Evaluation of Foundational Learning Programs in Uttar Pradesh

Midline Round – Topline Findings

Dec 2023 - Mar 2024

Language and Learning foundation
Strong Foundation, Stronger Future

CENTRAL SQUARE FOUNDATION

Educational Initiatives is now Ei

Executive Summary

1. A large positive impact was seen on NIPUN-focused tasks (ORF) in Cohort 1 for the High-Touch Demo group.

2. Similar to Cohort 1, NIPUN-focused tasks (ORF and Subtraction) showed a higher difference between the performance of the demo and non-demo groups in Cohort 2.

3. A medium positive impact was seen on NIPUN-focused tasks (Subtraction) in Cohort 1 for the High-Touch Demo group.

4. A small positive shift was seen from baseline in key FLN practices like highlighting the sound of a letter/ matra, showing strokes of the letter/ matra, asking open and close-ended questions, introducing new vocabulary, etc.

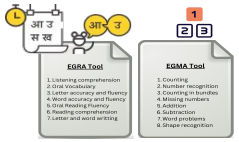
5. While asking CFU questions as a practice has improved slightly, the practice of giving clear instructions show a slight decline.

Overview of Evaluation Design

The study has a quasi-experimental design, with demonstration (demo) and non-demonstration (non-demo) sites matched based on similar characteristics, and covered 3,190 Grade 1 students and 3,192 Grade 2 students from 327 schools in the midline round

Demonstration Sites
High Touch: Sewapuri
Low Touch: Rest of the Varanasi

Non-Demonstration Sites: Siddharth Nagar, Unnao, Mirzapur and Kushinagar



Contextualized EGRA (Literacy) & EGMA (Math) Tools

57 Enumerators, Supervisors, District Coordinators

MDES of 0.11

80% Power

Tangerine app for data collection

95% Confidence Interval, with 5% Margin of Error

Round	Group	Grade 1	Grade 2	Grade 3
Baseline (19th Sep - 22nd Oct 2022)	Demo-High Touch	100 Schools 1,084 Students		
	Dem-Low Touch	109 1,006 Students		
	Non-Demo	99 Schools 1,152 Students		
Midline (06th Feb - 13th Mar 2024)	Demo-High Touch	107 Schools 1,055 Students	104 Schools 1,093 Students	
	Demo-Low Touch	103 Schools 1,056 Students	102 Schools 1,066 Students	
	Non-Demo	115 Schools 1,079 Students	108 Schools 1,033 Students	
Endline (Dec 2024)	Demo-High Touch			
	Dem-Low Touch			
	Non-Demo			

Cohort 1 | Cohort 2

Additionally, this round of the evaluation also included a follow-up qualitative study, which focused on 15 teachers from demo districts whose classrooms had been observed during the baseline round of process evaluation, along with Academic Resource Persons (ARPs) and on-ground LLF members

Method	Sample Size	Sampling Methodology
Teacher Survey	218*	Sent via WhatsApp to all primary school teachers (whose phone numbers were available), across demo and non-demo districts, with responses considered only from teachers teaching FLN grades
In-depth interviews (IDIs) with Implementation partners	6	As suggested by RTR, from the list of partners working in the demonstration districts
Classroom Observations (COs)	15	Purposively selected from the pool of teachers whose classrooms were observed in the baseline round of the process evaluation, based on the observed levels of implementation fidelity (high, medium, low)
Document review + In-depth interviews (IDIs) with teachers	15	Conducted with the teachers whose classrooms were observed during this follow-up study
Focus Group Discussion (FGD) with teachers	1	Purposively selected based on the COs & teacher IDIs conducted during this follow-up study
Joint visits + Document review + In-depth interviews with ARPs	9	Purposively selected based on the availability of the ARPs, and the COs & teacher IDIs conducted during this follow-up study

The study was conducted in only one demo site - Rest of Varanasi. The plan was to cover a mix of G1 (5/15) and G2 (10/15) classes, which was possible in literacy. However, in numeracy, we had to observe an additional G2 class in place of G1, due to significant teacher shortage on the day of the visit.

**166 of these are from high-touch, 25 from low-touch, and 27 from non-demo sites*

Key Findings for Cohort 1

Grade 1 in the Baseline Round to Grade 2 in the Midline Round

The intervention had a significant impact on student performance in literacy tasks such as Non-Word Reading (Fluency), ORF, and Reading Comprehension, across both the High-Touch and Low-Touch sites, with a greater effect size in Sewapuri across all literacy tasks

Task	Unit	Midline Average			Baseline Average			DiD Effect Size		Delta	Delta	Delta
		ND	D-HT	D-LT	ND	D-HT	D-LT	D-HT	D-LT	ND	D-HT	D-LT
Listening Comprehension	Percentage	77%	81%	77%	62%	59%	62%	0.24*	0.01	15%	22%	15%
Oral Vocabulary	Percentage	94%	96%	95%	95%	96%	96%	0.14*	-0.01	-1%	0%	-2%
Initial Sound Identification	Percentage	68%	82%	69%	12%	16%	19%	0.29*	-0.15*	55%	66%	50%
Letter Reading (Accuracy)	Percentage	77%	91%	85%	29%	35%	42%	0.28*	-0.15*	48%	56%	44%
Letter Reading (Fluency)	Count per minute	51.6	70.0	62.1	14.8	16.9	20.1	**0.85***	0.25*	36.8	53.2	42.0
Word Reading (Accuracy)	Percentage	70%	90%	83%	8%	13%	14%	**0.69***	0.29*	62%	78%	69%
Word Reading (Fluency)	Count per minute	21.6	34.9	29.3	5.1	6.2	7.5	**0.96***	0.40*	16.5	28.7	21.8
Non-Word Reading (Fluency)	Count per minute	19.9	33.2	27.8	2.7	3.5	4.5	**1.19***	**0.54***	17.2	29.7	23.4
Oral Reading Fluency	Count per minute	34.8	61.1	50.4	2.6	3.0	4.0	**1.22***	**0.63***	32.3	58.2	46.4
Reading Comprehension Passage 1	Percentage	60%	88%	78%	2%	4%	4%	**0.97***	**0.52***	57%	84%	73%
Letter Writing	Percentage	71%	80%	71%	22%	25%	32%	0.21*	-0.30*	48%	55%	39%
Word Writing	Percentage	58%	77%	71%	8%	9%	11%	**0.73***	0.37*	50%	68%	59%

The DiD effect size was calculated based on: ***[avg_delta_demo (Δi) - avg_delta_non-demo (Δc)] / SD_pooled (Pooled Standard Deviation)***

*For all tasks, the statistical significance of the difference was determined through Welch's unpaired t-test assuming unequal variance t-test . For t-test, one doesn't reject the null hypothesis if p-value is less than 0.05. *represents that the difference between means is significant.*

The High-Touch intervention model had a reasonably large impact on student performance across all numeracy tasks in Sewapuri. In the Low-Touch districts, the intervention had a small impact on student performance across most numeracy tasks, with the exception of Number Recognition (Fluency), where it had a reasonably large impact

Task	Unit	Midline Average			Baseline Average			DiD Effect Size		Delta ND	Delta D-HT	Delta D-LT
		ND	D-HT	D-LT	ND	D-HT	D-LT	D-HT	D-LT			
Number Recognition (Fluency)	Count per minute	22.8	29.9	28.5	12.4	11.4	13.0	**0.58**	0.37	10.5	18.5	15.4
Number Recognition (Accuracy)	Percentage	73%	83%	79%	26%	28%	31%	0.34	0.03	47%	55%	48%
Counting in Bundles	Percentage	63%	77%	67%	18%	16%	18%	**0.52***	0.11*	45%	60%	49%
Missing Number	Percentage	40%	51%	43%	14%	13%	16%	**0.50***	0.04*	26%	37%	27%
Addition (Accuracy)	Percentage	77%	91%	85%	33%	33%	41%	0.40*	0.02*	44%	58%	44%
Subtraction (Accuracy)	Percentage	67%	85%	74%	20%	18%	24%	**0.59***	0.08	47%	67%	50%
Word Problems	Percentage	56%	69%	60%	30%	25%	31%	**0.61***	0.11*	26%	44%	29%

Number Comparison and Shape recognition task was not reported in the baseline due to incorrect administration of this task.

The DiD effect size was calculated based on: ***[avg_delta_demo (Δi) - avg_delta_non-demo (Δc)] / SD_pooled (Pooled Standard Deviation)***

*For all tasks, the statistical significance of the difference was determined through Welch's unpaired t-test assuming unequal variance t-test . For t-test, one doesn't reject the null hypothesis if p-value is less than 0.05. *represents that the difference between means is significant.*

Detailed Findings in Literacy for Cohort 1

Grade 1 in the Baseline Round to Grade 2 in the Midline Round

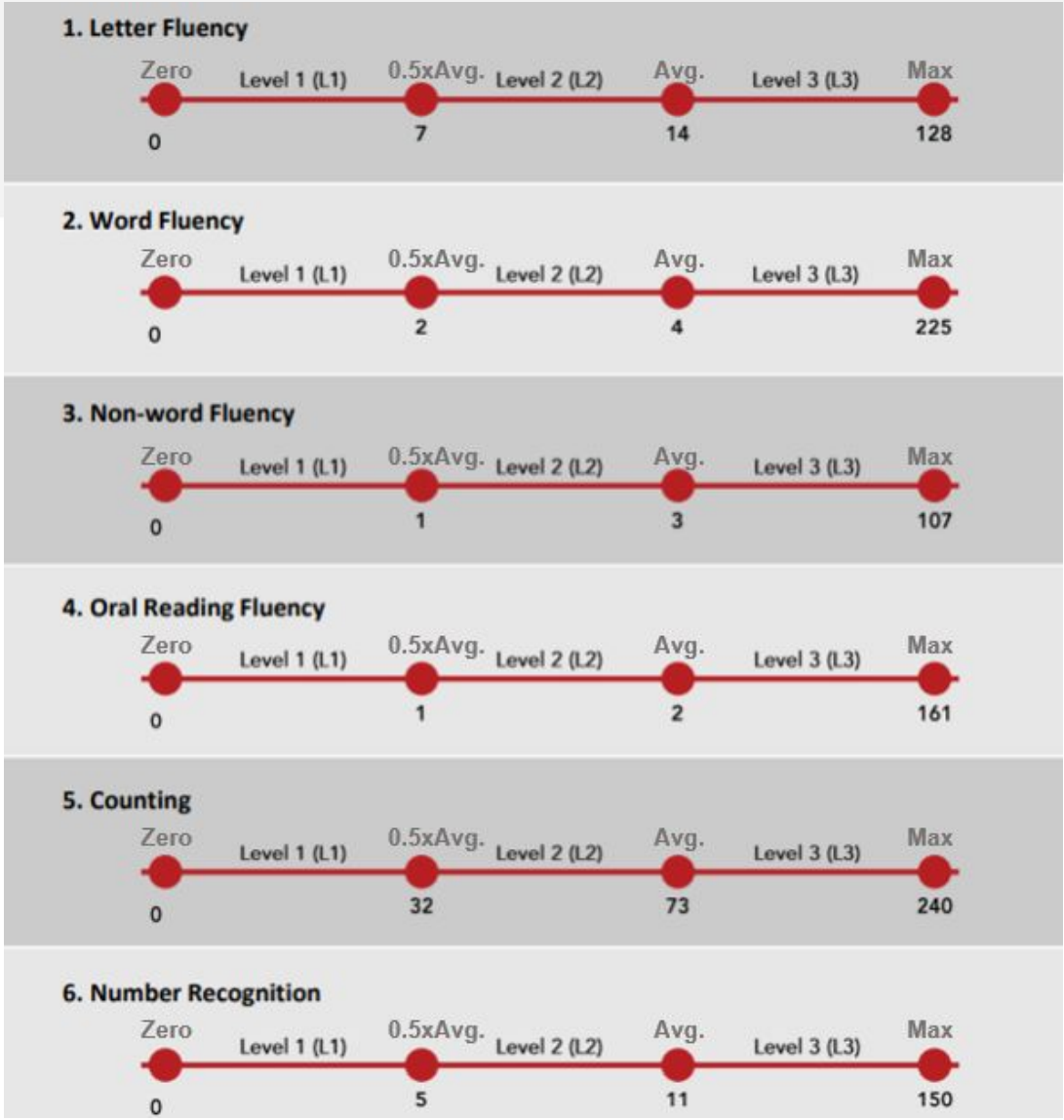
The student performance was also categorized into differentiated performance bands to provide detailed insights and analysis

Based on the student scores for cohort 1 in each task in the baseline round, 4 performance levels (L0 - L3) were created, with different calculations for timed and untimed tasks. The same performance levels have been used to represent the results of cohort 2 the midline round, for comparison.

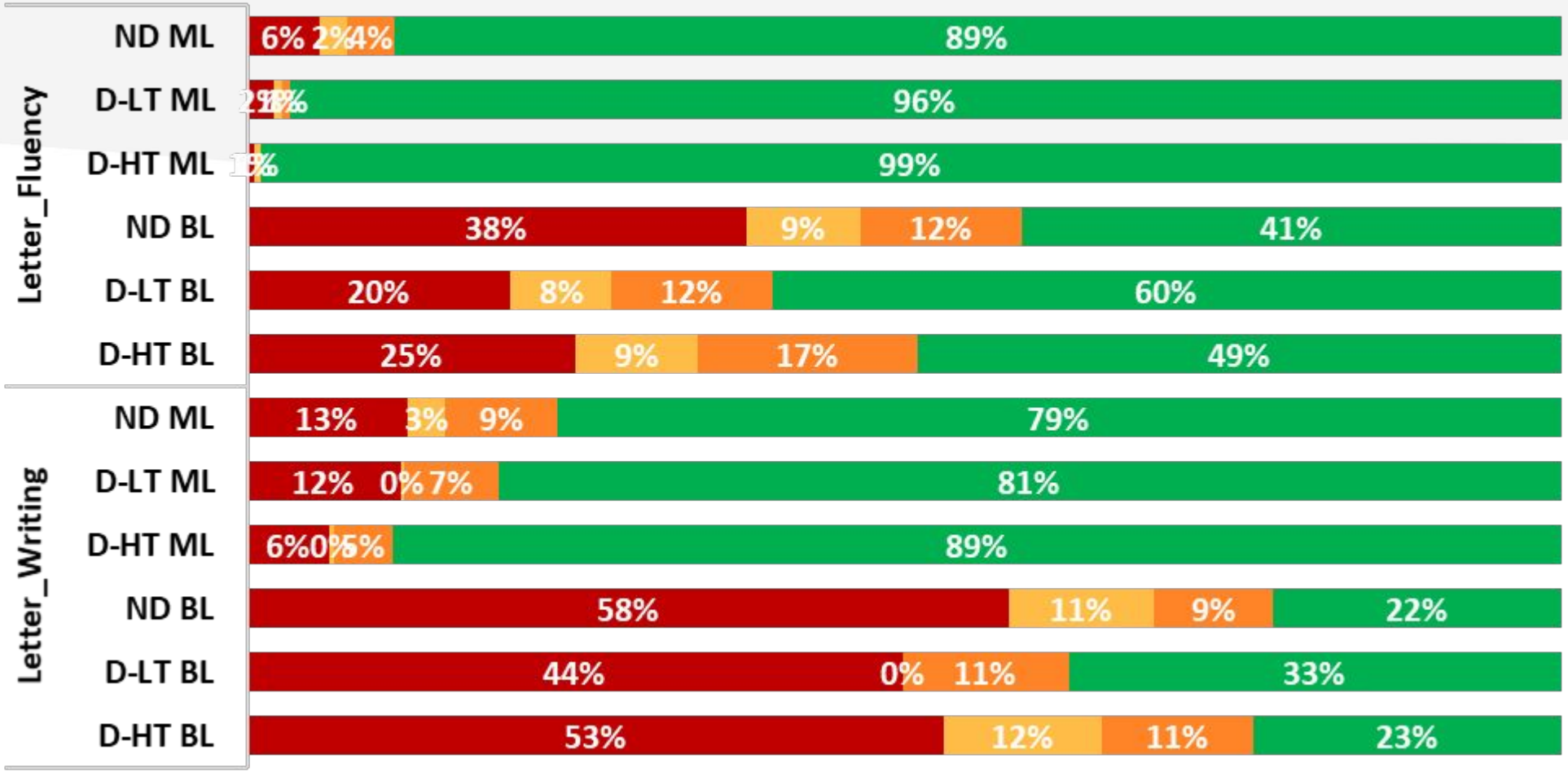
1. Untimed Tasks: Score = (no. of correct responses / total no. of items) * 100 = Accuracy Percentage

2. Timed Tasks: Score = no. of correct student responses per minute (e.g. no. of letters or words read per minute)

Level	Untimed Tasks	Timed Tasks
Level 0 (L0)	0%	0
Level 1 (L1)	> 0% – 25%	> 0 to (0.5 * Average)
Level 2 (L2)	> 25% – 50%	> (0.5 * Average) to Average
Level 3 (L3)	> 50% – 100%	> Average to Maximum

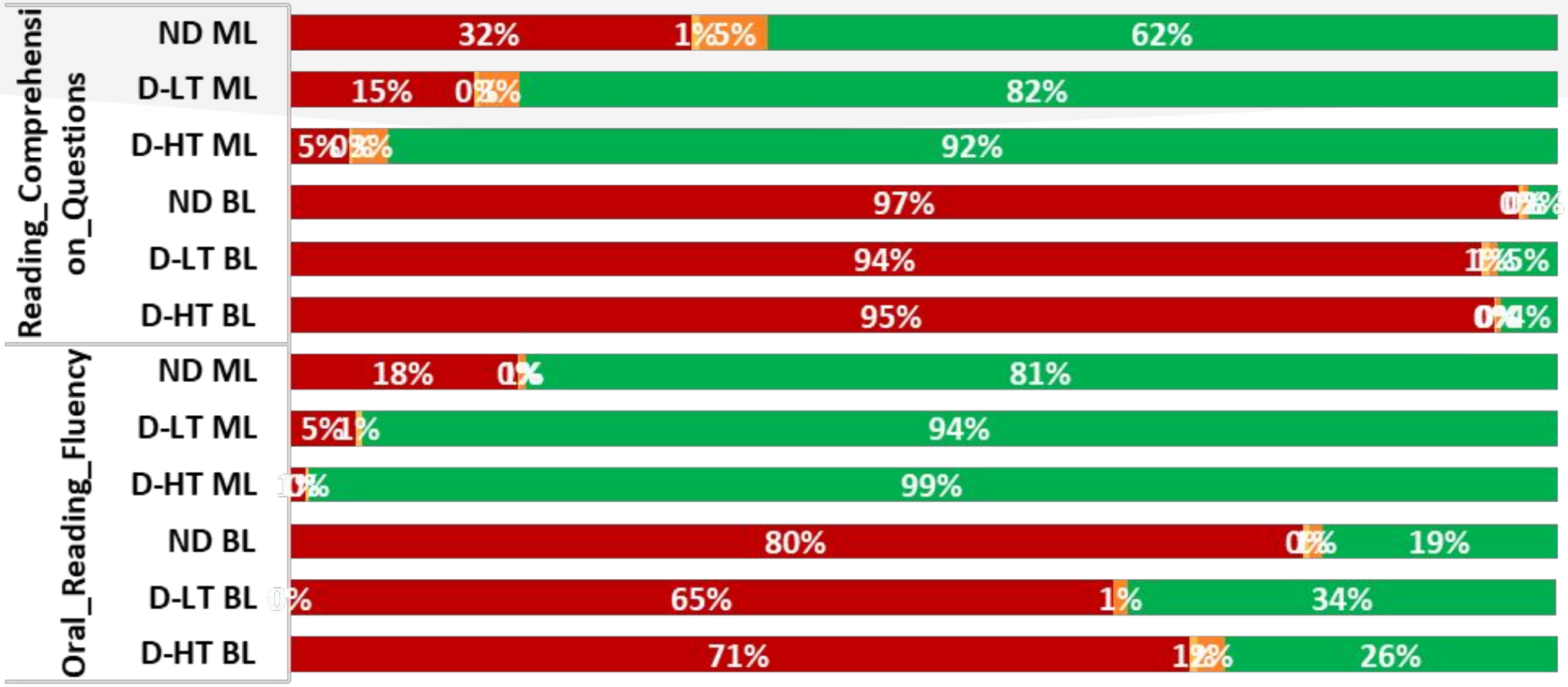


In the Letter Writing task, there has been a significant improvement from the baseline. In the Letter Naming Fluency task, more than 85% of the students have moved to highest learning level in all the groups



	L0	L1	L2	L3
Accuracy Tasks	0%	> 0% – 25%	> 25% – 50%	50% – 100%
Fluency Tasks	0	> 0 to (0.5 * Avg.)	> (0.5 * Avg.) to Avg.	> Avg. to Max.

Higher performance on Reading Comprehension and ORF indicates students' improved skills in reading a text with an understanding



	L0	L1	L2	L3
Accuracy Tasks	0%	> 0% – 25%	> 25% – 50%	50% – 100%
Fluency Tasks	0	> 0 to (0.5 * Avg.)	> (0.5 * Avg.) to Avg.	> Avg. to Max.

Apart from Listening Comprehension, Oral Vocabulary, and Initial Sound Identification, schools in the high-touch demo group maintained the performance patterns seen in the baseline round, though the top and bottom performing schools scored closer to the average in the midline round (1/2)

D-HT

Task	BL Category	# Schools	# Students_BL	BL_Avg	# Students_ML	ML_Avg
Listening Comprehension	Top 10%	10	123	73%	115	82%
	Mid 10%	11	131	55%	108	80%
	Bottom 10%	10	112	48%	93	76%
Oral Vocabulary	Top 10%	10	123	99%	115	96%
	Mid 10%	11	131	96%	108	96%
	Bottom 10%	10	112	92%	93	96%
Initial Sound Identification	Top 10%	10	123	37%	115	84%
	Mid 10%	11	131	11%	108	87%
	Bottom 10%	10	112	5%	93	71%
Letter Reading Fluency	Top 10%	10	123	30.1	115	75.6
	Mid 10%	11	131	17.5	108	69.8
	Bottom 10%	10	112	6.3	93	65.7
Letter Reading Accuracy	Top 10%	10	123	55%	115	94%
	Mid 10%	11	131	36%	108	94%
	Bottom 10%	10	112	13%	93	90%
Word Reading Fluency	Top 10%	10	123	13.1	115	39.3
	Mid 10%	11	131	6.5	108	35.8
	Bottom 10%	10	112	0.6	93	31.8
Word Reading Accuracy	Top 10%	10	123	32%	115	95%
	Mid 10%	11	131	13%	108	92%
	Bottom 10%	10	112	1%	93	86%

BL Category	Avg Decile Movement*
Top 10%	-2.6
Mid 10%	0.6
Bottom 10%	3.4

- Schools in the bottom 10% performance category in the baseline round for literacy improved significantly in most tasks in the midline round, apart from Initial Sound Identification, Reading Comprehension, and Letter and Word Writing.

- The performance of most top performing schools in the baseline round for literacy dropped significantly in the midline round in all tasks except Word Reading Accuracy.

Mean school score is in the top 25% of school scores	Mean school score is in the second 25% of school scores	Mean school score is in the third 25% of school scores	Mean school score is in the bottom 25% of school scores

- Schools were ranked based on the mean school score in each task, i.e., the average of the scores of all students in the relevant grade in that school.
- An aggregate rank was created for each school across all literacy / numeracy tasks by adding the average ranks for each task, based on which the top 10%, middle 10%, and bottom 10% schools were selected.
- Average scores were calculated by taking the simple average of the mean school score in that task for all schools in that performance category.

** The average decile movement is the average change in the deciles of all the schools in each performance category, from the baseline to the midline round, with deciles determined based on the aggregate rank of the school.*

Apart from Listening Comprehension, Oral Vocabulary, and Initial Sound Identification, schools in the high-touch demo group maintained the performance patterns seen in the baseline round, though the top and bottom performing schools scored closer to the average in the midline round (2/2)

D-HT

Task	BL Category	# Schools	# Students_BL	BL_Avg	# Students_ML	ML_Avg
Non-Word Reading Fluency	Top 10%	10	123	8.5	115	37.7
	Mid 10%	11	131	3.6	108	33.7
	Bottom 10%	10	112	0.3	93	30.6
Oral Reading Fluency (ORF)	Top 10%	10	123	9.2	115	70.4
	Mid 10%	11	131	2.3	108	63.8
	Bottom 10%	10	112	0.2	93	53.1
Reading Comprehension Questions	Top 10%	10	123	17%	115	92%
	Mid 10%	11	131	3%	108	91%
	Bottom 10%	10	112	0%	93	79%
Letter Writing	Top 10%	10	123	50%	115	86%
	Mid 10%	11	131	20%	108	81%
	Bottom 10%	10	112	6%	93	71%
Word Writing	Top 10%	10	123	25%	115	83%
	Mid 10%	11	131	4%	108	78%
	Bottom 10%	10	112	0%	93	70%

- *Schools were ranked based on the mean school score in each task, i.e., the average of the scores of all students in the relevant grade in that school.*
- *An aggregate rank was created for each school across all literacy / numeracy tasks by adding the average ranks for each task, based on which the top 10%, middle 10%, and bottom 10% schools were selected.*
- *Average scores were calculated by taking the simple average of the mean school score in that task for all schools in that performance category.*

** The average decile movement is the average change in the deciles of all the schools in each performance category, from the baseline to the midline round, with deciles determined based on the aggregate rank of the school.*

BL Category	Avg Decile Movement*
Top 10%	-2.6
Mid 10%	0.6
Bottom 10%	3.4

- Schools in the bottom 10% performance category in the baseline round for literacy improved significantly in most tasks in the midline round, apart from Initial Sound Identification, Reading Comprehension, and Letter and Word Writing.

- The performance of most top performing schools in the baseline round for literacy dropped significantly in the midline round in all tasks except Word Reading Accuracy.

Mean school score is in the top 25% of school scores	Mean school score is in the second 25% of school scores	Mean school score is in the third 25% of school scores	Mean school score is in the bottom 25% of school scores

Decile Movement from Baseline to Midline for the Top 10%, Middle 10%, and Bottom 10% Schools at the Baseline, for the High-Touch Demo Group of Cohort 1, in Literacy

D-HT

Decile Movement from Baseline to Midline for the Top 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
1	1	3	-2
2	1	9	-8
3	1	3	-2
4	1	1	0
5	1	1	0
6	1	5	-4
7	1	5	-4
8	1	4	-3
9	1	2	-1
10	1	3	-2
Average	**1**	**3.6**	**-2.6**

Decile Movement from Baseline to Midline for the Middle 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
45	5	5	0
46	5	7	-2
47	5	-	-
48	5	8	-3
49	5	8	-3
50	6	1	5
51	6	9	-3
52	6	8	-2
53	6	2	4
54	6	1	5
55	6	1	5
Average	**5.5**	**5.0**	**0.6**

Decile Movement from Baseline to Midline for the Bottom 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
90	10	10	0
91	10	2	8
92	10	7	3
93	10	1	9
94	10	9	1
95	10	10	0
96	10	5	5
97	10	-	-
98	10	5	5
99	10	10	0
Average	**10**	**6.6**	**3.4**

No consistent performance patterns are seen at the school level across the baseline and midline round in the literacy tasks for the low-touch demo group; in fact, both top and bottom performing schools in the baseline round scored much closer to the average in the midline round (1/2)

D-LT

Task	BL Category	# Schools	# Students_BL	BL_Avg	# Students_ML	ML_Avg
Listening Comprehension	Top 10%	10	100	73%	113	77%
	Mid 10%	10	102	65%	120	78%
	Bottom 10%	10	100	49%	104	78%
Oral Vocabulary	Top 10%	10	100	99%	113	95%
	Mid 10%	10	102	98%	120	95%
	Bottom 10%	10	100	94%	104	96%
Initial Sound Identification	Top 10%	10	100	35%	113	69%
	Mid 10%	10	102	18%	120	77%
	Bottom 10%	10	100	4%	104	67%
Letter Reading Fluency	Top 10%	10	100	33.0	113	65.1
	Mid 10%	10	102	20.6	120	68.4
	Bottom 10%	10	100	9.1	104	62.3
Letter Reading Accuracy	Top 10%	10	100	66%	113	87%
	Mid 10%	10	102	44%	120	90%
	Bottom 10%	10	100	21%	104	85%
Word Reading Fluency	Top 10%	10	100	16.1	113	28.5
	Mid 10%	10	102	6.6	120	33.0
	Bottom 10%	10	100	1.9	104	29.6
Word Reading Accuracy	Top 10%	10	100	33%	113	84%
	Mid 10%	10	102	10%	120	88%
	Bottom 10%	10	100	2%	104	83%

BL Category	Avg Decile Movement*
Top 10%	-4.2
Mid 10%	1.2
Bottom 10%	4.1

- The average scores of schools from all three performance categories in the baseline was between the first and third quartile in the midline round.

- Schools in the bottom 10% performance category in the baseline round for literacy showed improvement in the midline round.

- The performance of most top performing schools in the baseline round for literacy dropped significantly in the midline round.

Mean school score is in the top 25% of school scores	Mean school score is in the second 25% of school scores	Mean school score is in the third 25% of school scores	Mean school score is in the bottom 25% of school scores

- Schools were ranked based on the mean school score in each task, i.e., the average of the scores of all students in the relevant grade in that school.
- An aggregate rank was created for each school across all literacy / numeracy tasks by adding the average ranks for each task, based on which the top 10%, middle 10%, and bottom 10% schools were selected.
- Average scores were calculated by taking the simple average of the mean school score in that task for all schools in that performance category.

** The average decile movement is the average change in the deciles of all the schools in each performance category, from the baseline to the midline round, with deciles determined based on the aggregate rank of the school.*

No consistent performance patterns are seen at the school level across the baseline and midline round in the literacy tasks for the low-touch demo group; in fact, both top and bottom performing schools in the baseline round scored much closer to the average in the midline round (2/2)

D-LT

Task	BL Category	# Schools	# Students_BL	BL_Avg	# Students_ML	ML_Avg
Non-Word Reading Fluency	Top 10%	10	100	11.4	113	28.9
	Mid 10%	10	102	3.5	120	30.5
	Bottom 10%	10	100	0.9	104	27.4
Oral Reading Fluency (ORF)	Top 10%	10	100	12.2	113	48.9
	Mid 10%	10	102	2.3	120	57.5
	Bottom 10%	10	100	0.5	104	49.9
Reading Comprehension Questions	Top 10%	10	100	20%	113	77%
	Mid 10%	10	102	0%	120	87%
	Bottom 10%	10	100	0%	104	76%
Letter Writing	Top 10%	10	100	55%	113	70%
	Mid 10%	10	102	33%	120	75%
	Bottom 10%	10	100	10%	104	67%
Word Writing	Top 10%	10	100	29%	113	68%
	Mid 10%	10	102	10%	120	74%
	Bottom 10%	10	100	1%	104	68%

BL Category	Avg Decile Movement*
Top 10%	-4.2
Mid 10%	1.2
Bottom 10%	4.1

- The average scores of schools from all three performance categories in the baseline was between the first and third quartile in the midline round.

- Schools in the bottom 10% performance category in the baseline round for literacy showed improvement in the midline round.

- The performance of most top performing schools in the baseline round for literacy dropped significantly in the midline round.

- Schools were ranked based on the mean school score in each task, i.e., the average of the scores of all students in the relevant grade in that school.
- An aggregate rank was created for each school across all literacy / numeracy tasks by adding the average ranks for each task, based on which the top 10%, middle 10%, and bottom 10% schools were selected.
- Average scores were calculated by taking the simple average of the mean school score in that task for all schools in that performance category.

** The average decile movement is the average change in the deciles of all the schools in each performance category, from the baseline to the midline round, with deciles determined based on the aggregate rank of the school.*

Mean school score is in the top 25% of school scores	Mean school score is in the second 25% of school scores	Mean school score is in the third 25% of school scores	Mean school score is in the bottom 25% of school scores

Decile Movement from Baseline to Midline for the Top 10%, Middle 10%, and Bottom 10% Schools at the Baseline, for the Low-Touch Demo Group of Cohort 1, in Literacy

D-LT

Decile Movement from Baseline to Midline for the Top 10% Schools at the Baseline

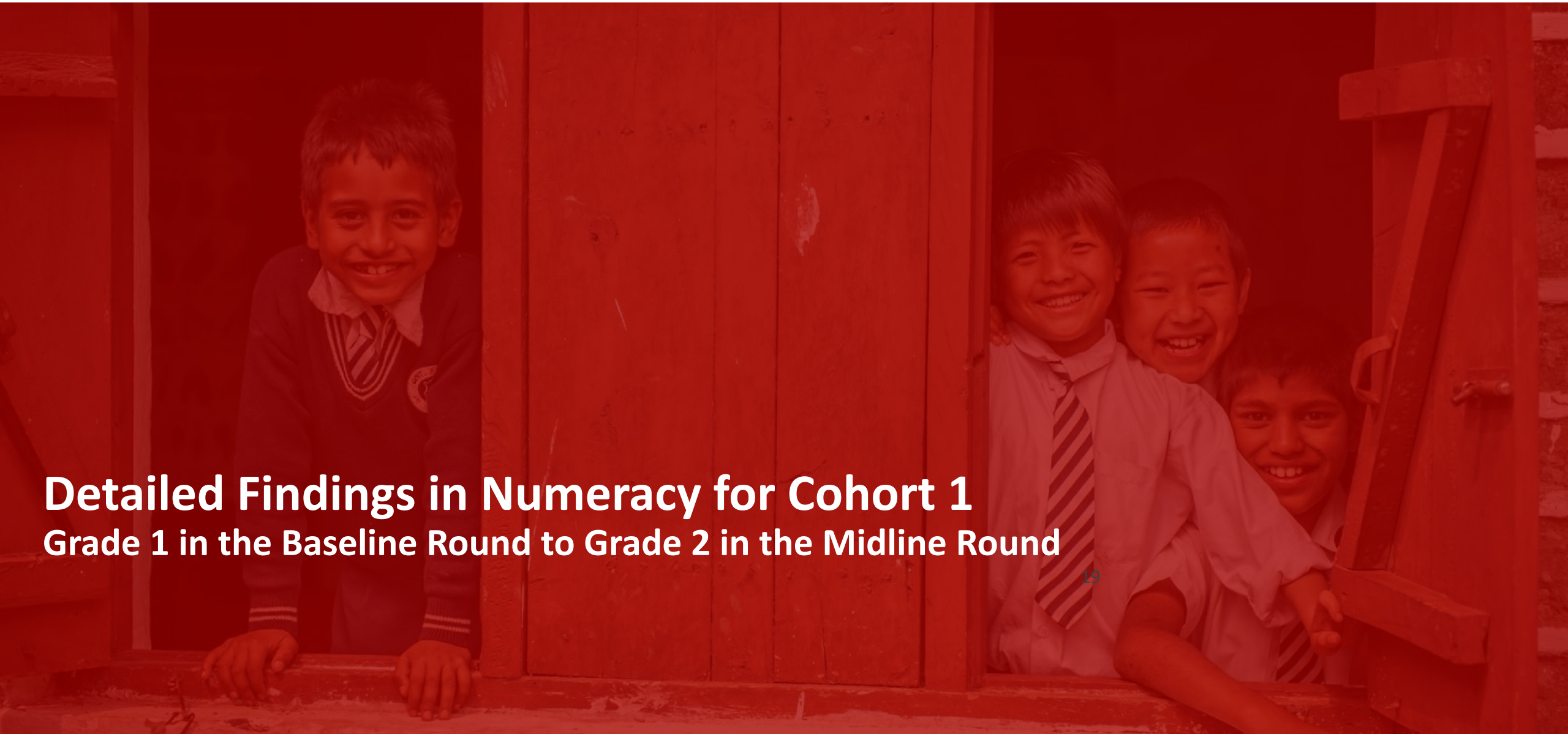
School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
1	1	5	-4
2	1	1	0
3	1	10	-9
4	1	6	-5
5	1	3	-2
6	1	2	-1
7	1	1	0
8	1	9	-8
9	1	5	-4
10	1	10	-9
Average	**1**	**5.2**	**-4.2**

Decile Movement from Baseline to Midline for the Middle 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
46	5	4	1
47	5	1	4
48	5	6	-1
49	5	9	-4
50	5	3	2
51	6	6	0
52	6	3	3
53	6	1	5
54	6	5	1
55	6	5	1
Average	**5.5**	**4.3**	**1.2**

Decile Movement from Baseline to Midline for the Bottom 10% Schools at the Baseline

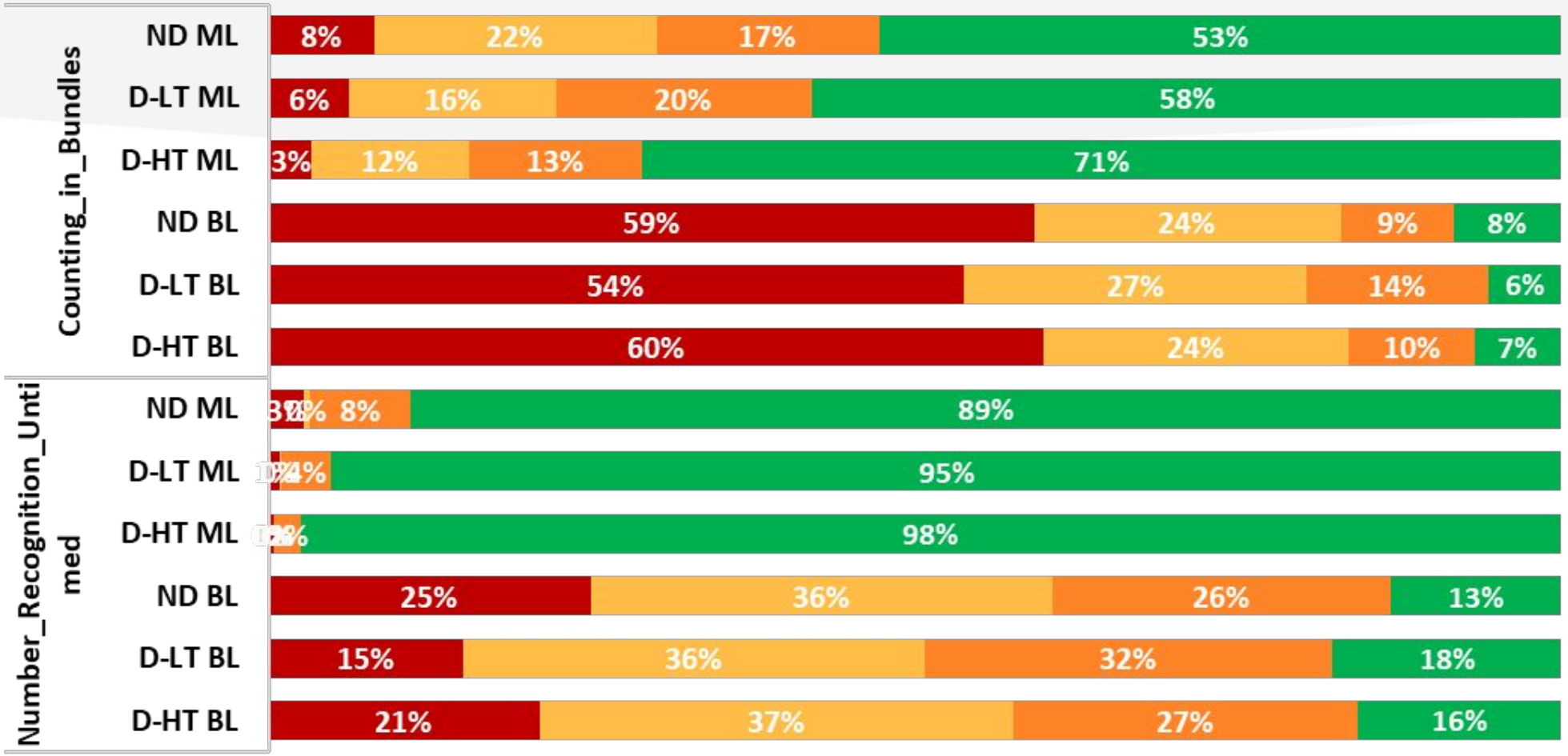
School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
91	10	6	4
92	10	5	5
93	10	4	6
94	10	9	1
95	10	9	1
96	10	2	8
97	10	8	2
98	10	5	5
99	10	7	3
100	10	4	6
Average	**10**	**5.9**	**4.1**



Detailed Findings in Numeracy for Cohort 1

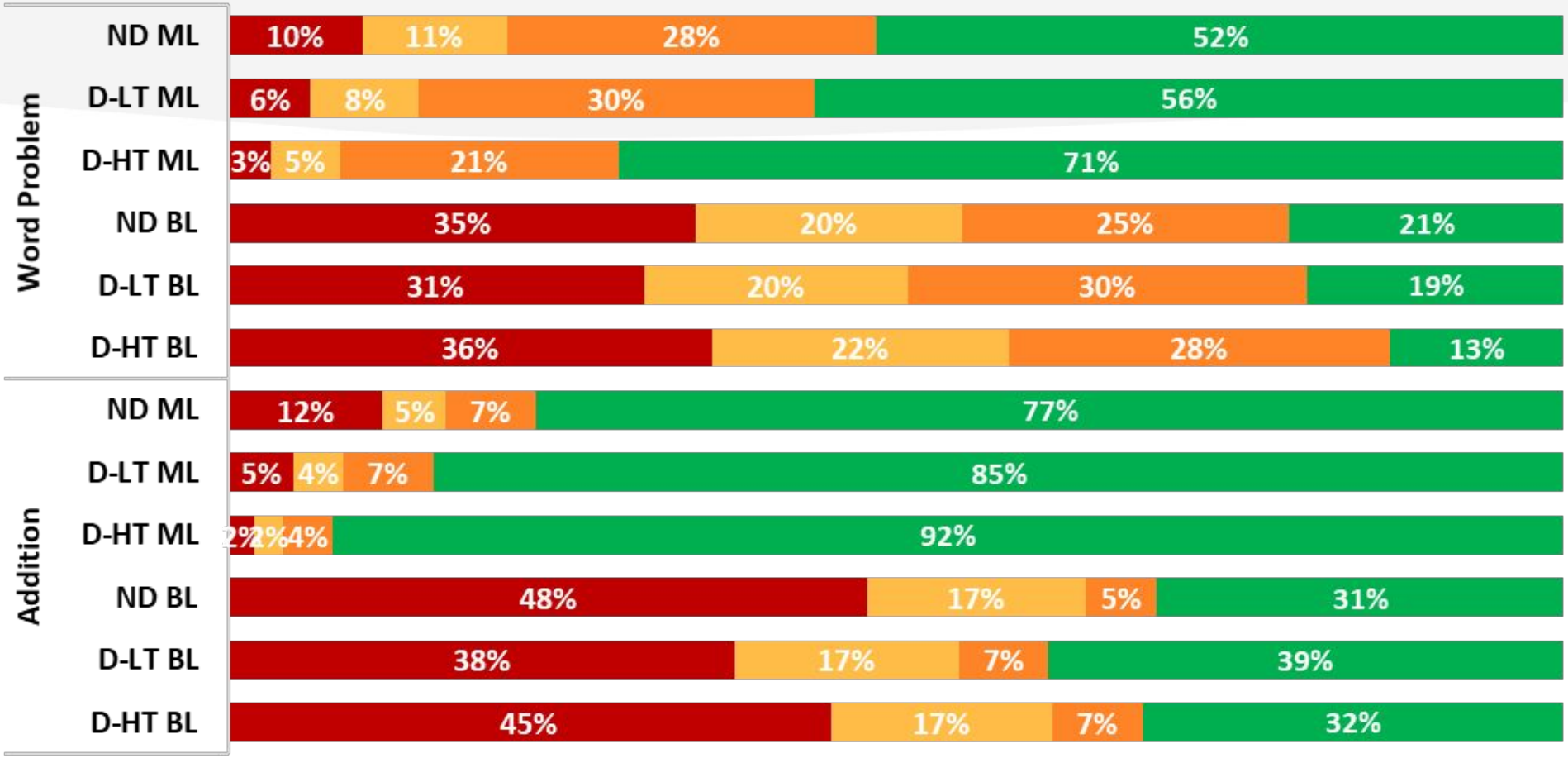
Grade 1 in the Baseline Round to Grade 2 in the Midline Round

The number of zero scorers has significantly reduced in the Number Recognition (Accuracy) and Counting in Bundles tasks. Most students performed at the L3 level in the Number Recognition (Accuracy) task, indicating that they were able to perform better than the average score for this task in the baseline round



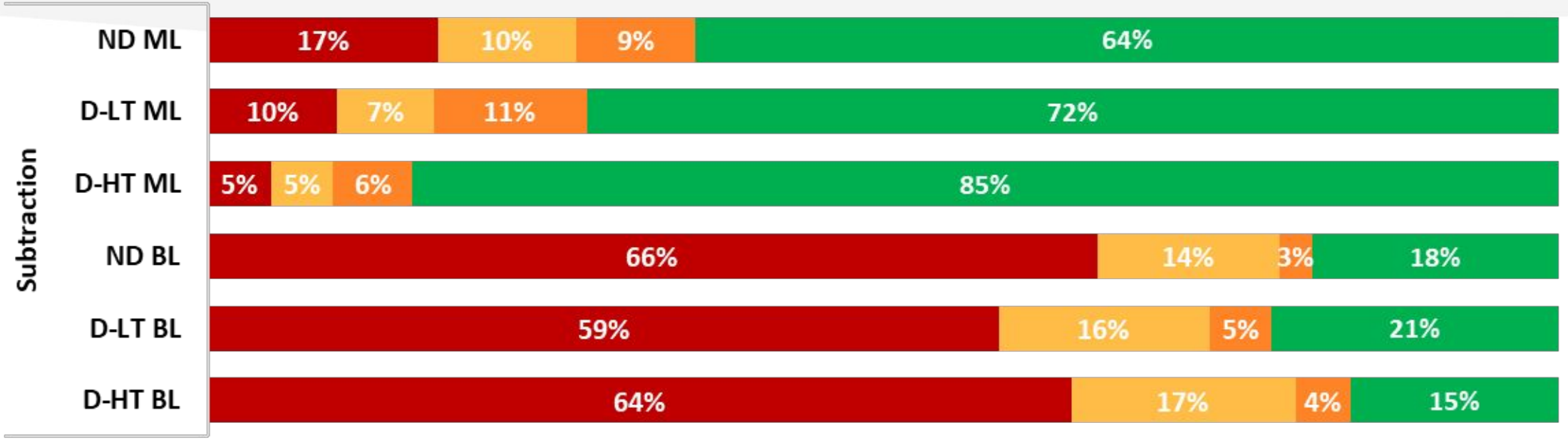
	L0	L1	L2	L3
Accuracy Tasks	0%	> 0% – 25%	> 25% – 50%	50% – 100%
Fluency Tasks	0	> 0 to (0.5 * Avg.)	> (0.5 * Avg.) to Avg.	> Avg. to Max.

In the Operations based task, a significant improvement is observed in the highest performance level compared to the baseline. In the Addition task, more than 92% of the students in the Demo sites performed at the L3 level, indicating that they have high proficiency in addition.



	L0	L1	L2	L3
Accuracy Tasks	0%	> 0% – 25%	> 25% – 50%	50% – 100%
Fluency Tasks	0	> 0 to (0.5 * Avg.)	> (0.5 * Avg.) to Avg.	> Avg. to Max.

Around 85% of students in High-Touch demo group were able to answer more than 50% of questions indicates a significant jump in the performance from the baseline for this performance level



	L0	L1	L2	L3
Accuracy Tasks	0%	> 0% – 25%	> 25% – 50%	50% – 100%
Fluency Tasks	0	> 0 to (0.5 * Avg.)	> (0.5 * Avg.) to Avg.	> Avg. to Max.

No consistent performance patterns are seen at the school level across the baseline and midline round in the numeracy tasks for the high-touch demo group; both top and bottom performing schools in the baseline round scored much closer to the average in the midline round

D-HT

Task	BL Category	# Schools	# Students_BL	BL_Avg	# Students_ML	ML_Avg
Number Recognition (Fluency)	Top 10%	10	102	18.9	113	33.5
	Mid 10%	11	127	12.6	81	28.6
	Bottom 10%	10	117	5.6	109	29.7
Number Recognition (Accuracy)	Top 10%	10	102	48%	113	88%
	Mid 10%	11	127	32%	81	81%
	Bottom 10%	10	117	12%	109	82%
Counting in Bundles	Top 10%	10	102	34%	113	86%
	Mid 10%	11	127	16%	81	79%
	Bottom 10%	10	117	5%	109	75%
Missing Numbers	Top 10%	10	102	29%	113	58%
	Mid 10%	11	127	12%	81	51%
	Bottom 10%	10	117	3%	109	47%
Addition	Top 10%	10	102	65%	113	94%
	Mid 10%	11	127	31%	81	90%
	Bottom 10%	10	117	8%	109	92%
Subtraction	Top 10%	10	102	48%	113	87%
	Mid 10%	11	127	14%	81	86%
	Bottom 10%	10	117	4%	109	82%
Word Problems	Top 10%	10	102	44%	113	72%
	Mid 10%	11	127	24%	81	70%
	Bottom 10%	10	117	12%	109	68%

BL Category	Avg Decile Movement*
Top 10%	-3.1
Mid 10%	0.3
Bottom 10%	4.4

- The average scores of schools from all three performance categories in the baseline was between the first and third quartile in the midline round.

- The performance of schools in the bottom 10% category in the baseline round for numeracy improved in the midline round for all tasks

- The performance of most top performing schools in the baseline for numeracy dropped significantly in the midline round for all tasks

Mean school score is in the top 25% of school scores	Mean school score is in the second 25% of school scores	Mean school score is in the third 25% of school scores	Mean school score is in the bottom 25% of school scores

- Schools were ranked based on the mean school score in each task, i.e., the average of the scores of all students in the relevant grade in that school.
- An aggregate rank was created for each school across all literacy / numeracy tasks by adding the average ranks for each task, based on which the top 10%, middle 10%, and bottom 10% schools were selected.
- Average scores were calculated by taking the simple average of the mean school score in that task for all schools in that performance category.

** The average decile movement is the average change in the deciles of all the schools in each performance category, from the baseline to the midline round, with deciles determined based on the aggregate rank of the school.*

Decile Movement from Baseline to Midline for the Top 10%, Middle 10%, and Bottom 10% Schools at the Baseline, for the High-Touch Demo Group of Cohort 1, in Numeracy

D-HT

Decile Movement from Baseline to Midline for the Top 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
1	1	2	-1
2	1	8	-7
3	1	1	0
4	1	1	0
5	1	3	-2
6	1	2	-1
7	1	5	-4
8	1	7	-6
9	1	6	-5
10	1	6	-5
Average	**1**	**4.1**	**-3.1**

Decile Movement from Baseline to Midline for the Middle 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
45	5	2	3
46	5	9	-4
47	5	-	-
48	5	-	-
49	5	3	2
50	6	3	3
51	6	-	-
52	6	8	-2
53	6	10	-4
54	6	-	-
55	6	2	4
Average	**5.5**	**5.3**	**0.3**

Decile Movement from Baseline to Midline for the Bottom 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
90	10	1	9
91	10	9	1
92	10	6	4
93	10	3	7
94	10	6	4
95	10	8	2
96	10	3	7
97	10	10	0
98	10	1	9
99	10	9	1
Average	**10**	**5.6**	**4.4**

Apart from the Number Recognition (Fluency) and Word Problems tasks, schools in the low-touch demo group maintained the performance patterns seen in the baseline round, though the top and bottom performing schools scored closer to the average in the midline round

D-LT

Task	BL Category	# Schools	# Students_BL	BL_Avg	# Students_ML	ML_Avg
Number Recognition (Fluency)	Top 10%	10	101	21.4	107	30.0
	Mid 10%	10	101	13.6	118	31.3
	Bottom 10%	10	98	5.7	103	29.4
Number Recognition (Accuracy)	Top 10%	10	101	55%	107	82%
	Mid 10%	10	101	30%	118	81%
	Bottom 10%	10	98	13%	103	77%
Counting in Bundles	Top 10%	10	101	34%	107	69%
	Mid 10%	10	101	20%	118	75%
	Bottom 10%	10	98	3%	103	60%
Missing Numbers	Top 10%	10	101	33%	107	46%
	Mid 10%	10	101	15%	118	45%
	Bottom 10%	10	98	5%	103	38%
Addition	Top 10%	10	101	68%	107	89%
	Mid 10%	10	101	37%	118	88%
	Bottom 10%	10	98	12%	103	77%
Subtraction	Top 10%	10	101	51%	107	78%
	Mid 10%	10	101	17%	118	80%
	Bottom 10%	10	98	6%	103	62%
Word Problems	Top 10%	10	101	55%	107	61%
	Mid 10%	10	101	28%	118	66%
	Bottom 10%	10	98	13%	103	57%

BL Category	Avg Decile Movement*
Top 10%	-3.5
Mid 10%	1.2
Bottom 10%	2.9

- The performance of schools in the bottom 10% category in the baseline round for numeracy improved significantly in the midline round for all tasks except Subtraction

- The performance of most top performing schools in the baseline for numeracy dropped significantly in the midline round for all tasks, especially Word Problems

Mean school score is in the top 25% of school scores	Mean school score is in the second 25% of school scores	Mean school score is in the third 25% of school scores	Mean school score is in the bottom 25% of school scores

- Schools were ranked based on the mean school score in each task, i.e., the average of the scores of all students in the relevant grade in that school.
- An aggregate rank was created for each school across all literacy / numeracy tasks by adding the average ranks for each task, based on which the top 10%, middle 10%, and bottom 10% schools were selected.
- Average scores were calculated by taking the simple average of the mean school score in that task for all schools in that performance category.

** The average decile movement is the average change in the deciles of all the schools in each performance category, from the baseline to the midline round, with deciles determined based on the aggregate rank of the school.*

Decile Movement from Baseline to Midline for the Top 10%, Middle 10%, and Bottom 10% Schools at the Baseline, for the Low-Touch Demo Group of Cohort 1, in Numeracy

D-LT

Decile Movement from Baseline to Midline for the Top 10% Schools at the Baseline

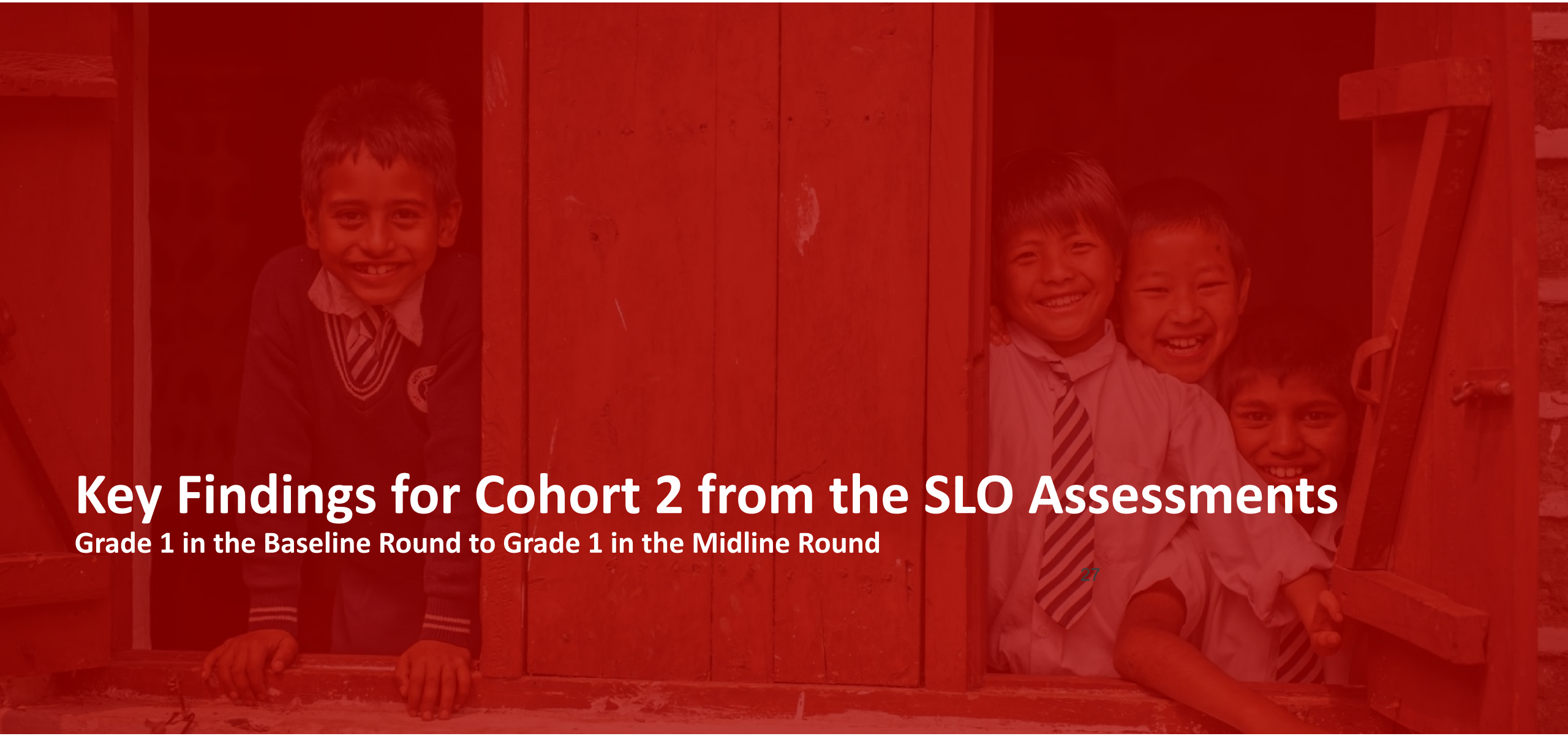
School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
1	1	6	-5
2	1	4	-3
3	1	10	-9
4	1	1	0
5	1	2	-1
6	1	8	-7
7	1	1	0
8	1	4	-3
9	1	6	-5
10	1	3	-2
Average	**1**	**4.5**	**-3.5**

Decile Movement from Baseline to Midline for the Middle 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
46	5	4	1
47	5	9	-4
48	5	2	3
49	5	3	2
50	5	4	1
51	6	7	-1
52	6	9	-3
53	6	1	5
54	6	2	4
55	6	2	4
Average	**5.5**	**4.3**	**1.2**

Decile Movement from Baseline to Midline for the Bottom 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
91	10	7	3
92	10	5	5
93	10	6	4
94	10	10	0
95	10	7	3
96	10	9	1
97	10	9	1
98	10	5	5
99	10	5	5
100	10	8	2
Average	**10**	**7.1**	**2.9**



Key Findings for Cohort 2 from the SLO Assessments

Grade 1 in the Baseline Round to Grade 1 in the Midline Round

The intervention had a significant impact on the performance of Cohort 2 across both the High-Touch and Low-Touch sites, when compared to the non-demo sites, in literacy tasks such as Letter Reading, Word Reading, Non-Word Reading, and ORF, with a greater effect size in Sewapuri across most literacy tasks

Task	Unit	Midline - Average			Effect Size	
		ND	D-HT	D-LT	D-HT	D-LT
Listening Comprehension	Percentage	71%	74%	66%	0.10	-0.16
Oral Vocabulary	Percentage	95%	96%	94%	0.06	-0.10
Initial Sound Identification	Percentage	45%	63%	46%	0.40	0.02
Letter Reading (Accuracy)	Percentage	60%	83%	73%	**0.76**	0.39
Letter Reading (Fluency)	Count per minute	33.6	51.4	40.4	**0.79**	0.29
Word Reading (Accuracy)	Percentage	54%	80%	67%	**0.88**	0.41
Word Reading (Fluency)	Count per minute	14.5	27.1	21.0	**0.86**	0.20
Non-Word Reading (Fluency)	Count per minute	10.5	21.2	15.0	**0.87**	0.33
Oral Reading Fluency (ORF)	Count per minute	15.2	32.3	20.7	**0.71**	0.24
Reading Comprehension Passage 1	Percentage	34%	59%	38%	**0.62**	0.10
Letter Writing	Percentage	61%	77%	64%	0.46	0.07
Word Writing	Percentage	45%	70%	58%	**0.75**	0.35

The DiD effect size was calculated using the formula: **[avg_demo (Δi) - avg_non-demo (Δc)] / SD_pooled (Pooled Standard Deviation)**

The intervention also had a significant impact on the performance of Cohort 2 in the High-Touch sites, when compared to the non-demo sites, in numeracy tasks such as Addition and Subtraction. The midline performance of Cohort 2 in the Low-Touch sites was also noticeably better than the non-demo sites

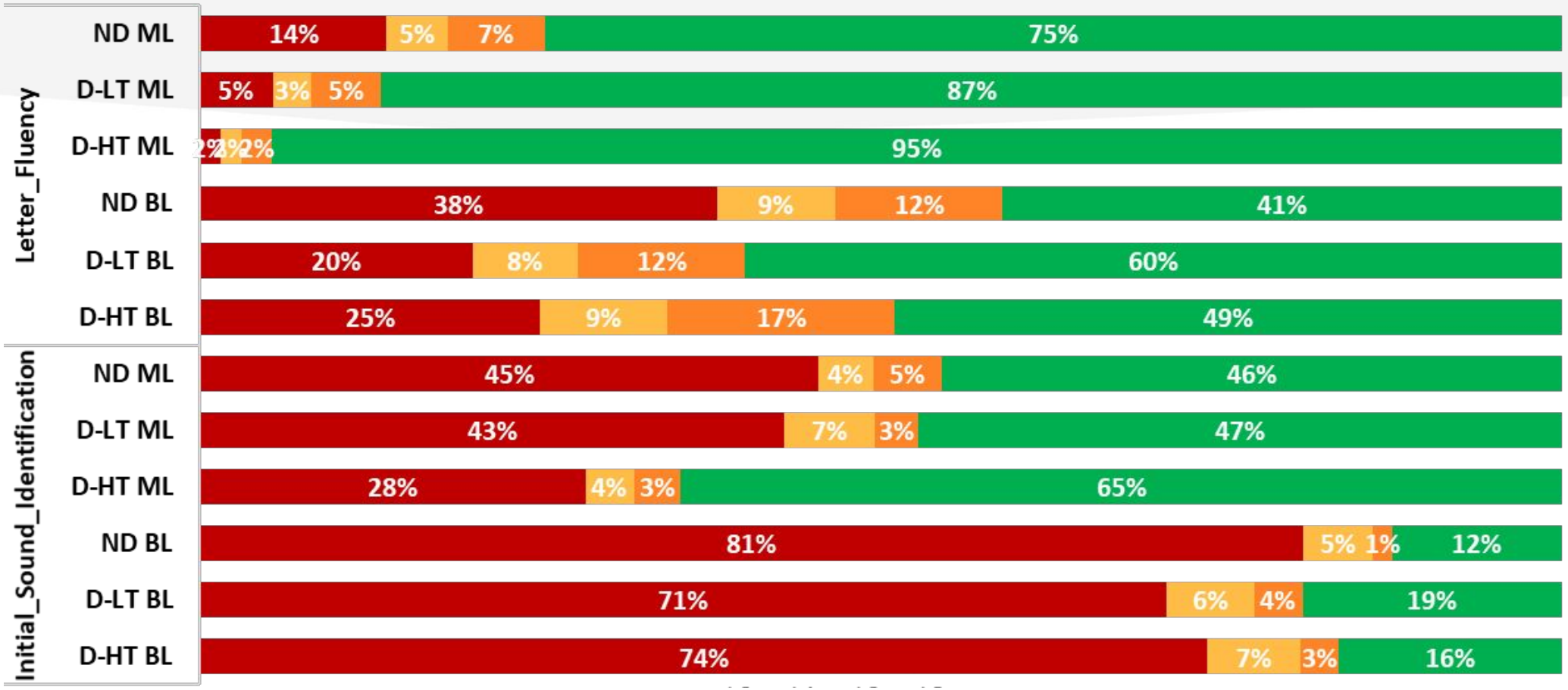
Task	Unit	Midline - Average			Effect Size	
		ND	D-HT	D-LT	D-HT	D-LT
Counting	Count per minute	98.9	108.9	102.2	0.28	0.09
Number Recognition (Fluency)	Count per minute	19.0	23.1	21.3	0.29	0.16
Number Recognition (Accuracy)	Percentage	61%	69%	65%	0.38	0.16
Number Comparison	Percentage	45%	57%	50%	0.37	0.16
Counting in Bundles	Percentage	45%	55%	45%	0.30	0.01
Missing Number	Percentage	34%	48%	39%	0.49	0.17
Addition Level 1 (Accuracy)	Percentage	64%	85%	72%	**0.58**	0.21
Subtraction Level 1 (Accuracy)	Percentage	51%	76%	60%	**0.64**	0.22
Word Problems	Percentage	55%	66%	54%	0.35	-0.02
Shape Recognition - Circle	Percentage	30%	26%	34%	-0.19	0.20
Shape Recognition - Rectangle	Percentage	53%	49%	53%	-0.21	0.03

The DiD effect size was calculated using the formula: ***[avg_demo (Δi) - avg_non-demo (Δc)] / SD_pooled (Pooled Standard Deviation)***

Detailed Findings in Literacy for Cohort 2

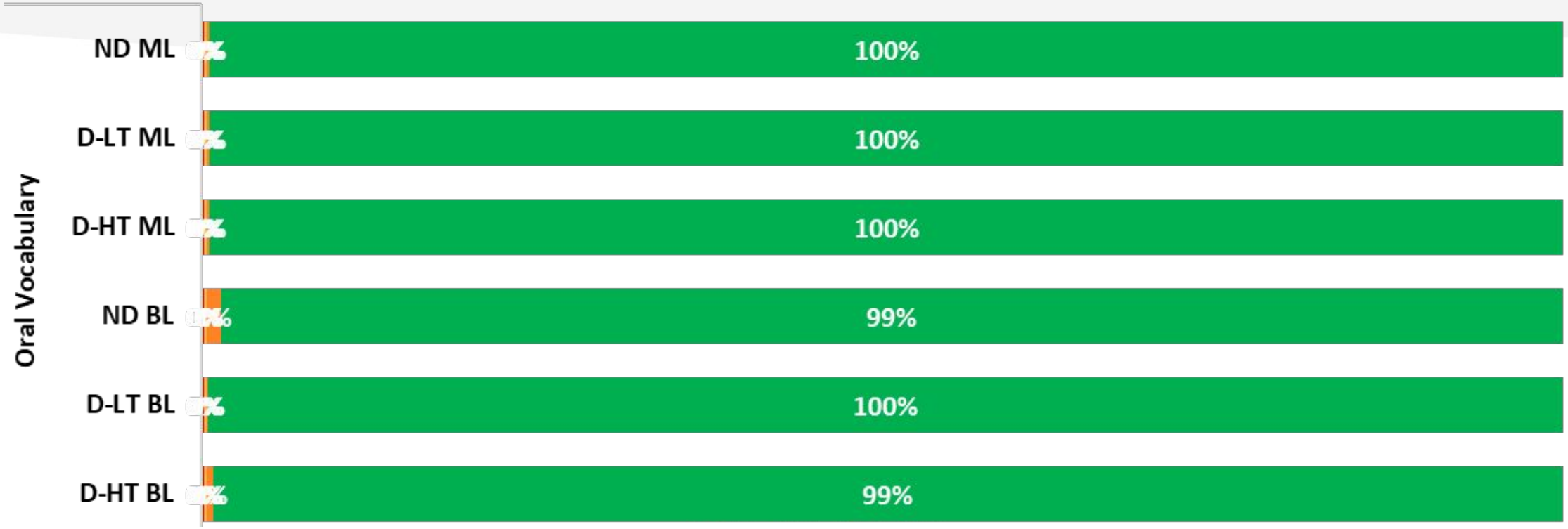
Grade 1 in the Midline Round

As compared to cohort 1, the number of students at L3 is significantly higher in the Initial Sound Identification task in this cohort. In the Letter Naming Fluency task, a large majority of students were able to perform better than the average score for this task in the baseline round



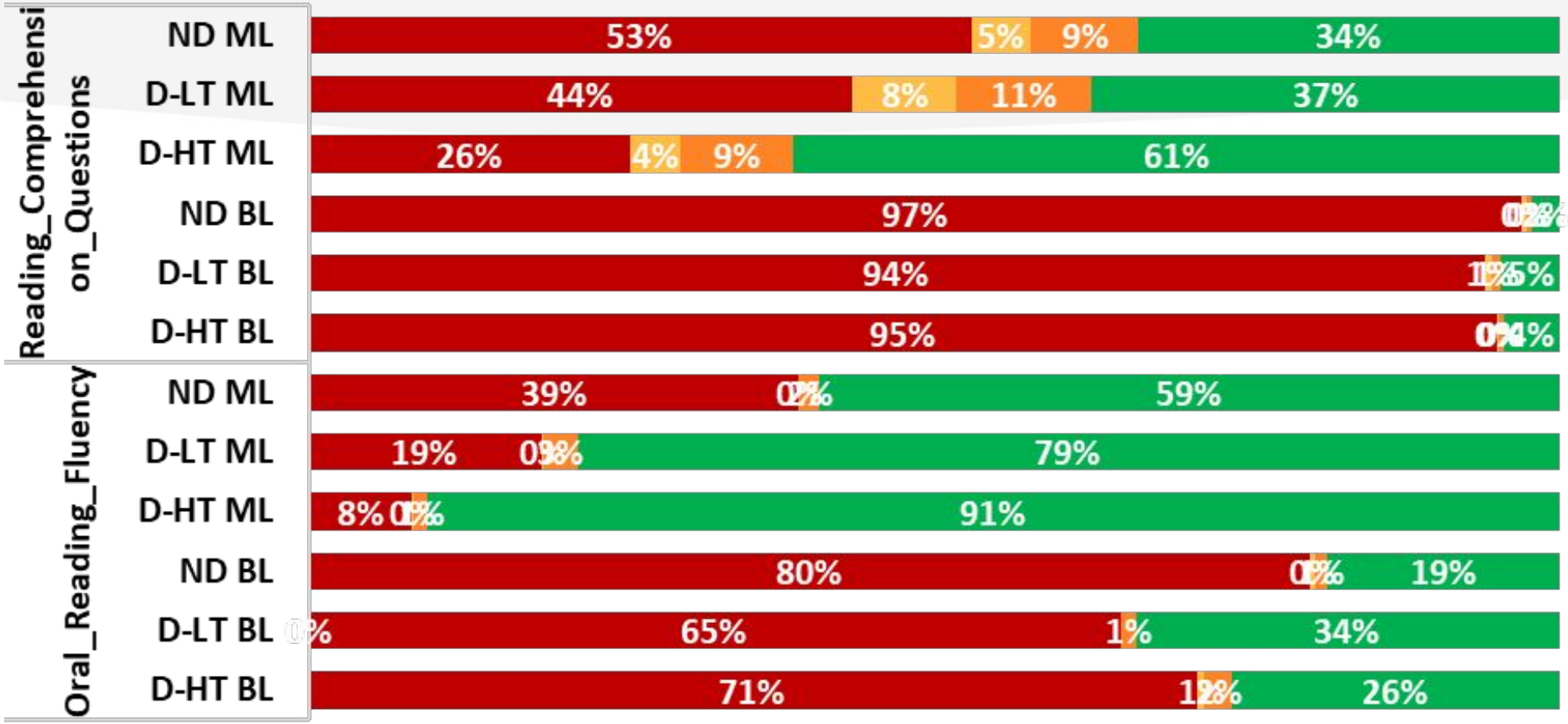
	L0	L1	L2	L3
Accuracy Tasks	0%	> 0% – 25%	> 25% – 50%	50% – 100%
Fluency Tasks	0	> 0 to (0.5 * Avg.)	> (0.5 * Avg.) to Avg.	> Avg. to Max.

The results of the Oral Vocabulary task show that it's a ceiling task for cohort 2 as well, with 100% of students across all groups achieving the highest learning level.



	L0	L1	L2	L3
Accuracy Tasks	0%	> 0% – 25%	> 25% – 50%	50% – 100%
Fluency Tasks	0	> 0 to (0.5 * Avg.)	> (0.5 * Avg.) to Avg.	> Avg. to Max.

Students in this Cohort from the High-Touch Demo group performed higher on both the tasks (RC and ORF) indicating their improved skills related to reading a text and making meaning of it



	L0	L1	L2	L3
Accuracy Tasks	0%	> 0% – 25%	> 25% – 50%	50% – 100%
Fluency Tasks	0	> 0 to (0.5 * Avg.)	> (0.5 * Avg.) to Avg.	> Avg. to Max.

Apart from the Listening Comprehension and Oral Vocabulary, schools in the high-touch demo group of Cohort 2 maintained the performance patterns seen in the baseline round for Cohort 1, though the top and bottom performing schools scored closer to the average in the midline round (1/2)

D-HT

Task	BL Category	# Schools	# Students_BL	BL_Avg	# Students_ML	ML_Avg
Listening Comprehension	Top 10%	10	123	73%	121	65%
	Mid 10%	11	131	55%	107	74%
	Bottom 10%	10	112	48%	87	72%
Oral Vocabulary	Top 10%	10	123	99%	121	96%
	Mid 10%	11	131	96%	107	96%
	Bottom 10%	10	112	92%	87	95%
Initial Sound Identification	Top 10%	10	123	37%	121	68%
	Mid 10%	11	131	11%	107	61%
	Bottom 10%	10	112	5%	87	50%
Letter Reading Fluency	Top 10%	10	123	30.1	121	56.9
	Mid 10%	11	131	17.5	107	47.4
	Bottom 10%	10	112	6.3	87	42.1
Letter Reading Accuracy	Top 10%	10	123	55%	121	88%
	Mid 10%	11	131	36%	107	81%
	Bottom 10%	10	112	13%	87	74%
Word Reading Fluency	Top 10%	10	123	13.1	121	30.3
	Mid 10%	11	131	6.5	107	24.4
	Bottom 10%	10	112	0.6	87	21.2
Word Reading Accuracy	Top 10%	10	123	32%	121	86%
	Mid 10%	11	131	13%	107	78%
	Bottom 10%	10	112	1%	87	71%

BL Category	Avg Decile Movement*
Top 10%	-3.3
Mid 10%	-0.4
Bottom 10%	3.5

- Schools in the bottom 10% performance category in the baseline round for literacy improved significantly in most tasks in the midline round, apart from Letter Reading Fluency and Accuracy, and Word Reading Accuracy.

- The performance of most top performing schools in the baseline round for literacy dropped significantly in the midline round in all tasks.

Mean school score is in the top 25% of school scores	Mean school score is in the second 25% of school scores	Mean school score is in the third 25% of school scores	Mean school score is in the bottom 25% of school scores

- Schools were ranked based on the mean school score in each task, i.e., the average of the scores of all students in the relevant grade in that school.
- An aggregate rank was created for each school across all literacy / numeracy tasks by adding the average ranks for each task, based on which the top 10%, middle 10%, and bottom 10% schools were selected.
- Average scores were calculated by taking the simple average of the mean school score in that task for all schools in that performance category.

** The average decile movement is the average change in the deciles of all the schools in each performance category, from the baseline to the midline round, with deciles determined based on the aggregate rank of the school.*

Apart from the Listening Comprehension and Oral Vocabulary, schools in the high-touch demo group of Cohort 2 maintained the performance patterns seen in the baseline round for Cohort 1, though the top and bottom performing schools scored closer to the average in the midline round (2/2)

D-HT

Task	BL Category	# Schools	# Students_BL	BL_Avg	# Students_ML	ML_Avg
Non-Word Reading Fluency	Top 10%	10	123	8.5	121	23.8
	Mid 10%	11	131	3.6	107	19.6
	Bottom 10%	10	112	0.3	87	17.1
Oral Reading Fluency (ORF)	Top 10%	10	123	9.2	121	37.0
	Mid 10%	11	131	2.3	107	26.5
	Bottom 10%	10	112	0.2	87	23.4
Reading Comprehension Questions	Top 10%	10	123	17%	121	65%
	Mid 10%	11	131	3%	107	52%
	Bottom 10%	10	112	0%	87	48%
Letter Writing	Top 10%	10	123	50%	121	86%
	Mid 10%	11	131	20%	107	78%
	Bottom 10%	10	112	6%	87	70%
Word Writing	Top 10%	10	123	25%	121	76%
	Mid 10%	11	131	4%	107	67%
	Bottom 10%	10	112	0%	87	63%

BL Category	Avg Decile Movement*
Top 10%	-3.3
Mid 10%	-0.4
Bottom 10%	3.5

- Schools in the bottom 10% performance category in the baseline round for literacy improved significantly in most tasks in the midline round, apart from Letter Reading Fluency and Accuracy, and Word Reading Accuracy.

- The performance of most top performing schools in the baseline round for literacy dropped significantly in the midline round in all tasks.

- Schools were ranked based on the mean school score in each task, i.e., the average of the scores of all students in the relevant grade in that school.
- An aggregate rank was created for each school across all literacy / numeracy tasks by adding the average ranks for each task, based on which the top 10%, middle 10%, and bottom 10% schools were selected.
- Average scores were calculated by taking the simple average of the mean school score in that task for all schools in that performance category.

** The average decile movement is the average change in the deciles of all the schools in each performance category, from the baseline to the midline round, with deciles determined based on the aggregate rank of the school.*

Mean school score is in the top 25% of school scores	Mean school score is in the second 25% of school scores	Mean school score is in the third 25% of school scores	Mean school score is in the bottom 25% of school scores

Decile Movement from Baseline to Midline for the Top 10%, Middle 10%, and Bottom 10% Schools at the Baseline, for Grade 1 in the High-Touch Demo Group of Cohorts 1 and 2, in Literacy

D-HT

Decile Movement from Baseline to Midline for the Top 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
1	1	3	-2
2	1	8	-7
3	1	5	-4
4	1	3	-2
5	1	3	-2
6	1	7	-6
7	1	5	-4
8	1	3	-2
9	1	2	-1
10	1	4	-3
Average	**1**	**4.3**	**-3.3**

Decile Movement from Baseline to Midline for the Middle 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
45	5	9	-4
46	5	9	-4
47	5	-	-
48	5	10	-5
49	5	2	3
50	6	4	2
51	6	8	-2
52	6	5	1
53	6	6	0
54	6	6	0
55	6	1	5
Average	**5.5**	**6.0**	**-0.4**

Decile Movement from Baseline to Midline for the Bottom 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
90	10	6	4
91	10	10	0
92	10	10	0
93	10	2	8
94	10	7	3
95	10	9	1
96	10	1	9
97	10	2	8
98	10	8	2
99	10	10	0
Average	**10**	**6.5**	**3.5**

No consistent performance patterns are seen at the school level in literacy across the baseline round for Cohort 2 and the midline round for Cohort 1, for the low-touch demo group; both top and bottom performing schools in the baseline scored much closer to the average in the midline (1/2)

D-LT

Task	BL Category	# Schools	# Students_BL	BL_Avg	# Students_ML	ML_Avg
Listening Comprehension	Top 10%	10	100	73%	111	67%
	Mid 10%	10	102	65%	122	69%
	Bottom 10%	10	100	49%	98	60%
Oral Vocabulary	Top 10%	10	100	99%	111	95%
	Mid 10%	10	102	98%	122	95%
	Bottom 10%	10	100	94%	98	94%
Initial Sound Identification	Top 10%	10	100	35%	111	40%
	Mid 10%	10	102	18%	122	55%
	Bottom 10%	10	100	4%	98	39%
Letter Reading Fluency	Top 10%	10	100	33.0	111	43.0
	Mid 10%	10	102	20.6	122	43.9
	Bottom 10%	10	100	9.1	98	29.5
Letter Reading Accuracy	Top 10%	10	100	66%	111	75%
	Mid 10%	10	102	44%	122	77%
	Bottom 10%	10	100	21%	98	60%
Word Reading Fluency	Top 10%	10	100	16.1	111	20.3
	Mid 10%	10	102	6.6	122	20.8
	Bottom 10%	10	100	1.9	98	13.5
Word Reading Accuracy	Top 10%	10	100	33%	111	69%
	Mid 10%	10	102	10%	122	72%
	Bottom 10%	10	100	2%	98	54%

BL Category	Avg Decile Movement*
Top 10%	-4.8
Mid 10%	1.3
Bottom 10%	2.1

- Schools in the bottom 10% performance category in the baseline round for literacy showed improvement in the midline round in all tasks except Letter Reading Accuracy, Word Reading Accuracy, and Oral Reading Fluency (ORF).

- The performance of most top performing schools in the baseline round for literacy dropped significantly in the midline round, performing worse than the middle 10% performance category across all tasks, on average.

Mean school score is in the top 25% of school scores	Mean school score is in the second 25% of school scores	Mean school score is in the third 25% of school scores	Mean school score is in the bottom 25% of school scores

- Schools were ranked based on the mean school score in each task, i.e., the average of the scores of all students in the relevant grade in that school.
- An aggregate rank was created for each school across all literacy / numeracy tasks by adding the average ranks for each task, based on which the top 10%, middle 10%, and bottom 10% schools were selected.
- Average scores were calculated by taking the simple average of the mean school score in that task for all schools in that performance category.

** The average decile movement is the average change in the deciles of all the schools in each performance category, from the baseline to the midline round, with deciles determined based on the aggregate rank of the school.*

No consistent performance patterns are seen at the school level in literacy across the baseline round for Cohort 2 and the midline round for Cohort 1, for the low-touch demo group; both top and bottom performing schools in the baseline scored much closer to the average in the midline (2/2)

D-LT

Task	BL Category	# Schools	# Students_BL	BL_Avg	# Students_ML	ML_Avg
Non-Word Reading Fluency	Top 10%	10	100	11.4	111	14.4
	Mid 10%	10	102	3.5	122	16.4
	Bottom 10%	10	100	0.9	98	10.0
Oral Reading Fluency (ORF)	Top 10%	10	100	12.2	111	18.8
	Mid 10%	10	102	2.3	122	21.7
	Bottom 10%	10	100	0.5	98	11.0
Reading Comprehension Questions	Top 10%	10	100	20%	111	34%
	Mid 10%	10	102	0%	122	46%
	Bottom 10%	10	100	0%	98	24%
Letter Writing	Top 10%	10	100	55%	111	59%
	Mid 10%	10	102	33%	122	77%
	Bottom 10%	10	100	10%	98	54%
Word Writing	Top 10%	10	100	29%	111	51%
	Mid 10%	10	102	10%	122	64%
	Bottom 10%	10	100	1%	98	46%

BL Category	Avg Decile Movement*
Top 10%	-4.8
Mid 10%	1.3
Bottom 10%	2.1

- Schools in the bottom 10% performance category in the baseline round for literacy showed improvement in the midline round in all tasks except Letter Reading Accuracy, Word Reading Accuracy, and Oral Reading Fluency (ORF).

- The performance of most top performing schools in the baseline round for literacy dropped significantly in the midline round, performing worse than the middle 10% performance category across all tasks, on average.

- Schools were ranked based on the mean school score in each task, i.e., the average of the scores of all students in the relevant grade in that school.
- An aggregate rank was created for each school across all literacy / numeracy tasks by adding the average ranks for each task, based on which the top 10%, middle 10%, and bottom 10% schools were selected.
- Average scores were calculated by taking the simple average of the mean school score in that task for all schools in that performance category.

** The average decile movement is the average change in the deciles of all the schools in each performance category, from the baseline to the midline round, with deciles determined based on the aggregate rank of the school.*

Mean school score is in the top 25% of school scores	Mean school score is in the second 25% of school scores	Mean school score is in the third 25% of school scores	Mean school score is in the bottom 25% of school scores

Decile Movement from Baseline to Midline for the Top 10%, Middle 10%, and Bottom 10% Schools at the Baseline, for Grade 1 in the Low-Touch Demo Group of Cohorts 1 and 2, in Literacy

D-LT

Decile Movement from Baseline to Midline for the Top 10% Schools at the Baseline

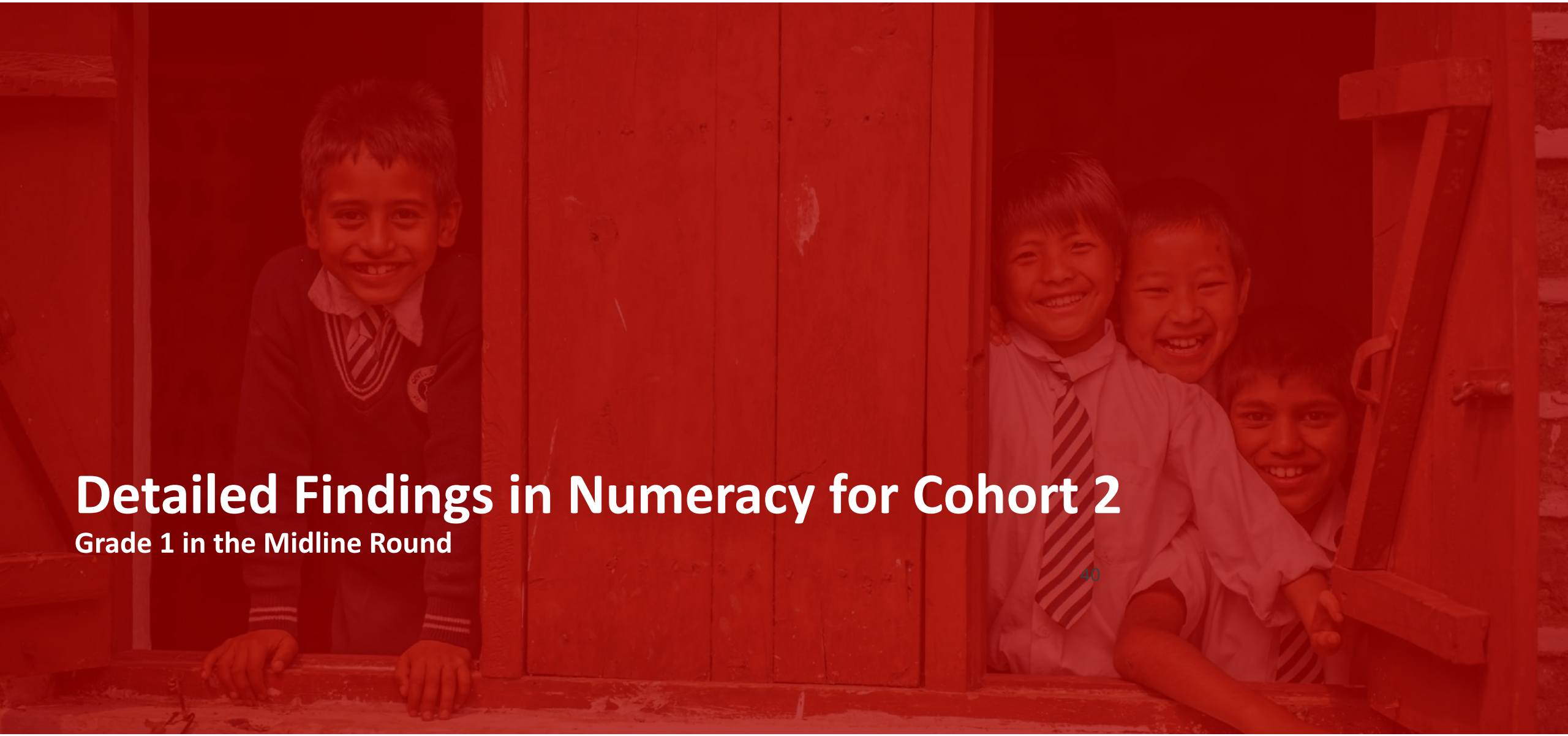
School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
1	1	5	-4
2	1	10	-9
3	1	10	-9
4	1	7	-6
5	1	1	0
6	1	1	0
7	1	8	-7
8	1	3	-2
9	1	5	-4
10	1	8	-7
Average	**1**	**5.8**	**-4.8**

Decile Movement from Baseline to Midline for the Middle 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
46	5	1	4
47	5	1	4
48	5	6	-1
49	5	10	-5
50	5	3	2
51	6	4	2
52	6	4	2
53	6	5	1
54	6	7	-1
55	6	1	5
Average	**5.5**	**4.2**	**1.3**

Decile Movement from Baseline to Midline for the Bottom 10% Schools at the Baseline

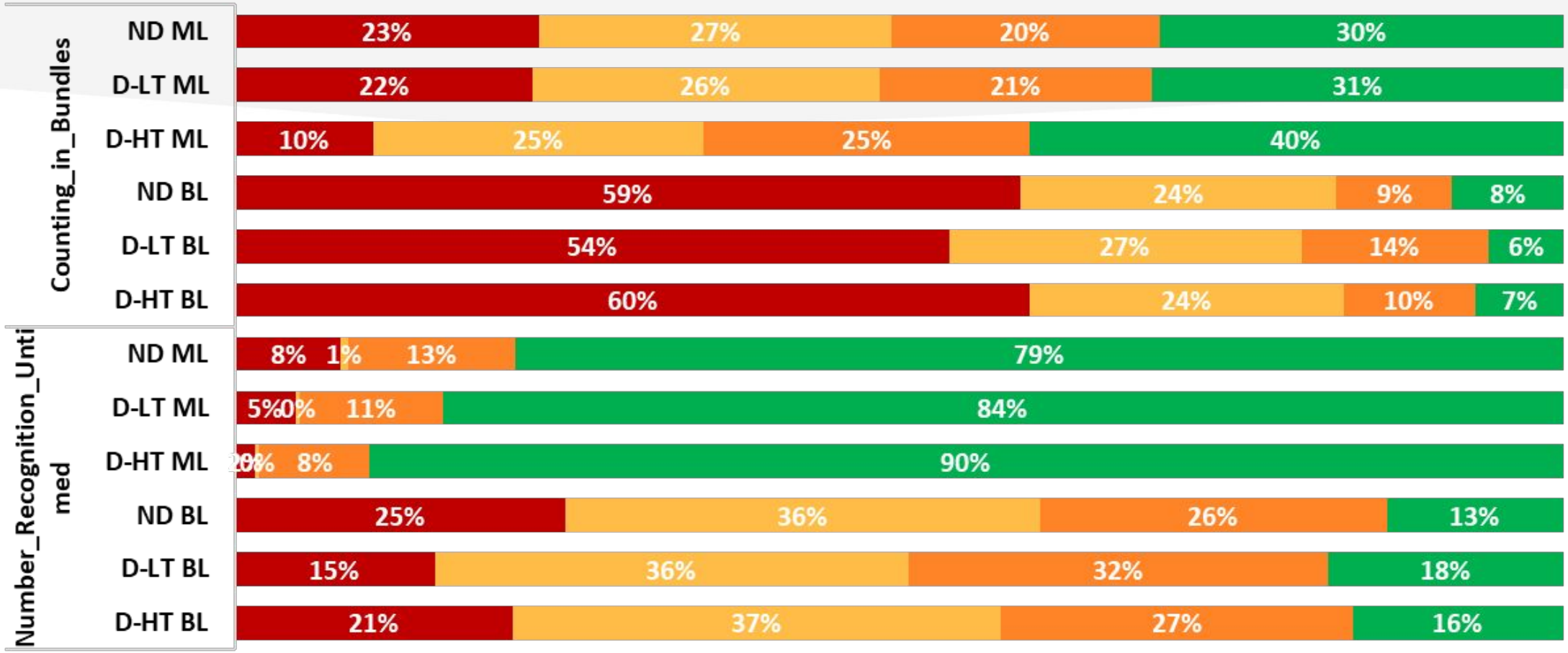
School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
91	10	9	1
92	10	9	1
93	10	10	0
94	10	6	4
95	10	10	0
96	10	8	2
97	10	10	0
98	10	4	6
99	10	5	5
100	10	8	2
Average	**10**	**7.9**	**2.1**



Detailed Findings in Numeracy for Cohort 2

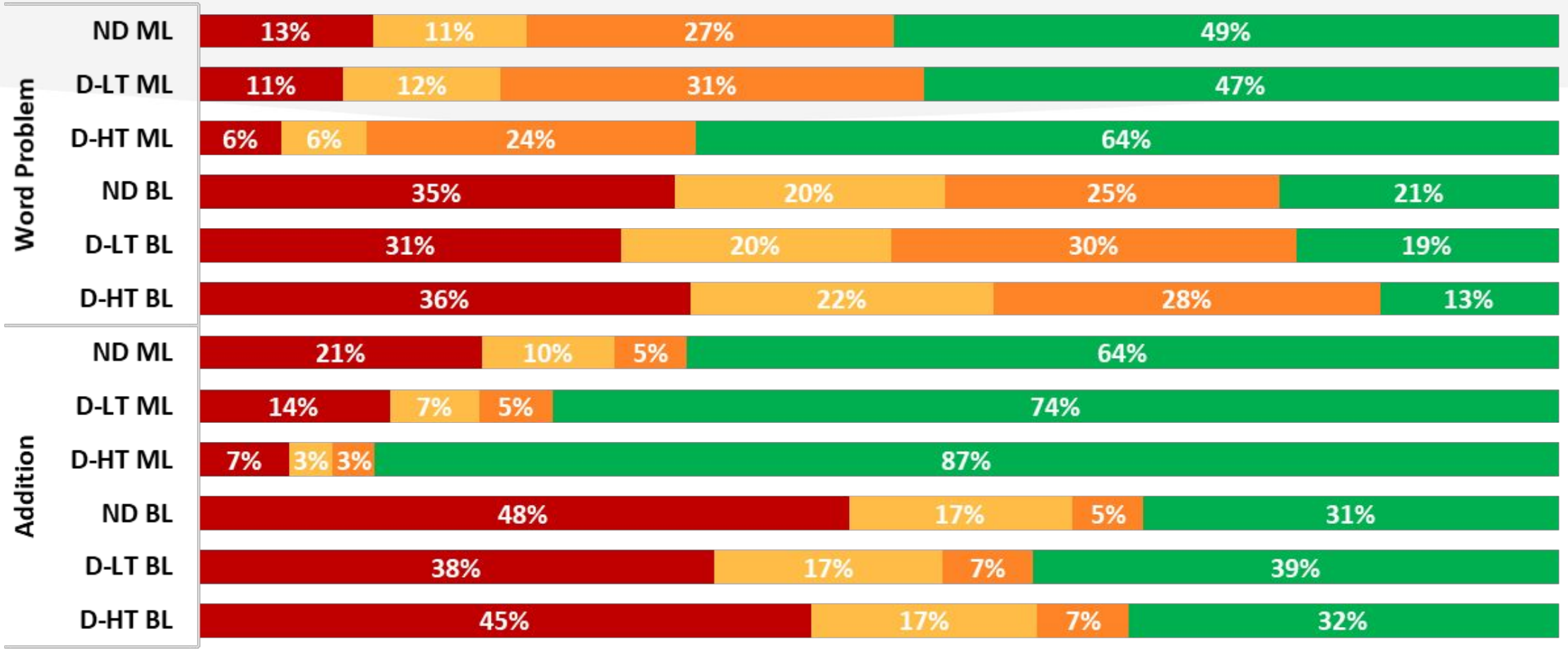
Grade 1 in the Midline Round

In the Number Recognition task, there's a noticeable shift toward the highest level (L3) in comparison with Cohort 1.



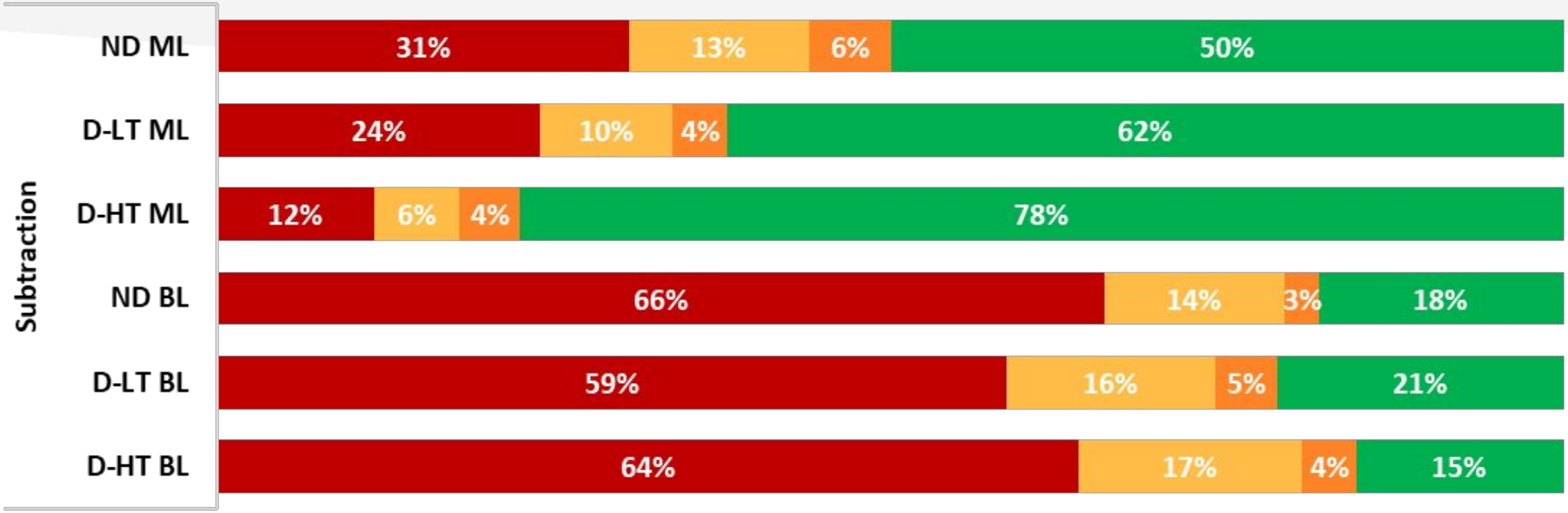
	L0	L1	L2	L3
Accuracy Tasks	0%	> 0% – 25%	> 25% – 50%	50% – 100%
Fluency Tasks	0	> 0 to (0.5 * Avg.)	> (0.5 * Avg.) to Avg.	> Avg. to Max.

In this cohort, most students are are at the L2 and L3 performance level in the word problem task, while the majority of them (<60%) are at the L3 level in the Addition task



	L0	L1	L2	L3
Accuracy Tasks	0%	> 0% – 25%	> 25% – 50%	50% – 100%
Fluency Tasks	0	> 0 to (0.5 * Avg.)	> (0.5 * Avg.) to Avg.	> Avg. to Max.

In High-Touch demo group, 78% of students were able to answer more than 50% of questions in this tasks indicating student's developing proficiency on high-order task



	L0	L1	L2	L3
Accuracy Tasks	0%	> 0% – 25%	> 25% – 50%	50% – 100%
Fluency Tasks	0	> 0 to (0.5 * Avg.)	> (0.5 * Avg.) to Avg.	> Avg. to Max.

No consistent performance patterns are seen at the school level across the baseline for Cohort 1 and the midline for Cohort 2 in the numeracy tasks, for the high-touch demo group; both top and bottom performing schools in the baseline scored much closer to the average in the midline for numeracy

D-HT

Task	BL Category	# Schools	# Students_BL	BL_Avg	# Students_ML	ML_Avg
Number Recognition (Fluency)	Top 10%	10	102	18.9	117	21.8
	Mid 10%	11	127	12.6	73	22.5
	Bottom 10%	10	117	5.6	94	22.1
Number Recognition (Accuracy)	Top 10%	10	102	48%	117	69%
	Mid 10%	11	127	32%	73	67%
	Bottom 10%	10	117	12%	94	69%
Counting in Bundles	Top 10%	10	102	34%	117	51%
	Mid 10%	11	127	16%	73	52%
	Bottom 10%	10	117	5%	94	55%
Missing Numbers	Top 10%	10	102	29%	117	46%
	Mid 10%	11	127	12%	73	43%
	Bottom 10%	10	117	3%	94	47%
Addition	Top 10%	10	102	65%	117	89%
	Mid 10%	11	127	31%	73	86%
	Bottom 10%	10	117	8%	94	80%
Subtraction	Top 10%	10	102	48%	117	81%
	Mid 10%	11	127	14%	73	83%
	Bottom 10%	10	117	4%	94	69%
Word Problems	Top 10%	10	102	44%	117	64%
	Mid 10%	11	127	24%	73	67%
	Bottom 10%	10	117	12%	94	62%

BL Category	Avg Decile Movement*
Top 10%	-4.8
Mid 10%	0.0
Bottom 10%	4.2

- The average scores of schools from all three performance categories in the baseline was between the first and third quartile in the midline round.

- The performance of schools in the bottom 10% category in the baseline round for numeracy improved in the midline round for all tasks

- The performance of most top performing schools in the baseline for numeracy dropped significantly in the midline round for all tasks

Mean school score is in the top 25% of school scores	Mean school score is in the second 25% of school scores	Mean school score is in the third 25% of school scores	Mean school score is in the bottom 25% of school scores

- Schools were ranked based on the mean school score in each task, i.e., the average of the scores of all students in the relevant grade in that school.
- An aggregate rank was created for each school across all literacy / numeracy tasks by adding the average ranks for each task, based on which the top 10%, middle 10%, and bottom 10% schools were selected.
- Average scores were calculated by taking the simple average of the mean school score in that task for all schools in that performance category.

** The average decile movement is the average change in the deciles of all the schools in each performance category, from the baseline to the midline round, with deciles determined based on the aggregate rank of the school.*

Decile Movement from Baseline to Midline for the Top 10%, Middle 10%, and Bottom 10% Schools at the Baseline, for Grade 1 in the High-Touch Demo Group of Cohorts 1 and 2, in Numeracy

D-HT

Decile Movement from Baseline to Midline for the Top 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
1	1	4	-3
2	1	9	-8
3	1	6	-5
4	1	4	-3
5	1	5	-4
6	1	4	-3
7	1	8	-7
8	1	9	-8
9	1	5	-4
10	1	4	-3
Average	**1**	**5.8**	**-4.8**

Decile Movement from Baseline to Midline for the Middle 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
45	5	8	-3
46	5	6	-1
47	5	-	-
48	5	-	-
49	5	5	0
50	6	7	-1
51	6	-	-
52	6	9	-3
53	6	2	4
54	6	-	-
55	6	2	4
Average	**5.5**	**5.6**	**0.0**

Decile Movement from Baseline to Midline for the Bottom 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
90	10	7	3
91	10	2	8
92	10	1	9
93	10	8	2
94	10	10	0
95	10	9	1
96	10	1	9
97	10	10	0
98	10	1	9
99	10	9	1
Average	**10**	**5.8**	**4.2**

No consistent performance patterns are seen at the school level across the baseline for Cohort 1 and the midline for Cohort 2 in the numeracy tasks, for the low-touch demo group as well; both top and bottom performing schools in the baseline scored much closer to the average in the midline

D-LT

Task	BL Category	# Schools	# Students_BL	BL_Avg	# Students_ML	ML_Avg
Number Recognition (Fluency)	Top 10%	10	101	21.4	104	22.1
	Mid 10%	10	101	13.6	105	21.8
	Bottom 10%	10	98	5.7	93	18.6
Number Recognition (Accuracy)	Top 10%	10	101	55%	104	65%
	Mid 10%	10	101	30%	105	62%
	Bottom 10%	10	98	13%	93	62%
Counting in Bundles	Top 10%	10	101	34%	104	44%
	Mid 10%	10	101	20%	105	42%
	Bottom 10%	10	98	3%	93	37%
Missing Numbers	Top 10%	10	101	33%	104	39%
	Mid 10%	10	101	15%	105	41%
	Bottom 10%	10	98	5%	93	30%
Addition	Top 10%	10	101	68%	104	73%
	Mid 10%	10	101	37%	105	71%
	Bottom 10%	10	98	12%	93	68%
Subtraction	Top 10%	10	101	51%	104	58%
	Mid 10%	10	101	17%	105	58%
	Bottom 10%	10	98	6%	93	51%
Word Problems	Top 10%	10	101	55%	104	56%
	Mid 10%	10	101	28%	105	55%
	Bottom 10%	10	98	13%	93	46%

BL Category	Avg Decile Movement*
Top 10%	-4.3
Mid 10%	-0.1
Bottom 10%	3.3

- The average scores of schools from all three performance categories in the baseline was between the first and third quartile in the midline round.

- The performance of schools in the bottom 10% category in the baseline round for numeracy improved in the midline round for all tasks

- The performance of most top performing schools in the baseline for numeracy dropped significantly in the midline round for all tasks

Mean school score is in the top 25% of school scores	Mean school score is in the second 25% of school scores	Mean school score is in the third 25% of school scores	Mean school score is in the bottom 25% of school scores

- Schools were ranked based on the mean school score in each task, i.e., the average of the scores of all students in the relevant grade in that school.
- An aggregate rank was created for each school across all literacy / numeracy tasks by adding the average ranks for each task, based on which the top 10%, middle 10%, and bottom 10% schools were selected.
- Average scores were calculated by taking the simple average of the mean school score in that task for all schools in that performance category.
** The average decile movement is the average change in the deciles of all the schools in each performance category, from the baseline to the midline round, with deciles determined based on the aggregate rank of the school.*

Decile Movement from Baseline to Midline for the Top 10%, Middle 10%, and Bottom 10% Schools at the Baseline, for Grade 1 in the Low-Touch Demo Group of Cohorts 1 and 2, in Numeracy

D-LT

Decile Movement from Baseline to Midline for the Top 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
1	1	8	-7
2	1	5	-4
3	1	10	-9
4	1	2	-1
5	1	5	-4
6	1	6	-5
7	1	2	-1
8	1	6	-5
9	1	6	-5
10	1	3	-2
Average	**1**	**5.3**	**-4.3**

Decile Movement from Baseline to Midline for the Middle 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
46	5	8	-3
47	5	4	1
48	5	6	-1
49	5	4	1
50	5	10	-5
51	6	10	-4
52	6	8	-2
53	6	3	3
54	6	2	4
55	6	1	5
Average	**5.5**	**5.6**	**-0.1**

Decile Movement from Baseline to Midline for the Bottom 10% Schools at the Baseline

School Aggregate Rank	Decile at Baseline	Decile at Midline	Decile Movement
91	10	7	3
92	10	8	2
93	10	7	3
94	10	3	7
95	10	9	1
96	10	10	0
97	10	3	7
98	10	9	1
99	10	6	4
100	10	5	5
Average	**10**	**6.7**	**3.3**

Key Findings from the Follow-Up Qualitative Study

A mix of classroom-level and systemic factors identified as key aiding and hindering factors in programme implementation

1. Belief in inherent student capabilities and consistent checking of workbooks by ARPs might be making teachers conduct the You-do as a We-do

2. Teachers, ARPs and LLF members indicate pressure to achieve NIPUN Goals, which may lead to teachers and ARPs prioritising learning outcomes over structured pedagogy

3. There is a high focus on reading-related sub-sections in G2 classrooms, most likely because NIPUN Lakshya App assessments focus only on reading skills

4. Length of the numeracy lesson plan and amount of preparation required for the ‘Math Games’ section may be leading to teachers skipping it altogether

5. Most ARPs conducted classroom observations and spot assessments, but many key guidelines were not followed

6. Generic feedback from ARPs, as well as lack of demos and written feedback makes teachers think that ARPs offer ‘suggestions’, rather than ‘sahyog’

Key Takeaways from the Follow-Up Process Evaluation for LLF

What could be improved?

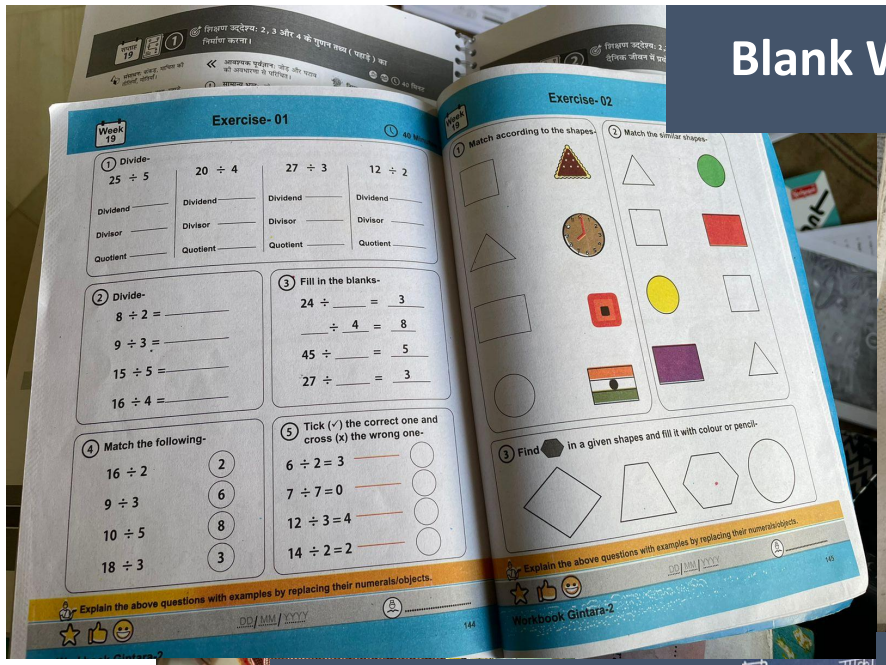
- Late delivery of materials is affecting teaching across all schools, including English-medium schools
- Teachers are finding it difficult to implement the entire lesson plan as per the suggested time in the TG
- In literacy, teachers find teaching matras/ half letters difficult, and Math games is being skipped across most classrooms.
- You-Do is being conducted as a We-Do in many classrooms.
- Although student are told whether their responses are correct/ incorrect, in many classrooms, they aren't being told 'why' their responses are correct or incorrect.
- In many cases, ARPs interrupted classroom teaching to ask students multiple questions.
- Many ARPs did not choose students randomly and did not complete the entire assessment with the number of required students during spot assessment.
- The average time spent on giving feedback to the teacher was about 1/4th (11 minutes) of the suggested time (40 minutes), and the generic nature of feedback as well as lack of demos make teachers think that it is ineffective.
- No tracking or utilization of CO, teacher feedback data mentioned, which may contribute to gaps in pedagogical practices.

Summary: Teachers find TLMs and discussion with other teachers in Sankul meetings particularly useful; certain support-based and systemic factors could be key barriers to success

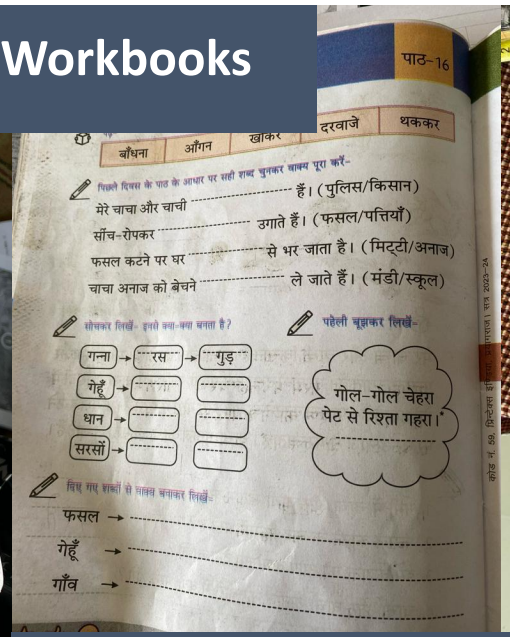
	Factors Aiding Program Success			Barriers to Program Success		
Teacher Mindset	Students learn faster in Numeracy	Absence of gender-biased actions		Perceived inherent variation in child ability	You-Do being done as We-Do	Gender-biased perceptions
Program Design	Engaging TLMs	Easy to use lesson plans		Length of the lesson plans		
Support & Monitoring	Frequent visits by ARPs	Usefulness of Sankul Meetings	Systematic tracking of student assessment data	Generic ARP feedback	Lack of Demos by ARPs	No analysis of CO, teacher feedback data
Systemic factors				Pressure to achieve NIPUN goals	Late Delivery of Resources	

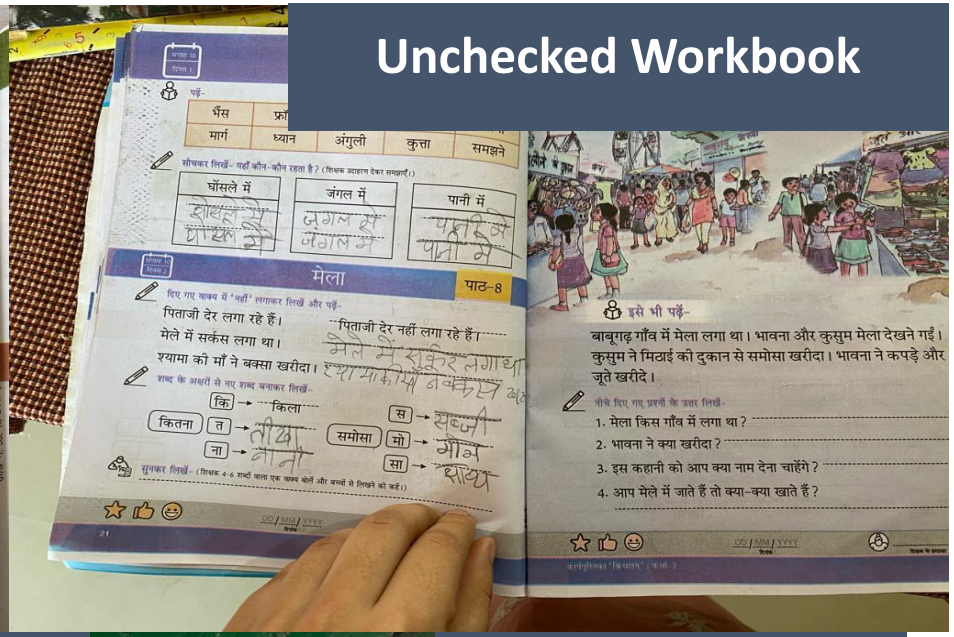
Legend: Highly Aiding | Moderately Aiding | Slightly Aiding | Mild Barrier | Moderate Barrier | Strong Barrier

Some Photos from field



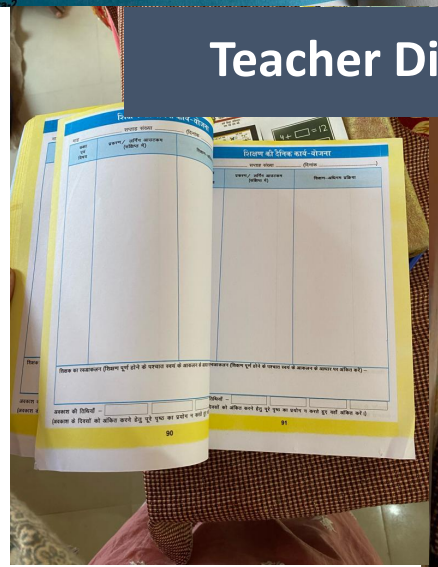
Blank Workbooks



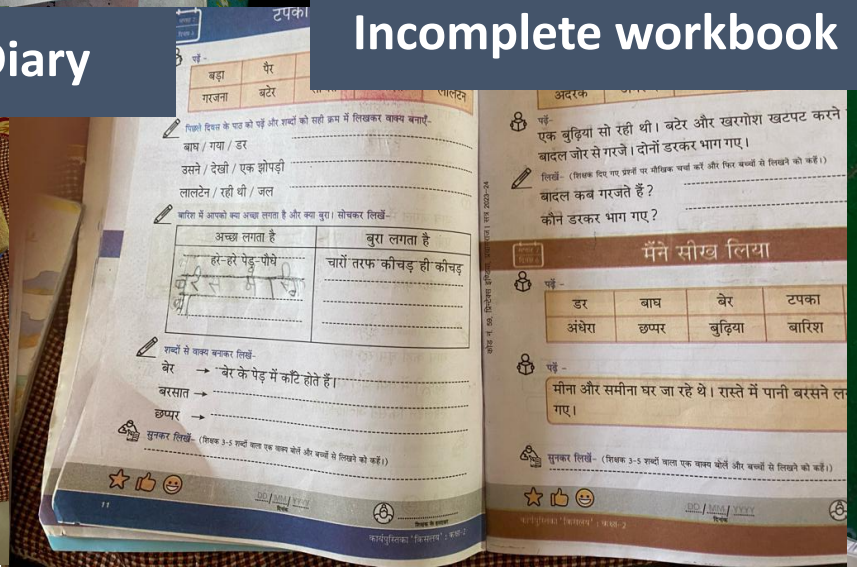


Unchecked Workbook

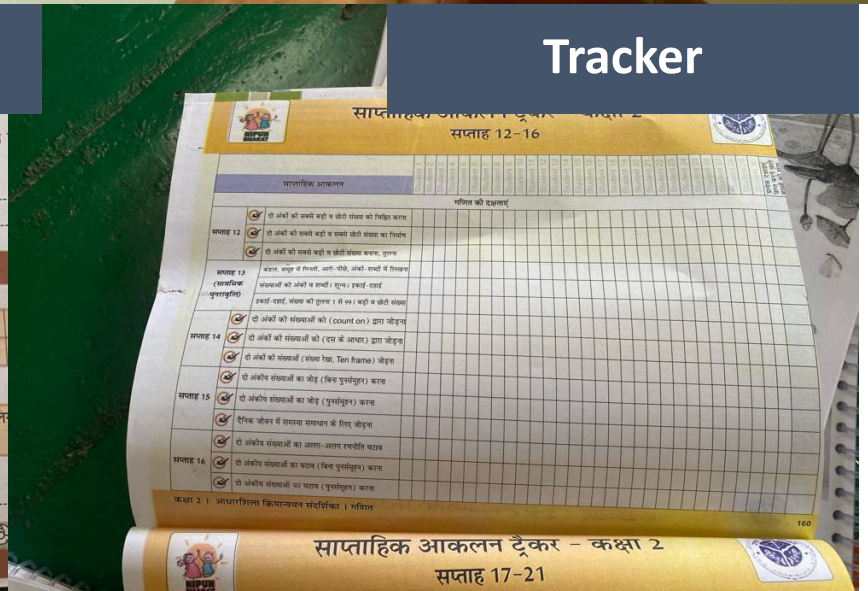
Teacher Diary



Incomplete workbook



Tracker



In numeracy classrooms, instances of teachers asking CFU questions rose significantly in Grade 2, but fell in Grade 1; there was also a drop in the instances of teachers giving clear instructions across both grades

In Literacy classrooms:

- A small positive shift seem in key FLN practices like highlighting the sound of a letter/ matra, showing strokes of the letter/ matra, asking open and close-ended questions, introducing new vocabulary, etc.
- Other teacher practices like giving clear instructions, monitoring student participation, giving feedback to students, and asking check-for-understanding (CFU) questions were found to be at levels similar to baseline.

In Numeracy classrooms:

- While asking CFU questions as a practice has improved slightly, the practice of giving clear instructions show a slight decline.
- Most other key teaching practices like monitoring student participation, sharing feedback with students, and key FLN section-related actions like using concrete objects/ real-life examples to demonstrate concepts, asking questions related to the activity have remained at the same level.

Literacy-related Findings from the Follow-Up Qualitative Study

Grade 1 Literacy Findings

Sub-Section Name	Conducted in ___ classrooms
OLD SEL	Many
OLD Story / Poem	-
OLD Story Vocab	-
OLD Story Discuss	Most
OLD Game	No
OLD WB	Most
PA LI/B	Many
PA LW	A Few
PA W/SR	Many
PA WB	Most
R Prac	Many
IR	-

OLD - Oral Language Development; SEL - Social and emotional Learning; PA - Phonological Awareness; LI - Letter Identification; B - Blending; LW - Letter Writing; SR - Sentence Reading; WB - Workbook; R Prac - Reading Practice; IR - Independent Reading

High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

Grade 1 Literacy Findings: Sub-Section Wise Overview

Category / School Code	OLD SEL	OLD Story / Poem	OLD Story Vocab	OLD Story Discuss	OLD Game	OLD WB	PA LI/B	PA LW	PA W/SR	PA WB	R Prac	IR
ROV_1	-	-	-		-			-				-
ROV_2	-	-	-		-	-				-		-
ROV_3		-	-	-		-						-
ROV_4	-	-	-		-			-				-
ROV_6		-	-	-		-						-

OLD - Oral Language Development; SEL - Social and emotional Learning; PA - Phonological Awareness; LI - Letter Identification; B - Blending; LW - Letter Writing; SR - Sentence Reading; WB - Workbook; R Prac - Reading Practice; IR - Independent Reading

Note:

i) For every sub-section, a list of teacher actions and corresponding student responses was created as indicators for the classroom observation tool, based on the teacher guides, and general best practices.

ii) These teacher actions were then studied for each subsection to identify how many of them a teacher be expected to perform on average in a classroom.

iii) Classrooms where teachers performed more than this average range of actions per sub-section were classified as 'High Fidelity', and those where teachers performed fewer than this range were classified as 'Low Fidelity', with the rest classified as 'Medium Fidelity'

		Grade 1	Grade 2
High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

Grade 2 Literacy Findings

Sub-Section Name	Conducted in ___ classrooms
OLD Ideal_R	Most
OLD R/S_Teach	-
OLD Vocab	Many
OLD R_Group	Many
OLD Discuss_Teach	Most
OLD Discuss_Group	-
OLD W	Many
WB Acti WB1	Many
WB Acti Ideal_R	-
WB Acti Story_Discuss	No
WB Acti WB2	Most
WB Acti R_Group	No

Sub-Section Name	Conducted in ___ classrooms
WB Acti Act1_Vocab You Do	Some
WB Acti Act2	Many
WB Acti Act3	Many
WB Acti Act4	Many
WB Acti Act5	Many
WB Acti NB_Act1	Many
WB Acti NB_Act2	Most
WB Acti Story_W	-
WB Acti IR	-
R Prac	Most
IR - P	-

OLD - Oral Language Development; R/S_Teach - Reading/ Sharing by Teacher; Vocab - Vocabulary; R_Group - Guided Reading in Student Groups; Discuss_Teach - Discussion based on poem/ story/ experiences with the teacher; Discuss_Group - Discussion in Student Groups; W - Writing Activity; WB - Workbook; WB Acti - Workbook-based Activities, Act - Activity, RC - Reading Practice; IR - Independent Reading

High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

Grade 2 Literacy Findings: Sub-Section Wise Overview (1/2)

Category/ School Code	OLD Ideal_R	OLD R/S_Teach	OLD Vocab	OLD R_Group	OLD Discuss_Teach	OLD Discuss_Group	OLD W	WB Acti WB1	WB Acti Ideal_R	WB Acti Story_Discuss	WB Acti WB2
ROV_5		-	-			-			-	-	
ROV_7		-	-			-			-	-	
ROV_8		-	-			-			-	-	
ROV_9		-		-		-			-	-	
ROV_10		-	-			-			-	-	
ROV_11		-		-		-			-	-	
ROV_12		-	-			-			-	-	
ROV_13		-	-			-			-	-	-
ROV_14		-	-			-			-	-	-
ROV_15	-	-	-	-	-	-	-	-	-	-	-

In 1 classroom, the teacher taught a lesson plan from Week 21 (revision week), and hence, it has been left blank; OLD - Oral Language Development; R/S_Teach - Reading/ Sharing by Teacher; Vocab - Vocabulary; R_Group - Guided Reading in Student Groups; Discuss_Teach - Discussion based on poem/ story/ experiences with the teacher; Discuss_Group - Discussion in Student Groups; W - Writing Activity; WB - Workbook; WB Acti - Workbook-based Activities, Act - Activity R Prac - Reading Practice; IR - Independent Reading

Key Finding: There is a high focus on reading-related sub-sections in G2 classrooms, most likely because NIPUN Lakshya App assessments focus only on reading skills

	Grade 1	Grade 2
High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

Grade 2 Literacy Findings: Sub-Section Wise Overview (2/2)

Category/ School Code	WB Acti R_Group	WB Acti Act1_Vocab You Do	WB Acti Act2	WB Acti Act3	WB Acti Act4	WB Acti Act5	WB Acti NB_Act1	WB Acti NB_Act2	WB Acti Story_W	WB Acti IR	R Prac	IR
ROV_5	-	-	-		-	-		-	-	-		-
ROV_7	-	-	-		-	-		-	-	-		-
ROV_8	-	-	-	-	-	-			-	-		-
ROV_9	-	-	-	-	-	-			-	-		-
ROV_10	-	-	-		-	-		-	-	-		-
ROV_11	-	-	-	-	-	-			-	-		-
ROV_12	-	-	-	-	-	-			-	-		-
ROV_13	-					-	-	-	-	-		-
ROV_14	-					-	-	-	-	-		-
ROV_15	-	-	-	-	-	-	-	-	-	-	-	-

OLD - Oral Language Development; R/S_Teach - Reading/ Sharing by Teacher; Vocab - Vocabulary; R_Group - Guided Reading in Student Groups; Discuss_Teach - Discussion based on poem/ story/ experiences with the teacher; Discuss_Group - Discussion in Student Groups; W - Writing Activity; WB - Workbook; WB Acti - Workbook-based Activities, Act - Activity R Prac - Reading Practice; IR - Independent Reading

Key Finding: There is a high focus on reading-related sub-sections in G2 classrooms, most likely because NIPUN Lakshya App assessments focus only on reading skills

Grade 1	Grade 2

High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

Numeracy-related Findings from the Follow-Up Qualitative Study

Numeracy Findings

Sub-section Name	No. of schools this subsection was conducted in, for	
	Grade 1	Grade 2
Mathematical Conversation - I Do + We Do (Through a story/ other activities)	Many	Most
Skill Building (1) - I Do + We Do	Many	Almost All
Skill Building (2) - I Do + We Do	No	All
Workbook Practice - You Do	Most	Almost All
Math Games - We Do + You Do	Some	A Few

Key Finding: Length of the numeracy lesson plan and amount of preparation required for the 'Math Games' section may be leading to teachers skipping it altogether.

In numeracy, findings have been reported in percentages and not a number count because the subsections to be done in each class varied depending on which lesson plan was being taught

High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

Numeracy Findings: Sub-Section Wise Overview

School Code	MC + Skills_I Do + We Do	MC + Skills_Skill (1)	MC + Skills_Skill (2)	WB_You Do	Math Games
ROV_1	-	-	-	-	-
ROV_2					
ROV_3	-	-	-	-	-
ROV_4					
ROV_5					
ROV_6					
ROV_7					
ROV_8					
ROV_9					
ROV_10					
ROV_11					
ROV_12					
ROV_13	-	-	-	-	-
ROV_14	-	-	-	-	-
ROV_15	-	-	-	-	-

6 G1 and 9 G2 classrooms were observed. In 5 classrooms across the two grades, teachers taught a lesson plan from Day 4 or Day 6 of the week, where they were either conducting their own activities or conducting assessments + remediation; MC = Mathematical Conversation, WB - Workbook

Grade 1	Grade 2

High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

Common Findings across Subjects

Findings across Subjects

1. Belief in inherent student capabilities and consistent checking of workbooks by ARPs might be making teachers conduct the You-do as a We-do

2. Teachers, ARPs and LLF members indicate pressure to achieve NIPUN Goals, which may lead to teachers and ARPs prioritising learning outcomes over structured pedagogy

3. While many teachers expressed positive opinions about the TG, late delivery of programme materials is still a key issue

4. No significant gender-biased actions observed across classrooms, however, some teachers seem to think of girls as more obedient and boys as more confident

Findings related to ARP support from the Follow-Up Qualitative Study

Findings related to ARP Support: Classroom observations and spot assessments conducted by most ARPs, Feedback to the teacher can be strengthened

OVERVIEW		
Activity Name	**Conducted by ___ ARPs**	**Average Time Spent (In minutes)**
Classroom Observation	Almost All	24 (Range - 5 to 42)
Spot Assessment	Most	27 (Range - 15 to 70)
Conversation with Teacher	Many	11 (Range - 5 to 27)
Conversation with HM	Most	29 (Range - 5 to 60)

Note:

i) For each section, a list of ARP actions and corresponding teacher responses was created as indicators for the joint visit tool, based on the supportive supervision guide, and general best practices.

ii) Joint visits where ARPs performed 0 - 30% of the expected actions were classified as 'Low Fidelity', and those where they performed more than 60% of the expected actions were classified as 'High Fidelity'. The rest were classified as 'Medium Fidelity'.

JOINT VISIT FIDELITY TO EXPECTED ACTIONS				
ARP Code	**Classroom Observation**	**Spot Assessment**	**Conversation with Teacher**	**Conversation with HM**
ROV_1				
ROV_2				
ROV_3				
ROV_4				
ROV_5				
ROV_6				
ROV_7				
ROV_8				
ROV_9				

High fidelity **Medium Fidelity** **Low Fidelity**

Findings related to ARP Support

1. While all ARPs said that they visit the mandated 30 schools in a month, lack of sufficient time and drawbacks of NIPUN Lakshya App reported as impediments to effective school visits

2. Most ARPs conducted classroom observations and spot assessments, but many key guidelines were not followed

3. Generic feedback from ARPs, as well as lack of demos and written feedback makes teachers think that ARPs offer ‘suggestions’, rather than ‘sahyog’

4. Data collection seems to be a high-priority for many ARPs, and spot assessment data is used to provide differentiated support

Baseline v/s Now - What has changed in classrooms?

The 15 schools from the qualitative study did not show significant improvement on average in most SLO literacy tasks from baseline to midline, except ORF, with clear school-wise performance patterns visible

School Code	Listening Comprehension	Oral Vocabulary	Initial Sound Identification	Letter Reading (Fluency)	Letter Reading (Accuracy)	Word Reading (Fluency)	Word Reading (Accuracy)	Non-Word Reading (Fluency)	Oral Reading Fluency	Reading Comprehension Passage 1	Letter Writing	Word Writing
ROV_1	-0.41	-2.43	-1.57	-0.93	-0.28	-0.44	-0.53	-0.53	-0.92	-1.09	-0.42	-0.47
ROV_2	0.23	0.23	-0.12	0.30	0.28	0.63	0.10	0.67	0.97	0.45	0.74	0.52
ROV_3	0.23	0.61	0.37	0.81	0.15	0.82	0.59	1.05	1.26	1.15	0.16	1.01
ROV_4	-0.98	-0.91	-2.16	-0.80	-0.90	-0.57	-0.43	-0.43	-0.31	-0.61	-2.39	-0.55
ROV_5	0.31	0.23	-0.23	0.71	0.80	0.69	0.81	0.99	0.98	0.98	0.68	0.37
ROV_6	0.31	-0.19	-0.66	0.04	-0.59	0.57	0.52	1.04	1.36	0.61	-0.55	0.18
ROV_7	-0.32	-0.33	-0.54	-0.39	0.17	-0.43	-0.20	-0.37	-0.62	-0.67	-0.77	-0.23
ROV_8	-0.17	-0.19	0.04	-0.27	-0.49	0.23	0.00	0.17	0.33	0.45	-0.77	0.14
ROV_9	0.38	-0.06	-0.04	0.31	-0.34	0.16	0.71	-0.32	0.87	1.14	-0.30	0.11
ROV_10	0.77	-0.18	0.75	1.11	-0.11	1.05	0.23	1.27	2.03	0.90	-0.19	0.80
ROV_11	0.15	-0.08	0.59	0.65	-0.20	0.53	0.66	0.73	0.88	1.09	-0.52	0.22
ROV_12	-0.74	-0.19	-1.89	-0.07	-1.24	-0.27	-1.56	0.25	0.73	-0.29	-1.20	-0.05
ROV_13	-0.17	-0.70	-1.53	-1.67	-1.19	-1.97	-2.31	-1.99	-1.57	-1.60	-1.02	-2.21
ROV_14	0.42	0.80	-0.74	-0.26	-0.29	-0.10	0.41	0.11	0.04	0.36	-1.02	0.04
ROV_15	0.35	-0.03	-0.69	0.76	1.03	0.51	0.93	0.46	0.44	0.94	0.44	0.39
All 15 Schools	**0.05**	**-0.20**	**-0.51**	**0.03**	**-0.21**	**0.11**	**0.02**	**0.22**	**0.45**	**0.27**	**-0.46**	**0.04**

** All performance data on this table is in DiD effect size, calculated using the formula: [avg_delta_demo (Δi) - avg_delta_non-demo (Δc)] / SD_pooled (Pooled Standard Deviation)*

These schools also did not show a significant performance change on average in most SLO numeracy tasks from baseline to midline, with school-wise performance patterns visible in numeracy as well

School Code	Number Recognition (Fluency)	Number Recognition (Accuracy)	Counting in Bundles	Missing Number	Addition (Accuracy)	Subtraction (Accuracy)	Word Problems
ROV_1	0.02	-1.06	-1.06	-1.05	-0.59	-0.96	-0.02
ROV_2	0.06	0.01	0.07	-0.57	0.15	0.00	0.40
ROV_3	0.91	-0.32	0.32	0.10	-0.57	-0.17	-0.14
ROV_4	-0.26	-1.03	-0.90	-0.69	-1.17	-1.09	-1.75
ROV_5	0.45	0.24	0.56	-0.01	-0.12	0.35	0.62
ROV_6	-0.15	-0.05	0.72	0.65	0.07	-0.09	-0.28
ROV_7	-0.52	-0.43	-0.17	-0.43	-0.32	-0.70	0.19
ROV_8	0.40	0.10	0.64	0.82	0.82	0.44	0.56
ROV_9	0.07	0.30	0.86	0.34	0.08	0.50	0.79
ROV_10	0.33	0.48	0.48	0.55	0.52	0.72	0.08
ROV_11	0.13	-0.54	-0.17	-0.43	0.27	0.62	-0.02
ROV_12	-0.05	-0.17	0.56	-0.20	-0.90	-1.36	-0.44
ROV_13	-0.52	-1.53	-1.71	-1.55	-1.64	-1.44	-1.59
ROV_14	0.46	0.78	0.69	0.82	0.22	0.47	0.46
ROV_15	0.55	0.30	0.15	0.42	0.27	0.26	0.46
All 15 Schools	**0.15**	**-0.16**	**0.10**	**-0.04**	**-0.19**	**-0.15**	**-0.02**

** All performance data on this table is in DiD effect size, calculated using the formula: [avg_delta_demo (Δi) - avg_delta_non-demo (Δc)] / SD_pooled (Pooled Standard Deviation)*

Question by LLF team

Competency and how to read the data (Progression of skills)

- Slide 11: This slide needs some clarity and info to read the data as it does not mention what is the level of counting and number recognition and how to read the data for counting and number recognition. Some need answers to the following questions.

Do you mean the level of questions? It was counting upto 20 and Number recognition up to 99 for G1 and upto 999 for G2.

- Can you provide a detailed description of what constitutes different levels of competency in counting and number recognition fluency?

Please specify this? Do you mean the difficulty level of different levels of items? The tools and levels of competency are aligned with NIPUN goals

Performance Interpretation:

- How should we interpret the average performance figures in terms of competency levels? For example, what does an average count per minute of 108.91 in High-Touch sites indicate about students' counting proficiency?

  Students were supposed to count upto 20 cwpm but they are counting 108 cwpm on average so proficiency is really high. Please consider that the assessment was conducted after 9 months of being in school and part of intervention so such results are expected.

- Are there specific thresholds or cut-off scores that distinguish between different levels of proficiency (e.g., basic, proficient, advanced)
- Were there any notable patterns or trends observed in the data that might indicate specific areas of strength or weakness among the students?

Question by LLF team

- How much progress did students make from the baseline to the midline in counting and number recognition?
- Slide 32: details data for counting and number recognition (fluency is missing while counting in bundles and number recognition (Accuracy is given)

  In Number recognition fluency, proportion of zero scorers is less than 1% for high-touch group and low-touch group. Most of the students in L1 level while in baseline, higher proportion of students were zero scorers and around 45% were L4 level.

  Counting was not presented for G2 in midline so this task was not compared.

- Number recognition (Accuracy) data is showing 98% in High touch. It is quite high. So eager to know the level or range of numbers for competencies (like 1-20 or 21-99 or else) and how many numbers (items) were given to check the accuracy. Similarly addition level 1 data is also very high. It is 99%. What is L3 in number recognition?

  For G2, the numbers recognition included the identification of numbers upto 3 digit. 10 items were kept for accuracy task and 40 in fluency task

  In Addition, the performance is 92% for the High-touch group. It was single-digit addition facts and a medium difficulty level task for this grade. Hence a performance is observed.

Question by LLF team

Response:

- Slide 35: what does it mean by midline average score? how to read it?

  It is the total scores by all students divided by total number of students.

- Slide 58 : Have both classes 1 and 2 been observed for the same teacher? Skill building should be written as skill teacher and why this has been split in to parts 1 and 2.
- Slide 59: slide 59 is saying that- "*6 G1 and 9 G2 classrooms were observed. In 5 classrooms across the two grades, teachers taught a lesson plan from Day 4 or Day 6 of the week, where they were either conducting their own activities or conducting assessments + remediation". Data is showing less use of the workbook. I think we should not do observation on day 6 as there is no worksheet for day 6.*
- Slides 63: data is telling that boys outperform in addition and subtraction than girls but what about other competencies?

Annexure

Annexure 1: Executive Summaries

Detailed Summary of Findings (1/3)

About the study: The study has a quasi-experimental design, with demonstration (demo) and non-demonstration (non-demo) sites matched based on similar characteristics, and covered 3,190 Grade 1 students and 3,192 Grade 2 students from 327 schools in the midline round.

Findings for Cohort 1:

1. The average performance of the demo groups showed greater improvements than the non-demo group, especially the High-Touch demo group, whose improvement in performance surpassed both the Low-Touch demo and non-demo groups across all literacy tasks from baseline to midline.
   a. The **performance of the High-Touch demo group in literacy showed an effect size >= 0.7 SD for 7 out of 12 tasks**, compared to the non-demo group, and between 0.15 and 0.28 SD for the remaining tasks. In comparison, **the Low-Touch demo group showed an effect size between 0.57 and 0.66 SD for 3 tasks, between 0.26 and 0.39 SD for 4 tasks**, and either a negative or a negligibly positive effect size for the remaining tasks.
   b. The increase in performance for the High-Touch demo group is between 22 to 84 percentage points for accuracy-based tasks, compared to 15 to 74 percentage points for the Low-Touch group and 15 to 62 percentage points for the non-demo group.
   c. For fluency-based tasks, the increase ranges from 28 to 58 correct words per minute (cwpm) for the High-Touch demo group, compared to 21 to 46 cwpm for the Low-Touch group and 17 to 36 cwpm for the non-demo group.

2. **Performance on higher-order tasks like Non-Word Reading, Reading Comprehension, and Oral Reading Fluency (ORF) has significantly improved** for both the demo and non-demo groups; however, the gain is higher in the demo groups.
   a. For example, in ORF, students from the High-Touch demo group were able to read 58 cwpm in the midline assessment, which is a significant gain from the baseline, where they were able to read only 3 cwpm on average. In comparison, the performance of the Low-Touch demo group increased from 4 to 47 cwpm, and the non-demo group's performance increased from 3 to 32 cwpm.

3. **Performance on lower-order tasks such as Letter and Word Reading improved notably** in the demo groups compared to the non-demo group in this round. The High-Touch demo group saw an increase of around 71 percentage points for these tasks, while the Low-Touch demo and non-demo groups experienced an increase of around 59 percentage points.

Detailed Summary of Findings (2/3)

4. Overall, there is an increase in the average performance of all three groups from baseline to midline in all numeracy tasks.
   a. **The performance of the High-Touch demo group in numeracy showed an effect size between 0.34 and 0.59 SD**. In comparison, **the Low-Touch demo group showed an effect size of 0.36 SD for the Number Recognition (Fluency) task**, and between 0.02 and 0.14 SD for the remaining tasks.
   b. The increase in performance for accuracy-based tasks for the High-Touch demo group is 29 to 67 percentage points, for the Low-Touch demo group, it is 27 to 51 percentage points, and for the non-demo group, it is 26 to 47 percentage points.

5. **The demo groups showed the highest increase in average performance for tasks such as Number Recognition (accuracy), Counting in Bundles, Addition, and Subtraction**. The difference in average performance for these tasks from baseline to midline was 55, 61, 58, and 67 percentage points, respectively, for the High-Touch demo group; 48, 49, 44, and 50 percentage points, respectively, for the Low-Touch demo group; and 47, 45, 44, and 47 percentage points, respectively, for the non-demo group.

6. **There is a notable performance gap between boys and girls on the Addition (4 percentage points) and Subtraction (4 percentage points) tasks** in the midline round across both groups, with boys outperforming girls. This gap is significantly higher compared to other tasks.

Findings for Cohort 2:

1. **The demo groups have performed significantly better than the non-demo group across most literacy and a few numeracy tasks for Cohort 2**, which is a consequence of the first assessment for this cohort being conducted almost at the end of Grade 1, and also shows the positive impact of the intervention.
   a. **The High-Touch demo group performed significantly better** (i.e., difference in average score >= 10 percentage points) **than the non-demo group across almost all literacy and numeracy tasks,** with a comparable performance only in the Listening Comprehension and Oral Vocabulary tasks in literacy, and the Number Recognition (Accuracy) task in numeracy.

Detailed Summary of Findings (3/3)

b. The Low-Touch demo group performed much better than the non-demo group in the Letter Reading, Word Reading, Non-Word Reading, ORF and Reading Comprehension tasks in literacy, and the Number Recognition (Fluency) task in numeracy.

c. The High-Touch demo group also performed much better than the Low-Touch demo group in all the literacy tasks except Listening Comprehension and Oral Vocabulary, as well as the Missing Number, Subtraction, and Word Problems tasks in numeracy.

2. **In this round, Cohort 2's average performance surpassed Cohort 1's in the previous round in all literacy tasks, across both demo and non-demo sites.** Specifically, the High-Touch demo group of Cohort 2 shows a difference in performance ranging from 15 to 67 percentage points compared to Cohort 1's High-Touch demo group. Similarly, the difference in performance for the Low-Touch and non-demo groups of Cohort 2, compared to Cohort 1's equivalent groups, ranges from 4 to 57 percentage points and 9 to 32 percentage points, respectively, across almost all tasks.

3. In the current round, the proportion of zero scorers in literacy tasks decreased compared to the previous round for all three groups. Notably, in this first round for Cohort 2, **the proportion ranged from 3% to 52% for accuracy-based tasks**. In contrast, in the previous round, the proportion of zero scorers for Cohort 1 was higher, ranging from 10% to 86% for all three groups.

4. There is an **increase in the average performance of the High-Touch demo groups in this cohort compared to the group's performance in the last round in Numeracy**. The change in performance for High-Touch is between 19 to 50 percentage points for accuracy-based tasks. This change in performance for the Low-Touch demo is between 14 to 44 percentage points for all accuracy-based tasks.

5. **The gap in the performance of boys and girls has increased specifically in tasks like Addition and Subtraction** in this round across demo and non-demo groups. The gap is 5 and 6 percentage points, respectively, for these tasks.

Key Takeaways from the Follow-Up Process Evaluation

What's working well?

- Many teachers expressed positive opinions about the TG and the TLMs.
- Many key teacher practices seem to have improved since the baseline. Some of these include asking check-for-understanding questions, giving feedback to students, using TLMs/ stories/ examples given in the TG, and giving clear instructions.
- High focus on reading-related sub-sections in G2 classrooms may be linked to significant gains in SLO performance.
- No significant gender-biased actions observed across classrooms.
- All ARPs and many teachers reported that ARPs visit every month, and spend 2 hours at school.
- Most ARPs conducted at least 3/4 activities during school visits - classroom observation, spot assessment, and conversation with the HM and other teachers.
- Many teachers find the monthly Sankul meetings helpful to engage in discussions about various teaching practices.
- Student assessment data is systematically tracked at the ARP, Sankul, Block level, which may be aiding student outcomes.

Key Takeaways from the Follow-Up Process Evaluation

What could be improved?

- Late delivery of materials is affecting teaching across all schools, including English-medium schools
- Teachers are finding it difficult to implement the entire lesson plan as per the suggested time in the TG
- In literacy, teachers find teaching matras/ half letters difficult, and Math games is being skipped across most classrooms.
- You-Do is being conducted as a We-Do in many classrooms.
- Although student are told whether their responses are correct/ incorrect, in many classrooms, they aren't being told 'why' their responses are correct or incorrect.
- Teachers, ARPs, and LLF members indicate pressure to achieve NIPUN goals, which might lead to actors prioritising learning outcomes over pedagogical principles.
- In many cases, ARPs interrupted classroom teaching to ask students multiple questions.
- Many ARPs did not choose students randomly and did not complete the entire assessment with the number of required students during spot assessment.
- The average time spent on giving feedback to the teacher was about 1/4th (11 minutes) of the suggested time (40 minutes), and the generic nature of feedback as well as lack of demos make teachers think that it is ineffective.
- While many ARPs noted technical issues with the App, some also highlighted their concerns with questions repeating on App assessments, leading to students memorising answers, rather than answering with understanding.
- No tracking or utilization of CO, teacher feedback data mentioned, which may contribute to gaps in pedagogical practices.

Annexure 2: Research Questions

Key Research Questions for the Evaluation of Foundational Learning Programs in Uttar Pradesh

S. No.	Research Questions
1	**What is the impact of the implementation of the FLN Programmes in the demonstration sites, vis-à-vis the comparison geography, on student learning outcomes?**
2	**How is the programme implemented vs designed, and what are the shifting classroom practices along with factors that aided or hindered implementation?**
3	**What are the design and implementation successes across different states/demonstration sites to indicate transferability for scale-up within and across states?**

Key Research Questions for the Follow-Up Process Evaluation

S. No.	Primary Research Questions	Secondary Research Questions
1	**How are the FLN programmes implemented in the demonstration sites?**	How accurately are the programme goals understood by teachers and ARPs?
		What is the degree of implementation fidelity of the programme at various levels?
		What factors aided or hindered the implementation of the programme at various levels? Why are certain teachers, ARPs implementing the programme with higher/ lower fidelity than others?
2	**What kind of shifts are visible in the roles and responsibilities of the Academic Resource Persons (ARPs) in the demonstration sites?**	How do ARPs understand their role in supporting teachers, and what is teachers' understanding of ideal support?
		What changes can be observed in the frequency of the academic support and mentorship provided by ARPs to teachers?
		What changes can be observed in the quality of the academic support and mentorship provided by ARPs to teachers?
3	**What kind of shifts are visible in the use of data by government education officials to achieve foundational learning in the demonstration sites?**	How often do ARPs collect data? How do they analyse and utilise the data they collect?
		What factors aided or hindered the use of data by ARPs to improve implementation of the programme?

Annexure 3: Tools and Analysis Approach

Learning levels in Literacy (Hindi) were measured through student learning outcome (SLO) assessments aligned with the globally accepted Early Grade Reading Assessment (EGRA) tools

EI's Foundational Literacy Assessment Tool covers listening, speaking, reading, and writing through 14 sub-tasks that evaluate both accuracy (correctness with which students answer irrespective of the time taken) and fluency (correct answers per minute). The tool is aligned with the global EGRA framework, and tailored to the local context in Uttar Pradesh.

1

Listening Comprehension

2

Oral Vocabulary

3

Initial Sound Identification

4

Letter Reading (Accuracy)

5

Letter Reading (Fluency)

6

Word Reading (Accuracy)

7

Word Reading (Fluency)

8

Non-Word Reading (Fluency)

9

Oral Reading Fluency

10

Reading Comprehension 1

11

Reading Comprehension 2

12

Dictation - Letter Writing

13

Dictation - Word Writing

14

Dictation - Sentence Writing

Similarly, learning levels in Numeracy were measured through student learning outcome (SLO) assessments aligned with the globally accepted Early Grade Mathematics Assessment (EGMA) tools

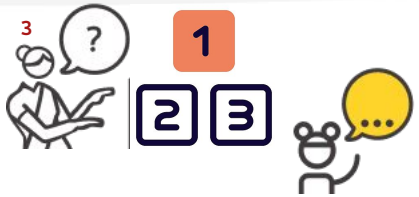
EI's Foundational Numeracy Assessment Tool covers number and shape recognition, counting, and basic operations through 11 sub-tasks that evaluate both accuracy (correctness with which students answer irrespective of the time taken), and fluency (correct answers per minute). The tool is aligned with the global EGMA framework, and tailored to the local context in Uttar Pradesh

1 

Number Recitation

2 

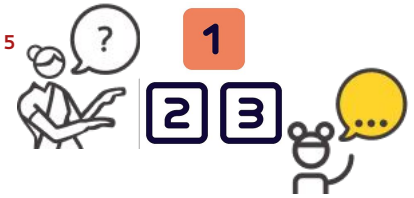
Counting

3 

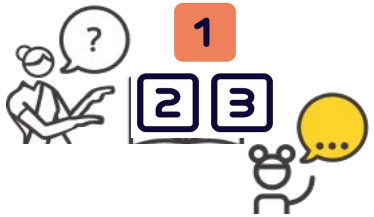
Number Recognition (Accuracy)

4 

Number Recognition (Fluency)

5 

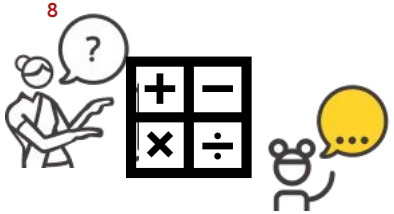
Number Comparison

6 

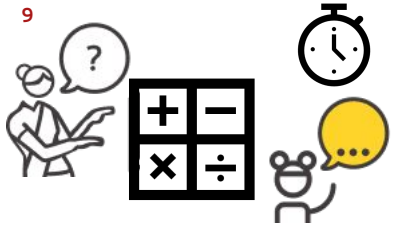
Counting in Bundles

7 

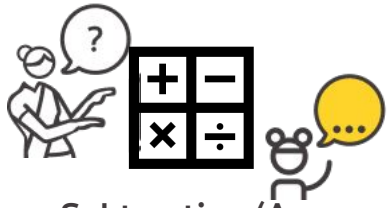
Missing Number

8 

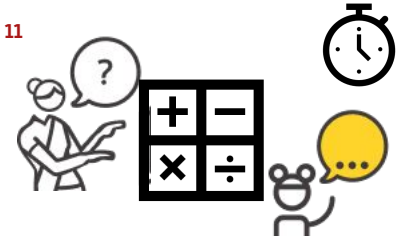
Addition (Accuracy)

9 

Addition (Fluency)

10 

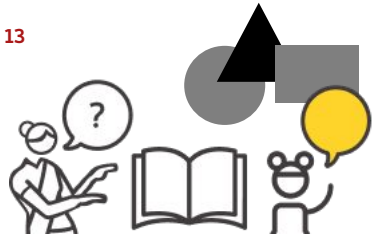
Subtraction (Accuracy)

11 

Subtraction (Fluency)

12 

Word Problems

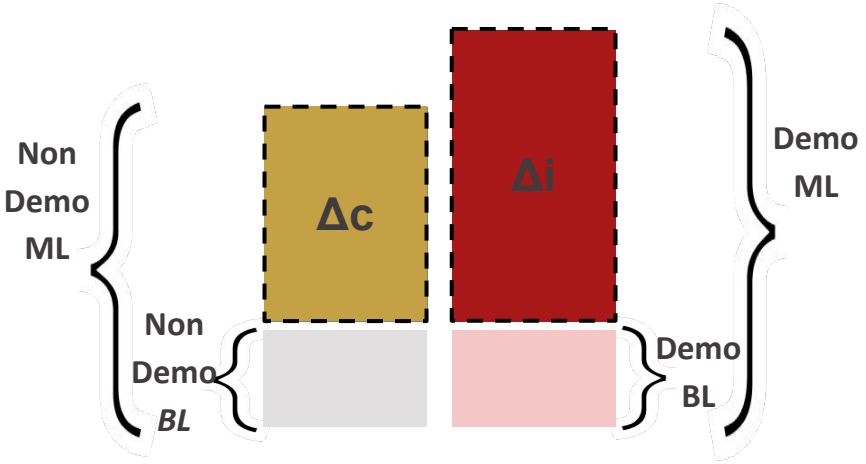
13 

Shape Recognition

The Difference-in-Differences (DiD) method was used to quantify the learning gains made by the demonstration group over the non-demonstration group from the baseline round to the midline round of the student learning outcome (SLO) assessments

Assessment data for each sub-task was analyzed based on the difference-in-differences approach to ascertain the magnitude of impact of the intervention by Room to Read (RTR)

Learning Measurements for Demo and Non-Demo groups, at Baseline (BL) and Midline (ML)



Difference-in-Differences (DiD)

DiD

DiD = Δi - Δc

DiD Effect Size

$$\sigma = \frac{\text{DiD}}{\text{Pooled Standard Deviation}}$$

In-depth Interview with Teachers - Tool Summary

Section	Content	Mapping to Tertiary RQs/ Details
A	Opening, Consent	1. Introduction to interviewers and Research objectives 2. Broad sections to be covered in the next 40-45 minutes 3. Verbal consent 4. Rapport building
B	Programme and subject-related beliefs	1. Do teachers prefer teaching one subject over another? If so, why? 2. Do teachers and ARPs believe that the structured pedagogy approach will lead to the intended student learning outcomes? Why or why not?
C	Classroom practices	1. Based on the classroom observation, why do or why don't teachers follow exactly what is prescribed in the teacher guide? 2. To what extent are the practice/ You-Do sections implemented as prescribed in the teacher guide? Why or why not? 3. To what extent are teachers adhering to the suggested time mentioned in the teacher guide? Why or why not? 4. Are the teacher guides, student workbooks, and TLMs available and utilized? Why or why not? 5. To what extent are teachers following the key activities under assessment-informed-instruction? Why or why not?
D	Support from ARPs + Implementation Partners	1. Based on the number of schools mapped to each ARP, how often do ARPs visit each school? 2. On average, how long are these visits? 3. What are the top 2-3 activities that ARPs conduct on school visits? Differentiate these in terms of compliance and pedagogical support related activities. 4. What are the key 3-4 ways in which ARPs analyze and use the data they collect? Does this match what is prescribed in the supportive supervision guide? 5. Is the support provided through coaching activities relevant, specific, and action-oriented? Why or why not? 6. How is the data collected by ARPs used in cluster-level and block-level meetings? 7. To what extent is the working relationship between teachers and ARPs authoritative or supportive in nature? What changes have been seen in this over time? 8. What kind of support do teachers want from ARPs? 9. How satisfied are teachers with the overall support given by ARPs? Why or why not? What changes have been seen in this over time? 10. How effective do teachers find the support offered by on-ground implementation partners? What are the key reasons for their response?
E	Closing	1. Gratitude 2. Reiterate the point about confidentiality 3. Answer any questions the respondent has.

In-depth Interview with ARPs - Tool Summary

Section	Content	Mapping to Tertiary RQs/ Details
A	Opening, Consent	1. Introduction to interviewers and Research objectives 2. Broad sections to be covered in the next 40-45 minutes. 3. Verbal consent 4. Rapport building
B	Awareness of + Belief in the Programme	1. Are ARPs aware of the goals of NIPUN UP/ Mission Ankur? 2. Are the teacher guides, student workbooks, and TLMs available? 3. Do ARPs believe that structured pedagogy approach will lead to the intended student learning outcomes? Why or why not?
C	On-ground support for teachers	1. Based on the number of schools mapped to each ARP, how often do ARPs visit each school? 2. On average, how long are these visits? 3. How often are cluster-level and block-level meetings organized? 4. What are the top 2-3 activities that ARPs conduct on school visits? Differentiate this in terms of compliance with pedagogical support activities. 5. To what extent do ARPs follow exactly what is prescribed in the supportive supervision guide? Why or why not? 6. Is the support provided through the activities relevant, specific, and action-oriented? Why or why not? 7. Do ARPs provide stronger support in one subject over another? If yes, why? 8. To what extent are teachers following exactly what is prescribed in the teacher guide?
D	Collection and Utilization of Data	1. What kind of data do ARPs collect, and how often? Does this match what is prescribed in the supportive supervision guide? 2. Which components of the data collected are visible to ARPs? Can they access this instantly or afterwards? 3. What are the key 3-4 ways in which ARPs analyse and use the data they collect? 4. Specifically, how is the data collected used in cluster-level and block-level meetings?
E	Definition of successes in their role + Support from RTR	1. How do ARPs define their role under NIPUN UP/ Mission Ankur? 2. How effective do teachers and ARPs find the support offered by on-ground implementation partners? What are the key reasons for their response?
F	Closing	1. Gratitude 2. Reiterate the point about confidentiality 3. Answer any questions the respondent has

Annexure 4: Findings from Cohort 1

A significant increase in average student performance is observed from the baseline round across both demo and non-demo sites. However, the increase in the High-Touch demo sites is higher. Additionally, there's a noticeable decrease in standard deviation from baseline to midline, indicating less variability in the scores.

Task	Unit	Midline - Average			Midline - SD			Baseline-Average			Baseline - SD		
		ND	D-HT	D-LT	ND	D-HT	D-LT	ND	D-HT	D-LT	ND	D-HT	D-LT
Listening Comprehension	Percentage	77%	81%	77%	27%	23%	25%	62%	59%	62%	34%	32%	32%
Oral Vocabulary	Percentage	94%	96%	95%	8%	7%	9%	95%	96%	96%	11%	8%	8%
Initial Sound Identification	Percentage	68%	82%	69%	43%	35%	42%	12%	16%	19%	30%	33%	35%
Letter Reading (Accuracy)	Percentage	77%	91%	85%	28%	13%	20%	29%	35%	42%	35%	35%	36%
Letter Reading (Fluency)	Count per minute	51.6	70.0	62.1	25.1	17.8	23.4	14.8	16.9	20.1	17.3	15.7	16.4
Word Reading (Accuracy)	Percentage	70%	90%	83%	31%	14%	21%	8%	13%	14%	19%	23%	24%
Word Reading (Fluency)	Count per minute	21.6	34.9	29.3	15.2	13.6	14.5	5.1	6.2	7.5	12.1	9.2	9.7
Non-Word Reading (Fluency)	Count per minute	19.9	33.2	27.8	14.7	12.2	14.1	2.7	3.5	4.5	6.9	6.5	7.5
Oral Reading Fluency	Count per minute	34.8	61.1	50.4	30.3	27.0	29.2	2.6	3.0	4.0	10.5	8.9	10.9
Reading Comprehension Passage 1	Percentage	60%	88%	78%	44%	24%	35%	2%	4%	4%	14%	18%	19%
Reading Comprehension Passage 2**	Percentage	49%	75%	62%	41%	29%	35%	-	-	-	-	-	-
Letter Writing	Percentage	71%	80%	71%	32%	25%	30%	22%	25%	32%	32%	34%	36%
Word Writing	Percentage	58%	77%	71%	31%	20%	23%	8%	9%	11%	21%	23%	26%
Sentence Writing**	Percentage	60%	69%	63%	25%	25%	27%	-	-	-	-	-	-

***These tasks are included in the G2 tool, but are not part of the G1 tool. Since the baseline assessment was conducted with Grade 1 students, the baseline data for these tasks is not available.*

A significant increase in average performance from baseline for all tasks across the groups have been observed however the performance of students in High-Touch has increased more compared to Low-Touch

Task	Unit	Midline - Average			Midline - SD			Baseline - Average			Baseline - SD		
		ND	D-HT	D-LT	ND	D-HT	D-LT	ND	D-HT	D-LT	ND	D-HT	D-LT
Number Recognition (Fluency)	Count per minute	22.8	29.9	28.5	14.3	16.9	16.0	12.4	11.4	13.0	12.0	11.2	11.2
Number Recognition (Accuracy)	Percentage	73%	83%	79%	21%	17%	19%	26%	28%	31%	25%	26%	26%
Number Comparison*	Percentage	67%	83%	78%	32%	24%	27%	-	-	-	-	-	-
Counting in Bundles	Percentage	63%	77%	67%	35%	30%	32%	18%	16%	18%	27%	25%	24%
Missing Number	Percentage	40%	51%	43%	27%	27%	26%	14%	13%	16%	21%	19%	20%
Addition Level 1 (Fluency)**	Count per minute	13.1	17.6	14.5	8.2	8.1	7.9	-	-	-	-	-	-
Addition Level 1 (Accuracy)	Percentage	77%	91%	85%	34%	20%	27%	33%	33%	41%	41%	39%	41%
Addition Level 2 (Accuracy)**	Percentage	59%	82%	69%	37%	27%	33%	-	-	-	-	-	-
Subtraction Level 1 (Fluency)**	Count per minute	9.7	12.4	10.3	6.8	5.8	6.2	-	-	-	-	-	-
Subtraction Level 1 (Accuracy)	Percentage	67%	85%	74%	39%	27%	33%	20%	18%	24%	35%	32%	37%
Subtraction Level 2 (Accuracy)**	Percentage	48%	70%	55%	37%	31%	32%	-	-	-	-	-	-
Word Problems	Percentage	56%	69%	60%	32%	27%	29%	30%	25%	31%	32%	27%	30%
Shape Recognition - Circle*	Percentage	17%	13%	18%	20%	15%	20%	-	-	-	-	-	-
Shape Recognition - Rectangle*	Percentage	25%	22%	26%	20%	19%	21%	-	-	-	-	-	-
Shape Recognition - Triangle**	Percentage	26%	23%	28%	15%	13%	17%	-	-	-	-	-	-

**The Number Comparison and Shape Recognition tasks were not reported in the baseline round due to the incorrect administration of these tasks.*

***These tasks are included in the G2 tool, but are not part of the G1 tool. Since the baseline assessment was conducted with Grade 1 students, the baseline data for these tasks is not available.*

Effect Size Calculations in Literacy for Cohort 1, using the Pooled Standard Deviation only from the Midline Round

Task	Unit	Midline Average			Baseline Average			Pooled SD		DiD Effect Size		Pooled SD (ML)		DiD ES (ML SD)		DiD ES Delta	
		ND	D-HT	D-LT	ND	D-HT	D-LT	D-HT	D-LT	D-HT	D-LT	D-HT	D-LT	D-HT	D-LT	D-HT	D-LT
Listening Comprehension	Percentage	77%	81%	77%	62%	59%	62%	30%	30%	0.24	0.01	25%	26%	0.28	0.01	0.04	0.00
Oral Vocabulary	Percentage	94%	96%	95%	95%	96%	96%	9%	9%	0.14	-0.01	8%	9%	0.17	-0.01	0.02	0.00
Initial Sound Identification	Percentage	68%	82%	69%	12%	16%	19%	35%	38%	0.29	-0.15	39%	43%	0.26	-0.13	-0.03	0.02
Letter Reading (Accuracy)	Percentage	77%	91%	85%	29%	35%	42%	30%	31%	0.28	-0.15	22%	25%	0.37	-0.18	0.10	-0.04
Letter Reading (Fluency)	Count per minute	51.6	70.0	62.1	14.8	16.9	20.1	19.3	20.9	0.85	0.25	21.8	24.3	0.75	0.21	-0.10	-0.03
Word Reading (Accuracy)	Percentage	70%	90%	83%	8%	13%	14%	23%	24%	0.69	0.29	24%	27%	0.64	0.27	-0.05	-0.03
Word Reading (Fluency)	Count per minute	21.6	34.9	29.3	5.1	6.2	7.5	12.7	13.1	0.96	0.40	14.4	14.9	0.85	0.36	-0.12	-0.05
Non-Word Reading (Fluency)	Count per minute	19.9	33.2	27.8	2.7	3.5	4.5	10.6	11.4	1.19	0.54	13.5	14.4	0.93	0.43	-0.26	-0.11
Oral Reading Fluency	Count per minute	34.8	61.1	50.4	2.6	3.0	4.0	21.2	22.4	1.22	0.63	28.7	29.8	0.90	0.48	-0.32	-0.16
Reading Comprehension Passage 1	Percentage	60%	88%	78%	2%	4%	4%	27%	30%	0.97	0.52	35%	39%	0.74	0.40	-0.22	-0.12
Letter Writing	Percentage	71%	80%	71%	22%	25%	32%	31%	32%	0.21	-0.30	29%	31%	0.22	-0.31	0.02	-0.01
Word Writing	Percentage	58%	77%	71%	8%	9%	11%	24%	25%	0.73	0.37	26%	28%	0.67	0.34	-0.06	-0.03

The DiD effect size was calculated based on: ***[avg_delta_demo (Δi) - avg_delta_non-demo (Δc)] / SD_pooled (Pooled Standard Deviation)***

Effect Size Calculations in Numeracy for Cohort 1, using the Pooled Standard Deviation only from the Midline Round

Task	Unit	Midline Average			Baseline Average			Pooled SD		DiD Effect Size		Pooled SD (ML)		DiD ES (ML SD)		DiD ES Delta	
		ND	D-HT	D-LT	ND	D-HT	D-LT	D-HT	D-LT	D-HT	D-LT	D-HT	D-LT	D-HT	D-LT	D-HT	D-LT
Number Recognition (Fluency)	Count per minute	22.8	29.9	28.5	12.4	11.4	13.0	13.7	13.6	0.58	0.37	15.6	15.2	0.51	0.33	-0.07	-0.04
Number Recognition (Accuracy)	Percentage	73%	83%	79%	26%	28%	31%	23%	23%	0.34	0.03	19%	20%	0.40	0.03	0.06	0.00
Counting in Bundles	Percentage	63%	77%	67%	18%	16%	18%	29%	30%	0.52	0.11	33%	34%	0.47	0.10	-0.05	-0.01
Missing Number	Percentage	40%	51%	43%	14%	13%	16%	24%	24%	0.50	0.04	27%	26%	0.43	0.04	-0.06	0.00
Addition (Accuracy)	Percentage	77%	91%	85%	33%	33%	41%	35%	36%	0.40	0.02	28%	31%	0.49	0.02	0.09	0.00
Subtraction (Accuracy)	Percentage	67%	85%	74%	20%	18%	24%	33%	36%	0.59	0.08	34%	36%	0.59	0.08	0.00	0.00
Word Problems	Percentage	56%	69%	60%	30%	25%	31%	30%	31%	0.61	0.11	29%	30%	0.61	0.12	0.00	0.00

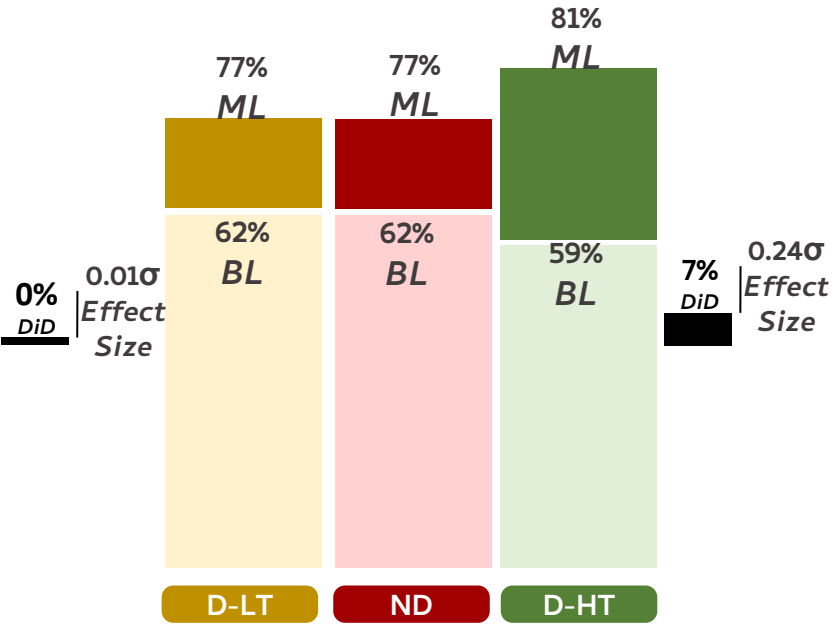
The DiD effect size was calculated based on: ***[avg_delta_demo (Δi) - avg_delta_non-demo (Δc)] / SD_pooled (Pooled Standard Deviation)***

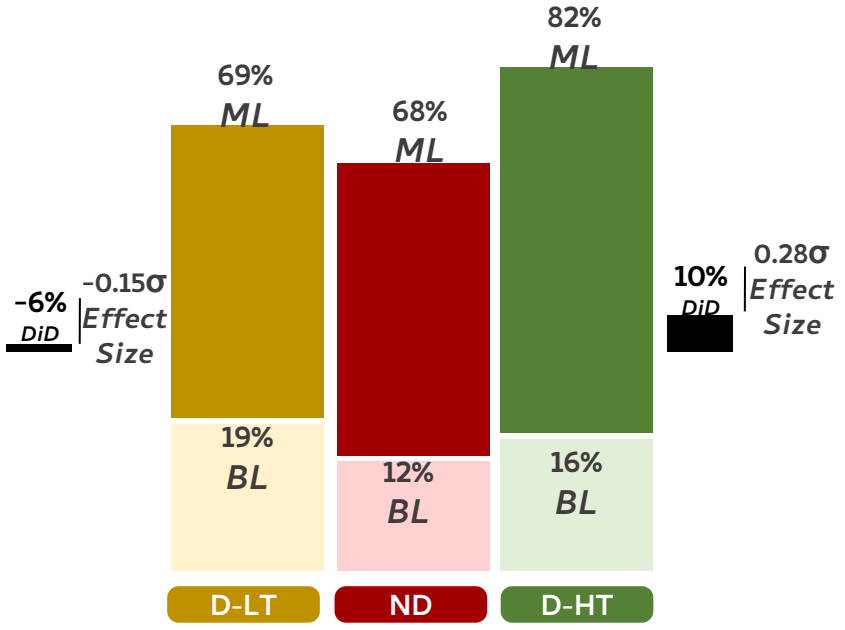
Significant learning gains are observed in students' performance in lower-order literacy tasks such as Initial Sound Identification and Letter Reading in High-Touch Demo Sites

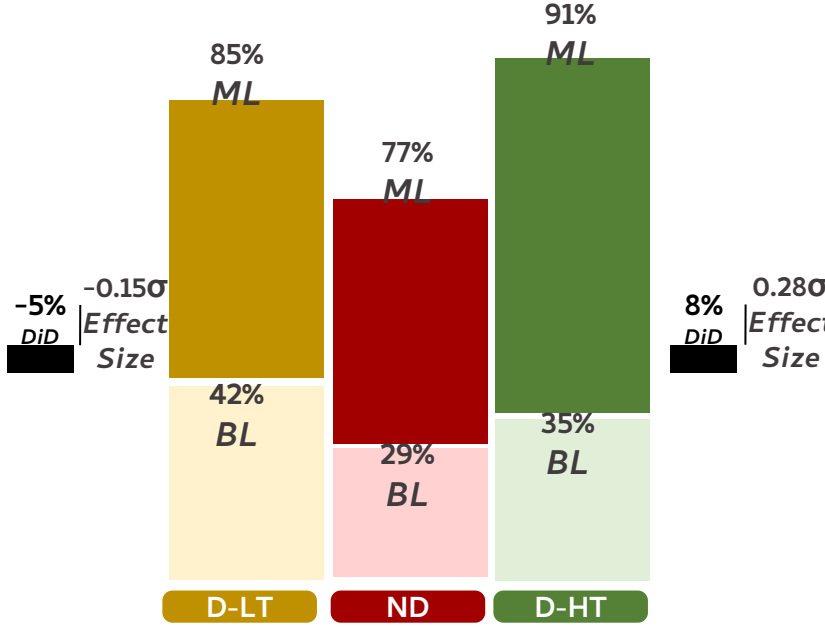
Listening Comprehension

Initial Sound Identification

Letter Reading Accuracy





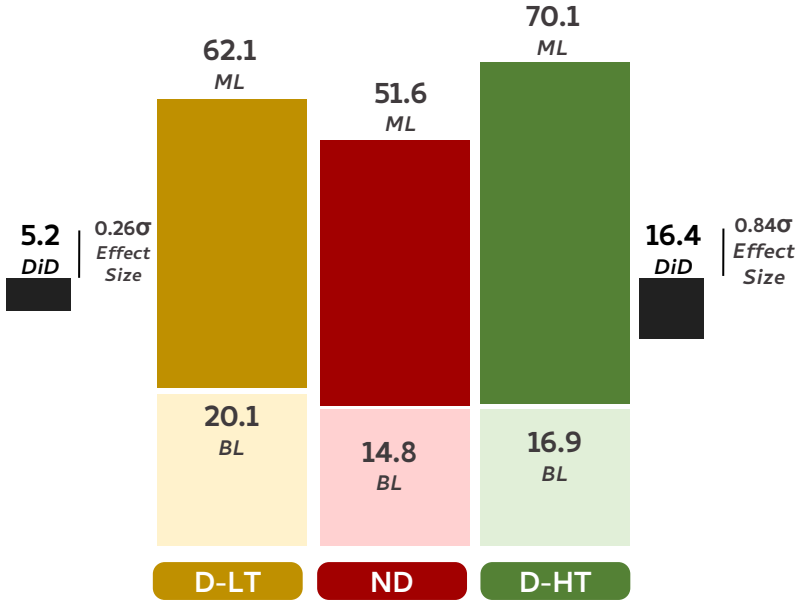


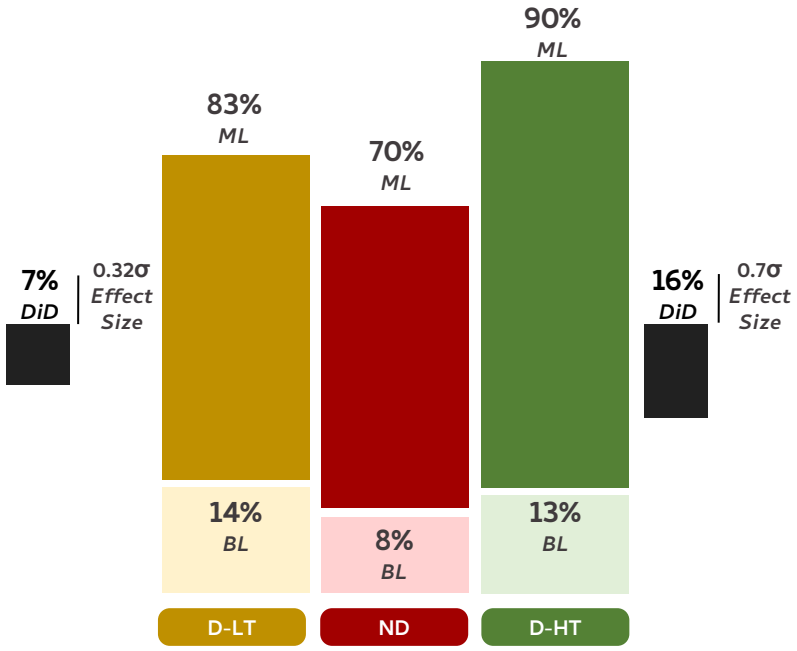
The High-Touch Demo group has shown a noteworthy performance in Letter Reading (Fluency) and Word Reading (Fluency and Accuracy) Tasks

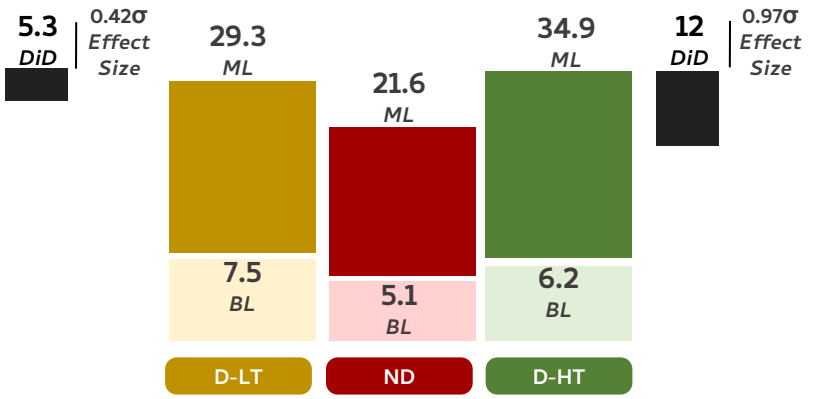
Letter Reading Fluency

Word Reading Accuracy

Word Reading Fluency





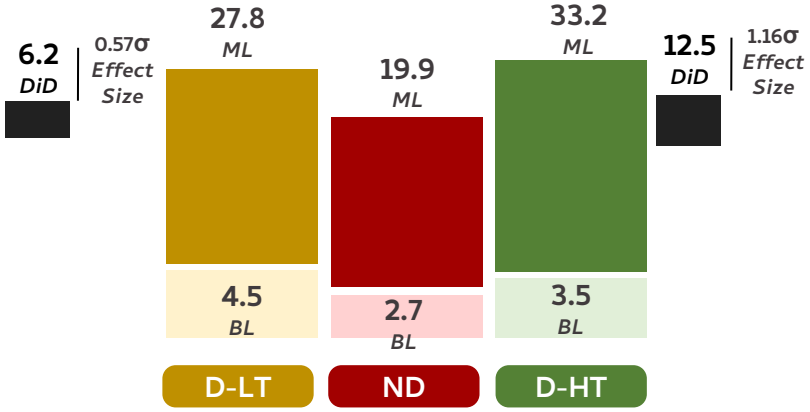


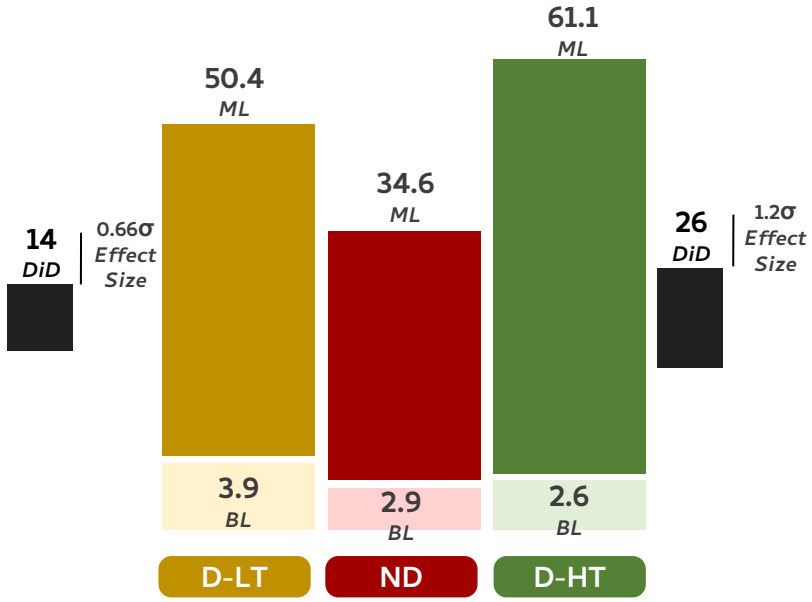
Student performance in higher-order tasks like ORF, and Reading Comprehension saw a significant increase from baseline. Higher performance in these skills indicates students' increased ability to read the text fluently with understanding, enabling them to answer questions based on the text correctly

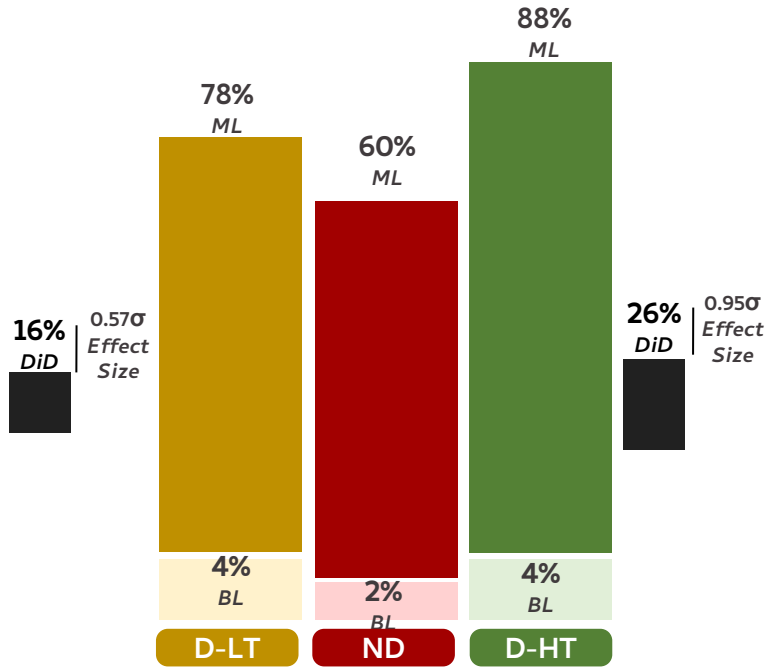
Non-Word Reading Fluency

Oral Reading Fluency

Reading Comprehension 1

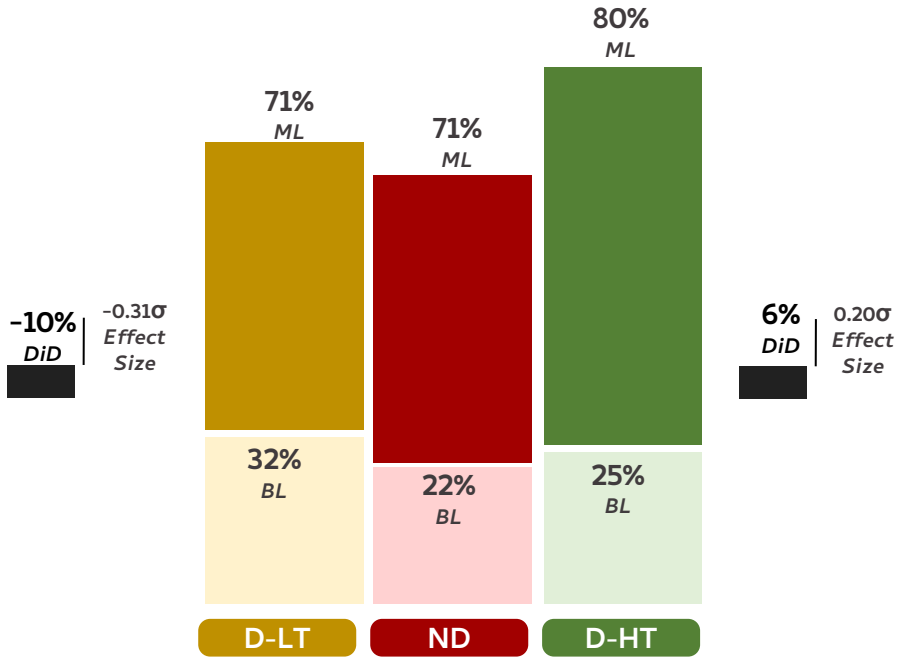




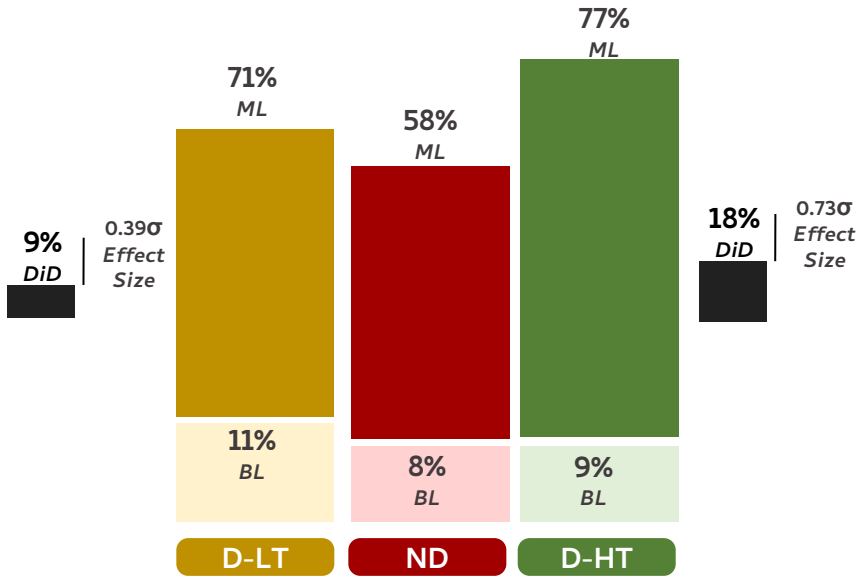


The writing skills of students in both the High and Low-Touch Demo groups have significantly improved in Letter Writing and Word Writing tasks.

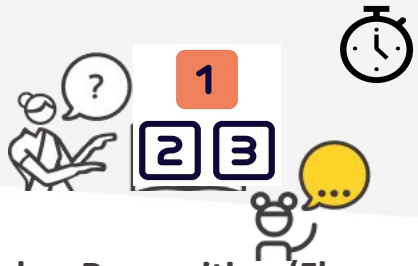
Dictation - Letter Writing



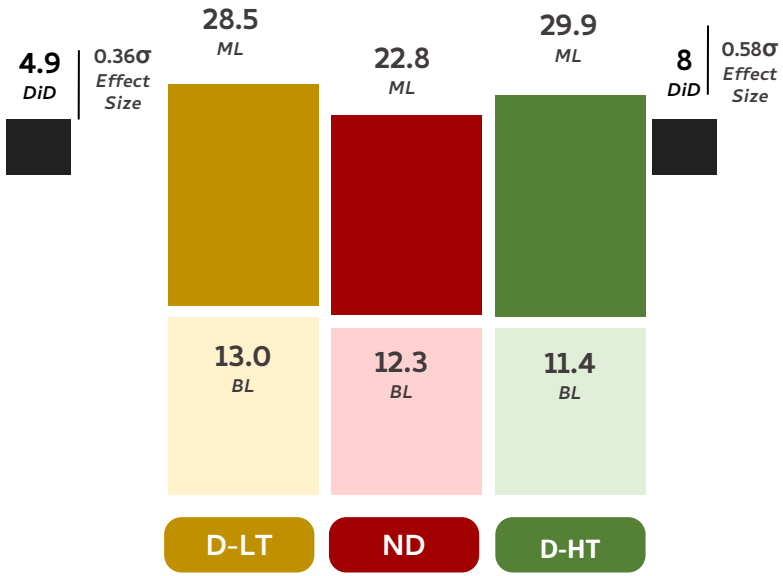
Dictation - Word Writing



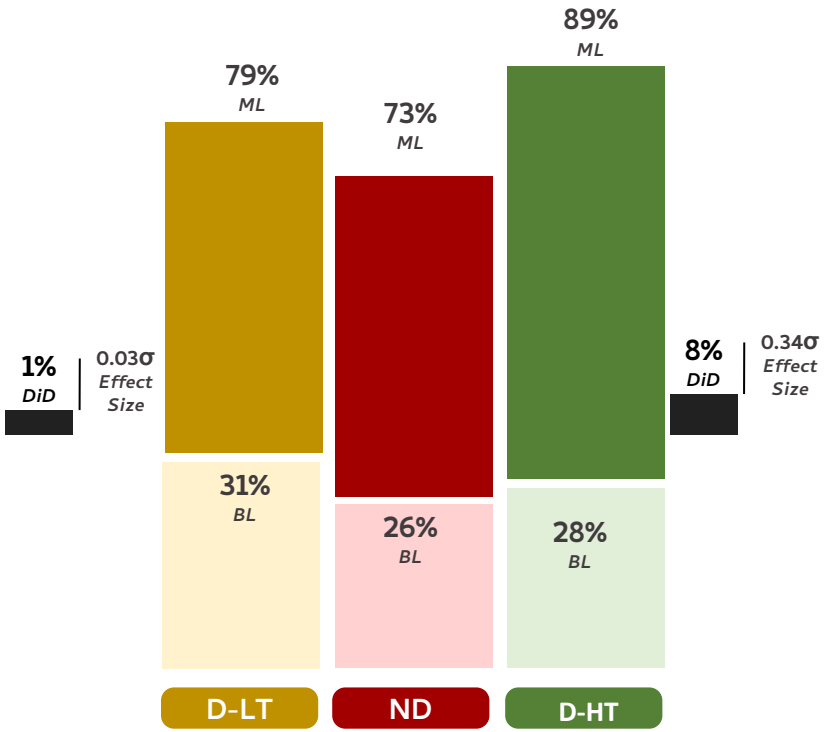
Significant gains have been observed in both foundational and complex tasks in both the High and Low-Touch demo groups, with consistent changes in the performance of all 3 groups from baseline to midline

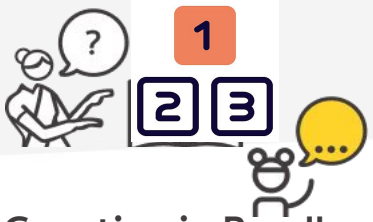


Number Recognition (Fluency)

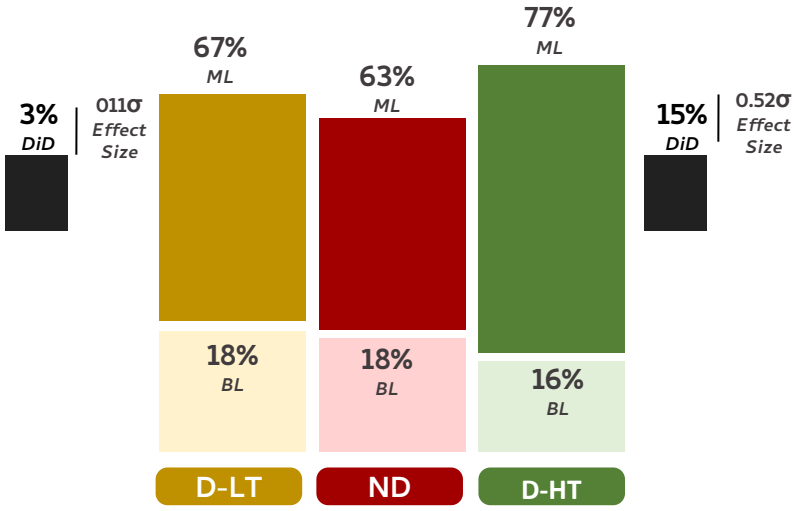


Number Recognition (Accuracy)



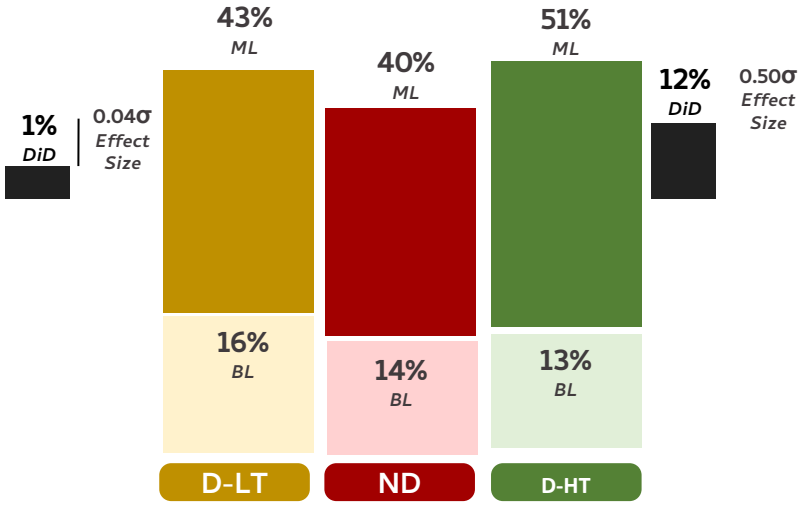


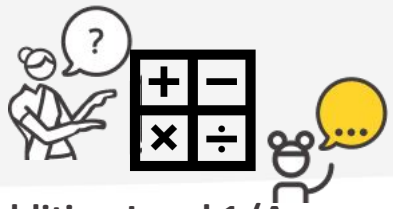
Counting in Bundles



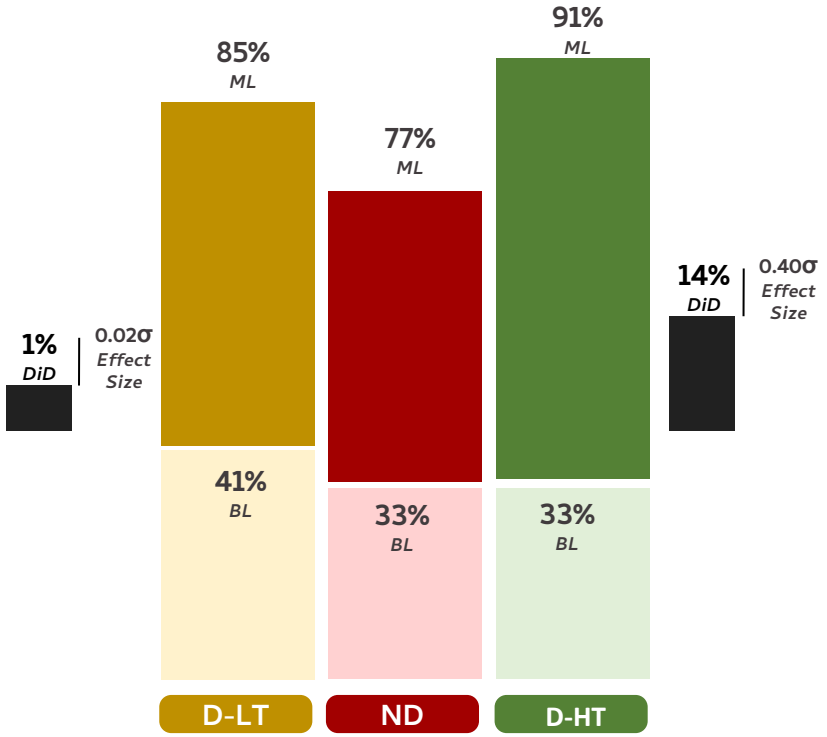
Significant changes have been observed in the performance of students from the baseline round to the midline round in operations related tasks, with large gains seen in the High-Touch demo group's performance

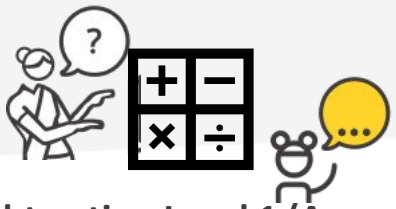
Missing Number



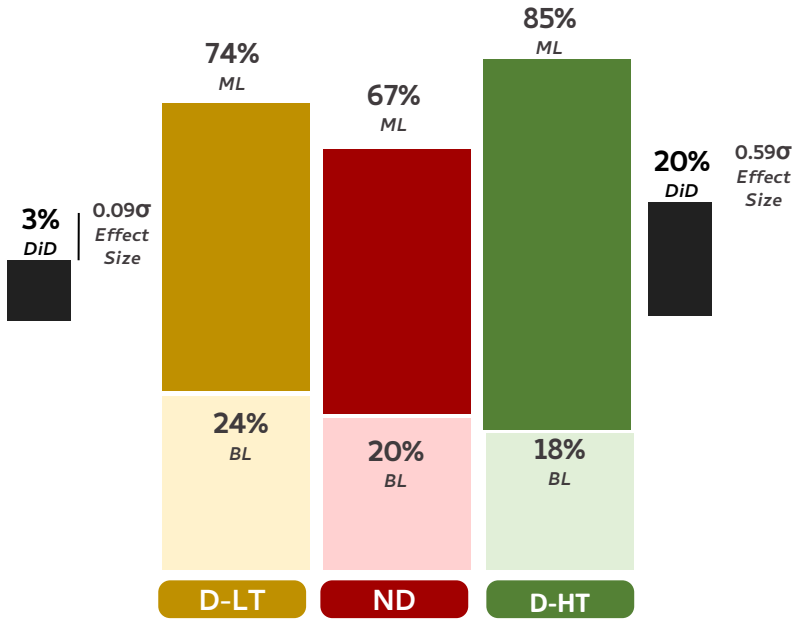


Addition Level 1 (Accuracy)





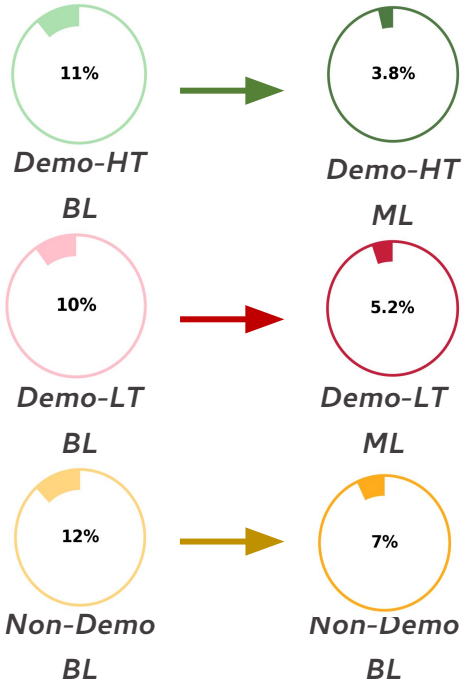
Subtraction Level 1 (Accuracy)



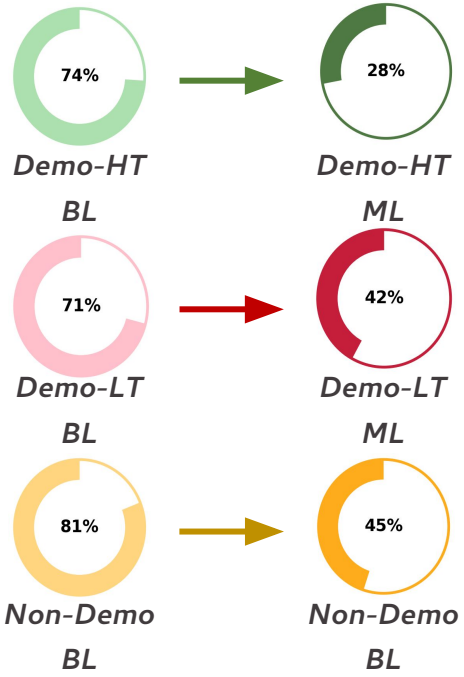
A major reduction was observed in the proportion of students with zero scores in all tasks from the baseline round to the midline round, with a greater change across the demo sites as compared to the non-demo sites

A lower proportion of zero scorers was observed in the Letter Reading task, indicating that students are developing foundational skills that can contribute to the development of reading and other higher order skills

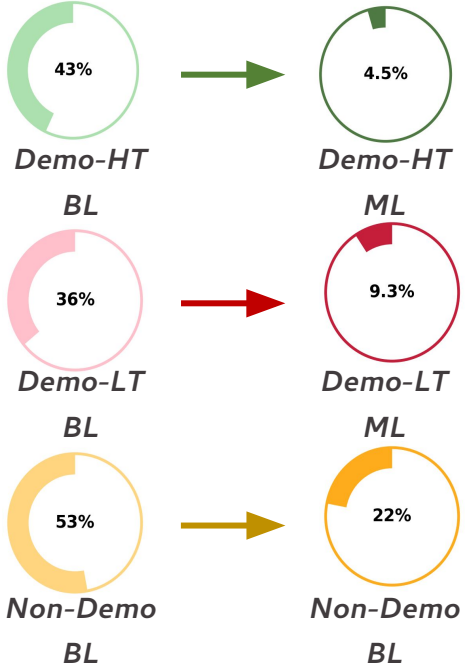
Listening Comprehension



Initial Sound Identification



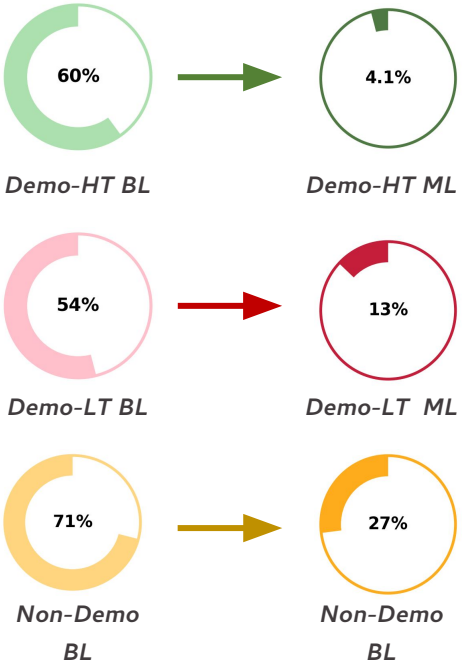
Letter Reading (Accuracy)



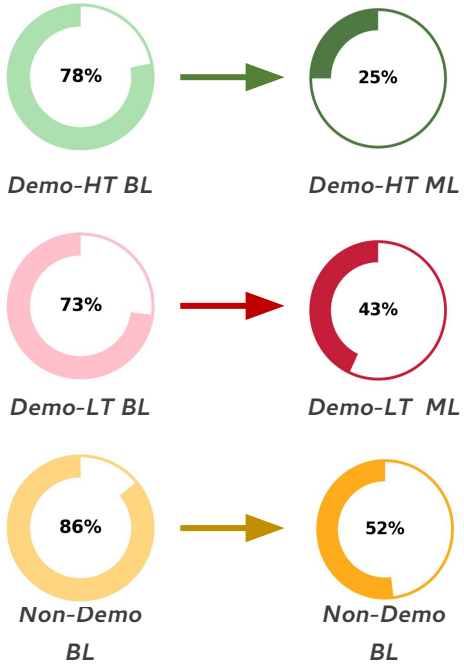
A major reduction was observed in the proportion of students with zero scores in the writing tasks, both Letter and Word Writing, in this Cohort

- Similar to Letter Reading, a low proportion of zero scorers in Word Reading Accuracy indicates that students are developing their reading skills
- A high proportion of zero scorers in this Cohort in Reading Comprehension indicates that students are still developing their skills in reading and understanding text meaning

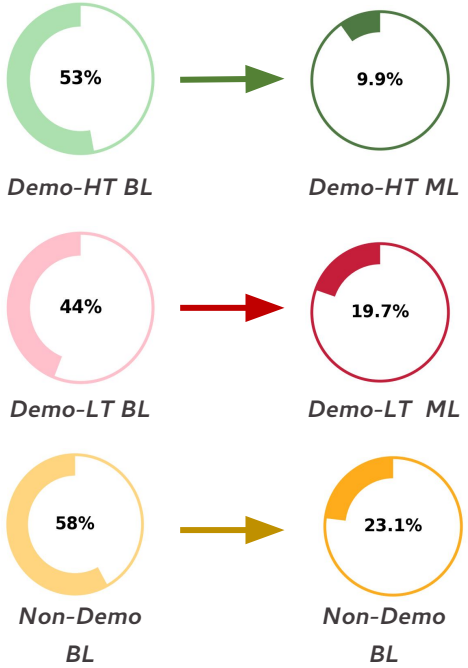
Word Reading Accuracy



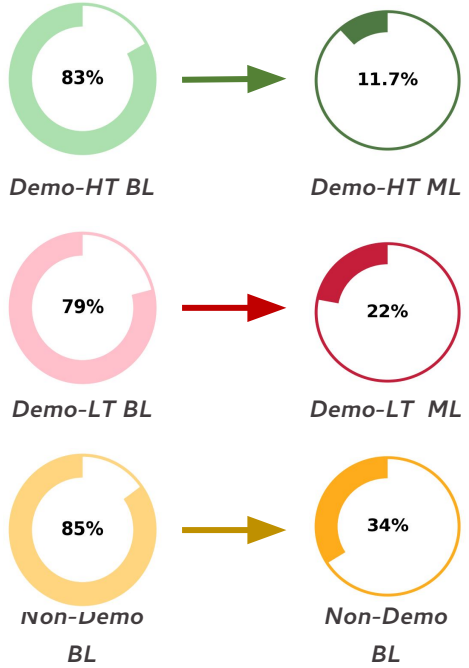
Reading Comprehension



Letter Writing



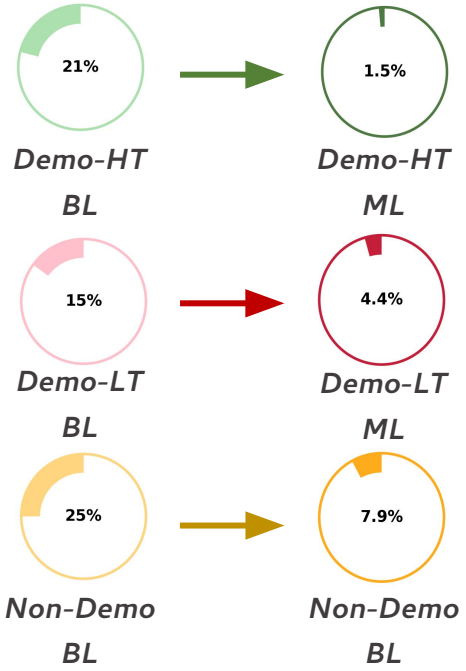
Word Writing



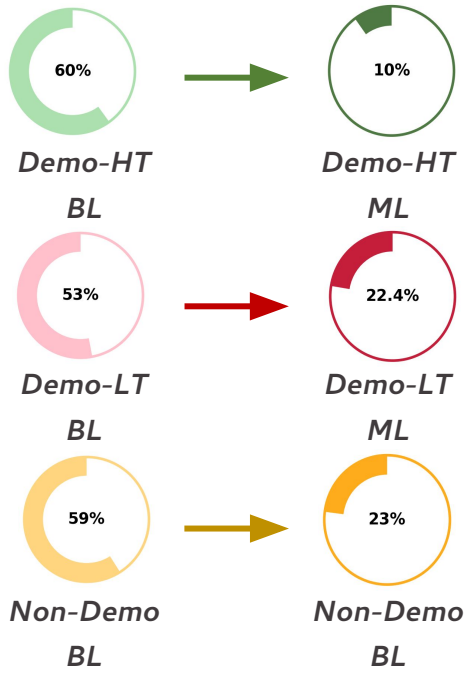
The proportion of zero scorers decreased to 1.5% in the High-Touch Demo group for the Number Recognition (Accuracy) task

A significant decrease in missing number tasks is also observed, and students' proficiency in number recognition tasks may have contributed to this reduction

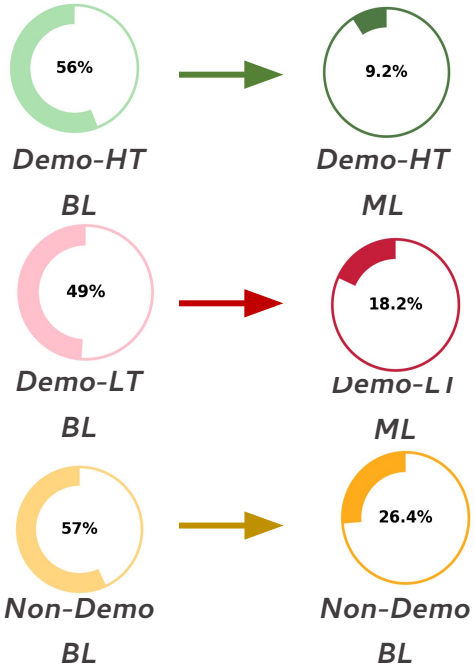
Number Recognition (Accuracy)



Counting in Bundles



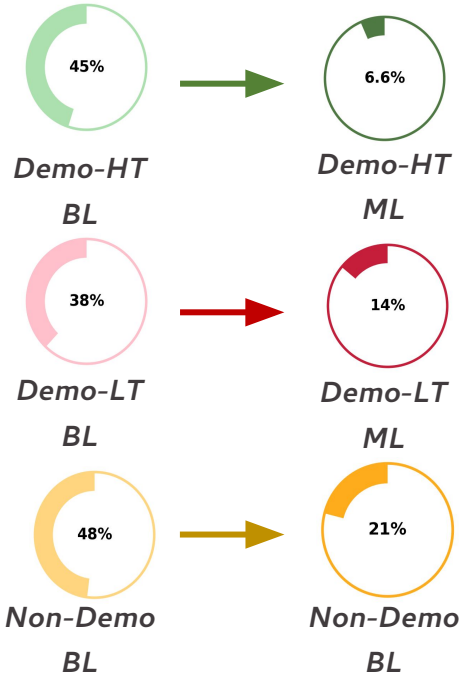
Missing Numbers

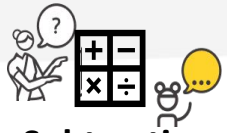


A major reduction was observed in the proportion of students with zero scores in higher order tasks from the baseline round to the midline in the High-Touch Demo group

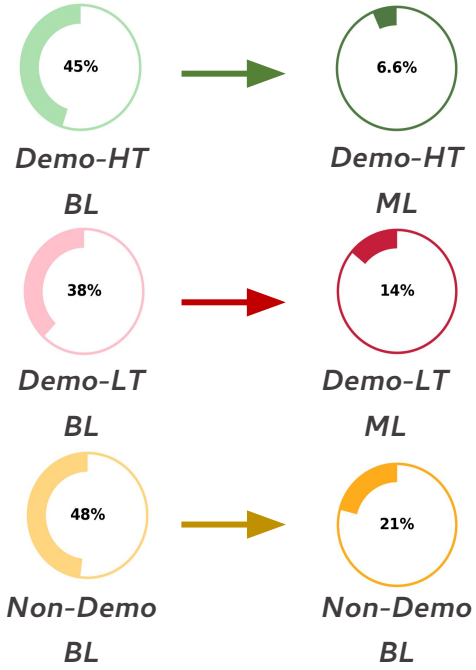
After the number recognition task, Word Problems, Addition, and Subtraction are the next tasks where the proportion of zero scorers is reduced

Addition

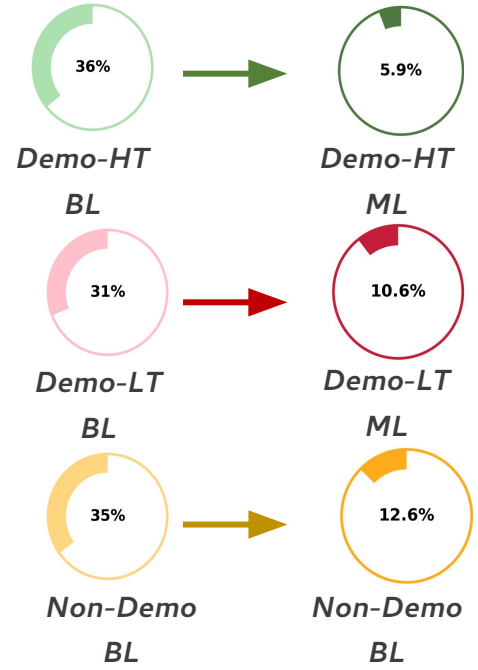




Subtraction



Word Problem



Similar to the baseline round, the performance of boys and girls remained comparable in the midline round as well, with girls performing marginally better than boys in 9 out of 12 literacy tasks

Oral Vocabulary

Boys | Girls → +1.2 (→ 0)

94.7% ML

95.6% ML

Listening Comprehension

Boys | Girls → -3.1 (→ +3)

79.9% ML

76.8% ML

Word Reading (Accuracy)

Boys | Girls → -1 (→ 0)

82% ML

81% ML

Initial Sound Identification

Boys | Girls → -0.9 (→ +1)

73.5% ML

72.6% ML

Letter Reading (Accuracy)

Boys | Girls → +1 (→ -1)

84% ML

85% ML

Reading Comprehension Passage

Boys | Girls → +4 (→ 0)

73% ML

77% ML

The Delta in Girls' and Boys' Scores from the BL has only been added for accuracy based tasks with %age scores, to ensure comparability

Girls' Scores_BL – Boys' Scores_BL → Girls' Scores_ML – Boys' Scores_ML

The performance gap between boys and girls in the midline round was greatest in tasks like Reading Comprehension and Oral Reading Fluency

Letter Writing

Boys: 73% ML

Girls → +1 (→ -2): 74% ML

Letter Reading (Fluency)

Boys: 60 ML

Girls → +2: 62 ML

Non Word Reading Fluency

Boys: 26 ML

Girls → +1: 27 ML

Word Writing

Boys: 67% ML

Girls → +2 (→ 0): 69% ML

Word Reading (Fluency)

Boys: 27 ML

Girls → +2: 29 ML

Oral Reading Fluency

Boys: 47.2 ML

Girls → +3.6: 50.8 ML

The Delta in Girls' and Boys' Scores from the BL has only been added for accuracy based tasks with %age scores, to ensure comparability

Girls' Scores_BL – Boys' Scores_BL

Girls' Scores_ML – Boys' Scores_ML

Boys' performance has significantly increased in most tasks in midline, resulting in a considerable gap between boys and girls

Number Recognition (Accuracy)

Boys: 80% ML

Girls → -4 (→ +1): 76% ML

Number Comparison

Boys: 78% ML

Girls → -5: 73% ML

Counting in Bundles

Boys: 72% ML

Girls → -7 (→ +5): 65% ML

Missing Numbers

Boys: 46% ML

Girls → -4 (→ +6): 42% ML

Addition Level 1 (Accuracy)

Boys: 86% ML

Girls → -4 (→ +7): 82% ML

Subtraction Level 1 (Accuracy)

Boys: 77% ML

Girls → -4 (→ +2): 73% ML

The Delta in Girls' and Boys' Scores from the BL has only been added for accuracy based tasks with %age scores, to ensure comparability

Girls' Scores_BL – Boys' Scores_BL → Girls' Scores_ML – Boys' Scores_ML

In baseline, girls' performance was higher or similar to boys in 6/10 task however in midline there are only 3 tasks where girls performed better than boys

Word Problems

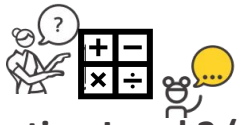
Boys | Girls → -6 (→ +6)

64% *ML* ♀ | ♂ 58% *ML*

Addition Level 2 (Accuracy)

Boys | Girls → -1

70 *ML* ♀ | ♂ 69% *ML*



Subtraction Level 2 (Accuracy)

Boys | Girls → -4

59% *ML* ♀ | ♂ 55% *ML*

The Delta in Girls' and Boys' Scores from the BL has only been added for accuracy based tasks with %age scores, to ensure comparability

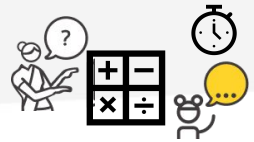
Girls' Scores_BL – Boys' Scores_BL → Girls' Scores_ML – Boys' Scores_ML

For the tasks where girls performed better than boys, the gain in performance is high

Number Recognition (Fluency)

Boys | Girls → -4

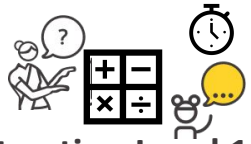
29 ML | 25 ML



Addition Level 1 (Fluency)

Boys | Girls → -3

16 ML | 13 ML



Subtraction Level 1 (Fluency)

Boys | Girls → -1

11 ML | 10 ML

The Gain in Girls' Scores from the Baseline Round has only been added for accuracy based tasks, with %age scores, to ensure comparability

Gain in Girls' Scores_BL → Gain in Girls' Scores_ML

District-Wise Average Scores For Cohort 1

Task- Literacy	Midline - G2 Average					Baseline - G1 Average				
	Varanasi	Siddharth Nagar	Unnao	Mirzapur	Kushinagar	Varanasi	Siddharth Nagar	Unnao	Mirzapur	Kushinagar
Listening Comprehension	79%	72%	83%	-	-	60%	62%	60%	69%	74%
Oral Vocabulary	95%	92%	96%	-	-	96%	96%	94%	96%	97%
Initial Sound Identification	76%	63%	72%	-	-	18%	13%	11%	20%	15%
Letter Fluency	66	48	55	-	-	18.4	15.6	11.8	24.7	22.2
Letter Accuracy	88%	72%	81%	-	-	38%	28%	26%	47%	43%
Word Fluency	32	19.2	23.8		-	7	5	3	14	8
Word Accuracy	87%	66%	74%	-	-	13%	9%	5%	20%	14%
Non Word Fluency	31	17	22	-	-	4.0	2.8	1.6	7.1	4.9
Oral Reading Fluency	56	30	39	-	-	3.5	2.7	1.6	6.8	4.7
Reading Comprehension Questions	83%	52%	67%	-	-	4%	3%	1%	7%	4%
Letter Writing	75%	66%	75%	-	-	28%	20%	20%	36%	36%
Word Writing	74%	54%	62%	-	-	10%	9%	4%	19%	12%

Task-Numeracy	Midline - G2 Average					Baseline - G1 Average				
	Varanasi	Siddharth Nagar	Unnao	Mirzapur	Kushinagar	Varanasi	Siddharth Nagar	Unnao	Mirzapur	Kushinagar
Number Recognition Timed	29.2	23.8	21.8	-	-	12	12	11	19	17
Number Recognition Untimed	81%	73%	72%	-	-	29%	26%	22%	37%	38%
Counting in Bundles	72%	64%	62%	-	-	17%	22%	12%	27%	31%
Missing Numbers	46%	35%	44%	-	-	15%	14%	12%	21%	21%
Addition	88%	75%	79%	-	-	37%	38%	25%	48%	60%
Subtraction	79%	44%	51%	-	-	21%	21%	16%	23%	44%
Word Problems	64%	52%	59%	-	-	28%	32%	26%	41%	46%

There is not a high difference in the performance of boys and girls on mean scores. However, the impact of the intervention appears to be higher among boys in Number Recognition (Accuracy), while the improvement in the performance of girls was much higher in Word Problems

Task Name	Midline-Boys			Baseline-Boys			DiD ES	Midline-Girls			Baseline-Girls			DiD ES
	Average			Average				Average			Average			
	D-HT	D-LT	ND	D-HT	D-LT	ND	Boys	D-HT	D-LT	ND	D-HT	D-LT	ND	Girls
Number Recognition Timed	32.58	29.93	24.96	12.41	13.61	13.33	0.58	27.27	27.04	20.67	10.4	12.51	11.33	0.61
Number Recognition Untimed	85%	80%	75%	30%	33%	27%	0.32	81%	77%	70%	26%	29%	24%	0.42
Counting in Bundles	82%	69%	67%	19%	21%	20%	0.51	72%	65%	59%	14%	16%	16%	0.53
Missing Numbers	54%	44%	43%	15%	17%	15%	0.45	47%	42%	37%	12%	15%	13%	0.5
Addition	93%	85%	80%	36%	47%	35%	0.35	90%	84%	74%	31%	35%	31%	0.48
Subtraction	86%	74%	70%	19%	27%	21%	0.53	83%	74%	63%	16%	21%	18%	0.67
Word Problems	71%	62%	60%	28%	35%	33%	0.53	67%	57%	51%	23%	27%	28%	0.74

The Number Comparison and Shape Recognition tasks were not reported in the baseline round due to the incorrect administration of these tasks.

*For all tasks, the statistical significance of the difference was determined through Welch's unpaired t-test assuming unequal variance t-test . For t-test, one doesn't reject the null hypothesis if p-value is less than 0.05.*represents that the difference between means is significant.*

There is not a high difference in the performance of boys and girls on mean scores. However, the impact of the intervention appears to be higher among boys only in Oral Vocabulary, while the improvement in the performance of girls was much higher in most other tasks

Task Name	Midline-Boys Average			Baseline-Boys Average			DiD ES	Midline-Girls Average			Baseline-Girls Average			DiD ES
	D-HT	D-LT	ND	D-HT	D-LT	ND	Boys	D-HT	D-LT	ND	D-HT	D-LT	ND	Girls
Listening Comprehension	84%	78%	79%	62%	64%	64%	0.23	79%	76%	76%	56%	60%	60%	0.24
Oral Vocabulary	96%	94%	94%	96%	97%	96%	0.16	96%	95%	94%	96%	96%	95%	0.13
Initial Sound Identification	83%	69%	68%	17%	21%	13%	0.33	81%	69%	68%	16%	18%	11%	0.24
Letter Naming Fluency	68.99	60.18	51.19	17.01	19.09	14.37	0.79	71.09	63.91	52.05	16.76	21.05	15.28	0.91
Letter Naming Accuracy	91%	84%	76%	35%	41%	30%	0.31	91%	86%	77%	34%	43%	29%	0.28
Familiar Word Reading Fluency	34.64	27.66	21	6.23	7.24	4.55	0.99	35.11	30.89	22.17	6.09	7.72	5.62	0.95
Familiar Word Reading Accuracy	90%	82%	70%	12%	15%	8%	0.75	91%	85%	71%	14%	14%	8%	0.65
Non-Word Reading Fluency	32.83	26.61	19.49	3.45	4.37	2.48	1.18	33.6	29.04	20.25	3.54	4.59	2.9	1.2
Oral Reading fluency	59.86	47.32	33.89	2.64	4.09	2.84	1.24	62.39	53.41	35.81	3.27	3.89	2.35	1.21
Reading Comprehension Questions	88%	75%	58%	4%	5%	2%	1.03	88%	81%	61%	4%	4%	2%	0.96
Letter Writing	80%	70%	70%	24%	31%	21%	0.25	80%	72%	71%	26%	33%	23%	0.21
Word Writing	77%	68%	58%	9%	13%	8%	0.76	77%	73%	58%	9%	10%	7%	0.71

*For all tasks, the statistical significance of the difference was determined through Welch's unpaired t-test assuming unequal variance t-test . For t-test, one doesn't reject the null hypothesis if p-value is less than 0.05. *represents that the difference between means is significant.*

Annexure 5: Findings from Cohort 2

The average performance of Cohort 2 in the midline is not comparable to the baseline performance of Cohort 1 across both demo and non-demo sites, which can be attributed to the fact that the midline assessment was carried out close to the end of the academic year, and the impact of the intervention in the demo sites

Task	Unit	Midline Average			Midline- SD			Baseline - Average			Baseline- SD		
		ND	D-HT	D-LT	ND	D-HT	D-LT	ND	D-HT	D-LT	ND	D-HT	D-LT
Listening Comprehension	Percentage	71%	74%	66%	31%	28%	29%	62%	59%	62%	34%	32%	32%
Oral Vocabulary	Percentage	95%	96%	94%	8%	7%	8%	95%	96%	96%	11%	8%	8%
Initial Sound Identification	Percentage	45%	63%	46%	46%	44%	46%	12%	16%	19%	30%	33%	35%
Letter Reading (Accuracy)	Percentage	60%	83%	73%	35%	22%	28%	29%	35%	42%	35%	35%	36%
Letter Reading (Fluency)	Count per minute	33.6	51.4	40.4	24.5	20.4	21.7	14.8	16.9	20.1	17.3	15.7	16.4
Word Reading (Accuracy)	Percentage	54%	80%	67%	35%	22%	30%	8%	13%	14%	19%	23%	24%
Word Reading (Fluency)	Count per minute	14.5	27.1	21.0	14.7	14.6	44.4	5.1	6.2	7.5	12.1	9.2	9.7
Non-Word Reading (Fluency)	Count per minute	10.5	21.2	15.0	12.7	12.0	14.6	2.7	3.5	4.5	6.9	6.5	7.5
Oral Reading Fluency (ORF)	Count per minute	15.2	32.3	20.7	22.8	25.1	23.1	2.6	3.0	4.0	10.5	8.9	10.9
Reading Comprehension Passage 1	Percentage	34%	59%	38%	41%	41%	40%	2%	4%	4%	14%	18%	19%
Letter Writing	Percentage	61%	77%	64%	38%	30%	37%	22%	25%	32%	32%	34%	36%
Word Writing	Percentage	45%	70%	58%	36%	30%	35%	8%	9%	11%	21%	23%	26%

Even in numeracy, the average performance of Cohort 2 in the midline is not comparable to the baseline performance of Cohort 1 across both demo and non-demo sites, due to the fact that the midline assessment was carried out close to the end of the academic year, and the impact of the intervention in the demo sites

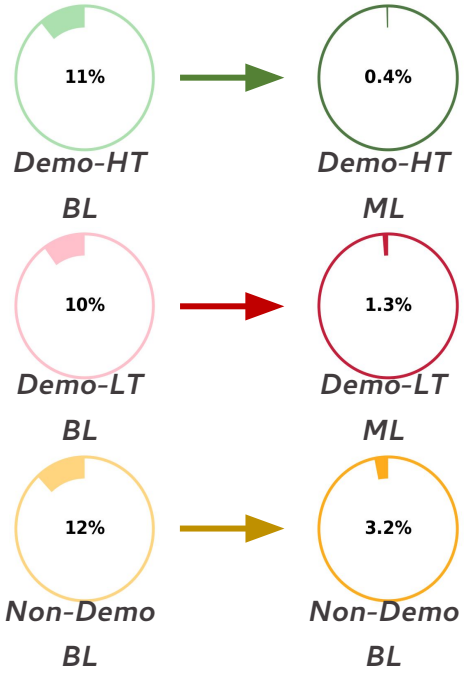
Task	Unit	Midline-Average			Midline-SD			Baseline - Average			Baseline - SD		
		ND	D-HT	D-LT	ND	D-HT	D-LT	ND	D-HT	D-LT	ND	D-HT	D-LT
Counting	Count per minute	98.9	108.9	102.2	38.2	33.7	34.8	81.0	75.4	76.5	36.3	39.5	33.1
Number Recognition (Fluency)	Count per minute	19.0	23.1	21.3	14.1	14.1	14.5	12.4	11.4	13.0	12.0	11.2	11.2
Number Recognition (Accuracy)	Percentage	61%	69%	65%	24%	19%	21%	26%	28%	31%	25%	26%	26%
Number Comparison*	Percentage	45%	57%	50%	34%	31%	33%	-	-	-	-	-	-
Counting in Bundles	Percentage	45%	55%	45%	36%	34%	36%	18%	16%	18%	27%	25%	24%
Missing Number	Percentage	34%	48%	39%	29%	27%	28%	14%	13%	16%	21%	19%	20%
Addition Level 1 (Accuracy)	Percentage	64%	85%	72%	41%	29%	37%	33%	33%	41%	41%	39%	41%
Subtraction Level 1 (Accuracy)	Percentage	51%	76%	60%	43%	36%	42%	20%	18%	24%	35%	32%	37%
Word Problems	Percentage	55%	66%	54%	34%	30%	32%	30%	25%	31%	32%	27%	30%
Shape Recognition - Circle*	Percentage	30%	26%	34%	20%	17%	23%	-	-	-	-	-	-
Shape Recognition - Rectangle*	Percentage	53%	49%	53%	18%	15%	25%	-	-	-	-	-	-

**The Number Comparison and Shape Recognition tasks were not reported in the baseline round due to the incorrect administration of these tasks.*

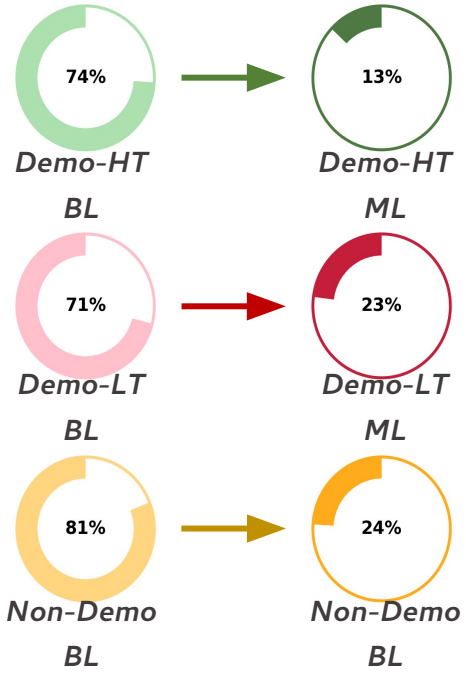
A major reduction was observed in the proportion of students with zero scores in all tasks from the baseline round to the midline round, with a greater change across the demo sites as compared to the non-demo sites

A commonly used indicator to gauge the prevalence of struggling learners is the proportion of students achieving zero marks on a particular task. For instance, at baseline, the proportion of students with zero scores in the letter accuracy task stood at 43% for high touch demo group, which came down to 0.8% at the midline.

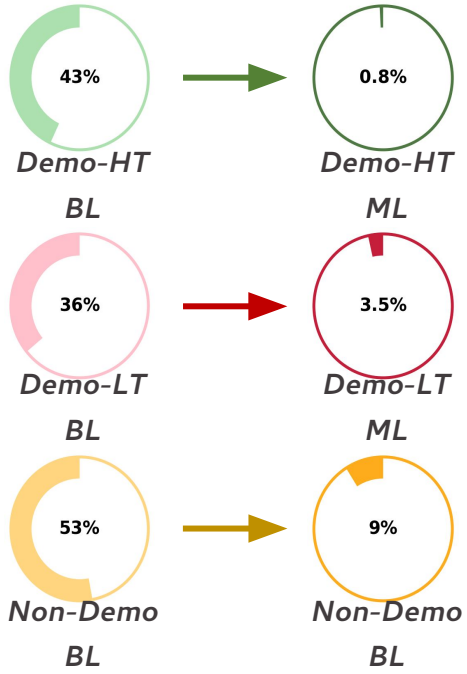
Listening Comprehension



Initial Sound Identification



Letter Reading (Accuracy)



A major reduction of zero scorers is observed in the Word Writing and Reading Comprehension tasks, especially for the High-Touch demo group

- In the task of word accuracy, a low proportion of zero scorers at midline across all groups indicates students' learning of foundational literacy skills
- Interestingly, the reduction in zero scorers in the Word Writing task is higher than in the Letter Writing task in High-Touch demo group

Word Reading Accuracy

Reading Comprehension

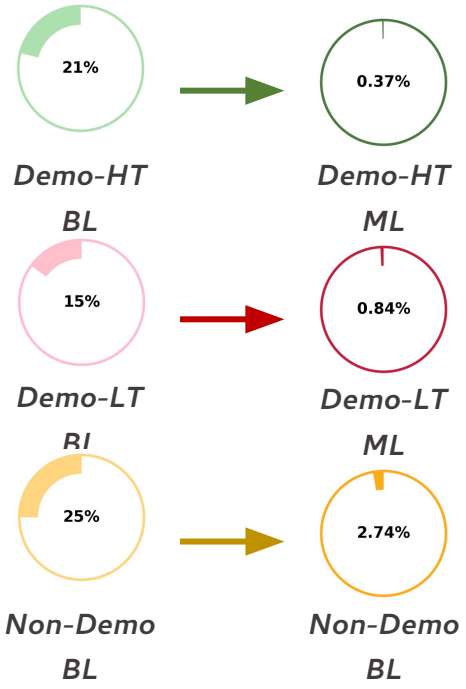
Letter Writing

Word Writing

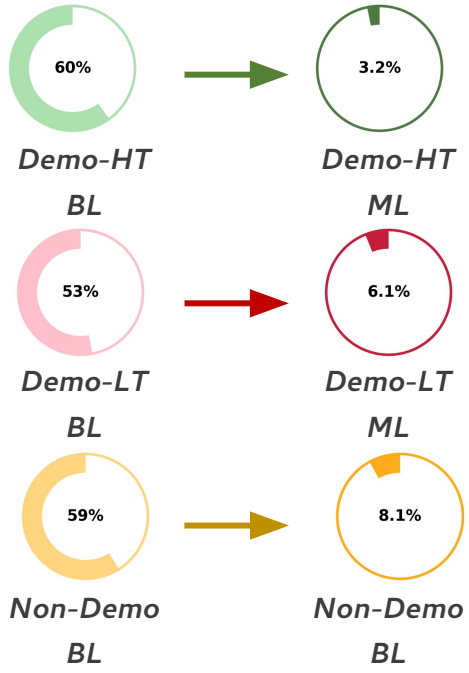
A major reduction was observed in the proportion of students with zero scores in tasks from the baseline round to the midline round across all sites

In number recognition task, zero scorers came down to around 0.3% and 0.8% for both High-Touch and Low-Touch sites shows students developing mastery on this competency

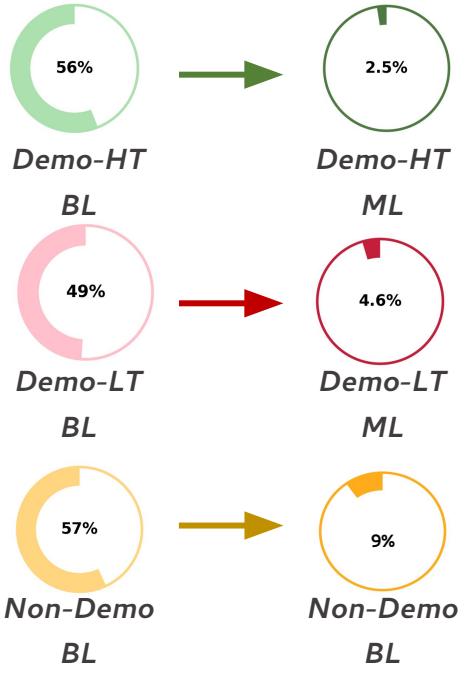
Number Recognition (Accuracy)



Counting in Bundles



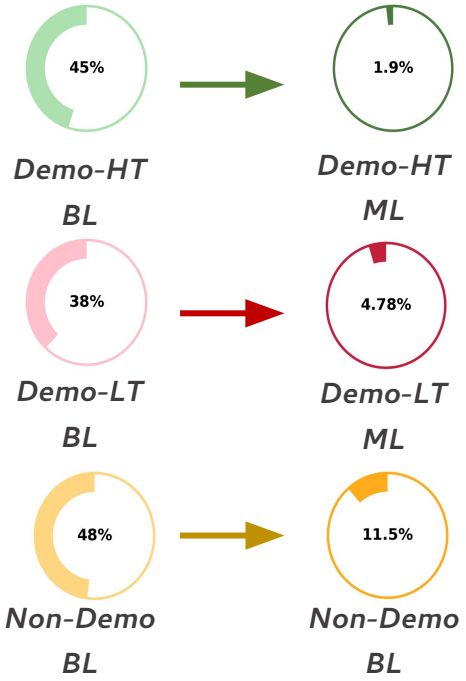
Missing Numbers

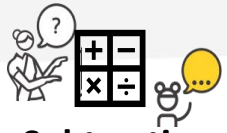


Reduction of zero scorers in tasks related to operations in indicates huge improve in higher order numeracy skills

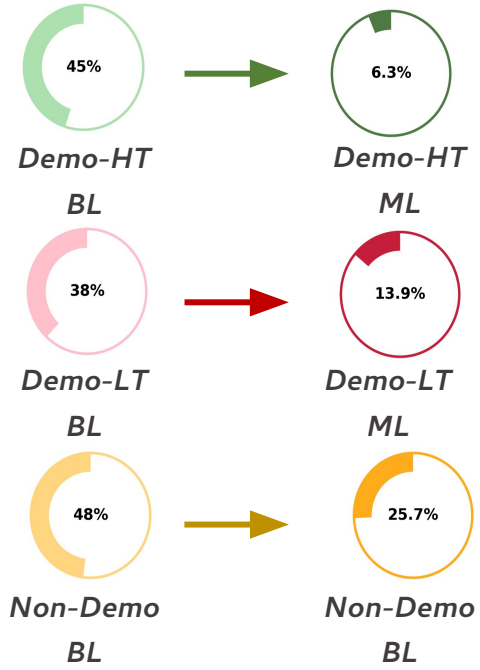
Addition and word problems had less zero scorers compared to subtraction in both High-Touch and Low-Touch demo sites

Addition

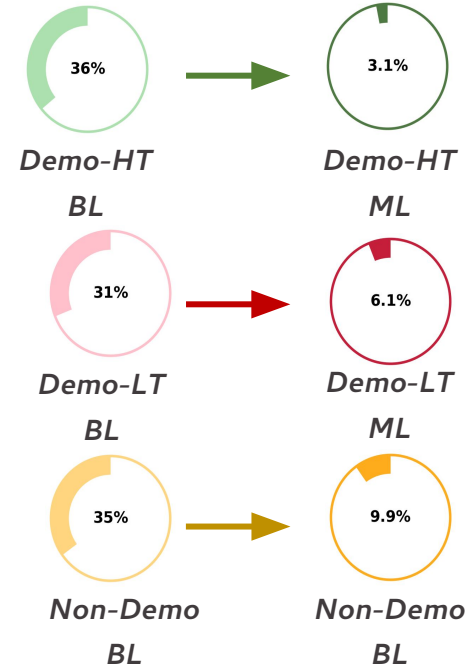




Subtraction



Word Problem



Similar to the baseline, girls continued to outperform boys in 10 out of 12 literacy tasks

Listening Comprehension

Boys: 70.8% ML

Girls → -0.5 (→ +3): 70.3% ML

Initial Sound Identification

Boys: 51% ML

Girls → +0.9 (→ +1): 51.9% ML

Word Reading (Accuracy)

Boys: 65% ML

Girls → +2 (→ 0): 67% ML

Oral Vocabulary

Boys: 95.2% ML

Girls → -1.2 (→ 0): 94% ML

Letter Reading (Accuracy)

Boys: 70% ML

Girls → +2 (→ -1): 72% ML

Reading Comprehension

Boys: 43% ML

Girls → +1 (→ +0): 44% ML

The Delta in Girls' and Boys' Scores from the BL has only been added for accuracy based tasks with %age scores, to ensure comparability

Girls' Scores_BL – Boys' Scores_BL → Girls' Scores_ML – Boys' Scores_ML

There is a significant performance gap in tasks such as Letter and Word Reading, as well as Writing, with higher performance observed among girls

Letter Writing

Boys | Girls → +3 (→ -2)

65% *ML*

68% *ML*

Letter Reading (Fluency)

Boys | Girls → +2

40 *ML*

42 *ML*

Non Word Fluency

Boys | Girls → +1

15 *ML*

16 *ML*

Word Writing

Boys | Girls → +1 (→ 0)

57% *ML*

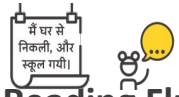
58% *ML*

Word Reading (Fluency)

Boys | Girls → +2

27 *ML*

29 *ML*



Oral Reading Fluency

Boys | Girls → +2

21 *ML*

23 *ML*

The Delta in Girls' and Boys' Scores from the BL has only been added for accuracy based tasks with %age scores, to ensure comparability

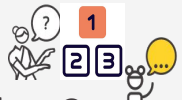
Girls' Scores_BL – Boys' Scores_BL → Girls' Scores_ML – Boys' Scores_ML

There is a significant gap between the performance of boys and girls in most numeracy tasks, which contrasts with the baseline results from last year (for cohort 1), where girls performed better on most tasks

Number Recognition (Accuracy)

Boys: 66% ML

Girls → -3 (→ +1): 63% ML



Number Comparison

Boys: 54% ML

Girls → -7: 47% ML

Number Recitation (Fluency)

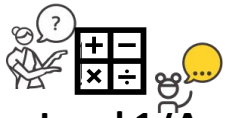
Boys: 104 ML

Girls → -7: 97 ML

Counting

Boys: 101 ML

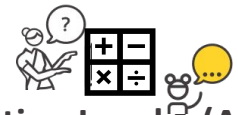
Girls → +1: 102 ML



Addition Level 1 (Accuracy)

Boys: 76% ML

Girls → -5 (→ +7): 71% ML



Subtraction Level 1 (Accuracy)

Boys: 65% ML

Girls → -6 (→ +2): 59% ML

The Delta in Girls' and Boys' Scores from the BL has only been added for accuracy based tasks with %age scores, to ensure comparability

Girls' Scores_BL – Boys' Scores_BL → Girls' Scores_ML – Boys' Scores_ML

The gap in performance in higher order tasks increased with the complexity of the tasks

Word problems

Boys

61% *ML*

Girls → -6 (→ +6)

55% *ML*

Missing Numbers

Boys

41% *ML*

Girls → -2 (→ +6)

39% *ML*

Counting in Bundles

Boys

51% *ML*

Girls → -7 (→ +5)

44% *ML*

Number Recognition (Fluency)

Boys

22 *ML*

Girls → -3

19 *ML*

The Gain in Girls' Scores from the Baseline Round has only been added for accuracy based tasks, with %age scores, to ensure comparability

Gain in Girls' Scores_BL → Gain in Girls' Scores_ML

District-Wise Average Scores For Cohort 2

Task- Literacy	Midline - G1 Average					Baseline - G1 Average				
	Varanasi	Siddharth Nagar	Unnao	Mirzapur	Kushinagar	Varanasi	Siddharth Nagar	Unnao	Mirzapur	Kushinagar
Listening Comprehension	70%	64%	80%	-	-	60%	62%	60%	69%	74%
Oral Vocabulary	95%	94%	96%	-	-	96%	96%	94%	96%	97%
Initial Sound Identification	55%	39%	54%	-	-	18%	13%	11%	20%	15%
Letter Fluency	46	30.8	36.9	-	-	18.4	15.6	11.8	24.7	22.2
Letter Accuracy	78%	54%	68%	-	-	38%	28%	26%	47%	43%
Word Fluency	24	13	16	-	-	7	5	3	14	8
Word Accuracy	74%	49%	59%	-	-	13%	9%	5%	20%	14%
Non Word Fluency	18.1	10	12	-	-	4.0	2.8	1.6	7.1	4.9
Oral Reading Fluency	26.4	14	16	-	-	3.5	2.7	1.6	6.8	4.7
Reading Comprehension Questions	49%	26%	44%	–	-	4%	3%	1%	7%	4%
Letter Writing	71%	56%	67%	-	-	28%	20%	20%	36%	36%
Word Writing	64%	41%	50%	-	-	10%	9%	4%	19%	12%
Task-Numeracy	**Midline - G1 Average**					**Baseline - G1 Average**				
	Varanasi	Siddharth Nagar	Unnao	Mirzapur	Kushinagar	Varanasi	Siddharth Nagar	Unnao	Mirzapur	Kushinagar
Number Recognition Timed	22.2	17.8	20.3	-	-	12.2	12.3	10.9	18.6	18.6
Number Recognition Untimed	67%	60%	63%	-	-	29%	26%	22%	37%	37%
Counting in Bundles	50%	44%	46%	-	-	17%	22%	12%	27%	27%
Missing Numbers	43%	30%	39%	-	-	15%	14%	12%	21%	21%
Addition	79%	63%	66%	-	-	37%	38%	25%	48%	48%
Subtraction	68%	47%	56%	-	-	21%	21%	16%	23%	23%
Word Problems	60%	54%	56%	-	-	28%	32%	26%	41%	41%

Annexure 6: Literacy-related Findings from the Follow-Up Qualitative Study

G1 Literacy Findings

Sub-Section Name	Conducted in ___% of 15 classrooms
OLD SEL	50%
OLD Story / Poem	-
OLD Story Vocab	-
OLD Story Discuss	67%
OLD Game	0%
OLD WB	67%
PA LI/B	60%
PA LW	33%
PA W/SR	60%
PA WB	80%
R Prac	60%
IR	-

OLD - Oral Language Development; SEL - Social and emotional Learning; PA - Phonological Awareness; LI - Letter Identification; B - Blending; LW - Letter Writing; SR - Sentence Reading; WB - Workbook; R Prac - Reading Practice; IR - Independent Reading

High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

G2 Literacy Findings

Sub-Section Name	Conducted in ___% of 15 classrooms
OLD Ideal_R	78%
OLD R/S_Teach	-
OLD Vocab	50%
OLD R_Group	43%
OLD Discuss_Teach	67%
OLD Discuss_Group	-
OLD W	56%
WB Acti WB1	56%
WB Acti Ideal_R	-
WB Acti Story_Discuss	0%
WB Acti WB2	71%
WB Acti R_Group	0%

Sub-Section Name	Conducted in ___% of 15 classrooms
WB Acti Act1_Vocab You Do	33%
WB Acti Act2	50%
WB Acti Act3	40%
WB Acti Act4	50%
WB Acti Act5	50%
WB Acti NB_Act1	43%
WB Acti NB_Act2	75%
WB Acti Story_W	-
WB Acti IR	-
R Prac	78%
IR - P	-

OLD - Oral Language Development; R/S_Teach - Reading/ Sharing by Teacher; Vocab - Vocabulary; R_Group - Guided Reading in Student Groups; Discuss_Teach - Discussion based on poem/ story/ experiences with the teacher; Discuss_Group - Discussion in Student Groups; W - Writing Activity; WB - Workbook; WB Acti - Workbook-based Activities, Act - Activity, RC - Reading Practice; IR - Independent Reading

High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

G1 Literacy Findings: Sub-Section Wise Overview

Category / School Code	OLD SEL	OLD Story / Poem	OLD Story Vocab	OLD Story Discuss	OLD Game	OLD WB	PA LI/B	PA LW	PA W/SR	PA WB	R Prac	IR
ROV_1	-	-	-	0%	-	0%	0%	-	0%	0%	0%	-
ROV_2	-	-	-	58%	-	-	18%	0%	0%	43%	24%	-
ROV_3	58%	-	-	-	0%	-	50%	0%	52%	36%	53%	-
ROV_4	-	-	-	79%	-	83%	55%	-	70%	82%	24%	33%
ROV_6	0%	60%	-	-	0%	21%	0%	39%	53%	29%	0%	-

OLD - Oral Language Development; SEL - Social and emotional Learning; PA - Phonological Awareness; LI - Letter Identification; B - Blending; LW - Letter Writing; SR - Sentence Reading; WB - Workbook; R Prac - Reading Practice; IR - Independent Reading

		Grade 1	Grade 2
High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

G2 Literacy Findings: Sub-Section Wise Overview (1/2)

Category/ School Code	OLD Ideal_R	OLD R/S_Teach	OLD Vocab	OLD R_Group	OLD Discuss_Teach	OLD Discuss_Group	OLD W	WB Acti WB1	WB Acti Ideal_R	WB Acti Story_Discuss	WB Acti WB2
ROV_5	0%	-	31%	0%	0%	-	0%	0%	-	-	0%
ROV_7	16%	-	-	0%	42%	-	61%	60%	41%	-	56%
ROV_8	82%	-	-	46%	50%	-	71%	78%	-	-	46%
ROV_9	84%	-	0%	-	79%	-	0%	68%	4	-	63%
ROV_10	86%	-	-	64%	58%	-	50%	75%	-	-	60%
ROV_11	18%	-	0%	-	0%	-	0%	0%	-	-	0%
ROV_12	52%	-	-	0%	46%	-	61%	78%	-	-	14%
ROV_13	39%	-	31%	0%	0%	-	0%	0%	-	0%	-
ROV_14	0%	-	-	43%	79%	-	64%	0%	-	0%	-
ROV_15	-	-	-	-	-	-	-	-	-	-	-

OLD - Oral Language Development; R/S_Teach - Reading/ Sharing by Teacher; Vocab - Vocabulary; R_Group - Guided Reading in Student Groups; Discuss_Teach - Discussion based on poem/ story/ experiences with the teacher; Discuss_Group - Discussion in Student Groups; W - Writing Activity; WB - Workbook; WB Acti - Workbook-based Activities, Act - Activity R Prac - Reading Practice; IR - Independent Reading

		Grade 1	Grade 2
High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

G2 Literacy Findings: Sub-Section Wise Overview (2/2)

Category/ School Code	WB Acti R_Group	WB Acti Act1_Vocab You Do	WB Acti Act2	WB Acti Act3	WB Acti Act4	WB Acti Act5	WB Acti NB_Act1	WB Acti NB_Act2	WB Acti Story_W	WB Acti IR	R Prac	IR
ROV_5	-	-	-	0%	-	-	0%	-	-	-	16%	-
ROV_7	-	-	-	0%	-	-	0%	-	14%	-	15%	-
ROV_8	-	-	-	-	-	-	49%	55%	-	-	34%	-
ROV_9	-	-	-	-	-	-	65%	40%	-	-	46%	-
ROV_10	-	-	-	49%	-	-	62%	-	-	-	54%	-
ROV_11	-	25%	-	-	-	-	0%	0%	-	-	14%	-
ROV_12	-	-	-	-	-	-	0%	62%	-	50%	10%	-
ROV_13	0%	0%	0%	0%	0%	0%	-	-	-	-	0%	-
ROV_14	0%	0%	38%	43%	43%	43%	-	-	-	-	0%	-
ROV_15	-	-	-	-	-	-	-	-	-	-	-	-

OLD - Oral Language Development; R/S_Teach - Reading/ Sharing by Teacher; Vocab - Vocabulary; R_Group - Guided Reading in Student Groups; Discuss_Teach - Discussion based on poem/ story/ experiences with the teacher; Discuss_Group - Discussion in Student Groups; W - Writing Activity; WB - Workbook; WB Acti - Workbook-based Activities, Act - Activity R Prac - Reading Practice; IR - Independent Reading

Key Finding: There is a high focus on reading-related sub-sections in G2 classrooms, most likely because NIPUN Lakshya App assessments focus only on reading skills

Grade 1	Grade 2

High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

There is a high focus on reading-related sub-sections in G2 classrooms, most likely because NIPUN Lakshya App assessments focus only on reading skills

- **In the first period - Oral Language Development - Ideal Reading was the most common sub-section done, in about 80% of classrooms. The average time spent was 9 mins.**
  - Some teachers gave a positive opinion on OLD in terms of its ease, preference or importance.
- **In the workbook-based activities period, the sub-sections most commonly implemented across classrooms were 'Workbook Activity 2' and 'Notebook-based Activity 2 in many and most of classrooms respectively. The average time spent was 11 mins and 16 mins respectively.**
  - In the FGDs, teachers said that students are very interested in working on their workbooks.
- **Reading practice in the third period was implemented in most of classrooms with an average time of 12 mins.**

"Haanji maukhik bhaag acha lagta hai aur thoda board pe, black board pe hum ko samjhana zyaada better lagta hai. Sir jaise ki saare bacche uspe focus karte hai, black board pe."

Teacher, during interview

"Kyunki mujhe nahi lag raha hai ki bachche itne interest se workbook bharte hain ki unko nahi samajh me aa raha hain."

Teacher, during FGD

Annexure 7: Numeracy-related Findings from the Follow-Up Qualitative Study

Numeracy Findings

Sub-section Name	% of schools this subsection was conducted in, for	
	Grade 1	Grade 2
Mathematical Conversation - I Do + We Do (Through a story/ other activities)	50%	67%
Skill Building (1) - I Do + We Do	50%	83%
Skill Building (2) - I Do + We Do	0%	100%
Workbook Practice - You Do	75%	83%
Math Games - We Do + You Do	25%	17%

Key Finding: Length of the numeracy lesson plan and amount of preparation required for the 'Math Games' section may be leading to teachers skipping it altogether.

In numeracy, findings have been reported in percentages and not a number count because the subsections to be done in each class varied depending on which lesson plan was being taught

High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

Numeracy Findings: Sub-Section Wise Overview

School Code	MC + Skills_I Do + We Do	MC + Skills_Skill (1)	MC + Skills_Skill (2)	WB_You Do	Math Games
ROV_1	-	-	-	-	-
ROV_2	78%	65%	0%	20%	0%
ROV_3	-	-	-	-	-
ROV_4	0%	39%	0%	63%	51%
ROV_5	0%	0%	0%	0%	0%
ROV_6	63%	0%	0%	88%	0%
ROV_7	0%	0%	66%	53%	0%
ROV_8	0%	22%	50%	53%	0%
ROV_9	80%	50%	51%	78%	0%
ROV_10	64%	79%	53%	48%	0%
ROV_11	75%	74%	47%	0%	47%
ROV_12	25%	64%	41%	93%	0%
ROV_13	-	-	-	-	-
ROV_14	-	-	-	-	-
ROV_15	-	-	-	-	-

6 G1 and 9 G2 classrooms were observed. In 5 classrooms across the two grades, teachers taught a lesson plan from Day 4 or Day 6 of the week, where they were either conducting their own activities or conducting assessments + remediation; MC = Mathematical Conversation

Grade 1	Grade 2

High fidelity	Medium Fidelity	Low Fidelity	Not a part of the day's LP

1 Length of the numeracy lesson plan and amount of preparation required for the ‘Math Games’ section may be leading to teachers skipping it altogether

- **'Math Games - We Do + You Do' conducted in upto some of the expected G1 & G2 classrooms. Other sub-sections conducted in many the classrooms.**

- **Wherever conducted, teachers spent only 7 - 8% of the total Numeracy class time on ‘Math Games’ section.**

- **Key reasons cited by teachers for not conducting this section:**
  - Instructions not understandable.
  - Long prep time, and so is sometimes, skipped.
  - Students take more time than suggested in the TG to grasp the concepts in periods 1 and 2, leaving little time for period 3. On the same lines, many teachers said they are unable to complete the lesson plan in the suggested time.

- **However, some teachers said that children learn faster in Maths compared to Hindi.**
  - In Hindi, a few teachers and ARPs mentioned that children struggle with understanding alphabets and maatras (specifically ए , ऐ, ओ and औ).

“...chhote bachhe hai na sab, chote bacchon ko sikhane me samay lagta hain to pehla dusra period me humne mera samay tisre period ka bhi ho gaya to ab hum usko chod diye”

- Teacher, during FGD

"..Kyunki khel gatividhiyan unko khud se tayyar karni hai. To thoda unko…jaise gaon ki bhasha mein hamare mehnat kehte hain…Aur bhasha mein kya hai sab barabar hai tisra class bhi remediate ka aaram se kar lete hain ganit mein bhi 2 class karne mein koi dikkat nahi hai tisre mein…."

- LLF Member, during interview

Annexure 8: Common Findings across Subjects

1 Belief in inherent student capabilities and consistent checking of workbooks by ARPs might be making teachers conduct the You-do as a We-do

- **The 'Workbook Practice - You Do' sub-section was the only section that was conducted in a majority of the expected Grade 1 and Grade 2 classrooms.**
- **In more than one-third of all classes, students did not fill their workbooks independently. This is a negative shift from the baseline.**
  - Teachers either directly shared the answers, or got some students to solve the questions on the board, while others copied.
  - Some teachers in interviews and FGDs, highlighted the need for 'assistance' in You-Do. Since students are young or all students are not the same and all students cannot be expected to be "IAS officers". The latter, to them, is a well-accepted fact.
  - Another reason for this could be the consistent checking of workbooks by ARPs as reported by teachers in the FGD. A few ARPs checked workbooks during our joint visits.

"To kya aap jis family se aate hain sare log ek hi jagah pe hain 1 hai IAS ban gaya to sare log IAS ban jate hain..to fir aap kalpana kaise kar sakti hai ki wo baccha wo sikh liya to ye bhi sikh liya.....Aur ye prakriti ke niyam hai aur prakriti ke niyam se ladai kaise ki sare bachche barabar ho jayenge..How this is possible?"

- Teacher, during FGD

"Jaise ARP aaye…toh kahenge....aapne workbook nahi bharwaya…to ho sakta hai unke class me zyada bacche ho wo kya hain ki jaldi usko pressure me jaldi jaldi usko karwane ke liye usko aise likh di ki bacche fat se utaar lenge kyunki dekh ke hi utaar na hai..."

- Teacher, during FGD

2 Teachers, ARPs and LLF members indicate pressure to achieve NIPUN Goals, which may lead to teachers and ARPs prioritising learning outcomes over structured pedagogy

- **Many teachers, ARPs and some LLF members said there is pressure to achieve NIPUN Goals.**

- **A few teachers said:**
  - Officials check whether students are NIPUN by Sep-Oct, much before the end of the academic year. An ARP confirmed this.
  - There is focus to make students practice questions from mobile assessments because they repeat across tests.

- **A few LLF members said:**
  - Due to the pressure, teachers do not follow the TG and end up teaching students via rote-learning method.
  - BOs pressurise ARPs to assess students every 15 days, instead of monthly, because they need to share these numbers with the BSAs. A few ARPs said that they were instructed to only assess NIPUN students to keep results high.

- **Another indicator of pressure is the presence of multiple assessments a student has to take every month.**
  - They include daily App assessments conducted, weekly TG assessments, Spot Assessments by ARP, and third-party assessments by DIET.
  - A few LLF members said that in smaller classrooms, the same set of students are assessed multiple times every week.

"…hum log ke ek aur dikkat hai ki yahan per saath hi dabav pad jata hai. November tak kariye, December tak kariye, January tak kariye kyunki kahin election jana hai…toh woh bhi samasya hai…..bhai, saal bhar kiya hua hai toh beech mein karenge toh thoda usko bhi dikkat hoga aur isko bhi dikkat hoga."

- ARP, during interview

"…Sir ne bataya tha ki aise, NIPUN main jo App mein jo kahaniya hai aur questions hai jo main kahaniya hai unko 45 minutes padhna hai, usko aap kya kariye ki aap charts pe likh dijiye, charts pe likh ke saamne board pe chipka dijiye, bache dekhte dekhte usko practice karenge usko padhenge toh jaldi se aa jayega. ."

- Teacher, during interview

3 While many teachers expressed positive opinions about the TG, late delivery of programme materials is still a key issue

- **While many teachers find the TG easy to use because of a clear structure, some find TLMs to be particularly effective or engaging.**
  - According to a few ARPs, and on-ground observations, one TLM, however, **the Math Kit, is not being used by some teachers,** either because they haven't found them to be effective or because they're afraid of damaging it.
- **These same materials, unfortunately, were delivered late in most schools, as reported by most teachers, and all ARPs.**
  - In a few newly converted English schools, English versions of materials are still not available, forcing them to use notebooks, instead of workbooks.
  - According to a few ARPs, the reason behind this is that these schools are sometimes not updated on the Block Resource Centre's list.
- **One finding that may explain higher programme fidelity in G2 rather than G1 is the amount of time spent teaching the content in the TG for the two subjects across grades.**
  - For Grade 1, ~60% of the time spent on the TG is for literacy.
    For Grade 2, on the other hand, it is the opposite: ~60% of the time spent on the TG is for numeracy.

" Jo shikshak sandarshika hai isme sari chizein itne ache se clear kar ke diya hai ki hum ko kuch sochna nahi padta hai ki hum ko aaj kya karna hai….. Isme saari chizein sequence se itne ache se likhi gayi hai."

- Teacher, during interview

"Keval english medium schoolon ki thodi problem karya pustika ko le kar… Material jaise hi aaya hamare BRC pe…distribute kar diye gaye..English medium schoolon ki thodi si problem hai, unko English medium kar diya gaya, jab paathye pustikaayein aati hain wahan se toh na unka hota naam hi nahi hota hai list mein..

- ARP, during interview

4 No significant gender-biased actions observed across classrooms, however, some teachers seem to think of girls as more obedient and boys as more confident

- **Enrollment numbers varied across the two grades, but not significant difference between the attendance of boys and girls**
  - 60% of the students enrolled in the G1 classrooms were girls, compared to only 46% in the G2 classrooms.
  - On average, 73% and 66% of the students enrolled in the G1 and G2 classrooms observed respectively were present on the day of the classroom observation.

- **Overall, it seems like teachers in FLN grades do not overtly show any gender-based biases.** This is in line with the baseline finding.
  - All teachers asked questions to both girls and boys during the class.
  - In all classrooms, boys and girls answered roughly an equal number of questions.
  - However, students were sitting according to their gender in a few classrooms.

- **The mindsets of some teachers, however, tell a slightly different story.**
  - Some teachers said that boys and girls participate equally in class, while a few said that boys engage more in numeracy, and girls engage more in literacy.
  - The few teachers who think that girls engage more across subjects attribute it to their obedience towards the teacher.
  - In terms of performance, while many teachers thought it to be equal across genders, a few of them reported that boys learn and perform better than girls in Math.
  - A few also remarked that girls fumble when answering questions in class while boys are more confident with their responses.

Annexure 9: Findings related to ARP support from the Follow-Up Qualitative Study

Findings related to ARP Support: Classroom observations and spot assessments conducted by most ARPs, Feedback to the teacher can be strengthened

OVERVIEW		
Activity Name	**Conducted by_/ 9 ARPs**	**Average Time Spent (In minutes)**
Classroom Observation	8	24 (Range - 5 to 42)
Spot Assessment	7	27 (Range - 15 to 70)
Conversation with Teacher	5	11 (Range - 5 to 27)
Conversation with HM	6	29 (Range - 5 to 60)

JOINT VISIT SCORES				
ARP Code	**Classroom Observation**	**Spot Assessment**	**Conversation with Teacher**	**Conversation with HM**
ROV_1	33%	0%	0%	0%
ROV_2	50%	33%	0%	50%
ROV_3	33%	100%	30%	33%
ROV_4	67%	83%	70%	83%
ROV_5	67%	67%	0%	0%
ROV_6	0%	67%	60%	33%
ROV_7	17%	0%	0%	0%
ROV_8	67%	50%	60%	67%
ROV_9	50%	33%	60%	50%

High fidelity **Medium Fidelity** **Low Fidelity**

1 While all ARPs said that they visit the mandated 30 schools in a month, lack of sufficient time and drawbacks of NIPUN Lakshya App reported as impediments to effective school visits

- **All ARPs said they visit 30 schools in a month. Some communicate more frequently with the adopted "god liye hue" schools.**

- **Many ARPs mentioned a lack of sufficient time to perform all activities/ visit all schools in a month.**
  - Some mentioned they are unable to visit all schools in a day, either due to several meetings, or the long travel time between two schools.
  - Some ARPs mentioned that conducting assessment is time consuming.
  - During joint visits, ARPs spent half an hour on average with the HM collecting data, which only 11 minutes on average were spent with teachers. This was confirmed by a few LLF members.

- **Most ARPs pointed out issues with the NIPUN Lakshya App.**
  - Over half of them mentioned that due to technical or network issues, the App does not declare deserving students NIPUN.
  - A few ARPs said that the App does not catch students' voice properly and another said it does not consider speech impediments.
  - A few ARPs mentioned that the questions in the App are repetitive, leading to students memorising answers for the assessments.

"... ARP hamare jo hain abhi jo academic monitoring, academic support karna chahiye wo academic nahi kar paate hain kam kar pa rahe hain. Kyunki unka jo tool hai itna lamba hai unko data lene mein 1 ghante lag jata hain…..

- LLF Member, during interview

"...Baccho ne rat rakha hai…..abhi jo wo padh rahe the 120 125 130 ki speed pe padh rahe the, jab nayi kahani aati hai to 60 65 67 aise phaunch paate hain"

- ARP, during interview

2 Most ARPs conducted classroom observations and spot assessments, but many key guidelines were not followed

- **Most ARPs observed classrooms but spent, on average, a little more than half the recommended time conducting this activity.**
  - ARPs, on average, spent 24 minutes observing classrooms, while a few teachers said ARPs spend 10-30 mins in class.
  - **Most ARPs also interrupted the class,** usually to interact with students and ask them questions related to the topic being taught.
- **Some ARPs did not choose students randomly for Spot Assessments. This was confirmed by many LLF members.**
  - Most of them directly asked the teacher to select students for them.
- **Some ARPs did not conduct the entire assessment with the required number of students**.
  - While some assessed less than 5 students, others did not ask all the 3 digit-addition and subtraction questions.
  - One of them cited lack of space in the classroom for all 5 students to work out questions in their notebook as a reason for the above.
- **Some ARPs helped students with answers during the assessment.**
  - Moreover, a few ARPs marked all questions related to a topic correct if they thought that the child seemed to have understood the concept.

"Dusra example ye hai ki jab woh baccho ka assessment karte hain toh generally, teachers kehte hain ki sir ye baccha bahut tez hai, is bacche ko utha lijiye, sir ye bachcha bahut tez hai. Toh ARP ko bhi lagta hai ki chalo mera data accha jayega."

- LLF member, during interview

3 Generic feedback from ARPs, as well as lack of demos and written feedback makes teachers think that ARPs offer ‘suggestions’, rather than ‘sahyog’

- **ARPs spent, on an average of 11 mins conversing with teachers, while the stipulated time is 40 minutes.**
  - The few ARPs who spent the most amount of time (about 30 minutes minutes) were the only ones who gave demos.
- **Many ARPs gave verbal positive and critical feedback to the teachers, while no ARPs gave written feedback.**
  - However, most verbal feedback was generic in nature, and not focused on specific teaching practices.
- **Only some teachers find ARP’s advice to be useful.**
  - In the FGD, teachers talked about how feedback from ARPs are merely operational suggestions, and not really advice.
  - Some of this feedback is also seen as knowledge that teachers already possess.
- **During the FGD, teachers clearly pointed out they want ARPs to give specific solutions as well as demonstrate these solutions.**
  - Demos need to happen with students in classrooms, rather than in cluster-level meetings, to help ARPs understand the challenges of a teacher.

“Mera ye kahne ka matlab ki puri class me ye observe kare jo meri kami ho us kami ko usko kar ke bataye… agla jo humko observe kar raha hai to meri jo kami hai vo humko bataaye, aur usko is tareeke se aap kariye usko sujaav de, kar ke dikhaaye”

- Teacher, during FGD

“Ek baat samjhaiye madam teachero ko bacha samajh ke karna aur bachon ke saath ab khud relate hona bahut antar hain..aap claas se jab karenge na connect tab pata chalta hai ...vo humari class me aa kar ek baar demo de.”

- Teacher, during FGD

4 Data collection seems to be a high-priority for many ARPs, and spot assessment data is used to provide differentiated support

- **Many ARPs had a conversation with the HM, and spent on average, 29 minutes on it, which is almost 3x that of the average time spent with the teacher (11 minutes).**
  - Some ARPs talked about how far the school is from achieving the goals.
  - Only one ARP asked the HM/ other teachers about the challenges faced in NIPUN. This was more of a monologue than a dialogue.
- **A significant part of conversations with HMs included data collection.**
  - Data points included student attendance, information on teachers, parent registers, sports equipment and other materials, photos of library, etc.
- **Some ARPs said that the spot assessment data or the attendance data collected is used to categorise schools based on performance such as Green-Yellow-Red or A-B-C.** The weaker performing schools are then provided extra support.
- **A few ARPs mentioned discrepancy between data seen online and on-ground**. For e.g, an ARP said that a school that is in red category according to him is shown in green category, number of resources delivered on-ground vs online do not match.

"..Jaise maan lijiye hum kisi vidayalay ko de rahe hai ki wahan per sabke pass sandharshika hai. Aur aapke yahan se agar data aa raha hai ki nahi itne logon ke paas sandarshika nahi hai, ya toh kahin na kahin se koi na koi vyakti usme kuch na kuch kar pa raha hai……unka base kya hai yeh aaj tak hum logon ko samajh nahi aata ki woh nikalte kaise hai.."

- ARP, during interview

Annexure 10: Inputs from the Implementation Partners (LLF) from the Follow-Up Qualitative Study

At the classroom level, many of LLF's broad observations areas align with our findings

- **4/6 members said that teachers are struggling with or not implementing Ganit Khel well. The reasons are:**
  - One said teachers have to put in effort in creating activities and lack of teachers in some schools.
  - Another ARP said teachers are treating Ganit Khel as a leisurely game, like in Literacy. They aren't aware of conceptual benefits of learning Maths through games.
  - The third ARP said teachers do not understand how to associate different items in Math Kit with different competencies.
  - An ARP said that teachers sometimes do not use materials thinking that most students have understood the concept.
  - Two ARPs said Ganit Khel is not being done as per TG because trainings haven't happened.

- **A few members said that teachers are not using TLMs.**
  - One said it is because they haven't understood how to use some of them and non acceptance of newer methods.
  - The other said that teachers are skipping usage of TLMs even though TG has clearly mentioned using it. He further goes on to say that teachers think that as long as students identify numbers they have understood the concept. They do not teach the other aspects of understanding - *quantity, symbol and association*. Failure to accept newer methods of teaching is the cause behind it.

- **A few LLF members said that teachers help students during the You-Do sections.**
  - One of them said that teachers give instructions to students before beginning WB work. In some schools, teachers whilst observing students, support those who are struggling to do WB work.
  - The other member said that teachers first teach students the concept via We-do and then ask them to write answers on their own.

At the classroom level, many of LLF's broad observations areas align with our findings

- **Almost all members said that the lack of training affected the way teachers conducted the Lesson Plan in classrooms.**
  - Most of those members said that -Teachers struggle with new content added in the TG.
  - One member cited 'Ganitya Khel' and other 'gatividhi' as examples.
  - Another member mentioned that teachers are struggling with 'Khoje aur Jaane'. Teachers are confused about how to conduct 'Khoje Aur Jaane' activities without giving homework since it is not allowed to give HW to Grade 1 & 2. Overall, 2 LLF members said that teachers are struggling with 'Khoje aur Jaane'.
  - A third LLF member cited a lack of structure in remedial classes.

- **2/6 members said that children haven't understood the concept of 2-digit abstract addition. They consider it as addition of two single digit numbers. One of these two members said that another way children add 2-digit numbers in an abstract form is by drawing lines for each of the 2-digit numbers and then counting the total.**

- **Half the LLF members pointed out issues with Literacy implementation in classrooms. Some of them are:**
  - One member said that teachers skip open and close ended questions in Maukhik Bhasha because they do not prep for it.
  - A member said that children copy teachers in the Pathan period, often just lip-syncing what she is saying instead of actually learning with understanding.