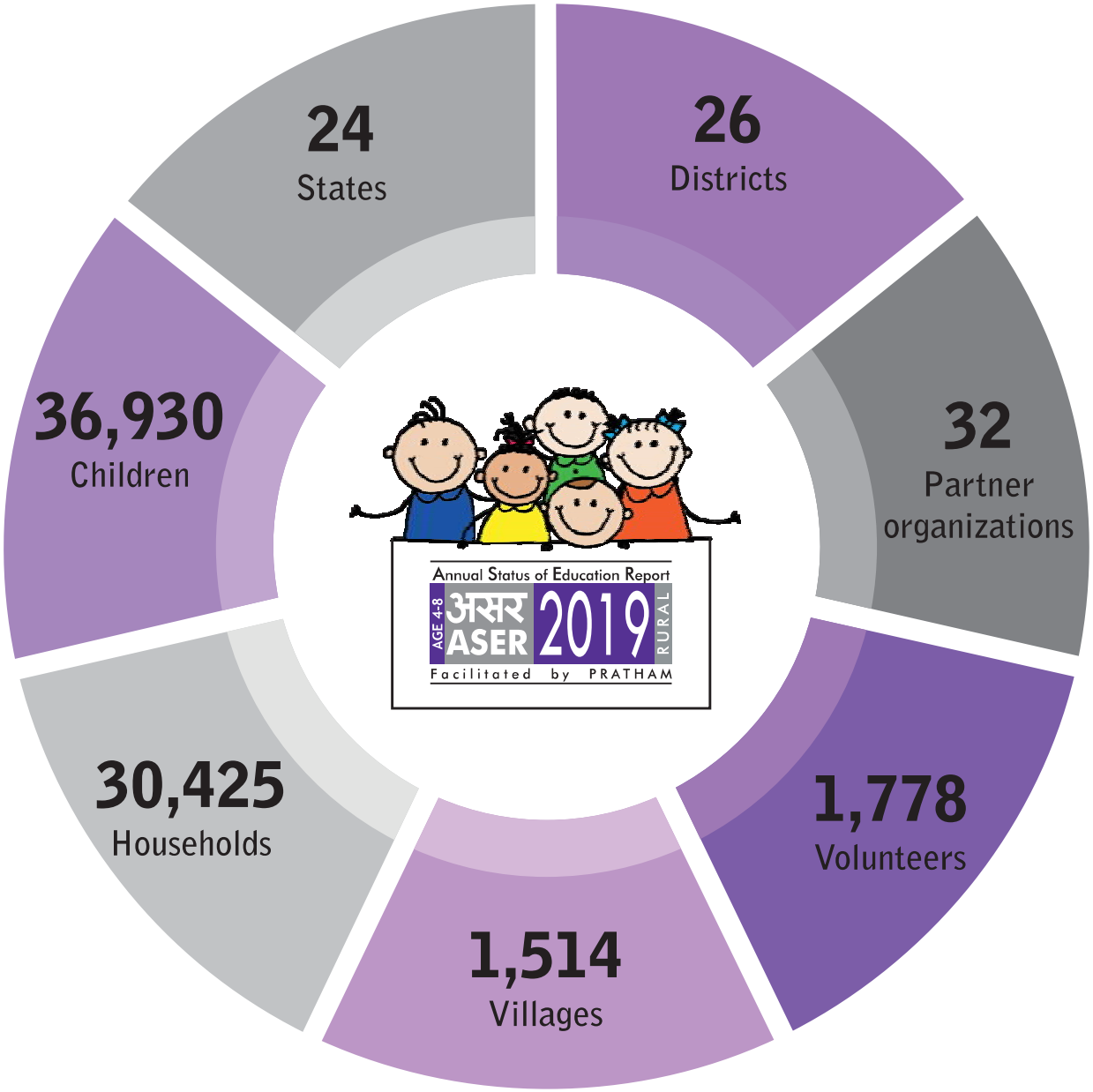
About ASER 2019 'Early Years'

ASER 2019 'Early Years' – Coverage



ASER 2019 'Early Years' – Summary of the survey process

A team of two volunteers goes to the village assigned to them by their ASER state team. They carry the material required to conduct the survey.

Once in the village, the volunteers meet the Sarpanch/Village representative. They explain what ASER is and ask for support to conduct the survey in the village.

The volunteers then walk around the entire village. They make a map showing the important landmarks in the village and fill the Village Information Sheet based on the facilities they observe. Next, the volunteers select the households. To do this, the volunteers:

- Divide the map into 4 sections or select 4 hamlets.
- Randomly select 5 households with children in the age group 4 to 8 from each hamlet/section using the 'every 5th household rule'.
- Follow this process in every hamlet/section in order to survey a total of 20 households in the village.

Volunteers record some basic information about all the households they visit during the sampling process. This includes: hamlet/section number, locked or no response household, name of head of the household, number of children in the age group 4 to 8 who regularly live in the household, and mobile number of the household.

In each sampled household the volunteers do the following tasks:

- Record information about children in the age group 4 to 8.
- Use the Testing Tool to assess children in the age group 4 to 8 on cognitive development, early language, early numeracy, and social and emotional development.
- Record children's response for each test item administered.
- Record information about parents' education and household assets.

After all 20 households are surveyed, the volunteers check and then submit the completed Survey Booklet to their ASER state team.

ASER 2019 'Early Years' – Early learning tasks

Young children's development can be assessed along many dimensions. While all of these are important, the guiding principles while creating an assessment for the ASER survey require it to be rapid, easy to administer consistently to children in the target age group in the household, and easy to understand by all stakeholders: children, volunteers doing the administration, and parents or community observing the assessment as well as researchers, educationists, and policy makers. The tasks for the ASER 2019 assessment were created keeping these principles in mind.

At the outset, a mapping of children's learning trajectories, policy guidelines, curricular frameworks, and learning outcomes defined for pre-school and early primary school (Std I and II) was conducted. Based on this exercise, an initial set of domains and tasks was developed, piloted extensively, and refined over several months. For the final assessment, four key domains were selected: cognitive development, early language, early numeracy, and social and emotional development. These broadly seek to address four key questions: Can children do simple problem-solving tasks? Do children have early language skills? Do children have a sense of numbers and quantity? Are children able to identify emotions? On average, four competencies within a domain were retained to provide an overview of what young children in India can and cannot do. The complexity of tasks within each domain is varied to accommodate expectations from the youngest child (age 4) to the oldest (age 8) in the survey age group.

Like every ASER, the assessment was conducted at home rather than in school, so as to include children not enrolled anywhere as well as children attending different types of pre-schools and schools. All children in the 4 to 8 age group in a sampled household were assessed using the same tool, irrespective of grade or schooling status. The ASER assessment process incorporated various elements aimed at ensuring that it captured the best that each child can do. Volunteers were trained to build rapport with children to create a relaxed and encouraging environment. The test was adapted in 14 Indian languages. The volunteers were trained to speak slowly and clearly to ensure that all children are able to fully understand the expectation from the task. Children were given the time they needed to do each task. The assessment process was adaptive to the child's ability, so that she did not have to attempt all levels. Thus, the child's comfort and the commitment to accurately record her best possible response was at the core of the test design. The following pages describe the testing process for tasks in each domain in ASER 2019.

I. Cognitive development

Cognitive ability helps build problem-solving, memory, logical reasoning, and creative thinking skills in young children. These skills help children think, explore, and make meaning of the world around them. They are fundamental to any kind of learning in school or in life. ASER 2019 tasks for this domain comprised sorting by color, spatial awareness, seriation, pattern recognition, and puzzle.

1. Sorting by color

Children are given four shapes of different sizes and colors. As an example for the child, the volunteer separates the blue shapes from the rest. Then the child is asked to sort the remaining shapes into separate groups by color. This question is administered to children in the age group 4 to 5 only.

2. Spatial awareness

The volunteer shows the picture to the child and asks two questions:

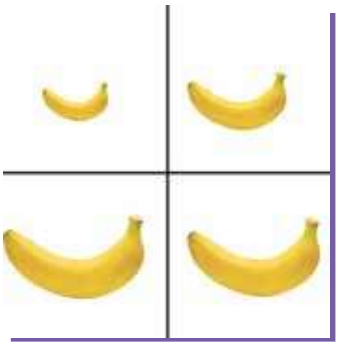
i. Which boy is on top of the tree?

ii. Which boy is farthest from the tree?

This question is administered to children in the age group 4 to 5 only.

3. Seriation

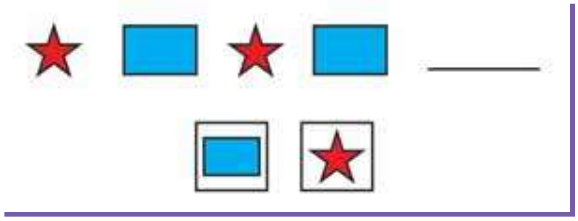
The volunteer shows the child the four cards and asks her to arrange them in a line from the smallest to the biggest banana. Even if the child correctly arranges the cards from the biggest to the smallest banana, this is considered correct.



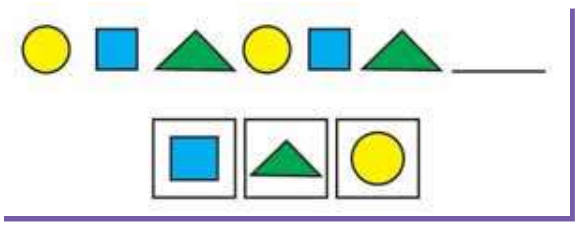
4. Pattern recognition

The volunteer describes a two-picture pattern to the child as an example: "Look, the star comes first, then the box comes next, then the star comes again, and then the box comes again". Then the child is shown the three-item pattern question and asked to point to the shape (from the options given) that should fill the blank space.

Example



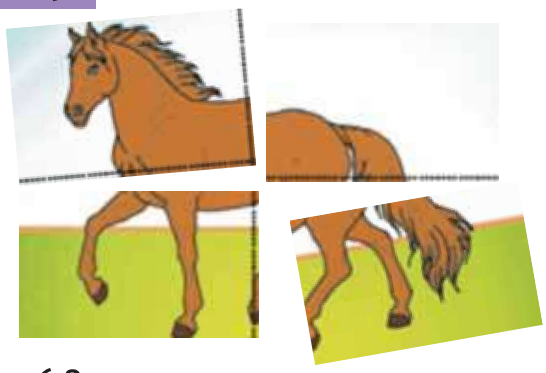
Question



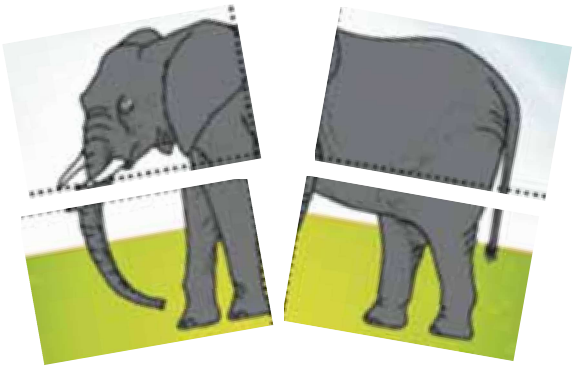
5. Puzzle

The volunteer solves the horse puzzle in front of the child, as an example of how to solve a puzzle. Then the child is given the elephant puzzle to solve. Children in the age group 4 to 5 are given a four-piece puzzle and children in the age group 6 to 8 are given a six-piece puzzle to solve. A separate picture of the elephant remains in front of the child while she is fitting the puzzle pieces together.

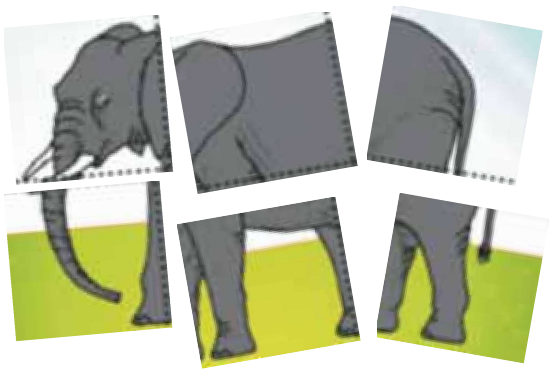
Example



Age 4-5



Age 6-8



II. Early language

Language is an essential skill for communication, needed to express and communicate thoughts, feelings and needs, as well as to understand and have conversations with others. Language is also the foundation for literacy. The ability to read begins when children learn to connect shapes of written letters with their sounds. Comprehension is essential to make meaning of any communication. The following tasks were included in ASER 2019 to provide a snapshot of early language acquisition in young children: picture description, listening comprehension, reading (letters, words, a short paragraph equivalent to a Std I level text), and reading comprehension.

1. Picture description

The volunteer shows a picture to the child and asks her two questions based on it:

i. What are people doing here?

ii. What can you see outside the window?

A comprehensive repository of answer options collated from extensive piloting was available with the volunteer for grading. Correct answers in the local dialect and synonyms were marked as correct.

2. Listening comprehension

The volunteer reads out a four-sentence story. Each sentence is short, comprising about six simple words.

The volunteer then asks two simple questions on the story to the child:

> Geeta wanted to fly a kite. Geeta made a kite with paper.
> Then she tied a thread to the kite. Geeta was happy to fly the kite.

i. Who wanted to fly a kite?

ii. What did Geeta tie to the kite?

In view of the short attention span of young children, the story can be read up to three times to a child. The volunteer is trained to take cues from the child on her comprehension of the task and make decisions about repeating the story.

3. Reading

Three types of reading tasks are administered to children.

i. **Letters:** A set of commonly used letters is shown to the child. To be marked correct, she should read at least 5 out of 8 letters correctly.

ii. **Words:** If the child is able to read at least 5 letters correctly, she is asked to read words. Common, familiar words with two letters and one or two matras are shown to the child. To be marked correct, she should read at least 5 out of 8 words correctly.

iii. **Std I level text:** If the child is able to read at least 5 words correctly, she is asked to read a Std I level text. A set of four simple linked sentences, each having no more than six simple words is shown to the child. The words and sentence construction are compatible with the content of Std I textbooks. To be marked correct, she should read the Std I level text fluently and make three or less than three mistakes while reading it.

i	ii	iii
क ल स न य ड म व	पीला मैना दिन आग बूढ़ा तोता मोर खुश	रूपा बाहर खेल रही थी। खेलते–खेलते रात हो गई। फिर रूपा अपने घर चली गई। वह खाना खाकर सो गई।

4. Reading comprehension

If the child reads the Std I level text, the volunteer asks two simple comprehension questions to the child. She is free to read the text a second time after listening to the questions.

> बाहर कौन खेल रहा था?
> रूपा ने खाना खाने के बाद क्या किया?

III. Early numeracy

Counting, measurement and sense of quantity are essential early numeracy skills. A strong foundation in early numeracy helps not only in school math but also in everyday life. The following tasks were included in ASER 2019 assessment to provide a snapshot of young children's early numeracy skills: counting objects, relative comparison of objects, oral word problems, single- and double-digit number recognition, relative comparison, and numerical operations.

1a. Counting objects

This question is administered to children in the age group 4 to 5 only.

The volunteer shows the picture to the child and asks her to identify each object. The volunteer names any object that she is unable to name, and then asks four questions:

i. Count how many umbrellas are there?
ii. Count how many balls are there?
iii. Count how many fish are there?

1b. Relative comparison of objects

iv. Which of these objects is the most in number?

2. Oral word problems

One at a time, the volunteer reads out two word problems. These tasks involve simple single-digit addition and single-digit subtraction, so that child can solve them mentally or by using her fingers.

Seema has 3 toffees. Her sister gave her 5 more toffees. Tell me, how many toffees does Seema have in total?

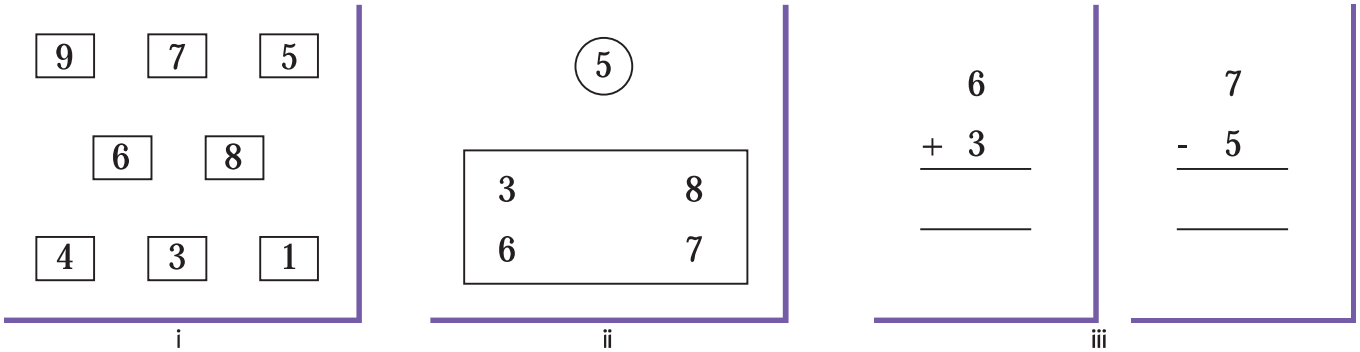
Reena has 7 bananas. She gave 3 bananas to her friend. Tell me, how many bananas are left with Reena?

3. Single-digit arithmetic tasks

i. **Number recognition:** A set of 8 numbers from 1-9 is shown to the child. She is asked to read them aloud. If the child is able to recognize at least 5 numbers correctly, she is asked to do the next two tasks involving single-digit relative comparison and single-digit numerical operations.

ii. **Relative comparison:** The child is asked to look at all the numbers in the question. The volunteer points at the number in the circle on top. He then asks the child to look at the numbers in the box below and point to the number that is smaller than the one in the circle.

iii. **Numerical operations:** One at a time, the volunteer copies each numerical problem onto a sheet of paper and asks the child to add (or subtract) the numbers. Oral answers are also accepted.



i ii iii

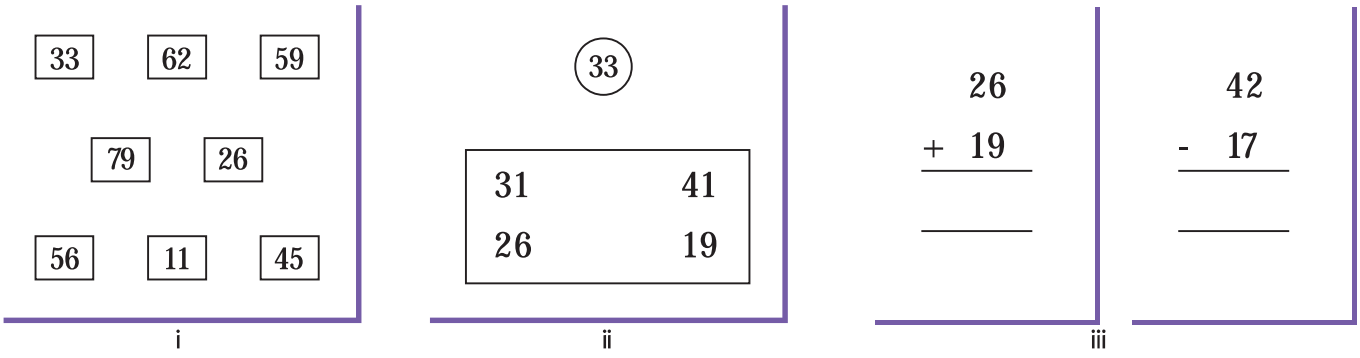
4. Double-digit arithmetic tasks

If the child is able to recognize at least 5 single-digit numbers correctly in task 3(i), she is asked to do the double-digit number recognition task.

i. **Number recognition:** A set of 8 numbers from 11-99 is shown to the child. She is asked to read them aloud. If the child is able to recognize at least 5 numbers correctly, she is asked to do the next two tasks involving double-digit relative comparison and double-digit numerical operations.

ii. **Relative comparison:** The child is asked to look at all the numbers in the question. The volunteer points at the number in the circle on top. He then asks the child to look at the numbers in the box below and point to the number that is greater than the one in the circle.

iii. **Numerical operations:** One at a time, the volunteer copies each numerical problem onto a sheet of paper and asks the child to add (or subtract) the numbers. The addition problem is with carry and the subtraction problem is with borrowing.



IV. Social and emotional development

The ability to identify their own as well as others' emotions and perspectives, establish and maintain relationships, and take responsible decisions are important skills for all human beings. Helping young children develop these skills early gives them important tools for the future. ASER 2019 tasks in this domain consisted of emotion identification, situation to emotion mapping, and situation reaction test.

1. Emotion identification

The child is shown four face cards with different emotions. She is asked to answer four questions by pointing to a card:

i. Which of these girls is sad?

ii. Which of these girls is angry?

iii. Which of these girls is afraid?

iv. Which of these girls is happy?

2. Situation to emotion mapping

Leaving the cards in front of the child, the volunteer reads out four hypothetical situations one-by-one to the child:

i. Your mother/father bought you a new toy. Point to any one of these and show me how you would feel.

ii. Your friend ate all your candy. Point to any one of these and show me how you would feel.

iii. You were sleeping at night. You heard a sound coming from the dark. Point to any one of these and show me how you would feel.

iv. You were about to play with your toy but your father stopped you from playing. Point to any one of these and show me how you would feel.

The volunteer records the code given on the card as the child's answer.

3. Situation reaction test

One by one, the volunteer reads out two hypothetical situations to the child. One situation is based on conflict and one on empathy:

i. You had only one toy. You were playing with your friend. While playing, your friend snatched your toy. What will you do now?

ii. You bought a toffee. You saw that a child was crying because his toffee fell in the mud. What will you do now?

A comprehensive repository of answer options collated from extensive piloting is available with the volunteer for grading.

Sample design of rural ASER 2019 'Early Years'

Wilima Wadhwa, Director ASER Centre

ASER 2019, also referred to as 'Early Years', is an assessment of 4 to 8 year olds in the ASER architecture. In other words, it is a rapid assessment of children, done in households, by ordinary citizens. Each of these elements creates its own challenges for the design of the assessment. The assessment was done simultaneously across the country in October and November 2019.

Like 'basic' ASER, the 'Early Years' survey also has a two-stage sample design. In the first stage, in each surveyed district, villages are randomly selected from the Census 2011 village directory. In the second stage, households are randomly selected in each of the villages selected in the first stage. This sampling strategy generates a representative picture of each district. One rural district has been surveyed in each major state, with the exception of Uttar Pradesh and Madhya Pradesh where two rural districts have been surveyed. While this is not a nationally representative sample, the size and geographical spread of the sample enables the estimates to be aggregated to get an overall picture of the rural population in India.

The ASER 2019 sample consists of 26 districts spread across 24 states[1]. The districts were not sampled randomly and care was taken not to choose districts that were anomalous in terms of their learning outcomes as measured in ASER 2018. First, since the main focus of the survey is on the learning levels of young children, districts where average learning levels were more than 5 percentage points higher or lower than the state average were not considered[2]. Second, in each state, partner organizations were approached in districts that had learning levels close to the state average. Eventually, the final districts to survey were chosen for logistical convenience based on partner availability and their ability to provide volunteers during the survey period.

In each surveyed district, 60 villages[3] are sampled from the Census 2011 frame using the probability proportional to size (PPS) sampling method in the first stage. This method allows villages with larger populations to have a higher chance of being selected in the sample. It is most useful when the first stage sampling units vary considerably in size, because it ensures that households in larger villages have the same probability of getting selected into the sample as those in smaller villages, and vice versa[4]. In the second stage, households/children in the age group of 4-8 years are surveyed in each sampled village.

There are various issues that complicate the second stage sampling. First is the issue of sparse populations. ASER as well as our research studies have shown that simply sampling households in the village may not result in sufficient sample sizes of children in the age group of 4-8 years. For instance, in ASER 2018, where 20 households were randomly selected in a village, only about 8 households had children in the age group of 4-8. The best solution to this problem of sparse populations of interest is to create a listing of the target population (for a particular cluster) and sample from that, thus, employing a stratified sample. However, given the rapid nature of assessment in ASER and several resource constraints (time, people, money), ASER does not stratify at the second stage - there is no houselisting done at the village level.

Second, the absence of a houselisting creates additional problems in surveys that are representative at multiple levels of aggregation. In these surveys estimates have to be weighted with appropriate weights[5] to account for different underlying population sizes - a more populous state like Uttar Pradesh will have a higher weight in the national estimate than a state like Himachal Pradesh. The calculation of these weights requires the underlying population proportion of the target group of interest. So, if the household were the unit of sampling then we would need the number of households in the village to calculate the weights. On the other hand, if children in the age group of 4-8 years were our target population, we would need the total number of such children in the village to calculate the weights. A houselisting of the village would provide not only the frame for sampling these children, but also the total number of such children in the village.

ASER resolves both these problems by sampling households. Household weights are easy to calculate since the Census provides the village population of households. Therefore, the sample in ASER is defined in terms of households and not children. In 'basic' ASER, all children in the age group of 3-16

[1]States and Union Territories not represented in the sample are Chandigarh, Delhi, Sikkim, Jammu and Kashmir, Arunachal Pradesh, Mizoram, Daman and Diu, Dadra and Nagar Haveli, Goa, Lakshadweep, Puducherry, and Andaman and Nicobar Islands.

[2]Estimates from ASER 2018 were used to compare district learning levels of 4- to 6-year-olds with the state average.

[3]We decided to double the number of villages per district as compared to ASER, so as to get a larger district sample since the 'Early Years' survey is being done in only one district per state.

[4]Most large household surveys in India, like the National Sample Survey and the National Family Health Survey also use this two-stage design and use PPS to select villages in the first stage.

[5]The weight associated with each sampling unit, household in ASER, is the inverse of the probability of it being selected in the sample.

years living in the sampled households are surveyed. So as to get a representative sample of the household distribution, even households with no children in the target age group are counted as part of the sample. The number of households and villages in ASER has remained more or less unchanged since 2006, though the number of children surveyed fell by about 25% between 2006 and 2014[6]. However, given the scale of ASER and large household sizes in rural India, this strategy yields large enough samples to do age-wise or grade-wise analysis at the state level.

Given that we wanted to retain as much of the rapid nature of assessment of the ASER architecture[7] as possible in ASER 'Early Years', houselisting at the village level was not an option. Following the ASER sampling strategy would have given us a representative distribution of households but may not have generated a large enough sample size for the target population. ASER 2018 data suggested that we would get about 11 children in the 4-8 year age group by sampling 20 households per village. This is a sample size problem and can be overcome in a number of ways - for example, by sampling more villages per district and/or more households per village[8]. However, both strategies have consequences: increasing the number of villages has cost implications and increasing the household sample in a village does not necessarily result in higher precision if the intra-cluster correlation is high. In any case, going by the ASER numbers to get a reasonable sample size (about 1,000 children per district) would require sampling close to 100 villages in each district and about 40 households in each village.

Another strategy could be to sample only households with members of the target population as is done in the National Family Health Survey. However, as discussed earlier, this would require creating a frame of the target population in the village, which would be used to both sample and calculate weights.

Finally, we adopted a sampling strategy that modified the ASER approach, so as to get sufficient sample sizes and be able to calculate weights without doing a houselisting in the village. The standard ASER sampling approach in the village is to mimic simple random sampling without doing a houselisting. Volunteers walk around the village, make a map, divide the village into four parts, and sample 5 households using the 'every 5th household rule', in each part, to get 20 households in the village. Households with no children in the target age group count as part of the sample since the aim is to get a representative picture of the household distribution.

In the ASER 2019 'Early Years' survey we modified this approach so as to capture sufficient numbers of 4- to 8-year-old children. The process is described below:

1. Walk around the village and make a map and divide the village into four parts.
2. In each part go to a central location and use the 'every 5th household rule' starting from the left to sample households.
3. If the household has children in the 4-8 age group currently residing in the household, record the household number, and the number of such children. Administer the survey to all children in the target age group in the household and collect information about the household. Proceed to the next 5th household.
4. If the household has no children in the 4-8 age group, record the household number and the fact that it has no children in the target age group and move to the next household. Note that unlike 'basic' ASER, we do not record household characteristics in households with 'no children'[9].
5. If the household is locked or does not want to participate in the survey, record the household number and the fact that it was locked or a 'no response' household and move to the next household.
6. Continue this process until you have administered the survey in 5 households with children in the 4-8 age group in each of the four sections of the village.

At the end of the survey in the village, this procedure will yield 20 households with completed survey information, as well as

[6]The drop in the number of sampled children is probably due to the increase in the number of rural households since 2006. Census 2011 notes that there was a 24% increase in rural households since Census 2001. Yet, the rural population increased by only 12% during the same period, implying that the average rural household size has gone down resulting in fewer children per household.

[7]Household based assessment of children; activity based assessment with easy to understand tools; community (volunteer) involvement in the actual survey; quick availability of the estimates; and rigorous methodology yielding reliable estimates at the state level.

[8]For example, in most of our research studies we sample 60 villages per district.

[9]This was done to save time since the survey is much longer as compared to the 'basic' ASER.

the total number of households visited to achieve this. The latter is needed for the calculation of correct weights[10].

To summarize, ASER 2019 'Early Years' employs a two-stage clustered design. In the first stage 60 villages are sampled from the Census 2011 village directory using PPS. In the second stage, 20 households with children in the age group of 4-8 years are surveyed in each sampled village. This gives a sample of at least 1,200 children in each district[11, 12].

The ASER 2019 'Early Years' report presents district report cards for each of the 26 surveyed districts. While the sample size of about 1,200 in each district is sufficient to present estimates for the target population as a whole or disaggregating by two sub-populations like gender or enrollment status, it is not sufficient for reliable estimates of smaller sub-populations. However, with a sample size of at least about 1,200 children per district the full sample is about 7,000 children. Therefore, we also present findings (appropriately weighted) based on the entire sample. Even though the geographical coverage of the ASER 2019 assessment of 'Early Years' competencies for young children is limited, since these districts are spread across the country in every major state, the full sample does give a snapshot of the national picture. So as to get a more balanced sample, two districts were surveyed by design in Uttar Pradesh and Madhya Pradesh, the two states with the largest number of districts in India.

---

[10]The probability that household j gets selected in village i (p_{ij}) is the product of the probability that village i gets selected (p_i) and the probability that household j gets selected ($p_{j(i)}$). This is given by:

$p_{ij} = p_i \; p_{j(i)} = \frac{n_v \, vpop_i}{dpop} \frac{n_{hi}}{vpop_i} = \frac{n_v n_{hi}}{dpop}$, where n_v is the number of villages sampled in the district, $vpop_i$ is the household population of village i, $dpop$ is the number of households in the district, and n_{hi} is the number of households visited in the village (to get the 20 sampled households). The weight associated with each sampled household within a district is the inverse of the probability of selection. Note that the sum of the weights of the households will give the district population and the sum of the weights for all children in the sample will approximate to the population of children in the 4-8 year age group in the district.

[11]All children in the target age group are surveyed in the sampled households. Therefore, the sample size in terms of children, by design, is at least 1,200 per district. However, it is possible to get fewer than 1,200 children, if some of the sampled villages do not have 20 households with children in the 4-8 age group– this is possible in smaller villages.

[12]For a two-stage sample design, sample size calculations have to take into account the design effect, which is the increase in variance of estimates due to departure from simple random sampling. This design effect is a function of the intra-cluster correlation. The greater this correlation, the larger is the design effect implying a larger sample size for a given level of precision. For a given margin of error (me), the sample size can be backed out from

$me = \frac{2\sigma}{p} = \frac{2\sqrt{\frac{d\,p(1-p)}{N-1}}}{p}$ where d is the design effect, p is the incidence in the population, σ its standard error and N the sample size. Since learning levels are unknown for the competencies being tested in 'Early Years', one has to start with some assumption about p. The largest uncertainty is around $p = 0.5$, and therefore that yields the largest sample size. Assuming no design effect, a margin of error of 10% and incidence of 0.5, gives a sample size of 400. A design effect of 2 would double that sample size. Therefore, a sample size of at least 1,200 in each district should give reasonable levels of precision. Of course, the combined estimates with a sample size of about 36,000 would be far more reliable.

State: District	Surveyed villages	Surveyed households	Surveyed children (age 4-8)	Surveyed children by age					Surveyed children (age 4-8) by sex	
				Age 4	Age 5	Age 6	Age 7	Age 8	Boys	Girls
Andhra Pradesh: Srikakulam	60	1172	1382	198	224	261	364	335	688	694
Assam: Kamrup (Rural)	60	1212	1308	238	249	264	304	253	704	604
Bihar: Nalanda	60	1202	1665	247	366	355	314	383	837	821
Chhattisgarh: Mahasamund	60	1202	1503	267	286	305	336	309	749	754
Gujarat: Mehsana	60	1200	1442	292	300	259	309	282	750	692
Haryana: Hisar	59	1203	1415	237	300	242	326	310	754	661
Himachal Pradesh: Kangra	60	1076	1334	260	270	279	287	238	679	655
Jharkhand: Ramgarh	59	1189	1456	237	309	275	317	318	728	722
Karnataka: Chamarajanagar	60	1192	1378	178	258	313	308	321	691	687
Kerala: Thrissur	49	1223	1417	250	304	271	305	287	700	715
Madhya Pradesh: Bhopal	60	1208	1568	244	308	316	351	349	827	741
Madhya Pradesh: Satna	60	1097	1365	291	282	252	280	260	662	703
Maharashtra: Nagpur	60	1212	1474	267	337	273	306	291	752	722
Manipur: Bishnupur	35	1202	1305	230	277	259	311	228	663	641
Meghalaya: East Khasi Hills	60	1137	1448	256	281	313	312	286	716	724
Nagaland: Dimapur	56	995	1172	210	251	236	241	234	603	566
Odisha: Khordha	60	1159	1252	227	222	240	282	281	653	599
Punjab: Bathinda	60	1207	1468	240	335	250	336	307	769	699
Rajasthan: Ajmer	60	1191	1660	263	346	368	356	327	813	836
Tamil Nadu: Vellore	60	1183	1550	302	298	291	332	327	760	790
Telangana: Karimnagar	60	1201	1426	230	262	305	308	321	746	679
Tripura: South District	60	1202	1257	198	279	270	268	242	657	600
Uttar Pradesh: Lucknow	60	1207	1494	240	340	290	308	316	793	701
Uttar Pradesh: Varanasi	60	1201	1615	232	397	328	341	317	858	757
Uttarakhand: Dehradun	56	985	1252	238	243	252	246	273	605	647
West Bengal: Bankura	60	1167	1324	250	268	251	299	256	667	657
All districts	1514	30425	36930	6322	7592	7318	8047	7651	18824	18067