Special Proceeding of the 22nd Annual Conference of SSCA held at Savitribai Phule Pune University, Pune, during January 02-04, 2020; pp 133-144

Large Scale Assessment Survey to Evaluate Learning Level of Students

Vishal D. Pajankar
National Council of Education Research and Training, New Delhi, India

Received: 17 July 2020; Revised: 12 August 2020; Accepted: 16 August 2020

Abstract

Many countries have started assessing their students through different assessment surveys to know the learning level of students, to know what they can do and what they know? In India, Large scale educational assessment survey named, National Achievement Survey (NAS) is being conducted at grades 3, 5, and 8 of the elementary stage under the flagship program of the Government of India. In 2017, the first time the NAS was conducted for grades 3, 5, and 8 on a single day i.e. November 13, 2017, in all 701 districts of 36 states and Union territories in India. About 2.2 million students from 1.2 lacs schools across the country were participated (NCERT, 2020). Schools in each distract were sampled using a stratified sampling procedure and section and students by NCERT designed a random sampling process. In this round, the district was the reporting unit of the study. NAS test items were constructed based on different competencies at different grades and linked them with learning outcomes instead of content-based. In this paper, we discussed how the NAS was implemented in the country and how the students' performed in the NAS. This paper highlights the performance of the students in different grades and different subjects. The comparisons between gender, school location, and school management are also discussed in the paper.

Key words: National achievement survey; IRT; Learning outcomes; Test items; Cohen's D.

1. Introduction

Large scale Assessment in education is one such tool that obtains information to assess the health of education systems and try to know whether the students meet curricular standards. Since the mid-1980s, the interest of measuring, comparing, and monitoring educational standards is growing in almost all countries. So, in the global countries have started assessing their students through different assessment surveys to know the learning level of students, to know what they can do and what they know? Many countries are taking participating in intercountry assessment surveys (large scale); for example PISA (Programme for International Students Assessment), TIMSS (the Trends in International Mathematics and Science Study), and PIRLS (the Progress in International Reading Literacy Study), *etc*. Some countries are conducting their assessment surveys to judge educational standards against national expectations. (NCERT, 2015a). The Sustainable Development Goal for education (SDG 4) is also called for an increased focus on learning outcomes, with five of the ten targets highlighting the learning skills and outcomes of children and adults (UNESCO, 2018).

In India, the large scale assessment survey is being conducted periodically since the seventies under different schemes and program of Govt. of India. In 2001, it is named as National Achievement Survey (NAS). The main aim of NAS is to provide reliable information about the achievement of students in the different grades of education in government and government aided schools. The NAS report gives a national and state-level picture, rather than scores for individual students or schools. The purpose of this survey is to obtain an overall picture of what students in specific classes know and can do and to use these findings to identify gaps and diagnose areas that need improvement. This information can then be used to impact policies and interventions for improving children learning in the country.

The data from National Achievement Survey gives the policymakers, curriculum specialists, researchers, and, most importantly, school principals and teachers a 'snapshot' of what students are achieving in key subjects at a particular point in time. By repeating such measurements at regular intervals, trends can be explored providing an invaluable perspective from which to consider educational reform and improvement. It does not give scores to individual students or schools (Pajankar, 2019).

2. History of National Achievement Survey (NAS)

India has a long history of conduct achievement (or assessment) surveys. The first notable survey was conducted by Kulkarni (1970) to know the achievement of students at different stages of school education in Mathematics. Another important study was untaken by Dave (1988) in NCERT under the project of Primary Education Curriculum Renewal (PECR) in 22 States at the primary stage in Language, Mathematics, and Environmental Studies. The third major study at the primary stage in Language and Mathematics was initiated by Shukla (1990) in NCERT and was completed in 1994 in 22 States and UTs. This was followed by district-specific surveys in primary classes under the District Primary Education Programme (DPEP) as the baseline, midterm, and terminal cycles (Dave, 1988 and NCERT, 2011).

Under the Sarva Shiksha Abhiyan (SSA) flagship program of Govt. of India, the survey is restructured and then named as National Achievement Survey (NAS). From 2001, NAS has been conducted in the different cycles in the country. The level was class III, class V, and class VIII in 2-3 years. Till 2017, 4 cycles of each grade have been conducted. In these cycles, the reporting unit was State and districts were sampled from each state. So NAS reported the learning level status of state only (Pajankar, 2019). The time-line of the conduct of NAS is given in table 1.

In 2017, the structure and nature of NAS was again changed. In 2017, the test was conducted on a single day November 13, 2017, in class III, V, and VIII. For NAS 2017, schools sample drawn through the Population Proportionate to Size (PPS) procedure includes nearly 2.2 million children from 1,10,000 schools spread across all districts in India. The salient features of this NAS 2017 were as below (NCERT, 2017):

- National Achievement Survey was linked to the learning outcomes;
- Assessment was being conducted for classes 3, 5 and 8 on a single day across the country;
- District was the unit for reporting;
- Automated reports were generated at the district level;
- Pedagogical interventions were provided in the same academic year.

National Achievement Survey 2017 was achieved by administering standardized tests to students of classes III, V, and VIII. NAS 2017 has contributed several new elements and gave remarkable momentum to the development of competency-based assessment. One of the main virtues of NAS 2017 is that it is embedded in an extremely rich system of background variables. The results help to accurately discover the students' performance in different learning outcomes vis-à-vis the contextual variables. The synthesis of the results of the national level provides a rich repository of evidence for developing and designing the future course of action for the Indian education system (Pajankar, 2019). Internationally accepted technical standards and practices were being adhered to while planning, designing, and implementing the NAS to ensure its robustness and sustainability (NCERT, 2017).

Table 1: The time period of the conduct of National Achievement Survey

Survey Cycle	**Class V**	**Class VIII**	**Class III**
Cycle I	2001-02	2002-03	2003-04
Cycle II	2005-06	2007-08	2007-08
Cycle III	2009-11	2010-13	2012-13
Cycle IV	2013-15	2014-16	2014-16
Subjects Tested	Mathematics Language Environmental-Studies	Mathematics Language Science Social Science	Mathematics Language Environmental-Studies
Background Questionnaires	Pupil Questionnaire (PQ)	Teacher Questionnaire (TQ)	School Questionnaire (SQ)

3. Sampling Design

In this NAS 2017, the target population was the students from classes III, V, and VIII from Government and Government aided schools. In earlier surveys, the state was reporting unit and 'n' number of districts was sampled. In NAS 2017, the district was reporting unit. So, all districts from 36 states/Union Territories were taken into consideration. 703 districts were listed in the sampling frame. But due to political reasons, 2 districts could not participate in the survey. Finally, the NAS was conducted in 701 districts in the country.

Sampling was done in three stages; first stage: school sampling; second stage: section sampling and third stage: students sapling. At the first stage, schools from each district were sampled using the Probability Proportional to Size (PPS) sampling procedure. Two types of stratification were used namely (i) explicit stratification (for the district) and (ii) implicit stratification (for block, area, school management, type of schools and medium of instructions). This exercise was conducted for each class III, V, and VIII, separately. The target sample size was 61 schools for class III and V and 51 schools for class VIII. Two additional lists of sampled schools, parallel to the original sampled list was prepared to replace the schools only if the original school does not exist or enrolment less than 5 or destroyed in natural calamities/ Naxal attack.

A second stage, one section of class III, V, and VIII was selected from each sampled school. If the sampled school has the only section of either class then the section was considered as selected. In the third stage, 30 students were selected from each selected section of sampled schools. The maximum target of students was 30. If in a class total attendance is less than 30 then all students were considered. The selection of sections in sampled schools and students from the selected section was conducted by a simple random procedure designed by the NCERT team.

Accuracy of a sample statistic as an estimate of an unknown population parameter is assessed through standard errors. Standard errors are computed through the following formula:

$$\sigma_{(\hat{\theta})} = \sqrt{\sigma^2_{(\hat{\theta})}}.$$

This formula assumes the use of Simple Random Sampling (SRS). Large scale assessments including NAS use complex sampling procedures. To ensure unbiased estimates of Standard Errors (SE) are generated, SEs are computed using the Jackknife Repeated Replication technique (JRR) for ability θ (NCERT, 2020).

4. Methodology

In this National Achievement survey 2017 (NAS 2017), schools were sampled using probability proportional to size procedure. In this process about 1,10,000 government and government aided schools were sampled from 701 districts of all states and union territories in the country. On November 13, 2017; the test was administered in all over the country About 2.2 million students from these sampled schools participated in NAS 2017. The students of classes III, V, and VIII were tested in different subjects through two sets of test booklets as shown below in Table 2.

Table 2: Class wise test booklets along with subjects and number of items

Class	Subjects	No. of items	Total items in Test Booklet	Number of Sets
III	Language* Mathematics Environmental Science	15 15 15	45	Two Sets 31 & 32
V	Language* Mathematics Environmental Science	15 15 15	45	Two Sets 51 & 52
VIII	Language* Mathematics Science Social Science	15 15 15 15	60	Two Sets 81 & 82

** Language used in a state as a local or regional language*

In NAS 2017, all subjects were tested through two test booklets for each class. Each subject had 15 items. So, 45 items in class III and class V and 60 items in class VIII. To

maximize the coverage of the test, two sets of test booklets were constructed. To establish a link between test booklets and to put them in a common platform, 05 items of each subject were common in both sets. Items were constructed with different competencies and linked with the learning outcomes (LOs) developed by the NCERT at elementary stage (classes I to VIII) in 2016. The items were piloted and removed all non-functioning items before finalizing the test booklets. The test booklets were then translated into 20 modern Indian languages. For quality check, translation was verified by experts and by back translation activity (with limited items).

Other booklets: three questionnaires *i.e.* Student Questionnaire (PQ), Teacher Questionnaire (TQ), and School Questionnaire (SQ) were also prepared for this NAS 2017. The objective of these questionnaires was to analyse the associations between the achievement and the background variables.

4.1. Analysis procedure

The data was collected through two sources; one was test booklets i.e. achievement data and another was questionnaires *i.e.* information of background factors collected from school heads, teachers, and pupils through interview mode. The achievement data of the students was analysed by classical test theory and item response theory. However, questionnaire data were analysed by classical test theory. Two different approaches were used to analyse and for reporting at a different level.

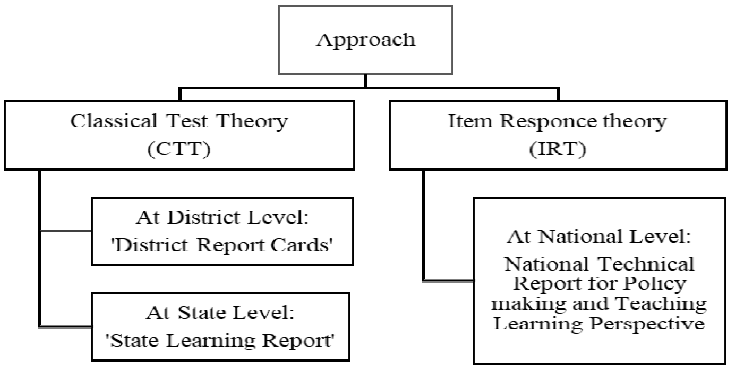


Figure 1: Different approaches used for the analysis at district level, state level and national level

Under the classical test theory, raw percentages of correct responses were used to measure students' abilities and item difficulties. With the classical test theory, the district report cards (DRCs) and state learning reports (SLRs) were generated for all districts and states/union territories with a record period of 2.5 months and 5 months respectively from the date of NAS 2017 administered. The district report cards were prepared in such a way that it can be easily read and understood by a layperson. The main objective behind it was that every parent/guardian can understand the learning level of his/her child. 10 independent report cards for each subject of each class of each district were generated. All DRCs and SLRs are available in the public domain at the NCERT web portal at link http://www.ncert.nic.in/programmes/NAS/NAS.html.

Table 3: List of district report cards for a district

Sl. No.	Class	Subject	No. of pages*
1	Class III	Language	2
2	Class III	Mathematics	2
3	Class III	Environmental Sciences (EVS)	2
4	Class V	Language	2
5	Class V	Mathematics	2
6	Class V	Environmental Sciences (EVS)	2
7	Class VIII	Language	2
8	Class VIII	Mathematics	2
9	Class VIII	Sciences (Sci)	2
10	Class VIII	Social Sciences (SSc)	2

**Few reports may have 3 pages*

Item Response Theory (IRT) approach was used in NAS 2017. Major large scale assessment studies conducted at international levels such as PISA (Programme for International Students Assessment), TIMSS (the Trends in International Mathematics and Science Study) and PIRLS (the Progress in International Reading Literacy Study), etc., are also using Item Response Theory (IRT). IRT measures the learning ability of students by calculating the probability of a student to respond to an item correctly. IRT analysis places students and test items on the same numerical scale and this helps us to create meaningful 'maps' of item difficulties and student abilities. In IRT, the difficulty of an item does not depend on the group of test-takers. Multiple test booklets can be used in IRT to increase the measurement points in any subject and the booklets can also be linked (NCERT, 2020).

IRT uses mathematical models that ensure the statistical connection between the difficulty level of the test item, the ability of the student, and the probability of that student being successful on a particular item. For example, students with higher ability scores are more likely to succeed on any item than their peers of lower ability Therefore, analysis in IRT is more complex than traditional methods like CTT. IRT uses the concept of an Item Characteristic Curve (ICC) to show the relationship between students' ability and performance on an item (NCERT, 2015b).

The two-parameter model (2PL) to the items was applied to analyse the data. The 2PL model associates student's ability to both item difficulty and item discrimination. The model includes difficulty (b) and discrimination (a) of the item. The expression for P_{ij}, the probability of the i^{th} examinee, ability θ_i, being successful on the j^{th} item, difficulty b_j is given by Thissen and Wainer (2001)

$$P_{ij} = \frac{\exp[a_j(\theta_i - b_j)]}{1 + exp[a_j(\theta_i - b_j)]}$$

$$P_{ij} = \frac{1}{1 + exp[-a_j(\theta_i - b_j)]}$$

where, P_{ij} is the probability of the i^{th} examinee, ability θ_i, being successful on the j^{th} item, difficulty b_j.

Test Reliability was estimated using the following formula

$$\bar{\rho} = \frac{\sigma_\theta^2 - \sigma_e^2}{\sigma_\theta^2}.$$

where, σ_θ^2 is the variance of the test score scale in the sample, and σ_e^2 is the mean error variance of scores. The values of both were estimated from BILOG software (Zimowski *et. al.*, 1996). At item and tests level, quality of achievement indices (or instruments) such as Item difficulty indices (*p*-value), Item discrimination indices (DI), Options analysis or Distractor analysis (DE), Differential Item Functioning (DIF) and Test reliability; were conducted using Classical Test Theory (CTT) and Item Response Theory (IRT) approach.

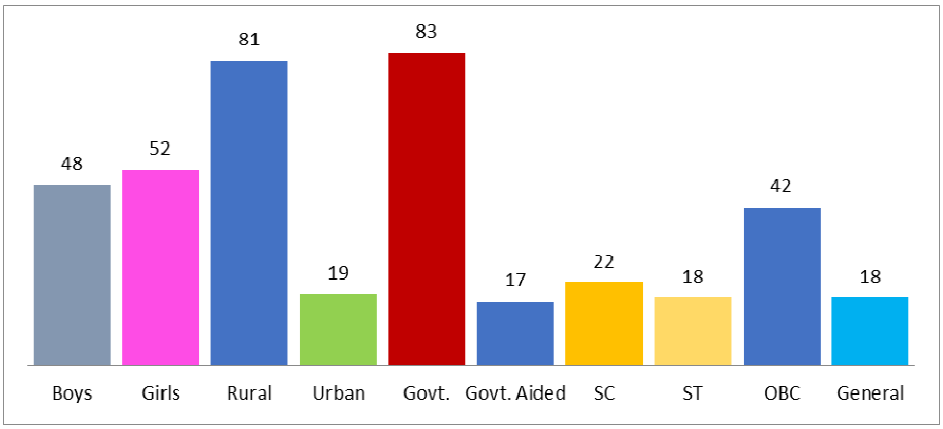
4.2. The reporting scale

IRT approach uses scale scores for reporting the results. In scaling, raw scores were transformed into a new set of scores by using either linear or nonlinear methods. The converted scores called Scaled Scores. The IRT scores were initially generated in the logit metrics, and then they were linearly converted into a scale that facilitates score interpretation. The reporting scale was set to the range of 100 - 500 with a mean of 300 and standard deviation of 50. Thus, the linear transformation from ability estimates expressed on the logit scale to the reporting scale scores was conducted using the expression: *Scale Score = Logit Score* * 50 + 300. Scaled scores were computed by statistically adjusting and converting raw scores into a common scale to account for differences in difficulty across different test forms (NCERT 2020 and 2014a).

5. Major Findings and Discussion

The National Achievement Survey (NAS) 2017 was conducted in India on dated November 13, 2017, in class III, V, and VIII. It was the first time when NAS for different classes administered on a single day. About 2.2 million from 1,10,000 government and government aided schools participated in this mega event. It may be the first kind of mega activities conducted on the globe in such a large magnitude. It includes participation from different sections: gender, location, management of schools, and social groups. Figure 2 shows the participation statistics.

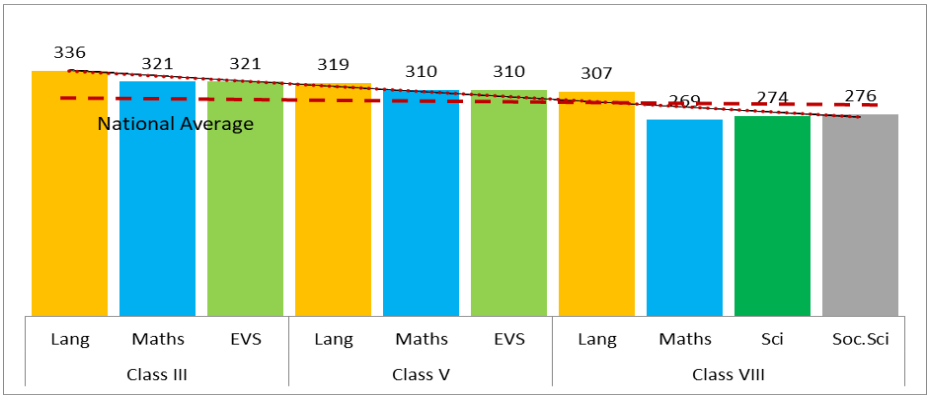
For gender, the participation of boys and girls was almost equal. Whereas, participation from rural-urban and government – government aided schools were very unequal. Participation was a cumulative representation in major social groups SC, ST, OBC and General as 22 %, 18 %, 42 % and 18 % respectively.



Source: NCERT (2020)

Figure 2: Participation by gender, location, school management and social groups in classes III, V, and VIII (figures in percentage)

Figure 3 shows the performance of students in classes III, V and VIII in the different subjects. The performance in different subjects was given in scale score values. It shows that the average achievement of students in class III was 326, the average achievement of students in class V was 313 and the average achievement of students in class VIII was 282. The overall national average of 300. Class III preference was much better than national performance and class V performance was close to the national average. However, the performance of class VIII was much below the national average. From the figure, it is concluded that with higher classes the performance of the students was decreasing.



Source: NCERT (2020)

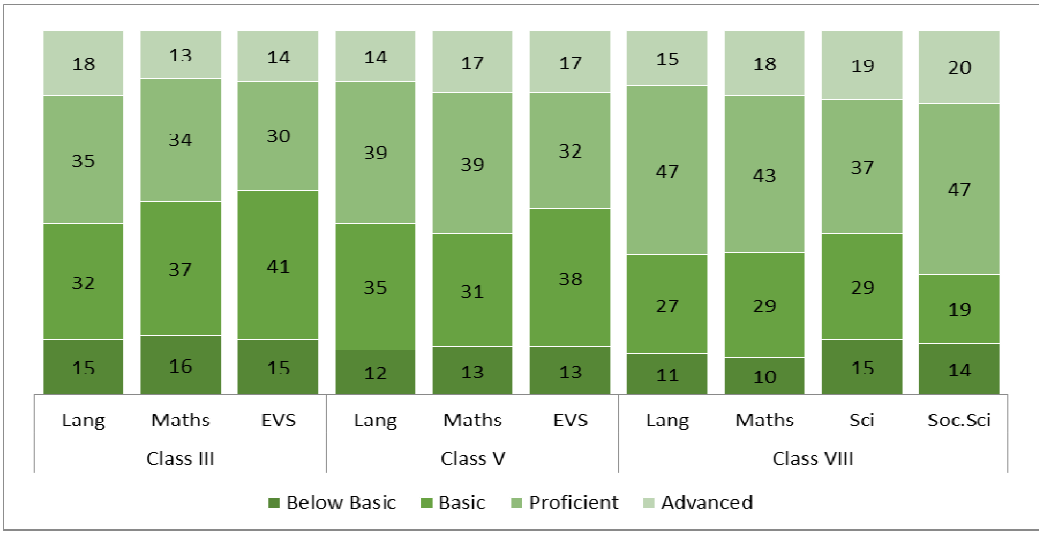
Figure 3: Students' performance at national level with class wise and subject wise and national average

Table 4 shown below presents the cut scores for each class III, V, and VIII for NAS. In Figure 4, students' performance at respective classes were given in percentage.

Table 4: Final cut scores for National Achievement Survey (NAS) tests

Test	Basic	Proficient	Advanced
Class III Language	268	315	370
Class III Mathematics	285	339	395
Class III Environmental Studies	263	315	375
Class V Language	260	306	370
Class V Mathematics	264	320	383
Class V Environmental Studies	261	315	375
Class VIII Language	255	320	370
Class VIII Mathematics	225	275	340
Class VIII Science	228	275	333
Class VIII Social Science	236	298	338

Source: NCERT (2020)



Source: NCERT (2020)

Figure 4: Percentage of students in each performance level (National Results)

Table 5: Performance of students in class III by gender, area and school management

Class III		Lang	Sig.	Cohen's D	Maths	Sig.	Cohen's D	EVS	Sig.	Cohen's D
Gender	Boys	335	**	-0.06	321	**	-0.01	320	**	-0.04
	Girls	338			321			322		
Area	Rural	336	**	-0.05	321	**	-0.02	320	**	-0.07
	Urban	336			322			324		
Management	Govt.	335	**	-0.12	320	**	-0.12	319	**	-0.22
	Govt. Aided	342			326			330		

Source: NCERT (2020)

*Note: Lang – Language, Maths – Mathematics and EVS – Environmental Science . * Statistically significant at p< 0.05; ** Statistically significant at p<0.01. The sizes of statistically significant differences are expressed by Cohen's D (Cohen, 1988). The size of the difference that is lesser than D=0.20 is considered small and practically irrelevant.*

Table 6: Performance of students in class V by gender, area and school management

Class V		Lang	Sig.	Cohen's D	Maths	Sig.	Cohen's D	EVS	Sig.	Cohen's D
Gender	Boys	317	**	-0.06	310	**	-0.04	309	**	-0.03
	Girls	320			311			310		
Area	Rural	318	**	-0.06	312	**	0.10	311	**	0.08
	Urban	321			306			306		
Management	Govt.	317	**	-0.12	311	**	0.07	311	**	-0.06
	Govt.Aided	324			308			307		

Source: NCERT (2020)

** Statistically significant at p< 0.05; ** Statistically significant at p<0.01. The sizes of statistically significant differences are expressed by Cohen's D (Cohen, 1988). The size of the difference that is lesser than D = 0.20 is considered small and practically irrelevant.*

Table 7: Performance of students in class VIII by gender, area and school management

Class VIII		Lang	Sig.	Cohen D	Maths	Sig.	Cohen D	Sci	Sig.	Cohen D	S.Sci	Sig.	Cohen D
Gender	Boys	306	**	-0.04	269	**	-0.01	275	**	0.02	278	**	0.01
	Girls	308			269			274			279		
Area	Rural	306	**	-0.09	271	**	0.18	276	**	0.16	280	**	0.13
	Urban	311			262			267			273		
Management	Govt.	305	**	-0.11	271	**	0.11	277	**	0.14	282	**	0.19
	Govt. Aided	311			265			269			271		

Source:NCERT (2020)

** Statistically significant at p< 0.05; ** Statistically significant at p<0.01. The sizes of statistically significant differences are expressed by Cohen's D (Cohen, 1988). The size of the difference that is lesser than D = 0.20 is considered small and practically irrelevant.*

6. Conclusion

From the overall analysis of NAS result, the following result was concluded as –

- Performance of Girls in the National Achievement Survey 2017 (NAS 2017) was slightly higher than boys' performance in almost all the classes.
- In class III, the achievement between urban and rural students was not distinguishable. In class V, urban students were performed statistically higher in language and rural students were performed higher in mathematics and environmental science. In class VIII, rural students were performed statistically higher in mathematics and sciences and urban students were higher in language.
- In class III, the performance of Government aided schools was statistically higher. In class V, Government schools were performed statistically higher in Mathematics and EVS whereas, in Government aided schools performance was higher in language. In class VIII, Government aided schools were performed statistically higher in Language, and Government in Mathematics, Science, and Social Science. However, in every class, the difference between the government and aided schools was very small.
- In class III, the performance of general category students was higher followed by OBC. In class V and VIII, general and OBC groups were performing slightly higher than other two social groups (SC and ST) in all subjects except language. In language, the OBC category was performed better than the general group in both classes. However, the differences were relatively small.
- A final remark on, if we considered the average of scale score achieved in all subjects in each class then there are no significant differences between gender (boys and girls) and areas (rural and urban schools). in all classes except management *i.e.*, between government and government aided schools however different was very small in class VIII. It means that boys and girls, and rural schools and urban schools were equally performed in the NAS 2017. But the performance of government schools and government aided schools was not the same.

The main objective of the study is to know the learning level of the student at different competency levels and to identify the gap in their learning. Ranking the states/Union Territories based on the performance of their students was not the objective of NAS 2017. Only to know the Top and Low performing states/union territories, ten names are: top ten states/union territories are Rajasthan, Karnataka, Chandigarh, Andhra Pradesh, Jharkhand, Dadra and Nagar Haveli, Assam, Gujarat, Kerala and Uttarakhand and low ten states/union territories are Arunachal Pradesh, Delhi, Puducherry, Meghalaya, Lakshadweep, Daman and Diu, Uttar Pradesh, Sikkim, Punjab, and Nagaland.

Acknowledgement

The author is indeed thankful to the Head and faculty members of the Educational Survey Division, NCERT, for providing the report and valuable guidance and support in the preparation of the paper. Thanks are also due to them for making available Figures and Tables from the technical report of NAS 2017 for the preparation of this paper.

References

Dave, P.N. (1988). *Achievement Under Project PECR*. New Delhi: NCERT (Mimeographed).

NCERT (2011). *Introduction of Item Response Theory in National Achievement Surveys Under Sarva Shiksha Abhiyan*. Concept Note. New Delhi, NCERT.

NCERT (2014 a). *National Achievement Survey*: *Class VIII* (*Cycle* 3). New Delhi, NCERT.

NCERT (2014 b). *National Achievement Survey* (*Cycle III*) *CLASS III*: *Achièvement Highlights*. New Delhi, NCERT.

NCERT (2015 a). *What Students Know and can do*: *A summary of National Achievement Survey Class X*. New Delhi, NCERT and MHRD, Government of India.

NCERT (2015 b). *What students of Class V know and can do*: *A Summary of India's National Achievement Survey, Class V* (*Cycle* 4). New Delhi, NCERT.

NCERT (2017). *Operational Guidelines-cum-Training Manual*: *National Achievement Survey - NAS* (2017) *Classes III, V and VIII*. New Delhi, NCERT and MHRD.

NCERT (2018). *Post NAS Intervention*: *Communication and Understanding of Districts Report Cards for NAS 2017*. New Delhi, NCERT.

NCERT (2020). *National Report to Inform Policy, Practices and Teaching Learning*: *National Achievement Survey* 2017 *Grades III, V and VIII. New Delhi*, NCERT and UNICEF (Unpublished).

Pajankar, V. D. (2019). Assessment of students' learning in different subjects at different levels of grades in Maharashtra (India). *International Journal of Research*, **9**(**1**), 73–89.

Thissen, D. and Wainer, H. (2001). *Test Scoring. Mahwah*, New Jersey, Lawrence Erlbaum Associates.

UNESCO (2018). *The Impact of Large-Scale Learning Assessments*. UNESCO and UIS, March 2018.

Zimowski, M., Muraki, E., Mislevy, R. and Bock, D. (1996). BILOG-MG (Computer program). *Available from Scientific Software International*, Inc. 7383 N. Lincoln Avenue, Suite 100 Lincolnwood, IL 60712-1747 USA.