequency of zero-readers significantly reduced from pre- to post-test in decoding phonemes in given words [$p < 0.001$, $\phi = 0.20$]; reading non-words accurately in the given time [$p < 0.001$, $\phi = 0.19$]; reading the given text accurately in the given time [$p < 0.01$, $\phi = 0.15$]; and, comprehending that text [$p < 0.001$, $\phi = 0.22$]. That said the magnitude of the difference was small. In terms of percentage differences, SRP intervention had the most influence on reading comprehension (22%), followed by non-word fluency (19%), phonemic awareness (17%), and oral reading fluency (14%). No significant progression was observed in *expressive vocabulary* (describing the given picture by identifying action, subject and object) and *listening comprehension* (comprehending the text read to them).

Figure 16: Overall comparisons on EGRA-Urdu: Differences across and progression within group



EGRA Urdu – Overall Comparisons

Figure 16 displays an overall comparison between (SRP vs. Non-SRP) and within group (progression over time). Results have shown that non-SRP had an edge over SRP at the outset by 5%; however, the difference was not significant ($p > 0.05$). Interestingly, after the intervention, SRP surpassed and achieved 3% better results than their non-SRP counterparts; however, the difference was still not significant. Although, the SRP group began with comparatively lower scores, the intervention seems to have helped in improving the reading skills with a significant difference over time ($p < 0.01$, $r = 0.17$). In terms of percentage difference, SRP show a gain of 7% in their scores from pre- to post-test. On the other hand, non-SRP who started off with better scores show a declining trend on post test.

EGRA Urdu – Comparing Performance on Reading Domains

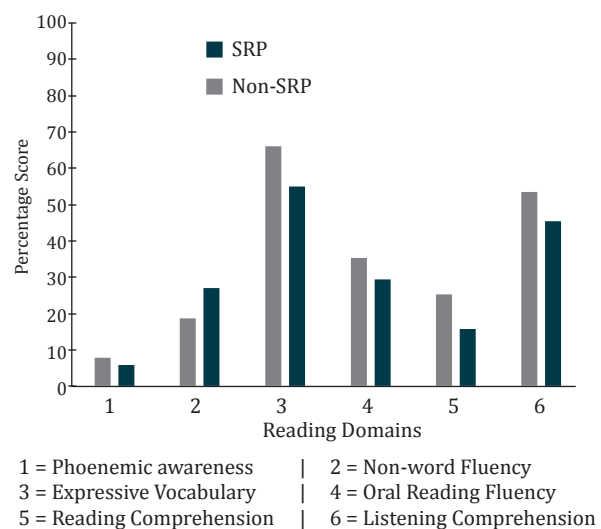
The total EGRA-Urdu score does not provide details about students' performance on various aspects of reading skills (i.e. reading domains) assessed in the test. Therefore, mean scores were computed for six domains to compare performance across groups and progression within each group over time.

Comparisons between groups: Non-SRP vs SRP

This section presents a comparative overview of students' performance on EGRA-Urdu before and after the intervention.

Comparison of groups on pre-test: Figure 17 presents a comparative overview of students' performance on pre-test across group. Of the six comparisons made at pre-test, the two groups matched only on one domain (i.e. *phonemic awareness*), whereas, in the other four reading domains including – *oral reading fluency*, *reading comprehension*, *listening comprehension*, and *expressive vocabulary* – the non-SRP group started off with a better performance; however, the difference was found to be significant only in *expressive vocabulary* ($p < 0.01$, $r = 0.23$) with 11% higher score. Nevertheless, the magnitude of the difference was small. On the other hand, in *non-word fluency*, SRP showed an edge over non-SRP at the outset; however, the difference was not found to be statistically significant.

Figure 17: Students' performance on EGRA-Urdu in pre-test across group



The pattern of performance across groups on six domains revealed that on pre-test Grade 3 children showed better skills in *expressive vocabulary* by describing the action, subject and object presented to them in a picture form. They also seem to have better listening skills as compared to the other domains. Interestingly, children in both groups have shown better performance in *oral reading fluency* than *reading comprehension* which indicates that children were able to read the given paragraph without necessarily comprehending the message given in the text. In *non-word fluency*, children faced difficulty in

reading the words accurately in the given time. Though SRP has an edge over non-SRP in this domain, none of them managed to score more than 30%. On the other hand, children's performance on *phonemic awareness* stays at the bottom end regardless of group, indicating that, like the Sindhi language children, they found it challenging to identify phonemes in the given words in Urdu as well. An overview of the performance across domains on pre-test revealed that *phonemic awareness* received the lowest scores followed by *reading comprehension*, *non-word fluency*, *oral reading fluency*, *listening comprehension*, and *expressive vocabulary*.

Comparison of groups on post-test: Figure 18 presents comparative overview of students' performance on post-test across group. It is evident that on average SRP shows a consistent edge on non-SRP across five domains; however, the difference was not found to be significant on any of the domains ($p>0.05$). On the other hand, non-SRP shows better performance than SRP in *listening comprehension*; however, the difference is still not significant ($p>0.05$).

The pattern of performance across groups on six domains revealed that on post-test Grade 3 children gained the lowest scores in *phonemic awareness* regardless of group. Moreover, the SRP group had an

edge over non-SRP in accurately reading 'non-word' in the given time (*non-word fluency*), reading the given text (*oral reading fluency*) accurately in the given time followed by *reading comprehension* which indicates that children's reading skills are better than comprehension regardless of group. Interestingly, non-SRP repeated the trend of surpassing SRP in comprehending the given text (*listening comprehension*); however, the difference was not significant. A scan of the performance across domains revealed a similar pattern as for the pre-test where *phonemic awareness* stayed at the bottom followed by *non-word fluency*, *reading comprehension*, oral reading, *listening comprehension* and *expressive vocabulary*.

Progression over time: from pre- to post-test

This section presents progression over time in EGRA-Urdu mean score for six reading domains from pre- to post-test in SRP and the non-SRP group separately.

Progression over time – the non-SRP group: Figure 19 presents progression over time for the non-SRP group.

Of the six comparisons made to examine the progression in the non-SRP group over time, improvement was observed in two domains (i.e., *non-word fluency* and *listening comprehension*); however, the difference was not found to be significant. Surprisingly,

Figure 18: Students' performance on EGRA-Urdu in post-test across group

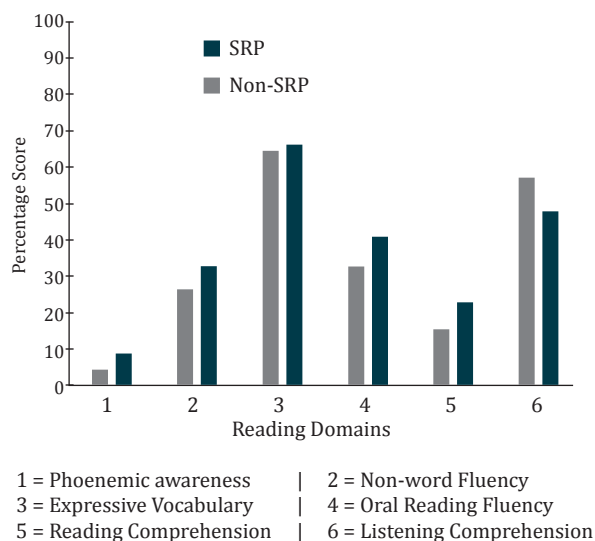
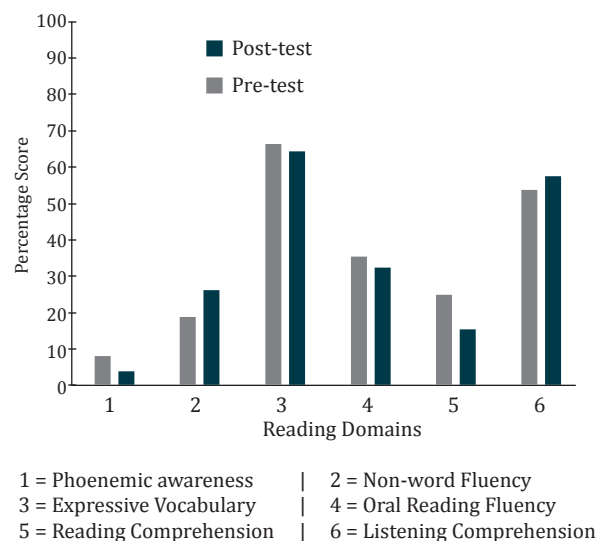


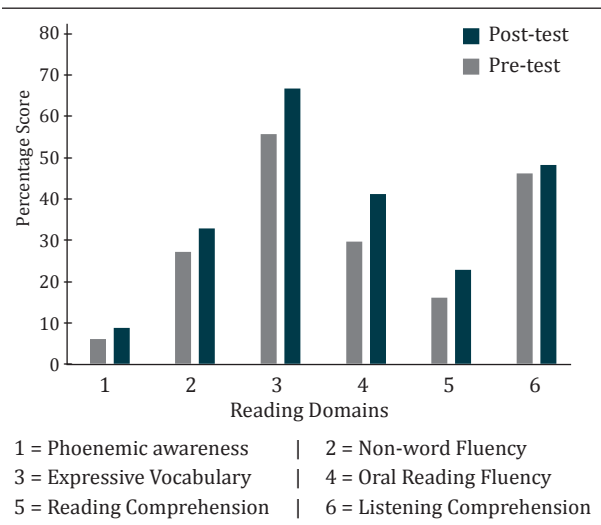
Figure 19: Non-SRP group: progression from pre- to post-test on EGRA-Urdu



on the rest of the four domains (i.e. *phonemic awareness, expressive vocabulary, oral reading fluency, reading comprehension*), the performance of non-SRP marginally declined over the period.

In general, children in the non-SRP group seem to have better skills in describing the given picture by identifying the subject, object and action (*expressive vocabulary*) and comprehending the given text (*listening comprehension*) as compared to the other reading domains. Interestingly, children show better performance in *oral reading fluency* than *reading comprehension* which indicates that children were able to read the given paragraph without necessarily comprehending the message given in the text. Results show a similar trend for *phonemic awareness* where children gained the lowest scores.

Figure 20: The SRP group: progression from pre- to post-test on EGRA-Urdu



Progression over time – the SRP group: From pre- to post-test, the SRP group shows improvements in all reading domains; however, the difference was found to be significant only in two domains (i.e. *expressive vocabulary, oral reading fluency*) as shown in Figure 20.

Over the period of time, students in the SRP group progressed in all reading domains; however, the difference was found to be significant only in the domain of *expressive vocabulary* ($p < 0.001$, $r = 0.24$) and *oral reading fluency* ($p < 0.05$, $r = 0.15$), which indicates that over time students show better performance in

describing the given picture by identifying action, subject and object and reading the text given to them; however, the magnitude of difference remained small for both domains. In terms of percentage difference, students show a gain of 11% in expressive vocabulary and 12% in oral reading fluency from pre- to post-test.

A scan of performance across domains over time revealed that children in the SRP group gained the highest score in *expressive vocabulary* followed by *listening comprehension*. Similar to Sindhi language, children show better skills in reading the given text (*oral reading fluency*) than comprehending the message in the text (*reading comprehension*). Results show a similar trend for *phonemic awareness* across languages.

EGRA Urdu – Zero vs Non-Zero Readers on Reading Domains

Zero readers – comparison across group

The data were reanalyzed to compare ratio of zero readers (those who cannot read a single word accurately) across group and over time. Table 14 presents a summary of zero readers’ comparison within and across groups.

Zero vs non-zero – comparison of pre-test across group:

Of the six comparisons made at pre-test, the two groups matched on four reading domains including – *phonemic awareness, non-word fluency, expressive vocabulary, and listening comprehension* ($p > 0.05$ for all). In *oral reading fluency*, the SRP group had 8% more zero readers as compared to their non-SRP counterparts. Similarly, in *reading comprehension*, SRP had 14% more zero readers as compared to the non-SRP group. However, the differences were not significant ($p > 0.05$ for both).

An overview of zero readers across domains revealed that at the outset, an overwhelming majority (more than 80%) of children fell into the category of zero readers in phonemic awareness, regardless of group. In non-word fluency, more than 40% of the children scored zero across two groups. In oral reading fluency, almost half of the children in non-SRP and a little more than half in SRP were unable to read a single word

Table 14: Zero readers – comparison across group and progression over time

Reading Domain	Non-SRP		SRP	
	Pre % (n)	Post % (n)	Pre % (n)	Post % (n)
Phonemic awareness	85% (39)	91% (53)	83% (92)	77% (95)
Non-Word Fluency	44% (20)	43% (25)	41% (45)	37% (46)
Expressive Vocabulary	4% (2)	7 % (4)	4% (4)	6% (7)
Oral Reading Fluency	48% (22)	45 % (26)	56% (62)	36% (44)
Reading Comprehension	57% (26)	69 % (40)	71% (79)	60% (74)
Listening Comprehension	15% (7)	12 % (7)	19% (21)	18% (22)

accurately in the given text. A little more than half (non-SRP) and nearly three-quarters (SRP) of the children across groups did not manage to comprehend the given text. Interestingly, both groups exhibited the lowest ratio of zero readers on expressive vocabulary (less than 10%) followed by listening comprehension (less than 20%).

Zero vs non-zero – comparison of post-test across group: Of the six comparisons made at post-test, SRP has the edge over non-SRP on five reading domains in terms of reducing zero readers; however, the difference was found to be significant only on phonemic awareness. The intervention seems to have influenced *phonemic awareness* significantly [$p < 0.05$, $\phi = 0.20$] with small magnitude. In terms of percentages, the SRP group managed to reduce zero readers by 14% as compared to their non-SRP counterparts. On the other hand, non-SRP surpassed SRP in reducing the number of zero readers in the domain of *listening comprehension* by 6% though, the difference was not significant.

An overview of zero readers across domains revealed that although the intervention helped to reduce the number of zero readers in phonemic awareness, the number of zero readers still remained highest in this domain. Moreover, almost 40% of the children scored zero across two groups in the domain of non-word

fluency. In oral reading fluency, almost half of the children in non-SRP and a little more than one-third in SRP were unable to read a single word accurately in the given text. Furthermore, almost two-thirds of the children across groups still did not manage to comprehend the given text. Interestingly, both groups exhibited the lowest ratio of zero readers on expressive vocabulary (less than 10%). Also, students show low zero readers in the domain of reading comprehension (less than 20%).

Zero readers – progression over time

This section presents progression over time in terms of reducing the frequency of zero readers in the SRP and non-SRP group separately.

Zero vs non-zero in the non-SRP group – progression over time: Of the six comparisons made to gauge progression in the non-SRP group, the numbers of zero readers reduced marginally in the three domains from pre- to post-test including *non-word fluency*, *oral reading fluency*, and *listening comprehension*; however, the difference was not significant on either domain. Surprisingly, in the other three domains (*phonemic awareness*, *expressive vocabulary*, *reading comprehension*) the numbers of zero readers increased over time; however, the difference was not found to be significant on either domain.

Zero vs non-zero in the SRP group – progression over time: Of the six comparisons made to gauge progression in the SRP group, the intervention seems to have managed to reduce the number of zero readers in five domains; however, only one reading domain exhibited a significant decline in zero readers over time. In the SRP group, the frequency of zero-readers reduced from pre- to post-test by 20% in the domain of *oral reading fluency* [$p < 0.01$, $\phi = 0.20$]; however, the magnitude of the difference was small. In terms of percentage difference, SRP intervention helped in reducing the number of zero readers by 20%.

CONCLUSION: COHORT-1

In order to examine EGRA scores across group (non-SRP and SRP) and progression over time (pre- to post-test), a total mean EGRA score was computed for each group separately. In addition, the ratio of zero and non-zero across group and progression over time was also computed. Results revealed that on average the two groups exhibited similar performance at the outset. After the intervention the two groups show significant difference in reading performance which was in favor of SRP. Evidently, both groups made progression over time in their reading skills. However, SRP shows statistically significant improvement with a small effect size. On the other hand, the progression made by the non-SRP is negligible. Results for Sindhi language followed the overall trend – SRP shows statistically better performance than non-SRP. However, for Urdu language, SRP shows some positive trends, but the differences were not found to be significant for most of the comparisons.

Both groups started with comparable frequencies of zero-readers except oral reading fluency where SRP started with a significantly lower frequency. Both made progress in terms of reducing zero-readers; however, SRP showed a significant drop in the frequency of zero-readers in most of the domains. On the other hand, non-SRP exhibited a nominal fall in the frequency of zero-readers over-time. The intervention seems to contribute significantly to reducing the frequency of zero-readers in all domains except expressive vocabulary and listening comprehension. It is important to note that the intervention contributed to

reducing zero-readers; however, only a fraction qualified to be fluent readers in most of the domains.

Exploring the factors which explains difference in students' performance across language was beyond the scope of this study. However, some speculations can be made to explain this difference. Perhaps, the sample drawn for Urdu EGRA was more heterogeneous in terms of languages. In other words, majority of children in Urdu-medium schools have mother-tongue which is different from Urdu. Therefore, they may only get an opportunity to practice Urdu in schools. On the other hand, Sample drawn for Sindhi EGRA comes from homogeneous group in terms of language. Majority of them speak the same language they were tested in.

A scan of the performance across domains in terms of mean scores revealed that listening comprehension skills have not been affected by the SRP intervention and neither has the expressive vocabulary. Interestingly, phonemic awareness stayed at the bottom end followed by non-word fluency, reading comprehension, oral reading, listening comprehension and expressive vocabulary.

Interestingly, children in both groups show better performance in oral reading than reading comprehension. In other words, children were able to read the given paragraph without necessarily comprehending the message given in the text. Despite a comparable performance on oral reading across group, SRP intervention contributed significantly to improving comprehension skills of students.

Performance on phonemic awareness stays at the bottom end indicating that students found it challenging to identify phonemes in the given words. Arguably, SRP intervention focuses heavily on phonemes, yet it remains one of the weakest areas on EGRA performance. Interestingly, students performed better in the rest of the five domains regardless of groups and the time test was administered. The questions raised here are: *what is the role of phonemic awareness in improving students reading skills? And can students read an unknown text and comprehend it without having adequate skills in decoding phonemes?*

PART-B: COHORT-2 (GT MODEL)

OVERALL EGRA RESULTS: NON-SRP VS SRP – COHORT-2

This section presents results of EGRA assessment for cohort-2. The section is organized into 5 main subsections including (i) overall comparison; (ii) comparing performance on reading domains across groups; and (iii) zero vs. non-zero readers on reading domains across groups.

EGRA Scores: Overall Comparisons

In order to examine EGRA scores across group (non-SRP and SRP), a total mean EGRA score was computed for each group separately. Figure 21 depicts an overall comparison between groups. Results reveal that on average the two groups exhibited significant difference in reading performance which was in favor of SRP ($p < 0.001$, $r = 0.12$).

While SRP intervention seems to contribute significantly in improving reading skills of participating students, the magnitude of the difference is rather small. On average, EGRA score of the SRP group was 5% higher than its non-SRP counterpart.

EGRA Scores: Comparing Performance on Reading Domains

Figure 22 presents comparative overview of students' performance on EGRA across groups. It is evident that on average SRP shows a consistent edge on non-SRP across all six domains; however, the differences were found to be significant in three domains (expressive vocabulary, reading comprehension and listening comprehension).

Students in the SRP group were significantly better in describing the given picture by narrating action along with object and subject ($p < 0.001$, $r = 0.16$). In terms of percentage difference, SRP scored 8% higher than their non-SRP counterparts in expressive vocabulary. Furthermore, EGRA score of the SRP group was 5% higher than the non-SRP group in comprehending the given text ($p < 0.05$, $r = 0.08$). The SRP students also demonstrated better skills in comprehending the text

Figure 21: Students performance on EGRA in cohort-2 across group

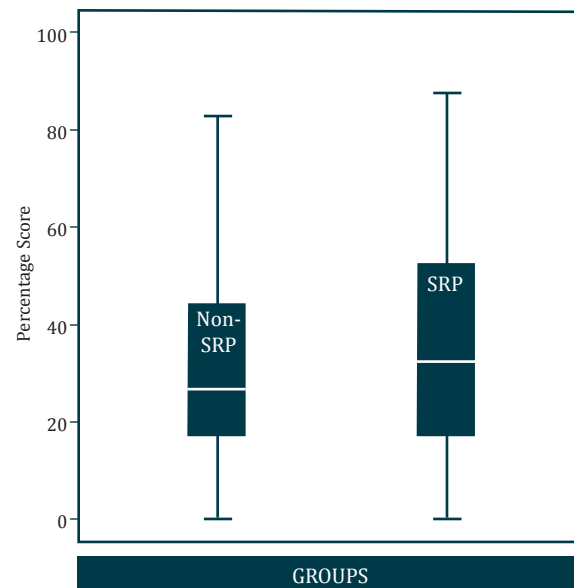
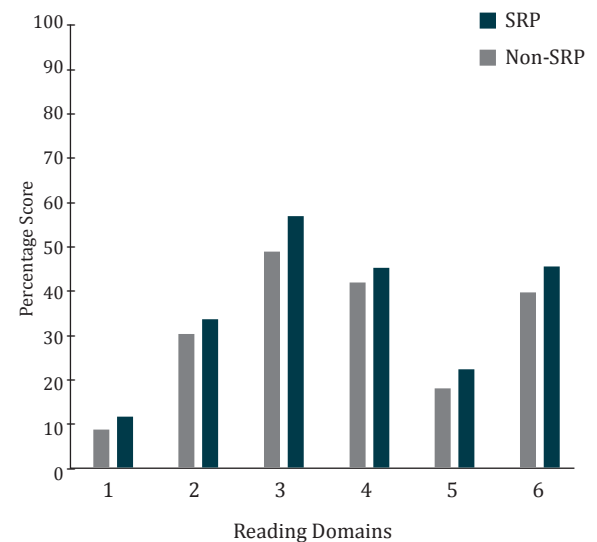


Figure 22: Students' performance on EGRA reading domains across group - Cohort-2



- 1 = Phonemic awareness | 2 = Non-word Fluency
3 = Expressive Vocabulary | 4 = Oral Reading Fluency
5 = Reading Comprehension | 6 = Listening Comprehension

read to them ($p < 0.05$, $r = 0.10$) with a 6% difference. Nevertheless, the magnitude of the difference was small for all three domains. No significant differences were observed in the two groups in reading, in identifying the last sound of the given word (phonemic

awareness) non-words accurately in the given time (non-word fluency) and in reading the given text (oral reading fluency). A scan of the performance across reading domains revealed that the pattern where phonemic awareness stayed at the bottom end followed by reading comprehension, non-word fluency, oral reading, listening comprehension and expressive vocabulary.

EGRA Scores: Zero vs Non-Zero Readers on Reading Domains

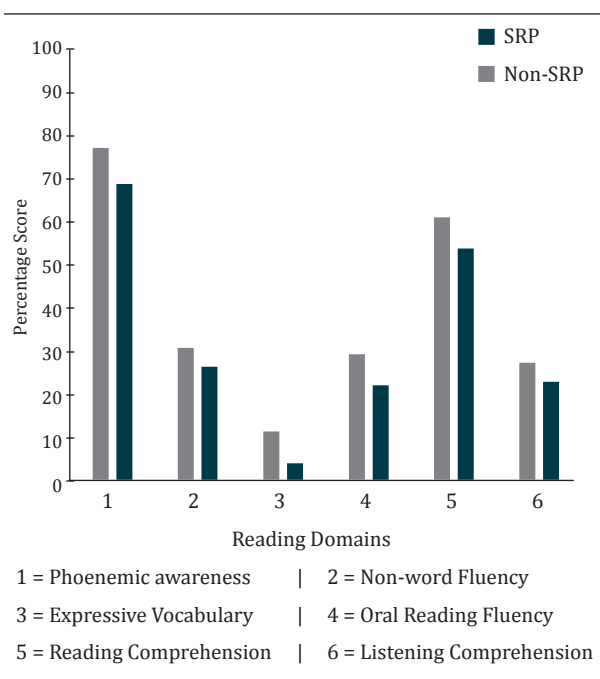
Zero-readers – comparison across group

The data were reanalyzed to compare the ratio of zero readers (those who cannot read a single word accurately) across groups. Figure 23 presents a summary of zero readers’ comparison within and across groups.

Of the six comparisons made on the EGRA scores, the differences were found to be significant (in favor of SRP) on three reading domains. The intervention seems to influence phonemic awareness [$p < 0.05$, $\phi = 0.10$] in reducing zero readers in SRP by 8% as compared to the non-SRP. Similarly, the difference was

found to be statistically significant in favor of SRP for expressive vocabulary [$p < 0.01$, $\phi = 0.12$] and for oral reading fluency [$p < 0.01$, $\phi = 0.08$] by 7% less zero readers as compared to the non-SRP. In other words, the frequency of zero readers is significantly lower in SRP as compared to the non-SRP for decoding the solicited phonemes in the given words (*phonemic awareness*) in describing the given picture by narrating action along with object and subject (*expressive vocabulary*) and in reading the given text accurately in the assigned time (*oral reading fluency*). The rest of the three comparisons followed the same trend; however, the ratio of zero readers was not significantly different in reading non-words fluently in the given time (*non-word fluency*); responding to the questions to comprehend the text after reading it (*reading comprehension*) and comprehending the text read to students (*listening comprehension*). A scan of the performance across reading domains revealed that regardless of the group there are overwhelmingly highest number of zero-readers in phonemic awareness and lowest in expressive vocabulary.

Figure 23: Zero readers on EGRA in cohort-2 across group



EGRA scores: comparing distribution of readers across groups

After analysis of mean scores and zero readers’ ratio, four performance categories were defined to demonstrate levels of ‘readers’: level 1 (1% to 25.9% score); level 2 (26% to 50.9% score); level 3 (51% to 75.9% score) and level 4 (76%-100% score). Performance of those children who ‘qualified’ to be readers (non-zero readers) was compared against these four levels for overall scores in six EGRA domains. A detailed analysis was undertaken to determine students’ performance in the six domains defined for EGRA. This section presents the results of readers’ scores against the four levels of performance across groups (see Table 15 for reference).

Phonemic awareness: It appears that SRP has an edge over the non-SRP although both groups followed the same trend. In other words, regardless of group, most of the ‘readers’ remained in the lower categories. Only

Table 15: Students' performance in the various domains of EGRA

Levels of Reading	Distribution of readers across group				
	Non-READERS	READERS			
	Zero Reader	Level 1 (1-25%)	Level 2 (26-50%)	Level 3 (51-75%)	Level 4 (76-100%)
Phonemic awareness					
SRP Children (%)	69	15	7	6	3
Non-SRP Children (%)	77	11	6	4	2
Non-word fluency					
SRP Children (%)	26	25	21	12	16
Non-SRP Children (%)	31	19	25	14	11
Expressive vocabulary					
SRP Children (%)	4	5	36	32	23
Non-SRP Children (%)	10	10	35	30	15
Oral reading fluency					
SRP Children (%)	22	18	18	12	30
Non-SRP Children (%)	29	17	13	11	30
Reading comprehension					
SRP Children (%)	54	15	11	11	9
Non-SRP Children (%)	61	13	13	7	6
Listening comprehension					
SRP Children (%)	23	0	13	35	12
Non-SRP Children (%)	27	0	36	30	7

a fraction of sample children managed to decode the solicited phonemes in most of the given words (i.e. eight to ten words) while a majority of readers remained at the bottom end who managed to decode phonemes only in one to two words.

Non-word fluency: Interestingly the percentage of the SRP readers is higher at two extremes (i.e. Level 1 and Level 4) than the non-SRP. It indicates that quite a few of the SRP readers only managed to read less than ten words with accuracy in the given time while less than one-fifth of the sample children read thirty-one or more words accurately. On the other hand, distribution of the readers in non-SRP exhibits a higher ratio of children in level 1 and level 2.

Expressive vocabulary: As described earlier, expressive vocabulary is one of the highest scoring domains on

EGRA for the current study. Distribution of the readers across four categories in both groups explains the trend of scoring on expressive vocabulary. The ratio of level 1 readers is higher for non-SRP while the trend is reversed for level 4. For the two middle categories, the trend is comparable across groups. Evidently almost a quarter of the SRP children managed to describe seven to ten pictures accurately. On the other hand, less than one-fifth of the non-SRP sample managed to qualify for the highest level.

Oral reading fluency: Performance of children was found to be comparable across groups and the distribution of readers across four categories explains the trend. Comparatively a higher number of students fall in the highest category for both groups. In other words, almost one-third of the sample children across

groups managed to read forty-two or more words accurately in one minute.

Reading comprehension: Children found comprehension more difficult than reading the given text; however, SRP seems to show comparatively better comprehension skills than their non-SRP counterparts. A higher ratio of SRP children in level 3 and 4 indicates that relatively more children managed to respond to three to five questions, hence better comprehension skills.

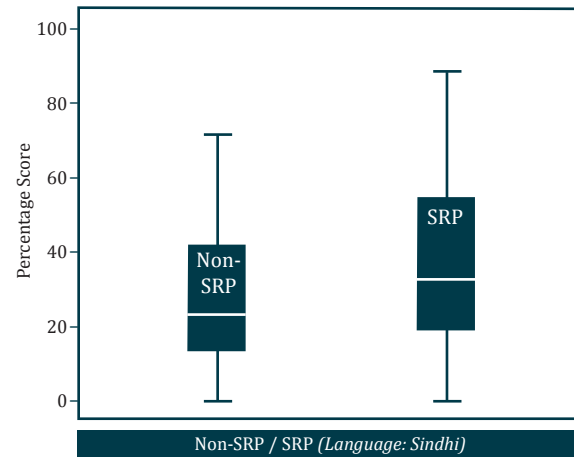
Listening comprehension: It is important to note that listening comprehension domains is based on three questions and the results can only be classified in three categories. However, 4 levels of performance have been defined for all six EGRA domains for consistency. As discussed earlier, the SRP children demonstrated better skills in comprehending the text and the trend is manifested in readers' distribution for this domain too – higher ratio of the SRP children in level 3 and 4. Readers' distribution indicates that relatively more children managed to respond to 2 (level 3) or 3 (level 4) questions to comprehend the text read to them.

In summary, the distribution of readers on six domains of EGRA highlights that SRP intervention has helped to decrease ratio of 'zero readers' in the SRP group; however, only a fraction qualifies to be fluent readers (level 4) in most of the domains. While it is important to 'monitor' drop in zero-readers, it is imperative to scrutinize the distribution of readers across all four levels. The latter may help to make targeted amendments in various components of the intervention (i.e. teachers' training, follow-up, adequacy and quality of material, validity of assessments).

COMPARISON BY LANGUAGE: NON-SRP VS SRP – COHORT-2

This section presents results of comparison of EGRA assessment for cohort-2 by language i.e., EGRA-Sindhi and EGRA-Urdu. These sections are further divided into 3 main subsections including; (i) overall comparison; (ii) comparing performance on reading

Figure 24: Overall comparisons on EGRA-Sindhi on cohort-2



domains; (iii) zero vs non-zero readers on reading domains.

EGRA-Sindhi – Overall Comparisons

In order to examine EGRA scores by Sindhi language across groups (Non-SRP and SRP), a total mean EGRA-Sindhi score was computed for each group separately. Figure 24 depicts an overall comparison between groups by Sindhi language. Results indicates that on average the two groups exhibited significant difference in reading performance which was in favor of SRP ($p < 0.001$, $r = 0.23$).

In terms of percentage difference, SRP has scored 9% higher than their non-SRP counterparts. However, the magnitude of the difference is rather small.

EGRA-Sindhi: Comparing Performance on Reading Domains

Figure 25 presents the comparative overview of students' performance on EGRA-Sindhi across reading domains. It is evident that on average SRP shows a consistent edge over non-SRP across all six domains with significant differences on all except phonemic awareness.

Significant differences were also observed in reading non-words accurately in the given time (*non-word fluency*) favoring the SRP group by approximately 9% ($p < 0.01$, $r = 0.16$). For *expressive vocabulary*, students in the SRP group has performed better by 11% in

Figure 25: Students' performance on EGRA-Sindhi in cohort 2 across reading domains

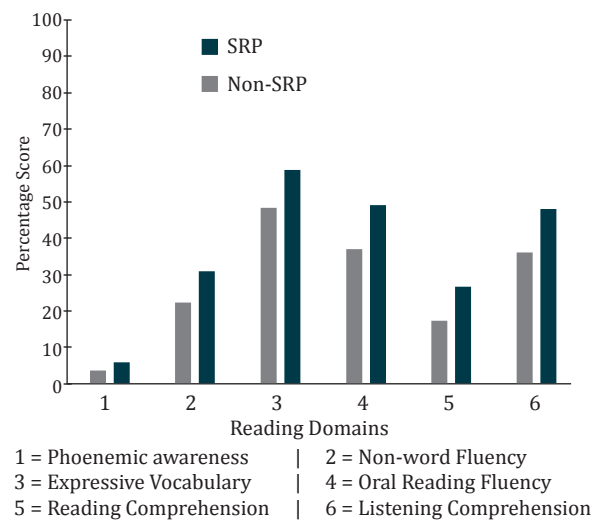
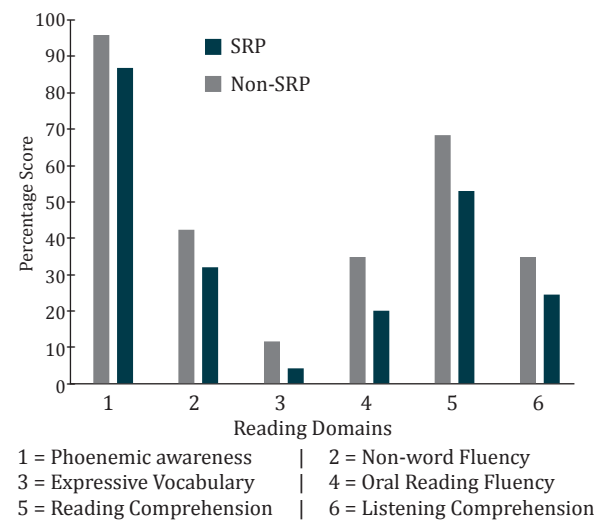


Figure 26: Zero readers on EGRA-Sindhi in cohort 2 across group



describing the given picture by narrating action along with object and subject ($p < 0.001$, $r = 0.22$). In *oral reading fluency*, the SRP children performed better than their counterparts by 12% ($p < 0.001$, $r = 0.16$). Students in the SRP group scored 9% higher than the non-SRP group in comprehending the given text ($p < 0.01$, $r = 0.16$). Similarly, students in the SRP group also scored 12% higher than the non-SRP group in comprehending the text read to them ($p < 0.001$, $r = 0.19$). Furthermore, a scan of the performance across reading domains revealed a pattern where *phonemic awareness* stayed at the bottom end followed by *reading comprehension*, *non-word fluency*, *listening comprehension*, *oral reading fluency*, and *expressive vocabulary*. Having said that, SRP had an edge in all reading domains with a significant difference; however, the magnitude of difference was found to be small in all cases.

EGRA-Sindhi: Zero vs Non-Zero Readers on Reading Domains

The data were re-analysed to compare the ratio of zero readers (those who cannot read a single word accurately) across groups in EGRA-Sindhi. Figure 26 presents a summary of zero readers' comparison within and across groups.

Of the six comparisons made on the EGRA-Sindhi scores, the differences were found to be significant on five reading domains; whereas the sixth domain had a close tendency towards significance. The intervention seems to influence in reducing zero scorers in *phonemic awareness* [$p < 0.05$, $\phi = 0.10$] by 8% among the SRP group as compared to the non-SRP. Similarly, the difference was found to be significant and in favor of SRP for *expressive vocabulary* [$p < 0.01$, $\phi = 0.12$]. In terms of percentages, SRP intervention seems to have contributed in reducing the ratio of zero readers by 6% as compared to non-SRP. In other words, the frequency of zero readers is significantly lower in SRP as compared to non-SRP for decoding the solicited phonemes in given words (*phonemic awareness*) and in describing the given picture by narrating action along with object and subject (*expressive vocabulary*). Furthermore, there was a significant difference in favor of SRP for *oral reading fluency* [$p < 0.01$, $\phi = 0.14$], *reading comprehension* [$p < 0.01$, $\phi = 0.13$] and *listening comprehension* [$p < 0.05$, $\phi = 0.10$] with a percentage difference of 13% (*Oral reading and reading comprehension*) and 9% (*listening comprehension*). In other words, the frequency of zero readers is significantly lower in SRP as compared to non-SRP for their ability to read the given text fluently in given time

(oral reading comprehension); responding to the questions to comprehend the text after reading it (reading comprehension) and comprehending the text read to students (listening comprehension). A marginally significant difference was found in the domain of non-word fluency which favored SRP [$p=0.051$, $\phi= 0.09$] with 9%. A close look of the performance across reading domains revealed that, regardless of the group, there are overwhelmingly highest number of zero-readers in phonemic awareness and lowest in expressive vocabulary.

EGRA-Urdu – Overall Comparisons

Figure 27 presents an overall comparison between groups by Urdu language. Results showed that in Urdu intervention, Non-SRP had an edge over SRP; however, the difference was not significant.

EGRA-Urdu: Comparing Performance on Reading Domains

Figure 28 depicts a comparative overview of students' performance on EGRA-Urdu across the groups. The graph shows that in the four domains (non-word fluency, oral reading fluency, reading comprehension, and listening comprehension), non-SRP has an edge over SRP; whereas, in the rest of the two domains (phonemic awareness and expressive vocabulary), SRP shows better performance than their counterparts; however, the difference was found to be significant only in the domain of oral reading fluency.

Among all domains, students achieved lower scores in the domain of reading comprehension and higher in expressive vocabulary and reading comprehension. Interestingly, non-SRP had a significant edge over SRP in reading the assigned text accurately in the given time (oral reading fluency) by 15% ($p<0.01$, $r=0.19$). An overview of the performance across reading domains revealed a mixed pattern where non-SRP is leading in four reading domains, whereas SRP showed better performances in two domains; however, the difference was found to be significant in only one domain (i.e., oral reading fluency). Moreover, comparison across domains revealed that reading comprehension stayed at the bottom followed by

Figure 27: Overall comparisons on EGRA-Urdu in cohort-2

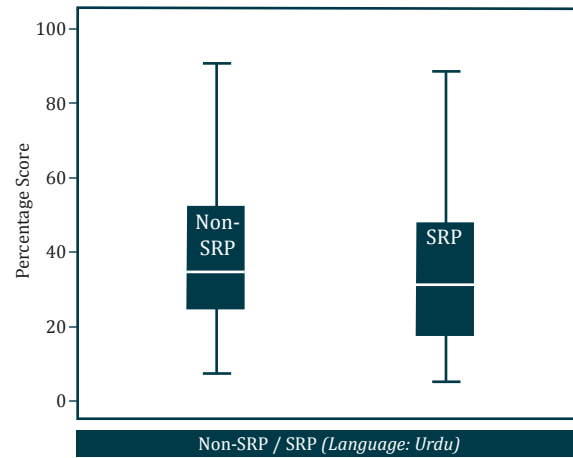
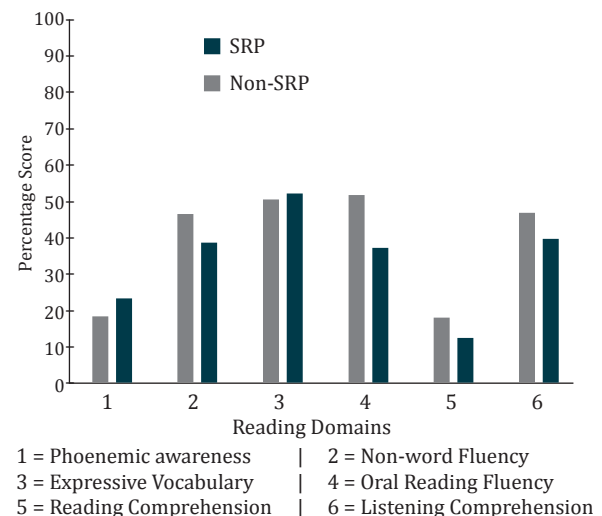


Figure 28: performance on EGRA-Urdu in cohort-2 across reading domains



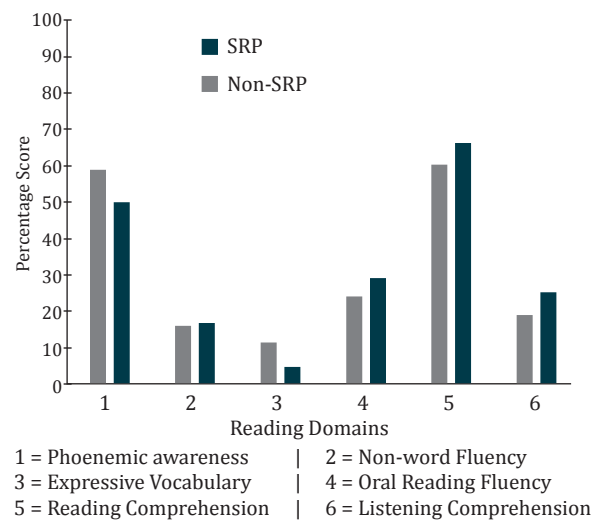
phonemic awareness, non-word fluency, listening comprehension, oral reading fluency, and expressive vocabulary.

EGRA-Urdu: Zero vs Non-Zero Readers on Reading Domains

The data were reanalyzed to compare the ratio of zero readers (those who cannot read a single word accurately) across groups. Figure 29 presents a summary of zero readers' comparison within and across groups.

Of the six comparisons made on the EGRA-Urdu scores, the frequency of zero readers is found to be lower in

Figure 29: Zero readers on EGRA-Urdu in cohort-2 across group



non-SRP as compared to SRP in four domains (i.e., *non-word fluency, oral reading fluency, reading comprehension, and listening comprehension*); however, the difference was not found to be significant. For the rest of the two domains, the frequency of zero readers is lower in SRP as compared to their non-SRP counterparts (i.e., *phonemic awareness and expressive vocabulary*); however, the difference is not significant. In EGRA-Urdu, there is an overwhelming number of zero readers in the domains of *phonemic awareness and reading comprehension* regardless of the group. On the other hand, there were lowest numbers of zero readers in the *expressive vocabulary* regardless of the group.

CLASSROOM PRACTICE: NON-SRP VS SRP – COHORT-2

Overall Quality of Classroom Practice

In total, 104 lessons of the early grade teachers were observed who were recruited from the target districts. An almost equal number of the lessons was observed for non-SRP (n=50; 48%) and SRP (n=54; 52%) teachers. This section presents a summary of the quality of classroom practice across groups using the overall ERCOP ratings as well as ranking on generic pedagogical strategies and specific reading strategies.

On average, the SRP (M=1.81; SE=0.02) classrooms exhibited relatively better practices as compared to the

non-SRP (M= 1.54; SE=0.01) counterparts and the difference was found to be significant (p<0.001; r=0.45). Results indicate that the SRP classrooms have shown a difference of 0.27 points on a 3-point scale. Results followed the same pattern for the two distinct categories defined in ERCOP to represent generic pedagogy (ERCOP-pedagogy) and reading specific pedagogy (ERCOP-reading). On average, the SRP (M=1.91; SE=0.02) classrooms have exhibited relatively more participatory practices as compared to the non-SRP (M= 1.62; SE=0.19) classes and the difference was found to be significant (p<0.001; r=0.43). Similarly, SRP (M= 1.73; SE=0.02) has a significant edge in the use of reading specific strategies as compared to the non-SRP (M= 1.48; SE=0.01) classrooms (p<0.001; r=0.42). The SRP intervention showed marginally higher contribution in teachers' classroom practice for the use of generic pedagogy (difference = 0.29 on a 3-pointscale) than reading specific strategies (0.25 on a 3-point scale).

In order to examine the characteristic of the classroom practices, quality was defined as three broader categories. Any score less than or equal to 1.5 was defined as weak practice, a score between 1.51 and 2.5 was defined as mediocre while a score of more than 2.5 was categorized as good practice. Table 16 presents the percentage of lessons by ranges of ERCOP-overall, ERCOP-pedagogy and ERCOP-reading scores across groups.

In placing these results within this framework, it is important to note that a majority of lessons for the SRP group qualified for the middle band of quality for all three comparisons. Furthermore, an ERCOP score of

Table 16: Percentage of lessons by ranges of ERCOP score across group

	Level/ Group	Weak (≤1.50)	Mediocre (1.51-2.5)	Good (>2.5)
ERCOP Overall	Non-SRP	29 (58%)	21 (42%)	0 (0%)
	SRP	07 (13%)	47 (87%)	0 (0%)
ERCOP Pedagogy	Non-SRP	16 (33%)	32 (63%)	02 (04%)
	SRP	10 (18%)	42 (77%)	02 (05%)
ERCOP Reading	Non-SRP	33 (65%)	17 (35%)	0 (0%)
	SRP	16 (30%)	38 (70%)	0 (0%)

less than 1.51 for quite a few of the observed lessons in non-SRP and small number of lessons in SRP pointed to weak classroom practices. At the highest end of quality, both non-SRP and SRP received comparable ratings for a fraction of lessons qualifying for good practice only in the use of generic pedagogical strategies (e.g. questioning, interactions, active involvement of children). Conversely, none of the classrooms across groups qualified for the highest band of rating for reading specific pedagogical strategies (e.g. phonemic awareness, teaching vocabulary, reading to students, and reading in groups).

Quality of Classroom Practice by ERCOP Items

ERCOP pedagogy

Table 17 presents item-wise comparison of classroom quality across groups for generic pedagogical skills. The scoring across items followed a similar pattern which was observed for overall comparison of generic pedagogical skills across both SRP and non-SRP groups. Evidently, the SRP group has exhibited mediocre practice on all aspects except *child-child interaction*. Similarly, the practices in the non-SRP group also fell in the range of mediocre practice on all aspects except two (*child-child interaction* and *material use*). In other words, interactions between children

were almost non-existent and the pattern appears to be similar across groups. Similarly, *supervision* remains at the lower end across groups after child-child interaction. Supervision does not mean ‘patrolling’ the class and telling students what to do. On the contrary, it is to draw on students’ skills and expertise by giving constructive feedback on their work, by acknowledging their effort, by clarifying tasks, and ensuring correct content. Glimpses of this type of supervision were rarely observed across groups.

All eight comparisons made, a majority (n=6; 75%) favored the SRP group indicating a significantly better quality of general teaching learning practices in the SRP classes. Teachers in the SRP classes did not only manage to use basic material (e.g. chalk, board, textbook), but in quite a few observed lessons, teachers also used some subject related supplementary material (e.g. flash cards, story books) to enrich children’s learning (*use of material*). Furthermore, teachers in the SRP classes displayed a more supportive gesture by exhibiting a calmer tone, listening to children and treating them fairly (*teacher-child interaction*). A majority of the teachers in SRP group was observed to maintain control in their classes without using discipline techniques, such as corporal punishment or excluding children from activities for a long period of time (*discipline*). In most of the cases, children were seen actively engaged in classroom activities in the SRP group. In other words, they were not observed to be day dreaming, socializing with each other or reading totally irrelevant material (*on-task behavior*). *Active involvement of children* was another area where SRP has shown better practices as compared to the non-SRP. This aspect requires students’ engagement in activities through asking and answering questions, problem solving and by demonstrating a certain level of creativity. Some of the requirements were met in some of the classes in the SRP group for this particular aspect. In some of the instances the SRP teachers were observed asking a series of questions where he/she responds to the students with further questions. However, such examples were hardly observed in the non-SRP group.

Table 17: Quality of classroom practice – mean score on ERCOP pedagogy items

Constructs/ Groups	Non-SRP	SRP	Differences <i>Non-SRP-SRP</i>
Material Use	1.41	1.93	-0.52**
Teacher-Child Interaction	1.84	2.19	-0.35**
Child-Child Interaction	1.12	1.13	ns
Discipline	1.71	2.06	-0.35**
On task Behavior	1.93	2.19	-0.26**
Supervision	1.63	1.75	ns
Active Involvement	1.67	2.00	-0.33**
Questioning	1.62	2.01	-0.39**

*p<0.05 | **p<0.01 ns = difference is not statistically significant

Minus sign indicates that SRP demonstrated better practice as compared to non-SRP.

Table 18: Quality of classroom practice – mean score on ERCOP reading items

Constructs/ Groups	Non-SRP	SRP	Differences Non-SRP-SRP
Reading aloud	2.58	2.60	ns
Phonological awareness activities	1.11	1.41	-0.30**
Use of common text (storybooks, poetry)	1.19	1.79	-0.60**
Teaching new vocabulary	1.31	1.55	-0.24**
Vocabulary in the text	1.33	1.50	-0.17**
Students read to the teacher (Flash cards, text)	1.65	1.96	-0.31**
Students reading to the whole class	1.43	1.90	-0.25**
Students reading in small groups	1.05	1.06	ns
Students reading to self (silently)	1.05	1.11	ns
Students listening to teacher	2.14	2.41	-0.27**

* $p < 0.05$ | ** $p < 0.01$ ns = difference is not statistically significant

Minus sign indicates that SRP demonstrated better practice as compared to non-SRP.

That said, examples of students raising questions were almost non-existent in both groups (*questioning*).

ERCOP reading

Of the ten ERCOP-reading items, the SRP classes scored within the range of mediocre or acceptable practices (1.51-2.50) on five items as shown in Table 18. On the other hand, score of the non-SRP classes fell in the range of weak practices for seven items. Interestingly, *reading aloud* exhibited good practice (≥ 2.5) across groups. Reading aloud refers to typically a whole class activity in which the teacher read aloud to students using age appropriate text in the form of stories, words and sentences. This was one of the most commonly used techniques recorded during classroom observations. On the other hand, *students reading in small groups* and *students self-reading to* were identified as two common lowest scoring items. On these two items, lessons scored within the range of weak quality of reading instructions across groups.

Reading in small group refers to an activity where students read a text to other members of the group or with the group. This finding coincides with the scoring pattern on *child-child interaction* in ERCOP-pedagogy. There were not many examples of planned opportunities for children to interact with each other. Even in terms of academic interactions, children were hardly given a chance to read to each other from their books. Moreover, children were hardly given a chance of independent reading where they can vocalize text to themselves or walk around the class reading captions and labels on the walls.

All ten comparisons made, a majority ($n=7$; 70%) favored the SRP group indicating a significantly better quality of reading instructions in the SRP classes. Teachers in the SRP classes exhibited an edge over their non-SRP counterparts in using *phonological awareness activities*. That is to say, teachers in some of the SRP classes used activities to enhance awareness of sounds within spoken words through teaching of letter sounds, word sounds and rhymes. Furthermore, teachers in the SRP classes were more inclined to use common text (*use of common text*) such as story books, poetry and flash cards to draw students' attention on comprehension and on specific features such as word building, punctuations and layout. Teaching of vocabulary in the target language was observed from two interlinked perspectives including *teaching new vocabulary* and understanding new *vocabulary in the text*. Observation in the SRP classes provided some examples of teaching word meaning through written expressions or non-linguistic way such as pictures (*teaching new vocabulary*). Similarly, in few SRP classes, teachers were also observed helping students to identify new words in the text and practise it (*vocabulary in the text*). That said, teaching of vocabulary barely met the standard of mediocre practice for introducing new vocabulary or remained in the range of weak practice for helping students understand new vocabulary in the text that they are reading. On average, more examples were observed in the SRP classes where students were given opportunities to read to the teacher (*students reading*

to the teacher) or read to the whole class (*students reading to the whole class*). This aspect of practice gives an edge to the SRP group in getting more avenues to practice and improve reading skills. However, in most of the observed examples, textbooks were used as a reading text. None of the observations recorded children reading their 'own writing' to the teacher or to the whole class. Group reading (*students reading in small groups*) and silent reading (*students reading to self*) fell in the range of weak practice across groups. Scoring pattern on *reading in small group* is in line with scoring on one of the relevant aspects of generic pedagogy (i.e. child-child interaction). Regardless of the group, the scoring on the latter fell in the category of weak practice. Evidently, peer interaction is one of the pre-requisite for encouraging children to read to peers. Teacher-child interaction remained a prevalent observed practice therefore, limited opportunities for peer reading was not surprising. Arguably, a collaborative effort would help children to improve their reading skill. *Reading to self* is one of the critical aspects of reading instructions; however, an overwhelming majority of the classes did not demonstrate any example of silent reading. It is argued that independent reading is the foundation for creating independent readers. As part of the silent reading, children may read books according to their own level and at their own pace. They can visualize and interpret the words in their own way followed by sharing their thoughts on the books that they read. This type of reading instructions would help in enhancing comprehension skills – one of the most important outcomes of reading instructions. Lack of this practice in the SRP classes will have implications on students' reading skills. *Student listening to the teachers* was one of the highest scoring aspects of reading instructions across groups with SRP having a significant edge over its non-SRP counterparts. In quite a few of the SRP classes students were observed to be listening to the teacher while he or she reads from textbooks or supplementary text.

ERCOP items 7 and 8 – variety of active methods and classroom management strategies: A detailed analysis

Table 19: Variety of teaching strategies

Teaching strategies	Non-SRP	SRP
Questioning	79	96
Stories	28	68
Picture/Drawing	8	57
Drama/Role Play	10	5
Songs/Poems	4	0
Games	6	18
Others	17	8

Table 20: Classroom management strategies

Management strategies	Non-SRP	SRP
Whole Class Strategies	100	94
Group / Pair work	0	0
Combination	0	6

was undertaken for items 7 and 8 of ERCOP to assess the range of teaching strategies and variety of classroom management techniques used during the observations. The results are presented in Table 19 and Table 20.

It is important to note that items 7 and 8 were used only to record the frequency of teaching and management strategies used in the classroom. The intention was not to gauge the quality of these methods. The other items in ERCOP-Pedagogy (e.g. *questioning, active involvement, and child-child interaction*) were defined to measure quality in terms of children's involvement in the teaching and learning processes. Of the 104 lessons observed, *questioning* strategies were used in a majority of classes across groups. However, a variety of teaching strategies (e.g. *stories, pictures/drawing, and games*) were observed in most of the SRP classrooms as compared to non-SRP. Interestingly, non-SRP shows a marginal edge on SRP in using *role play and songs* as teaching strategies.

Item 8 encompasses three modes of classroom management strategies including *whole-class teaching, group/pair work or combination of both*. The results are presented in Table 20. Evidently, *whole-class teaching* was used as the only classroom management strategy in the non-SRP classes. Similarly, an

Table 21: Home literacy environment – comparison across group

Statements	Non-SRP	SRP	Differences
Literacy Involvement overall	2.25	2.54	-0.29
Listen to my child read aloud	2.52	3.12	0.60
Talk with my child about things we have done	2.65	2.64	0.01
Talk with my child about what he/she is reading on his/her own	2.87	3.36	-0.49
Discuss my child's classroom reading work with him/her	2.70	3.03	-0.33
Go to the library or book store with my child	1.43	1.47	-0.04
Help my child with reading for school	2.39	2.91	-0.52
Read books to my child	1.91	2.19	-0.28
Tell stories to my child	2.30	1.94	0.36
Sing poems to my child	1.91	2.39	-0.48
Help my child write letters or words or sentences	1.78	2.16	-0.38

* $p < 0.05$ | ** $p < 0.01$ ns = difference is not statistically significant

Minus sign indicates that SRP demonstrated better practice as compared to non-SRP.

overwhelming majority of the SRP classes used *whole-class teaching* while a *combination of both strategies* was also used in some classrooms.

HOME LITERACY ENVIRONMENT: NON-SRP VS SRP – COHORT-2

A questionnaire comprising 10 statements about home literacy practices was administered with parents of students from SRP and non-SRP groups through a face-to-face structured interview. Table 21 presents a comparative overview of home literacy environment as reported by the parents. Evidently, in overall home literacy environment, SRP has an edge on non-SRP which means that the parents of SRP students reported that they provide more opportunities to their children's in literacy enhancement activities. However, the difference was not statistically significant ($p > 0.05$). A detailed analysis was carried out to explore differences for ten components of home literacy environment. SRP has shown an edge over their non-SRP counterparts in eight components; however, the difference was significant only for one aspect of home literacy environment. Parents in the SRP group claimed to have more frequent engagement (difference of 0.60 points on a 4-point scale) in *listening to their child when he/ she reads* than the non-SRP group; however, the difference was not significant.

Interestingly, it was found that parents in the non-SRP group *tell stories to their children* more often than their non-SRP counterparts; however, the difference was not significant. The results also revealed that frequency of *parents' talk with their child about things that they have done in life* was almost similar across groups. Within group scanning of patterns exhibited that the SRP parents engage their children more frequently by *listening to them when they read aloud, talking to them about their own reading and discussing with them about classroom reading*. Like the SRP group *talking with their child about his/her own reading, discussing classroom reading* remain relatively most frequently reported practice in the non-SRP group as well. It was not surprising to observe that visit to a library or a book store was reported to be a rare practice. Arguably, there was hardly any provision of library in the SRP target areas which limits this important 'out of school' opportunity to augment children's reading skills.

CONCLUSION: COHORT-2

In order to examine EGRA scores across groups (non-SRP and SRP) in Cohort-2 a total mean EGRA score was computed for each group separately. In addition, the ratio of zero and non-zero and the distribution of readers across groups was also computed. The results indicated that the two groups have shown significant

difference in reading performance in favor of SRP. The results for Sindhi language followed the overall trend – SRP has shown statistically better performance than non-SRP. However, for Urdu language, SRP has shown some positive trends, but the differences were not significant for most of the comparisons. Exploring the reasons for non-difference in performance in Urdu language was beyond the scope of the study. However, some speculations were made and have been discussed earlier (Conclusion Cohort-1, p-40).

The intervention seems to have contributed significantly in reducing the frequency of zero-readers in phonemic awareness and expressive vocabulary. The results followed a similar trend in all other domains; however, the difference was not significant. Since the intervention contributed in reducing zero-readers, only a fraction has qualified to be fluent readers in most of the domains.

A scan of the performance across domains revealed that there was no significant difference between the performance of children in two groups on non-word fluency and oral reading fluency. It means that the intervention does not seem to have contributed significantly in enhancing children's skills in reading non-word (non-word fluency) as well as reading the given text (oral reading fluency) accurately in the given time. Interestingly, phonemic awareness stayed at the

bottom end followed by reading comprehension, non-word fluency, oral reading fluency, listening comprehension and expressive vocabulary. A similar pattern was followed in both groups.

Interestingly, students in both groups showed relatively better performance in oral reading than reading comprehension. In other words, students were able to read the given paragraph without necessarily comprehending the message given in the text. Despite a comparable performance on oral reading across group, SRP intervention contributed significantly in improving comprehension skills among students.

The performance on phonemic awareness stays at the bottom end indicating that students struggled to identify phonemes in the given words. Arguably, SRP intervention focuses heavily on phonemes yet it remains one of the weakest areas on EGRA performance. Interestingly, students performed better in other five domains regardless of groups and the time of test administration. This situation raises important questions such as; *what is the role of phonemic awareness in improving students reading skills? Can students read an unknown text and comprehend it without having adequate skills in decoding phonemes?*

5

Findings: Qualitative Analysis

EGRA SCORES AND FACTORS EXPLAINING STUDENTS' PERFORMANCE

Within-Case Analysis

Research questions one and two focus on understanding the factors that may have contributed to students learning. More importantly, the questions are concerned with finding out how intervention affected those factors. Thus we collected quantitative and qualitative data from various stakeholders, looking at both processes and outcomes.

The findings from interviews with heads and teachers revealed that they found SRP intervention helpful in terms of providing material and pedagogical techniques. The training particularly focused on sounds and phonemic awareness. The pedagogical techniques that were highlighted were use of flash cards, games, stories, pictures and role play. Teachers and head teachers acknowledged the usefulness of SRP material i.e. lesson plans and stories.

One of the teachers shared, “My school has an SRP project running. I noticed that when an interesting thing (learning material) is placed in front of the child s/he learns quickly. They also start reading it”. They also considered training beneficial. They think that SRP emphasizing on phonemic method does improve students reading skills, and that pictures and stories help to develop their interest. The majority of the teachers were of the view that they used the phonemic method and they observed positive results in terms of children’s reading, interest and their participation in class. However, teachers found continuing this practice (teaching through phonemes) cumbersome as it took more time and there was limited support in terms of implementing the strategies teachers had

learnt during SRP training. Teachers also struggled with the phonemic method at a personal level, because they viewed that reading through the hijjay method (zair, zabr and paish) was more effective than the phonemic method as the former was their modus operandi and students were more acquainted with it. One of the teachers explained, “We shared with the SRP team that this [the phonemic method] is creating problems At the time when we studied, in that environment, alif, bay, pay was ‘the method’ to learn alphabets and not aa, baa, taa”. There were further issues which hindered the proper implementation of SRP intervention, such as lack of/limited supply of reading material, and there was little additional support in terms of teaching aids. In addition, some of the heads and teachers who went through training under cohort-1 reported that they had discontinued the use of SRP material. When the researchers inquired, they could not elaborate on various aspects of their training which happened in 2015.

The guide teachers echoed teachers’ views in terms of usefulness of SRP intervention for enhancing teachers’ practices. They emphasized that SRP material was complimentary to the Sindh Textbook Board (STBB), though it was difficult for teachers to understand in reality as they had a ‘different’ textbook with them and they struggled to find linkages. A review of SRP lesson plans and reading material revealed that it would be an overestimation of teachers’ skills to expect them to integrate SRP material in their routine teaching after participating in a three to six-day training session. Furthermore, guide teachers felt that “teachers and students both were accustomed to old methods and, therefore, it was hard for them to change”. Although the guide teachers visited schools to provide additional

classroom support, the visits were still considered insufficient by the teachers to reconceptualize and integrate new ways of teaching reading into their existing scheme of thinking and work routines.

The SRP field officers reported that they had established reading corners and also shared materials which included read aloud, level readers and lesson planners. This provision of material in classes positively affected pedagogy and students' learning. In particular, the use of pictures helped in recognition of words by students. They further explained that teachers were supported to practise learnt skills during training through model lessons and micro teaching. They agreed to some of the concerns shared by the other stakeholders, which may have hindered the implementation of SRP project, such as: lack of teaching aids, limited understanding of teachers to link teaching with students' level and short duration of training.

DEOs and TEOs appreciated the project and acknowledged the usefulness of phonemic method to enhance reading skills. They felt that teachers' preference for the traditional method of teaching reading (the hijjay method) was one of the main hindering factors in not only implementing but also sustaining the project's intervention. One of them explained it as "[Teachers think that] the old method [*hijjay*] gives results; [therefore], old method is better". They argued that Grade 1 students were still able to read despite the use of traditional methods, hence this raises question about the efficacy of the phonemic method as the core intervention of the project.

The mismatch between STBB and SRP intervention was also highlighted by the provincial officials from the Bureau of Curriculum (BoC), the Sindh Text Book Board (STBB) and the Directorate of Education. One of them explained the mismatch in terms of the expectations from STBB and SRP material as, "The SRP material only focused on reading, whereas STBB emphasizes both reading and writing". However, they acknowledged the effort of SRP in contextualizing the language material in collaboration with the education department.

According to SRP Central officials, the material was developed in close consultation with the relevant departments of the Government of Sindh. They claim that the value addition by their project was the reading performance standards, which had never been a part of the national curriculum; thus it was a major policy shift. Explaining the reason behind the limited readers provided to the classes, they shared that they provided two types of readers including level readers and read aloud. The level readers were supposed to be used by students in a group situation, while level readers were to be used by teachers for whole class instruction. Teachers might have thought differently in terms of the adequacy of the SRP material. Nevertheless, SRP prides itself in pioneering the phonemic method in teaching Urdu and Sindhi languages, which has stirred a healthy language learning debate in the circles of practitioners and policy makers.

Cross Case Analysis

Content and SRP material

All stakeholders agree that the material provided by SRP is attractive and useful. It is considered useful in enhancing students' reading skills and engagement. However, they also observed the limited availability of material as a barrier to effective implementation of reading intervention in the classroom. Teachers found it difficult to align SRP lesson plans and reading material with the STBB. Although the SRP team and Guide Teachers/Mentors explained and showed the teachers that the given material is complementary to STBB, teachers struggled to find linkages. This may be caused by the short-term training the teachers went through along with limited follow-up, which was also discontinued after the first round of implementation.

Classroom Teaching

Generic pedagogical techniques used by teachers, as it appeared from quantitative data, indicate that SRP teachers tend to use more student-centered approaches as compared to non-SRP. Some of the examples reported by the teachers include storytelling, questioning, flashcards, games, role play and book reading. Similarly, reading instructions were also reported to be better in SRP classes. Some of the

examples include use of phonological awareness activities and use of story books provided by the SRP. The qualitative data indicates that teachers acknowledge that the use of SRP material and techniques help in improving students reading skills and the availability of SRP material and stories positively affected teachers' pedagogy and students' learning. However, they were found to be time consuming and sustaining them is challenging over the longer run without integrating them in the mainstream curriculum and without providing additional administrative support. Teachers found lesson plans helpful, but also lengthy as a result of which there was a dissonance in teachers' responses. While they accept that the phonemic method would help students in their reading, they also feel reluctant to continue it for longer duration due to it being more time consuming and due to the unavailability of support for implementation.

Physical infrastructure

The research shows that the physical infrastructure of the classroom affects teaching and learning positively. The SRP intervention did not include improved physical infrastructure as a main input. Therefore, there was little classroom infrastructure development by SRP. There were, however, smaller initiatives mainly by the teachers, SMC and some SRP field officers through which we observed some physical improvements, such as furniture purchase, classroom repair, solar panel installation, the establishment of reading corners, etc. These did have a positive effect on the overall teaching and learning environment, but these were not very prevalent themes in the qualitative data. Having said that, teachers and HMs raised concerns over SRP's approach of not providing any infrastructure support which at times became a bottleneck.

Parental engagement

The SRP interventions mainly tried to develop students' reading abilities through supplementary material and teacher training. The parental engagement was not a direct intervention focus of SRP.

Thus, the research did not find much evidence of how parents were engaged by SRP in the improvement of their children's reading abilities at home. There was quantitative data collected to see parents' home practices that support reading which shows positive results of parental engagement on children's reading though the sample was short and cannot be generalized. The qualitative data also suggest that some of the SRP selected schools were also part of the Community Mobilization Program (CMP). However, the focus of CMP was more on the increase in enrolment and community engagement and not on reading enhancement of children. Teachers and HMs did emphasize the engagement of parents' in children's learning (see more on parents' involvement under the theme of 'role of parents and community').

TLA VS GT MODEL: MOTIVATIONAL/ STRUCTURAL BOTTLENECKS

Within-Case Analysis

Research questions 3 and 4 focus on the lessons learnt from the TLA model and the GT model. The emphasis of the questions remains on finding the structural and motivational bottlenecks of both the models. The following paragraphs present the findings from qualitative data gathered from various stakeholders including: teachers, head teachers, guide teachers, SRP field staff, Government district officials, SRP central office staff and Government central officials.

Teachers and head teachers found lack of coordination to be one of the major reasons pertaining to SRP training. They observed that training nominations for teachers were not discussed with schools and instead came from central offices. As a result, there were discrepancies in nominations, for example many times higher grade teachers were nominated instead of Grade 1 and Grade 2. They further raised many concerns about training. Firstly, the duration of training of the SRP was very short. Teachers found training to be inadequate in making any lasting change. One of the teachers remarked "What would we learn in two days training?" Many teachers noticed that in order to

reconceptualize and adopt a new method they need more time. Secondly, they raised questions about the competency of the trainers and the adequacy of their training. Finally, they were unhappy, because they found training centers were far from their locations and they were not given any TA/DA support. Teachers and head teachers found the SRP materials to be less in quantity as compared to the number of students in their classes, and that the large classes required more material. Smaller schools also struggled in terms of space for storing the SRP material, as students were not allowed to take books home. Furthermore, teachers did not consider SRP material complementary to STBB; rather they took it as an additional subject and did not find it in cohesion with STBB. Teachers paid more attention to STBB as their children are examined on the basis of STBB by the education officials. Head teachers considered that SRP also had a weak monitoring structure and they could not involve government officials in monitoring, which resulted in inconsistency in the implementation of SRP interventions.

The teachers and head teachers also pointed to various structural bottlenecks that hindered them in implementing the SRP intervention. During the project implementation time the teachers were also asked to perform several other duties. For example, elections, census, polio vaccinations, exam invigilation etc. The teachers found it hard to deliver lessons to a large number of students and believed that the students-teacher ratio was beyond the ability of a teacher to teach effectively; some classes had as many as 80 children in a single room. A good number of primary schools also lacked an adequate number of teachers and had to teach multigrade classes. The SRP material and training did not help them deal with multigrade teaching effectively. Due to multigrade classes, it was hard for teachers to maintain students' interest and attendance. In addition, the poor infrastructure (such as poor toilets, lack of boundary walls and absence of classroom facilities) in many schools also affected teachers' morale in general negatively, which also has ramification for the teaching and learning process in school generally.

Guide teachers noticed a lack of interest amongst the teachers during training. Teachers felt disappointment as they were only provided with a traveling allowance, but not a food allowance; as a result they could not make teachers stay longer. One of the guide teachers shared, "Teachers showed less interest during training and therefore used to leave early". Guide teachers also felt that their own training was inadequate as they were not shown a model lesson by the SRP team. The timing of the training and material distribution was also found to be problematic as the training took place in the middle of school session and materials to schools were also delayed due to various reasons. One of the guide teachers explained the problems related to the material distribution and training timings as, "We did good work under SRP, but books and lesson plans did not reach us in time and training was short". Guide teachers also highlighted that many primary schools have shortage of teachers; many schools have only one or two teachers. Therefore, only one teacher could be sent for training at any given time from two-teacher schools. Sometimes it was difficult for the guide teachers to follow up with a teacher in school if he/she was absent during the visit. If a teacher was absent, the guide teacher could then only visit the next month due to his/her overall mentoring load and this delay resulted in low achievement. Guide teachers felt that the GT model was better than the TLA model, perhaps because the TLA model did not involve government trainers.

The SRP field officers, however, felt that the TLA model was better than the GT model as they could better coordinate things when all the trainers were from the private sector. Working with the government has its own challenges, for example, they could only have guide teachers for 8-10 days in a month as per the agreed TORs with the education department. In managing guide teachers, the field officers also had to rely on the TEOs and DEOs, which brought its own challenges. However, this model was considered more sustainable as it built capacity of the system. The field officers also had little understanding of the rationale behind changing the models in each round of training and they

also considered the training duration to be insufficient to achieve lasting changes. The changes in teachers' mindset require longer exposure under the guidance of exemplary mentors. The field officers considered the design of SRP somewhat problematic in recruiting trainers afresh for each round of training instead of having a permanent pool - finding experts each time was not always easily possible. The involvement of guide teachers in training and mentoring did not work well as per their understanding, as there were low financial incentives for them and, therefore, they showed less zeal in training and mentoring. Regarding the supply of material, they felt that it was shared as per the desired number, but sometimes they could not fulfill that number in large classes. They further explained that teachers, despite explanation, felt confusion in terms of using material. The level readers were provided for teachers for reading aloud activities, while stories were to be read in groups. Thus the material was not supposed to be one-book-for-each-child. Although they agreed with the timing issues, they felt that if the training and material had been provided at the beginning of the session, it would have been better. Due to this, the teachers might have felt SRP to be an additional book instead of complementary to Urdu or Sindhi textbooks. They did receive complaints from monitoring teams which indicated that some teachers, particularly older ones, did not show much desire to continue with SRP lessons and reverted to traditional styles. They might have conducted the classes as per the new method only during monitoring visits. SRP district officials considered that had HTs and government field officers supported the intervention and enquired from teachers about the new methods, the results would have further improved. These are the lessons that need to be incorporated in the next rounds.

The DEOs and TEOs shared their own challenges pertaining to SRP interventions. They showed their concerns about the planning and scheduling of SRP training where they have little involvement. As a result, the SRP training sometimes clashed with their other planned activities. Teachers are often called for various other activities e.g. polio vaccinations, census, visits of

judges. This affected the release of teachers for training and further follow-ups. One of the Regional Directors that we were able to meet during field visits felt lack of coordination between the SRP and the education department particularly at district level. Neither did the DEOs and TEOs like the TLA model where training was outsourced to NGOs as their own understanding was limited and it affected their relationships with teachers. The GT model did try to involve district and taluka officials and teachers in the mainstream activity, but generally their role was limited to attending some formal events or releasing teachers. One of them shared this concern as, "[the] DO Office's role is to cooperate ... we are called into the certificate distribution ceremony on the last day of the event". They felt that they were not involved in the monitoring of the project intervention. The DEOs and TEOs were concerned about several changes to the SRP model, which affected the desirable outcomes from project work. They also shared that they had no say in nominating the teachers; the names of teachers came from central government offices and they were obliged to release them for the said days. One of the district government officials highlighted this issue as,

...as far as nominations [of teachers for SRP training] is concerned I receive a list either from the Directorate or from the SRP ... they should first talk to us, tell us their requirements, then we will tell them who are the teachers of class 1 and 2. I would want to nominate such teachers who I can nominate as master trainers. This way the project can be sustained. I did not get such a chance to identify.

They felt that SRP monitoring of the field was weak; as a result they could not see much difference and the results of Sindhi and Urdu were still weak. Neither did TEOs check about the SRP work during their school visits and instead focused on the completion of the official STBB syllabus. The zero period as mandated by the official notification of the Secretary School Education was followed weekly in schools, as there was little check from supervisory officers. DEOs and TEOs explained that SRP requirements in the reading lessons

did not consider the variety of contexts that exist in Sindh schools e.g. large classes, multi-grade teaching and one or two teacher schools. It was a challenge in districts such as Kashmore to release teachers for the workshop, where most schools are one- or two-teacher schools. The lessons were considered to be time consuming. The TEOs complained that guide teachers who were appointed by the government were not provided with sufficient travel budget and allowances which hampered their school visits. Additionally, there was no provision for TEOs visiting schools to monitor progress. The other major bottleneck was the coordination with the bio-metric team. The guide teachers and other teachers need to be away from schools to attend training and sometimes the bio-metric teams did not have information about their participation; hence they were marked absent.

According to central government officials, the teachers reacted differently to SRP material and training: some considered it complementary while others considered it as an additional burden. One of the central government officials shared; “[the] SRP has provided lesson plans, but we do not provide lesson plans, we only provide textual material. The materials taken to school were with the lesson plans and teachers would have definitely found it different and useful”.

One of the Directors in the Education Department thought that had they been taken into confidence at the beginning of the project they would have been more engaged and would have monitored the teachers who reverted to the old methods. Another issue with the training was its timing, which came in the third quarter of the academic session, hence teachers struggled to implement the new method and material in the middle of the session when they had already completed three-quarters of the STBB syllabus. Another line department complained that the SRP contacted them at the last moment for expert selection while the department was not deeply involved in material planning and development. The training plans were not shared with them on time which sometimes resulted in disturbing or overlapping with their own training plans. One of the government officials shared;

“We had lack of coordination with SRP. They were working in isolation. During our training, they would come and ask us for sparing our teachers for training. I would not spare the teachers as they were engaged in various activities. If I send a government teacher for [SRP training] their own classes would suffer ... this program should have been aligned with the CPD program of government, it would have been sustainable”.

Central government officials also considered the GT model better as it has continuity with the government guide teachers, while the TLA model was not preferred due to engagement of private mentors, who would not stay with the system.

A conversation with members of the SRP central office helped us understand the reason behind various changes in the training models. According to SRP central officials, the conceived model was different from the one which was initially implemented. The TLA model was designed keeping ‘Supervisors’ as the core cadre of government field officials who would play the role of mentor. However, when the project formally launched, the cadre of ‘supervisors’ was abolished. As a result, they had to engage private trainers and mentors in the TLA model. The change from the TLA model to the GT model was circumstantial. Reasons for the changes made in the design included changed focus of USAID from maths and upper classes to only languages in lower grades i.e., 1 and 2. The change from TLA to guide teachers was due to the realization that the government teachers need to play the role of mentor if the capacity of government needs to be built. The SRP officials were also mindful of the short duration of the training and shorter follow-ups of 4-5 months after training. They were, however, hopeful that some of the schools would be part of Cohort-1, Cohort-2 and Cohort-3 and thus will have better chances to display changed pedagogical practices towards reading. The SRP officials also felt that at times the government officials were too busy and they did not correspond in stipulated time or did not respond at all, which created

a communication gap and often only superficial understanding. SRP officials also felt that low scores of children may have been caused by teachers not implementing the lessons as per their spirit in the classroom. Regarding the TLA and GT models, the SRP officials also aligned with the government counterparts and considered that the TLA model was not sustainable as it relied on outsourced TLAs who were not from the government system and that the model was also fairly expensive. Although, it was not the focus of research; the SRP officials shared that the model is now further changed in the third cohort and engages head teachers more centrally, due to the unwillingness of the government to spare around 400 teachers as guide teachers for another 6-8 months. These contingencies from the government and other circumstances made the SRP change its model at each cohort. The SRP officials were mindful that teachers took SRP material as additional, because of the lack of ownership by district officials despite the SRP's best efforts to involve them. There were other challenges that created further hurdles for the SRP such as the biometric system marking teachers absent when they were actually on training and teachers' demand for extra compensation to attend training.

Cross-Case Analysis

TLA vs GT: a comparison

Different stakeholders hold different points of view regarding the challenges and opportunities in both the TLA and GT models. TLA is considered better than GT mostly by the SRP officials, because it was found more effective as they could include competent private trainers and monitoring staff, who may not necessarily have been available with the government, though they acknowledged that it had challenges of sustainability. On the other hand, the GT model was considered better mostly by the government stakeholders as it engaged government teachers as guide teachers and also built their mentoring capacity to provide classroom support. Some of the government officials wanted to have a greater role for TEOs in monitoring the SRP intervention. The GT model therefore tended to be

more sustainable as more government teachers and staff were involved, which builds system capacity better than the TLA model.

Lack of coordination

Lack of coordination amongst various stakeholders was highlighted as one of the major factors that hindered effective implementation of the SRP intervention. The lack of coordination particularly at sub-district and school level resulted in nomination of irrelevant teachers for the training. Sometimes due to the lack of coordination, SRP training coincided with the training planned by government line departments, resulting in an unpleasant situation and unavailability of teachers for either of the partners. An issue of biometric team was also highlighted by teachers suggesting that sometimes the biometric monitoring teams considered them absent while they were away from schools to attend SRP-related activities.

Training issues

Stakeholders highlighted various issues pertaining to the SRP training which negatively affected the intervention. Some of the major problems included the short duration of the training as it comprised of 3-6 days of initial training followed by classroom support 1-2 days a month for up to 6 months, which was considered short by the majority of the teachers. The teachers also highlighted the low competency of the trainers as another training-related challenge. Many teachers complained about the non-provision of TA/DA support during training as a challenge, because they had to afford their transport cost on their own. A major training issue was related to the timings of the training; the teachers, GTs and even SRP district officials felt that training needed to be synchronized with the school academic year, the TLA and GT training was carried out in the middle of the academic year and ran through examination time and during student transiting to higher grades. This non-alignment of training timings was a troubling issue for all.

Inadequacy of material

It was noted by most of the teachers that the SRP material was inadequate in terms of quantity. They

further felt that the material was not aligned with their curriculum and textbook. The SRP teams explained that level readers and read aloud stories were supposed to be used in groups, but perhaps many teachers did not understand the message clearly. Teachers also felt challenged to apply cooperative learning strategies due to overcrowded or multigrade classrooms. Lack of storage to keep the material safe was also mentioned as an issue by HMs and teachers as they were supposed to keep the material in school for use; children were not allowed to take material home.

Monitoring mechanism

Despite the efforts of SRP to involve government, the monitoring of SRP initiatives by the DEOs and TEOs remained weak. In the GT model, the involvement of GTs from the government teachers did improve the ownership at the department level, but it was still weak in terms of monitoring the intervention at taluka level. Even when TEOs visited the schools, they mainly focused on the STBB syllabus coverage. The stakeholders think that the program would have been sustained better had the SRP involved district level government officials more in the monitoring of the program during the rest of the intervention. Due to lack of involvement of DEOs/TEOs, policies like the reading periods are not being implemented fully in schools.

Large classes and multigrade teaching

A concern raised by many of the stakeholders was that SRP intervention had weak consideration towards some contextual issues in the public sector, such as large classes, multigrade teaching and schools having only one or two teachers. Teachers found it hard to follow the SRP methods and materials in the classes which were overcrowded. Due to the large number of students, it became very difficult for every child to gain benefit from the SRP material. Teachers reported that SRP material and training could not help them deal with multigrade teaching while ensuring the SRP methods were used.

Lack of motivation and ownership

Most of the government officials and school teachers were reported to have shown less motivation and

ownership towards the SRP intervention. Some of the senior teachers were found to have reverted to the traditional methods of teaching soon after the conclusion of monitoring visits by the GTs. Teachers felt that their responsibilities were only until they had visits by the GTs or the SRP team. They shared that their students are assessed on the STBB textbooks and therefore they teach as per the requirement of that syllabus. On the other hand, SRP methods required more resources and efforts, which were difficult for them to continue with.

Weak infrastructure

Generally, the infrastructure of the schools which came under the sample was not too bad although there were issues with non-functional toilets, shortage of rooms, contaminated/non-availability of drinking water, non-availability of electricity etc.). The weak infrastructure was demotivating for teachers in their teaching and learning situation. In more remote areas, there were shortage of desks which also caused trouble for teachers during teaching. The provision of physical infrastructure as a main input was beyond the scope of SRP intervention, therefore there was little classroom infrastructure development by SRP in the intervention schools. However, some teachers and SMC led initiatives were observed that enhanced physical infrastructure (e.g. reading corners, furniture purchase, classroom repair, solar panel installation) in schools.

IMPLEMENTATION OF SRP: ROLE OF STAKEHOLDERS

Within-Case Analysis

Role of teachers and HM in implementing the SRP: The teachers' prime role was to be a competent reading teacher. According to HM and teachers, most of the teachers found SRP as additional and difficult task to carry out and found it incoherent with the STBB, hence they often switched to the traditional methods of teaching in order to complete the STBB course within the stipulated time. However, a few teachers adjusted and taught using SRP method for certain period of time. Some teachers were not clear on how to use the SRP material and which students to provide the books as the

books were less in number. Some teachers were strictly told not to hand over material to students. However, for equal distribution, they managed by giving one book to two students. Similarly, there were some self-motivated teachers who delivered all lessons of the SRP planner while some others succeeded taking two to three lesson plans per week using the SRP method and resources. Teachers tended to continue the interventions by making pictures, getting students to color them and writing stories. Some took initiatives by inviting parents to school either by going to parents themselves or by sending notices through their children.

Head teachers had a very limited or no role in the SRP interventions for Cohort-1 and Cohort-2. Although they were not involved, they generally interacted with the SRP teams when they visited or cooperated with guide teachers during their class observations and children's assessment. The head teachers complained that they were not included in SRP and were thus unaware of SRP intervention. However, some Head teachers were observed playing active role in the intervention despite having no formal role in SRP intervention. The HMs believed in effectiveness of parental involvement in the school, hence they asked parents to meet teachers once a week and checked on their children's copies and learning in school. In order to increase the attendance of the students some HMs claimed to have provided incentives to students. A few HMs also performed the duties as guide teachers. Since the monitoring was weak, one HM claimed that "when SRP comes to our school, we teach through their method and once they leave we use old method". Similar claims were made by many other HMs and teachers as well.

According to the SRP field officers, the HMs were mainly responsible for ensuring high enrolment and full attendance of teachers. A head teacher from Cohort-1 school complained that SRP had directly involved teachers and excluded HMs and thus that HM disassociated himself from SRP training. The SRP field officers opined that the HMs had conventional approach and the relationship with teachers was not very cordial which resulted in discouraging teachers.

With respect to teachers' role, the SRP field officers observed that teachers were overburdened; despite material being provided to them, they were still using old methods of teaching and were unable to use SRP material. The SRP field officers believed that teachers like these eventually made a difference by their own initiatives.

According to DEOs and TEOs, those HMs who were also GTs could not play active role in their own school as they were from campus schools and would spend only one day in their own school while 5 days per month they were in other schools of their cluster. Moreover, some old HMs were considered politically active and refrained from academic activities, whereas the new HMs were inexperienced though with positive mindset. DEOs and TEOs were of the opinion that teachers faced difficulties in teaching through the SRP method as they took the additional responsibility of developing lesson plans. According to provincial government officials, it was teachers' responsibility to develop lesson plans according to their contexts.

Role of parents, community and SMC in implementing the SRP: According to HMs and teachers, parents were illiterate but very supportive. Despite being illiterate themselves, a few parents visit school and ask about the progress of their children. However, in rural settings, most of the villagers are illiterate; therefore they could not support their children at home and rely on tuitions in order to support them. Although very few parents were educated, those who claimed that they provide support to their kids at home, did so in terms of replica of the way children were taught in schools i.e. traditional style, which did not benefit their children. Most of the teachers and HMs repeatedly told that the parents were not involved by SRP, thus their involvement was limited and only restricted to appreciation and limited feedback.

A few HMs from cohort-1 commented that the USAID Community Mobilization Program (CMP) and the parents-teachers meeting were successful as they helped in bringing parents to school. This is, however, the case in such schools where, along with SRP, CMP

project is also running. According to SRP central officials, there was no coordination with CMP and they were unaware if CMP was focusing on reading. With regards to parental engagement the SRP officials claimed to engage parents in cohort-2, but only in a limited way. Now SRP is trying to engage parents in cohort-3. Generally, the teachers and HMs believed that home environment of the children was not conducive to reading as their poor parents involved them in domestic chores which resulted in low attendance. In most of the cases, parents' participation was limited to SMC involvement, which also was minimal. Except one school, where two teachers were provided, SMC could hardly play any further role in academic development beyond increasing children's attendance, consulting parents and using SMC budget for infrastructure/maintenance. With regards to reading, there was insufficient support from SMC. According to DEOs/TEOs, difference in home and school languages also create reading challenges for children in school which needs to be addressed.

Role of government etc. (SRP) in implementing the intervention of SRP: The teachers and HMs claimed that DEO had a major role in nominating the teachers for the training. They also felt that apart from weak support and nomination, the DEOs and TEOs were not deeply involved and there was no significant support from DEOs in academic and infrastructural matters other than SMC funds. Similarly, there was no input and follow-up of DEOs and TEOs regarding reading and SRP intervention. As one teacher remarked, "No one from TEO office comes to ask about this new intervention method." It was also believed by the teachers that SRP did not coordinate with district and Taluka officers and due to this lack of coordination between government and SRP, the district and taluka officers did not play an active part. According to Teachers and HMs, the major focus of government officials was on timely completion of the STBB syllabus and TEO is more connected with school as compared to DEO. It would have been useful if TEOs would have been given a more active role by SRP. HMs and teachers also complained that SRP collected data and made

promises but did not work effectively while considering the ground realities of the schools, thus, activities became more of a game than learning.

A majority of guide teachers were actually government teachers, recruited recently through the aptitude tests conducted by the IBA-Sukkur or NTS, and were engaged by SRP in the training of Cohort-2. The guide teachers' role was to conduct cluster training and provide support to the teachers in their respective schools. During their schools visits, the guide teachers supported, observed and then discussed feedback with teachers and students. Guide teachers shared that they taught through stories and maps to deliver model lessons and which enhanced students' interests. During the school visits, the guide teachers would first discuss the plan with teachers and then test students and discuss on how to improve. Guide teachers in this manner used to facilitate teachers in classrooms. During this process, the GTs would guide on how to improve lessons for example by teaching through play, sound, writing etc. During training, the guide teachers used to tell teachers to teach through gestures and expressions. The GTs claimed that they were trained for three days which focused on lesson planning, observation and use of tablet. They worked for six months as GTs where each GT had six schools and worked for twenty days per month in their respective schools while for remaining ten days they performed SRP duties. GTs visited twice per month per school, one visit for assessment and the other visit for observation and classroom support. According to GTs their roles included visiting schools, training teachers, conducting pre and post class conferences with the teachers, delivering model lessons, observing them and giving them feedback as well as presenting their reflections in SRP meetings. Guide teachers revealed that SRP was responsible for selection of GTs and also conducting assessment. GTs were also required to inform the teachers before their visits to prepare 10 lessons and deliver one for observation. Their focus in the observation was on utilization of the provided material and student engagement in the class. However, GTs did not emphasize on the last sound of the word which was assessed in the tests at the end.

The SRP field officers shared that recruiting GTs from the government teachers' cadre was an intentional move to develop pool of government mentors. GTs were responsible for conducting meetings with their SRP district official once in a month. The SRP district officers would also visit GTs and teachers in schools. However, the government officials did not frequently visit the schools and followed up with the GTs. The SRP field officers mentioned that they did not liaise effectively with the PMIU which could have brought more coordination at district level with other components.

The major role of DEOs and TEOs were in the selection of GTs and coordinating for conducting workshops and trainings. All GTs were NTS/IBA-Sukkur recruited teachers and received extra allowances for the project work. In some of the districts TEOs paid surprise visits and visited up to 3 schools each day. However, they did not monitor and observe classes of the SRP teachers as in their opinion the monitoring of SRP intervention was being done by the SRP team only. DEOs/TEOs were sometimes not very pleased by SRP, because of mismanagement in the training of SRP and non-seriousness of the projects staff.

The central government officials at the Education Secretariat played a key role of overseeing and greater coordination. According to central government officials, SRP provided them their material, shared the training plans, list of schools, teachers and guide teachers for training. These were then approved and shared with the line departments and division and district officers. The STBB, PITE and BoC were also engaged in the provincial consultations as participants. However, they did not play a greater role in material production and designing and conducting training. Thus, the relationship with line departments and SRP remained that of participant and not of co-developer. The material produced was duly approved by the STBB and BoC. The district officials at Taluka were also involved in the distribution of material from district offices to schools.

According to central SRP officials, Sindh government gave ample importance to the SRP reading

intervention. Government counterparts (BoC, PITE, STEDA, STBB) were all taken on board and there were regular communication and meetings to critique and make changes accordingly. Various committees at different levels were actively working to facilitate the project. For example, reading standing committee and provincial steering committee were both working at the central offices with ample participation of various government officials. Material developed and tested in the field was brought back to these committees for further review and finalization before being rolled out in the field. The provincial government also played positive role in strategizing training, developing material, reviewing curriculum and approving interventions. The Curriculum Wing at the Education Secretariat played the central coordinating role from the government side for ensuring smooth coordination while selecting schools and teachers for SRP intervention and training. Talking about their role and contributions, SRP officials stated "reading-specific SLOs (reading performance standards) were not developed in our curriculum and now we have added reading-specific SLOs in the curriculum and submitted it to the Curriculum Wing to get reviewed and approved". During the life of the project, SRP ensured to work with specialists from all stakeholders to design tools and develop material and remained in close coordination with the Pakistan Reading Program; together they helped design the CPD Framework and Reading Standards.

Cross-Case Analysis

Role of teachers and head teachers

Teachers, especially those who received SRP training, were the major stakeholders in the SRP intervention and were responsible for applying the SRP methods and materials in their classrooms according to their respective contexts. Some teachers who were self-motivated tried to adopt the intervention whole-heartedly, made serious changes in the teaching and continued the SRP methods even after the intervention. Some of the teachers and HTs played active role in mobilizing and engaging the community

in children's education; however, this mobilization was mainly in terms of their general educational attainment and not specific to reading enhancement. HTs were not assigned any specific role in SRP intervention for the TLA and GT models, except those HTs who were selected to be GTs. Despite this, the HTs were found to have played effective role in various aspects of school development. HTs, along with other stakeholders of their schools, ensured attendance from students and teachers and provided supportive environment for teaching and learning. One of their main roles in terms of SRP intervention was to allocate the right classes with manageable teaching load to the SRP trained teachers and keep them assigned through the intervention period.

Role of parents and community

There are very few parents who, despite being illiterate, get themselves engaged in the schools' matters and support their children in their education. Most parents being from working class background and illiterate, are neither able to visit the schools nor support their children's education at home. Instead, they send their children for tuition, if they can afford it. Even if the parents support children's reading at home, they will use the traditional method, rather than phonemic method, as they themselves are not oriented towards this method. HTs consider the parents-teacher meetings and the presence of SMC useful for improving community participation, but very few parents attended such meetings and their role remained limited to mere appreciation. A few parents are part of the SMC, but their role is minimal regarding the SRP intervention. SMC generally works on improving students' enrollment, parental participation, school budget utilization and arranging additional teachers. So far, SRP has not fully engaged community in the intervention, although it can play a useful role. The SRP schools, that happened to be the schools where the CMP program was also working, had more active community due to the construction of a state of the art school buildings, but the focus on community mobilization was weak in terms of the objectives of SRP intervention.

Role of government stakeholders

The government stakeholders operated largely on two levels: central level and district/taluka level. The central government officials from the Education Secretariat provided essential support and ensured smooth coordination between SRP and other related departments. The STBB, PITE and BoC were some of the major government departments which were engaged as stakeholders in the intervention. At the district and taluka levels, the DEOs and TEOs remained as participant; facilitative but not very active. The DEOs/TEOs made visits to the schools but their focus remained on the STBB syllabus coverage and not on the SRP intervention. GTs were another group of stakeholders whose role demanded school visits, training teachers, conducting pre and post class conferences with the teachers, delivering model lessons, observing them and giving them feedback. According to SRP officials, Sindh government and its other counterparts (BoC, PITE, STEDA, STBB) gave due importance to the SRP reading intervention and were also involved in the implementation of intervention.

TECHNOLOGY-BASED ASSESSMENT

Within-Case Analysis

According to cohort-1 and cohort-2, HMs and teachers, TLAs or GTs made regular visits and took students tests on tablets. Teachers did not share any technology based support by the SRP in teaching learning process. The main purpose of using tablets was to carry out technology based tests. Although HMs were unaware of the tablet based tests, they felt confident that children were able to perform well as the tablets were user-friendly technology.

GTs admit that in the training they learned how to use tablets for carrying out formative assessments and also classroom observations. The SRP field officials complained that the number of tablets was less and sometimes they faced problems due to tablets not working at the last moment. Many GTs were also uncomfortable with technology-based assessment and encountered initial challenges. One of them shared,

“...in some schools, the use of tablet was difficult for children. They were struggling and some children even did not see the computer before. So, the tablet based assessment was difficult. They were familiar with teaching through books but going through assessment using tablet was difficult.”

DEOs/TEOs mentioned that the government guide teachers were trained to handle tablets and used them for mentoring which positively affected reading skills of the students. Likewise, the government central officials appreciated effectiveness of the assessment conducted by SRP. According to the SRP central officials, apart from development of reading performance standards, the SRP team worked on ICT-based formative assessment, which was not a deliverable of the project, but helped in overall intervention. The TELE TALEEM introduced a software for formative assessment and online automatic feedback for which the mechanism was not there in cohort-1. SRP team says that TBA was a novel idea. It helped in giving immediate feedback and helped revisit the strategy. Moreover, ICT-based formative assessment remained successful in terms of improving students learning. It has a bright future as it is appreciated and owned by Sindh Government.

Cross-Case Analysis

Limited use of technology

The SRP intervention included the use of technology in assessing students learning. The tablets were used to make digital assessments of students. There were two assessment rounds carried out at baseline and midline and another will be carried out as end-line. The intervention team also used tablets for formative assessments in order to understand students' learning. The TLAs and guide teachers used tablets for carrying out formative assessments. However, there was no use of technology by the teachers in enhancing their pedagogy and the use of technology remained limited to the automation of assessment rather than in improving pedagogy.

Technical glitches in handling technology

The TLAs and mentors felt comfortable with the use of technology. However, the interviews with the guide teachers revealed that they faced some glitches with the technology, for example during trying out assessments; they accidentally submitted half-filled formative assessment forms. At other times, some tablets encountered technical faults which hindered the assessment process.

SUSTAINABILITY AND SCALING UP OF SRP

Within-Case Analysis

HMs and Teachers of cohort-2 suggested various ways to improve, sustain and scale up the SRP intervention. They considered SRP and STBB materials to be different from each other and not aligned, hence aligning the two will improve the sustainability and effectiveness. The SRP lesson plans should be developed in accordance with STBB. Some HMs and teachers made an interesting demand that the government should print زبر and زير in books, in color, which would perhaps improve the phonetic awareness. In fact, some of the books observed in the field work by the research team had زبر and زير marked by the teacher to facilitate students' reading. Some HMs and teachers demanded GTs to have frequent visits to the schools, at least twice a month, to see how SRP is taught, while others recommended that neutral teams should monitor the schools on regular basis. A few of them opined that exam committee by TEOs should emphasize on SRP material assessment, which will have a backwash effect. With regard to material resources, majority of the HMs and teachers identified shortage and/or inadequacy of resources as a major reason hampering their teaching learning process. Teachers and HMs demanded more ready-made materials like books, charts, flashcards and CDs to teach better. They said that SRP will improve further if the intervention is initiated in the beginning of the academic year and the SRP materials taught as compulsory subjects. It was highlighted by the teachers that due to the large number of students they require more material for effective classroom teaching.

HMs and teachers felt that the duration of training was not enough. They recommended that the training should last at least for 15 consecutive days instead of 2 to 3 days. Furthermore, training should be conducted on regular basis and SRP should increase its monitoring by appointing separate GT for each school. A teacher felt strongly about longer trainings and said “6 months’ intervention [training] is not enough. It should be at least for one to two years”. Teachers also mentioned that the training centers should be situated nearer and separate for the female teachers as they face traveling issues and are not comfortable with their male counterparts. Moreover, some teachers requested not to conduct the trainings during vacations and should be informed 2-3 days prior to the training. Further, the teachers wanted to be compensated through TA/DAs. Some of the teachers suggested that the training should be provided to all teachers of in the selected schools since the classes keep on rotating amongst all the teachers. They suggested that the HMs should also get the trainings so that they can monitor the teaching and learning effectively. A few teachers suggested that one class should be handed over to one teacher for the whole day to teach everything (including SRP material) while others believe that the SRP teacher should only be focused on SRP task and the reading classes be given half an hour daily. It was observed that the reading period was loosely followed in the school; the research team did see some timetables mentioning the reading period every day, but it was mostly symbolic with little implementation. Teachers and HMs strongly suggested that their nominations should not be done by TEOs or higher ups, it should be done by the HMs as they know the school better and are responsible for supervising teachers upon their return. The guide teachers can be nominated by the TEOs.

The teachers and HMs shared various important suggestions for sustainability of the SRP intervention. The most important one was the suggestion to solve the issue of teachers’ shortage so that dedicated teachers can be provided for reading classes. It was recommended that HMs greater involvement should

be ensured so that there is leadership support for the SRP trained teachers and HMs can hold teachers accountable to deliver SRP material and follow lesson plans as envisaged originally. The teachers and HMs felt that it would have been useful to train all available primary teachers in the school to overcome teachers’ shortage and high turnover. The HMs have to involve teachers in all classes and sometimes the SRP trained teachers have to be moved to higher grades. In order to sustain the SRP lesson plans and material, it would be important to align and if possible make it part of the STBB textbooks. In addition, they suggested more materials to be provided to the students including books that contain moral stories and advice. Regarding the duration of the program, teachers and HMs suggested to increase it beyond six months, at least for 4 to 5 years to be more sustainable and repeat the training every year, with regular monitoring by the central team along with government officials until it becomes the habit of the teachers.

The guide teachers also felt that the 3-day training was not enough as it was too short. Some of them suggested increasing the duration from 3 days to 10-15 days each month. They said that the teachers were experienced, but they needed to learn new methods of teaching. Most of the GTs wanted the higher authorities to repeat and sustain the intervention over a longer duration. They added that the program is going to be beneficial even if 60% to 70% of the initial target is achieved. The guide teachers suggested combining SRP material with Urdu textbooks of STBB. They strongly suggested to start the training and intervention in April so that it is aligned with the academic calendar of the Government of Sindh; otherwise teachers already start teaching following traditional methods and it is difficult to make them switch in the middle of the academic year; it is equally difficult for the students to adjust to it too.

The SRP field officials shared that the TLA model had better training in terms of number of days which were around 6, while trainers had more time available for the training. The quality of trainers was also considered better in the TLA model. The officials

suggested continuous monitoring and evaluation and more commitment from TEOs and DEOs for sustainability. Moreover, they suggested that the government should plan for regular capacity building of teachers to sustain SRP interventions and the phonemic method of teaching language. The SRP field officials suggested that duration of the training should be increased and can be sustained after 2-3 continuous cycles. They also considered involvement of HMs in training in the upcoming cohorts as positive. Most importantly, the SRP field officials considered that there was a strong need for better synchronization at the district level and recommended that the PMIU should involve the education department more closely at the district level.

The DEOs/TEOs recommended that the IBA and NTS recruited teachers should be trained as master trainers due to their merit and experience. DEOs/TEOs were also of the opinion that duration of the teachers training should be increased from 3 to 6 days or if possible, two weeks. The DEOs/TEOs suggested that all main stakeholders should be taken on board and involved all relevant officers (including HM, DEO and TEO) so that a continuous follow up is ensured and teachers could work better. One of the officials said, "They [SRP] should train fewer teachers, and train them well and for longer duration – 1-2 months long training. We can effectively use such trained teachers in the future". Moreover, the SRP material and STBB books be merged so that a sustainable change could happen through changes in the existing books/curriculum. They opined that the new method of phonics for Urdu and Sindhi should be made mandatory. Several steps were taken at DEO and TEO levels to make the program successful such as encouraging teachers and HMs to work together in the same school and banning the transfer of SRP trained teachers by the TEO to ensure continuity. One of the DEOs thinks that the GT model was better than the TLA model as it involved government teachers more. However, to decide on the best model, SRP should first experiment, look at the results and then implement the program in schools. One of the DEOs stressed on this

aspect and said, "We should first experiment at a small scale and only after successful result, we should expand. We need to know if students are learning and then scale-up the intervention." Also, they highlighted that the SRP district coordinators should be made stronger and the intervention should work the whole year and they should not leave without ensuring sustainability.

According to central government officials, for sustainable intervention, the education department and teachers must be taken on board before the intervention. They recommended that SRP should include all schools of the districts so that more people from each district take part in the trainings and not only favored by the project people. The project design of the SRP should be aligned with CPD framework of STEDA as it would provide long term sustainability of the intervention. One of the central government officials highlighted, "I find it effective to focus on assessment, however, the work on CPD is also very important. They have to pay attention to their sustainability". Moreover, they emphasized that SRP needs to be more organized in its work and that all future projects should coordinate with ground level departments during planning stage. The central government officials commended the materials developed by SRP and its effectiveness towards enhancing students learning. The officials of central government, showed commitment on behalf of the government to adopt the SRP material and were ready to explore the possibility of STBB printing the material developed by SRP for school distribution after the project completion. However, the STBB officials were not sure about getting the required financial support as they do not have any additional budget available for this purpose. One of the central government officials shared his concern as follows,

"...the textbooks printing require budget support which comes from the government. We can only work within the available budget. For any additional work [e.g. SRP material] we require additional budget ... this is the main problem, we can do [printing and] supply material but we have the problem of budget."

Besides, they were of the opinion that the SRP materials/methods can inform textbooks which can provide basic guidelines for the teachers to follow phonemic method. There was some disagreement among the central government officials about aligning SRP material with the STBB, while some felt that it is aligned, others wished that SRP should have been advised to align their material with national curriculum and suggested that there should be a session with subject experts of STBB to sort out important tips from SRP lesson plans and get them reviewed by the BoC. One of the government officials suggested,

“[In order to insert reading related tips in the textbooks] you can have a session with subject experts of STBB and can discuss the things that can be inserted at appropriate places in the lessons. These additions need to be sent to the Bureau of Curriculum, which will reply after its thorough review.”

Also, it was highlighted that the SRP material exclusively focused on ‘reading’, while the national curriculum and textbooks aim to develop other language competencies side by side e.g. writing. They also suggested teaching *بارع* ‘earaab’ to children from early childhood. One of the Central Government Officials, pointed out that the SELD has prepared a revised Scheme of Studies for grades 1-5, which has assigned a separate library period. This period can be used for improving reading and can provide a sustainable option for SRP efforts after the project life.

The central Government officials opined that efforts should be made by both SRP and the government to sustain the project interventions for improving language; the government is serious about sustainability. According to the officials, follow-up is important which should be done by the Government master trainers. The education department is motivated to take up the work of SRP as it is aligned with various other government initiatives. The PEAC can take up the technology based assessment, ECE policy can take up literacy initiatives of SRP for early grade reading.

According to the SRP central officials, SRP has done its level best to ensure sustainability and long term effects of the intervention. Some of the government people have been trained to analyze and report assessment data which was previously done with the help of home office. SRP is currently working on Cohort-3 which is focusing on the sustainability and tries to remain in line with the CPD framework that was initiated by the Government of Sindh. The SRP officials believe that the alignment of SRP material in the text books should be made part of the scheme of study, which is already with the SELD. The SRP has also contributed to the national curriculum through developing reading performance standards which will be one way the project initiatives will sustain. The reading performance standards have been developed with PRP, validated and approved and will be implemented in two subjects next year. The SRP and PRP have been working together to enrich existing textbooks of STBB and prepare supplementary reading material, but the work is still in process. SRP is also working on a proposal in coordination with the government counterparts and due to submit to SELD (school education literacy department) for inclusion in 2018-19 budgets so that regular budgets can be allocated for the material and trainings through STBB and PITE. Moreover, SRP is working in partnership with the Sindh Government to replicate early grade reading and ICT based assessment in 4500 schools. Sindh government also wants SRP to support these 4500 schools on high priority and cover 70 -80% students enrolled in these schools. Overall, central SRP officials are hopeful about sustainability of the program as all the departments are on board and experts have been identified and trained to keep on working. For core funding for HR development, proposal has been submitted to the relevant department while experts from BoC, DCAR etc. are working on the issue of phonemic way of teaching language. After SRP project finishes, some bits of it will continue by PRP and SBEP until 2020.

Cross-Case Analysis

Alignment between SRP and STBB material

Most of the stakeholders are of the view that the SRP

material and the textbooks of STBB are not-aligned, especially at school level. The teachers who were the main implementers consider SRP material to be 'additive' and not complementary. Stakeholders at school and district levels are of the opinion that the SRP methods of teaching through phonics, will sustain only if the required changes will be made in the textbooks of Sindhi and Urdu subjects. Moreover, it was highlighted that the SRP material was more focused on 'reading', while the national curriculum and textbooks required the teachers to work on many other skills (e.g. writing) of the students as well.

Timings for intervention

It was suggested by the stakeholders that SRP intervention will sustain and improve further if the trainings will be conducted in the beginning of the academic session. The academic year starts in April while the SRP interventions initiated around December. The stakeholders also mentioned that since the teachers and students become habitual of the traditional methods by mid of the academic session, it is difficult to change their perspective and make them follow a new method for the rest of the academic year. It was also difficult for the students to experience a sudden change in mid-year.

District / school level coordination

It was emphasized by stakeholders at district and school level that all relevant stakeholders, including the HMs, should be involved in the intervention for better and sustainable results. It was further suggested that for sustainable intervention it is necessary to take the education department and teachers on board before the intervention. The SRP field officials felt the need for better coordination at the district level and suggested that the PMIU should involve the education department more closely at district level to ensure long lasting results. The coordination between the SRP and SELD at central level was better and the SELD and other line departments were more involved. However, some coordination issues did arise related to various aspects of SRP intervention.

Duration and follow-up of the training

Majority of the stakeholders, particularly at school and district level, considered the training duration as insufficient to prepare the teachers to teach through the phonemic method. They recommended that the training should be conducted at least for fifteen consecutive days instead of three or six days. In addition, it was suggested to increase the intervention beyond six months, at least for four to five years and the refresher training be conducted every year with regular supervision from all parties until it becomes deeply entrenched in teaching culture. Various recommendations were shared by different stakeholders to strengthen the monitoring and follow up mechanisms before and after the SRP intervention. At the school levels, it was suggested that GTs and TEOs should carry out joint monitoring visits and ensure the implementation of SRP intervention. Moreover, it was also recommended that the government should plan for regular capacity building of teachers to sustain SRP interventions and the phonemic method of teaching language.

Adequacy of material

Shortage and inadequacy of resources was highlighted as a major reason by the teachers that is hampering the process of teaching and learning. Teachers and HTs therefore suggested that more ready-made materials such as books, charts, flashcards and CDs are required at school levels to ensure that teachers follow the SRP method and engage students in active learning. This is, however, in contrast with some other teachers who complained about the lack of storage for SRP material. It would be useful to provide training in the development of low-cost material utilising the resources from the surroundings. Moreover, it was also recommended that the exam committee which works at district level should also assess children in light of the phonics methods of reading which will improve the sustainability of the SRP intervention.

SRP officials' view about sustainability

The SRP officials said that they have done their level best to ensure sustainability of the intervention. Some

of their contributions included: capacity building of government officials, inclusion of the SRP recommendations in the revised scheme of study, development of reading performance standards, preparation of supplementary reading material and enrichment of textbooks. SRP is also working in partnership with Sindh Government to replicate early grade reading and ICT based assessment in future. The SRP officials are hopeful about sustainability of the program as all other stakeholders show willingness to continue with the initiative. They are sure that the

central government and other concerning departments are taking ownership of the SRP intervention and showing commitment in its continuation. It seems that the government is ready to give space for SRP recommendations in their revised scheme of studies to make the initiative sustainable through declaring library period.

6

Conclusion and Recommendations

STUDENTS PERFORMANCE ON EGRA

It is encouraging to note that most of the assessment results were in favor of the SRP group. In other words, the assessment results have managed to identify differences across groups (intervention and comparison groups). Having said that, the magnitude of the difference is rather small. It is important to note that regardless of cohorts (Cohort-1 & Cohort-2) and groups (SRP and non-SRP) the results have followed the same trend. Out of the six domains the children performed the best in 'expressive vocabulary' followed by 'listening comprehension', 'oral reading', 'non-word fluency' and 'reading comprehension', while phonemic awareness stayed at the bottom end. Interestingly, relatively more children have performed better in oral reading compared to reading comprehension, which means that they were able to read the given text in the available time, but did not manage to comprehend the message given in that text. That said, the overall performance of the children in reading comprehension is better in SRP group as compared to non-SRP. Another interesting pattern that emerged from the study, across groups and cohorts, is that despite lowest performance on 'phonemic awareness', many of them exhibited better fluency in 'oral reading'. This finding does not corroborate with existing literature which considers phonemic awareness as one of the pre-requisites to enhance fluency in oral reading in English (Slavin et. al., 2009). Slavin et. al. (2009) explored efficacy of phonemic awareness instructional strategies by comparing two different reading interventions that were using phonics instruction to those that were not and found the outcomes to be disappointing. The authors concluded that the addition of phonics to traditional basal instruction is not

sufficient to bring about widespread improvement in children's reading. Reading interventions need to consider other factors including overall teaching methods in order to improve children's reading. In the case of SRP intervention, which is the basis for this study, the results are in favor. However, the difference in performance produced by the intervention is rather small. Thus the lesson for SRP is that this intervention needs to run for a longer duration along with rigorous professional development of reading-teachers with classroom support. The question that arises here is what is the optimum duration of professional development for reading-teachers? Some local interventions can serve here as prototypes of teachers' professional development. The Strengthening Teachers Education Programme (STEP) provides one such example, where the research shows that in order for teachers to reconceptualize and sustain their changed-practices, they need to go through the following stages: initiation, implementation and institutionalization. These stages at least require 4 years to make an impression on teachers' classroom practice (Bhutta, Parveen and Ali, 2016).

Distribution of readers on six domains of EGRA highlights that the SRP intervention succeeded in bringing down the ratio of 'zero readers' in the SRP group; however, only a fraction qualified to be fluent readers (level 4) in most of the EGRA domains. While it is important to 'monitor' the drop in zero-readers, it is also imperative to scrutinize the distribution of readers across four levels. The latter may help to make targeted amendments in various components of the intervention (i.e. teachers' training, follow-up, adequacy and quality of material, validity of assessments).

FACTORS INFLUENCING EGRA SCORES

Teachers' Classroom Practices

The observed practices indicated difference both in Generic Pedagogical Skills and Reading Instructions which were in favor of SRP. Results indicate that strategies that teachers have learnt during the training have permeated to an extent into their classroom practice. In terms of generic pedagogical skills the observed practices of SRP revealed better classroom practices in the use of material, questioning, teacher-child interaction, progressive discipline, active involvement of children and keeping children on task. In terms of reading instruction, SRP classrooms have shown better practices in the use of common text, phonological awareness activities, student reading to the teacher, student listening to the teacher, student reading to the whole class and on the exposition of vocabulary in the text. However, these findings should be interpreted with caution given that the majority of the classroom practices remained in the band of mediocre practice, suggesting that the practices demonstrated by SRP classrooms have a lot to improve if they are to improve students' learning outcomes to the desired level. Arguably, intervention programs can be effective when they are not considered as merely short-term interventions, but rather as programs that are taken over by host-country governments in order to sustain them (Graham & Kelly, 2018).

Home Literacy Environment

A small sample of parents was recruited to investigate the home literacy environment. These parents were recruited from the children who appeared in the EGRA test. The children of SRP schools experienced a more enriched home literacy environment than non-SRP children. Although the limited sample size did not yield significant differences, the SRP parents have an edge over non-SRP in most aspects. Interestingly, SRP parents reported that they listen to their child while they read aloud more frequently than non-SRP parents and this is the only difference which was found to be significant. Additionally, the home literacy environment was positively associated with non-word

fluency, expressive vocabulary and oral reading fluency of the children and these associations were also statistically significant. In other words, higher engagement of parents in children's home literacy activities significantly enhances their language performance in three important aspects of reading. These findings are supported by a widespread literature that shows that involvement and praise by parents for reading affect children's performance in reading in school (Cook-Cottone, 2004; Dodici & Pertson, 2003; Rashid, 2005). The SRP intervention has mainly focused on material and teacher training; parents' involvement was beyond the scope of the intervention. It would have yielded positive results had the intervention engaged parents more actively. The research shows that the most effective interventions are those where parents and school personnel work together to implement interventions utilizing a two-way exchange of information (Cox, 2005). Thus, if the SRP intervention has to move into the next phase, it would be prudent to involve parents through an amalgam of school-based literacy program and family literacy program (Pikulski, 1994; Morrow & Young 1997) to help children make significant progress with their learning (Westbrook et. al., 2009).

SRP INTERVENTION AND FACTORS INFLUENCING EGRA SCORES

Content, SRP Material and Teachers' Professional Development

All stakeholders agreed that the materials provided by the SRP were well-designed and useful. They are considered useful in enhancing students' reading skills and engagement. However, they also observed the limited availability of material as a barrier to effective implementation of reading intervention in the classroom. One of the major findings from qualitative data shows that teachers were not given enough time to reconceptualize how students learn to read through the material being provided by the SRP. They continue to believe and practice the *hijay* method, which they considered a tried and tested way of learning to read and which yielded 'good results' for them. The

literature emerging from developing countries suggests that for any innovation to take root, teachers have to be exposed for a longer duration which may continue for some years (Graham & Kelly, 2018). They also need classroom support through initial mentoring and then through institutionalized practices at school/district levels (Bhutta et. al., 2016). Unfortunately, teachers in this case did not receive enough support for longer duration that could make them reconceptualize to make permanent changes in their schemata. They continued to consider it an additional load contradictory to the existing curriculum rather than complementary to it. Furthermore, teachers also found the use of SRP material and phonemic awareness activities to be time consuming. Hence it is recommended that in order to sustain the practices over the longer run the intervention needs to be integrated into the mainstream curriculum along with provision of system-wide administrative support.

Physical Infrastructure

Research shows that the physical infrastructure of the classroom affects teaching and learning positively (Cuesta, Glewwe & Krause, 2016; Glewwe, Hanushek, Humpage & Ravina, 2011). The provision of physical infrastructure as one of the main inputs was beyond the scope of SRP intervention. Therefore, there was little to no classroom infrastructure development by the SRP in the intervention schools. However, some teachers and SMC-led initiatives were observed which enhanced the physical infrastructure (e.g. reading corners, furniture purchase, classroom repair, solar panel installation) at school. These encouraging examples suggest that SRP interventions are being complemented by other initiatives by the government (e.g. SMC funds). In order for the SRP to be sustained, it is recommended to undertake a scoping exercise in the intervention area to map the nature and patterns of existing intervention by other development partners. It would help to avoid duplication and enhance the complementarities between various development initiatives.

Parental Engagement

Quantitative data was collected to see parents' home practices that support reading, which shows positive results of parental engagement on children's reading though the sample was short and cannot be generalized. Teachers and HMs emphasized the engagement of parents' in children's learning. The informal discussion with parents indicated that they had very little understanding about the components of SRP intervention. The qualitative data also revealed that some of the SRP selected schools were also part of the Community Mobilization Program (CMP). However, the focus of CMP was more on enrolment enhancement, particularly of girls and community engagement and not solely on reading enhancement.

TLA VS. GT MODEL COMPARISON

The quantitative findings show that both the TLA and GT models contributed to producing positive results in terms of students' EGRA scores compared to non-SRP intervention schools for most of the domains. While the results of Sindhi EGRA followed similar trends Urdu-EGRA did not show significant differences. In fact, the pattern is reversed for oral reading in Urdu. Table 22 shows the difference between the two models in terms of various reading domains.

The qualitative data exhibited divergent views of various stakeholders in terms of the efficacy of the two models. The TLA model was considered better mostly by the SRP officials, because it provided them with an opportunity to engage competent private trainers and monitoring staff. On the other hand, the GT model was found better mostly by the government stakeholders as it developed the capacity of the government sector in various ways. The GT model engaged Guide Teachers from amongst the competent government teachers who provided cluster based support to the SRP reading teachers. Hence, in terms of sustainability, the GT model has better prospects than the TLA model. That said, in order to sustain any model, it needs to be implemented for a longer period of time with a rigorous monitoring system at various levels (Ali, 2006; Shields, Marsh & Adelman, 1998; Weiss,

Table 22: TLA vs GT Model – comparing children’s performance across groups

Domains	TLA			GT			Comments
	Overall	Sindhi	Urdu	Overall	Sindhi	Urdu	
Phonemic Awareness	SRP**	SRP**					TLA model favored SRP and the results followed the same trend for Sindhi
Non-Word Fluency	SRP**	SRP**			SRP**		TLA favored SRP for overall. Both models favored SRP for Sindhi.
Expressive Vocabulary				SRP**	SRP**		Only GT favored SRP for overall and Sindhi.
Oral Reading	SRP**	SRP**			SRP**	Non-SRP**	TLA favored SRP for overall EGRA scores. Both models favored SRP for Sindhi. However, the trend was reversed for GT in Urdu.
Reading Comprehension	SRP**	SRP**		SRP**	SRP**		Bothe models favored SRP for overall and Sindhi.
Listening Comprehension				SRP**	SRP**		Only GT favored SRP for both overall and Sindhi.

* p<0.05 **p<0.01

Montgomery, Ridgway & Bond, 1998). However, the SRP had to alter its models for pragmatic reasons (e.g. discontinuation of the Supervisor position, and non-availability of government teachers for subsequent cycles of intervention), which affected its original plans and sustainability prospects. A comparative overview of the two models is presented in Table 23.

Table 23: TLA vs GT model: quality, sustainability and impact

	TLA	GT
Quality	High	Low
Sustainability	Low	High
Impact	Low	Low

MOTIVATIONAL/STRUCTURAL BOTTLENECKS

The motivational and structural bottlenecks were identified by stakeholders regardless of the TLA or GT models. Lack of coordination amongst stakeholders was highlighted as one of the major factors that hindered effective implementation of the program. This lack of coordination led to various discrepancies such as: inappropriate nominations of teachers for SRP training; parallel training organized by SRP and government line departments; marking absence of trainee/guide teachers by bio-metric monitoring teams. In terms of training-related challenges, various issues were highlighted by the stakeholders which

included short duration of training, incompetency of trainers and inappropriate timings of training – which began in the middle of the academic year. Regarding the provision of material, the stakeholders felt that the SRP material was not aligned with the mainstream curriculum and was also inadequate in terms of meeting the needs of large classes. There was a consensus amongst the stakeholders about weak monitoring mechanism particularly by the taluka and district education officials (TEOs and DEOs), whose involvement would have enhanced the prospects of sustainability of the SRP interventions. It was quite often the case that teachers had to deal with multigrade classes in the schools, yet the training was found to be geared towards single grade teaching. This gap disadvantaged the multi-grade teachers who were not adequately prepared to follow the SRP intervention in their unique situation. One of the pre-requisites to initiate and sustain change is self-motivation of teachers (Stallings & Krasavage, 1986), which was found to be inadequate amongst the SRP-trained teachers, many of whom felt that they are only supposed to teach reading through SRP strategies the day they were observed. It was a recurring theme run through the data and often such teachers reverted to the traditional methods of teaching soon after the implementation cycle was completed.

ROLE OF STAKEHOLDERS IN IMPLEMENTING SRP INTERVENTION

As mentioned earlier, the motivation of teachers plays a critical role in implementing and sustaining the SRP intervention. In SRP schools only a few teachers were found to be self-motivated and continued the SRP methods even after the intervention. As far as the role of head teachers was concerned, they were not assigned any formal role in both TLA and GT models. Arguably, teachers must be given feedback and reassurance by the leadership (Tschannen-Moran et al., 1998). The head teachers in SRP intervention schools were found to be engaged in their regular administrative duties (e.g. ensuring students' and teachers' attendance, assisting in managing teaching loads and maintaining relationships with the parents) which may have contributed positively towards implementing the SRP intervention. The SMC and parents' engagement were not a formal part of SRP intervention; as a result the role of the SMC and parents in children's learning remained unplanned. Due to the availability of SMC funds, SMC members in various schools provided some infrastructural support but were not found to have played any formal role in improving the reading of children. Many of the parents were found to be illiterate and engaged very little in their children's reading at home, but whoever was engaged did have positive impact on their children's reading outcomes (refer to quantitative result on home literacy environment). The government stakeholders operated at least at two levels, at central level and at district/taluka levels. The central government officials were found to be more aware and engaged with the SRP intervention. They supported the reading intervention by nominating teachers for training and liaising with line departments (PITE, BoC, STEDA, STBB) with the SRP team. The line departments were not playing any lead role in the SRP intervention, their role remained at the level of participant at best. At the district and taluka levels, the DEOs and TEOs also remained at participant level and often remained disengaged with the core objectives of the intervention. The change literature emphasizes a more participatory approach towards stakeholders' involvement at various levels of implementation in order to enhance the prospects of sustainability (Fullan, 2006).

TECHNOLOGY-ENABLED ASSESSMENT

The use of technology in the SRP intervention was limited to only assessments through tablets. There was no use of technology by the teachers in enhancing children's learning outcomes. The TLAs and Guide Teachers used tablets for carrying out formative assessments. In addition, the formal assessments by the central project were carried out at baseline, midline and end line. The interviews with the guide teachers revealed that they faced some glitches with the technology, for example during trying out assessments, they accidentally submitted half-filled formative assessment forms. At other times, some tablets encountered technical faults which hindered the assessment process. Thus the technology use was only limited to the automation of assessment rather than any technology enhanced learning. It is recommended that this piecemeal approach would not take technology enhanced assessment far enough. If technology were to be successfully implemented it would need to be integrated into training, classroom teaching, students' learning and assessment.

PROSPECTS OF SUSTAINABILITY FOR THE GT MODEL

This section discusses prospects for sustainability of SRP model based on research findings of the study.

Alignment between SRP and STBB Material

In order to sustain any change, it is recommended that the innovation is integrated well into the existing system (i.e. start from where 'you' are). With this backdrop, it is imperative to align SRP interventions with the existing curriculum and textbooks. The teachers who were the main implementers felt SRP material to be 'additive' and not complementary. In order to align the new initiative of SRP and the existing curriculum, there are two possibilities: firstly, enough regular budget be approved at the SELD to mass-produce the SRP material along with appropriate training of a longer duration (3-5 years with each cohort) with a rigorous monitoring system; secondly, to enrich STBB by incorporating phonemic instruction into the textbooks of Sindhi and Urdu. The latter might work better as it will align with teachers, head teachers,

parents and government officials because the textbook is considered to be the lone resource for teaching.

Timings for Intervention

Timings for intervention have been criticized by various stakeholders including the teachers as the intervention started in the middle of the academic year and ran through the examination period, seriously disrupting school routines. In order to be accepted by the implementers the intervention should start at the beginning of the academic year. This will help teachers initiate new teaching techniques from the outset with fresh cohorts of children and have enough time to absorb, change and settle with the new intervention. In addition, trained teachers should be specified for grade 1 and 2 to maximize the benefits of training for students in the early grades. A student who enters grade 1 should remain with one particular teacher throughout the year and then move to grade 2 with another trained teacher for that level for another full year. These two years with consistent teachers will help the students develop a strong foundation in reading and will lead them to become proficient readers by grade 3.

District / School Level Coordination

There was relatively better coordination at the central level; however, the same did not trickle down to school level. In order to sustain and scale the intervention it has to have meticulous coordination amongst various stakeholder right from the top to the teachers. How can better coordination be achieved? Perhaps an initial step would be to take a bottom-up approach i.e. from taluka level to provincial. This will enhance ground level understanding and also bring teachers, head teachers, TEOs and DEOs on board. This involvement should not only be tokenistic, rather it has to be an activity in co-planning and 'coordinating'. It would help in selecting appropriate teachers, enriched monitoring by district officials and an overall emphasis on objective achievement that would have higher chances of sustainability after the project.

Duration and Follow-up Training

Considerably short duration of training was questioned by almost all the stakeholders. Literature also

establishes a positive link between the intensity and duration of professional development and the extent of teacher change (Shields, Marsh & Adelman 1998; Weiss, Montgomery, Ridgway & Bond 1998). The quantitative data revealed differences in students' performance which were in favor of SRP; however, qualitative views question the long-term sustainability of these differences. In order to have greater chances of sustainability, it is highly recommended that the duration of training be increased (at least 3-5 years with each cohort) along with strong classroom support. Additionally, teachers need appreciation and recognition of their contribution; hence an appraisal system that acknowledges the teachers' efforts need to be put in place. The assessment system needs to be aligned with the intervention strategies. Hence a system that aligns teaching with learning and assessment would have a higher probability of being sustained.

Parental Involvement

Along with teachers, the parents play an important role in the enhancement of students' learning outcomes. Some projects (e.g. RCC) have in fact used school level interventions with children to enhance even parents' literacy through their involvement. Literature also strongly supports the role of parents in enhancing early grade reading and literacy in general (Cox, 2005; Pikulski, 1994; Morrow & Young 1997; Westbrook et. al., 2009). In order to enhance the prospects for sustainability, it is recommended that a blend of a school-based literacy program and a family literacy program is implemented to help children make significant progress in their reading skills.

Sensitization of Senior Political and Bureaucratic Leadership

Importance of reading cannot be overestimated. Sustainability of the interventions suggested above cannot be achieved without support from the top. Senior political leadership and management of SELD will need to be sensitized on the importance of specialized intervention in reading. Specialized consultation sessions and seminars should be arranged for this purpose.



References

- Abadzi, H. (2013). Literacy for all in 100 days? A research based strategy for fast progress in low income countries.
- Abadzi, H. (2017). Turning a molehill into a mountain? How reading curricula are failing the poor worldwide. *Prospects*, 46(3-4), 319-334.
- Adams, M. J. (1990). *Beginning to Read: Thinking and Learning about Print*. Cambridge, UK: MIT press.
- Adams, S., Frampton, I., Gilmore, G., & Morris, M. (2010, August). A novel parent supported emotional literacy programme for children. *Community Practitioner*, 83(8), 27-30. Retrieved from <http://go.galegroup.com/ps/i.do>
- Allan, E.J. & Horn, N. (2013). *Evaluation of the USAID funded textbooks and learning materials programs (TLMP) in Ethiopia, Ghana, Malawi, Senegal, South Africa and Tanzania*. Washington, DC: USAID.
- Almalki, S. (2016). Integrating Quantitative and Qualitative Data in Mixed Methods Research-Challenges and Benefits. *Journal of Education and Learning*, 5(3), 288-296.
- Anderson, J. (1995). How parents' perceptions of literacy acquisition relate to their children's emergent literacy knowledge. *Reading Horizons*, 35, 209-228.
- Backman, J., Bruck, M., Hebert, M., & Seidenberg, M. (1984). Acquisition and use of spelling-sound correspondences in reading. *Journal of Experimental Child Psychology*, 38, 114-133.
- Baker, L., Fernandez-Fein, S., Scher, D., & Williams, H. (1998). Home experiences related to the development of word recognition. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 263-287). Mahwah, NJ: Erlbaum
- Balsiger, L. (2010). The Importance of Early Intervention for Reading Problems. *Bend Language and Learning*. Retrieved from <http://www.bendlanguageandlearning.com/Early%20Intervention.pdf>
- Brooks, G. (2002). *What works for children with literacy difficulties?: the effectiveness of intervention schemes*. London: Department for Children, Schools and Families.
- Bruner, J. S. (1978). Learning the mother tongue. *Human Nature*, 1, 43-49
- Bus, A., & Van Ijzendoorn, M. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, 91, 403-414.
- Catts, H.W., Fey, M.E., Zhang, X., & Tomblin, J.B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading*, 3, 331-361.
- Clay, M.M. (1993b). *Reading Recovery, a guidebook for teachers in training*. Portsmouth, NH: Heinemann.

- Clark-Chiarelli, N. & Louge, N. (2016). Teacher quality as a mediator of student achievement. In *Understanding what works in oral reading assessments*. Montreal: UNESCO, 30-40.
- Cohen, V. L. & Cowen, J. E. (2007). *Literacy for children in an information age: Teaching reading, writing, and thinking*. Belmont, CA: Wadsworth Cengage Learning.
- Cook-Cottone, C. (2004). Constructivism in family literacy practices: Parents as mentors. *Reading Improvement*, 41(4): 208-216.
- Couston-Theoharis, J.N., Giangreco, M., Doyle, M., & Vadasy, P. (2007). Paraprofessionals: The ‘sous-chefs’ of literacy instruction. *Teaching Exceptional Children*, 40, 56–62.
- Cox, D. D. (2005). Evidence-based interventions using home–school collaboration. *School Psychology Quarterly*, 20(4), 473–497.
- Cuesta, A. W., Glewe, P., & Krause, B. (2016). School Infrastructure and Educational Outcomes: A Literature Review, with Special Reference to Latin America. *Economía*, 17(1), 95-130. Retrieved from <http://www.jstor.org/stable/economia.17.1.95>
- Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, 33, 934-945.
- Desimone, L., Porter, A., Garet, M., Yoon, K., & Birman, B. (2002). Effects of Professional Development on Teachers’ Instruction: Results from a Three-Year Longitudinal Study. *Educational Evaluation and Policy Analysis*, 24(2), 81-112. Retrieved from <http://www.jstor.org/stable/3594138>
- Dodici, B. J., Draper, D. C., & Peterson, C. A. (2003). Early parent—child interactions and early literacy development. *Topics in Early Childhood Special Education*, 23(3), 124-136.
- Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel’s meta-analysis. *Review of Educational Research*, 71(3), 393-447
- Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH, DHHS. (2000). *Report of the National Reading Panel: Teaching Children to Read: Reports of the Subgroups* (00-4754). Washington, DC: U.S. Government Printing Office.
- Fehrler, S., Michaelowa, K., & Wechtler, A. (2009). The effectiveness of inputs in primary education: insights from recent student surveys for Sub-Saharan Africa. *The Journal of Development Studies*, 45(9), 1545-1578.
- Fried, L., Konza, D., Mulcahy, P. (2012) Paraprofessionals implementing a research-based reading intervention. *Australian Journal of Learning Difficulties*, 17(1): 35–45. doi: 10.1080/19404158.2012.674052.
- Foorman, B. R., Breier, J. I., & Fletcher, J. M. (2003). Interventions aimed at improving reading success: An evidence-based approach. *Developmental Neuropsychology*, 24(2–3), 613–63
- Fricke, S., Bowyer-Crane, C., Haley, A. J., Hulme, C., & Snowling, M. J. (2013). Efficacy of language intervention in the early years. *Journal of Child Psychology and Psychiatry*, 54(3), 280-290.
- Field. A. P. (2013). *Discovering statistics using SPSS* (4th Edition). London: Sage.
- Fullan, M. (2001). *Leading in a culture of change*. San Francisco: Jossey-Bass.

- Fullan, M. (2006). Change theory: A force for school improvement. *Center for Strategic Education*, 157, 3-14.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38, 915-945.
- Glewwe, P. W., Hanushek, E. A., Humpage, S. D., & Ravina, R. (2011). *School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010* (No. w17554). National Bureau of Economic Research.
- Goleman, D., Boyatzis, R., & McKee, A. (2002). *Primal leadership: Realizing the power of emotional intelligence*. Boston: Harvard Business School Publishing.
- Goswami, U. (2005) Synthetic Phonics and Learning to Read: A Cross-language Perspective. *Education Psychology in Practice*, 21(4), 273-282, DOI: 10.1080/02667360500344823
- Graham, J. & Kelly, S., (2018). *How effective are early grade reading interventions?: A review of the evidence*. World Bank Group. Retrieved from <http://documents.worldbank.org/curated/en/289341514995676575/pdf/WPS8292.pdf>
- Greenberg, J., & Baron, R. A. (2000). *Behavior in organizations (7th ed.)*. Upper Saddle River, NJ: Prentice Hall.
- Kanu, Y. (2005). Tensions and dilemmas of cross-cultural transfer of knowledge: post-structural/postcolonial reflections on an innovative teacher education in Pakistan. *International Journal of Educational Development*, 25(5), 493-513.
- Kennedy, M. M. (1999). The role of preservice teacher education. *Teaching as the learning profession: Handbook of policy and practice*, 54-85.
- Kolba, J. Kimberly (2005). Introduction to No Child Left Behind (NCLB) and Reading First (RF). In Block, C. C., & Israel, S. E. *Reading first and beyond: The complete guide for teachers and literacy coaches*. Corwin Press.
- Lieberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, 18, 201-212.
- Little, A., (2006). Multigrade lessons for EFA: a synthesis. In: Little, A. (Ed.), *Education for All and Multigrade Teaching: Challenges and Opportunities*. Springer, Dordrecht, The Netherlands, pp. 301-348.
- Lockheed, M. E., Vail, S. C., & Fuller, B. (1986). How textbooks affect achievement in developing countries: Evidence from Thailand. *Educational Evaluation and Policy Analysis*, 8(4), 379-392
- Mahar, N.E., & Richdale, A.L. (2008). Primary teachers' linguistic knowledge and perceptions of early literacy instruction. *Australian Journal of Learning Difficulties*, 13, 17-37.
- Malik, S., B. Balfour, J. P. Nzabonimpa, S. Cozzolino, G. Dib, and A. J. Dowd. (2015). *Endline evaluation of Rwandan children's book initiative*. Kigali: Save the Children.
- Manis, F. & Morrison, F. (1985). Reading disability: A deficit in rule learning? In L. Siegel & F. Morrison (Eds.), *Cognitive development in atypical children* (pp. 1-26). New York: Springer-Verlag.
- McIntosh, K., Horner, R.H., Chard, D.J., Boland, J.B., & Good, R.H. (2006). The use of reading and behavior screening measures to predict non-response to school-wide positive behavior support: A longitudinal analysis. *School Psychology Review*, 35, 275-291. [Google Scholar](#), [ISI](#)

- Meier, M., Reid, K., McKenzie, P. & Mellor, S. (2013), *Literacy and numeracy interventions in the early years of schooling : a literature review: report to the Ministerial Advisory Group on Literacy and Numeracy*.
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: an expanded source book*. London: Sage.
- Miller, S.D. (2003). Partners-in-reading: Using classroom assistants to provide tutorial assistance to struggling first-grade readers. *Journal of Education for Students Placed at Risk*, 8, 333–349.
- Mohammad, R., & Kumari, R. (2007). Effective use of textbooks: A neglected aspect of education in Pakistan. *Journal of Education for International Development*, 3(1), 1-12.
- Monique, S., & LeFevre, J. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development*, 73 (2): 445–460.
- Muter, V., Hulme, C., Snowling, M.J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology*, 40, 665–681.
- Pelletier, L.G., Fortier, M.S., Vallerand, R.J., & Briere, N.M. (2001). Associations among perceived autonomy support, forms of self-regulation and persistence: A prospective study. *Motivation and Emotion*, 25, 279–306. doi: 10.1023/a:1014805132406
- Pikulski, J. (1994). Preventing Reading Failure: A Review of Five Effective Programs. *The Reading Teacher*, 48(1), 30-39. Retrieved from <http://www.jstor.org/stable/20201362>
- Rashid FL, Morris RD, Servik RA.(2005). Relationship between home literacy environment and reading achievement in children with reading disability. *Journal of Learning Disability*, 48(2): 245-360.
- Rayner, Keith et al. (2002) How should reading be taught? *Scientific American*. Vol. 286, No. 3 (84-91)
- Reynolds, A. J. (1998). Developing early childhood programmes for children and families at risk: Research-based principles to promote long-term effectiveness. *Children and Youth Services Review*, 20, 503–523.
- Richardson, V. (1998). How teachers change: What will lead to change that most benefits student learning. *Focus on basics*, 2(4), 7-11.
- Richardson, V., & Placier, P. (2001). Teacher change. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed.) (pp. 905-947). Washington, DC: American Educational Research Association.
- Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale survey research tells us about teacher effects on student achievement: Insights from the Prospects Study of Elementary Schools. *Teachers College Record*, 104, 1525–1567.
- Ruth, F. B. (2010). Australian pre-service teachers' knowledge of phonemic awareness and phonics in the process of learning to read. *Australian Journal of Learning Difficulties*, 15:1, 99-110, DOI: 10.1080/19404150903524606
- Tibi, S., & John, R. (2018). Investigating Phonological Awareness and Naming Speed as Predictors of Reading in Arabic. *Scientific Studies of Reading*, 22:1, 70-84, DOI:10.1080/10888438.2017.1340948
- Savin, H. B. (1972). What the child knows about speech when he starts to learn to read. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye* (pp. 319–326). Cambridge, MA: MIT Press.

- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143–174.
- Schweisfurth, M. (2011) Learner-centred education in developing country contexts: From solution to problem? *International Journal of Educational Development*, 31: 425–432.
- Shanahan, T., & Barr, R. (1995). Reading Recovery: An Independent Evaluation of the Effects of an Early Instructional Intervention for At-Risk Learners. *Reading Research Quarterly*, 30(4), 958-996. doi:10.2307/748206
- Shields, P. M., Marsh, J. A., & Adelman, N. E. (1998). *Evaluation of NSF's Statewide Systemic Initiatives (SSI) Program: The SSIs' impacts on classroom practice*. Menlo Park, CA: SRI
- Slavin, R. E., Karweit, N. L. & Wasik, B. A.(Eds.), (2001) *Preventing early school failure: Research on effective strategies* (pp. 102–122).
- Slavin, R.E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009).-evidence synthesis. *Review of Educational Research*, 79(4), 1391–1466. doi:[10.3102/0034654309341374](https://doi.org/10.3102/0034654309341374).
- Stallings, J., & Krasavage, E. (1986). Program implementation and student achievement in a four-year Madeline Hunter Follow-Through project. *Elementary School Journal*, 87(2), 117-138.
- Sylva, K., Hurry, J., Mirelman, H., Burrell, A., & Riley, J. (1999). Evaluation of a focused literacy teaching programme in reception and year 1 classes: Classroom observations. *British Educational Research Journal*, 25(5), 617-635.
- Tarnoff, C. (2005). Foreign aid: An introductory overview of US programs and policy. Library of Congress Washington DC Congressional Research Service.
- Teale, W. H., & Sulzby E. (1986). Emergent literacy as a perspective for examining how young children become readers and writers. In *Emergent Literacy: Reading and Writing*, (pp.7–25). Norwood, NJ: Ablex Publishing.
- Troia, G. (1999). Phonological awareness intervention research: A critical review of the experimental methodology. *Reading Research Quarterly*, 34, 28–52.
- Troia, G. A. (2004). Phonological processing and its influence on literacy learning. In C. Stone, E. Silliman, B. Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders* (pp. 271–301). New York, NY: Guilford Press.
- Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, 68(2), 202-248.
- UNESCO, (2004). *Education for All: The Quality Imperative*. EFA Global Monitoring Report. UNESCO, Paris.
- UNESCO. (2014). *Textbooks and learning resources: A global framework for policy development*. Education sector.
- UNESCO. (2016). *Every child should have a textbook*. Policy paper 23. Global education monitoring report. Retrieved from <http://unesdoc.unesco.org/images/0024/002433/243321E.pdf>
- USAID. (2016). *Early grade reading in Latin America and the Caribbean: A systematic review*. LAC READS capacity program.

- Verspoor, A. (1989). *Pathways to Change: Improving the Quality of Education in Developing Countries*. World Bank Discussion. 53, Washington, D.C. World Bank.
- Villegas, E., & Reimers, F. (2000). *The professional development of teachers as lifelong learning: models, practices, and factors that influence it*. The Board on International Comparative Studies in Education of the National Research Council: Washington D.C.
- Warwick, D.P., Reimers, F., McGinn, N., 1992. The implementation of educational innovations: lessons from Pakistan. *International Journal of Educational Development*, 12 (4), 297–307.
- Weiss, I. R., Montgomery, D. L., Ridgway, C. J., & Bond, S. L. (1998). *Local systemic change through teacher enhancement: Year three cross-site report*. Chapel Hill, NC: Horizon Research, Inc.
- Westbrook, J., Shah, N., Durrani, N., Tikly, C., Khan, W., Dunne, M., (2009). Becoming a teacher: transitions from training to the classroom in the NWFP, Pakistan. *International Journal of Educational Development*, 29, 437–444.
- Whitehurst, G. J., & C. J. Lonigan. (1998). Child development and emergent literacy. *Child Development*, 69: 848–872.
- Zimmerman, J. (2006), *Why some teachers resist change and what principals can do about it*. NASSP Bulletin, Vol. 90, No. 3, pp. 238-249. doi:10.1177/0192636506291521



Appendix: EGRA - Detailed Results of Statistical Analysis - Cohort-1

EGRA scores: Comparison of groups (Non-SRP vs SRP) on pre-test

Domain		Bootstrap Confidence (Interval BCa 95% CI)		t-equation	Effect Size (r)	Percentage Difference (%)	Trends
Overall	-2.95	BCa 95% CI	[-6.21, -0.18]	t (563)= -1.746, ns	ns	-	-
Phonemic Awareness	0.19	BCa 95% CI	[-3.15, -3.22]	t (563)= 0.119, ns	ns	-	-
Non-word Fluency	-6.42	BCa 95% CI	[-10.90, -1.56]	t (547)= -2.90, p<0.01	r=-0.12	6%	in favor of SRP
Expressive Vocabulary	-4.75	BCa 95% CI	[-9.09, -0.20]	t (563)= -2.30, p<0.05	r=-0.10	5%	in favor of SRP
Oral Reading Fluency	-5.33	BCa 95% CI	[-12.46, 1.47]	t (563)= -1.62, ns	ns	-	-
Reading Comprehension	-1.31	BCa 95% CI	[-6.21, 3.51]	t (563)= -0.55, ns	ns	-	-
Listening Comprehension	-0.09	BCa 95% CI	[-5.54, 5.38]	t (563)= -0.036, ns	ns	-	-

EGRA scores: Comparison of groups (Non-SRP vs SRP) on post-test

Overall	-6.39	BCa 95% CI	[-10.14, 2.53]	t (546)= -3.45, p<0.01	r=-0.15	6%	in favor of SRP
Phonemic Awareness	-5.90	BCa 95% CI	[-8.95, -2.83]	t (544)= -3.521, p<0.01	r=-0.18	6%	in favor of SRP
Non-word Fluency	-8.00	BCa 95% CI	[-12.73, -3.18]	t (546)= -3.010, p<0.01	r=-0.14	8%	in favor of SRP
Expressive Vocabulary	-3.03	BCa 95% CI	[-8.00, 2.41]	t (546)= 1.26, ns	Ns	-	-
Oral Reading Fluency	-9.94	BCa 95% CI	[-16.29, -3.78]	t (546)= -2.90, p<0.01	r=-0.16	10%	in favor of SRP
Reading Comprehension	10.14	BCa 95% CI	[-15.29, -4.90]	t (528)= -0.48, p<0.01	r=-0.20	10%	in favor of SRP
Listening Comprehension	-1.36	BCa 95% CI	[-6.95, 4.23]	t (546)= -0.48, ns	Ns	-	-

EGRA scores: Progression over time in Non-SRP group

Overall	3.35	BCa 95% CI	[-6.96, -0.020]	t (462)= 1.763, ns	ns	-	-
Phonemic Awareness	2.20	BCa 95% CI	[-1.12, 5.30]	t (462)= 1.294, ns	ns	-	-
Non-word Fluency	-6.58	BCa 95% CI	[-11.28, -2.23]	t (462)= -2.648, p<0.01	r=-0.12	7%	Gain overtime
Expressive Vocabulary	-4.27	BCa 95% CI	[-9.19, 0.036]	t (457)= -1.680, ns	ns	-	-
Oral Reading Fluency	-9.30	BCa 95% CI	[-16.196, -2.23]	t (462)= -2.58, p<0.01	r=-0.12	9%	Gain overtime
Reading Comprehension	-2.59	BCa 95% CI	[-7.65, 2.58]	t (462)= -0.970, ns	ns	-	-
Listening Comprehension	0.43	BCa 95% CI	[-4.64, 5.47]	t (462)= 0.145, ns	ns	-	-

EGRA scores: Progression over time in SRP group

Overall	-6.80	BCa 95% CI	[-10.22, -3.35]	t (638)= -4.177, p<0.001	r=-0.16	7%	Gain overtime
Phonemic Awareness	-3.90	BCa 95% CI	[-7.09, -0.288]	t (621)= -2.417, p<0.05	r=-0.10	4%	Gain overtime
Non-word Fluency	-8.16	BCa 95% CI	[-12.84, -3.42]	t (633)= -3.430, p<0.01	r=-0.14	8%	Gain overtime
Expressive Vocabulary	-2.55	BCa 95% CI	[-6.17, 1.17]	t (618)= -1.304, ns	Ns	-	-
Oral Reading Fluency	13.92	BCa 95% CI	[-20.66, -7.34]	t (647)= -4.525, p<0.01	r=-0.18	14%	Gain overtime
Reading Comprehension	-11.42	BCa 95% CI	[-16.47, -6.12]	t (612)= -4.66, p<0.001	r=-0.19	11%	Gain overtime
Listening Comprehension	-0.84	BCa 95% CI	[-5.60, 3.85]	t (647)= 0.145, ns	Ns	-	-

Analysis by Language - Cohort-1 SINDHI

EGRA-Sindhi scores: Comparison of groups on pre-test

Domain		Bootstrap Confidence (Interval BCa 95% CI)		t-equation	Effect Size (r)	Percentage Difference (%)	Trends
Overall	-5.64	BCa 95% CI	[-9.47, -1.62]	t(406)=-2.92, p<0.01	r=-0.14	6%	in favor of SRP
Phonemic Awareness	-4.90	BCa 95% CI	[-4.43, 3.94]	t(406)=-0.25, ns	ns	-	-
Non-word Fluency	-5.50	BCa 95% CI	[-10.69, -0.40]	t(405)=-2.13, p<0.05	r=-0.11	6%	in favor of SRP
Expressive Vocabulary	-10.66	BCa 95% CI	[-15.40, -5.56]	t(369)=-4.318, p<0.001	r=-0.22	11%	in favor of SRP
Oral Reading Fluency	-9.43	BCa 95% CI	[-17.45, -1.88]	t(406)=-2.506, p<0.05	r=-0.12	9%	in favor of SRP
Reading Comprehension	-4.22	BCa 95% CI	[-9.32, 1.39]	t(403)=-1.596, ns	ns	-	-
Listening Comprehension	-3.55	BCa 95% CI	[-9.51, 3.02]	t(406)=-1.17, ns	ns	-	-

EGRA-Sindhi scores: Comparison of groups on post-test

Overall	-8.50	BCa 95% CI	[-13.27, -3.80]	t(364)=-3.65, p<0.001	r=-0.19	9%	in favor of SRP
Phonemic Awareness	-7.08	BCa 95% CI	[-11.44, -2.61]	t(354)=-3.35, p<0.001	r=-0.18	7%	in favor of SRP
Non-word Fluency	-8.83	BCa 95% CI	[-15.44, -2.72]	t(364)=-2.749, p<0.01	r=-0.14	9%	in favor of SRP
Expressive Vocabulary	-3.22	BCa 95% CI	[-8.93, 2.98]	t(364)=-1.072, ns	Ns	-	-
Oral Reading Fluency	-12.64	BCa 95% CI	[-21.08, -4.35]	t(364)=-3.042, p<0.01	r=-0.16	13%	in favor of SRP
Reading Comprehension	-12.98	BCa 95% CI	[-19.82, -6.41]	t(364)=-3.811, p<0.001	r=-0.20	13%	in favor of SRP
Listening Comprehension	-6.24	BCa 95% CI	[-12.28, 0.17]	t(364)=-1.825, ns	Ns	-	-

EGRA-Sindhi scores: Progression over time in Non-SRP group

Overall	-4.47	BCa 95% CI	[-8.65, -0.19]	t(341)=-2.04, p<0.05	r=-0.12	4%	Gain
Phonemic Awareness	1.63	BCa 95% CI	[-1.97, 5.24]	t(358)=0.846, ns	Ns	-	-
Non-word Fluency	-6.23	BCa 95% CI	[-12.04, -0.37]	t(338)=-2.188, p<0.05	r=-0.12	6%	Gain
Expressive Vocabulary	-5.598	BCa 95% CI	[-10.88, -0.01]	t(358)=-1.913, ns	Ns	-	-
Oral Reading Fluency	-12.84	BCa 95% CI	[-20.54, -4.85]	t(358)=-3.12, p<0.01	r=-0.16	13%	Gain
Reading Comprehension	-5.87	BCa 95% CI	[-11.56, -0.12]	t(337)=-1.968, p<0.05	r=-0.11	6%	Gain
Listening Comprehension	2.08	BCa 95% CI]	[-4.25, 8.82]	t(358)=0.617, ns	Ns	-	-

EGRA-Sindhi scores: Progression over time in SRP group

Overall	-7.32	BCa 95% CI	[-11.03, -3.41]	t(385)=-3.53, p<0.001	r=-0.18	7%	Gain Overtime
Phonemic Awareness	-4.96	BCa 95% CI	[-8.97, -1.14]	t(382)=-2.34, p<0.05	r=-0.12	5%	Gain Overtime
Non-word Fluency	-9.54	BCa 95% CI	[-15.17, -3.61]	t(381)=-3.21, p<0.001	r=-0.16	10%	Gain Overtime
Expressive Vocabulary	1.85	BCa 95% CI	[-3.28, 6.66]	t(367)=0.72, ns	Ns	-	-
Oral Reading Fluency	-16.04	BCa 95% CI	[-23.65, -8.81]	t(412)=-4.23, p<0.001	r=-0.20	16%	Gain Overtime
Reading Comprehension	-14.63	BCa 95% CI	[-20.73, -9.11]	t(366)=-4.70, p<0.001	r=-0.24	15%	Gain Overtime
Listening Comprehension	-0.61	BCa 95% CI	[-6.92, 5.31]	t(398)=-0.20, ns	Ns	-	-

Analysis by Language: EGRA-Urdu

EGRA-Urdu scores: Comparison of groups on pre-test

Domain		Bootstrap Confidence (Interval BCa 95% CI)		t-equation	Effect Size (r)	Percentage Difference (%)	Trends
Overall	4.56	BCa 95% CI	[-1.87, -11.18]	t(115)=1.28, ns	Ns	-	-
Phonemic Awareness	1.79	BCa 95% CI	[-4.44, 8.47]	t(115)=0.57, ns	Ns	-	-
Non-word Fluency	-8.16	BCa 95% CI	[-16.62, 0.94]	t(122)=-1.92, ns	Ns	-	-
Expressive Vocabulary	10.84	BCa 95% CI	[2.80, 8.95]	t(155)=2.95, p<0.01	r=0.23	11%	in favor of non-SRP
Oral Reading Fluency	5.93	BCa 95% CI	[-6.44, 18.15]	t(155)=0.86, ns	Ns	-	-
Reading Comprehension	9.11	BCa 95% CI	[-0.65, 18.93]	t(75)=1.64, ns	Ns	-	-
Listening Comprehension	7.85	BCa 95% CI	[-3.67, 18.30]	t(155)=1.47, ns	Ns	-	-

EGRA-Urdu scores: Comparison of groups on post-test

Overall	3.26	BCa 95% CI	[-8.99, -2.89]	t(180)=-1.05, ns	Ns	-	-
Phonemic Awareness	-4.76	BCa 95% CI	[-9.98, 0.84]	t(134)=-1.65, ns	Ns	-	-
Non-word Fluency	-6.13	BCa 95% CI	[-15.34, 4.05]	t(180)=-1.24, ns	Ns	-	-
Expressive Vocabulary	-2.01	BCa 95% CI	[-10.56, 7.41]	t(180)=-0.49, ns	ns	-	-
Oral Reading Fluency	-8.72	BCa 95% CI	[-20.49, 3.02]	t(180)=-1.42, ns	Ns	-	-
Reading Comprehension	-7.25	BCa 95% CI	[-15.64, 2.03]	t(180)=-1.48, ns	Ns	-	-
Listening Comprehension	9.32	BCa 95% CI	[-0.41, 18.31]	t(180)=1.90, ns	Ns	-	-

EGRA-Urdu scores: Progression over time in Non-SRP group

Overall	1.09	BCa 95% CI	[-6.47, 8.79]	t(102)=0.28, ns	Ns	-	-
Phonemic Awareness	4.03	BCa 95% CI	[-4.11, 11.62]	t(86)=1.07, ns	Ns	-	-
Non-word Fluency	-7.60	BCa 95% CI	[-17.81, 2.83]	t(100)=-1.49, ns	Ns	-	-
Expressive Vocabulary	1.64	BCa 95% CI	[-7.50, 11.44]	t(102)=0.33, ns	Ns	-	-
Oral Reading Fluency	2.86	BCa 95% CI	[-11.16, 17.81]	t(102)=0.38, ns	Ns	-	-
Reading Comprehension	9.61	BCa 95% CI	[-0.98, 20.82]	t(102)=1.62, ns	Ns	-	-
Listening Comprehension	-3.99	BCa 95% CI	[-16.83, 10.04]	t(102)=-0.64, ns	Ns	-	-

EGRA-Urdu scores: Progression over time in SRP group

Overall	-6.73	BCa 95% CI	[-11.54, -2.41]	t(233)=-2.58, p<0.01	r=-0.17	7%	Gain Overtime
Phonemic Awareness	-2.51	BCa 95% CI	[-7.23, 2.19]	t(233)=-1.02, ns	Ns	-	-
Non-word Fluency	-5.57	BCa 95% CI	[-12.92, 2.27]	t(233)=-1.38, ns	Ns	-	-
Expressive Vocabulary	-11.21	BCa 95% CI	[-16.81, -6.16]	t(233)=-3.79, p<0.001	r=-0.24	11%	Gain Overtime
Oral Reading Fluency	-11.79	BCa 95% CI	[-21.66, -2.04]	t(233)=-2.28, p<0.05	r=-0.15	12%	Gain Overtime
Reading Comprehension	-6.74	BCa 95% CI	[14.84, 0.33]	t(233)=-1.70, ns	Ns	-	-
Listening Comprehension	-2.54	BCa 95% CI	[-10.40, 5.27]	t(233)=-0.64, ns	Ns	-	-

Analysis by Language: Cohort-2

EGRA scores: comparing performance on reading domains

Domain		Bootstrap Confidence (Interval BCa 95% CI)	t-equation	Effect Size (r)	Percentage Difference (%)	Trends
Overall	-4.82	BCa 95% CI [-8.43, -1.75]	t(607)= -3.01, p<0.01	r=-0.12	5%	in favor of SRP
Phonemic Awareness	-3.15	BCa 95% CI [-6.39, 0.033]	t(603)= -1.85, ns	ns	-	-
Non-word Fluency	-3.55	BCa 95% CI [-8.63, 1.33]	t(607)= -1.41, ns	ns	-	-
Expressive Vocabulary	-7.95	BCa 95% CI [-12.10, -3.76]	t(607)= -3.98, p<0.001	r=-0.16	8%	in favor of SRP
Oral Reading Fluency	-3.50	BCa 95% CI [-9.77, 2.44]	t(607)= -1.13, ns	ns	-	-
Reading Comprehension	-4.69	BCa 95% CI [-9.44, -0.29]	t(607)= -2.07, p<0.05	r=-0.08	5%	in favor of SRP
Listening Comprehension	-6.10	BCa 95% CI [-11.72, -0.94]	t(605)= -2.40, p<0.05	r=-0.10	6%	in favor of SRP

Analysis by Language: EGRA-Sindhi

EGRA-Sindhi scores: comparing performance on reading domains

Overall	-9.29	BCa 95% CI [-13.01, -5.32]	t(410)= -4.82, p<0.001	r=-0.23	9%	in favor of SRP
Phonemic Awareness	-2.31	BCa 95% CI [-5.04, 0.41]	t(411)= -1.85, ns	ns	-	-
Non-word Fluency	-8.92	BCa 95% CI [-14.20, -3.38]	t(409)= -3.18, p<0.01	r=-0.16	9%	in favor of SRP
Expressive Vocabulary	-10.95	BCa 95% CI [-15.75, 5.71]	t(413)= -4.52, p=0.001	r=-0.22	11%	in favor of SRP
Oral Reading Fluency	-11.93	BCa 95% CI [-18.07, -4.37]	t(413)= -3.23, p<0.01	r=-0.16	12%	in favor of SRP
Reading Comprehension	-9.41	BCa 95% CI [-14.61, -3.86]	t(411)= -3.30, p<0.01	r=-0.16	9%	in favor of SRP
Listening Comprehension	-12.21	BCa 95% CI [-18.27, -5.92]	t(413)= -3.97, p=0.001	r=-0.19	12%	in favor of SRP

Analysis by Language: EGRA-Urdu

EGRA-Sindhi scores: comparing performance on reading domains

Overall	4.75	BCa 95% CI [-0.71, 11.27]	t(192)= 1.73, ns	ns	-	
Phonemic Awareness	-4.82	BCa 95% CI [-12.91, 4.29]	t(192)= -1.15, ns	ns	-	
Non-word Fluency	8.06	BCa 95% CI [-1.79, 18.07]	t(192)= 1.74, ns	ns	-	
Expressive Vocabulary	-1.59	BCa 95% CI [-8.61, 5.48]	t(192)= -0.45, ns	ns	-	
Oral Reading Fluency	14.53	BCa 95% CI [4.14, 25.64]	t(192)= 2.62, p<0.01	r=0.19	15%	
Reading Comprehension	5.34	BCa 95% CI [-1.72, 12.91]	t(167)= 1.53, ns	ns	-	
Listening Comprehension	6.99	BCa 95% CI [-2.37, 16.94]	t(192)= 1.62, ns	ns	-	

