Endline Report

Reading Diagnostic Assessment In Bundi, Jhalawar and Kota Districts of Rajasthan

Educational Initiatives Private Limited
27/07/17
Contents

1 Executive summary .................................................................................................................................. 4

2 Design of the study .................................................................................................................................. 7
  2.1 Intervention model .............................................................................................................................. 7
  2.2 Test Design and Blueprint .................................................................................................................. 7
    2.2.1 Test Design .................................................................................................................................. 8
    2.2.2 Selection of 47 common questions ................................................................................................. 8
    2.2.3 Scoring and logic to categorize students ..................................................................................... 9
    2.2.4 Cut off range for each proficiency level ....................................................................................... 9
    2.2.5 Selection of rest of questions (adaptive component of the test) .................................................. 9
    2.2.6 Approach to item selection/development from baseline ............................................................. 10
  2.3 Endline Test Conduction: ................................................................................................................ 11

3 Quantitative Data Analysis .................................................................................................................... 12
  3.1 Performance of students on different levels of proficiency .............................................................. 12
  3.2 Comparative performance of students in different age groups ....................................................... 17
    3.2.1 Allotment of language bucket in the fine test ............................................................................ 18
    3.2.2 Performance percentage of different age groups in all sub-skills .............................................. 18
  3.3 Comparative performance of boys and girls .................................................................................... 23
  3.4 Gain in student performance from Baseline to Endline ................................................................. 25
    3.4.1 Sub-Skill wise Average gain analysis ......................................................................................... 25
    3.4.2 Question wise gain ..................................................................................................................... 29
    3.4.3 Student wise gain analysis ......................................................................................................... 32
    3.4.4 Grade wise gain analysis ........................................................................................................... 33

4 Skill wise analysis .................................................................................................................................. 34
  4.1 Skill 1 – Letter Knowledge ................................................................................................................. 34
    4.1.1 Sub-skill 1 – Identifying Vowels ................................................................................................. 34
    4.1.2 Sub-skill 2 – Identifying Consonants .......................................................................................... 36
    4.1.3 Subs-skill 3 – Identifying Matras ............................................................................................... 40
    4.1.4 Sub-skill 4 – Identifying Syllable Blends .................................................................................. 42
  4.2 Vocabulary ......................................................................................................................................... 44
    4.2.1 Sub-skill 5 – Basic Vocabulary ................................................................................................. 44
    4.2.2 Sub-skill 6 – Word Meaning ....................................................................................................... 45
4.2.3 Sub-Skill 7 – Usage ......................................................................................................................... 48
4.2.4 Sub-skill 8 – Recall ....................................................................................................................... 50
4.2.5 Sub-skill 9 – Sentence Structure Knowledge ................................................................................. 52
4.2.6 Sub-skill 10 – Punctuation Knowledge ......................................................................................... 54
4.2.7 Sub-skill 11 – Grammar Knowledge ............................................................................................. 56
4.3 Skill – Reading Comprehension ....................................................................................................... 58
4.3.1 Sub-skill 12 – Passage Comprehension ........................................................................................ 59
1 EXECUTIVE SUMMARY

Reading Diagnostic Assessment is a teacher-supporting but non-teacher-dependent, personalized, engaging and scalable computer-based audio-visual solution for developing language reading skills among students. The complexities of the process of student learning are such that one has to rely on computer technology if one really wants every student to learn with understanding. This ensures consistency in providing high quality instruction and learning tailored to the needs of each student.

The study was rolled out to create a testing tool (tech-based) for early grade reading skills, a tool which is able to measure language learning level of the students and do fundamental research on reading & building remediation tools that will act as the building blocks for large scale impact.

Initially, as first version of diagnostic test, a common test was designed based on 4 core skills and 12 sub skills. The test had 230 questions spread across classes, skills and sub-skills. During the baseline, this version was tested on over 400 students in 18 schools in three districts (Kota, Bundi and Jhalawar) of Rajasthan. The results were analysed and feedback was taken from an advisory board comprising eminent experts at national level. During the review, it was observed that

- students who were good at the language was attempting maximum questions and also it was taking more than 2.5 hours for the student to finish the test, which was a lot.
- though it is good to make an ability test, which doesn’t care about the questions’ grade level but then as an significant stakeholder, a teacher, would rather want to know what and how much, the students of her class are able to do of that grade level content and not higher than that.

Based on these inputs, the test was truncated scientifically based on the following:

a. The first part is a crude test being mandatory for all to do (restricted to 47 questions) irrespective of their grades. This part determines the overall language of the child and then further testing is based on that level, assigned to the child.

b. The second part, fine test, which gets into depths of diagnosing the issues, each student is facing at the assigned level and reasons of it. The file level of the test was of three type – Beginner, Intermediate and proficient.

c. Questions with performance less than 20% and more than 80% were not selected for the endline assessment. Also the question which had PBC value of less than 0.1 were dropped.

d. The test was attempted to be made for duration of 60 minutes.

e. Questions with 4 options preferred over questions with lesser options.

f. Questions with single correct option preferred. Only single attempt was allowed to answer a question.

g. Over all difficulty level of the test (item set) was kept around 50% respectively for the 3 levels (Students of grade 2 and 3 should have 50% performance on “Beginner” test set, Students of grade 3 and 4 should have 50% performance on “Intermediate” test set, Students of grade 4 and 5 should have 50% performance on “Proficient” test set)

Every student was expected to give crude level test comprising 47 questions. If student performance is up to 30% then the student will appear for Beginner level, if the performance is 31 to 55 then the student will appear for intermediate level and if the performance is above 55 then the student will appear for proficient level of test.
**A Beginner level** student is at the basic level of language learning, the child has automatized all alphabets but few, can read and understand basic sight words with one or max. two simple matras and is able to understand basic questions and answer them correctly based on a contextual sentence read. This level had 48 questions.

**An intermediate level** student has automatized all alphabets, with correct sound and symbol relation, can spell the a given word correctly and use it in the proper context, can reconstruct a sentence in the meaningful manner and read a short para to answer factual retrieval questions based on it. This level had 40 questions.

**A proficient level** student has good hold on syntactic and semantic knowledge to use words and sentences in the right context with proper spellings and grammar, the child can read a story and make out inferential conclusions. This level had 55 questions.

This was test was administered to more than 370 students in the same schools where baseline study was conducted. The endline test was conducted from 13\textsuperscript{th} February to 28\textsuperscript{th} February. The test was administered students from 2\textsuperscript{nd}, 3\textsuperscript{rd}, 4\textsuperscript{th} and 5\textsuperscript{th} grades; the students were kept same from baseline to endline. Due to absenteeism students did not come to school and testing was repeated to cover such students. Over 370 students were tested in endline study.

The study shows the following findings:

- More than 60\% Students in grade 2 and 3 were found to be in intermediate level, while in grade 4 and 5 numbers of students in proficient level is more.

- The performance in grade 3 questions on identify consonants, basic vocabulary, recall and reading comprehension were specifically found to be low in crude level of students which were allotted beginner language level. Also, Students in beginner language level have found content of grade 4 difficult, across sub-skills.

- The performance of student on reading comprehension sub-skill is exceptionally low across Beginner and Intermediate language levels. This improves for Proficient level of students though.

- According to the age distribution, maximum number of students, across age brackets, have performed between 30 to 55\% on the crude test and hence were allotted Intermediated language bucket in the finer test.

- **Overall % performance of boys and girls on crude test** - Girls have performed 4.5\% better than boys on the endline diagnostic test. A total of 162 boys and 167 girls participated in the study.
Grade wise % performance of boys and girls on crude test - Girls are performing better than boys across grades with 3.64% in grade 2, 2.09% in grade 3, 2.84% in grade 4 and 6.28% in grade 5. Test was not significant for all the classes for 95% confidence level.

The performance has increased from baseline to endline for 64.5% students. 35.4% students have underperformed on the endline test. The overall gain comparison is based on the data of 370 students and 45 common test items in baseline & endline test versions. Performance gain of 32.1% students is in the range of 0 to 10% and the performance gain of 21.6% students is in the range of 10% to 20%. About 10.8% students have shown performance gain above 20% from baseline to endline.

Grade wise gain analysis - 70% of grade 5 students showed performance gain on endline diagnostic test. With 67% and 66% of grade 2 and 4 students, respectively, showing overall similar performance gain, the least performance gain is observed in grade 3 with 60% students performing better on endline test.
2 Design of the Study

Reading Diagnostic Assessment is a teacher-supporting but non-teacher-dependent, personalized, engaging and scalable computer-based audio-visual diagnosis for understanding the language reading skills among students. The complexities of the process of reading assessment is such that if one really wants to identify the exact learning level of a child and help the child grow better with reading skills, one has to rely on computer technology. Objective of the study

- To create a comprehensive adaptive testing tool (tech-based) for early grade reading skills, which clearly defines the possible pitfalls and indicates accurately the skill deficit (if any) to achieve optimum reading fluency with comprehension
- To provide a tool which adapts to the pace & learning level of the students and accommodates for different learning styles. This will enable customized instruction and practice time optimized based on the gaps diagnosed
- To do fundamental research on reading & building remediation tools that will act as the building blocks for large scale impact
- To conduct a baseline and endline assessments with a light touch intervention with teachers through capacity building workshop on finding from baseline assessments.

2.1 Intervention Model

The assessment tool was developed to measure the student learning level in Hindi language. The assessments were carried out twice, first at beginning and second end of the academic year. In between the two rounds of assessment, a capacity building workshop was provided to teachers to enable them to use the findings from baseline in their day-to-day teaching. This intervention was followed-up by reminding them to use insights via whatsapp messages.

2.2 Test Design and Blueprint

Initially, as first version of diagnostic test, a common test was designed based on 4 core skills and 12 sub skills. The test had sufficient number of questions at the grade, skill and sub-skill level. The analysis of the results was on the basis of the ability to answer the test at grade as well higher grade level and to the accuracy of answering the question.

Following this model, it was realised that the student who is good at the language was attempting maximum questions and also it was taking more than 2.5 hours for the student to finish the test, which was a lot. Another point highlighted by one of the advisory board member was, though it is good to make an ability test, which doesn’t care about the questions’ grade level but then as an significant stakeholder, a teacher, would rather want to know what and how much, the students of her class are able to do of that grade level content and not higher than that. Based on these inputs, the test was truncated scientifically, by splitting the test into two parts. The first part is a crude test being mandatory for all to do (restricted
to 47 questions) irrespective of their grades. This part determines the overall language of the child and then further testing is based on that level, assigned to the child. This is followed by the second part, fine test, which gets into depths of diagnosing the issues, each student is facing at the assigned level and reasons of it. Following section defines the test design of the version 2 diagnostic test, which was used for end line testing.

2.2.1 Test Design

- Test has 90 to 100 items (decided based on maximum test time of 60 minutes and 0.7 min average time to answer a question)
- Test does not include question on speaking or writing skills.
- Pool of 47 common questions which will be given to each student to gauge his overall proficiency in Language which will help us categorize student in “Low”, “Medium” and “High” level
- Rest of the 50 to 60 questions set will be different for these 3 categories

2.2.2 Selection of 47 common questions

These 47 common questions are spread across 12 sub-skills and 4 grades. The spread was decided based on weightage given by language experts to each sub-skill for the 4 grades respectively. The weights were given such that total is close to 100 for each grade. This was then consolidated –

<table>
<thead>
<tr>
<th>Sub-Skill Name</th>
<th>Grade-2</th>
<th>Grade-3</th>
<th>Grade-4</th>
<th>Grade-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify vowels</td>
<td>12</td>
<td>8</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Identify consonants</td>
<td>12</td>
<td>8</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Identify Matras</td>
<td>8</td>
<td>12</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Identify syllable blends</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Basic Vocab</td>
<td>15</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Word meaning</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Usage</td>
<td>10</td>
<td>10</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Recall</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Sentence structure Knowledge</td>
<td>5</td>
<td>7</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Punctuation Knowledge</td>
<td>0</td>
<td>5</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Grammar Knowledge</td>
<td>4</td>
<td>7</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Passage comprehension</td>
<td>8</td>
<td>8</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

Using these weights, number of question in each sub-skill and grade was decided –

<table>
<thead>
<tr>
<th>Sub-Skill Name</th>
<th>Grade-2</th>
<th>Grade-3</th>
<th>Grade-4</th>
<th>Grade-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify vowels</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Identify consonants</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
These questions will be from “Medium” difficulty level in each grade.

### 2.2.3 Scoring and logic to categorize students

The students were categorised into "Low", "Medium" and "High" proficiency – Once student answers these 47 question, these were scored to decide his overall proficiency level. Here, for assigning a student his level, questions will be assigned differential marking

- Grade 2 question – 1 mark
- Grade 3 question – 2 marks
- Grade 4 question – 3 marks
- Grade 5 question – 4 marks

### 2.2.4 Cut off range for each proficiency level

It was decided based on, what maximum % of question of each grade level, we expect the student (of that category) to answer –

Expected accuracy for each proficiency bucket assuming 4MCQs (10% guessing)

<table>
<thead>
<tr>
<th>Category</th>
<th>Grade-2</th>
<th>Grade-3</th>
<th>Grade-4</th>
<th>Grade-5</th>
<th>Expected score</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginner</td>
<td>70%</td>
<td>40%</td>
<td>20%</td>
<td>10%</td>
<td>29.9</td>
<td>1 to 30</td>
</tr>
<tr>
<td>Intermediate</td>
<td>90%</td>
<td>70%</td>
<td>40%</td>
<td>30%</td>
<td>55.7</td>
<td>31 to 55</td>
</tr>
<tr>
<td>Proficient</td>
<td>100%</td>
<td>90%</td>
<td>70%</td>
<td>50%</td>
<td>81.7</td>
<td>Above 55</td>
</tr>
</tbody>
</table>

### 2.2.5 Selection of rest of questions (adaptive component of the test)

Once the students have been categorized in the 3 categories, each one be given rest of the question as per their proficiency level. The below table shows which skills will be covered for which proficiency level
<table>
<thead>
<tr>
<th>Sub-Skill Name</th>
<th>Beginner</th>
<th></th>
<th>Intermediate</th>
<th></th>
<th>Proficient</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of Questions</td>
<td>Spread across Grades</td>
<td>No. of Questions</td>
<td>Spread across Grades</td>
<td>No. of Questions</td>
<td>Spread across Grades</td>
</tr>
<tr>
<td>Identify vowels</td>
<td>6</td>
<td>2,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identify consonants</td>
<td>12</td>
<td>2,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identify Matras</td>
<td>8</td>
<td>2,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identify syllable blends</td>
<td>8</td>
<td>2,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic Vocab</td>
<td>12</td>
<td>2,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word meaning</td>
<td>9</td>
<td>2,3,4</td>
<td>12</td>
<td>4,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usage</td>
<td></td>
<td></td>
<td>12</td>
<td>4,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recall</td>
<td>12</td>
<td>2,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentence structure Knowledge</td>
<td>9</td>
<td>3,4,5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Punctuation Knowledge</td>
<td></td>
<td></td>
<td>12</td>
<td>4,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammar Knowledge</td>
<td></td>
<td></td>
<td>12</td>
<td>4,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passage comprehension</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>3,4</td>
<td>7</td>
<td>5</td>
</tr>
</tbody>
</table>

### 2.2.6 Approach to item selection/development from baseline-

Based on the criterion the number of questions needed was calculated. The questions were selected from baseline test or new items were developed. **Before starting selection, items were removed from the pool of 231 questions based on following conditions** –

- Grade-wise performance on each question was seen, those questions which had average performance less than 20% or more than 80% were removed for the test. Discrimination (PBC) value below 0.1
- **After that from the remaining valid pool (for each grade sub skill), we will select items with the following conditions to meet the blueprint mentioned in point 1.2.2 (Question framework in crude test) and 1.2.5 (Question framework in fine test)**
  - Items with 4 options preferred
  - Items with single correct option preferred
  - Over all difficulty level of the test (item set) should be 50% respectively for the 3 levels (Students of grade 2 and 3 should have 50% performance on “Beginner” test set, Students of grade 3 and 4 should have 50% performance on “Intermediate” test set, Students of grade 4 and 5 should have 50% performance on “Proficient” test set)
2.3 **Endline Test Conduction:**

Sampling of schools was done manually during the baseline study. The following points in mind:
- Schools from Rural and Urban areas were selected for the study.
- Each classroom had roughly an equal number of male and female students.
- As per DEO’s instructions students from grade 1 were not tested. Hence, Only Grade 2nd to 5th were covered in sampling.
- 10 schools were randomly sampled from each district (Kota, Bundi and Jhalawar) to participate in the study. Out of these, 6 schools were kept as the main sample and rest 4 schools were kept as replacement schools.

The test was conducted during 13th February to 28th February. It was targeted to test 24 students from each school – 6 each from 2nd, 3rd, 4th and 5th who were tested in baseline. But due to absenteeism students did not come to school and testing was repeatedly done in the same school for even few students coming to school. Over 370 students were tested in endline study.

Challenges Faced
- Students were absent on the day of assessment, so the team had to visit their home and bring them to school for testing.
3 Quantitative Data Analysis

The data obtained from both the tests, baseline and endline have been analysed section wise in this chapter. The endline test was designed to have 3 levels of proficiency – Beginner level, Intermediate level, and Proficient level. Each student appeared for a crude level test and then based on her performance the next level was assigned.

The crude level test, assessed a student on all the sub-skills under four major skills i.e. Letter knowledge, vocabulary, syntactic & semantic knowledge and reading comprehension. Depending on the student’s performance on these, then he/she was assigned the overall language level.

A Beginner level student is at the basic level of language learning. The child has automatized all alphabets but few, can read and understand basic sight words with one or maximum two simple matras and is able to understand basic questions and answer them correctly based on the contextual sentence read out to them.

An Intermediate level student has automatized all alphabets, with correct sound and symbol relation, can spell the a given word correctly and use it in the proper context, can reconstruct a sentence in the meaningful manner and read a short para to answer factual retrieval questions based on it.

A proficient level student has good hold on syntactic and semantic knowledge to use words and sentences in the right context with proper spellings and grammar. Such child can read a story and make out inferential conclusions.

The chapter describes various data and the insights drawn from them. The data has also been analysed based on gender and age of the student. The data of baseline and endline was compared based on the common questions to calculate the performance gain in percentage of students.

This chapter presents the findings in the following sections:

1. Performance of Students on different levels of proficiency
2. Comparative performance of different age groups
3. Comparative performance of boys and girls
4. Gain in student performance from Baseline to Endline

3.1 Performance of students on different levels of proficiency

Finding 1: The crude level assessment is showing performance gain as we move from grade 2 to grade 5

All students from grade 2 to grade 5 appeared for crude level assessment. The performance of students in crude level was utilized as the basis to allot the next level of assessment. The results of the crude test showed that the performance of students was increasing as we moved from grade 2 to grade 5. The average performance was found increasing by 4% from grade 2 to 3, 11% from grade 3 to grade 4 and 6% from grade 4 to grade 5.
Finding 2: More than 60% Students in grade 2 and 3 were found to be in intermediate level, while in grade 4 and 5 number of students in proficient level is more.

Based on the performance of students in crude assessments, they were allotted the next level of test-Beginner, Intermediate or Proficient. The data shows that 73% of students in grade 2 and 62% of students in grade 3 were allotted intermediate level assessment. As expected, the percentage of student getting Intermediate assessments was found to be reducing from grade 2 to grade 5. In addition to this, the number of students getting proficient level was also found increasing from grade 2 to grade 5. This data showed that as the students moved from grade 2 to grade 5, they were showing more familiarity and proficiency in Hindi Language. This also in a way, corroborated that the diagnostic tool was able to analyse and classify students at all levels.

As can be seen from the graph, with each grade advancing, the percentage of students in the Beginner and Intermediate level is getting reduced. By grade 5, there are more number of proficient children and very few beginners. This is an indication that the test is able to differentiate between students at various grades and students within the grade depending on their mastery level.
Finding 3: The performance in grade 3 questions on identifying consonants, basic vocabulary, recall and reading comprehension were specifically found to be low in crude level of students which were allotted beginner language level. Also, Students in beginner language level have found content of grade 4 difficult, across sub-skills. The data show in below table.

The Crude level test was analyzed to understand the performance of beginner language level students across sub-skills. This analysis helps us understand what kind of questions is dragging the performance lower and at what level. This insight will help modify the test for better detection of Hindi learning levels among children.

The beginner language level students have performed low across grade level on questions pertaining to matras, word meaning, usage, and sentence structure knowledge. They have performed above 60% on the sub-skills questions belonging to grade 2. This explains why some of the students are still behind in their learning levels. As many students struggle with basic skills of identifying letters, matras and words, their low performance is expected due to the weak foundation.

<table>
<thead>
<tr>
<th>Sub-skill</th>
<th>Grade 2 Ques.</th>
<th>Grade 3 Ques.</th>
<th>Grade 4 Ques.</th>
<th>Grade 5 Ques.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify Vowels</td>
<td>60%</td>
<td>33%</td>
<td>32%</td>
<td></td>
</tr>
<tr>
<td>Identify Consonants</td>
<td>54%</td>
<td>17%</td>
<td>66%</td>
<td></td>
</tr>
<tr>
<td>Identify Matras</td>
<td>33%</td>
<td>49%</td>
<td>23%</td>
<td></td>
</tr>
<tr>
<td>Identify Syllabic blends</td>
<td>40%</td>
<td>60%</td>
<td>17%</td>
<td></td>
</tr>
<tr>
<td>Basic Vocabulary</td>
<td>66%</td>
<td>23%</td>
<td>32%</td>
<td></td>
</tr>
<tr>
<td>Word Meaning</td>
<td>30%</td>
<td>21%</td>
<td>8%</td>
<td>20%</td>
</tr>
<tr>
<td>Usage</td>
<td>34%</td>
<td>57%</td>
<td>17%</td>
<td>29%</td>
</tr>
<tr>
<td>Recall</td>
<td>48%</td>
<td>12%</td>
<td>12%</td>
<td></td>
</tr>
</tbody>
</table>
Here some of the questions are picked to showcase the data at the question level for the finding 4 discussed above.

**Sample question 1:** This is a grade 3 level question from sub-skill Identify consonants. Only 20% students from grade 2 to 4 were able to answer this question, while the performance was around 40% in grade 5. The performance is overall low on this question.

In this question, the instruction was to click on a word which has letter ज्ञ in it. The instruction was spoken, but the options did not have any voice over i.e. students had to read & identify the words and their units themselves. The average score is indicating that students across grades are struggling to identify such sanyuktakshar when there is no support of voice over.

<table>
<thead>
<tr>
<th>Q-code</th>
<th>Category</th>
<th>Q grade</th>
<th>Sub-skill</th>
<th>options</th>
<th>PBC</th>
<th>Overall Avg. Performance</th>
<th>Beginner level avg</th>
<th>Gr 2</th>
<th>Gr 3</th>
<th>Gr 4</th>
<th>Gr 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>16104</td>
<td>Beginner</td>
<td>3</td>
<td>Identify Consonants</td>
<td>3</td>
<td>0.38</td>
<td>23%</td>
<td>8%</td>
<td>20%</td>
<td>23%</td>
<td>20%</td>
<td>40%</td>
</tr>
</tbody>
</table>

**Sample question 2:** This is a grade 3 level question from sub-skill Recall. The performance was found to be 33% at the grade 3 level which was seen going to 60% which was significantly high as compared to lower classes.

In this question, students had to identify and select the spoken sentence (हाथी खेल करता है) on voice over. The options also had the voice overs and hence students had the support of identifying the sentence here on the basis of the listening skill. Considering there was a listening aid available, the performance could have been higher. However, the
distractors are too close and that might have made some students make a mistake. This could be due to the distracters (other similar options) or in remembering 4 words sentence and then later clicking on it.

<table>
<thead>
<tr>
<th>Q-code</th>
<th>Category</th>
<th>Q grade</th>
<th>Sub-skill</th>
<th>#options</th>
<th>PBC</th>
<th>Overall Avg. Performance</th>
<th>Beginner level avg</th>
<th>Gr 2</th>
<th>Gr 3</th>
<th>Gr 4</th>
<th>Gr 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2833</td>
<td>Beginners</td>
<td>3</td>
<td>Recall</td>
<td>3</td>
<td>0.06</td>
<td>40%</td>
<td>27%</td>
<td>33%</td>
<td>36%</td>
<td>44%</td>
<td>60%</td>
</tr>
</tbody>
</table>

**Sample Question 3:** This is a grade 3 level question from sub-skill Recall. The performance ranges from 25% to 40% question from grade 2 to grade 5. In this question, students had to choose the spoken sentence on voice over. The options also had the voice over which means if the student would take the cursor over each of these options, he/she could hear the sentences. This could be due to the distracters (other similar options) or in remembering 6 words sentence and then later clicking on it. Interesting to note that the performance as the length of the sentence increases.

<table>
<thead>
<tr>
<th>Q-code</th>
<th>Category</th>
<th>Q grade</th>
<th>Sub-skill</th>
<th>#options</th>
<th>PBC</th>
<th>Overall Avg. Performance</th>
<th>Beginner level avg</th>
<th>Gr 2</th>
<th>Gr 3</th>
<th>Gr 4</th>
<th>Gr 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2678</td>
<td>Beginners</td>
<td>3</td>
<td>Recall</td>
<td>3</td>
<td>0.32</td>
<td>30%</td>
<td>15%</td>
<td>25%</td>
<td>36%</td>
<td>22%</td>
<td>40%</td>
</tr>
</tbody>
</table>

**Sample Question 4:** This is a grade 3 level question from sub-skill of Reading Comprehension. The performance ranges from 24% to 40% for grades 2 to 5. In this question, students were to select the right question for the given statement. The fact the question making skill isn’t practiced often in the classrooms but the students are always in the habit of answering them, higher numbers of students were found to be struggling. The low PBC (0.1) in this question also suggests that even the good performers on the test weren’t able to perform on this question.

<table>
<thead>
<tr>
<th>Q-code</th>
<th>Category</th>
<th>Q grade</th>
<th>Sub-skill</th>
<th>#options</th>
<th>PBC</th>
<th>Overall Avg. Performance</th>
<th>Beginner level avg</th>
<th>Gr 2</th>
<th>Gr 3</th>
<th>Gr 4</th>
<th>Gr 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>4647</td>
<td>Beginners</td>
<td>3</td>
<td>Reading Comprehension</td>
<td>3</td>
<td>-0.10</td>
<td>29%</td>
<td>29%</td>
<td>24%</td>
<td>29%</td>
<td>33%</td>
<td>40%</td>
</tr>
</tbody>
</table>

**Finding 4:** The Beginner and Intermediate language level children are seen to be struggling with the reading comprehension skill as compared to the proficient level students.
The graph below shows that the performance of students on grade 3 level questions has been particularly low across Beginner, Intermediate and Proficient levels.

In the proficient level as well, the Grade 4 students show higher performance than grade 5 students on usage and reading comprehension. This can be due to higher presence of more proficient level children at grade 4 level.

### 3.2 COMPARATIVE PERFORMANCE OF STUDENTS IN DIFFERENT AGE GROUPS

As this diagnostic test has been constructed for early grade readers, it is targeting students from the age bracket 6 to 10 years. As per government’s education policy, children are admitted in a particular class based on their age in elementary school years. This section is exploring insights based on the age brackets of the test takers to explore if there is any performance pattern between the language bracket allocated and sub-skill wise performance.
3.2.1 Allotment of language bucket in the fine test

According to the age distribution, maximum number of students, across age brackets, have performed between 30 to 55% on the crude test and hence were allotted Intermediated language bucket in the finer test. As we move from ‘Below 8 years’ to ‘10-11’ years’ age bracket higher percentage of student that were allotted proficient level. There were only 23 students above 11 years; hence they were ignored from this analysis.

<table>
<thead>
<tr>
<th>Age brackets</th>
<th>&lt;8yrs</th>
<th>8 to 9 yrs</th>
<th>10 to 11 yrs</th>
<th>&lt;8</th>
<th>8-9</th>
<th>10-11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginners</td>
<td>18%</td>
<td>18%</td>
<td>8%</td>
<td>9</td>
<td>27</td>
<td>8</td>
</tr>
<tr>
<td>Intermediate</td>
<td>72%</td>
<td>52%</td>
<td>46%</td>
<td>36</td>
<td>79</td>
<td>48</td>
</tr>
<tr>
<td>Proficient</td>
<td>10%</td>
<td>30%</td>
<td>46%</td>
<td>5</td>
<td>45</td>
<td>48</td>
</tr>
</tbody>
</table>

3.2.2 Performance percentage of different age groups in all sub-skills

Performance has dipped considerably across age brackets as we move from skill 1 to skill 4. Students below the age of 8 years are performing with more than 50% accuracy; across sub-skills under skill 1 i.e. letter knowledge. This indicates that the knowledge of basic units of the written script which is foundational is mastered by students by the end of grade 2 (as the child is 7 years old in grade 2, as per government’s education policy). They are also performing upto 48% in the basic vocabulary questions which is primarily built by listening and decoding pictures at early age to understand the written text. Best performers across sub-skills are of the age bracket 10-11 years. Though one sub-skill, punctuation knowledge has got low accuracy even with age bracket 10-11 years, which means not enough exposure is provided to students in the later years of elementary schooling to understand the basic punctuation marks to form syntactically meaningful sentences.

♦ In skill 1 - Letter Knowledge, students below 8 years have performed least at identifying vowels and best at identifying syllabic blends, students between 8-9 years have performed least at identifying consonants and best at identifying syllabic blends and students between 10-11 years have performed least at identifying matras and best at identifying syllable blends. This indicates that early age children struggle with vowel identification, as more consonants are introduced, some are similar looking and some similar sounding, complex letters like sanyuktakshars and half letters are introduced later, students struggle with consonant identification. Matras seems to be a struggle even in later years of elementary education.

♦ In skill 2 - Vocabulary, least performance across age brackets is on word meaning. Next comes usage, with best performance at basic vocabulary which involves high frequency, sight word knowledge testing.

♦ In skill 3 - Syntactic/Semantic knowledge, least performance is at punctuation knowledge across age brackets then comes Sentence Structure Knowledge followed by recall which tests for working memory and surprisingly best performance at grammar knowledge.
In skill 4 - Reading Comprehension, performance percentage is understandably increasing as we move from below 8 to 10-11 years’ age bracket.

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Skill Name</th>
<th>Sub-skill Name</th>
<th>&lt; 8 years</th>
<th>8-9 years</th>
<th>10-11 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Letter Knowledge</td>
<td>Identify vowels</td>
<td>51.6%</td>
<td>67.0%</td>
<td>70.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identify consonants</td>
<td>54.1%</td>
<td>58.4%</td>
<td>65.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identify Matras</td>
<td>52.8%</td>
<td>61.2%</td>
<td>61.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identify syllable blends</td>
<td>59.3%</td>
<td>70.1%</td>
<td>80.0%</td>
</tr>
<tr>
<td>2</td>
<td>Vocabulary</td>
<td>Basic Vocabulary</td>
<td>48.5%</td>
<td>60.7%</td>
<td>70.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Word meaning</td>
<td>37.5%</td>
<td>44.5%</td>
<td>51.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Usage</td>
<td>45.0%</td>
<td>55.7%</td>
<td>67.8%</td>
</tr>
<tr>
<td>3</td>
<td>Syntactic/Semantic</td>
<td>Recall</td>
<td>32.6%</td>
<td>44.7%</td>
<td>52.5%</td>
</tr>
<tr>
<td></td>
<td>knowledge</td>
<td>Sentence structure</td>
<td>29.0%</td>
<td>39.3%</td>
<td>51.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Knowledge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Punctuation Knowledge</td>
<td>29.9%</td>
<td>32.6%</td>
<td>36.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammar Knowledge</td>
<td>38.2%</td>
<td>46.2%</td>
<td>55.4%</td>
</tr>
<tr>
<td>4</td>
<td>Reading Comprehension</td>
<td>Passage comprehension</td>
<td>38.1%</td>
<td>46.4%</td>
<td>56.4%</td>
</tr>
</tbody>
</table>

Sample question 5: This question belongs to the sub-skill of vowel identification. Students were to click on the sound of the vowel (ृ) spoken on voice over and click the corresponding symbol. As we move from below 8 years to 11 years the performance on this question has increased by more than 3%. Performance % does not have too much gap between that of age range below 8 and between 10 to 11 years. This highlights the need for better reinforcement and focused efforts on such vowels. Also, since Hindi language does not very specifically differentiate between the two vowels asked in the question in its colloquial form, this gap in identification is likely to increase with increase in age group.
Sample question 6: This question belongs to the sub-skill 2 - Identify consonants of skill 1, letter knowledge. Students were to click on the sound of the sankyunkakshar (श्र) spoken on voice over and click the corresponding symbol. As expected, performance on this question is increasing as we move from below 8 to 11 years of age ranges. This is corroborative of the understanding that these letters are more frequently used at higher grades and hence students between 9 to 11 years are able to do better performance. However, the performance is 54.6% which indicates that there are still significant numbers of higher grade students who are struggling with the identification of sanyuktaksars.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>&lt;8</th>
<th>8 to 9</th>
<th>10 to 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>34.0%</td>
<td>43.2%</td>
<td>54.7%</td>
</tr>
</tbody>
</table>

Sample question 7: This question belongs to the sub-skill 3 - Identify matras of skill 1, letter knowledge. Students were to click on the sound of the syllable (गि) spoken on voice over and click the corresponding symbol. Performance on this question has increased as we move from below 8 years to 9 years of age but has dipped for students in the age range of 10 to 11 years.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>&lt;8</th>
<th>8 to 9</th>
<th>10 to 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>52.0%</td>
<td>55.5%</td>
<td>44.6%</td>
</tr>
</tbody>
</table>
Sample Question 8: This question belongs to the sub-skill 4 - Identify syllabic blends of skill 1, letter knowledge. Students were to click on the sound of the word (निर्मल) spoken on voice over and click the corresponding word. As expected, performance on this question is increasing as we move from below 8 to 11 years of age ranges. The words with raf matras are introduced in the later years of elementary education, even though the words having them are introduced from grade 3 itself. 46% performance by students below 8 years suggests that most of them must have tried to attempt this question phonetically, by hearing the letters in the words (ि, र्, ल) they would have selected the correct word. They might have understood the matra or might have just followed the sounds phonetically to understand the correct answer.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>&lt;8</th>
<th>8 to 9</th>
<th>10 to 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>46.0%</td>
<td>60.5%</td>
<td>71.3%</td>
</tr>
</tbody>
</table>

Sample Question 9: This question belongs to the sub-skill 6 - Word Meaning of skill 2, Vocabulary. In this question, question stem was spoken on VO and students had to read the options themselves and select the right answer. The performance on this question is increasing as we move from below 8 to 11 years of age ranges. Low performance of students below 8 years of age may have struggled with words like (वस्तु, प्रयोग) to comprehend the question asked.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>&lt;8</th>
<th>8 to 9</th>
<th>10 to 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>46.0%</td>
<td>53.7%</td>
<td>63.7%</td>
</tr>
</tbody>
</table>
Sample Question 10: This question belongs to the sub-skill 10 i.e. Punctuation Knowledge of skill 3, Syntactic and Semantic Knowledge. In this question, question stem was spoken and the students had to select the right punctuation mark. Overall performance on this question across age ranges in low but given that, percentage is increasing as we move from below 6 years to 11 years’ range. Punctuations are not consciously taught and practiced at the classroom level and in that scenario, students often struggle with punctuations. Especially when the sentence structure is confusing and might requires complex comprehension, the students tend to struggle. The same is evident from the data here. These concepts can also be reinforced while reading aloud in story telling sessions where in reading with proper expressions and pauses helps in understanding the purpose of these symbols in language learning.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>&lt;8</th>
<th>8 to 9</th>
<th>10 to 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>16.3%</td>
<td>21.6%</td>
<td>24.8%</td>
</tr>
</tbody>
</table>

Sample Question 11: This question belongs to the sub-skill 11 - Grammar Knowledge of skill 3, Syntactic and Semantic Knowledge. In this question, question stem was spoken and the students had to read the options themselves and select the right answer. The performance on this question across age ranges is low but given that, percentage is increasing as we move from below 8 years to 11 years’ range. The concept of tense isn’t explicitly taught at elementary school but by reading and listening the syntactic framework of the sentence in the context of a particular time of reference and communicating in the same language at home, happens indirectly.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>&lt;8</th>
<th>8 to 9</th>
<th>10 to 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>42.9%</td>
<td>56.8%</td>
<td>64.4%</td>
</tr>
</tbody>
</table>

Sample Question 12: This question belongs to the sub-skill 12 - Passage Comprehension of skill 4, Reading comprehension. In this question (belonging to grade 4), question stem was spoken and the students had to read the options themselves and select the right answer. The performance on this question across age ranges is relatively. However, percentage is increasing as we move from below 8 years to 11 years’ range. This suggests the students in the age range between 10 to 11 years were able to read the passage with better understanding than other and hence the answer the question correctly.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>&lt;8</th>
<th>8 to 9</th>
<th>10 to 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>39.9%</td>
<td>48.3%</td>
<td>57.3%</td>
</tr>
</tbody>
</table>
3.3 **COMPARATIVE PERFORMANCE OF BOYS AND GIRLS**

This section contains test insights on the performance of boys and girls put separately. This analysis has following three sections to it, overall % performance of boys and girls on crude test, grade wise % performance of boys and girls on crude test, and sub-skill wise % performance of boys and girls on crude test.

- **Overall % performance of boys and girls on crude test** - Girls have performed 4.5% higher than boys on the endline diagnostic test. A total of 162 boys and 167 girls participated in the study.

- **Grade wise % performance of boys and girls on crude test** - Girls are performing higher than boys across grades with 3.64% in grade 2, 2.09% in grade 3, 2.84% in grade 4 and 6.28% in grade 5. T-Test was not significant for all the classes for 95% confidence level.

- **Sub-skill wise % performance of boys and girls on crude test** - Girls are in lead across sub-skill on the average performance, though the difference is limited between 3 to 6% with a difference of 7.6%, which is highest, on sub-skill ‘Usage’ under Vocabulary skill.
The maximum performance gap between girls & boys, has been noticed is of 7.57% in the sub-skill 7-Usage, under the Skill 2, Vocabulary. From the data it appears that girls have slightly better language comprehension capacity. However, the difference indicated is not very significant.

Under skill 3 - Syntactic & Semantic Knowledge, girls have shown maximum performance gap of 5.86% more from boys at the sub-skill 9 - sentence structure knowledge and least performance gap of 0.60% more from boys at the sub-skill 11 - grammar knowledge. This suggests that girls are better at placing words together to form a meaningful sentence but when asked questions on grammar skills like plural, gender, tense, they aren’t any better than boys but as weak at it as boys.

From the analysis, it can be seen that the girls are overall performing higher that the boys in the skill of reading comprehension. Various factors could be responsible for this and factors like overall attention span are worth getting into the depth of at the classroom level.

Sample question from few sub-skills where girls performed better than boys:
Sample Question 13: The question belongs to sub-skill 5 - Basic Vocabulary, at which girls have performed overall 6.1% higher than boys. Girls have performed 14% better on this question which indicates that they are more attentive to these nouns and their meaning which are introduced in many stories at early grades but also are part of the daily lives as an essential societal necessity.

Sample Question 14: This is the question from the sub-skill 7 i.e. Usage, where maximum performance gap has been noticed between girls and boys with girls in the lead by 7.57%. On this question girls have performed 9.13% better than boys.

Explanation: This question also belongs to the sub-skill 7 i.e. Usage under the skill 2 - Vocabulary. There is a 14.9% higher performance of girls which indicate that boys struggle a little more with identifying the correct context for the given question.

3.4 Gain in Student Performance from Baseline to Endline

3.4.1 Sub-Skill wise Average gain analysis

In end-line diagnostic test, the average performance gain is 5.1%. There is more than 15% gain in the student performance on end line test in three sub-skills Grammar Knowledge (Sub-skill code no. 11), Reading Comprehension Sub-skill code no. 12), and Basic Vocabulary (Sub-skill code no. 5). In end-line test performance has dipped at 2 sub-skills namely Punctuation Marks (Sub-skill code no. 10) and Identify Matras (Sub-skill code no. 3).
Question examples for skills where performance has dipped in Endline compared to Baseline

**Skill 1 - Identifying Vowels:** On this question there is a gain of 0.64% in end line test. This indicates that differentiating between long & short vowels is a major area of concern in early grade language learning which isn’t being given adequate attention. At the classroom level, this requires special emphasis.
### Skill 3 - Identifying matras:
In this question students were expected to click on the symbol of the sound गि. Students have underperformed on this question by 7.56% which indicates that in spite of baseline results clearly indicating that students across grades are struggling with matras especially differentiating long from short, not much practice, activities were carried which has retained the confusion with more children attempting it incorrectly.

As spoken Hindi does not always have the room to capture the difference between short and long pronunciation of the matra, specific and directed inputs to understand the role of matras is needed.

### Students were to click on the word corresponding to the picture shown. They have underperformed on this question by 1.30% which indicates that reading words with matras is still a major issue and not much emphasis has being given to practice this skill with interactive activities in the classrooms.

### Sub-skill - 10: Punctuation knowledge:
The concept of using punctuation mark in a sentence comes under syntactic knowledge & semantic knowledge but it has been consistently observed that students are highly confused and clueless on which punctuation mark to use given a contextual sentence. The sentence in the question was spoken as well for a child to listen to the tone and figure the type of sentence and select a mark. As mentioned earlier, students do struggle with punctuations and will find it really difficult to pick the specific punctuation when the context becomes the deciding factor.

Focused efforts are required at the classroom level. Exposure to reading can also help as children learn tacitly in the process.

**Question examples for skills where performance has increased in Endline over Baseline**
**Sub-skill 11 - Grammar Knowledge**

Explanation: Students have gained on the questions of syntactic knowledge under which comes concept of Kaarak. As there isn’t any ‘ko’ after the blank hence, it can’t be option 1 but option 2. At early grades, this concept isn’t explicitly taught but while teaching sentence construction and also reading many stories having complete meaningful sentences adds to towards imbibing this concept indirectly.

More students (by 15.98%) have picked on syntactic knowledge under which comes the concept of tense. This can also be due to cognitive development due to age and overall exposure to reading at the classroom level.

---

**Sub-skill - 12: Reading Comprehension**

Performance on this question has increased by 14.78% which is testing for the reading comprehension. Students were supposed to read the given sentence and answer basic why-questions based on it. More number of students was able to answer the questions despite the distractors (closely related words as other options). This is the first stage of comprehension which is more to do with factual retrieval.

16.81% performance gain on this question indicates that more students were able to correctly read & comprehend the sentences to put them in the sequence based on the context it contained. Sequencing is the second skill after factual retrieval under RC where in larger context is understood to the events in the correct order.
**Sub-skill - 5: Basic Vocabulary:** There is a 15.29% gain on this question which suggests that stories which have local connect are being discussed in the classroom for more students to pick up understanding on the nouns and relating them to their respective professions. Such discussions exposes the students with words in a context, used more in our colloquial language communication which enhances basic vocabulary skill.

Explanation: 12.94% performance gain in this question indicates more students were able to decode the pictures with right words from the pool of basic vocabulary and selected the correct sentence.

3.4.2 Question wise gain

Performance on 77.77% of common test items has shown gain in comparison between the two rounds.

Students have gained on the syntactic knowledge under which comes concept of Kaarak. As there isn’t any ‘ko’ after the blank hence, it can’t be option 1 but option 2. At early grades, this concept isn’t explicitly taught but while teaching sentence construction and also reading many stories having complete meaningful sentences adds to towards imbibing this concept indirectly.

<table>
<thead>
<tr>
<th>Item Code</th>
<th>Subskill</th>
<th>Student responses - Common</th>
<th>Performance (%) Round 1</th>
<th>Performance (%) Round 2</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>5768</td>
<td>11</td>
<td>180</td>
<td>37.22%</td>
<td>70.00%</td>
<td>32.78%</td>
</tr>
</tbody>
</table>
More students have picked on syntactic knowledge under which comes concept of tense. At early grades, this concept isn’t explicitly taught but while teaching sentence construction and also reading many stories having complete meaningful sentences adds to towards imbibing this concept indirectly.

<table>
<thead>
<tr>
<th>Item Code</th>
<th>Sub-skill</th>
<th>Student responses - Common</th>
<th>Performance (%) Round 1</th>
<th>Performance (%) Round 2</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>12045</td>
<td>11</td>
<td>108</td>
<td>40.70%</td>
<td>62.00%</td>
<td>21.30%</td>
</tr>
</tbody>
</table>

Explanation: Students were supposed to make a meaningful sentence from the given jumbled words. The sentence to be made was spoken on the VO to test the working memory of students, to see how many were able to recall the said sentence put the words in the exact same order. This suggests that more students have picked on syntactic knowledge & semantic knowledge under which comes concept of meaningful sentence making. Also, students have been more attentive to the voiceover to remember the said sentence and reproduce it in the correct manner. This can be related to an overall increased exposure, directed efforts by teacher and overall cognitive development due to time passage. The gain is considerably high.

<table>
<thead>
<tr>
<th>Item Code</th>
<th>Sub-skill</th>
<th>Student responses - Common</th>
<th>Performance (%) Round 1</th>
<th>Performance (%) Round 2</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>5743</td>
<td>9</td>
<td>157</td>
<td>24.8</td>
<td>45.9</td>
<td>21.1</td>
</tr>
</tbody>
</table>

Students were supposed to click on the said sentence which they heard. The sentence to be selected was spoken on the VO to test the working memory of students, to see how many were able to recall the said sentence and choose the same from the given options. Students had to read the sentences themselves, if need be by hovering the cursor over the options, they were able to listen to the option sentences. This suggests that more students have picked on syntactic knowledge & semantic knowledge under which comes concept of recall. Also, students have been more attentive to the voiceover to remember the said sentence and reproduce it in the correct manner. However,
there is further scope to make sure that more students understand this better and there is higher performance on the question like this which has a very simple sentence.

<table>
<thead>
<tr>
<th>Item Code</th>
<th>Sub-skill</th>
<th>Student responses - Common</th>
<th>Performance (%) Round 1</th>
<th>Performance (%) Round 2</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>14714</td>
<td>8</td>
<td>364</td>
<td>54.67</td>
<td>74.67</td>
<td>20.05</td>
</tr>
</tbody>
</table>

This is a reading comprehension question. To answer this, child had to read the short story without any voice over help. The question stem and options were audible if the cursor was taken over them. As can be seen from the data, more students have been able to do the question correctly this time. The question is from the text directly and also there is a familiarity angle which might be playing a role here. this

<table>
<thead>
<tr>
<th>Item Code</th>
<th>Sub-skill</th>
<th>Student responses - Common</th>
<th>Performance (%) Round 1</th>
<th>Performance (%) Round 2</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1158</td>
<td>12</td>
<td>58</td>
<td>69.9</td>
<td>87.9</td>
<td>18.97</td>
</tr>
</tbody>
</table>

Top 5 questions with negative improvement

The concept of using punctuation mark in a sentence comes under syntactic knowledge & semantic knowledge but it has been consistently observed that students are highly confused and clueless on which punctuation mark to use given a contextual sentence. The sentence in the question was spoken as well for a child to listen to the tone and figure the type of sentence and select a mark. Those who attempted this correctly last time weren’t also clear with the concept to choose a wrong one this time.

<table>
<thead>
<tr>
<th>Item Code</th>
<th>Sub-skill</th>
<th>Student responses - Common</th>
<th>Performance (%) Round 1</th>
<th>Performance (%) Round 2</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>16143</td>
<td>10</td>
<td>255</td>
<td>81.1</td>
<td>52.1</td>
<td>-29.0</td>
</tr>
</tbody>
</table>
In this question, students were to read the sentence which appeared on the screen for 5 seconds and then click on the corresponding picture out the 4 which appeared next on the screen. Students have underperformed on this by 10.44% this indicates that this skill requires further emphasis and retention. Students might have working memory to answer thing question, but it requires more training and conscious inputs.

<table>
<thead>
<tr>
<th>Item Code</th>
<th>Sub-skill</th>
<th>Student responses – Common</th>
<th>Performance (%) Round 1</th>
<th>Performance (%) Round 2</th>
<th>Difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12218</td>
<td>9</td>
<td>364</td>
<td>80.4</td>
<td>70.0</td>
<td>- 10.44</td>
</tr>
</tbody>
</table>

In this question, students were to select the right question based on the sentence given. Students have underperformed on this question by 10%. This is one area where focused classroom efforts can help and students can learn language more thoughtfully. Making students practice on making questions from the answer is a higher order skill at this grade level and teachers are encouraged to focus their efforts.

<table>
<thead>
<tr>
<th>Item Code</th>
<th>Sub-skill</th>
<th>Student responses – Common</th>
<th>Performance (%) Round 1</th>
<th>Performance (%) Round 2</th>
<th>Difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3667</td>
<td>12</td>
<td>120</td>
<td>39.1</td>
<td>29.1</td>
<td>- 10.0</td>
</tr>
</tbody>
</table>

3.4.3 Student wise gain analysis

The performance has increased from baseline to endline for 64.5% students. 35.4% students have underperformed on the endline test. The overall gain comparison is based on the data of 370 students and 45 common test items in baseline & endline test versions. Performance gain of 32.1% students is in
the range of 0 to 10% and the performance gain of 21.6% students is in the range of 10% to 20%. About 10.8% students have shown performance gain above 20% from baseline to end-line.

### 3.4.4 Grade wise gain analysis

70% of grade 5 students showed performance gain on endline diagnostic test. With 67% and 66% of grade 2 and 4 students, respectively, showing overall similar performance gain, the least performance gain is observed in grade 3 with 60% students performing better on endline test.

<table>
<thead>
<tr>
<th>Grade</th>
<th>No. of students</th>
<th>Number of students with performance gain</th>
<th>Percentage of Students with performance gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>91</td>
<td>61</td>
<td>67</td>
</tr>
<tr>
<td>3</td>
<td>93</td>
<td>56</td>
<td>60</td>
</tr>
<tr>
<td>4</td>
<td>94</td>
<td>62</td>
<td>66</td>
</tr>
<tr>
<td>5</td>
<td>92</td>
<td>65</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>370</td>
<td>244</td>
<td>65</td>
</tr>
</tbody>
</table>
4 SKILL WISE ANALYSIS

Qualitative skill wise analysis of data helps in linking the performances with learning trends/patterns of typical mental models exhibited by children in responding to certain type of questions. The qualitative insights serve as a feedback for teachers to focus on certain skills/sub-skills/question types or certain errors that a larger number of children consistently make. These errors/misconceptions create a cascading effect if not corrected and children might internalize these as their learning shortcomings.

While doing the Qualitative Analysis of the End line test results, we looked at the performance of children on all the questions and also the PBC value, an assessment parameter which highlights the performance on a particular question item with respect to performance on overall test. If this value is nearing 0.1 or below that means the best performers on the overall test have faced issue with this question and it helps us identify concept misconceptions which are existing amongst most of the students including the good performers which needs to be seriously dealt with in the classrooms. Studied performance accuracy to find insights skill-wise, grade wise, language level wise and tried to see the patterns like:

✓ on which type of questions students belonging to a specific level of grade or at a specific level of language learning were performing well
✓ on which type of questions children belonging to a specific level of grade or at a specific level of language learning were performing poorly
✓ what type of gaps in language acquisition were observed in the performance patterns

The insights from the qualitative analysis are summarized skill wise below:

4.1 SKILL 1 – LETTER KNOWLEDGE

4.1.1 Sub-skill 1 – Identifying Vowels

4.1.1.1 Children were seen struggling in recognizing ह्स्व and दीर्घ letters, which included स्वार as well as मात्रा, example- इ-ई, उ-ऊ, ओ-औ, ए-ऐ.
In this question, students were to identify the letter for the sound. Overall उ was selected by 33.5% students and the correct answer ऊ by 58.7% students. Grade wise, 2nd and 4th graders seemed to be more confused, as they scored 52% and 53% respectively; whereas, 3rd and 5th graders scored 64% and 66% respectively.

It is evident from the data that students of different grades have confusion between उ and ऊ.

In this question, again the students were expected to identify the correct letter for the sound. Though overall the correct answer ओ did get higher % (59%); but grade wise performance was not too satisfactory. 56% of grade 2 children, 69%, of grade 3 children and 50% and 60% of grades 4 & 5 children answered the question correctly.

As can be seen from the question, students are not able to differentiate between the sounds of इ and ई. They seem to be more familiar with ई sound, so even if they are shown इ, they select ई. The overall performance on both these options is almost the same. Grade 2 children seem to be struggling the most with performance as low as 20%. Performance of grades 3 to 5 children was between a broad range of 54% to 60%.

1.1.1 Students are able to score well on non-similar letters & sound or the given letter but can’t differentiate well between the similar sounding letters.
In this question where the child had to read the letter and select the corresponding sound for इ from आ & इ, 50% of children of grade 2 are able to answer the question correctly. The performance goes on increasing with increase in grade with grade 5 students performing at 88% accuracy.

Performance is increasing from 47% to 73% on this question across grades. Grade 4 students have performed less than grade 3 students. Also, result of students allotted beginner category has performed below that of student’s allotted Intermediate category, by almost half.

4.1.2 Sub-skill 2 – Identifying Consonants

4.1.2.1 Students across grades are struggling with identifying a specific sankyukt akshar from the pool

Overall performance % shows some confusion between with क्ष & श्र on this question. क्ष was spoken on voice over but around 46% students selected क्ष. 33% students still selected श्र which is significantly different from क्ष, both in its script and sound. Strangely here, grade 3 children showed lower performance (31%) as compared to grade 2 children at 41%. Grade 4 and grade 5 results were quite close i.e. 54% and 58% respectively.
4.1.2.2  Students are struggling in selecting from letters which are similar sounding

In this question, students were to hear both the sounds and click on green button if they are same and red if they are of different letters. Percentage of children selecting the wrong answer is 45% this indicates they think there's no difference between the sounds of ड and ध. Not much difference was seen in the grade wise performance with 57% in Grade 2, 47% in Grade 3, 57% in grade 4 and 58% in grade 5.

4.1.2.3  Students across grades are struggling with the questions checking for the correct sounds.

In this questions, students had to click on the icon of the sound file of ध्र. The option sounds were of letters श and श्र. Performance on both the options is almost 50%, showing very high confusion between the two letters. Grade wise, the lowest performance is by grade 2 (33%) and after that by grade 5 children (40%). Grade 3 the highest with 57% and grade 4 with 56%. This could also be because श्र has श in its sound. Children who do not have clear understanding of the mixed letters or Sanyuktakshars, would not be able to differentiate between the two.

In this question, students were instructed to select a word with letter ज्ञ. Average performance was seen to be low for the correct option, आज्ञा with 23%. Maximum number of children selected क्षमा (around 48%) and rest of them selected श्रम (around 24%). Grade wise performance % is below 25% in grades 2, 3 & 4 with an increase in grade 5 which had 40% accuracy on this question.
Students seem to be randomly selecting the answer here. Also, because the word doesn’t start with झ, students might have found it difficult to arrive at the correct answer.

4.1.2.4 Students across grades have confusion between letters, स & श.

In the above questions, students were to identify the given letter ‘स’ from the sounds स and श. Nearly 50% students have selected the sound of श as the correct answer. Grade wise performance is as follows with grade 2 (47%), grade 3 (50%), grade 4 (78%), grade 5 (60%), though the overall PBC value is less than 0.1 which suggests that the best performing students have attempted this incorrectly. The reason of this confusion could be the local dialect influence. Both the letters have similar sounds, they are similar in scripts and dialect wise, they are used interchangeably.

4.1.2.5 Students are struggling with questions in which they have to identify sanyukt akshars (compound letters) and ardhakshars (half letters).

In this question, students had to identify the spoken letter झ from the given letters- झ, श्र & ज. Around 19% and 24% students chose ज and श्र respectively. Grade-wise; grade 3 (38%) performed lower than grade 2 (47%) whereas grade 4 and grade 5 performed more than 80%. Overall, it can be seen that as grades increase, there is a higher clarity in terms of understanding of letters.
In this question, students had to determine which of these words have a half letter in it - चम्चा, चम्च, चम्का. Most common wrong answer is चम्चा. This confusion might have occurred due to incorrect pronunciation of this word which gives false sense that letter म is half in it. Grade 2 and 4 have performed with 33% and 44% respectively and grade 3 and 5 performance were much better, around 64% and 80% respectively.

In this question, students had to determine which letter is half in the word पत्थर - प, त, थ. The correct option त has 43% performance, थ and प, 36% and 22% respectively. As can be seen from the data, there is some confusion between त and थ, which can be due to the way the word is pronounced. Also, the grade level performances were very close to each other which is indicative of the wrong notion staying unchanged even with advancing grades. 40% children of grade 2, 46% of, grade 3, 44% of grade 4 and 40% of grade 5 children could do this question correctly.

In the question, students had to click on the sound of that word which contained a half letter. The options give were- नकळ or चप्पल. Performance on the both the options were very close with 49% and 51% on नकळ and चप्पल respectively. Grade wise performance was also very close to each other and not progressing with the increasing grade level. This suggest students are struggling with the concept of half letter as this concept isn't being explicitly practiced from a pronunciation perspective to make each sound making the clear to spell it correctly. The reason adding to this confusion might be because in this question students are not able to read the word and hence not able to pick the half letter from the script.
4.1.3 Subs-kill 3 – Identifying Matras

4.1.3.1 Students across grades are selecting the first option when they are confused about the correct matra

In this question, the students had to pick the letter क with appropriate Matra to make a meaningful word. Though the question has high performance as compared to others, children who are not able to opt the correct answer are by default picking the first option and hence the most wrong option was काताब.

The common wrong answer in this question is बटाआ with 27.59% students selecting it. This indicates if children do not recognise उ matra, they select the one they know which is of आ. Though the maximum students did select the right syllable, given this a high frequency word.

4.1.3.2 Students are struggling in differentiating between short & long vowels like इ/ई or उ/ऊ matra letter.
In this question, students were to select the first syllable to make the word लीची. There is a 10% difference between the attempt percentage of लीची (33%) and लीघि (around 22%) which indicates that students are struggling in differentiating between long and short vowels. Interesting thing to notice is that grade wise, grade 2 (31%) performed better than grade 3 (22%) and grade 4 (53%) performed better than grade 5 (30%), the more students are moving in the higher grades they are losing the touch with basics which might be due to not enough reinforcement.

In this question, students were to select the first syllable to make the word मूली. A clear confusion was noticed in the attempts on मूली and मुली which indicates that students are struggling in differentiating between long and short vowels. Another interesting thing to notice on this question was that grade 4 performed much better than grade 5. This indicates that as the grades increase, focusing on reinforcement can help mitigate such performance patterns.

4.1.3.3 Students across grades cannot differentiate between words having ए/ऐ matras. They understand ए only (Regional effect)

In this question, students were to select the first syllable to make the word सफेद. Difference of 10% between सफेद and सफ्द was observed in the attempt performances. There was no pattern observed in grade wise performance. Grade 2 scored quite close to grade 5 with 38% and 40%, respectively. Grade 3 scored 33% and grade 4 performed least with 29%.
In this question, students were to select the first syllable to make the word पैदल. Most students selected पेदल (around 37%), instead of पैदल (14%). Around 25% selected पदल. Performance on this question is really very low across grades. The main reason for this could be the impact of dialect on one’s language acquisition. पैदल is often pronounced as पेदल and that might have led students to select the wrong option.

4.1.4 Sub-skill 4 – Identifying Syllable Blends

4.1.4.1 Overall low performance across grades where words to be formed using given set of syllables.

In this question, students were supposed to make the word, धूष from the given grid of syllables. The word was spoken on the voice over, which they could hear as many times by clicking on the sound icon above the grid. This question requires good hold of spellings. Only 23% of students were able to spell the word correctly even grade wise students performed with 37% accuracy on this question. This indicated that spelling is a weak skill across grades which further involves strong hold on matras and their corresponding sounds.

In this question, students were supposed to make the word, मेदान from the given grid of syllables. The word was spoken on the voice over, which they could hear as many times by clicking on the sound icon above the grid. This question requires good hold of spellings. The average performance on this question is only 11%, with all grades performing 20% or less (grade 2 – 8%, grade 3 12%, grade 4 – 6% & grade 5 – 20%). This indicated that spelling is a weak skill across grades which further involves strong hold on matras and their corresponding sounds.
In this question, students were supposed to make the word, सिकलोिा from the given grid of syllables. The word was spoken on the voice over, which they could hear as many times by clicking on the sound icon above the grid. This question requires good hold of spellings. The average score on this question is 22% with grade 2, 3, 4 & 5 scoring 20%, 24%, 18% & 27%, respectively. This indicated that spelling is a weak skill across grades which further involves strong hold on matras and their corresponding sounds.

4.1.4.2 Students across grades have performed well where they had to select a word after listening it on voice over.

In this question students were to identify the correct sound of the word shown on the screen, प्रकाश. The % ratio of students selecting correct sound vs the wrong sound (प्रदशमि) is 70: 30, even grade 2 students have scored more than 50% on this question. Here, the students also had an advantage of both the words being of different length. Students were able to identify the three letter word from the sound.

In this question students were to identify the correct sound of the word shown on the screen, गिशूल. Students from all the three language level buckets beginner, intermediate and proficient have scored well on this question. This could also be because all the three words has letters like गि which can confuse children!
In this question students were to identify the correct sound of the word shown on the screen, ‘गिर्मल’. Average score on this question is 60%, both the other options निर्जल and निकल were selected by around 20% of students, each. Students allotted beginner-level have scored lowest (17%) on this question. Intermediate scored around 50% and Proficient scored really well with 94%.

4.2 Vocabulary

4.2.1 Sub-skill 5 – Basic Vocabulary

4.2.1.1 Students across grades have issue in spelling basic sight words.

The average performance on this question is 33%. Though the highest response was there for the correct answer माला, performance percentages for माल and मला were also high. Grade 2 and 3 have performed lowest on this question. This indicates that the basic knowledge on the sight words is weak amongst students and hence when they are given similar looking options, they get confused.

4.2.1.2 Students across grades have scored well on questions involving general knowledge
In this question, students had to listen to the question stem and choose the correct answer. Students across grades have performed well on the questions where basic vocabulary was getting tested in a given context. On the question, the grade wise performance was seen to be increasing as from the grades advanced (grade 2 – 72%, grade 3 – 86%, grade 4 – 90% & grade 5 – 91%).

**Explanation:** In this question, students had to listen to the question stem and choose the correct answer. Students across grades have performed well on the questions where basic vocabulary was getting tested in a given context. Though students which were allotted beginners and intermediate levels struggled at this question, but overall the performance result is seen to be good. Surprisingly, students of grade 2 have performed better than students of grade 3, with 66% and 53% respectively. Students in grade 5 have performed better than grades 4 (79%), 3 & 2 with 85% accuracy.

4.2.2 Sub-skill 6 – Word Meaning

4.2.2.1 Students across grades have poorly performance on match the following questions, testing for antonyms.
In this question, students were expected to read the words written in the balls above and pair each one of them with the words written on the stone beneath. The pairing was to be done for the appropriate antonym. The average performance on this question was 20% with grades 2, 3 & 4 all performing at 18% accuracy and grade 5 at 33%. This indicates that even at higher grades antonyms of the basic levels aren’t practiced enough. Hence the students were not able to do the simple pairing without contextual understanding as well. Also, students of lower grades would have faced issues with reading and understanding the meaning of the words like गंदा, कड़चा, अधिक. The way balls were to be dragged and dropped on to the circled position, could also been a factor for some students’ low performance on this question.

In this question, students were expected to read the words written in the balls above and pair each one of them with the words written on the stone beneath. The pairing was to be done for the appropriate antonym. The average performance on this question was 27% with grades 2, 3 & 4 performing at 26%, 27% and 25% accuracy, respectively and grade 5 performing highest at 33%. This indicates that even in higher grades, the antonyms of the basic levels aren’t practiced enough. Hence the students were not able to do the simple pairing without contextual understanding as well. The way balls were to be dragged and dropped on to the circled position, could also been a factor for some students’ low performance on this question.
In this question, students were expected to read the word written in the square on the left, change the word in the box on the right by pressing the arrows multiple times and then selecting the appropriate synonym to match the pair. The average performance on this question is 27% with grades 2, 3 & 4 performing at 28%, 27% and 20% accuracy, respectively and grade 5 performing highest at 33%.

From this question, we can say that children have not mastered the meaning which is a rote learning and reinforcement practice. They are finding it difficult to come up with the correct synonym even when the context is not needed. The students might have found the manoeuvring of arrow slightly tricky which also could be one reason for the low performance.

### 4.2.2.2 Students across grades have misconceptions regarding meaning of specific words

In this question, students were to select the right word to complete the right sentence. The correct answer र्ीठा has been selected by 41% students, but 32% selected कड़वा option, this indicated 32% students thinks that लड्डू is bitter in taste or they didn’t know the meaning of the word कड़वा.

Another reason of only 41% students attempting this question correctly could be that students weren’t able to read the लड्डू, as the sentence was not spoken on the voice over nor were the options. Grade wise performance has been 33% for grade 2, 43% for grade 3, 50% but surprisingly grade 5 has performed lower than grade 4 with 47% and 50%, respectively.
In this question, students were to choose the option which is of same meaning as the phrase written above on the screen. The average performance on this question is 24%. The percentage attempt on different options has shown that children were highly confused among the options and that they clearly didn’t know the meaning of this common phrase. The correct answer आराम से was selected by the lowest number of students (24%), 31% selected तेज़ी से and the maximum percentage of students selected ज़ोर से (44%). This misunderstanding or misconception with meaning of the phrase is across grades, all the grades have performed less than 30% on this question with grade performing the lowest at 18% and surprisingly, grade 2 performing the highest at 28%. One of the possible reasons could be that students might have just read a part of the phrase and might have selected the first option as that also indicates speed.

In this question, students were to select the option with the correct meaning of the word प्रसिद्ध (जाना-माना). The average performance on this question is 28%. It also had a low PBC value, which indicates that the best performing kids haven’t performed on this question. All the three options got around 30% attempts. This shows that students struggle with meanings of some common words as well and even when it is used in very familiar context. One of the reasons could be students incapacity to read प्रसिद्ध as it is difficult to decode. Grade 3 students have performed low on this question with 18% and highest by grade 5 with 50% accuracy.

4.2.3 Sub-Skill 7 – Usage
4.2.3.1 Students in the beginner and intermediate language levels have struggled with questions on word usage

In this question, students were to complete the above written sentence by picking up the appropriate word. Here, sentence comprehension, understanding of basic rules of grammar were also needed for a child to pick the relevant correct answer. The average performance on this question was observed at 50%. Students at the beginner level have performed with only 17% of accuracy and those in intermediate with 39% accuracy. The two incorrect options एक and सारा, got around 25% each response. Children from grade 2 & 3 both have performed with 35% of accuracy, grade 4 with 60% and grade 5 at highest with 68%. Using an appropriate word to complete a sentence, considering the overall context is a weak skill for students in the buckets beginners and intermediate. This might be getting limited due to addition need for sentence comprehension.

4.2.3.2 Students in the proficient language level have struggled with phrase usage

In this question, students were to complete the above written sentence. The average performance was seen to be 47%. Around 34% students selected छोटपटा उठा, this indicates that students across grades and language levels (as students in proficient also performed low on this question) require practice the usage of the verbs in the proper contextual format. Grade 3 performed lowest with 30% and grade 4 & 5 performed almost the same with 50% and 51%, respectively.
In this question, children had to complete the sentence written above by understanding the context given. Average performance for this question was observed to be around 49%. The PBC value of this question is -0.09, which is reflecting that even the high performing students found the question very difficult. Both the options have received 50% response. Grade 2 has understandably performed lowest on this question with 40% and grade 4 the highest with 41%. The question required finer mastery of the skill of sentence comprehension which many students seem to be struggling with.

In the question, students had to determine the correct usage of the phrase 'हृदय का पिघलना'. This question has received low performance with only 33% attempt on the correct answer (बुढे भिखारी की हालत देखकर सेठ का हृदय पिघल गया।) Most Common wrong answer for this question is आइसक्रीम खाते-खाते सेठ का हृदय पिघल गया।, which was selected by around 28%.

Grade 5 scored the best on the question with 45% and at lowest is grade 4 with only 21% accuracy. This is indicative of limited or weak understanding of the idiom usage in children.

4.2.4 Sub-skill 8 – Recall

4.2.4.1 Students in Beginners and Intermediate language levels struggled in recalling sentences. They could not differentiate between options with slight difference.
have scored 12%, those is the Intermediate, 26% and Proficient 60%, respectively. Grade 2 and 3 scored 22% each, grade 4, 46% and grade 5, 52%. The most common wrong answer was 'बरसात हो रही है' which was selected by almost 50% students. Students at the Beginner levels were found to be weak at the questions involving usage of working memory, with only 12% performance.

In the questions, students had to read (no VO here for the sentence) ‘नगरभर में मुरलीवाले के आने का समाचार फैल गया।' and select the same sentence from the given options. Average performance on this question is 25% and PBC value is 0.29. Overall performance in all the three categories and in the individual grades is low with students in beginner level scoring 12%, Intermediate 17% and Proficient scored 42%. Performance is increasing as we move from grade 2 to grade 5. The most common wrong answer was 'नगरभर में गुब्बारेरावले के आने का समाचार फैल गया।' which was selected by almost 34% children. Around 22% selected सारे शहर में मुरलीवाले के आने का समाचार फैल गया.

The sentence was slightly longer to memorize and unless children could comprehend and then visualise, it was difficult for them to arrive at the correct answer. Here, though these are recall questions, sentence comprehension is playing a subtle role and hence the performance is seen to be very low.

4.2.4.2 Students across grades & language levels performed low at questions with listening & recalling a longer sentence
In this question, students had to listen to the sentence spoken on the voice over and select it amongst the given options. Average performance on this question was 32% with low PBC value of 0.24. The performance on the correct answer was just 6.82%. Around 22.7% of the children selected 'आम में फलों का राजा कहा जाता है'. Even when seen grade wise, the performance is low across grades with grade 3 scoring 36% and grade 5 scoring only 20%. This clearly indicates that as the length of the sentence increases, the working memory isn’t able to recall the whole sentence even with a gap of only few seconds. Also, this isn’t only an issue with the students weak in language skills but also those who are best performers on this test, as indicated by low PBC value.

In this question, students had to listen to the sentence spoken on the voice over and select it amongst the given options. Average performance on this question was 30% with low PBC value of 0.32. Grade 5 has performed highest on this question with 40% and lowest by grade 4 with 22%. This clearly indicates that as the length of the sentence increases, the working memory isn’t able to recall the whole sentence even with a gap of only few seconds. Also, this isn’t only an issue with the students weak in language skills but also those who are best performers on this test, as indicated by low PBC value.

4.2.5 Sub-skill 9 – Sentence Structure Knowledge
4.2.5.1 Students in Beginners and Intermediate language level struggled in making meaningful sentences from given jumbled words

In this question, students had to make a meaningful sentence (‘पानी में रंग-बिरंगी मछलियाँ हैं’) using the set of jumbled words provided on the screen. Average performance on this question was 18%. Students in Beginner and Intermediate levels have poorly performed with as low as 0% and 5% respectively. Though, grade wise the performance was seen to be increasing, the overall performance was seen to be quite low with grade 2 students at 7%, grade 3 students at 13%, grades 4 & 5 at 22% and 29% respectively. This shows that syntactic and semantic knowledge of students across grades and language levels is extremely weak. This is in spite the fact that full stop is already put against the last word of the sentence as a hint.

There are limited activities for children which involve puzzles which helps in developing various cognitive language skills.

In this question, students had to make a meaningful sentence (‘नील की पत्तियाँ कड़वी लगती हैं’) using the set of jumbled words provided on the screen. Average performance on this question is 27%. Students in Beginner and Intermediate levels seem to be struggling a lot with performance as low as 2% and 6% respectively. Though, grade wise performance is increasing when we move ahead in grades, but performance percentages are quite low. Grade 2 students scored 6%, grade 3 students have scored 15%, grade 4 & 5, 34% and 52% respectively. This shows that syntactic and semantic knowledge of students across grades and language levels is extremely weak. This is in spite the fact that full stop is already put against the last word of the sentence as a hint for students. To be able to do this question correctly, students not only need the understanding of each word, but also require to understand the grammar rules to construct a
sentence and arrange the words in the correct order. This itself is a higher order skill.

4.2.6 Sub-skill 10 – Punctuation Knowledge

4.2.6.1 Students across grades by default selects ‘?’ or ‘|’, due to lack of concept understanding

In this question, students were to select the correct punctuation mark in accordance to the sentence written above. Average performance on the question is 51%. Students at Beginners level have scored 48%, students in Intermediate level scored 50% and those in Proficient level scored 55%. But the overall PBC is quite low of 0.07, which indicated that the best performing students on this test have not attempted this question properly. The grade wise score was also around 50% throughout. Around 27% selected ‘|’ and around 22% selected ‘;’ as their answer. This shows if students do not have the hold on the concept of the punctuation marks, they are selecting full stop also as the blank in at the end of the sentence.

In this question, students were to select the correct punctuation mark in accordance to the sentence written above. Average performance on the question was 40%. Students at Beginners level have scored 48%, students in Intermediate level scored 50% and those in Proficient level scored 55%. But the overall PBC is low of 0.30. 46% of the students have selected ‘?’, around 40% selected ‘|’ and around 14% selected ‘,’. This indicates that there is a gap in the conceptual understanding of punctuation and its correct usage. Grade 5 students, as expected, have performed the highest with 51% and grade 3 has performed lowest with 25% accuracy.
4.2.6.2 Students across grades are putting punctuation mark ‘?’ right after a question word (क्यों, कहा, कितने etc.) in the sentence.

In this question, students were to select the correct punctuation mark in accordance to the sentence written above. Only 19% could select the correct answer. The PBC was -0.07, which indicated that students having higher overall performance have also found this question difficult. 47% students selected question mark ‘?’ and 34% students selected ‘|’. This is clearly showing that students have extremely low clarity on punctuation concept, just because they read the word कितने in the sentence, they assumed the right answer is a question mark but the context of the sentence wasn’t considered at all.

In this question, students were to select correctly punctuated sentence from the options given. The average performance on this question is 52% and low PBC value of 0.39, which indicated that the best performing students on this test have not attempted this question properly. Around 27% selected क्या, मैं जो मागूँगा? वह आप देंगे. 21% student selected the option which had ‘?’ right after the word ‘क्या’. This is clearly showing that students have extremely low clarity on punctuation concept and they often decide the punctuations by identifying certain words like क्या in this question. Students are not able to comprehend the sentence and the intonation and hence are going merely by the presence of certain indicative words.
In this question, students were to select correctly punctuated sentence from the options given. The average performance on this question is 47% and low PBC value of 0.43, which indicated that the best performing students on this test have been able to do the question. This is also because this is a simpler sentence and doesn’t have context that can change the punctuation. 30% student selected the option which had '?' right after the word ‘क्या’ and around 23% selected the option with '?' right after ‘मनहर’. This is clearly showing that students have extremely low clarity on punctuation concept, just because they read the word क्या in the sentence and question mark put right after it, they assumed that’s the right answer as the placement doesn’t matter.

4.2.7 Sub-skill 11 – Grammar Knowledge

4.2.7.1 Students in the beginner language level are struggling with questions involving verb form related questions more than the other categories.

In this question, students were to select the right verb tense out of the given options, to complete the sentence. Average performance on this question is 58%. Students in Beginners level seem to be struggling a bit, with 34% performance, students in Intermediate level scored 48% and those in Proficient level, performed highest with 83%. Around 22% students selected ‘आना है’ and 20% selected ‘आएगी’. They didn't pay attention that they had to select a present tense option. The wrong choice of options could be possible due to incomplete understanding of the verb form or due to tendency to hurriedly respond.
In this question, students were to select the right verb tense out of the given options, to complete the sentence. The average performance on this question is 45%. Students in Beginner and intermediate level have performed with only 12% and 32%, accuracy, which is expected as kids in this bracket have low syntactic knowledge. Students in Proficient level have performed with 79% accuracy. Around 45% students selected the correct answer ‘सूझा’, 34% selected ‘सूझिा’ and around 22% selected ‘सूझेंिे’, this is showing that there is confusion between ‘सूझा’ and ‘सूझिा’ amongst kids indicating lower understanding of sentence semantics. The word चिन्मय itself is difficult and unfamiliarity of this word can be another reason for low performance.

In this question, students were to select the right verb tense out of the given options, to complete the sentence. The average performance on this question is 33% and low PBC value of 0.17. Students in Beginner and intermediate level have performed with only 21% and 31%, accuracy, which is expected as kids in this bracket have low syntactic knowledge. Students in Proficient level have performed with 42% accuracy. Grade wise, none of the grades scored more than 40% and all the three options- बांधती, बांधा, बांधेंिी got 33%, which clearly suggests that students have very limited semantic knowledge to appropriately use a given verb tense in the given context.
4.2.7.2  Students across grades & language level have confusion on the usage of है and हैं

In this question, students had to complete the short sentences with given set of words. The average performance on this question is 51% and PBC value is quite low at 0.33. Only around 16% actually filled up all the correct answers for the blanks (i.e. लड़का खेलता है | लड़के खेलते हैं | कमरा सुंदर है | कमरे सुंदर हैं). Around 57% students answered in the following incorrect manner - लड़का खेलता है | लड़के खेलते हैं | कमरा सुंदर है | कमरे सुंदर है | Students do not understand the concept of using है / हैं where है is used in a plural context. Such concepts are highly important for a child to comprehend the whole story line.

4.2.7.3  Students across grades and language level have struggled with question on Pronoun

In this question, student is getting tested on for the knowledge of pronoun (सर्वनाम). The average performance on this question is 59% with low PBC value of 0.57. Around 38% students filled the blanks with all the correct answers. This shows that students struggle when they have to build stories by selecting the correct pronouns so that the story retains its meaning. Teaching these things under specific grammar headings is not necessary but explaining through examples that where and how a noun is replaced with pronoun, can be done while narrating a story as well. Encouraging reading and story listening can also help in creating this comprehension in children.

4.3  Skill – Reading Comprehension
4.3.1 Sub-skill 12 – Passage Comprehension

4.3.1.1 Students in the Beginner language level are struggling in the questions with no voice over help

In this question, students had to answer a series of ‘Wh questions’ to reconstruct the sentence as written above. The average performance on this question is 62% with low PBC value of 0.61. The students in Beginner and intermediate level have scored 41%, and 53%, respectively. Those in Proficient level have scored highest with 85%. The performance on this question increased as we moved from grade 2 to 5. Though the interesting point to consider here is that, in the option-wise result, it was observed that only 30% students gave all the correct answers for all the asked questions, rest of them had at least one of their answers wrong. This indicates that factual retrieval is a weak skill amongst tested students.

In this question, students had to decode the picture to choose the corresponding sentence. There was no voice over for the options. The average score on this question is 51% with low PBC value 0.14. Around 49% selected the wrong option i.e. the sentence नाखून, this shows that they are selecting the answers randomly without any understanding.
In the question, students were to search the question for the given sentence 'कबूतर हमेशा समूह में रहते हैं'. The average score on this question is quite low at 29% with PBC value in negative, -0.10. This clearly suggests that the best performing students on this test have not been able to attempt the question well. Most of the children (around 38%) selected ‘समूह हमेशा कहाँ रहते हैं?’ and only around 28% selected the correct answer ‘कौन हमेशा समूह में रहते हैं?’. The skill of making a question or asking a question is a neglected skill and not enough focus or practice is given on this order of learning. Students can be made to practice or can be encouraged on making their own questions from the given which not only helps them understand the text better also express the areas of doubts.

In this question, students were to read the given passage to answer the questions followed. The average performance on this question is 27% with PBC value as low as 0.21. This clearly suggests that the best performing students on this test have not been able to attempt this question well. All of the correct answers were chosen by just 0.57% students. Most of them selected 'पानी की एक ही बूंद रही हुई' instead of 'पानी की एक ही बूंद रही'. The ability to answer this question is also dependent on the ability to comprehend the given text.